



Scire Vias

Humanidades Digitales
y Conocimiento

V Congreso de la Sociedad Internacional de Humanidades Digitales Hispánicas

Santiago de Compostela, 4 a 8 de octubre de 2021

www.hdh2021.org · congreso.hdh2021@usc.es · [@hdh2021](https://twitter.com/hdh2021)

Libro de resúmenes



Tabla de contenido

Sesión 1	La vía de la palabra: Lenguaje, discurso y argumentación	3
Panel 1	Proyecto Mnemosine: <i>Data Reperta et Historia Digitalis</i>	33
Sesión 2	La vía de la implementación: Herramientas e infraestructuras	40
Plenaria 1	<i>Guillermo Rojo</i> : La lexicografía electrónica	65
Panel 2	Experiencias en torno a la <i>Text Encoding Initiative</i>	66
Sesión 3	La vía de la difusión: Edición digital y medios digitales	72
Panel 3	Humanidades digitales y teatro	89
Sesión 4	La vía de lo femenino: Estudios culturales I	97
Plenaria 2	<i>Emilio Ros-Fábregas</i> : Hacia la codificación del patrimonio musical hispánico en el contexto de las Humanidades Digitales: desafíos y repercusión historiográfica, económica y cultural	110
Sesión 5	La vía del sentido: Estudios de corpus	111
Sesión 6	La vía del testimonio: Estudios culturales II	150
Sesión 7	La vía de la conservación: Digitalización de archivos, bibliotecas y museos	172
Plenaria 3	<i>Felipe Criado-Boado</i> : Cognición y materialidad	202
Panel 4	Aracne Nodus. Red de humanidades digitales y letras hispánicas	203
Panel 5	El proyecto «Mapping Chivalry. Spanish Romances of Chivalry from Renaissance to 21 st Century: A Digital Approach»	208
Panel 6	Redes Sociales del Pasado. Cartografía de la Modernidad Literaria Hispánica, 1898-1959	214
Sesión 8	La vía de la expresión: Estudios culturales III	223
Sesión 9	La vía de la civilización: Educación, enseñanza y participación ciudadana	245
Plenaria 4	<i>Patricia Murrieta-Flores</i> : La decolonización de lo imaginario: hacia la inclusividad y la consideración de lo subalterno en las Humanidades Digitales	270
Panel 7	INTELE: Red estratégica para integrarse en las infraestructuras digitales europeas de investigación en Humanidades	271
Panel 8	Historia del Arte Digital: conocimientos renovados, nuevos conocimientos	276
Panel 9	Eventos festivos históricos y Humanidades digitales: reconstrucción virtual y estudio interdisciplinar de la entrada de Isabel de Valois en Toledo (1560)	284
Sesión 10	La vía de la definición: Modelos, ontologías y simulaciones	289
Sesión 11	La vía de la visión: Representaciones 3D, realidad virtual y realidad aumentada	303
Plenaria 5	<i>Amparo Alonso Betanzos</i> : Las Humanidades en la era de la Inteligencia Artificial: cooperando para modelar sostenibilidad	334

Sesión 1

La vía de la palabra

Lenguaje, discurso y argumentación

Modera: César González-Pérez · Asiste: César González-Pérez

- 18 Entre la Filología y la Historia: el proyecto Cordican y el valor de la interdisciplinariedad en las HD
Roberto J. González Zalacain, Dolores Corbella Díaz, Ana Viña Brito

- 51 Una metodología semiautomática de anotación de entidades nombradas para la creación de un *gold standard*
Susana Sotelo Docío, Pablo Gamallo, Álvaro Iriarte Sanromán

- 48 El español de Cuba en el siglo XIX: estudio sociolingüístico y propuesta de red semántica a través de textos oficiales y egodocumentos
Eva Bravo-García

- 36 Pygmalion: una herramienta para la visualización interactiva del cambio semántico desde una perspectiva diacrónica
Helena Bermúdez Sabel

- 44 Construcción de un corpus de evaluación de la anotación léxico-gramatical del castellano medieval (siglos 13-15)
Matthias Gille Levenson, Olivier Brisville-Fertin, María Díez Yáñez, Simon Gabay

- 69 Rhetor (Red de Humanidades y Estadística Aplicada al Texto Oratorio): un proyecto interdisciplinar y transversal
Ángela Celis Sánchez, Víctor Casero-Alonso, María Lozano Zahonero

Entre la Filología y la Historia: el proyecto Cordican y el valor de la interdisciplinariedad en las HD

Roberto J. González Zalacain, Dolores Corbella Díaz, Ana Viña Brito

Universidad de La Laguna

Palabras clave: Edición digital de textos, corpus documental, interdisciplinariedad, lingüística de corpus, repositorio documental

Descripción del problema de investigación, metodología y resultados obtenidos.

La edición documental tiene una profunda tradición en el ámbito académico del archipiélago canario, con ejemplos muy significativos que han permitido avanzar en el conocimiento del pasado histórico y lingüístico de las islas. No obstante, esa labor siempre había sido concebida desde un paradigma de trabajo individual y con un marcado carácter disciplinar, desde la óptica ora de la Historia, ora de la Filología.

Esta inercia investigadora se ha roto en los últimos años. Merced al trabajo del grupo de investigación *LexHis* (<http://lexhis.webs.ull.es/>) se inició una línea de investigación interdisciplinar, que aunó a historiadores y filólogos, a través de la cual se abordó el análisis multifocal de la historia del azúcar en el ámbito atlántico en la época colonial y sus implicaciones en el campo lingüístico. El objetivo principal pasaba por acreditar que los hechos históricos permitían afirmar el proceso operado en las islas Canarias de españolización de la terminología azucarera atlántica. Esto permitió rebatir la postura clásica, sustentada en los repertorios lexicográficos americanos, que avalaban la interpretación tradicional de la adscripción brasileña de este conjunto léxico.

El trabajo documental requerido en esta inicial línea de investigación implicó la transcripción literal de decenas de registros manuscritos, conservados en diversos archivos, que contenían información acerca de la instalación de los ingenios azucareros en el archipiélago y, en general, del desarrollo de esta agroindustria tanto en Canarias como en América, con una cronología que abarcaba desde finales del siglo XV, cuando se producen las primeras instalaciones en las islas, hasta principios del siglo XVIII.

Más allá del resultado concreto de esta investigación, con la publicación de hasta tres volúmenes monográficos (dos de documentación canaria seleccionada en función de su cercanía al tema y otro con un análisis de la difusión atlántica del léxico azucarero), el principal valor de ese trabajo conjunto entre historiadores y filólogos fue la posibilidad de avanzar en la exploración del enorme potencial que tanto para los estudios históricos como para los filológicos implicaban la puesta a disposición de la comunidad científica de la colección documental íntegra manejada en estos análisis.

Es en este punto en el que entran en escena las Humanidades Digitales, y más concretamente las posibilidades que la edición en XML y el etiquetado TEI ofrecen a ambas disciplinas. Se trata de un campo de investigación científica muy transitado desde el campo de la Lingüística, pero mucho menos explorado desde la Historia, por lo que los avances que este tipo de proyectos pueden ofrecer a cada uno de los campos científicos, y muy especialmente las derivadas del trabajo interdisciplinar, son muy notables.

Para ello, partiendo de la base del programa informático diseñado para el desarrollo de la investigación ya reseñada, cuyos metadatos facilitaban tanto la gestión de búsquedas simples como combinadas, a partir de una

serie de campos predefinidos, se ha desarrollado el proyecto *Corpus documental de las islas Canarias* (CORDICan), en construcción, que parte de una solución informática con una visión integral de estas temáticas basada en Teitok (<http://teitok.corpuswiki.org/>), lo que permite una optimización de los procesos de trabajo y una potente visualización de las transcripciones. De este modo, las transcripciones que se vuelquen a través de ese portal ofrecen una triple posibilidad de consulta (transcripción textual, edición con desarrollo de abreviaturas y versión normalizada) a la que se le añaden las potenciales virtudes de la edición en XML para los estudios lingüísticos. Y, todo ello, con la posibilidad complementaria de cotejar las transcripciones facilitadas con la reproducción digital del documento original, lo que abre a su vez una dimensión de gran interés para el campo de la docencia universitaria, especialmente de la especialidad de Paleografía.

El impulso institucional a este proyecto, actualmente alojado en los servidores de la Universidad de La Laguna, ha venido de la mano de la obtención, en concurrencia competitiva, de un proyecto de investigación financiado por la Agencia Canaria de Investigación, Innovación y Sociedad de la Información del Gobierno de Canarias, que permitirá implementar las últimas mejoras en la visualización del entorno web y aumentar considerablemente los recursos documentales disponibles.

El objetivo de la presente comunicación es el de compartir con la comunidad científica interesada en las Humanidades Digitales la experiencia acumulada en el proyecto *Cordican*, y participar del debate académico en los campos de la digitalización de archivos, los estudios de corpus, la edición digital y la educación y enseñanza.

Bibliografía.

BARAIBAR, Á. (2014). «Las humanidades digitales desde sus centros y periferias». En *Humanidades Digitales: una aproximación transdisciplinar*. Janus, Anexo 2, pp. 7-15.

BARAIBAR, Á., y COHEN, S. (2012). «Nuevas tecnologías y redes sociales en la investigación en Humanidades», *La Perinola*, 16, pp. 155-164.

BERTOLOTTI, V., y COMPANY, C. (2014). «El Corpus Diacrónico y Diatópico del Español de América (CORDIAM). Una propuesta de tipología textual». En PARODI, C., y CARRERA DE LA RED, M. (eds.). *Cuadernos de la ALFAL*, 6 (número especial: *El español en América. Corpus y textos*), pp. 130-148.

BRITO GONZÁLEZ, A. (2006). «Entre el pasado y el futuro. Recursos en internet para la investigación histórica en Canarias». *Cartas Diferentes. Revista Canaria de Patrimonio Documental*, 2, pp. 95-118.

CALDERÓN CAMPOS, M. (2015). *El español del reino de Granada en sus documentos (1492- 1833)*. Berna: Peter Lang.

CORRALES, C., y CORBELLA, D. (2012). «Historia y documentación del canarismo léxico». *Cuadernos del Instituto de Historia de la Lengua*, 7, pp. 143-180.

CORRALES, C.; CORBELLA, D., y VIÑA BRITO, A. (2014). *Léxico azucarero atlántico*. San Millán de la Cogolla (Logroño): Cilengua.

ENRIQUE-ARIAS, A. (2012). «Dos problemas en el uso de corpus diacrónicos del español: perspectiva y comparabilidad». *Scriptum Digital*, 1, pp. 85-106.

FERNÁNDEZ IZQUIERDO, F. (2006). «Investigar, escribir y enseñar historia en la era de internet». *Hispania*, 66, nº 222, pp. 11-30.

GARCÍA MARCO, F. J. (2006). «Los sistemas de información histórica: una nueva frontera en la construcción científica de la Historia». *Aragón en la Edad Media*, 19 (*Homenaje a la profesora María Isabel Falcón*), pp. 213-234.

- IGLESIAS, F. (2006-2007). «La recuperación de documentos canarios en América». *Anuario Americanista Europeo*, 4-5, pp. 181-195.
- KABATEK, J. (2016). *Lingüística de corpus y lingüística histórica iberorrománica*. Berlín / Boston: De Gruyter.
- MAÍZ CHACÓN, J. (2005). «El crepúsculo tecnológico del medievalismo hispánico: nuevas tecnologías e historia medieval o el ocaso de lo desconocido». En *Pescar o navegar: la Edad Media en la red*. Zaragoza: Universidad de Zaragoza, pp. 67-83.
- PÉREZ HERRERO, E. (2009). *Historia de los archivos de Canarias. I*. Las Palmas de Gran Canaria: Anroart Ediciones.
- PONS, A. (2013). *El desorden digital. Guía para historiadores y humanistas*. Madrid: Siglo XXI.
- PRIANI SAISÓ, E., y otros (2014). «Las humanidades digitales en español y portugués. Un estudio de caso: Día HD/Día HD». *Anuario Americanista Europeo*, 12, pp. 5-18.
- RAMÍREZ, M. (2016). «Las ciencias y las técnicas historiográficas en el contexto de las Humanidades digitales: oportunidades para el desarrollo». En MARCHANT RIVERA, A. (ed.) y BARCO CEBRIÁN, L. (coord.). «*Dicebamus hesterna die...*» estudio en homenaje a los profesores Pedro J. Arroyal Espigares y M^a Teresa Martín Palma. Oviedo-Málaga: Editorial Encasa- Universidad de Málaga, pp. 365-392.
- ROJAS MAYER, E.M^a (2000). *Documentos para la historia lingüística hispanoamericana. Siglos XVI a XVIII*. Madrid. Anexo del Boletín de la Real Academia Española. Se había publicado un año antes en CD-ROM, Tucumán: Universidad Nacional de Tucumán y ALFAL.
- SÁNCHEZ-PRIETO BORJA, P. (2011). *La edición de textos españoles medievales y clásicos. Criterios de presentación gráfica*. San Millán de la Cogolla: Cilengua.
- SERRANO-TENLLADO, M^a A., CALVO, A., y ROMERO, J. A. (2001). «La representación de documentos históricos: aplicación de esquemas y XML para modelar y representar la información y el conocimiento descrito en las fuentes de protocolos notariales». En EXTEMEÑO PLACER, A. I. (coord.). *La representación y organización del conocimiento: metodologías, modelos y aplicaciones. Actas del V Congreso ISKO-España, 25-27 de abril de 2001*. Alcalá de Henares, s.p.
- TORRUELLAS CASAÑAS, J. (2017). *Lingüística de corpus: génesis y bases metodológicas de los corpus (históricos) para la investigación en lingüística*. Fráncfort: Peter Lang.
- VIÑA BRITO, A.; CORRALES, C., y CORBELLA, D. (2014). *Islas y voces del azúcar I (Tenerife, La Gomera y La Palma)*. San Cristóbal de La Laguna - Las Palmas de Gran Canaria: Gobierno de Canarias - Archivos Históricos Provinciales de Tenerife y de Las Palmas de Gran Canaria.
- VIÑA BRITO, A.; CORRALES, C., y CORBELLA, D. (2016). *Islas y voces del azúcar II (Gran Canaria)*. San Cristóbal de La Laguna - Las Palmas de Gran Canaria: Gobierno de Canarias - Archivos Históricos Provinciales de Tenerife y de Las Palmas de Gran Canaria.

Una metodología semiautomática de anotación de entidades nombradas para la creación de un *gold standard*

Susana Sotelo Docío
susana.sotelo.docio@usc.gal

Pablo Gamallo
pablo.gamallo@usc.gal

Alvaro Iriarte Sanromán
alvaro@ilch.uminho.pt

1. Introducción

El Reconocimiento de Entidades Nombradas (NER, *Named-Entity Recognition*) es una tarea del Procesamiento del Lenguaje Natural (NLP, *Natural Language Processing*) que tiene por objeto la detección y clasificación de entidades nombradas en un corpus textual. Esta tarea ha sido aplicada a múltiples dominios específicos, como pueden ser documentos médicos, jurídicos, de aviación, genómica o la extracción de nombres a partir de documentos periodísticos. Abordajes más recientes exploran la posibilidad de aplicar técnicas NER a nuevos campos tradicionalmente más alejados de la computación, tales como el análisis literario, dentro del contexto de la llamada *lectura distante*[2].

En este trabajo describimos el proceso de creación de un corpus multilingüe de 250 entrevistas telefónicas a visitantes de Santiago de Compostela, orientado al análisis de consumos y localizaciones en la ciudad y el Camino de Santiago. Con este propósito, desarrollamos una metodología de anotación semiautomática para entidades de tipo 'localización', 'organización', 'persona' y 'miscelánea', empleando la herramienta Linguakit[4] con revisión manual posterior. Este corpus, codificado en XML con esquema XSD propio, servirá de *gold standard* para el entrenamiento y evaluación de sistemas NER adaptados al dominio de la cultura y el turismo, así como para la geolocalización de los espacios referenciados.

El trabajo que proponemos presentar en el congreso se estructurará de la siguiente forma: en primer lugar describiremos las fuentes empleadas para la construcción del corpus, así como el proceso de transcripción y codificación. A continuación, introduciremos la metodología utilizada para el proceso de anotación de entidades nombradas, con el objetivo de construir un *gold standard* para NER en el ámbito del análisis cultural y el turismo ligado a la ciudad de Santiago de Compostela y el Camino. Finalmente, describiremos brevemente algunas posibilidades de explotación basadas en técnicas de análisis lexicométrico[7].

2. Descripción de las fuentes del corpus

El corpus de entrevistas usado como punto de partida para la elaboración de un *gold standard* para NER fue compilado en el marco del proyecto *Discursos, imágenes y prácticas culturales de Santiago de Compostela como meta del Camino*¹, en el que se aunaban dos niveles de análisis: el análisis de los productos culturales como objeto de conocimiento, y el de los discursos sociales (incluyendo prácticas y usos) hacia ese objeto de conocimiento. En este marco se llevó a cabo una recogida de datos de índole cuantitativa, a través de 2081 encuestas a visitantes de la ciudad de Santiago procedentes de Galicia, España, Portugal y Brasil. Sobre la base de estas encuestas se

¹ Discursos, imágenes y prácticas culturales sobre Santiago de Compostela como meta de los Caminos de Santiago. Proyecto de 3 años de duración financiado por la Subdirección General de Proyectos de Investigación. Dirección General de Investigación Científica y Técnica. Ministerio de Economía y Competitividad. Gobierno de España [Código: FFI2012-35521] (2012-2015).

hizo posteriormente una selección proporcional de perfiles, a los que se realizó una entrevista telefónica semiestructurada en una segunda fase de recogida de datos de índole cualitativa (tabla 1).

Estas entrevistas semiestructuradas, realizadas en tres idiomas (gallego, español y portugués), constituyen la fuente del corpus.

1

	Encuestas	Porcentaje	Entrevistas	Porcentaje
Galicia	398	19,13%	47	18,80%
España	878	42,19%	90	36,00%
Portugal	408	19,61%	54	21,60%
Brasil	396	19,03%	59	23,60%
NC	1	0,04%	0	0,00%
Total	2081	100,00%	250	100,00%

Cuadro 1: Proporción de encuestas y entrevistas telefónicas realizadas, segmentadas por origen (Fuente: [1])

3. Codificación del corpus

El proceso de codificación del corpus consta de varias fases:

- la transcripción a partir de las grabaciones telefónicas,
- la corrección y armonización de los textos, la
- codificación de los textos transcritos en XML.

En la tabla 2 se muestra la distribución de palabras del corpus en función de la lengua del texto. Se aprecia que el tamaño de los textos en portugués y español es similar, frente al gallego, que representa el 13,48% del corpus total.

Lengua	Palabras	Porcentaje
gallego	203.567	13,48%
español	653.563	43,27%
portugués	652.999	43,24%

Cuadro 2: Datos del corpus

4. Anotación del corpus

La metodología empleada para la anotación de entidades se basa en el uso de sistemas automáticos de detección de entidades nombradas, con una fase posterior de corrección manual. En nuestro caso la herramienta empleada

es Linguakit[4], que incluye soporte para las tres lenguas del corpus. En la figura 1 se muestra el flujo de procesamiento de cada documento.

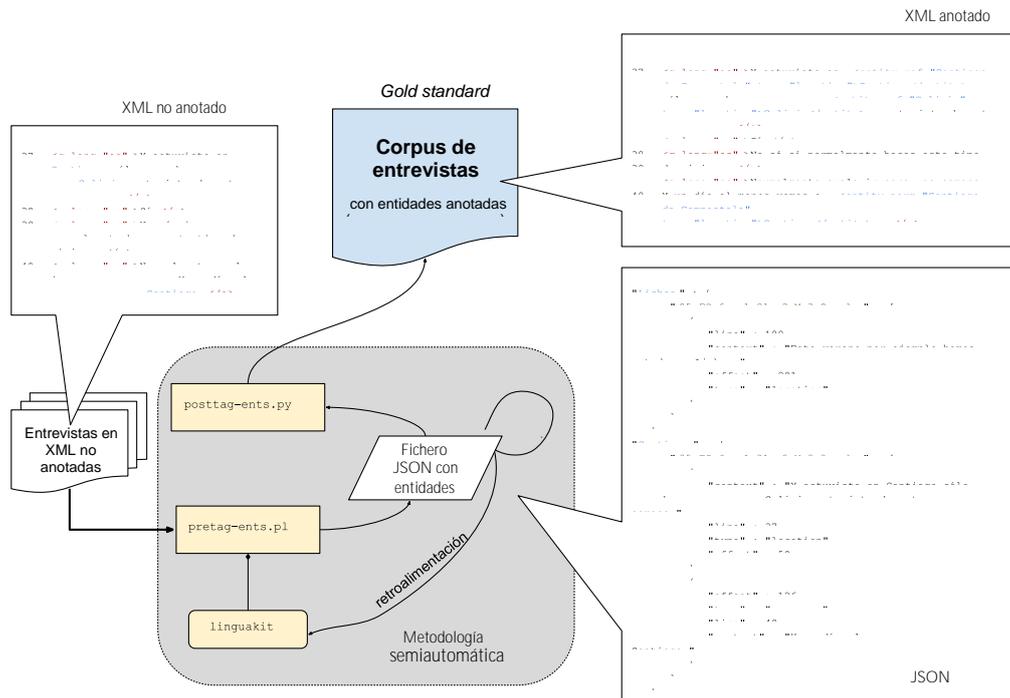


Figura 1: Flujo de trabajo empleado para la anotación semiautomática de entidades (elaboración propia)

Un script (**pretag-ents.pl**) extrae automáticamente las entidades nombradas mediante llamadas a la herramienta Linguakit, almacenando la información en un fichero JSON junto al contexto en que dichas entidades fueron extraídas y otros metadatos. Este fichero JSON es posteriormente editado por un anotador humano, que podrá hacer las correcciones necesarias allí donde el procedimiento automático haya incurrido en errores. Una vez finalizada esta corrección, un segundo script (**posttag-ents.py**) integra las entidades nombradas ya corregidas en el documento XML, que es la salida final del proceso y que constituye un *gold standard* anotado para NER.

Esta metodología asegura una alta precisión, ya que se realiza una corrección manual posterior de la clasificación automática de las entidades detectadas. Sin embargo, esta corrección manual no supone un incremento de la cobertura, que depende de los resultados producidos por Linguakit. Por ello, el proceso se completa con diversas heurísticas que incrementan la cobertura, tratando de detectar entidades que no fueron extraídas de forma automática.

4.1. Detección de entidades nombradas

El módulo NER de Linguakit identifica expresiones *numex* (de base numérica) y *enamex* (nombres propios) mediante máquinas de estados finitas, que tienen en cuenta tanto las formas ortográficas (uso de mayúsculas) como las palabras funcionales que puedan contener (*Universidad de Santiago de Compostela*). Una vez identificadas estas entidades, el módulo de clasificación aplica un método de supervisión distante que le permite clasificar las entidades en cuatro clases: *persona*, *organización*, *localización* y *miscelánea*. El sistema emplea listas de entidades ya conocidas (*gazetteers*) y un conjunto de reglas contextuales para desambiguar aquellas entidades que pueden aparecer en más de una lista, así como para determinar la clase más probable de las entidades desconocidas.

Este módulo NER fue evaluado para cuatro lenguas (inglés, portugués, español y gallego)[3, 5, 6], con resultados próximos a sistemas como Freeling o Stanford CoreNLP, y superando ampliamente los resultados de los modelos incluidos en OpenNLP.

5. Explotación del corpus

En este apartado discutiremos las posibilidades de explotación del corpus, por ejemplo mediante técnicas de análisis lexicométrico para la detección de especificidades lingüísticas de los diferentes participantes que intervienen en las entrevistas [7]. Se describirán también las futuras líneas de trabajo, tanto en lo que se refiere a mejoras en el corpus (geolocalización de las entidades nombradas detectadas y anonimización automática de datos personales para su disponibilización en abierto) como a sus posibles usos para entrenar y evaluar sistemas NER basados en aprendizaje automático (*machine learning*) y adaptados al ámbito específico del análisis cultural y el turismo, así como al lenguaje oral de las entrevistas, más informal que el texto escrito.

Referencias

- [1] Luís Gustavo Cuntín Docampo. Labores de asesoramento, recollida e rexistro mediante enquisa, e tratamento de información. Technical report.
- [2] Francesca Frontini, Carmen Brando, Joanna Byszuk, Ioana Galleron, Diana Santos, and Ranka Stanković. Named Entity Recognition for Distant Reading in ELTeC. In C. Navarretta and M. Eskevich, editors, *Proceedings of CLARIN Annual Conference 2020*, pages 37–41.
- [3] Pablo Gamallo and Marcos Garcia. A resource-based method for named entity extraction and classification. *LNCS*, 7026:610–623, 2011.
- [4] Pablo Gamallo and Marcos Garcia. Linguakit: uma ferramenta multilingue para a análise linguística e a extração de informação. *Linguamática*, 9(1), 2017.
- [5] M. Garcia, I. González, and I. del Río. Identificación e clasificación de entidades mencionadas em galego. *Estudos de Linguística Galega*, 4:13–25, 2012.
- [6] Marcos Garcia and Pablo Gamallo. Yet another suite of multilingual NLP tools. In *Languages, Applications and Technologies*, volume 563 of *Communications in Computer and Information Science*, pages 65–75, Switzerland, 2015. Springer. Revised Selected Papers of the Symposium on Languages, Applications and Technologies (SLATE 2015).
- [7] Alvaro Iriarte, Pablo Gamallo, and Alberto Simões. Estratégias Lexicométricas para Detetar Especificidades Textuais. *Linguamática*, 10:19–26.

El español de Cuba en el siglo XIX: estudio sociolingüístico y propuesta de red semántica a través de textos oficiales y egodocumentos

Eva Bravo-García

Universidad de Sevilla

Instituto Universitario de Estudios sobre América Latina (IEAL)

Resumen

Esta comunicación presenta algunos aspectos del proyecto Cuba y Andalucía en el siglo XIX: Estudio de los lazos lingüísticos y culturales desde las Humanidades Digitales (FEDER-US126310) de la Universidad de Sevilla². El objetivo general de este proyecto es ampliar las fuentes por las que obtenemos nuestro conocimiento en el ámbito de la historia de la lengua y en concreto del español en América en el siglo XIX, aplicando enfoques y técnicas de las HD.

1. Introducción

Los siglos XVIII y XIX conforman una época clave en la historia del español en América; además de que plantean una amplia gama de problemas, ya sea sobre fenómenos concretos, periodización o normalización lingüística (Bertolotti/Coll 2012), son los siglos en los que se gestó la identidad de las naciones americanas. Los cambios sociales y culturales que darán lugar a los procesos de independencia y al surgimiento de nuevos países se gestan entre 1770 y 1900, coincidiendo con la definitiva expansión de la lengua española en América.

Sin embargo, en contraste con la profundidad y precisión de la comprensión que hemos alcanzado sobre la lengua de los siglos XVI y XVII, el conocimiento de la lengua de los siglos XVIII y XIX es aun parcial y fragmentario, suplido, en ocasiones, mediante el expediente de trazar un puente entre lo que sabemos de los Siglos de Oro y del español actual, que sirva como base para la reconstrucción.

Se da, por otra parte, la circunstancia de que, como instrumento sumamente valioso para el análisis de la lengua de estos siglos, se encuentra accesible una extensa masa documental, que permanece en su mayor parte inédita, cuya publicación y estudio permitiría esclarecer los elementos de innovación lingüística que afloran a partir de las importantes transformaciones sociales de estas centurias.

1.1. Presupuestos e hipótesis

El proyecto asume la existencia del área variacional caribeña (que incluye la isla de Cuba, las Antillas y las regiones caribeñas de Colombia y Venezuela) como ámbito geográfico de una variedad hispanoamericana específica (Perl 2001: 655), con centro en Cuba. Sobre este presupuesto parte la hipótesis de que el contexto histórico y social de la Cuba de los ss. XVIII y XIX y, concretamente, una serie precisa de circunstancias económicas, sociales y culturales, son determinantes en la creación y definición específica de la lengua de Cuba, y con ella, del área variacional caribeña.

En este contexto, el proyecto proponiendo que es justamente la mezcla genuina de emigrantes e influencias externas a lo hispánico, de una parte, y de otra, la emergencia de componentes internos y la

² Programa Operativo FEDER 2014-2020 (Consejería de Economía, Conocimiento, Empresas y Universidad de la Junta de Andalucía). IP Eva Bravo-García, Co-IP Ana Mancera Rueda.

movilidad social lo que provoca el progresivo crecimiento de un germen identitario, dando lugar a un proceso original de nivelación; la investigación podrá así revelar el proceso de aparición de una koiné postcolonial con foco en la isla.

1.2. Objetivos

La investigación se centra en los siguientes objetivos generales:

1. Analizar usos y parámetros sociolingüísticos que evidencian un estado de lengua en la segunda mitad del siglo XIX, atendiendo al continuo dialectal Andalucía-Caribe.
2. Conceptualizar y formular nuevos parámetros y significaciones culturales y sociales relacionados con el proceso de independencia y la constitución de la identidad cubana.
3. Identificación de tendencias sociolingüísticas y culturales coincidentes entre Andalucía y Cuba y su manifestación en la prensa.

Con este fin, se llevarán a cabo investigaciones que se enmarcan en el seno de las *Humanidades Digitales*, especialmente mediante el tratamiento de texto utilizando TEI-TIOK y la creación de redes semánticas a través de Atlas.ti.

2. Descripción del Corpus

La investigación parte de la construcción de un corpus textual no cerrado (Caravedo 1999: 98)³, integrado por documentación relacionada con el proceso de independencia, en concreto se trata de “documentación tomada al enemigo”, material escrito recogido en el campo de batalla, del que se hace acopio para su posterior análisis. Esta circunstancia nos permite disponer de un conjunto sincrónico de documentos que abarca muy distinta tipología y finalidad comunicativa, realizados por emisores de diferente estatus social y formación. La imposición de una condición conceptual garantiza la homogeneidad conceptual de los materiales del corpus, que comparten un “universo de discurso”⁴: responden al mismo hecho histórico y deben ser interpretados conforme a idéntico contexto situacional. En concreto, formarán parte del corpus aquellos documentos que contengan:

- a) Referencias al proceso independentista: cartas de y documentos dirigidos a la ciudadanía de ambos bandos contendientes (img. 1, 3, 4, 6 y 12) y comunicaciones militares internas de instrucciones o informes oficiales, diarios de campaña (img. 5, 7, 8, 9, 21 y 22).
- b) Manifestación de opiniones particulares, escritas generalmente de forma rápida y poco elaborada: documentos cuya finalidad es manifestar puntos de vista individuales sobre acontecimientos (img. 13 y 16), reforzar las relaciones personales —amistades, familiares o amorosas (img. 15 y 17)— en tiempos de dificultad política o contienda bélica o dar avisos, frecuentemente en forma de anónimos (img. 11).
- c) Difusión de nuevas ideas por medios públicos y generalmente impresos (libelos y prensa), (img. 1 y 2).
- d) Expresión de las relaciones familiares y privadas en el contexto de la insurgencia, tales como cartas, poemas, billetes, etc. (img. 23 y 24).

³ La peculiar configuración factual del corpus permite eludir la tan cuestionada *teleología invertida* (Oesterreicher 2007) de los corpus históricos.

⁴ En el sentido definido por Coseriu (1973 [1955-1956]: 318).

Los fondos ya incorporados al corpus pertenecen a las siguientes instituciones: el Archivo General de Indias (AGI), de Sevilla; el Archivo Histórico Nacional (AHN) de Madrid; y el Archivo del Instituto de Historia y Cultura Militar (AIHCM) de Madrid.

AGI	660
AHN	120
AIHCM	220
TOTAL	1000

Tabla 1. Archivos y número de documentos. Elaboración propia.

Debido a su propio origen, el análisis de este material puede encontrar dificultades específicas (la documentación puede estar desordenada o incompleta; la identificación de autor o datación precisa, tópica o crónica, puede no ser accesible en algunos casos); a pesar de ello, es indudable que estos documentos conforman un corpus óptimo para la investigación filológica: el corpus resultante constituye una construcción coral, que no obedece a un diseño o propósito previo deliberado, refleja un *continuo concepcional* con una masa de producción suficiente para la reconstrucción de un amplio ámbito de uso de la lengua, que se extiende entre los polos extremos de informalidad o inmediatez comunicativa y formalidad o distancia comunicativa (Koch/Oesterreicher 1985, 1994, 1998); son documentos auténticos (ciertos), compuestos con la espontaneidad de quien supone que no va a ser leído más allá del receptor, redactados en muchos casos en condiciones de apremio y sin corrección posterior. En síntesis, el análisis de un corpus de estas características permite alcanzar como resultado una imagen extraordinariamente vívida y ceñida a la realidad tanto de la lengua general como de su estratificación social.

2.1. Selección de documentos y tratamiento

La selección y tratamiento de los manuscritos se ha hecho conforme a criterios paleográficos y filológicos rigurosos, utilizando siempre documentación original e inédita. El corpus documental resultante ha sido transcrito y está siendo editado siguiendo los principios propuestos por la Red CHARTA⁵, que permite obtener una edición paleográfica además de una edición crítica más adecuada para la difusión entre un público no experto. Dada la peculiaridad de la documentación, se ha diseñado una ficha de descripción pormenorizada que permitirá identificar materiales según distintos aspectos escriturarios y procedimentales⁶; esta descripción servirá de base para la elaboración de la tipología mencionada. En este proceso se atenderá a todos los elementos que propicien la *recontextualización*⁷ de los elementos del corpus.

El corpus documental será sometido a un proceso de marcación digital mediante el lenguaje TEI (*Text Encoding Initiative*). TEI permite la representación estructural y semántica de diversos fragmentos de un texto, mediante un gran número de etiquetas. Esto ayuda al editor a describir y procesar documentos de manera rigurosa y eficiente, adaptándose a las necesidades que requiera cada corpus (Cummins 2008).

⁵ Accesible desde: <https://www.redcharta.es>.

⁶ Incluyendo tres marcas de correlación fundamentales: 1) tipo de documento, definido en un contexto de evolución histórica, es decir, considerando los formatos tradicionales, sus transformaciones y la eventual aparición de nuevos formatos; 2) nivel sociolingüístico del autor (ya sea conocido independientemente del documento analizado o únicamente deducible a partir de él); 3) alcance de recepción del documento: oficial, social, personal, familiar o íntimo.

⁷ “[P]roceso de reconstrucción de las diversas relaciones semióticas del texto, por parte del observador” (Oesterreicher 2001: 212.). Oesterreicher incide en el que hecho de que disponemos de una documentación histórica de-contextualizada y de-especificada. “Por eso, es preciso someter al texto escrito a una recontextualización para disminuir las ‘pérdidas’ mencionadas y recuperar, a partir del texto, un máximo del contexto discursivo originario” (2001: 212-213).

Por ejemplo, la marcación semántica puede ayudar a recopilar en el corpus de documentos del siglo XIX información sobre personas, lugares o acontecimientos históricos que permitan desambiguar conceptos, enriquecer los elementos o establecer relaciones entre ellos y, además, realizar trabajos de investigación sobre prosopografía, toponimia, antroponimia, léxico específico, etc. Toda esta información se puede visualizar en índices que, a su vez, permiten realizar búsquedas en la totalidad del corpus.

En este proyecto la visualización y edición de los datos, así como su procesamiento lingüístico, se realizará a través de TEITOK, una plataforma en línea creada por Maarten Janssen (2014) y pensada para ver, crear y editar corpus que presentan simultáneamente marcación textual y anotación lingüística (<http://www.teitok.org/index.php?action=about>).

Este tratamiento se complementa con el uso de *Atlas.ti* una herramienta de apoyo para el estudio de investigación cualitativa que aplicamos a manuscritos y, especialmente, a prensa e imágenes.

3. Desarrollo de la investigación

La investigación pone el foco en los siguientes aspectos:

1. Descripción y evaluación en la cadena variacional de rasgos lingüísticos singulares (*Varietätenkette*, Koch/Oesterreicher 1990 [2006]), eje que sirve para ligar internamente las distintas variedades de una lengua en torno a una dirección tal que lo diatópico puede integrar a la variación diastrática, y esta a su vez a la diafásica, pero no en una dirección inversa. Para ello se elegirán aquellos fenómenos que (a) describen tendencias o se convierten en elementos identitarios del español cubano; (b) permiten esbozar una sociolingüística del español en la isla y (c) admitan la aplicación de parámetros de resistencia o variación respecto a la normalización de otras zonas hispánicas. La composición del corpus permite contar con escritores potencialmente protagonistas de *cambios desde abajo* (Lavob 1994), cuya actuación lingüística se da en formatos documentales preestablecidos.

La investigación buscará las correlaciones entre usos lingüísticos y discursivos y rasgos etnolingüísticos, con el fin de determinar la existencia o ausencia de dichas correlaciones y, en su caso, establecer las adscripciones etnolingüísticas que sean pertinentes. Este proceso permitirá también detectar procesos de innovación y el grado de penetración de los mismos en el español cubano de la época. Para cumplir este punto, así como el que se expone a continuación, se atenderá a los procesos de *graficación* y *escrituralización* (Oesterreicher 2001: 202) ya que el corpus recoge un buen número de documentos, manuscritos o impresos, que han sido dictados, otros han sido concebidos para lectura —en público o en privado— o incluso la declamación.

La Tabla 2 ofrece un resumen de los manuscritos oficiales ya catalogados y editados.

Documentos oficiales	civiles	militares
Informes	16	25
Proclamas	12	18
Cartas	150	145
Diarios de operaciones	0	15
Reglamentos y normativas	10	20
Órdenes y comunicados	31	48

Tabla 2. Documentos oficiales. Elaboración propia.

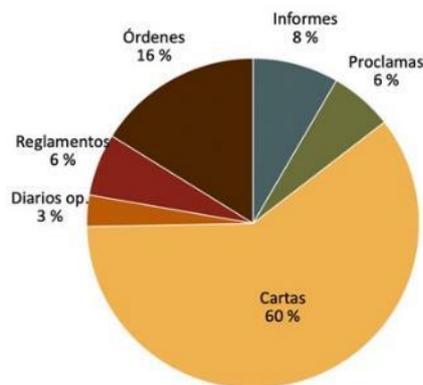


Figura 2. Porcentaje de documentos oficiales según tipo. Elaboración propia.

2. Especialmente interesante en esta masa documental es la correspondencia, que muestra un extenso rango de finalidad y formato: es el medio del que combatientes de uno y otro bando se sirven para comunicaciones de muy diversa índole (transmisión de órdenes militares, de nombramientos, saludo entre mandos, muestras de amistad o amor, etc.). La carta es un tipo muy utilizado en los siglos XVI y XVII, que exhibe un alto grado de flexibilidad para adaptarse en cada época a la diversa tipología de emisor-receptor y contenido; si a ello se añade el aumento de su difusión y la amplificación de sus funciones en época moderna (Castillo Gómez 2005), se sigue que el género epistolar conforma una sólida tradición discursiva, que permite estudiar la contigüidad en la historia del español.

A través de este material, aunque no exclusivamente en él, se diseñará un mapa relativo a las fórmulas de tratamiento y cortesía —pronombres, expresiones y formas verbales— que permitirán observar la articulación de la nueva sociedad, a través de la detección de procesos de rutinización expresiva y de innovación.

3. Una fracción significativa del corpus está constituida por documentos diseñados para ser transmitidos oralmente a sus destinatarios finales; el examen del proceso de oralización como parte de un continuo (que abraza desde el acto escrito hasta su transmisión oral) permite evaluar aspectos tan diversos como fidelidad a usos ortográficos, presencia de formas dialectales o creación de expresiones inéditas en la historia de la lengua. Dado que una buena parte de ellos son textos diseñados para incitar a la reflexión y la acción (no dialógicos), su análisis debe incluir el contexto social en el que se inserta y su intencionalidad práctica concreta.

4. Muchos de los materiales pertenecen a la categoría de fuentes históricas denominada *egodocumentos* (Presser) o *autodocumentos* (*Selbtszeugnisse*) Schulze considerando como tales memorias, diarios, cartas personales y textos afines en los que el autor expone de forma manifiesta sus circunstancias vitales y sentimientos⁸. El análisis y tratamiento desde este enfoque permitirá no solo ofrecer un material inédito para el estudio de historiadores⁹, sino disponer de un conjunto sólido para analizar las voces de sus protagonistas y, a partir de ellas, extraer los elementos que hacen posible realizar estudios de la lengua de los líderes, nuevos conceptos políticos y filosóficos, valores sociales, etc.

⁸ "el criterio común para designar textos como ego-documentos (*Ego-Dokumente*) debería ser el de aquellos que, libremente o de manera involuntaria, dan información personal de un ser humano, sobre su familia, su comunidad, su país, informan sobre su punto de vista social o reflejan su relación con estos sistemas y sus transformaciones. Deben justificar sus actitudes individuales y humanas, manifestar sus miedos, mostrar sus valores, reflejar sus experiencias vitales y sus esperanzas" (Schulze 2005: 110).

⁹ Sobre esta tipología, únicamente se ha publicado el trabajo de Aristizábal (2012) consistente en una recopilación de ediciones disponibles en Alemania.

La Tabla 3 muestra los egodocumentos localizados y editados.

Egodocumentos	
Diarios	22
Cartas	385
Notas y billetes	54
Anónimos	25
Relaciones e informes	24

Tabla 3. Egodocumentos. Elaboración propia.

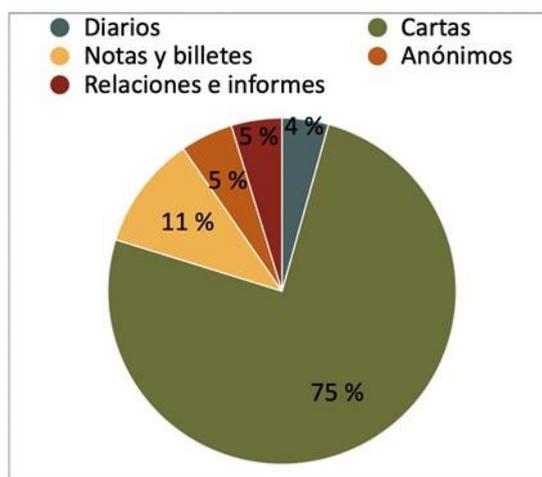


Figura 3. Porcentaje de egodocumentos según tipo. Elaboración propia.

4. Resultados

El proyecto ofrecerá a partir de marzo de 2022 resultados de carácter teórico, como la elaboración de una clasificación de la tipología documental de los documentos (manuscritos e impresos) o la identificación de fenómenos sociolingüísticos y culturales con contigüidad Andalucía-Cuba.

Como resultado ecdótico y textual está prevista la edición paleográfica y crítica de los documentos que conforman el corpus, así como la elaboración de una base de datos de consulta de léxico cultural convergente. Sobre la base de los materiales del proyecto, se realizarán análisis específicos; entre ellos está previsto:

- el estudio de marcadores sociales y familiares como las fórmulas de tratamiento;
- el análisis de procesos de creación léxica y resemantización relativa a conceptos culturales emblemáticos de la época;
- estudios particulares sobre la tipología documental (cartas, prensa, diarios, etc.).

Además de la presentación de resultados científicos en el ámbito académico, los resultados de divulgación quedarán disponibles a través de: a) una web del proyecto que incluirá acceso a los contenidos esenciales del corpus de trabajo; b) una base de datos con el corpus textual; c) una base de datos con el léxico y los

marcadores culturales definidos. Para un público más amplio se elaborarán libros electrónicos y materiales en soporte digital para la difusión y con descarga accesible. Con ello esperamos contribuir al conocimiento de la sociedad cubana del XIX y muy especialmente a los lazos culturales familiares con Andalucía.

5. Referencias citadas

- ARSITIZÁBAL, C. (2012). *Autodocumentos hispanoamericanos del siglo XIX. Fuentes personales y análisis histórico*, Berlin: LIT.
- CARAVEDO, R. (1999). "Lingüística del Corpus. Cuestiones teórico-metodológicas aplicadas al español". En J. de Koch (ed.) *Gramática española. Enseñanza e investigación. Apuntes metodológicos*, t. 1, vol. 6, Salamanca: Universidad de Salamanca.
- CASTILLO GÓMEZ, A. (2005). "«El mejor retrato de cada uno» la materialidad de la escritura epistolar en la sociedad hispana de los siglos XVI y XVII". *Hispania*, 65(221), 847-875.
- CUMMINGS, J. (2008). "The Text Encoding Initiative and the Study of Literature". En R. Siemens y S. Schreibman (eds.): *A Companion to Digital Literary Studies*, Oxford, Blackwell.
- KOCH, P. y W. OESTERREICHER (1985). "Sprache der Nahe- Sprache der Distanz. Mündlichkeit und Schriftlichkeit im Spannungsfeld von Sprachtheorie und Sprachgeschichte". *Romanistisches Jahrbuch* 36, 15-43.
- KOCH, P. y W. OESTERREICHER (2007 [1990]). *Lengua hablada en la Romania: francés, italiano, español*, Madrid: Gredos.
- LABOV, W (1994). *Principles of Linguistic Change, I: Internal Factors*, Oxford: Blackwell.
- JANSSEN, M. (2014). *TEITOK. A Tokenized TEI environment*. <http://teitok.corpuswiki.org/site/index.php>
- OESTERREICHER, W. (2001). "La 'recontextualización' de los géneros medievales como tarea hermenéutica". En D. Jacob y J. Kabatek (eds.), *Lengua medieval y tradiciones discursivas en la Península Ibérica. Descripción gramatical – pragmática histórica – metodología*, Frankfurt am Main/ Madrid: Vervuert/Iberoamericana, 199–231.
- PERL, M. (2001). "Die Sprachsituation in Kuba". En O. Erre y M. Franzbach (eds.). *Kuba heute. Politik, Wirtschaft, Kultur*. Frankfurt am Main/Madrid: Vervuert/Iberoamericana, 653-668.

Pygmalion: una herramienta para la visualización interactiva del cambio semántico desde una perspectiva diacrónica

Helena Bermúdez Sabel
Institut des sciences du langage, Université de Neuchâtel

Esta contribución presenta [Pygmalion](#),¹⁰ una herramienta que facilita la creación de mapas semánticos diacrónicos interactivos. Los desafíos abordados conciernen la elaboración de visualizaciones que puedan transmitir conceptos complejos, como la descripción de significados y el cambio semántico, así como el diseño de recursos que permitan crear este tipo de visualizaciones de una manera intuitiva. Este recurso ha sido desarrollado dentro del marco del proyecto *A World of Possibilities. Modal pathways over an extra-long period of time: the diachrony of modality in the Latin language (WoPoss)*¹¹ dirigido por Francesca Dell’Oro y que cuenta con la financiación del Fondo Nacional Suizo (FNS n° 176778).

Un mapa semántico es una representación geométrica de funciones interconectadas en el espacio semántico (Haspelmath 2003: 213). Con más de cincuenta años de historia,¹² los mapas semánticos han demostrado su versatilidad a la hora de condensar mucha información sobre configuraciones semánticas, incluyendo aspectos relacionados con cambios de significado (ver Fig. 1).

A pesar de estar basado en propuestas anteriores (Hjelmslev 1963; Lazard 1981; Anderson 1982; Traugott 1985; Croft et al. 1987; Haspelmath 2003; François 2008), nuestro modelo intenta integrar el aspecto lexicográfico a la representación de las relaciones semánticas. Así pues, este modelo presenta varios aspectos novedosos: 1) desde el punto de vista de las dimensiones representadas, nuestro diseño hace hincapié en el aspecto cronológico: no solo permite codificar visualmente las fechas exactas (o aproximadas) de la emergencia o desaparición de un significado, sino que también es posible enriquecer el mapa añadiendo información relevante como la primera atestación conservada de un significado (Fig. 2);¹³ 2) desde el punto de vista de la representación visual, ha sido concebido dentro del ecosistema digital, por lo que incluye indicios visuales que se activan a partir de las acciones del usuario; 3) otra novedad de nuestro diseño es que ofrecemos dos visualizaciones complementarias del espacio semántico: en una de ellas, el espacio semántico se activa a partir de significados específicos que se hacen visibles a la derecha del mapa; en la segunda visualización, el espacio semántico se configura a partir de un algoritmo basado en fuerzas (*force directed layout*). Cabe destacar también, que nuestro modelo también se puede adaptar a la representación de los significados de una palabra o función desde una perspectiva monolingüe o multilingüe (igual que lo habían hecho otros modelos como François 2008).

¹⁰ Ver <<https://woposs.unine.ch/pygmalion.php>> (fecha de consulta: 20/04/2021).

¹¹ Ver <<https://woposs.unine.ch>> (fecha de consulta: 20/04/2021).

¹² Ver Hjelmslev 1963 para una primera conceptualización.

¹³ Ver la versión interactiva del mapa diacrónico en <<https://woposs.unine.ch/semantic-maps/map-facultas.php>> (fecha de consulta: 20/04/2021).

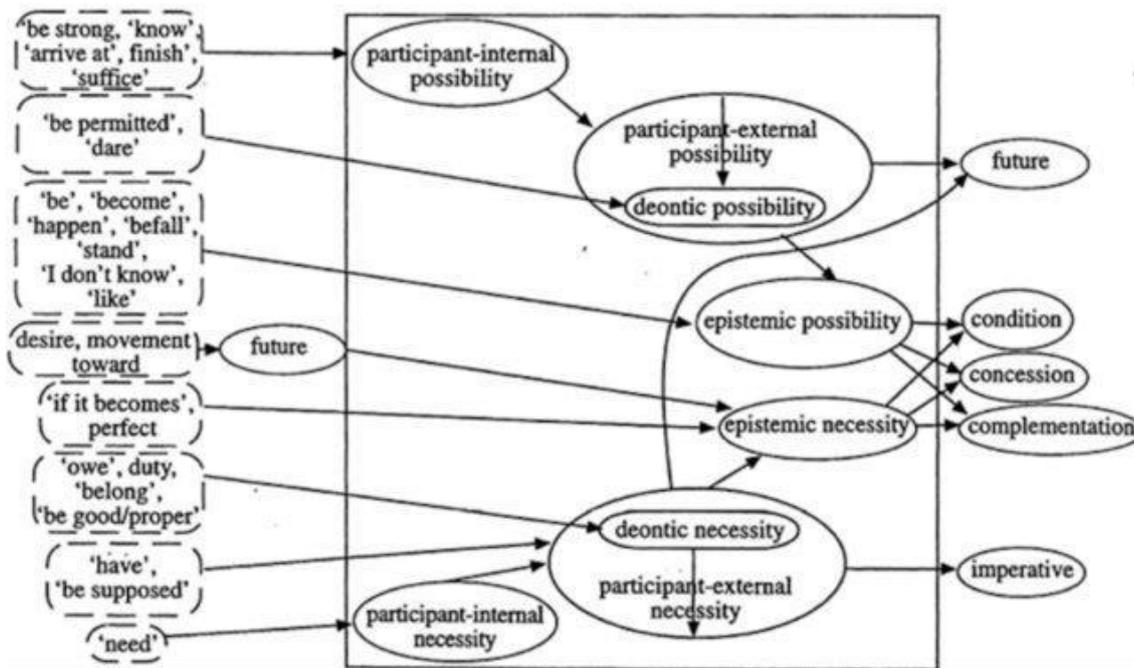


Fig. 1. "Unifying the possibility and necessity paths": Ejemplo de un mapa semántico que representa la evolución de la configuración de la posibilidad y la necesidad (van der Auwera and Plungian 1998: 98)

Pygmalion ofrece tres versiones diferentes: [Pygmalion-simple](#), la versión más básica de la herramienta con la que crear mapas semánticos diacrónicos (Dell'Oro et al. 2020b); [Pygmalion-colour](#), versión que permite añadir una dimensión a los mapas básicos que es codificada visualmente con el uso de colores (el usuario escoge cuál será la nueva dimensión, como por ejemplo, la lengua o la función de significados) (Dell'Oro et al. 2021b); y, por último, [Pygmalion-modal](#), concebido para el diseño de mapas diacrónicos modales (Dell'Oro et al. 2020a).

A través de un formulario (específico a cada una de las versiones de la herramienta), Pygmalion recoge toda la información que será representada en la visualización final. Esta visualización, además, se presenta en dos formatos diferentes: uno de ellos utiliza el eje horizontal para expresar la información cronológica y el otro consiste en un grafo de redes.

Los datos solicitados en la primera parte del formulario son:

- Lema (obligatorio)
- Etimología (optativo)
- Por cada significado o función que se añada:
- El significado, función o utilización que se describe (obligatorio)
- Análisis (disponible en Pygmalion-color) o descripción modal (disponible en Pygmalion-modal)
- Colocación / expresión (optativo)
- Grupo (semántico o una asociación de otro tipo) (optativo)
- Fecha¹⁴ de aparición (obligatorio)
- Fecha de desaparición (optativo)
- Primera atestación (optativo)

¹⁴ Las fechas pueden introducirse en diferentes formatos: años, décadas o siglos.

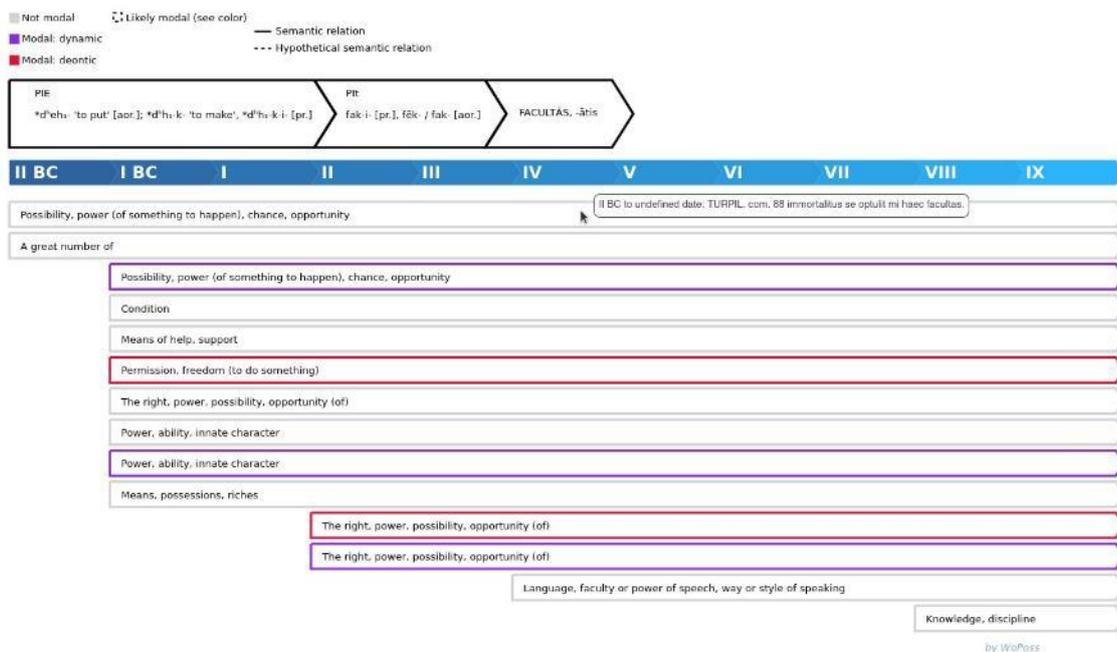


Fig. 2. Ejemplo de un mapa semántico diacrónico (modal) del lema latino *facultas* diseñado con Pygmalion (Marongiu & Dell’Oro 2021)

Una vez que se introducen todos los significados o funciones, la segunda parte del formulario permite determinar relaciones entre ellos, pudiendo establecer la dirección de la relación.

Todos los datos recogidos en el formulario son incluidos en la visualización. Se trata, pues, de una visualización a través de la cual se sintetiza mucha información, por lo que uno de los desafíos consiste en configurar visualmente los datos de una manera que no abrume al lector. Nuestra solución al problema es la de habilitar la visibilidad de determinada información según la demanda del usuario. A continuación se describen las funcionalidades que dependen de la interacción del usuario en la visualización cronológica horizontal (Fig. 2):

- La primera atestación de cada significado se hace visible cuando el usuario pasa el ratón por encima de la flecha correspondiente.
- Las colocaciones y grupos se configuran cuando el usuario selecciona la opción deseada en el menú desplegable. Este menú ofrece las opciones de ordenación posibles (por defecto, el mapa diacrónico presenta los significados ordenados cronológicamente).

Al hacer click en un significado, las relaciones del mismo se hacen visibles (los significados no relacionados desaparecen y las flechas que establecen la dirección de la relación se materializan a la derecha). Para tener una visión de conjunto de todas las relaciones, la visualización bajo la forma de un grafo de redes es más adecuada.

Un mapa semántico diacrónico es un recurso útil para la investigación en lingüística histórica, especialmente para la difusión de resultados, pero también puede ser utilizado en el entorno educativo para transmitir conceptos complejos de una manera didáctica gracias a su lectura fácil e inmediata. Debido a la abundante información que puede codificarse en una visualización de este tipo, el soporte digital se presenta como el medio más adecuado para poder diseñar un modelo de mapa semántico que sea legible y, por tanto, una visualización realmente eficaz como fuente de información lingüística. Con el fin de habilitar la creación automática de mapas semánticos diacrónicos e interactivos, se ha desarrollado Pygmalion. Esta herramienta, de fácil utilización, permite al usuario

organizar de una manera intuitiva la información a incluir en la visualización a través de un formulario. El usuario tiene el control sobre el volumen de datos que quiere introducir, y los campos de ayuda y mensajes de error específicos le ayudan a determinar los campos obligatorios y los optativos. Además, la conceptualización de la herramienta en tres versiones diferentes ayuda a que el proceso de entrada de datos esté más dirigido en relación al producto final, facilitando, por tanto, su uso.

Bibliografía

- Anderson, L. B. (1982). The 'perfect' as a universal and as a language-particular category. *Tense-aspect: Between semantics & pragmatics*. Amsterdam: Benjamins, 227-264.
- Croft, W., Shyldkrot, H. B.-Z., & Kemmer, S. (1987). Diachronic semantic processes in the middle voice. *7th international conference on historical linguistics*, 48, 179-192.
- Dell'Oro, F., Rimaz, L., Bermúdez Sabel, H., & Marongiu, P. (2020a). *Pygmalion-modal 1.0. A tool to draw interactive and diachronic semantic maps of modality*. WoPoss. A World of Possibilities. Swiss National Science Foundation. <https://woposs.unine.ch/pygmalion-modal.php>
- Dell'Oro, F., Rimaz, L., Bermúdez Sabel, H., & Marongiu, P. (2020b). *Pygmalion-simple 1.0. A tool to draw interactive and diachronic semantic maps*. WoPoss. A World of Possibilities. Swiss National Science Foundation. <https://woposs.unine.ch/pygmalion-simple.php>
- Dell'Oro, F., Rimaz, L., Bermúdez Sabel, H., & Marongiu, P. (2021b). *Pygmalion-colour 1.0. A tool to draw interactive and diachronic semantic maps of modality*. WoPoss. A World of Possibilities. Swiss National Science Foundation. <https://woposs.unine.ch/pygmalion-colour.php>
- François, A. (2008). Semantic maps and the typology of colexification: Intertwining polysemous networks across languages. En *From polysemy to semantic change: Towards a typology of lexical semantic associations* (pp. 163-215). <https://halshs.archives-ouvertes.fr/halshs-00526845/document>
- Haspelmath, M. (2003). The Geometry of Grammatical Meaning: Semantic Maps and Cross-Linguistic Comparison. En *The New Psychology of Language* (pp. 217-248). Psychology Press. <https://doi.org/10.4324/9781410606921-11>
- Hjelmslev, L. (1963). *Prolegomena to a theory of language*. University of Wisconsin Press. Marongiu, Paola; Dell'Oro, Francesca (2021). "*facultas*". v.1.0. WoPoss. <https://woposs.unine.ch/semantic-maps/map-facultas.php>
- Lazard, G. (1981). La quête des universaux sémantiques en linguistique. *Actes Sémiotiques (Bulletin)*, 19, 26-37.
- Traugott, E. C. (1985). Conditional markers. *Iconicity in syntax*, 289-307.
- van der Auwera, J., & Plungian, V. A. (1998). Modality's semantic map. *Linguistic Typology*, 2(1), 79-124.

Construcción de un corpus de evaluación de la anotación léxico-gramatical del castellano medieval (siglos XIII-XV)

Matthias GILLE LEVENSON

Casa de Velázquez, ENS de Lyon – CIHAM/UMR 5648
matthias.gille-levenson@casadevelazquez.org

Olivier BRISVILLE-FERTIN

Université Lyon 1 – CIHAM/UMR 5648 olivier.brisville-
fertin@univ-lyon1.fr

María DÍEZ YÁÑEZ

Universidad Complutense de Madrid mariadiezy@ucm.es

Simon GABAY

Université de Genève simon.gabay@unige.ch

Palabras clave: evaluación, lematización, anotación gramatical, castellano medieval, Freeling

Introducción

La utilización de herramientas de etiquetado gramatical (lemas y partes del discurso) es fundamental para analizar y procesar textos, sea para el análisis del discurso [Allés Torrent 2021], con herramientas como TXM por ejemplo [Quignard y col. 2019], colación automatizada [Camps, Ing y Spadini 2019], estilometría, reconocimiento de autor o estudios de historia de la lengua [Sánchez Marco 2012, Ueda, Borja y Sandoval 2020], entre otros.

De ahí que una anotación la más perfecta posible sea necesaria, pues de ella depende la producción de una edición crítica, el análisis filológico y lingüístico. Pero, ¿cómo evaluarla? La cuestión no es sencilla, puesto que no solo la evaluación en sí no es nada fácil sino que varía además entre herramientas. En efecto, un mismo útil bien puede resultar ser más eficiente con ciertos estados de lengua o géneros: la evaluación ha de tener en cuenta la heterogeneidad del corpus. Asimismo, ciertas herramientas por naturaleza no son generales, sino que necesitan datos de entrenamiento (en particular las vinculadas con redes de neuronas): por lo tanto resulta fundamental tener en cuenta tal parámetro en el proceso de evaluación.

En efecto, importa garantizar la ausencia de imprecisiones que, debido a la escala y sin retrospectiva, podrían aparecer como verdades lingüísticas a ojos de usuarios que no sean conscientes de posibles sesgos. Una evaluación permitiría identificar posibles correcciones, así como la utilización de porciones corregidas, corpus llamado «*gold*» para entrenar nuevas herramientas complementarias para tareas precisas (por ejemplo, anotación en diacronía o de *scriptae* complejas). Siguiendo el *workflow MATTER* (*Model, Annotate, Train, Test, Evaluate, Revise*) [Pustejovsky y Stubbs 2013, pág. 24], la anotación mejorada de manera iterativa se convierte tanto en un fin –la producción de corpus– como en un medio –entrenar modelos– que se alimentan el uno y el otro.

Así pues, este trabajo radica en la concepción de un indicador para estimar la eficacia de la herramienta de lematización actualmente disponible y así ayudar en el futuro la investigación en lingüística de corpus y en filología computacional. Dicho indicador tendrá que adoptar, en una primera etapa, la forma de un corpus abierto representativo del castellano bajomedieval (s. XIII-XV), antes de ampliarlo a otros estados de lengua.

1. Estado de la cuestión

1.1. Herramientas para el castellano bajomedieval

No conocemos otra herramienta de acceso libre de anotación gramatical para el castellano medieval que Freeling. Freeling es un proyecto maduro, de más de veinte años, desarrollado esencialmente por Lluís Padró, y que funciona con muchas lenguas. La lematización y el análisis morfológico son tan solo una pequeña parte de sus posibilidades de análisis. Desde 2010-2012, Freeling incluye la posibilidad de trabajar con corpus en castellano medieval, gracias al trabajo de Cristina Sánchez Marco [Sánchez Marco, Boleda y Padró 2011]. Freeling está escrito en C++, con un API python también disponible.

La adaptación de Freeling para el castellano antiguo se efectuó a partir de 2009-2010, en dos fases: en la primera, se preprocesaba el texto para trabajar con una versión modernizada [Sánchez Marco, Boleda, Fontana y col. 2010]. Las dificultades técnicas y las pérdidas filológicas hicieron que se prefiriera adaptar el etiquetador (*tagger*) para el tratamiento de un corpus de textos antiguos (s. XII-XVI) [Sánchez Marco, Boleda y Padró 2011] (tal adaptación se integró en Freeling a partir de la versión 3.0 [Padró y Stanilovsky

2012]). Gracias a un subcorpus de entrenamiento (*Gold Standard Corpus*) de 30 000 formas (*tokens*) y al establecimiento de una serie de reglas, se integraron variantes y se amplió el diccionario del español estándar con unas 32 000 entradas [Sánchez Marco, Boleda y Padró 2011, págs. 4-8].

1.2. Resultados anunciados por Freeling

Los resultados anunciados por Freeling son los que se presentaron en 2011 [ibíd.] y se integraron en la versión 3.0 [Padró y Stanilovsky 2012]. La evaluación de la precisión o exactitud para el español antiguo se hizo con respecto a los resultados para el español estándar. La ampliación del diccionario y el entrenamiento del módulo *Hidden Markov Model PoS Tagger* de Freeling con el *Gold Standard Corpus* permitieron ganar un promedio de 19 puntos, es decir, alcanzar un 92,6 % de exactitud léxica para el lema asignado, un 89,9 % para el etiquetaje morfológico detallado (PoS-1: etiqueta de información de 6 caracteres) y un 94,5 % para el etiquetaje morfológico simple (PoS-2: etiqueta de 1 carácter). Pese a no alcanzar los niveles para el español estándar (99,1 %, 94 % y 97,6 % respectivamente), los resultados anunciados ya eran de lo más satisfactorios. Sobre los errores, el análisis da cuenta de dos causas: la ambigüedad (morfosintáctica o por falta de acento gráfico), que explica un 90 % de los errores, y palabras que no se habían vertido al diccionario.

Es de precisar que Sánchez-Marco siguió trabajando sobre la lematización para su tesis de doctorado, utilizando también el segundo módulo lematizador que propone Freeling (*Relaxation Labelling PoS Tagger*). Dio resultados actualizados y mejorados, que reducían aún la diferencia con el español estándar contemporáneo: con ambos módulos consiguió alcanzar un 95,8 % de exactitud para los lemas; un 90,1 % con *hmm* y 92,6 % con *relax* para el PoS-1; y un 94,5 % y un 95,3 % para el PoS-2 [Sánchez Marco 2012, págs. 140-143]. Ahora bien, no se ha podido saber si dichas mejoras se habían integrado en una actualización de Freeling.

Queda necesario un corpus de evaluación que permita producir resultados de evaluación precisos, fuesen cualitativos o cuantitativos: el porcentaje de eficacia media global es significativo, sin ser suficiente: ¿es Freeling

eficaz sobre todos los géneros?; ¿todos los periodos?; a nivel gramatical, ¿todas las categorías gramaticales se analizan con la misma eficacia? ¿Con qué normas de transcripción es más eficiente? Estas métricas permitirán facilitar el uso y la mejora de la lematización del castellano medieval, sea mejorando la propia herramienta Freeling o proponiendo otras herramientas.

2. Metodología

Nuestra posición es la del usuario y de la editora que quieren anotar un texto o un corpus editado según las normas actuales (lo que supone, por ejemplo, acentuar las palabras). Por ello, descartamos del corpus los textos editados hace demasiado tiempo o con modernización excesiva de la lengua.

2.1. Producción del corpus de evaluación

El corpus ideado debe ser representativo de la producción literaria castellana medieval, en términos de siglo de producción y de géneros. Cada texto del corpus es bastante corto: más o menos 200 palabras. El corpus reunido hasta ahora es de más de 150 obras¹⁵; buscamos un equilibrio entre las épocas y los géneros para poder evaluar lo más precisamente posible (por fecha, por género y de manera global), tal y como hacen Camps, Gabay *et al.* en un artículo reciente sobre lematización del teatro clásico francés [Camps, Gabay y col. 2021, págs. 11, 12]. Este corpus en curso de selección y de adquisición (extracción de los textos, lematización, corrección manual de las anotaciones) se presentará detenidamente el día de la conferencia¹⁶.

2.2. Producción de las estadísticas

La metodología global de nuestro estudio es la siguiente: primero, producción del corpus «correcto»; luego, lematización de este corpus por la herramienta que se quiere evaluar (Freeling en nuestro caso); en tercer lugar, comparación de los resultados de la lematización con el texto correspondiente «correcto»; por fin, producción de estadísticas por texto individual y generales (por corpus). Este procedimiento se automatiza mediante un *script*, ya publicado en el repositorio gitlab susodicho.

3. Evaluación de Freeling: primeros resultados

3.1. Resultados cualitativos

Podemos proponer unos primeros resultados cualitativos obtenidos a partir del análisis de la edición (en curso) de la última parte del *Regimiento de los príncipes* por Matthias Gille Levenson, en el marco de su tesis de doctorado¹⁷; estos se podrán completar con el análisis completo del corpus una vez establecido.

Esta edición como primer subcorpus nos parece interesante porque cada manuscrito se transcribe individualmente. Por lo tanto, hemos podido efectuar una primera evaluación cualitativa de dos capítulos de 2.200 *tokens* por cada capítulo –multiplicado por siete manuscritos e incunable (en total, más o menos 30.000

¹⁵ Una lista de las obras así como la repartición por género y por siglo se presentan en el siguiente repositorio Gitlab: https://gitlab.huma-num.fr/mgillelevenson/paper_hdh_lematizacion. El último *commit* antes de someter nuestra contribución es el siguiente: 76bf995.

¹⁶ Para la corrección de las anotaciones, utilizamos la interfaz de post-corrección Pyrrha, desarrollada por el equipo de la École nationale des Chartes de París [Pinche 2019].

¹⁷ Este texto tiene su interés en la medida en que mezcla los géneros y las fuentes, entre materia técnica y militar sacada de Egidio Romano, ejemplos históricos compilados del *Communiloquium* de Juan de Gales y materia narrativa prosificada a partir del *Libro de Alexandre*.

tokens)– que permite tener una idea de la influencia de la variación gráfica en la eficacia de la anotación de Freeling.

3.1.1. Lematización

Nos ha llamado la atención algunos errores recurrentes y que consideramos estructurales; por ejemplo, algunas formas verbales no se reconocen: así, los imperfectos de indicativo del primer grupo, grafiados con <v>(«mandavanlos», «tiravan», «dava», etc); algunos condicionales no se reconocen correctamente («matarían», «farías»). A nivel gráfico, algunos grafías suelen dificultar el análisis: una forma grafiada con <v>como «cavallería» no se analiza, pero «caualleria» se analiza correctamente; formas con <rr->geminada inicial («rraviosa», «rrespondió»); formas con <ç>(«artefiços», «profeçión»¹⁸, «reççiba»); formas grafiadas con <y>: «lydiava», «respondyó». Algunos análisis nos parecen propiamente incorrectos: así, se lematiza «defendimiento» en «defensa», o «conquerir» en «conquistar» y «onde» en «donde». Asimismo, «fazienda» se analiza como «hacienda», cuando debería analizarse como «facienda»¹⁹. Cuando el corpus se haya producido enteramente, propondremos un balance de la lematización de formas ambiguas y de homógrafos.

3.1.2. Etiquetado morfológico (*PoStagging*)

Los primeros resultados cualitativos del etiquetado morfológico y de partes del discurso (*Part of Speech*) de Freeling²⁰ nos parecen de muy buena calidad, pero esperamos tener más datos para concluir con más precisión y seguridad. Sin embargo, notamos un análisis incorrecto y recurrente, el del verbo «deber»: por ejemplo, la forma «deve» se analiza como «VMMO2SO», *i.e.*, como verbo principal, en modo imperativo, segunda persona del singular, lo que morfológicamente no es.

3.2. Resultados cuantitativos

Los resultados que se dan no son de momento representativos de la calidad de la herramienta Freeling, dado que se han producido con un solo texto (fruto, como se ha dicho, de la transcripción de seis manuscritos y un incunable, que se distribuyen por el siglo XV, hasta 1494). Los resultados finales se darán para la fecha de la conferencia. Estos resultados se encontrarán en el repositorio cuya dirección se ha dado más arriba, en el directorio *results/*.

Se excluirán los nombres propios de la evaluación tanto cualitativa como cuantitativa en cuanto a lematización; en lo que toca a reconocimiento de los nombres propios (partes del discurso, por lo tanto), consideramos que sí pueden ser un criterio de evaluación de una herramienta de anotación gramatical.

3.2.1. La cuestión de los acentos gráficos y de la normalización

Por ahora podemos decir que funciona peor la lematización con corpus acentuados: puede ser problemático respecto a las pautas generales de la edición actual, en las que se suelen acentuar los textos editados.

	Lema	PoS	Global ²¹
Con acento gráfico	0.067	0.058	0.089

¹⁸ «profesión» se analiza correctamente.

¹⁹ La asignación de los equivalentes modernos a lemas, si bien arcaicos, que están presentes en el *DLE*, no vale para los análisis léxicos diacrónicos.

²⁰ Contrariamente a otros juegos de etiquetas (*tagset*), las etiquetas de partes del discurso y de morfología fusionan en Freeling, puesto que se retoma el juego de etiquetas de EAGLES.

²¹ Este resultado global corresponde a la proporción de análisis incorrectos en general, sea el lema, el PoS o ambos incorrectos.

Sin acento gráfico 0.054 0.039 0.068

Cuadro 1: Resultados con el capítulo 19 (commit 76bf995)

Sin embargo, también se tiene que tener en cuenta las evoluciones tecnológicas y la aparición de herramientas de transcripción automatizada eficaces y de uso fácil, como eScriptorium [Kiessling y col. 2019], que ya se pueden considerar como maduras. Estas herramientas probablemente afectarán las normas editoriales: quizá el acentuar los textos no siga como norma en el futuro (puesto que supondría acentuar los textos de manera un poco artificial, seguramente automáticamente). De hecho, los resultados con el texto del manuscrito A, que es el fruto de una transcripción automatizada mediante eScriptorium y de una regularización minimal, es muy bueno y linda con un 97 % de análisis correctos:

Lema	PoS	Global
0.031	0.033	0.048

Cuadro 2: Resultados con el capítulo 19, ms A. (commit 76bf995)

3.2.2. Lematización

Como se ve, con un corpus no acentuado se llega a un resultado medio de 94.6 % de análisis correctos con un corpus del siglo XV, lo que es un muy buen resultado, por ahora, y corresponde a lo que anunciaban los desarrolladores de Freeling.

3.2.3. Etiquetado morfológico (*PoS* tagging)

El etiquetado morfológico es muy eficiente con este primer corpus, y se alcanza una media de 96,1 % de eficacia. Con el corpus completo podremos proponer resultados por categoría gramatical, filtrando por género textual y por siglo de los textos/manuscritos.

4. Conclusiones

La actual situación de constitución de corpus históricos y necesidad de lematizarlos explica la importancia de las herramientas de anotación automatizada para el filólogo. De ahí surge la necesidad de criterios cuantitativos y cualitativos para evaluar y apreciar la eficacia de tales herramientas.

Para tal propósito, se ha ideado la elaboración de un corpus de evaluación para continuar con bases firmes en la anotación gramatical del castellano medieval. Mediante este corpus, se esperan dos resultados: el primero, detallar la eficiencia de la única herramienta disponible (Freeling), según distintos siglos, géneros y transcripciones; el segundo, percibir los posibles fallos y analizarlos. Así, este balance podrá dar cuenta de las posibles mejoras que se necesiten.

Finalmente, según los resultados y llegado el caso, se podrá comparar dicha herramienta con otras y así considerar los beneficios –o inconvenientes– de un cambio de tecnología en las herramientas de anotación, como el uso del aprendizaje supervisado y de redes neuronales, tendencia importante en el procesamiento actual de otras lenguas históricas [Manjavacas, Kádár y Kestemont 2019].

Referencias

- Allés Torrent, Susanna. «Digital humanities and the Iberian Middle Ages». En: *The Routledge Hispanic Studies Companion to Medieval Iberia*. Routledge, 2021, págs. 327-344. URL: <http://dx.doi.org/10.4324/9781315210483-27>.
- Camps, Jean-Baptiste, Simon Gabay y col. «Corpus and Models for Lemmatisation and POS-tagging of Classical French Theatre». En: *Journal of Data Mining & Digital Humanities 2021* (feb. de 2021). URL: <https://jdmhd.episciences.org/7161>.
- Camps, Jean-Baptiste, Lucence Ing y Elena Spadini. «Collating Medieval Vernacular Texts: Aligning Witnesses, Classifying Variants». En: *DH2019 Digital Humanities Conference*. Utrecht, 2019. URL: <https://dh-abstracts.library.cmu.edu/works/10074>.
- Kiessling, B. y col. «eScriptorium: An Open Source Platform for Historical Document Analysis». En: *2019 International Conference on Document Analysis and Recognition Workshops (ICDARW)*. Vol. 2. Sep. de 2019.
DOI: 10.1109/ICDARW.2019.10032.
- Manjavacas, Enrique, Ákos Kádár y Mike Kestemont. «Improving Lemmatization of Non-Standard Languages with Joint Learning». En: *Proceedings of the 2019 Conference of the North*. Minneapolis, Minnesota: Association for Computational Linguistics, 2019, págs. 1493-1503. DOI: 10.18653/v1/N19-1153. URL: <http://aclweb.org/anthology/N19-1153>.
- Padró, Lluís y Evgeny Stanilovsky. «FreeLing 3.0: Towards Wider Multilinguality». En: *Proceedings of the Language Resources and Evaluation Conference (LREC 2012)*. Istanbul, Turkey: ELRA, mayo de 2012.
- Pinche, Ariane. «Annoter facilement un corpus complexe : l'exemple de Pyrrha, interface de post-correction, et Pie, lemmatiseur et tagueur morphosyntaxique, pour l'ancien français». En: *Rencontres lyonnaises des jeunes chercheurs en linguistique historique*. Lyon, France, jun. de 2019. URL: <https://hal.archivesouvertes.fr/hal-02151796>.
- Pustejovsky, J. y Amber Stubbs. *Natural language annotation for machine learning*. en. OCLC: ocn794362649. Sebastopol, CA: O'Reilly Media, 2013.
- Quignard, Matthieu y col. «Textometric Exploitation of Coreference-annotated Corpora with TXM: Methodological Choices and First Outcomes». En: *Fourteenth International Conference on the Statistical Analysis of Textual Data*. Rome, 2019, págs. 610-615.
- Sánchez Marco, Cristina. «Tracing the development of Spanish participial constructions: An empirical study of semantic change». Tesis doct. Barcelona: Universitat Pompeu Fabra, 2012. URL: <https://www.tdx.cat/bitstream/handle/10803/97044/tcsm.pdf?sequence=1>.
- Sánchez Marco, Cristina, Gemma Boleda, Josep Maria Fontana y col. «Annotation and Representation of a Diachronic Corpus of Spanish». En: *Proceedings of the Seventh International Conference on Language Resources and Evaluation (LREC'10)*. Ed. por Nicoletta Calzolari y col. Valletta, Malta: European Language Resources Association (ELRA), mayo de 2010. URL: http://www.lrec-conf.org/proceedings/lrec2010/pdf/535_Paper.pdf.

Sánchez Marco, Cristina, Gemma Boleda y Lluís Padró. «Extending the tool, or how to annotate historical language varieties». En: *Proceedings of the 5th ACL-HLT workshop on language technology for cultural heritage, social sciences, and humanities*. 2011, págs. 1-9.

Ueda, Hiroto, Pedro Sánchez-Prieto Borja y Antonio Moreno Sandoval. «Lematización y visualización cartográfica del corpus CODEA». En: *Estudios de lingüística del español* 42 (2020), págs. 245-261. URL: <https://www.raco.cat/index.php/Elies/article/view/384939>.

Réthor (Red de Humanidades y Estadística Aplicada al Texto Oratorio): un proyecto interdisciplinar y transversal

Ángela Celis Sánchez (Departamento de Filología Hispánica y Clásica. UCLM)

Víctor Casero-Alonso (Departamento de Matemáticas. UCLM)

María Lozano Zahonero (Dipartimento di Studi Letterari, Filosofici e di Storia del'Arte, Università degli Studi di Roma Tor Vergata)

PALABRAS CLAVE: ANÁLISIS CON R, ANÁLISIS DEL DISCURSO, ANÁLISIS ESTILOMÉTRICO, ANÁLISIS DE SENTIMIENTOS, COMUNICACIÓN NO VERBAL.

KEY WORDS: R SOFTWARE, DISCOURSE ANALYSIS, STYLOMETRY, SENTIMENT ANALYSIS, NONVERBAL COMMUNICATION.

PROBLEMA DE INVESTIGACIÓN

Las humanidades y la estadística son ciencias transversales que han contribuido al progreso de la investigación y el conocimiento en áreas muy diferentes. Además, en la actualidad, el moderno desarrollo tecnológico ha creado un nuevo espacio de interrelación e interdependencia entre ambas disciplinas en el que han surgido nuevos objetos de estudio, interrogantes, enfoques y retos, así como nuevas respuestas a cuestiones clásicas. *Rhétor* ('Red de Humanidades y Estadística aplicada al Texto Oratorio') es un grupo de investigación de reciente creación que se sitúa en este espacio de confluencia con la finalidad de abordar el estudio del texto oratorio en toda su complejidad y variedad.

El ser humano es un ser social. Lo sabemos, como sabemos también que esa naturaleza social se materializa a través del lenguaje, que se pone de manifiesto no solo a través del lenguaje oral verbal sino también mediante toda una serie de señales orales no verbales y otras no verbales y no orales que conforman el denominado *comportamiento lingüístico* y que configuran el significado final del mensaje. Toda alocución lingüística oral conlleva, pues, la participación de una serie de elementos que tienden a ser decodificados en su conjunto, por mucho que tanto emisor como receptor lo realicen inconscientemente. En las últimas décadas, además, se ha afianzado un nuevo paradigma comunicativo, surgido a finales del siglo XX, en el que prima la multimodalidad y la hibridación entre transmisión oral y escrita.

En una primera aproximación podemos definir el texto oratorio como un texto oral planificado de carácter público cuyo objetivo último es la persuasión. El texto oratorio no es solo un producto, sino también un proceso comunicativo de doble dirección basado en la conexión-reacción racional y emocional.

El texto oratorio se configura, así, como una realidad poliédrica. Su estudio, por tanto, no puede ser distinto, de modo que para realizar un análisis crítico debemos atender necesariamente a los siguientes parámetros:

1. El propio texto o discurso, es decir, la expresión textual lineal y el contenido verbal de cada intervención, su género, tipología, estructura, función y forma.

2. El orador y el lenguaje no verbal, conformado por la kinesia, la postura del cuerpo y su orientación, los emblemas, ilustradores, reguladores, señales de afecto, adaptadores, los movimientos y microexpresiones faciales o la mirada, entre otros.
3. La puesta en escena y el entorno, con especial atención a las técnicas y recursos semióticos, comunicativos y persuasivos de carácter multimodal.
4. Los canales de comunicación oratoria, entre ellos las redes sociales y demás medios digitales de comunicación.
5. La construcción y la recepción del discurso oratorio, en particular las reacciones que suscita, todos ellos elementos claves del discurso retórico y ciberretórico, prevalentemente emotivo.

La labor de investigación de *Rhétor* encuentra aplicación en los múltiples contextos en los que está presente el texto oratorio (institucional, político, académico, sanitario, empresarial, publicitario, etc.), además de tener una aplicación didáctica (cursos, conferencias, seminarios, trabajos de fin de grado y de máster, tesis doctorales, talleres "profesionalizantes", etc.)

METODOLOGÍA

Teniendo en cuenta estas premisas, acercarse al estudio del texto oratorio desde cada uno de estos parámetros exige en la actualidad la necesaria confluencia de expertos en diferentes campos de análisis y la adopción de un enfoque en el que las herramientas "clásicas" de la retórica, la lingüística y el análisis del discurso han de ir acompañadas de instrumentos matemático-estadísticos e informáticos que nos permitan extraer, clasificar y visualizar grandes cantidades de datos con objeto de detectar los modelos y patrones subyacentes y sacar a la luz la dimensión emotiva y parte de lo infinitamente grande y lo infinitamente pequeño que queda oculto al ojo humano.

Siguiendo este enfoque, en *Rhétor* nos proponemos realizar un análisis computacional del texto oratorio desde cada uno de los parámetros antes citados, cuyo núcleo fundamental se cimienta en la aplicación de métodos de estilometría y análisis de sentimientos basados en el lenguaje de programación R y su entorno. Este lenguaje de alto nivel, orientado al análisis estadístico de datos masivos en todas sus fases (extracción, preparación, presentación, aplicación), ha adquirido recientemente una notable popularidad en el análisis de fuentes textuales, en particular de datos no estructurados en formatos múltiples o sin formato. Asimismo, se experimentarán y ensayarán nuevos marcos de análisis multimodal basados en inteligencia artificial.

RESULTADOS

Rhétor es un grupo de reciente creación, que surge precisamente de la confluencia de intereses de varios investigadores en relación al análisis del discurso desde los prismas expuestos más arriba. En el Congreso de Humanidades Digitales Hispánicas de 2021 presentaremos nuestras líneas de trabajo, objetivos y método, así como los principales resultados obtenidos, con el fin de ilustrar nuestra labor y ampliar la red de colaboraciones. Hasta el momento nuestro interés se ha centrado, principalmente, en el discurso político institucional, con el análisis de los discursos electorales de las primeras elecciones democráticas de España en 1977 y los discursos institucionales a propósito de la Declaración del Estado de Alarma y las comparecencias oficiales del Presidente del Gobierno en esta situación política y social excepcional. En una fase sucesiva, extenderemos nuestro modelo al análisis de mensajes de reacción ante el discurso institucional en redes sociales (Twitter).

BIBLIOGRAFÍA

- Alayón Gómez, J. (2009). "Retórica y Discurso Hipertextual: Del trovador oral al trovador hipermedial. Notas para un estudio". Biblioteca Virtual Miguel de Cervantes.
- Baayen, H. (2008). *Analyzing Linguistic Data: A Practical Introduction to Statistics Using R*, Cambridge: Cambridge University Press.
- Berlanga Fernández, I., y García García, F. (2014). *Ciberretórica: Aristóteles en las redes sociales. Manual de Retórica en la comunicación digital*. Madrid, Editorial Fragua.
- Colectivo DigiRhet.org (2006). "Teaching digital rhetoric: Community, critical engagement, and application", en *Pedagogy: Critical Approaches to Teaching Literature, Language, Composition, and Culture*, vol. 6, núm. 2, 231-259.
- Eder, M.; Rybicki, J. (2011). "Stylometry with R", *Digital Humanities 2011: Conference Abstracts*, Stanford University, Stanford, CA, 308-311.
- Eder, M.; Kestermont, M.; Rybicki, J. (2013). "Stylometry with R: a suite of tools", *Digital Humanities: Conference Abstracts*, University of Nebraska, Lincoln, NE, pp. 487-89.
- Eder, M.; Rybicki, J.; Kestemont, M. (2016). "Stylometry with R: a package for computational text analysis", *R Journal* 8 (1), pp. 107-121. <<https://journal.r-project.org/archive/2016/RJ2016-007/index.html>>.
- Eggs, E. (1999). "Ethos aristotélicien, conviction et pragmatique moderne", en Amossy Ruth (dir.) *Images de soi dans le discours. La construction de l'ethos*. Paris, Delachaux et Niestlé, 31-59.
- García Berrio, A. (1984). "Retórica como ciencia de la expresividad (presupuestos para una retórica general)", en *ELUA. Estudios de Lingüística*, (2), 7-59.
- Gries, S. (2010). *Statistics for Linguistics with R: A Practical Introduction*, Berlín: De Gruyter Mouton.
- Jockers, M. (2014). *Text analysis with R for students of literature*. Nueva York: Springer.
- Jockers, M. (2015). "Reveling Sentiment and Plot Arcs with the Syuzhet Package". <<http://www.matthewjockers.net/2015/02/02/syuzhet/>>
- Jockers, M. (2017). "Introduction to the Syuzhet Package", en <https://cran.rproject.org/web/packages/syuzhet/vignettes/syuzhet-vignette.html>.
- Jockers, M.; Thalken, R. (2020). *Macroanalysis: Digital Methods and Literary History*, (Topics in the Digital Humanities), Chicago: University of Illinois Press.
- Johnstone, B., y Eisenhart, C. (eds.) (2008). *Rhetoric in Detail*. Amsterdam, John Benjamins Publishing Company.
- Katz, E. y Lazarsfeld, P. F. (1965). *Personal influence: the part played by people in the flow of mass communications*. New York, Free Press.
- Liu, B. (2015). *Sentiment Analysis. Mining Opinions, Sentiments and Emotions*, Nueva York: Cambridge University Press.
- Lutolawski, W. (1898). "Principes de stylométrie", *Revue des Études Grecques*, 11 (41), pp. 6181.
- Menéndez, S. (2012). "Multimodalidad y estrategias discursivas: un abordaje metodológico". *Asociación Latinoamericana de Estudios del Discurso*, Vol: 12 (1), pp. 57-73.

- O'Halloran, K. (2011). "Multimodal Discourse Analysis". En *Companion to Discours*, editado por K. Hyland y B. Paltridge, 120-137. London and New York: Continuum.
- Perelman, C., y Olbrechts-Tyteca, L. (2009). *Tratado de la argumentación: La nueva retórica*. Madrid, Gredos.
- Poria, S.; Hussain, A.; Cambria, E. (2018): *Multimodal Sentiment Analysis*, Socio- Affective Computing Series, Vol. 8, Springer.
- Pujante Sánchez, D. (2003). "La operación dispositivo como base de la construcción del significado discursivo", en Tonos. *Revista Electrónica de Estudios Filológicos* (5).
- Pujante Sánchez, D. (2012). "La operación elocutio, ¿una reina destronable? Su complejo predominio en el discurso retórico", en M. del C. Ruiz de la Cierva y T. Albaladejo (Eds.), *Retórica y política. Los discursos de la construcción de la sociedad*, (175-188). Logroño, Instituto de Estudios Riojanos.
- RSTUDIO TEAM. (2017). *RStudio: Integrated development environment for R*. Boston, MA: RStudio Inc.
- Seargeant, P. y Tagg, C. (eds) (2014). *The language of social media: identity and community on the internet*. London, Palgrave Macmillan.
- Trester, A. M. y Tannen, D. (eds.) (2013). *Discourse 2.0: Language and New Media*. Washington, D.C., Georgetown University Press.
- Vela Delfa, C. (2014). "Estrategias de construcción y mantenimiento de la imagen pública en canales de chat", en *Aposta. Revista de Ciencias Sociales*, núm. 63.
<http://apostadigital.com/revistav3/hemeroteca/cvela1.pdf>
- Zappen, J. P. (2005). "Digital Rhetoric: Toward an Integrated Theory", en *Technical Communication Quarterly*, 14, 3, pp. 319-325.

Panel 1

Proyecto Mnemosine: *Data Reperta et Historia Digitalis*

Modera: Dolores Romero López · *Asiste:* Alexandra Iglesias Cortés

Panelistas: Jeffrey Zamostny; Joaquín Gayoso Cabada; José Miguel González Soriano; Juana María González García

Proyecto Mnemosine: Data Reperta et Historia Digitalis

Coordinadora

Dolores Romero López

(Universidad Complutense de Madrid)

Participantes

1.- José Luis Sierra Rodríguez (Universidad Complutense)
y Joaquín Gayoso Cabada (Universidad Politécnica de Madrid)

2.- José Miguel González Soriano (Universidad Complutense)

3.- Jeffrey Zamostny (University of West Georgia)

4.- Juana María González García (Universidad Internacional de La Rioja)

Palabras clave

Edad de Plata, Biblioteca Digital *Mnemosine*, Historia Digital de la Literatura Española,

Base de Conocimiento

Descripción de la investigación, objetivos, metodología y resultados
--

Mnemosine (Referencia: RTI2018-095522-B-100) es un proyecto de investigación multidisciplinar que se plantea superar los principios historiográficos tradicionales para proponer una 'historia digital' (Romero, 2018) de la llamada otra Edad de Plata (1868-1936) (Ena, 2013) acorde con los datos, materiales y fundamentos teóricos derivados del proceso de digitalización masiva de contenidos llevado a cabo en las últimas décadas. Sus objetivos son: (i) rescatar y ampliar el corpus de creadores, textos, fuentes y géneros 'raros y olvidados' de La otra Edad de Plata; (ii) importar datos y crear metadatos, completarlos, enlazarlos, abrir nuevas fichas y relaciones para desarrollar colecciones específicas dentro de la base de datos, utilizando herramientas que potencien su interoperatividad con otros sistemas y plataformas (Wikidata), así como con distintas instituciones (BDH, HathiTrust-UCM, Biblioteca de California e Instituto Iberoamericano de Berlín); (iii) formular un enfoque teórico sobre la Edad de Plata de la cultura española basado en los principios de la llamada 'historia digital' (Melo, 2011; Pons 2013, 2018) que permita realizar un cambio de paradigma historiográfico y docente; y (iv) facilitar la aplicación práctica de contenidos altamente especializados relacionados con el canon alternativo de la Edad de Plata con el fin de hacer factible su valor educativo en programas docentes y potenciar la difusión de la investigación a través de formatos digitales y multimedia para su conocimiento general (Romero, 2021).

La gestación y desarrollo de *Mnemosine*, Biblioteca digital de La otra Edad de Plata, ha permitido crear un corpus de metadatos a través del que se hacen visibles enfoques históricos y teóricos que responden a desafíos acordes con un mundo digital, plural y abierto siempre a nuevas preguntas. En este proyecto el grupo LOEP-UCM ha colaborado con el grupo ILSA-UCM para adoptar, como soporte a *Mnemosine*, la aplicación *Clavy*, un sistema de

creación y gestión de repositorio de objetos digitales reconfigurable que permite importar, preservar y editar información de colecciones académicas digitales para interoperar entre los repositorios de datos masivos y la creación de colecciones especializadas y enriquecidas con metadatos específicos. *Clavy* presenta un sistema básico de visualización de datos, edición, búsqueda y navegación. Al ser extensible permite la creación de metadatos ajustados a las necesidades de la investigación (Fernández Pampillón, Gayoso, Sarasa y Sierra, 2018). De esta forma, hemos sido capaces de enriquecer la Biblioteca Digital *Mnemosine* con nuevas colecciones que permiten ver mejor y de forma diferente la producción literaria de este período y avanzar hacia la historia digital. El resultado de esta investigación es visible en la web <http://repositorios.fdi.ucm.es/mnemosine/>. Gracias al proyecto arriba indicado estamos rediseñando la interfaz de la Biblioteca Digital *Mnemosien* y queremos presentárselo a la comunidad científica de humanistas digitales en primicia en el Congreso de HDH-2021.

Descripción de las intervenciones de cada participante

Joaquín Gayoso Cabada y José Luis Sierra (Universidad Politécnica y Universidad Complutense, Grupo ILSA)

Clavy: base de datos para el conocimiento experto

Mnemosine ha sido desarrollada con la aplicación *Clavy*, una plataforma experimental para la gestión de colecciones digitales altamente específicas y reconfigurables. *Clavy* incorpora un sistema muy expresivo para la navegación/filtrado/búsqueda, que dota a la herramienta de mecanismos de introspección de colecciones con un gran volumen de información y un gran número de relaciones entre los datos. Asimismo, permite al experto la creación de colecciones de carácter específico basándose en la importación de la información desde diferentes fuentes, edición de la colección para adaptarla a un esquema y propósito esperado, y presentación de la colección resultante. De esta forma, el sistema puede configurarse de manera sencilla para llevar a cabo la importación de información desde diferentes medios, transformación automática o dirigida por el experto de la información importada; y exposición de la información resultante en archivos y servicios que permiten conectarse con fuentes externas. Para ello, la plataforma adopta un enfoque ETL (*Extract, Transform and Load*) para el mantenimiento de colecciones de contenido heterogéneo. De esta forma:

- La aplicación *Clavy* incorpora una arquitectura de *extracción* basada en *plugins*, que permite crear colecciones ingiriendo información desde diferentes fuentes específicas, así como usar *plugins* de extracción genéricos que incorpora ya la propia herramienta (JSON, XML, MYSQL...).
- Partiendo de la colección creada por el proceso de extracción, el experto puede utilizar transformaciones especializadas para la manipulación automática de la colección, gracias a una arquitectura de *transformación*, basada, al igual que la de extracción, en *plugins*. También puede transformar manualmente la colección utilizando el editor incorporado en *Clavy*. Este editor está apoyado en un sistema de exploración que consta de visualización, navegación, filtrado y búsqueda básica para poder realizar la labor de edición.

- Cuando el experto considera oportuno puede utilizar el propio sistema de visualización para publicar la colección en *Clavy*, dar de alta diferentes servicios REST (*Representational State Transfer*) para la exploración de la colección mediante herramientas de visualización externas, o generar archivos de exportación mediante una arquitectura de *carga*, también basada en *plugins*. Estos *plugins* de carga permiten al experto cargar sus propios formatos de salida en fuentes externas, o exportar a formatos estándares mediante *plugins* genéricos ya proporcionados por la herramienta.

La herramienta *Clavy* ha sido desarrollada por el Grupo ILSA (<http://ilsa.fdi.ucm.es/>) gracias a proyectos financiados por la Fundación BBVA (Ayudas a Equipos de Investigación Científica 2014 – área de Humanidades Digitales) y por el Plan Estatal de I+D+i (proyectos RedR+Human, TIN201452010-R, y CetrO+Spec, TIN2017-88092-R).

Enlaces:

<http://clavy.fdi.ucm.es/>

<http://ilsa.fdi.ucm.es/>

José Miguel González Soriano (Universidad Complutense, Grupo LOEP): *La nueva Mnemosine*:

Colecciones y datos enlazados

Una vez pasada –en buena medida– una primera fase de digitalización masiva de fondos por parte de numerosas instituciones culturales, fenómeno que tuvo lugar en los primeros años del presente siglo, todo el vasto conjunto de objetos, recursos y metadatos contenidos en las nuevas bibliotecas electrónicas surgidas desde entonces precisa de ser organizado científicamente, para poder inferir nuevos conocimientos que generen investigación e innovación de calidad. Dentro de la Biblioteca Digital *Mnemosine*, tras considerarse científicamente relevante el volumen de contenidos de su repertorio, compuesto por objetos virtuales divididos en entidades o clases y enlazados entre sí, comenzó la creación en su seno de colecciones temáticas y de autores, definiendo un modelo que hace factible la búsqueda compleja de datos cruzados, modelando la estructura de los metadatos (su *dataset*) de acuerdo a los criterios y necesidades científicas de cada colección; unas necesidades definidas por los usuarios que pueden crear sus propias colecciones importando datos masivamente mediante la aplicación que compone su almacén interno, *Clavy*, la cual posibilita dicha importación –de un modo cada vez más automático– y la edición y creación de nuevos datos o metadatos, gracias a la extensibilidad del modelo que permite nuevos elementos de descripción así como definir nuevas clases o entidades entre sus objetos.

En *Mnemosine*, definiendo una información dentro de unos mismos metadatos entre aquellos objetos virtuales que habrán de componer la colección, como campos → valor de las búsquedas cruzadas en la base de datos, la misma se generará de un modo automático. Así, hemos sido capaces de enriquecer la Biblioteca Digital *Mnemosine* con nuevas colecciones a través de las cuales tenemos una nueva visión de la historia digital de la literatura de La otra Edad de Plata. El modelado de datos permite adaptar los metadatos a la naturaleza diversa de las colecciones, cada una con su genuina especificidad, que hacen visibles –por ejemplo– diversos géneros,

macrogéneros o subgéneros (Colección de Protociencia-Ficción, de Diálogo Literario), ejes temáticos novedosos (Madrid en la Literatura, La Mujer Intelectual, Traductoras de la Edad de Plata, Señas de Identidad Lésbica, Muertos en Conflicto (1936-1939), Exilio Norteamericano), variedad de lectores (Colección de Literatura Infantil, de Literatura de Quiosco), revistas culturales (*Índice Literario*), y monográficos de autores (Colección Carmen de Burgos), entre otras (González Soriano, 2021).

La Biblioteca Digital *Mnemosine* ha sido desarrollada gracias a proyectos financiados por el Plan Estatal de I+D+i (Ref. FFI2012-34666 y Ref. RTI2018-095522-B-I00) y la Consejería de Educación, Juventud y Deporte de la Comunidad de Madrid (H215/HUM-34-26).

Enlaces: <http://repositorios.fdi.ucm.es/mnemosine/>
<https://www.ucm.es/loep>

Jeffrey Zamostny (University of West Georgia, EEUU)

La Colección de Literatura de Quiosco en Mnemosine

En el primer tercio del siglo XX el quiosco fue un nodo para la convergencia y difusión de numerosos objetos impresos: postales, periódicos, revistas ilustradas y –de especial interés para este caso– colecciones literarias de gran tirada y diverso género que se vendían a precios asequibles para el entretenimiento de las masas. Hoy, en el siglo XXI, la Biblioteca Digital *Mnemosine* sirve un propósito similar: almacenar y divulgar textos de La otra Edad de Plata en el gran quiosco digital del Internet. La Colección de Literatura de Quiosco dentro de *Mnemosine* reúne metadatos y textos digitalizados correspondientes a una parte considerable del corpus de colecciones literarias que se multiplicaron en un cuarto de siglo tras el éxito de *El Cuento Semanal* (1907-1912) y que llegaron a publicar una enorme cantidad de textos novelísticos, teatrales, eróticos, políticos y cinematográficos (entre otros géneros y subgéneros), escritos por hombres y mujeres. Según se ha observado en otras ocasiones (Zamostny, 2017; Romero, Bueren y Gayoso, 2017), esta literatura fue excluida del canon literario de la Edad de Plata durante décadas por varios motivos, entre ellos el hecho de “haber cultivado un tipo de literatura no considerada de vanguardia en su época” (González Soriano, 2021). Irónicamente, pues, la recuperación de esta literatura en *Mnemosine* la hace objeto de algunos de los métodos más avanzados hoy por hoy dentro de los estudios culturales y literarios. La lectura distante, el macroanálisis o la minería de textos requieren un amplio corpus de textos y metadatos tales como los que se ofrecen en esta colección.

De momento, la Colección Literatura de Quiosco reúne datos y metadatos sobre aproximadamente trescientos cincuenta colecciones que se publicaron durante la Edad de Plata, la mayoría de ellas representadas por ejemplares físicos dentro de la Colección Ricardo Donoso-Cortés y Mesonero-Romanos de Literatura de Quiosco, albergada desde 2011 en la Universidad de California-Los Ángeles. En un futuro próximo, y con el permiso de UCLA, se espera importar los metadatos para casi 10.500 títulos específicos pertenecientes a estas colecciones, impulsando también la digitalización de los libros impresos. Esta presentación demostrará que, si bien queda mucho por hacer respecto a la importación de datos, con los metadatos ya existentes se puede iniciar el macroanálisis; por ejemplo, se puede usar Voyant para visualizar y analizar los títulos de obras pertenecientes a

colecciones cuyos datos bibliográficos ya están bien desarrollados en *Mnemosine: El Cuento Semanal* (1907-1912), *La Novela Roja* (1922-23) y *La Novela de Hoy* (1922-32).

Enlaces:

<http://repositorios.fdi.ucm.es/mnemosine/buscadorColeccion.php?name=Literatura%20de%20Quiosco&q=tipodecoleccion:quiosco>

Juana María González García (Universidad Internacional de La Rioja) De los metadatos al análisis cuantitativo de Índice Literario

La colección *Índice Literario* da a conocer obras literarias publicadas en España entre 1932 y 1936 reseñadas en la revista *Índice Literario*, principal fruto editorial de la sección “Archivos de Literatura Española Contemporánea” del Centro de Estudios Históricos (1910-1939), dirigida por Pedro Salinas y un grupo de colaboradores (González, 2018). Estas reseñas ofrecen la visión del Centro de Estudios Históricos, de la realidad cultural del periodo republicano. El diseño de los metadatos específicos de la ficha para la colección *Índice Literario* tiene en cuenta los datos bibliográficos de las obras reseñadas (autor, título, editorial, etc.), la ubicación de las reseñas dentro de la revista (sección, página, etc.), los datos de las reseñas (autor, medio, etc.), y los datos sobre la propia revista *Índice* (materia, periodicidad, director, etc.). Este modelo de datos permitirá búsquedas cruzadas de gran interés como, por ejemplo, la de los medios de comunicación que más se interesaron por la obra literaria de mujeres de la Edad de Plata o las editoriales más importantes de la realidad cultural española de los años 30, siempre desde la perspectiva de los colaboradores de *Índice*. Las potencialidades del análisis de los datos que ofrece la revista *Índice Literario* nos han llevado asimismo a un primer análisis cuantitativo de la misma (años 1932-1933) con el lenguaje de programación R (González, 2021), que ofrece conclusiones interesantes como las temáticas a las que dio prioridad la publicación en función del género de las obras reseñadas, su preferencia por la literatura popular, el sesgo ideológico de la mayor parte de las publicaciones o los temas más destacados en las reseñas de obras escritas por mujeres.

Enlaces:

Enlace aún no visible en la web

Bibliografía

Ena Bordonada, Ángela, ed. (2013). *La otra Edad de Plata. Temas, géneros y creadores*, Madrid, Editorial Complutense.

Fernández-Pampillón, Ana; Gayoso, Joaquín; Sarasa, Antonio; y Sierra, José Luis (2018). “De OdA a Clavy, dos aplicaciones orientadas a la creación de repositorios de objetos digitales académicos”. En Romero López, D. y Salamanca López, M. J. (coord.), *Entornos digitales: Humanidades y Ciencias Sociales en la Universidad Complutense de Madrid*, Ciudad de México, Red de humanidades Digitales, 61-83.

- González García, Juana (2018). "La sección de Archivos de Literatura Española Contemporánea (1932-1936) y su labor difusora de la obra creativa de la mujer en la Edad de Plata". *Bulletin of Hispanic Studies*, 95, 959-975. <https://doi.org/10.3828/bhs.2018.56>.
- González García, Juana (2021). "Quantitative analysis of the journal *Índice Literario* (1932-1936)". *Artnodes*, 27. <https://doi.org/10.7238/a.v0i27.374373>.
- González Soriano, José Miguel (2021). "Mnemosine: biblioteca digital de La otra Edad de Plata (orígenes, contenidos, perspectivas)". En Romero López, D. (ed.), "Archivos, repositorios y bibliotecas digitales de la Edad de Plata". *Signa. Revista de la Asociación Española de Semiótica*, 30, 31-58.
- Melo Flórez, Jairo Antonio (2011). "Historia digital". *Historia Crítica* 43, 82-103.
- Pons, Anaclot (2013). *El desorden digital. Guía para historiadores y humanistas*. Madrid: Siglo XXI.
- Pons, Anaclot y Eiroa, Matilde (2018). "Historia digital: una apuesta del siglo XXI". Monográfico. *Ayer* 110.
- Romero López, Dolores (2018). "La Edad de Plata en la tarima digital. Hacia la transdisciplinariedad y la cultura Smart". *Artnodes. Journal on Art, Science and Technology*, Universitat Oberta de Catalunya, 22, 110-117.
- Romero López, Dolores, José Luis Bueren Gómez-Acebo y Joaquín Gayoso Cabada (2017). "Modeling Kiosk Literary Collections for the *Mnemosyne Digital Library*". En Zamostny, J. y S. Larson (eds.), *Kiosk Literature of Silver Age Spain: Modernity and Mass Culture*. Bristol, Intellect, 397-417.
- Romero López, Dolores, ed. (2021). "Archivos, repositorios y bibliotecas digitales de la Edad de Plata". Sección Monográfica I. *Signa. Revista de la Asociación Española de Semiótica*, 30, 17162.
- Zamostny, Jeffrey (2017). "Introduction: Kiosk Literature and the Enduring Ephemeral". En Zamostny, J. y S. Larson (eds.), *Kiosk Literature of Silver Age Spain: Modernity and Mass Culture*. Bristol: Intellect, 1-27.

Sesión 2

La vía de la implementación

Herramientas e infraestructuras

Modera: Pablo Gamallo · *Asiste:* Brianda Otero Moreira

- 23 Exea: hacia una infraestructura de datos en Ciencias Sociales y Humanidades digitales en el CSIC
Isabel del Bosque, Lourdes Martín-Forero, Nuria Hermida, Carlos Fernández
- 16 Bases de datos coordinadas del teatro español del Siglo de Oro: segunda fase de ASODAT
Alejandra Ulla Lorenzo, Teresa Ferrer Valls, Alejandro García Reidy, Ariadna Fuertes, Lola Josa, Héctor Urzáiz, Jesús Tronch
- 37 RMSdatabase: una nueva forma de explorar la conquista romana del noroeste peninsular
José Manuel Costa García, Breogán Amoedo Fariña, Víctor Vicente García, Fernando Batlle López, Daniel Gayoso González
- 59 Orientaciones y evaluación de técnicas en humanidades digitales: de la estadística al deep-learning
Ana García Serrano, Antonio Menta Garuz

Exea: hacia una infraestructura de datos en Ciencias Sociales y Humanidades digitales en el CSIC

Isabel Del Bosque González. Centro de Ciencias Humanas y Sociales (CCHS). CSIC.

isabel.delbosque@csic.es

Lourdes Martín-Forero Morente. Centro de Ciencias Humanas y Sociales (CCHS). CSIC.

lourdes.martin-forero@cchs.csic.es

Nuria Hermida Jiménez. Centro de Ciencias Humanas y Sociales (CCHS). CSIC.

nuria.hermida@cchs.csic.es

Carlos Fernández Freire. Centro de Ciencias Humanas y Sociales (CCHS). CSIC.

carlos.ffreire@cchs.csic.es

Palabras Clave: Exea, Infraestructuras de investigación, datos abiertos, interoperabilidad.

Algunos de los desafíos que plantean las humanidades digitales (HD) requieren de la creación de infraestructuras de investigación que aborden los nuevos retos que supone el uso de las Tecnologías de Información, en todas las fases del proceso científico y para la generación de una ciencia “centrada en los datos”. Para ello, desde el Centro de Ciencias Humanas y Sociales (CCHS) del CSIC, la unidad de sistemas de información geográfica y humanidades digitales (SIGyHD)¹, ha trabajado para la creación de una infraestructura de datos científicos, con un objetivo ambicioso como es el facilitar el punto de inflexión “digital” en la investigación en humanidades y ciencias sociales en este Centro.

Esta infraestructura digital, entendida como espacio virtual de encuentro de distintos actores y temáticas, deberá facilitar la colaboración eficiente entre investigadores y la compartición de recursos tecnológicos. Asimismo, pretende proporcionar servicios digitales duraderos que permitan soportar las diversas etapas del ciclo de vida de los proyectos de investigación, gestionar grandes repositorios, permitir el procesamiento, la conservación, el acceso y la interoperabilidad de los datos científicos, mediante soluciones basadas en estándares y normas abiertas, que puedan mantenerse y perfeccionarse a lo largo del tiempo. Con estas premisas, se ha creado una Infraestructura de datos en el CCHS que hemos denominado “Exea”²: <http://exea.csic.es>

Exea es, por tanto, el resultado de un proyecto colaborativo, abierto, interdisciplinar y de enriquecimiento continuo. Cuenta con un repositorio de contenidos estructurados, catálogo de metadatos, visualización dinámica, servicios web y herramientas que permitan descubrir,

- 1 <http://unidadsig.cchs.csic.es/sig/index.html>
- 2 El término “*exea*” ya aparece en los textos de la Castilla del s. XV y es definido, en distintos diccionarios etimológicos, como “*el guía o explorador que acompañaba a los viajeros por los caminos*”.

compartir, visibilizar, comparar y reutilizar datos, resultados de investigación y procedimientos en ciencias humanas y sociales. En un contexto de “**ciencia abierta**”, y bajo un marco de interoperabilidad, que garantice el acceso, sin restricciones, a los datos de los proyectos y acreciente la difusión y el impacto de los mismos (Figura 1).

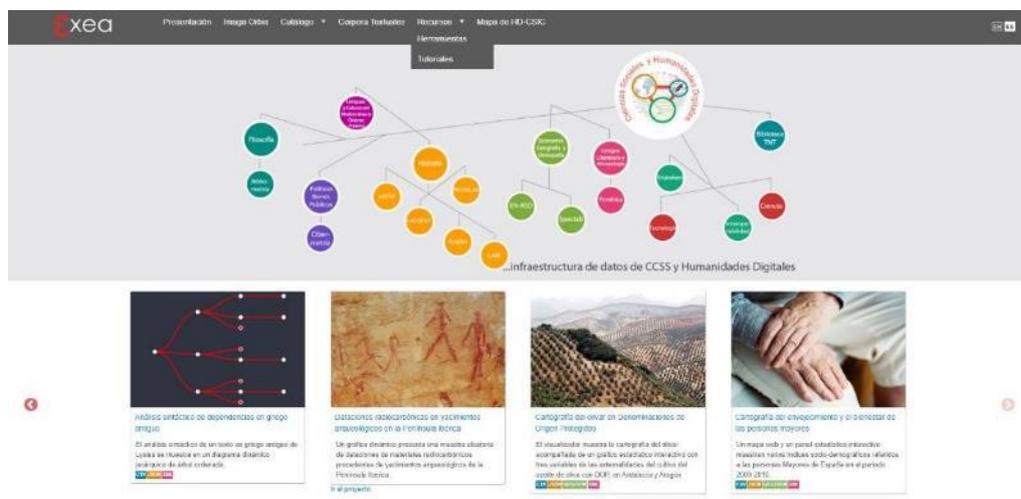


Fig.1. Imagen de la página de inicio de Exea

Esta Infraestructura de Datos se está llevando a cabo con financiación CSIC, con cargo a *un Proyecto Intramural Especial*. Está desarrollada en un entorno abierto y colaborativo, respetuoso con el trabajo de todos y que permite integrar información de diferentes orígenes, con un gestor de contenidos web libre, dinámico y altamente configurable como es *Drupal*.

Exea proporciona respuestas comunes a necesidades y usuarios diversos. Se sustenta en cinco pilares básicos, cuyos contenidos y funcionalidad deberán ir creciendo con el tiempo, y adaptándose a las necesidades de los investigadores de humanidades digitales (Figura 2):

1. Portal de inicio y de datos abiertos
2. Visualización dinámica de datos
3. *Corpora* textuales
4. Infraestructura de Datos Espaciales (IDE): *Imago Orbis*.
5. Grupo de interés o comunidad temática en HD



Fig.2. Estructura básica de Exea

1. Portal de inicio y de datos abiertos:

Desde Exea, convergiendo con los *principios FAIR* (acrónimo de *Findable, Accessible, Interoperable y Reusable*) que se impulsa desde los programas-marcos de investigación e innovación europeos y desde la propia institución CSIC (a través de un mandato institucional del 2017), se ha construido el entorno adecuado que garantice el acceso y libre descarga de datos de investigación de humanidades y ciencias sociales digitales, en formatos estándares, CSV, JSON, XML, etc.; respetando la autoría y el origen de los datos y bajo licencias de uso "*creative commons*"

2. Visualizador dinámico:

Un elemento muy importante de Exea lo conforma la visualización de los datos con distintas formas de representación. Los resultados de la mayor parte de los proyectos tienen un alto potencial para ser divulgados, en forma de mapas, redes, gráficos, etc., de manera dinámica e interactiva. Esto posee un gran interés para los investigadores de humanidades, ya que las nuevas formas de visualización de datos se convierten no solo en una manera de mostrar conclusiones, sino también en un medio para la investigación en sí mismo.

El visualizador dinámico de Exea cumple con esta necesidad, mostrando los resultados de los proyectos de una forma intuitiva y gráfica, vinculados a menudo con indicadores de tipo geográfico, socio-económico, cultural, etc. (Figura 3)

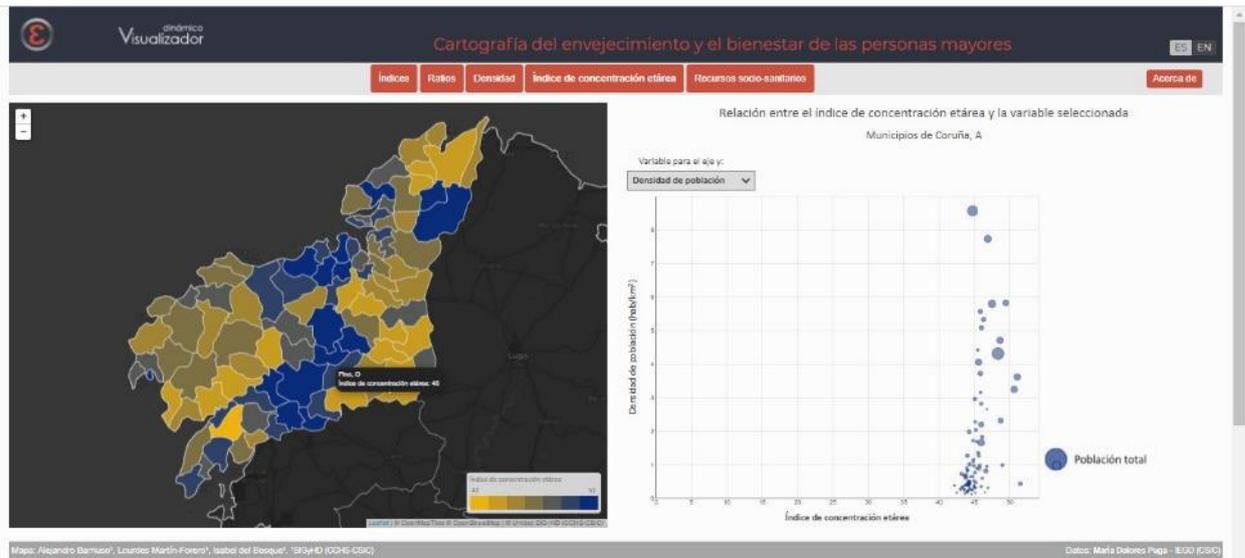


Fig.3. Ejemplo de visualización del índice de concentración etárea, por densidad de población, de A Coruña. Proyecto "Espacios de bienestar intergeneracional" de Dolores Puga (IEGD-CCHS-CSIC)

El visualizador dinámico ha sido desarrollado con la librería de JavaScript, Leaflet, para el diseño de los mapas web y D3.js para los paneles gráficos y estadísticos, lo que proporciona gran expresividad y dinamismo al mostrar la información, manteniendo, a su vez, los estándares existentes en la web como son HTML y CSS.

3. Corpora textuales:

En el caso de disciplinas como la filología o la lengua, el material sobre el que aplicar estos recursos son los textos orales o escritos. La posibilidad de acceder a gran cantidad de información, cuya codificación, ordenación y organización sólo es posible gracias a los recursos informáticos hace posible nuevas aproximaciones a viejos problemas, permitiendo contrastar afirmaciones o ideas que antes sólo eran conjeturas o especulaciones provenientes de impresiones más o menos fundadas de lingüistas. Para facilitar estos estudios se ha creado *Corpora* Textuales, como parte de Exea.

Corpora textuales permite almacenar y recuperar colecciones de documentos XML-TEI y el futuro desarrollo de aplicaciones y visualizaciones sobre las colecciones de documentos XML.

Actualmente, posee la siguiente estructura:

- Corpus de textos escritos
- Acceso a la interfaz directa a *Recogito*, herramienta de *Pelagios Network* que permite crear mapas a partir de textos. Anotar imágenes y conectar la investigación con otros recursos para que sea más visible y reutilizable.

- Gestión con eXist-DB, sistema de base de datos NoSQL orientado a documentos y datos en formato digital original o interoperable XML.

4. IDE *Imago Orbis*:

Desde la unidad SIGyHD del CCHS-CSIC se ha desarrollado, en los últimos años, aplicaciones *web mapping* y geoportales como BDTNS²², JAE EDUCA²³ HISDIMAD²⁴, IDEArq²⁵, DANZA en la Edad de Plata²⁶, etc. con temáticas y objetivos diversos, pero con el denominador común de servir de canales de transferencia de los resultados de investigación a la comunidad científica y a la sociedad en general.

Dada la envergadura y la relevancia de estas aplicaciones se hace necesaria la creación de una infraestructura que vertebre y ordene todos estos desarrollos, pero manteniendo los objetivos propios y las particularidades de cada una. De esta manera nace *Imago Orbis*, como la infraestructura de datos espaciales (IDE) en humanidades digitales, dentro de Exea.

Una IDE es una aplicación tecnológica, en la que a través de un geoportal se permite integrar en Internet: bases de datos científicas, metadatos, servicios web y otro tipo de información geospacial (o información temática georreferenciada) que se produzca en las distintas disciplinas de las humanidades digitales espaciales (*spatial humanities*), de tal forma que los conjuntos de datos generados sean, asimismo, descubribles, accesibles, combinables y compartidos. En este marco de trabajo, los distintos grupos interesados pueden incorporar sus datos científicos georreferenciados a través de una interfaz único, de una manera fácil, y de acuerdo a estándares y protocolos trabajando de forma normalizada, con el consiguiente aumento de la eficiencia y la correspondiente disminución de los costes imputables a cada uno de los proyectos (Figura 4).

²² <http://bdtns.filol.csic.es/mapa.php>

²³ <http://ceies.cchs.csic.es/visor/>

²⁴ <http://www.idehistoricamadrid.org/hisdimad/index.htm>

²⁵ http://www.idearqueologia.org/visualizador_idearq/

²⁶ <http://usig-proyectos.cchs.csic.es/investigacionendanza-map/>

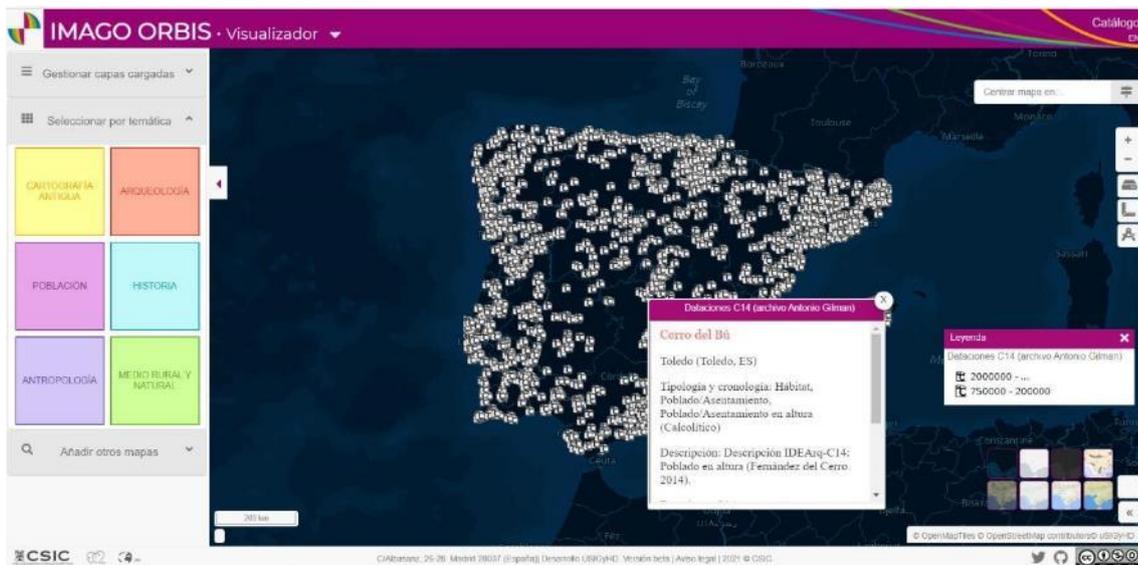


Fig.4. Ejemplo de la integración y vinculación de datos arqueológicos del proyecto IDEARQ (IH-CCHS-CSIC) en la IDE Imago Orbis.

La IDE Imago Orbis integra, por tanto, geodatos de diferentes disciplinas, consultables a través de un geoportal o “nodo central” que actúa también para sumar o vincular los geoportales o nodos de los proyectos de otros Centros e Institutos de Humanidades. Se trata de un desarrollo propio con tecnología *OpenLayers*.

5. Comunidad temática en CCSS y HD:

Para promover el desarrollo de los proyectos de CCSS y Humanidades Digitales, compartir experiencias y reutilizar información, herramientas, etc., se ha creado un espacio para el Grupo de trabajo interdisciplinar de Interés Temático (GIT).

A través de Exea se dinamizarán los encuentros y reuniones periódicas, en las que se expondrán ideas de diseño y contenidos de la Infraestructura, se consensuarán y aprobarán recomendaciones, se intercambiarán experiencias, formación, talleres, etc.

Actualmente, esta parte de Exea se encuentra en pleno desarrollo y se destina a dar visibilidad a la comunidad de científicos del CSIC que trabajan en CCSS y HD. A través de un mapa interactivo se puede acceder a la información de los proyectos, con la geolocalización de los investigadores, institutos, recursos o herramientas en dicho ámbito.

Referencias:

Carrara, W., Enzerink, S., Oudkerk, F., Radu, C., Steenbergen E. (2018): Open Data Goldbook for Data Managers and Data Holders. Practical guidebook for organisations wanting to publish Open Data. <http://www.europeandataportal.eu/>, 10/03/18, Bruselas. European Commission Directorate General for Communications Networks, Content and Technology.

Leech, G. (1992). "Corpora and theories of linguistic performance", en J. Svartvik (ed.), *Directions in corpus linguistics*. Berlín/New York: Mouton de Gruyter, 105-122.

Santana Marrero, T. M., Fernández Freire, C., & Bosque González, I. D. (2018). Visualización cartográfica de datos científicos en humanidades y ciencias sociales: mapas personalizados mediante teselas vectoriales.

Bases de datos coordinadas del teatro español del Siglo de Oro: segunda fase de ASODAT

Teresa Ferrer Valls (Universitat de València), Ariadna Fuentes (Universitat de València), Alejandro García Reidy (Universidad de Salamanca), Lola Josa (Universitat de Barcelona), Jesús Tronch (Universitat de València), Alejandra Ulla (Universidad Internacional de la Rioja), Héctor Urzaiz (Universidad de Valladolid)

Palabras clave: bases de datos, teatro áureo, representaciones teatrales, transmisión textual, censuras, música

El objetivo de la presente intervención consiste en presentar la segunda fase de ASODAT, un proyecto coordinado que reúne a varios grupos de investigación y tiene como finalidad la articulación de la información contenida en bases de datos relacionadas con el teatro clásico español y su documentación.

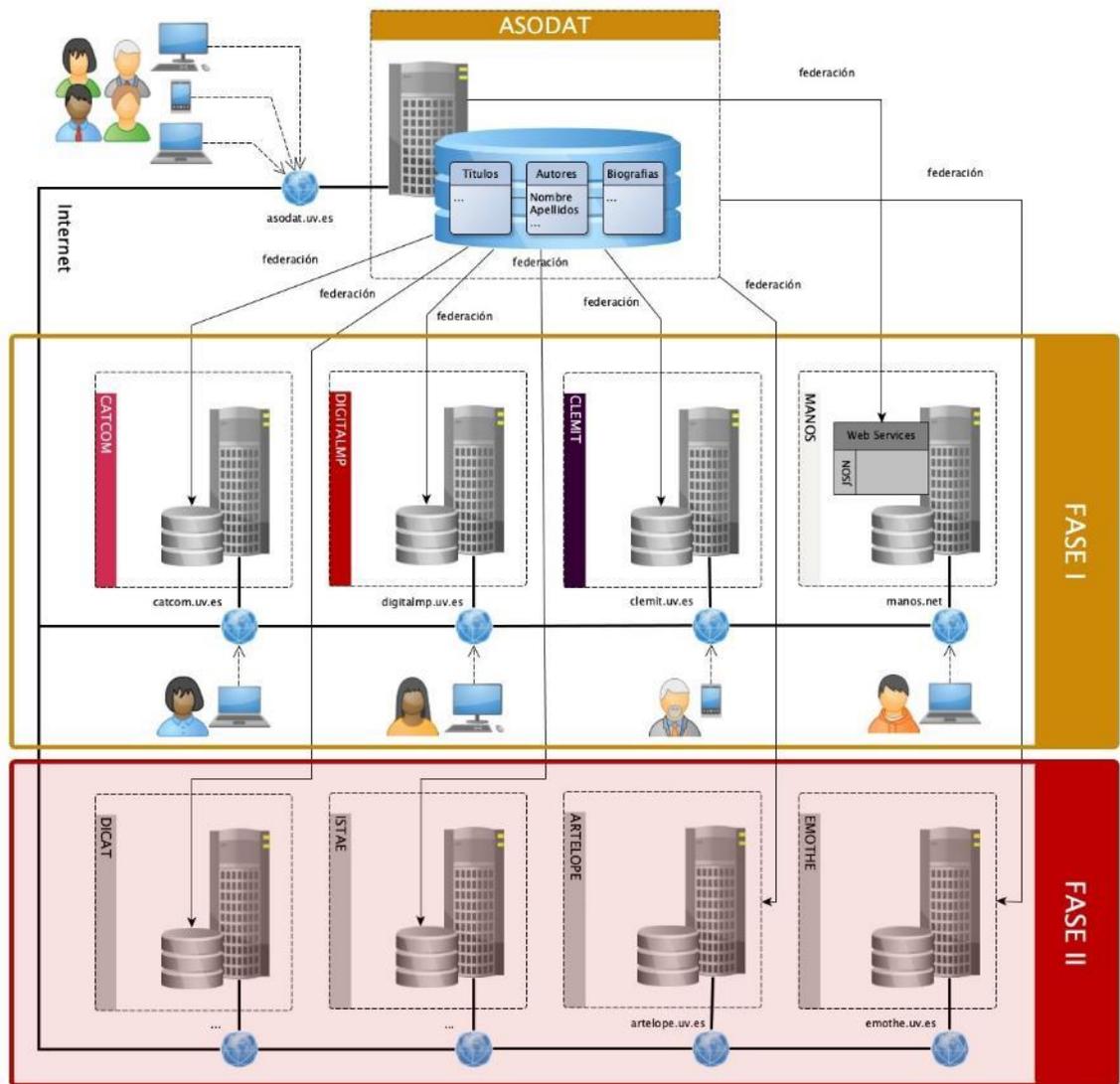
En su primera fase, el equipo partió de tres bases de datos “legadas” (i.e., CATCOM, MANOS y CLEMIT), que se caracterizaban por disponer de contenidos generados por trabajos iniciados algunos años antes, y de una base de datos de nueva concepción (i.e., DIGITAL MÚSICA POÉTICA), cuyos contenidos se han elaborado a la par que el sistema final. En aquel primer estadio se optó por escoger una arquitectura federada y un diseño de federación en dos niveles. En un primer nivel, que puede ser ampliado, se compartían determinados elementos comunes a todas las bases de datos. El segundo nivel de federación involucraba la relación de distintos aspectos particulares de cada una de las bases y nos permitía, por otra parte, explotar la información según el criterio independiente de los investigadores. Es importante indicar que, como paso previo a la federación de la información, en la fase anterior fue necesario estandarizar tanto los contenidos como los formatos de los datos de las diferentes bases.

En la segunda fase del proyecto, que aquí presentamos, se asocian al proyecto dos nuevos grupos ISTAE y Emothe. De este modo se cumple el objetivo general de ASODAT, que persigue que esta plataforma se convierta en un referente que, por su naturaleza, permita la ampliación de esta información con nuevas bases de datos y a la que se puedan ir incorporando otros grupos de investigación afines, asimilándose a los esquemas de adaptación y trabajo que se han desarrollado para la federación de la información del teatro clásico español dentro de ese nuevo marco colaborativo. Desde el punto de vista técnico, el objetivo general, que ya hemos puesto en marcha, consiste en realizar la verificación del funcionamiento de los protocolos creados en la fase anterior, y su adaptación o modificación si fuese necesario. Asimismo, estamos llevando a cabo el estudio de los requerimientos por parte de las nuevas bases de datos a incorporar y, asimismo, estudiando dentro de la base de datos ASODAT las nuevas características que puedan aportar dichas bases de datos de cara a la futura explotación conjunta. Esta tarea se llevará a cabo con la adaptación de la base de datos DICAT (Diccionario biográfico de actores del teatro clásico español) y la creación de la nueva base de datos ISTAE (Impresos sueltos del teatro antiguo español), ya en versión beta en estos momentos, para su integración en ASODAT. Además, se están estudiando las bases de datos EMOTHE (textos teatrales

españoles y europeos y sus traducciones) y ARTELOPE (argumentos y motivos del teatro de Lope de Vega) para fijar las líneas de actuación de cara a su asociación.

Además de estos objetivos generales comunes a ASODAT, es necesario considerar los objetivos específicos de cada uno de los proyectos que integran la federación y que, fundamentalmente, tienen que ver con el tratamiento y el incremento de la información relacionada con la investigación en cada uno de los campos de trabajo que atañen a los diferentes equipos.

La incorporación de los nuevos grupos al coordinado nos permitirá validar la metodología de trabajo implantada en el caso de las nuevas bases de datos que parten de características similares a algunas de las asociadas en la primera fase. Pero también estamos estudiando los nuevos casos, sin antecedentes en la fase anterior, para ir abriendo horizontes a la incorporación de nuevas bases de datos. En particular, el caso de la base de datos DICAT podría asimilarse al estado en el que se encontraba la base de datos CATCOM en la fase anterior, la cual también se encontraba desarrollada en Filemaker y hubo que convertirla. El caso de la base de datos ISTAE sirve para verificar los pasos que se han establecido a la hora de incorporar bases de datos de nueva creación, tal y como sucedió en el pasado con base de datos DIGITALMP. Estas correspondencias en cuanto al estado en el que se encuentran las bases de datos de nueva incorporación es lo que ha tratado de reflejarse en la figura que se ofrece a continuación. Por otra parte, el estudio de las bases de datos ARTELOPE y EMOTHE nos permite terminar de fijar los pasos a seguir en la federación de bases de datos para adaptarnos a la tecnología legada.



Esta federación surge, por tanto, de la convicción de que la puesta en contacto de la información perteneciente a las diferentes bases de datos permitirá hallazgos y ofrecerá respuestas a cuestiones que, estudiadas de forma aislada, no resultarían visibles.

Bibliografía

Bibliografía básica desde el punto de vista técnico:

Adamson. C., *Mastering data warehouse aggregates*, Wiley Publishing, 2006.

Stephens, R. *Beginning Database Design Solutions*. Wiley Publishing, 2008.

Celma, M., Casamayor, J.C., y Mota, L., *Bases de datos relacionales*, Pearson Education, 2003.

Connolly, T., y Carolyn Begg, C., *Database systems. A practical approach to design, implementation and Management*, Pearson Education, 2015, 6ª ed.

Date, C. J., *Introducción a los sistemas de bases de datos*, Pearson Education, 2001.

Date, C.J., *SQL and Relational Theory*, O'Reilly Media, Inc., 2015 3ª ed.

- Inmon, H., *Building the Data Warehouse*, John Wiley & Sons, 2005.
- Fowler M., Rice D., Foemmel M., Heatt E., Mee R., and Stafford R. *Patterns of Enterprise Application Architecture*, Addison-Wesley, .2003.
- Kleppmann, M. *Designing Data-Intensive Applications: The Big Ideas Behind Reliable, Scalable, and Maintainable Systems*, O'Reilly Media Ed., 2016.
- Marz, N., Warren, J. *Big Data: Principles and best practices of scalable realtime data systems*, Hanning, 2015.
- Opel, Andy, *Data Modeling*, Mc Graw Hill Ed., 2010.
- Özsu, M. Tamer, y Valduriez, Patrick. *Principles of distributed database systems*, Springer, 2011.
- Ramakrishnan, Raghu, y Gehrke, Johannes, *Database Management Systems*, 2003, 3ª ed. Reese, G., *Database programming with JDBC & Java*, O'Reilly Media Ed., 2003.
- Taniar, D., Leung, Clement H.C., Rahayu, Wenny, y Goel, Sushant. *High-performance parallel database processing and grid databases*, Wiley-Blackwell, 2008.
- Django: the Web framework for perfectionists with deadlines.
<https://www.djangoproject.com/start/overview/>
- FileMaker: An Apple Subsidiary. <https://www.filemaker.com/>
- Galera Cluster for MySQL. <https://galeracluster.com>
- MySQL Data Base. <https://www.mysql.com/>
- PHP: Hypertext Preprocessor. <https://www.php.net/>
- PostgreSQL: The World's Most Advanced Open Source Relational Database.
<https://www.postgresql.org/>
- Python. <https://www.python.org>
- Ruby on Rails: A web-application framework that includes everything needed to create database-backed web applications according to the Model-View-Controller pattern. <https://rubyonrails.org/>
- CATCOM
- Badía Herrera, J. "Bases de datos para el estudio del teatro clásico español. Propuesta de modelo para el análisis de microsecuencias", en Rebeca Lázaro Niso (coord.), *Corpus y bases de datos para la investigación en literatura*, 2017, pp. 13-33.
- Ferrer Valls, T., "Hacia un sistema integrado de bases de datos sobre teatro clásico español y práctica escénica: el proyecto coordinado ASODAT", *Cuadernos AISPI. Estudios de lenguas y literaturas hispánicas*, nº 11 (2018), pp. 115-30.
- García Reidy A., "The Technological Environment of the Early Modern Spanish Stage," in E. García Santo-Tomás (ed.), *Science on Stage in Early Modern Spain*, Toronto, University of Toronto Press, 2019, pp. 58-78.
- García Reidy, Alejandro, "Las comedias religiosas de Calderón en la escena barroca: posibilidades y límites a partir de la base de datos CATCOM," *Criticón*, 130 (2017), pp. 93-107.

- García-Reidy, Alejandro (2019): "La presencia escénica barroca de Andrés de Claramonte a partir de las bases de datos Manos y CATCOM", *Bulletin of the comediantes*, 71. 1-2: 135-54.
- Greer, M. G., "Manos Teatrales: la historia contada por los manuscritos," in *TeaPal 7* (2013).
- Oleza, Joan, "Una base de datos para una travesía compleja". *Renovación en el Siglo de Oro: repertorio e instrumentos de investigación. Cuadernos de Teatro Clásico*, nº 29, (2014), 209-240.
- Oleza, Joan, "Exploring a multilingual digital edition of Early Modern European Theatre." *Journal for Early Modern Cultural Studies*, vol. 13-14, nr. 4, (Fall 2013), pp. 152-154.
- Joan Oleza, "A Description of ARTELOPE: a database containing information, plots and texts related to Lope de Vega's theater", *Journal for Early Modern Cultural Studies*, vol. 13-14, nr. 4, (Fall 2013), pp. 150-152.
- Tronch, Jesús. "Displaying Textual and Translational Variants in a Hypertext and Multilingual Edition of an Early Modern Play." In *Early Modern Studies after the Digital Turn*, 88–116. Iter Press, 2016.

RMSdatabase: una nueva forma de explorar la conquista romana del noroeste peninsular

José M. Costa-García Departamento de Historia. Universidade de Santiago de Compostela.
josemanuel.costa@usc.es

*Breogán Amoedo Fariña*² Lapisoft Projects S.L. breo@lapisoft.es

Víctor Vicente García Departamento de Historia. Universidade de Santiago de Compostela. victor.vicente.garcia@usc.es

Fernando Batlle López Lapisoft Projects S.L. fernando@lapisoft.es

*Daniel Gayoso González*² Lapisoft Projects S.L. daniel@lapisoft.es

Palabras clave: arqueología, conquista romana, datos geospaciales, base de datos, plataforma web colaborativa.

En la última década, el acceso a nuevas tecnologías, herramientas digitales y conjuntos de datos geospaciales ha supuesto una verdadera revolución en arqueología [1-6]. La aparición de repositorios en abierto de fotografía aérea o datos LiDAR, la popularización de determinados dispositivos de detección remota (drones, equipos de prospección geofísica) o la consolidación del uso de sistemas de información geográfica (SIG) como herramienta analítica son algunos de los avances que han transformado el modo en el que detectamos y estudiamos los yacimientos arqueológicos, resultando en una disciplina cada vez más tecnificada.

En particular, los estudios sobre la expansión del Estado romano se han beneficiado enormemente de estos avances. En tiempos recientes ha aumentado exponencialmente el número de yacimientos que pueden relacionarse con la presencia del ejército romano en la Península Ibérica [7,8] (Fig. 1) y en Europa [9-13]. La acumulación de nuevos datos arqueológicos nos permite ahora construir una imagen más detallada y vívida de los procesos de conquista, desafiando las narrativas históricas tradicionales [14,15]. Sin embargo, la velocidad a la que se ha acumulado nueva información ha imposibilitado la articulación de estas nuevas visiones que tengan en cuenta asimismo el rol jugado por las poblaciones indígenas que habitaban en estos territorios antes de la llegada de los romanos [16].

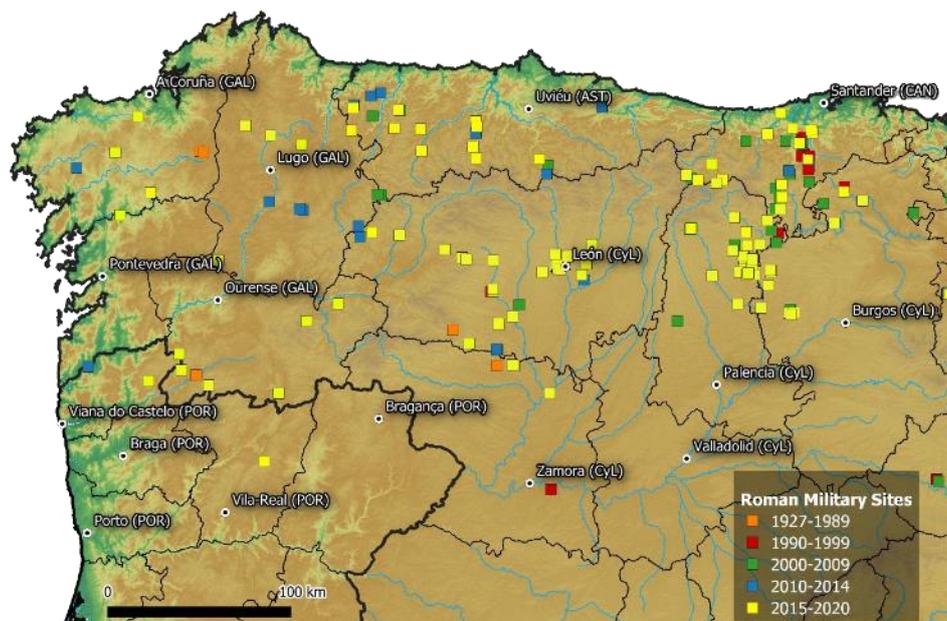


Fig.1. Asentamientos militares romanos en el noroeste ibérico. Descubrimientos recientes.

No es esta la única sombra proyectada por este proceso de renovación metodológica. La mayor parte de estos yacimientos arqueológicos son extremadamente frágiles [17] y las herramientas digitales muestran la destrucción de muchos de ellos en las últimas décadas debido a distintas actividades antrópicas [18] que en buena medida pueden relacionarse con el proceso de despoblación del mundo rural y el abandono de los sistemas tradicionales de explotación del territorio (mecanización agrícola, políticas de reforestación, instalación de eólicos, etc.) [19,20]. Urge, por tanto, explorar nuevas formas que permitan dar a conocer su existencia y concienciar a la ciudadanía de su valor patrimonial y cultural.

El proyecto A-REDE (*Archaeologies of the Roman Expansion in the Digital Era*) aglutina a un equipo interdisciplinar integrado por arqueólogos, ingenieros de telecomunicaciones, informáticos y especialistas en comunicación. Siguiendo los principios de la ciencia en abierto y de la arqueología pública, en la actualidad desarrollamos *RMSdatabase*, una plataforma web interactiva, de acceso abierto que permite el almacenamiento y uso de datos arqueológicos con diferentes fines. Su vocación abierta y colaborativa permitirá servir de punto de encuentro y estudio tanto a la comunidad investigadora como a diversas audiencias interesadas.

Como herramienta de investigación, permite la revisión, análisis y búsqueda cruzada de datos a una escala sin precedentes. Para ello, se ha desarrollado un modelo de datos especialmente adaptado a los yacimientos objeto de análisis –en la versión piloto, los asentamientos militares romanos- y a las narrativas que se pretenden desarrollar a partir de los mismos –la conquista y ocupación romana del noroeste ibérico-. Partiendo de una clasificación tipológica, morfológica y locacional pionera [21,22], el modelado de datos se ha depurado y perfeccionado para adaptarse a las posibilidades de la tecnología sin renunciar a la minuciosidad del registro arqueológico (Fig. 2).

Para el diseño de la plataforma, se aplican las estrategias técnicas necesarias para garantizar la coherencia e integridad de los datos, así como las políticas de copias de seguridad de respaldo por una eventual necesidad de recuperación. Se garantiza el acceso concurrente a los mismos, se establecen mecanismos de publicación jerarquizados y revisión y moderación de los datos a publicar.

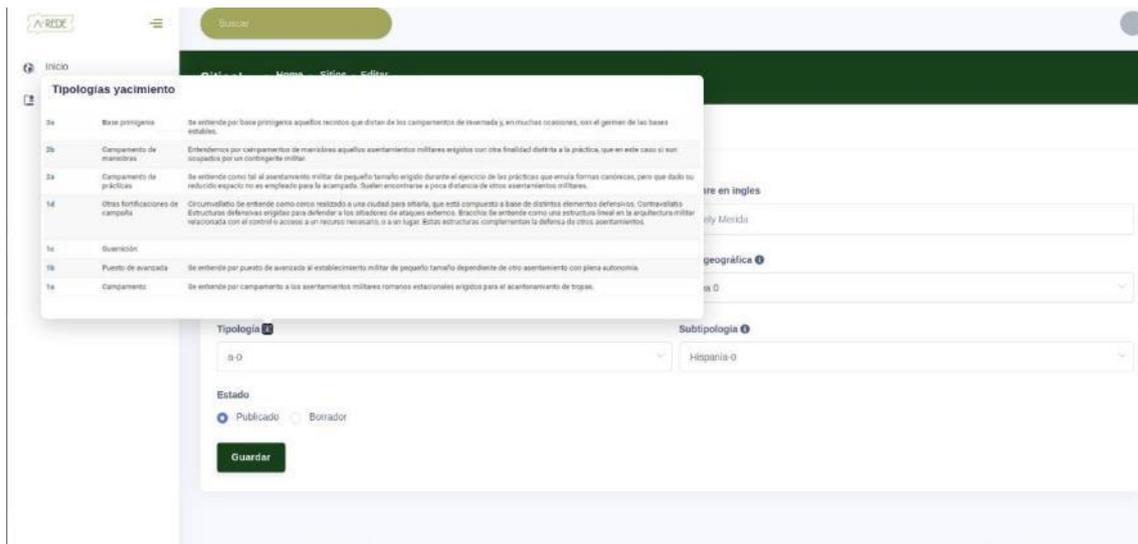


Fig. 2. Plataforma web. Selección de tipología del yacimiento (WIP).

La plataforma aspira asimismo a ser una innovadora forma de mostrar a la sociedad cómo la arqueología puede reconstruir episodios olvidados de nuestro pasado mediante el uso de nuevas tecnologías. Con este fin se han incorporado asimismo numerosos recursos gráficos (planimetrías, fotografías, mapas interactivos, reconstrucciones 3D) (Fig. 3) que permiten al usuario a entender y diferenciar las distintas realidades arqueológicas. Sin embargo, nuestro principal objetivo es desarrollar narrativas atractivas a partir de los datos generados, de modo que la descripción de yacimientos arqueológicos es solamente un primer escalón dentro de esta experiencia digital. La combinación de los diferentes sitios permite recrear escenarios más complejos (como batallas, asedios o rutas a través del territorio) y estos a su vez episodios históricos más amplios (como campañas militares, la organización del territorio durante la posguerra o la logística de las guarniciones militares durante la *pax romana*).

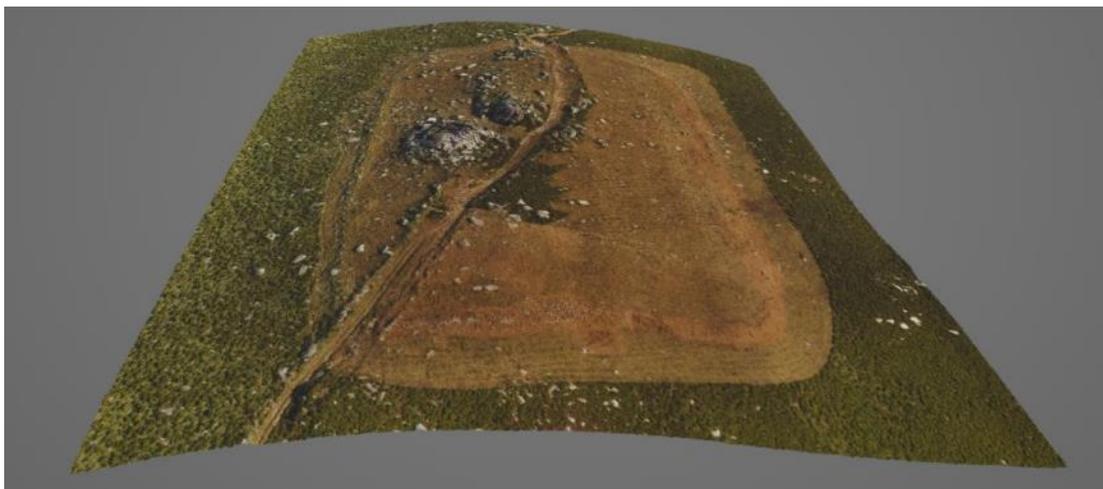


Fig. 3. La plataforma integra una amplia selección de recursos. Incluyendo modelados 3D a partir de datos LiDAR o reconstrucciones fotogramétricas.

De este modo, el usuario puede escoger el grado de detalle deseado en su consulta, desde las grandes narrativas hasta los estudios más técnicos y detallados. En el futuro, pretendemos generar un espacio para el diálogo activo

y animar a la gente a formar parte del proceso de investigación. Al subir sus propias contribuciones y materiales gráficos, cualquier ciudadano interesado contribuye al conocimiento y defensa de nuestro patrimonio arqueológico.

Bibliografía

1. Kokalj, Ž.; Somrak, M. Why Not a Single Image? Combining Visualizations to Facilitate Fieldwork and On-Screen Mapping. *Remote Sensing* 2019, 11, 747.
2. Kokalj, Ž.; Hesse, R. *Airborne laser scanning raster data visualization: a Guide to good practice*; ZRC: Ljubljana, 2017; 10.3986/9789612549848.
3. Corsi, C.; Slapšak, B.; Vermeulen, F. *Good Practice in Archaeological Diagnostics. Noninvasive Survey of Complex Archaeological Sites*; Springer: Cham, 2013.
4. Adamopoulos, E.; Rinaudo, F. UAS-Based Archaeological Remote Sensing: Review, Meta-Analysis and State-of-the-Art. *Drones* 2020, 4, 46.
5. Campana, S. Drones in Archaeology. State-of-the-art and Future Perspectives. *Archaeological Prospection* 2017, 10.1002/arp.1569, 275-296, doi:10.1002/arp.1569.
6. Cerrillo Cuenca, E.; López López, A. Evaluación y perspectivas del uso del LiDAR en la arqueología española. *Boletín del Museo Arqueológico Nacional* 2020, 39, 221-238.
7. Menéndez Blanco, A.; García Sánchez, J.; Costa-García, J.M.; Fonte, J.; GonzálezÁlvarez, D.; Vicente García, V. Following the Roman Army between the Southern Foothills of the Cantabrian Mountains and the Northern Plains of Castile and León (North of Spain): Archaeological Applications of Remote Sensing and Geospatial Tools. *Geosciences* 2020, 10, n. 485, doi:10.3390/geosciences10120485.
8. Camino Mayor, J.; Peralta Labrador, E.; Torres Martínez, J.F., Eds., *Las Guerras AsturCántabras*, KRK Ediciones: Gijón, 2015.
9. Oltean, I.; Hanson, W. Conquest strategy and political discourse: New evidence for the conquest of Dacia from LiDAR analysis at Sarmizegetusa Regia. *Journal of Roman Archaeology* 2017, 30, 429-446.
10. Driver, T.G.; Burnham, B.C.; Davies, J.L. Roman Wales: Aerial Discoveries and New Observations from the Drought of 2018. *Britannia* 2020, 10.1017/S0068113X20000100, 1-29, doi:10.1017/S0068113X20000100.
11. Bernardini, F.; Vinci, G.; Horvat, J.; De Min, A.; Forte, E.; Furlani, S.; Lenaz, D.; Pipan, M.; Zhao, W.; Sgambati, A., et al. Early Roman military fortifications and the origin of Trieste, Italy. *Proceedings of the National Academy of Sciences* 2015, 112, 1520-1529, doi:10.1073/pnas.1419175112.
12. Komoróczy, B.; Rajtár, J.; Vlach, M.; Hüssen, C.-M. A companion to the archaeological sources of Roman military interventions into the Germanic territory north of the Danube during the Marcomannic Wars. In *Marcomannic Wars and Antonine Plague. Selected essays on two disasters that shook the Roman World*, Edrich, M., Komoróczy, B., Madejski, P., Vlach, M., Eds. Czech Academy of Sciences - Instytut Archeologii: Brno-Lublin, 2020; Vol. Vol. 1, pp. 173-254.
13. Groh, S.; Sedlmayer, H. *Expeditiones Barbaricae. Forschungen zu den Römischen Feldlagern von Engelharstetten, Kollnbrunn und Ruhhof, Niederösterreich*; Landessammlungen Niederösterreich / Donau-Universität Krems: Krems, 2015.
14. Costa-García, J.M.; Fonte, J.; Gago, M. The reassessment of the Roman military presence in Galicia and Northern Portugal through digital tools: archaeological diversity and historical problems. *Mediterranean Archaeology and Archaeometry* 2019, 19, 17-49, doi:doi.org/10.5281/zenodo.3457524.
15. Peralta Labrador, E.J.; Camino Mayor, J.; Torres-Martínez, J.F. Recent research on the Cantabrian Wars: the archaeological reconstruction of a mountain war. *Journal of Roman Archaeology* 2019, 32, 421-438, doi:https://doi.org/10.1017/S1047759419000217.

16. González-Álvarez, D.; Costa-García, J.M.; Menéndez Blanco, A.; Fonte, J.; Álvarez Martínez, V.; Blanco-Rotea, R.; Gago Mariño, M. La presencia militar romana en el noroeste ibérico hacia el cambio de era: estado actual y retos de futuro. In *Accampamenti, guarnigioni e assedi durante la Seconda Guerra Punica e la conquista romana (secoli III-I aC): prospettive archeologiche*, Vallori Márquez, B., Bellón Ruíz, J.P., Rueda Galán, C., Eds. Edizioni Quasar di Severino Tognon: Roma, 2019; pp. 127138.
17. Peralta Labrador, E. Los campamentos romanos de campaña (castra aestiva): evidencias científicas y carencias académicas. *Nivel Cero* 2002, 10, 49-87.
18. Costa-García, J.M.; Fonte, J. Scope and limitations of airborne LiDAR technology for the detection and analysis of Roman military sites in Northwest Iberia. In *Archaeology and Geomatics. Harvesting the benefits of 10 years of training in the Iberian Peninsula (2006-2015)*, Mayoral Herrera, V., Parcero-Oubiña, C., FábregaÁlvarez, P., Eds. Sidestone Press: Leiden, 2017; pp. 57-73.
19. Pazos Lamoso, M.Á.; Palmeiro Ramos, X.M.; Conde Gómez, D. Del minifundio al megaincendio. Transiciones, rupturas y nuevos retos cara a un desarrollo sostenible en el medio rural gallego. *Sémata: Ciências Sociais e Humanidades; No 30 (2018): Bioeconomía y Memoria Ecológica de los Territorios: Transdisciplinarietà para un Futuro Sostenible* 2018, 10.15304/s.30.5402, doi:10.15304/s.30.5402.
20. Corbelle Rico, E.; Crecente Maseda, R. Land Abandonment: concept and consequences. *Revista Galega de Economía* 2008, 17, 1-13.
21. Costa-García, J.M. Roman Camp and Fort Design in Hispania: An Approach to the Distribution, Morphology and Settlement Pattern of Roman Military Sites during the Early Empire. In *Limes XXIII. Proceedings of the 23rd International Limes Congress in Ingolstadt 2015*, Sommer, C.S., Matešić, S., Eds. Nünnerich-Asmus: Mainz, 2018; Vol. Vol. 2, pp. 986-993.
22. Costa-García, J.M. *Arqueología de los asentamientos militares romanos en la Hispania altoimperial (27 a. C. – ca. 280 d. C.)*; USC - Tesis Doctoral Inédita: Santiago de Compostela, 2013; 10.13140/RG.2.2.14403.66084/1.

Orientaciones y evaluación de técnicas en humanidades digitales: de la estadística al *deep-learning*

Ana García Serrano y Antonio Menta Garuz

ETSI Informática de la UNED, Madrid, España (agarcia@lsi.uned.es)

Resumen: Cuando en Humanidades Digitales (HD) se dispone de material digital en un proyecto de investigación o profesional, es necesario identificar las herramientas más prometedoras para alcanzar los objetivos planteados. En este trabajo, aportamos algunas orientaciones metodológicas para conocer las técnicas y la tecnología más innovadora en el campo de las HD en lo que respecta al análisis automático de información. Las técnicas estadísticas son importantes para las HD, así como las de aprendizaje automático y quizá lo sean las de *deep learning*. Utilizamos algunos casos de éxito para hacer un breve recorrido de estas tecnologías utilizadas en HD. Finalmente, hacemos una mención especial a los denominados “foros de evaluación” como herramienta metodológica de investigación.

Palabras Clave: recursos y herramientas informáticas; gestión automática de datos; procesamiento automático del lenguaje natural.

1. Introducción

La definición de Humanidades Digitales (HD) en la wikipedia²⁷ hace converger a las humanidades y la informática al centrarse en el “patrimonio cultural digitalizado” y en “el uso de herramientas informáticas”. Desde 2011 hay una recomendación europea²⁸ para que en 2025 todo el patrimonio cultural (institucional) europeo esté digitalizado. Sin embargo, de acuerdo con [Tasovac et al., 2020], en 2017 solo hay un 30% digitalizado y en 2019 un 20% de este patrimonio digitalizado ya no es de calidad para su uso con los procesos automáticos avanzados actuales. Pero en el 2020 hay 500.000 investigadores de Humanidades en Europa que demandan esta información digitalizada y el uso de tecnología informática en su investigación o profesión. Actualmente, además necesitamos que el patrimonio cultural (institucional) europeo digital sea interoperable (intercambiable) y su acceso esté garantizado. En algunos casos no es interoperable, porque no se digitalizó con estándares básicos y en otros muchos casos sí, pero son recursos privados, están accesibles bajo licencia o pago, o están en abierto y son accesibles a través de aplicaciones web, pero ¿son accesibles para cualquier usuario?

Otro de los objetivos importantes en HD es conocer qué herramientas son las más prometedoras en un proyecto y son los humanistas digitales los que deben hacer de “intermediarios” entre los informáticos y los humanistas para apoyar en esta decisión. En lo que sigue, aportamos algunas orientaciones metodológicas para conocer las técnicas y la tecnología más innovadora en el campo de las HD en lo que respecta al análisis automático de información. A continuación, discutimos sobre el interés de los foros internacionales de evaluación, para concluir con algunos comentarios relacionados con las herramientas tecnológicas en las HD.

²⁷ https://es.wikipedia.org/wiki/Humanidades_digitales

²⁸ Recommendation of 27 October 2011 on Digitisation and Online Accessibility of Cultural Material and Digital Preservation (REC 2011/711/EU)

2. Soporte informático a la investigación humanística: ¿qué técnica puede ser interesante en mi investigación?

En el panorama actual se pueden encontrar multitud de herramientas disponibles, que se basan en algoritmos que implementan técnicas que se pueden utilizar en un problema concreto de análisis automático de información. Siempre, como paso previo a la investigación de los humanistas con corpus digitales, es necesario seleccionar las herramientas más “adecuadas” para los objetivos planteados.

Los conceptos dato, información y conocimiento, son pertinentes cuando nos referimos a las “Humanidades digitales y el conocimiento”. Aparecen en la pirámide DICS de datos-información-conocimiento-sabiduría [García-Marco, 2011], de forma que el conocimiento surge de la aplicación de la experiencia a la información, que a su vez es la interpretación de un conjunto de datos estructurados y organizados. Un dato hace referencia a la unidad mínima semántica, se representa a través de un símbolo y por sí solos no aportan nada.

En informática, el objetivo de la estructuración de los datos es identificar, ordenar y crear modelos de conocimiento sobre la base de técnicas de: clasificación, *clustering*, *redes neuronales* y otras [García Serrano y Castellanos, 2016].

Las técnicas de clasificación permiten organizar colecciones de objetos, identificar sus características comunes y discriminantes y denotar los objetos de forma consistente sobre la base de las etiquetas. Un modelo de clasificación muy utilizado es la taxonomía. Las ontologías son modelos de conocimiento que estructuran la información en una jerarquía de clases. Son difíciles de construir, pero, si se construyen y reutilizan, son muy útiles. Sin embargo, en ambos casos, para su construcción es necesario disponer previamente del conocimiento relacionado con los objetos a clasificar.

Las *técnicas de clustering* persiguen agrupar los datos de manera automática de acuerdo con similitudes, correferencias o paralelismos que se puedan encontrar entre ellos. Por ejemplo, es esperable que dos documentos compartiendo un vocabulario similar pertenezcan a la misma categoría. Relacionadas también con la organización automática de los datos, están las técnicas probabilísticas, técnicas que asumen que existe una estructura latente en los datos basada en el conjunto de temáticas que estos abordan. Así mismo cada una de las temáticas es definida por la distribución de términos (de entre los incluidos en los datos o descripciones) que mejor la representa. Por lo tanto, cada uno de los documentos individuales puede ser expresado mediante una distribución probabilística a lo largo del conjunto de temáticas latente.

A diferencia de la técnica de clasificación mediante ontologías, las aproximaciones probabilísticas o las basadas en *clustering* no requieren de información previa. Es decir, no es necesario que se establezca a priori las clases o se defina el modelo ontológico necesario para estructurar los documentos. Sin embargo, con estas técnicas es necesario parametrizar el funcionamiento de los algoritmos para crear la representación final de los datos, por ejemplo, el número de *clusters* (o grupos) en los cuales se desea estructurar los datos. Aquellos algoritmos que no lo requieren, sí que necesitan establecer de manera manual algún criterio de parada o de selección final de los grupos a generar.

Por otra parte, los resultados obtenidos por la mayoría de las técnicas anteriores son difíciles de entender por las personas, ya que la estructura latente detectada a menudo no se corresponde con conceptos reales y la organización resultante tiene que ser analizada por expertos para interpretar su significado. Y este problema se incrementa con las técnicas basadas en redes neuronales y *deep learning*.

3. Análisis de datos: de las técnicas estadísticas al Deep Learning

Gran parte de los corpus disponibles en las HD tienen un tamaño pequeño desde el punto de vista informático. Esto dificulta su utilización con determinadas técnicas basadas en el análisis de patrones estadísticos de los datos. Por ejemplo, cuanto mayor es el corpus, más veces aparece una determinada palabra en uno o varios contextos y se obtiene una mejor representación distribuida de la misma al aplicar técnicas como los *word embeddings*. Estos son un conjunto de modelos centrados en el estudio de la semántica distribucional que tuvieron una gran repercusión tras la aparición del modelo *word2vec* creado por un equipo de Google dirigido por Tomas Mikolov [Mikolov et al., 2013]. Para ser capaces de representar correctamente la información, los modelos necesitan ser entrenados de forma no supervisada (sin necesidad de anotación y otros procesos previos) en un gran corpus con millones de tokens. Gracias a que se comparten en la comunidad, se dispone de modelos pre-entrenados de los mismos, por lo que no es necesario entrenarlos por los investigadores cuando se quieren utilizar en un ámbito general. Para su utilización en campos específicos es necesario realizar tareas de ajuste o el entrenamiento desde cero mediante un corpus específico de la materia. En la mayoría de las ocasiones, el vocabulario o el contenido semántico de las palabras es muy diferente de un modelo pre-entrenado con la Wikipedia al utilizado en manuscritos históricos [Maupéou, Le Men, 2019].

A esta circunstancia hay que añadir la dificultad tanto en tiempo como económica de etiquetar de forma manual grandes cantidades de texto. Sobre todo, para textos pertenecientes a una temática muy específica que implica un conocimiento de esta para saber identificar correctamente las entidades a etiquetar. Para mejorar esta circunstancia varias investigaciones han publicado herramientas o técnicas centradas en la etiquetación de corpus pertenecientes a dominios con escasez de estos. En [Lin, Lu, 2018] se propone la utilización de técnicas de adaptación al dominio, utilizando redes neuronales para solventar la falta de recursos en las Humanidades Digitales. Para conseguirlo, modifican la arquitectura de la red neuronal añadiendo capas de adaptación entre los datos entrenados en el dominio de origen, donde existen grandes cantidades de *datasets* etiquetados y el dominio final donde los recursos son escasos.

En los últimos años, la utilización de técnicas basadas en el aprendizaje profundo (*Deep Learning*) ha supuesto una revolución en muchas de las tareas de procesamiento del lenguaje natural en las que previamente se utilizaban técnicas de aprendizaje automático que necesitaban la creación de características de forma manual para obtener buenos resultados. Era necesario un estudio en detalle de los datos para encontrar patrones importantes para añadir al modelo. Desde la aparición del aprendizaje profundo, el propio modelo es capaz de encontrarlos mediante la representación y modificación de un espacio latente gracias a la sucesión de capas neuronales que lo forman.

Un ejemplo dentro del mundo de las Humanidades Digitales es la clasificación de textos históricos según su época de publicación. En [Liebeskind, Liebeskind, 2019], se hace un estudio del corpus *Responsa* que contiene textos históricos de cuatro épocas diferentes: el siglo XI hasta finales del siglo XV, el siglo XVI, los siglos XVII al XIX y el siglo XX hasta hoy. Para el estudio se proponen varios métodos diferentes, como las Redes Neuronales Convolucionales (en inglés CNN) en las que el modelo es capaz de aprender el contexto anterior en una determinada oración, lo que mejora los resultados en tareas como la clasificación o el reconocimiento de entidades, donde las palabras anteriores tienen gran importancia para la detección de las mismas. También se utilizan las Redes Neuronales Recurrentes (RNN), que destacan especialmente en el procesamiento de datos secuenciales. De ahí su utilización en el tratamiento de textos, donde tiene gran importancia las palabras anteriores o posteriores a la hora de tratarlo. En las conclusiones de la investigación muestran la superioridad de los resultados obtenidos por los modelos de aprendizaje profundo sobre otros enfoques del aprendizaje automático.

Recientemente, los modelos de lenguaje neuronales han desplazado a las RNN en el estado del arte de la mayoría de las tareas de procesamiento del lenguaje natural, en gran parte debido a la aparición de los modelos basados en *Transformer* [Vaswani et al., 2017]. Modelos como BERT [Devlin et al., 2018] han provocado una revolución vertiginosa y se han creado redes cada vez con más parámetros y más capas profundas capaces de almacenar más información en su espacio latente.

El *Deep Learning* mejora la calidad y la eficiencia de problemas ya resueltos (como el caso de la búsqueda en imágenes y textos) o para afrontar problemas o tareas muy complejas no resueltas como es identificar la autoría de textos o el plagio entre textos [García-Serrano, Menta-Garuz, 2019], o calcular la similitud semántica entre palabras [Lastra, García-Serrano, 2015] o entre las oraciones [Lara-Clares et al., 2021], pero detectamos dos grandes problemas con esta tecnología, que exponemos en el siguiente apartado.

4. Foros internacionales **de evaluación de los algoritmos: ¿cómo asegurar cierta “calidad” en los nuevos algoritmos y aplicaciones?**

Tanto en el caso de aproximaciones de acceso abierto, como en el de las comerciales, siempre habrán pasado o estarán en un periodo de “consolidación” posterior a su aparición en los foros de investigación. Es una etapa de evaluación de la calidad y la eficiencia de las técnicas implementadas por algoritmos que resuelven tareas concretas y que han de ser evaluadas en corpus o colecciones de datos o textos de suficiente tamaño y calidad. Para ello se organizan los denominados foros de evaluación en los que los investigadores y profesionales se enfrentan a tareas concretas en los mismos corpus y cuyos resultados comparan entre sí y con respecto a la solución aportada por los organizadores de tarea concreta. Ejemplos son el foro internacional CLEF³ o IberLEF⁴, principalmente para corpus en español. A modo de ejemplo aclaratorio, algunas de las 12 tareas de CLEF 2021⁵ son: *Answer Retrieval for Questions on Math*; *BioASQ: Large-scale biomedical semantic indexing and question answering*; *Early Risk prediction on the Internet*; *ImageCLEF Multimedia Retrieval Challenge in CLEF*; *PAN Lab on Digital Text Forensics and Stylometry*; *SimpleText: (Re)Telling right scientific stories to non-specialists via text simplification*.

Una vez que una técnica o algoritmo ha sido evaluado y obtiene resultados de calidad, si además resuelve una **tarea de interés profesional o investigador**, pasa a la etapa que en informática se denomina “de producción”. Tiene que pasar por una serie de etapas hasta que forme parte de una herramienta o software accesible desde otros programas informáticos (mediante API⁶, o *endpoints*⁷) o a través de aplicaciones web.

En el área GLAM²⁹ la información que se maneja es multimodal, generalmente textual y gráfica (imágenes), aunque actualmente se esté incorporado mucha información en modo audio y video. Ejemplos bien

³ Desde 2013 la web es <http://www.clef-initiative.eu/>. CLEF comienza oficialmente en octubre de 2001 con el proyecto europeo del FP-5 <https://cordis.europa.eu/project/id/IST-2000-31002/es>

⁴ <https://sites.google.com/view/iberlef2021/>

⁵ <http://clef2021.clef-initiative.eu/index.php?page=Pages/labs.html>

⁶ Una API es un conjunto de definiciones y protocolos que se utiliza para desarrollar e integrar el software de las aplicaciones. API significa interfaz de programación de aplicaciones.

https://es.wikipedia.org/wiki/Interfaz_de_programaci%C3%B3n_de_aplicaciones

²⁹ GLAM es un acrónimo de “galerías, libraries (bibliotecas), archivos y museos” para referirse a instituciones culturales que recopilan y mantienen materiales del patrimonio cultural de interés público.

[https://en.wikipedia.org/wiki/GLAM_\(industry\)#:~:text=GLAM%20is%20an%20acronym%20for,materials%20in%20the%20public%20interest](https://en.wikipedia.org/wiki/GLAM_(industry)#:~:text=GLAM%20is%20an%20acronym%20for,materials%20in%20the%20public%20interest)

- 7 Un endpoint es un punto de acceso a un servicio, herramienta o aplicación en la red. A veces utilizan un lenguaje concreto para “solicitar un servicio” como es el caso del SPARQL endpoint <https://ceweb.br/guias/web-semantic/es/capitulo-8/>

conocidos son la biblioteca digital EUROPEANA³⁰, o la web del Museo del Prado y en concreto su biblioteca digital³¹ en la que se pueden explorar muchas de sus obras.

Los buscadores sobre imágenes con texto asociado son una muestra de lo que ocurre tras alcanzar en investigación un resultado interesante y que en futuros estadios pase a “producción”. Google puso en producción este tipo de búsqueda con una versión inicial beta (esto es, una versión del algoritmo no totalmente acabada ni libre de errores) que testamos los propios usuarios. Una década después de este caso, la tecnología que se utiliza no es la misma, pero hubo que andar un camino que es muy común en el desarrollo de la tecnología.

La investigación sobre buscadores de imágenes con textos asociados también recorrió la etapa de evaluación. Forma parte de la experiencia de la primera autora en las campañas de evaluación CLEF, en los que participa desde antes del 2006. Es a finales de la primera década del siglo XXI, cuando su trabajo queda el segundo tras el grupo de Xerox Europa en una evaluación de búsqueda de información multimodal (imágenes anotadas con texto), cuando pasamos a usar con cierta seguridad y buenos resultados la técnica de “mezclar la evidencia de los indicadores provenientes de la información textual con la de los de información visual de las imágenes” para decidir la respuesta a una necesidad de información multimodal como “necesito fotos con vestidos rojos” [Martinez-Fernandez et al., 2006] [Garcia-Serrano et al., 2009] [Benavent et al., 2010].

Este es solo un ejemplo de cómo en el desarrollo de la tecnología que se incorporará en las futuras herramientas, el participar de forma continuada en estos foros de evaluación, la calidad de las

aproximaciones va mejorando y tanto la descripción teórico-práctica de la técnica utilizada como el software implementado, está disponible para la comunidad investigadora. En nuestro conocimiento, es en la edición de 2019 de CLEF donde se organizó HIPE³² (*identifying Historical People, Places and other Entities*) una campaña de evaluación de la tarea de anotación de entidades nombradas en periódicos antiguos (en francés, inglés y alemán).

5. Comentarios finales

De acuerdo con [McGillivray et al., 2020] el recorrido actual de la investigación en informática y del procesamiento del lenguaje natural (PLN) en concreto, puede no ser de gran interés para los humanistas. Con lo que la dificultad de colaboraciones aumenta. Con el análisis de datos (*Data Mining*) o con los algoritmos de la web semántica sobre datos enlazados, los corpus (colecciones de documentos normalmente anotados) tienen una cierta magnitud, que permite a los humanistas supervisar los corpus, anotarlos con información significativa para ellos y tras procesarlos automáticamente, obtener información descriptiva “oculta” previamente.

Sin embargo, tras la incursión del *Deep Learning* en muchas áreas de investigación, surge una gran dificultad, que debería resolverse primero en el ámbito informático. Es el tamaño de los corpus necesarios para este tipo de aproximaciones. ¿Dispone el humanista de estos corpus?, pues de momento no, como tampoco están disponibles en un área tan bien financiada como es la biomedicina, ¿Se pueden invertir recursos para producirlos? Pero, es más, si no se dispone de una ingente cantidad de información digitalizada ¿creemos que el esfuerzo de su desarrollo supondrá un avance real en las humanidades? Y los autores se preguntan, ¿hemos utilizado de todas las

³⁰ <https://es.wikipedia.org/wiki/Europeana>

³¹ <https://www.museodelprado.es/coleccion/obras-de-arte>

³² <https://impresso.github.io/CLEF-HIPE-2020/tasks.html>

formas posibles y útiles para los humanistas, las herramientas software disponibles? ¿Son los resultados de las herramientas explicables y entendibles por todos los interesados?

Bibliografía

- [Benavent et al., 2010] Benavent, J; Benavent, X; de Ves, E; Granados, R; García-Serrano, A. (2010) Experiences at ImageCLEF 2010 using CBIR and TBIR Mixing Information Approaches. M. Braschler, D. Harman, E. Pianta (Eds.): CLEF 2010 LABs and Workshops, Notebook Papers, ISBN 978-88-904810-0-0, 22-23 September 2010, Padua, Italy. CEUR Workshop Proceedings, Volume 1176. ISSN1613-0073 2010. <http://ceur-ws.org/Vol-1176/CLEF2010Own-ImageCLEF-BenaventEt2010.pdf>
- [Devlin et al., 2018] Jacob Devlin, Ming-Wei Chang, Kenton Lee, Kristina Toutanova (2018). BERT: Pretraining of Deep Bidirectional Transformers for Language Understanding, Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1.
- [García-Marco, 2011] García-Marco, Francisco-Javier. "La pirámide de la información revisitada: enriqueciendo el modelo desde la ciencia cognitiva". *El profesional de la información*, 2011, enero-febrero, v. 20, n. 1, pp. 11-24. DOI: 10.3145/epi.2011.ene.02
- [García Serrano y Castellanos, 2016] García Serrano, Ana; Castellanos, Ángel. Representación y organización de documentos digitales: detalles y práctica sobre la ontología DIMH. *Revista de Humanidades Digitales*, [S.l.], v.1, p.314-344, oct. 2017. ISSN 2531-1786. <https://doi.org/10.5944/rhd.vol.1.2017.17155>.
- [García-Serrano et al., 2009] Ana García-Serrano, Xaro Benavent, Rubén Granados and José Miguel GoñiMenoyo. (2009) Some Results Using Different Approaches to Merge Visual and Text-Based Features in **CLEF'08 Photo** Collection. In *Evaluating Systems for Multilingual and Multimodal Information Access*. Lecture Notes in Computer Science LNCS, Volume 5706. #pp: 4 Páginas: 568-571. Springer-Verlag Berlin. ISSN: 03029743 (Print) 1611-3349 (Online). ISBN: 978-3-642-04446-5. https://link.springer.com/chapter/10.1007%2F978-3-642-04447-2_69
- [García-Serrano, Menta-Garuz, 2019] García-Serrano A., Menta Garuz A. (2019) Detection of translingual plagiarism with semantic graphs: Experimenting with open resources [Detección de plagio translingüe con grafos semánticos: Experimentando con recursos en abierto]. *Procesamiento del Lenguaje Natural V. 62*, pags: 61-68, ISSN (edición impresa): 1135-5948. ISSN (edición electrónica): 1135-5948, DOI: 10.26342/201962-7
- [Lara-Clares et al., 2021] Lara-Clares, Alicia; Lastra-Díaz, JJ; García-Serrano, A (2021) Protocol for a reproducible experimental survey on biomedical sentence similarity. *PLOS ONE- Registered Report Protocol*. DOI: <https://doi.org/10.1371/journal.pone.0248663>. ISSN: 1932-6203
- [Lastra, García-Serrano, 2015] Lastra, J and García-Serrano, A (2015) A new family of Information Content models with an experimental survey on WordNet. *Knowledge-Based Systems*. V 89 PP: 509-526 Ed: ELSEVIER SCIENCE. ISSN: 0950-7051 Nov 2015. <http://dx.doi.org/10.1016/j.knosys.2015.08.019>
- [Liebeskind, Liebeskind, 2019] Chaya Liebeskind and Shmuel Liebeskind (2019). Deep Learning for Period Classification of Historical Texts. hal-02324617. [https:// hal. archives-ouvertes.fr/hal-02324617](https://hal.archives-ouvertes.fr/hal-02324617)
- [Lin, Lu, 2018] Bill Yuchen Lin and Wei Lu (2018). Neural adaptation layers for cross-domain named entity recognition, Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing, EMNLP 2018, pags 2012-2022, 2018. doi: 10.18653/v1/d18-1226. arXiv: 1810.06368.

[Martinez-Fernandez et al., 2006] Martínez-Fernández, J.L., J. Villena Román, A. M. García-Serrano and J. C. González-Cristóbal (2006) Combining Textual and Visual Features for Image Retrieval Accessing Multilingual Information Repositories Lecture Notes in Computer Science, LNCS V 4022 PÁGINAS: 80-691 Springer-Verlag Berlin ISBN/ISSN 978-3-540-45697-1 / 0302-9743.

https://link.springer.com/chapter/10.1007/11878773_75

[Maupeou, Le Men, 2019] Félicie Faizand de Maupeou and Ségolène Le Men (2019). The Artist Libraries Project in the Labex Les passés dans le present. Journal of Data Mining and Digital Humanities, Episciences.org, 2019.

[McGillivray et al., 2020]. McGillivray, Barbara and Poibeau, Thierry and Ruiz, Pablo. (2020) Digital Humanities and Natural Language Processing: ``Je t'aime... Moi non plus'' Digital Humanities Quarterly, V 14, N 2. <https://hal.archives-ouvertes.fr/hal-02970302>

[Mikolov et al., 2013] Tomas Mikolov, Ilya Sutskever, Kai Chen, Greg Corrado, Jeffrey Dean (2013).

Distributed Representations of Words and Phrases and their compositionality, Advances in Neural Information Processing Systems, pags. 3111-3119, 2013. arXiv: arXiv:1310.4546v1

[Tasovac et al., 2020] Toma Tasovac, Sally Chambers, Erzsébet Tóth-Czifra. Cultural Heritage Data from a Humanities Research Perspective: A DARIAH Position Paper. 2020. hal-02961317

[Vaswani et al, 2017] Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N.

Gomez, Lukasz Kaiser, Illia Polosukhin (2017). Attention Is All You Need. <http://arxiv.org/abs/1706.03762.pdf>

La lexicografía electrónica

Moderadora: Fátima Díez Platas · *Asiste:* Alexandra Iglesias Cortés

- La difusión de la informática como un caso de revolución instrumental.
- Los grandes cambios en la lexicografía contemporánea:
 - De la prescripción a la descripción
 - El empleo de corpus textuales
 - La utilización de recursos electrónicos en la confección y difusión de los diccionarios
- Los diccionarios electrónicos
- Las bases de datos lexicográficas (BDL)
- La explotación de los diccionarios electrónicos

Panel 2

Experiencias en torno a la *Text Encoding Initiative*: Retos de codificación, plataformas de publicación y creación de corpus

Modera: Antonio Rojas Castro · *Asiste:* César González-Pérez

Panelistas: Leyre Martín Aizpuru; Marta López Izquierdo; Víctor Caballero Gómez

Experiencias en torno a la *Text Encoding Initiative*: Retos de codificación, plataformas de publicación y creación de corpus

Palabras clave: *Text Encoding Initiative*, edición digital, etiquetado de textos, anotación semántica

Coordinadora: Susanna Allés Torrent, University of Miami

Participantes:

Antonio Rojas Castro, Berlin-Brandenburgische Akademie der Wissenschaften y Universidad Internacional de la Rioja

Leyre Martín Aizpuru y Víctor Caballero Gómez, Universidad de Sevilla

Marta López Izquierdo, Université Paris 8

Este panel presenta diversas iniciativas relacionadas con el uso de la Text Encoding Initiative (TEI) con el objetivo de reflexionar sobre los retos afrontados a lo largo de su adopción y aplicación: ya sea en su concepción, el modelado de los datos, la creación de infraestructuras para la publicación, la colaboración en el etiquetado de los textos, o incluso en la exploración de los datos textuales desde un punto de vista cuantitativo y translingüístico. La sesión promueve un diálogo constructivo que ilumine los escollos y las carencias del uso de TEI en el seno de la comunidad de usuarios en español, al mismo tiempo que evidencie los nuevos derroteros por los que esta práctica debe discurrir.

La TEI es una de las iniciativas académicas en el ámbito de las Humanidades Digitales más conocidas. Su trayectoria, ahora ya con más de treinta años, y una comunidad global con miles de usuarios, hace de esta iniciativa una de las más sólidas para la codificación informática de textos humanísticos y del ámbito de las Ciencias Sociales. Aunque su uso y su sistema de marcado se ha aplicado especialmente en el campo de la edición de textos, independientemente de la escuela o tradición de crítica textual, cada vez son más los escenarios en los que se adopta TEI, como puede ser la creación de corpus lingüísticos e incluso para la estructuración de contenidos textuales a modo de base de datos.

La TEI se ampara en un consorcio compuesto por personas e instituciones, principalmente norteamericanas y europeas. Este consorcio es el responsable de desarrollar y mantener actualizadas las Guías directrices (*Guidelines*) para el marcado (también llamado etiquetado y codificación) de los textos en formato digital. Estas guías son una especie de manual de codificación o de recomendaciones de buenas prácticas que especifican una metodología concreta de marcado de los textos de manera que puedan ser procesados informáticamente³³. El espíritu de acceso abierto y colaboración del consorcio hace que sean los mismos usuarios los que tengan la capacidad para proponer mejoras y modificaciones a través de su repositorio en GitHub. Este espíritu de colaboración y de difusión se manifiesta también a través de un diálogo constante a través de la lista de discusión³⁴ y de la organización anual de una conferencia. El consorcio, en fin, crea grupos de trabajo y de interés para ciertas disciplinas, como pueden ser diccionarios, lingüística de corpus, descripción de manuscritos, etc. Todo ello hace que la comunidad TEI sea cada vez más amplia, y al mismo tiempo que más especializada.

³³ <https://www.tei-c.org/release/doc/tei-p5-doc/en/html/index.html>

³⁴ Más información sobre la lista de discusión está disponible en: <https://tei-c.org/support/>

En el ámbito del español, la TEI se ha utilizado y estudiado desde finales de los ochenta. Entre las publicaciones pioneras se encuentra las de Faulhaber (1994), Marcos Marín (1986, 1994) o Irizarry (1997), que entendieron ya con precisión las potencialidades de un estándar que podía crear textos enriquecidos, interoperables, preservables y con funcionalidades de búsqueda semántica. A lo largo de los 2000 también fueron apareciendo, aunque tímidamente, escritos divulgativos o centrados en proyectos concretos de investigación (Fradejas Rueda, 2009; o los múltiples trabajos de Alex Bias). También aparecieron las primeras publicaciones sobre las posibilidades de la edición digital académica, que desde entonces han tenido una cierta continuidad (Lucía Megías, 2008). Y más recientemente, han visto la luz otros artículos de alcance más general y centrados sobretudo en España (Spence, 2014; Allés Torrent, 2017; Rojas Castro, 2017), o en casos de estudio de proyectos en curso (Revenga, 2014; González-Blanco et al., 2014; González-Blanco & Rodríguez, 2015; del Río Riande & Zubillaga, 2015; Navarro Colorado 2015; Rojas Castro, 2017). En lo que respecta específicamente a TEI, y sus especificaciones técnicas la literatura en español es todavía poco extensa, aunque esta tendencia va mejorando significativamente (Fradejas, 2009; Allés Torrent, 2015; Rojas Castro, 2018; Schöch et al. 2019).

De manera paralela, fueron surgiendo también desde los años 90 diferentes proyectos que adoptaron TEI como método de marcado de textos. Algunos de los ejemplos más conocidos son la Biblioteca Virtual Miguel de Cervantes (1999), que ofrecía facsímiles con imágenes digitales con cierta calidad. También los corpus del CREA y del CORDE de la Real Academia de la Lengua Española llevó a cabo una codificación y ejecución similar. En ambos casos, desafortunadamente, nunca pusieron a disposición de los usuarios los archivos XML-TEI en abierto para que toda la comunidad pudiera usar como referencia o tratar con ellos como lo que eran, datos informáticos para ser procesables. Desde hace unos años, la comunidad internacional ha puesto de relieve la necesidad de compartir los archivos codificados y esta práctica empieza a ser también adoptada por los proyectos en español. A día de hoy, empiezan a ser más numerosos los ejemplos tanto de ediciones digitales como de corpus lingüísticos que, además de una documentación sobre el tipo de marcado, ofrecen en abierto los archivos fuente en XML-TEI³⁵. Entre los primeros, y especialmente fértil entre los estudios del Siglo de Oro (Allés Torrent, 2017) encontramos los proyectos editoriales de *La Dama Boba* (ed. Prolope, 2015)³⁶, las *Soledades* de Luís de Góngora (ed. Rojas Castro, 2016)³⁷, el corpus de Poesía medieval (2020)³⁸ o las *Siete Partidas Digital* (ed. Fradejas Rueda, 2018)³⁹. Entre los corpus, también se encuentran iniciativas sumamente relevantes como el CLiGS textbox⁴⁰ o DRACOR⁴¹, de carácter internacional y multilingüe y que contienen una sección de corpus etiquetados en español.

La relativa lentitud en la adopción de TEI para proyectos en español se debe también a la escasez de materiales para su aprendizaje. A diferencia de la comunidad anglófona, que desde hace ya más de diez años tiene tutoriales en línea, como el TEI By Example⁴², no ha ocurrido así en nuestro caso. Recientemente se han traducido tutoriales de TEI del inglés (Schreibman et al., 2020), pero el reto es todavía el de la creación de materiales que realmente se adapten a la tradición cultural y tomen ejemplos concretos en esta lengua. Con tal fin, han aparecido iniciativas como el *Text Technologies Hub. Recursos sobre tecnologías del texto y edición digital* (TTHub)⁴³ que ofrece lecciones en línea, de diferente nivel, una bibliografía sobre TEI en español, e incluso una base de datos *in fieri* de los proyectos que ofrecen ejemplos concretos de marcado.

³⁵ Véase una lista con ejemplos sobretudo en español, en <https://tthub.io/recursos/ejemplos-tei/>

³⁶ <http://damaboba.unibo.it/>

³⁷ <http://soledades.uni-koeln.de/>

³⁸ <http://hdlab.space/Poesia-Medieval/>

³⁹ <https://7partidas.hypotheses.org/>

⁴⁰ <https://github.com/cligs/textbox>

⁴¹ <https://dracor.org/span>

⁴² <https://teibyexample.org/>

⁴³ <https://tthub.io/>

Apostando por la diversidad lingüística y cultural, el Consorcio (TEI-C) está explorando actualmente estrategias para aumentar la accesibilidad para aquellos usuarios que trabajan con lenguas diferentes del inglés. Para ello ha creado un grupo de trabajo para la internacionalización (*Internationalization Working Group*) que, entre otros objetivos, tiene el de la creación de un glosario multilingüe de términos utilizados en las guías directrices de TEI y un repositorio de ejemplos en lenguas y orígenes diferentes. Este trabajo irá acompañado de la reestructuración de la página web para que pueda dar una mayor visibilidad a los contenidos y a la dimensión internacional del Consorcio, como por ejemplo, páginas de bienvenida en diferentes lenguas. Otra de las líneas de trabajo es la redacción de materiales introductorios en diferentes lenguas para acercar más fácilmente a los usuarios de otras comunidades lingüísticas, así como el desarrollo de una infraestructura para hacer la traducción de las especificaciones de TEI más ágil. Como es sabido, tanto el sitio de la TEI, como sus guías directrices solo son consultables en inglés, y a día de hoy la traducción en español sigue siendo solo parcial⁴⁴.

Bajo el auspicio de esta iniciativa del Consorcio TEI, financiada este año por la fundación Mellon, este panel pretende analizar el uso y la vitalidad del marcado con TEI en español. En esta ocasión, se agrupan experiencias de investigadores españoles, activos en diferentes centros nacionales e internacionales. Cada una de las intervenciones trata un tema clave para el desarrollo del uso de TEI en español con diferentes niveles de familiaridad y de complejidad. La primera intervención (Rojas Castro) aborda el tema crucial de la infraestructura y la publicación de los archivos en TEI y trae a colación el problema lingüístico en la interfaces web. Las otras dos presentaciones se centran en la creación de corpus textuales a partir del estándar propuesto por la TEI y con materiales textuales de los siglos XIX y XX. Ambos proyectos, CODHECUN y CAREXIL-FR, ofrecen un corpus bien delimitado y han optado por el uso de la plataforma TEITOK, una plataforma web para la visualización, creación y edición de corpus con marcado semántico (como XML-TEI) y anotación lingüística. Estas cuatro iniciativas pioneras ofrecerán una visión de conjunto del estado actual sobre edición digital y creación de corpus y abren las fronteras para que los humanistas digitales puedan finalmente transitar y aprovechar al máximo los datos producidos a través de herramientas de análisis textual.

Ich verstehe gar nichts... Sobre la localización de ediarum al español

Antonio Rojas Castro

Ediarum es un framework de edición en formato TEI diseñado e implementado por TELOTA en la Academia de Ciencias y Humanidades de Berlín-Brandeburgo (BBAW, Alemania) basado en dos componentes principales: una base de datos XML de código abierto (eXistDB) y un editor XML comercial ampliamente utilizado (Oxygen XML editor). Estos dos componentes se comunican mediante una interfaz en alemán utilizada en proyectos de Humanidades Digitales desarrollados en la Academia de Berlín. En 2020 el proyecto de cooperación cubano-alemán Humboldt Digital (ProHD) emprendió un proceso de adaptación de la interfaz de ediarum con el objetivo de facilitar la tarea de edición de textos en formato TEI: por un lado, se desarrollaron las funciones y el código para que fueran independientes de la lengua (internacionalización) y, por el otro, se localizó la interfaz al español mediante la creación de un fichero con las traducciones (localización). Como resultado de este proceso, Proyecto Humboldt Digital ha desarrollado un nuevo framework llamado ediarum.PROHD.edit disponible en Github. El

⁴⁴ Verónica Zumárraga y Marcela Tabanera tradujeron una primera versión de las guías directrices de la TEI en su *Manual para la codificación e intercambio de textos informatizados. Normas de la Text Encoding Initiative* (Marcos Marín, 1994: 89). Más tarde, Alejandro Bia y Carmen Arronis Llopis empezaron la traducción de las guías directrices en su versión P5 pero estas permanecen inconclusas.

objetivo de esta contribución es presentar ediarum.PROHD.edit y reflexionar sobre los retos más importantes encontrados durante la localización.

Corpus Documental y Hemerográfico de la Cuba del Novecientos (CODHECUN): diseño, elaboración y explotación de un corpus en la plataforma TEITOK

Leyre Martín Aizpuru y Victor Caballero Gómez

En el marco del proyecto “Cuba y Andalucía en el siglo XIX: estudio de los lazos lingüísticos y culturales desde las Humanidades Digitales” (<https://institucional.us.es/cuba19/>), estamos realizando la transcripción y edición digital de un corpus de acuerdo con los criterios de la Red CHARTA en su versión digital (Isasi *et alii*, 2020) y en el marco de la plataforma TEITOK, desarrollada por Maarten Janssen (2014). Además de ofrecer este corpus tokenizado, normalizado, lematizado y anotado morfosintácticamente, nos hemos propuesto marcar aspectos léxico-semánticos, con el propósito de elaborar índices temáticos. Ello nos permitirá poner a disposición de otros investigadores los documentos del corpus, así como llevar a cabo un análisis de los usos lingüísticos y discursivos del español del siglo XIX, tomando Cuba como referencia geográfica y centro del espacio variacional caribeño, con el propósito de definir la especificidad del español cubano y sus lazos con la sociedad y la lengua andaluza del Novecientos.

TEI para un proyecto transdisciplinar: construyendo CAREXIL-FR

Marta López Izquierdo

El proyecto en construcción [CAREXIL-FR](https://carexil.univ-paris8.fr/) (CARTas de REpublicanos Españoles REFugiados y EXILIados en FRancia, <https://carexil.univ-paris8.fr/>) nace como un empeño transdisciplinar de edición y estudio de un corpus epistolar inédito, para responder a las necesidades de un equipo que reúne a filólogos, lingüistas, historiadores, cartógrafos, antropólogos, especialistas en humanidades digitales, en literatura y en arte. Por un lado, hemos desarrollado un protocolo de transcripción, edición y anotación basado en los estándares TEI, siguiendo la pauta de otros proyectos de corpus epistolares digitales (Post-scriptum), pero adaptándola a las necesidades del estudio transdisciplinar específico de CAREXIL-FR. Por otro lado, la adopción del entorno TEITOK (Maarten Janssen, 2014) conlleva sus propias normas de modelado y anotación de los documentos que es necesario integrar en la fabricación del corpus. En nuestra intervención, expondremos algunos de los escollos que hemos encontrado a lo largo de nuestro trabajo y las soluciones que estamos proyectando.

Bibliografía

Allés Torrent, S., y del Río Riande, G. (2018). Enseñar edición digital con TEI en español. Aprendizaje situado y transculturación. En *Humanidades Digitales: construcciones locales en contextos globales*. Actas del I Congreso Internacional de la Asociación Argentina de Humanidades Digitales (AAHD), Buenos Aires: Editorial de la Facultad de Filosofía y Letras Universidad de Buenos Aires.

Allés Torrent, S. (2015). Edición digital y algunas tecnologías aliadas. *Insula*, 822: 18–21.

Allés Torrent, S. (2017). Tiempos hay de acometer y tiempos de retirar: Literatura áurea y edición digital. *Studia Aurea*, 11: 13–30.

- Faulhaber, Ch. (1994). La Text Encoding Initiative y su aplicación a la codificación textual y explotación. En *Actas del Congreso de la lengua española: Sevilla, 7 al 10 Octubre, 1992*, 331–40. Madrid: Instituto Cervantes. https://datospdf.com/download/la-text-encoding-initiative-y-su-aplicacion-a-la-codificacion-textual-y-explotacion-_5a44beefb7d7bc422b88bbd1_pdf
- Fradejas Rueda, J. M. (2009-10). La Codificación XML/TEI de Textos Medievales. *Memorabilia*, 12: 219–47. <https://parnaseo.uv.es/Memorabilia/Memorabilia12/PDFs/Codificacion.pdf>
- González-Blanco, E., y Rodríguez, J. L. (2014). ReMetCa: A Proposal for Integrating RDBMS and TEIVerse. *Journal of the Text Encoding Initiative*, 8. <https://doi.org/10.4000/jtei.1274>.
- Irizarry, E. (1997). *Informática y Literatura. Análisis de Textos Hispánicos*. Madrid: Anthropos.
- Lucía Megías, J. M. (2008). Las relaciones entre la bibliografía textual y la informática humanística: el incunable del hipertexto. *Tipofilología. Rivista Internazionale di Studi Filologici e Linguistici sui testi a stampa*, 1: 119–38.
- Marcos Marín, F. (1986). Metodología informática para la edición de textos. *Incipit*, 6: 185–97.
- Marcos Marín, F. (1994). *Informática y Humanidades*. Madrid: Gredos.
- Navarro Colorado, B. (2015). A Computational Linguistic Approach to Spanish Golden Age Sonnets: Metrical and Semantic Aspects. En *Proceedings of the Fourth Workshop on Computational Linguistics for Literature*. Denver. http://www.dlsi.ua.es/~borja/navarro2015_GoldenAgeSonnets.pdf.
- Revenge García, N. (2014). La edición crítica digital de textos teatrales. La estrella de Sevilla como ejemplo. *Anuario de Lope de Vega. Texto, Literatura, Cultura*, 20: 99–121.
- Rio Riande, G., y Zubillaga, C. (2015). Un proyecto de edición digital académica en Argentina. Diálogo Medieval. En *Actas I Jornadas de Humanidades Digitales de la AAHD*. Buenos Aires: Editorial de la Facultad de Filosofía y Letras. <https://www.aacademica.org/gimena.delrio.riande/15>.
- Rojas Castro, A. (2018). La edición académica digital. De las teorías del texto a la visualización de la información. En I. Galina Russell, J. F. Barrón Tovar, M. Peña Pimentel, E. Priani Saisó, D. Domínguez Herbón, and A. Álvarez Sánchez (Eds.), *Humanidades Digitales: Edición, Literatura y Arte*. CDMX: Bonilla Artigas.
- Rojas Castro, A. (2017). La edición crítica digital y la codificación TEI. Preliminares para una nueva edición de las *Soledades* de Luis de Góngora. *Revista de Humanidades Digitales*, 1.0: 4–19. <https://doi.org/10.5944/rhd.vol.1.2017.16379>.
- Schöch, C., Calvo Tello, J., Henny-Krahmer, U., y Popp, S. (2019). The CLiGS Textbox: Building and Using Collections of Literary Texts in Romance Languages Encoded in TEI XML. *Journal of the Text Encoding Initiative*. <https://doi.org/10.4000/jtei.2085>.
- Schreibman, S., y Bleier, R. (2020). La codificación textual y la Iniciativa TEI. *dariahTeach. DH Teaching material*. <https://teach.dariah.eu/mod/page/view.php?id=739>
- Spence, P. (2014). Edición académica en la era digital: modelos, difusión y proceso de investigación. *Anuario Lope de Vega. Texto, literatura, cultura*, 20.0: 47–83. <https://doi.org/10.5565/rev/anuariolopedevega.74>.

Edición digital y medios digitales

Modera: Sagrario López Poza · *Asiste:* César González-Pérez

- 56 Del texto impreso antiguo a la edición académica digital. Avances en la configuración del repositorio documental BIDISO TEXTOS
Nieves Pena, Carlota Fernández Travieso, Manuel Garrobo Peral, Stefano Bazzaco, Mónica Martín Molares
- 3 Una propuesta de edición digital de sermones medievales hispanos
Oriol Catalán
- 33 Interfaces de ediciones digitales académicas: modelos y expectativas
María Luisa Alvite Díez
- 66 Razones y reflexiones para una meta-historia digital de las literaturas ibéricas usando TEI y *Open Linked Data*
Santiago Pérez Isasi
- 74 ENTRIB Entremeses Ibéricos: hacia una edición sinóptica digital
José Pedro Sousa

Del texto impreso antiguo a la edición académica digital. Avances en la configuración del repositorio documental BIDISO TEXTOS

Palabras clave: Edición digital, Transcripción automática, Modelización, Textualidad digital, Repositorio documental, Relaciones de sucesos

Nieves Pena Sueiro (UDC)

Stefano Bazzaco (Universidad de Verona) y Mónica Martín Molares (UDC)

Carlota Cristina Fernández Travieso (UDC) y Manuel Garrobo Peral (Universidad de Verona)

Descripción del problema de investigación, metodología y resultados

El proyecto *Biblioteca Digital Siglo de Oro* ofrece al investigador interesado cinco bibliotecas digitales sobre Literatura Emblemática, Relaciones de Sucesos, Divisas o empresas históricas, Polianteas e inventarios y bibliotecas. Las bibliotecas digitales dan acceso a bases de datos (ocho de fuentes primarias y una que almacena bibliografía o fuentes secundaria) que contienen, además de los datos bibliográficos y la localización actual de los ejemplares, imágenes o enlaces a ediciones facsimilares, de gran interés para la investigación en la Literatura, la Historia, la Historia del libro y bibliotecas y la Historia del Arte de los siglos XVI y XVII.

Una necesidad urgente para los humanistas, constatada y comentada con otros investigadores, es el acceso en línea a las fuentes primarias a partir de textos filológicamente fiables. Es ya un lugar común el hecho de que la digitalización masiva ha facilitado el acceso a los textos, pero ¿a qué textos? Es cierto que fácilmente encontramos imágenes facsimilares de algunas obras y también transcripciones u adaptaciones de otras que, lamentablemente, difieren mucho de lo que escribió su autor (bien por una falta de cuidado en la transcripción, bien por la ausencia o dejadez de la puntuación, bien por haber pasado un OCR y no haber sido revisado correctamente). En los trabajos humanísticos es fundamental aplicar el rigor filológico y acudir, como base de la investigación científica, a fuentes fiables y, por lo tanto, a ediciones elaboradas con criterios fundamentados.

Todos estos motivos fueron los que nos llevaron a crear un corpus documental o banco de textos de acceso libre, que permitiese no solo la lectura facsimilar, sino también la lectura modernizada y la descarga en diversos formatos. Así nació el proyecto BIDISOTextos, un repositorio donde se reunirán, para su lectura, uso, difusión y explotación, textos transcritos en formato digital, enriquecidos con marcación formal y semántica, garantizando así su preservación. Sostenibilidad y utilidad para la comunidad.

Las colecciones que integrarán el repositorio BIDISO-Textos estarán vinculadas a las fuentes primarias de nuestras bibliotecas digitales de BIDISO, añadiendo ahora, al formato digital facsimilar el incremento del texto transcrito, incluso con marcación, que será de enorme interés para explotar el contenido desde distintos enfoques y áreas de conocimiento (Historia, Historia de la lengua, Literatura, Historia del periodismo, Historia del libro...)

El objetivo principal de BIDISO-Textos es proporcionar el acceso abierto y gratuito a los textos transcritos de un amplio corpus que permitan responder a las necesidades planteadas por la comunidad investigadora en un formato buscable, susceptibles de ser analizados sistemáticamente, pero, sobre todo, estableciendo un texto base con suficientes garantías desde un punto de vista filológico. Pero, si bien en los últimos años se ha producido un importante avance en los estudios sobre edición académica digital (*vid.*, entre otros, Allés Torrent, 2015; Pierazzo, 2015; Driscoll y Pierazzo eds., 2016; Rojas Castro, 2016; Fiormonte, 2018, etc.) para la Literatura Española se echan en falta todavía guías procedimentales y metodologías y, sobre todo, la utilización de un

modelo estándar, la adopción de criterios comunes por quienes deciden sumarse a la edición digital, y el uso de herramientas de *software* libre (con todo lo que esto implica en cuanto a costes, mantenimiento y dependencia tecnológica), que es todavía escaso. Fue aquí donde encontramos los principales escollos, en la búsqueda de un modelo idóneo, de un estándar abierto y compatible con la comunidad.

Para la realización de este repositorio se han establecido las siguientes fases:

- Definir el objeto y finalidad de nuestro proyecto editorial: delimitar un corpus y explicitar qué queremos ofrecer a los usuarios.
- Localización de las fuentes primarias: impresos de los siglos XVI a XVIII, ediciones diplomáticas o ediciones críticas –basadas en un único testimonio o en varios– publicadas o inéditas, de las que podamos disponer legalmente para nuestro proyecto, ya por haber considerarse obras en dominio público, ya por contar con la autorización expresa de sus responsables intelectuales, ya por ser editores de las mismas.
- Obtención de textos transcritos a partir de sistema OCR, digitalización HTR o transcripción manual, según el tipo de material de partida. Codificación de Relaciones de Sucesos: selección de etiquetas TEI y flexibilidad según la naturaleza de los materiales.
- Visualización y descarga: diferentes soluciones (TXT, RTF, PDF, XML-TEI, FACSIMILAR)

El proyecto de repositorio tiene como objetivo ofrecer textos relativos a las líneas de investigación y contenidos de las bibliotecas digitales que contiene BIDISO (y enlazados con estas), organizados en tres colecciones:

- Colección 1, de *Relaciones de sucesos* breves (pliego suelto).
- Colección 2, de *Relaciones de sucesos en formato libro*
- Colección 3, de *libros de emblemas*

Hemos comenzado a trabajar con una selección de textos noticieros breves (relaciones de sucesos), estableciendo un pequeño corpus de prueba. Tras localizar las fuentes primarias, hemos dado los primeros pasos en la digitalización y trabajo con OCR, así como en la selección de etiquetas TEI y codificación.

Este panel presenta los avances realizados en la transcripción automática de relaciones de sucesos y en la edición digital de relaciones de sucesos, así como los problemas que se fueron planteando por las características de estos textos y las soluciones adoptadas.

Transcripción semi-automática de Relaciones de Sucesos con la herramienta de HTR Transkribus (READ Coop))

Recientemente la transcripción automatizada de *historical texts* ha experimentado unos significativos avances gracias a la difusión de herramientas de HTR (*Handwritten Text Recognition*) que emplean redes neuronales y consienten entrenar la máquina por medio de *Golden Standard transcriptions* (Mühlberger *et al.*, 2019; Bazzaco, 2018 y 2020). Dentro de este conjunto, la plataforma *Transkribus* (READ Coop), se reveló como un instrumento fundamental para la transcripción de las Relaciones de Sucesos que constituyen el corpus de estudio del proyecto BIDISO. En la presente comunicación se describen las fases principales de este proceso, indicando las funcionalidades que ofrece la plataforma, cuáles han sido los resultados obtenidos y su posible explotación para futuros avances del proyecto.

En este marco se ha realizado un corpus piloto de Relaciones de Sucesos con el que se ha estado trabajando en Transkribus, prestando atención especial los elementos identificadores de las relaciones escogidas: el tipo de letra del impreso, la disposición del texto o los elementos que configuran la plana (reclamos, firmas, letras capitulares, etc.), entre otros aspectos. Se describirán los principales escollos derivados del corpus de trabajo que seleccionamos a la hora de llevar a cabo la experimentación con *Transkribus* y las soluciones planteadas.

Avances en el modelo de codificación de la edición académica digital de Relaciones de Sucesos en la Biblioteca Digital del Siglo de Oro.

La concreción de un modelo de codificación es una fase esencial en el establecimiento de una metodología para la creación de una edición académica digital (Mancinelli y Pierazzo, 2020: 92). En esta tarea, la utilización del formato de la Text Encoding Initiative (TEI) <<https://tei-c.org>> basado en el metalenguaje de marcado XML se ha mostrado muy eficaz, convirtiéndose en el estándar más utilizado en el ámbito de las Humanidades. Si bien este estándar cuenta con límites, presenta ventajas concretas para nuestro proyecto como la de ofrecer, con sus 21 módulos y en torno a 500 elementos, etiquetas suficientes para atender a los requerimientos de nuestro proyecto o la posibilidad de aprovechar soluciones estandarizadas que facilitan una visualización amigable de los archivos XML, como EVT.

Esta comunicación partirá de la consideración de la dificultad que supone tomar como punto de partida para la obtención de los textos materiales que requieren diversos tipos de tratamiento (OCR, HTR o transcripción manual) y podrán atender a diferentes criterios de representación gráfica y acercamiento a las fuentes. Seguidamente, nos centraremos en los metadatos y fenómenos textuales que creemos preciso poner de relieve en las obras del género y en qué etiquetas del esquema TEI seleccionamos para marcarlos, circunscribiendo los planteamientos propuestos por Fernández Travieso (2013) a los objetivos actuales de nuestro proyecto de edición digital de relaciones de sucesos. Así mismo, mostraremos, a través de los ejemplos del corpus de prueba seleccionado, cómo nuestro modelo de codificación está dotado de una modularidad que nos permite incluir textos de Relaciones de sucesos que provienen de fuentes primarias de diferentes tipos en un mismo proyecto de edición, estableciendo una codificación sistemática, pero diferenciada. Se atenderá, entre otros, al caso de los textos provenientes del tratamiento con *Transkribus* o de la transcripción manual llevada a cabo por Garrobo Peral (2020) al realizar la edición crítica de la *Fiesta que se hizo en*

Aranjuez a los años del rey nuestro señor don Felipe III... , escrita por D. Antonio de Mendoza...

Bibliografía citada

- Allés Torrent, Susanna, "Edición digital y algunas tecnologías aliadas", *Ínsula*, 822 (2015), 18-21.
- Baena Sánchez, F.; Fernández Travieso, C.; Espejo Cala, C. y Díaz Noci, J. (2014). «Codificación y representación cartográfica de noticias. Aplicación de las humanidades digitales al estudio del periodismo de la Edad moderna». *El profesional de la información* 23, 5: 519526.
<https://doi.org/10.3145/epi.2014.sep.09>
- Bazzaco, Stefano, «El Progetto Mambrino y las tecnologías OCR: estado de la cuestión», *Historias Fingidas* 6 (2018), 257-272 < <http://historiasfingidas.dlcs.univr.it/index.php/hf/article/view/89/135> >.
- Bazzaco, Stefano, «El reconocimiento automático de textos en letra gótica del Siglo de Oro: creación de un modelo HTR basado en libros de caballerías del siglo XVI en la plataforma Transkribus», *Janus*, 9 (2020), 534-561 < <https://www.janusdigital.es/articulo.htm?id=160> >.
- Bordalejo, Barbara, «The texts we see, the works we imagine. The Shift of Focus of Textual Scholarship in the Digital Age», *Ecdotica*, VII (2010), 64-75.
- Driscoll, Matthew y Elena Pierazzo (eds.), *Digital Scholarly Editing. Theory, Practice and Future Perspectives*, Cambridge, Open Book Publishers, 2016.

- Fernández Travieso, Carlota, *Estudio de codificación XML/TEI para Relaciones de Sucesos españolas*, A Coruña, SIELAE, 2013 <<https://www.bidiso.es/sielae/upload/estaticas/file/FTCXMLTEI1.pdf>>.
- Fiormonte, Domenico, *Per una critica del testo digitale. Letteratura, filologia e rete*, Roma: Bulzoni, 2018.
- Garrobo Peral, M. (2020), *Fiesta que se hizo en Aranjuez por los años de Felipe IV en 1622. Hacia una edición digital académica de una relación de sucesos* [TFM inédito del Máster en Literatura Cultura y Diversidad de la Universidade da Coruña, dirigido por Nieves Pena Sueiro y Stefano Neri].
- Italia, Paola, *Editing 2000. Per una filologia dei testi digitali*, Roma, Salerno Editrice, 2020.
- Kichuk, Diana, «Quantità e qualità dei testi online: il problema della digitalizzazione di massa», en *Teoria e forme del testo digitale*, introducción, edición y notas de Michelangelo Zaccarello, Roma, Carocci Editore, 2019, pp. 135-166.
- Mancinelli, Tiziana y Elena Pierazzo, *Che cos'è un'edizione scientifica digitale*, Roma: Carocci, 2020.
- Mühlberger, Günter *et al.*, «Transforming scholarship in the archives through Handwritten Text Recognition. Transkribus as a case study», *Journal of Documentation - Emerald Publishing*, vol. 75, n. 5 (2019), pp. 954-976 <<https://www.emerald.com/insight/content/doi/10.1108/JD-07-2018-0114/full/html>>.
- Pierazzo, Elena. *Digital Scholarly Editing: Theories, Models and Methods*. University of Grenoble: Hal, 2015 <hal-01182162>.
- Rojas Castro, Antonio, "La edición crítica digital y la codificación TEI. Preliminares para una nueva edición de las Soledades de Luis de Góngora", *Revista de Humanidades Digitales* n.1 (2016), 4-19.

Una propuesta de edición digital de sermones medievales hispanos

Oriol Catalán. Investigador posdoctoral.
Universitat Pompeu Fabra, Barcelona

El proyecto de edición digital Antología de sermones Medievales Hispanos es continuación y está vinculado con la base de datos [PredMedDB](#) de sermones medievales hispanos iniciada hace algunos años y con la que comparte características.

Ambos proyectos surgen de una investigación posdoctoral en el marco de un proyecto del MINECO y de la necesidad de construir infraestructuras digitales tanto para desarrollar la investigación como para divulgarla. Otra característica común de ambos proyectos es la elaboración de todo el código por parte del investigador, y por lo tanto, con un coste prácticamente nulo. A pesar de la modestia del proyecto, se ha intentado estar al día de las tendencias en Edición digital (Driscoll–Pierazo 2016). El proyecto está alojado, como la base de datos comentada, en los servidores de la Universitat Pompeu Fabra.

El proyecto de edición se encuentra actualmente en fase de desarrollo en <http://predmed.upf.edu/edismhi.html> y prevé tener editados unos 15 sermones para octubre de 2021. El proyecto tiene una doble vertiente de investigación y divulgación. Por un lado, se pretende editar o reeditar sermones importantes de la Edad Media hispana. Algunos sermones están ya magníficamente editados, pero muchos otros no han sido editados (por ejemplo, se intentará editar sermones de Nicolau Eimeric y de Alonso de Espina), o lo han sido de una forma poco adecuada o en publicaciones de difícil acceso. Por otro lado, se quiere editar los sermones medievales hispanos más destacados con un mismo criterio editorial y en un único espacio para facilitar y promover su consulta y aprovechamiento.

En un primer momento, no se veía posible presentar toda la información deseada en pantalla de una forma estéticamente agradable. Se pensó entonces en mantener una edición en PDF paralela a la edición en línea. La evolución del proyecto, sin embargo, y especialmente el uso de las funciones de ocultar contenidos, ha hecho posible visualizar un gran número de opciones en un formato de lectura agradable, y ha hecho abandonar por lo tanto la idea de una edición en PDF de los textos.

Una de las ventajas, y no menor, de esta decisión, es la de reducir el trabajo de conversión del etiquetado y de revisión de la edición. Pero no es únicamente una decisión basada en la necesidad de reducir el volumen de trabajo. Se pretendía también, que la edición fuese totalmente digital, y no una edición tradicional digitalizada, y que condensara la mayor cantidad posible de información en un único formato, que permite al mismo tiempo ofrecer toda la información necesaria y tener una agradable experiencia de lectura.

Una edición digital en pantalla con XML y Javascript.

La primera de las ediciones previstas, y la que ha finalmente prosperado, es una edición en línea a partir de un etiquetado XML y una búsqueda de etiquetas mediante Javascript. Esta edición se basa en el etiquetado TEI según las instrucciones del *TEI P5: Guidelines for Electronic Text Encoding and Interchange*, alejándose de este etiquetado solo en los casos en que Javascript daba problemas de lectura de TEI (Por ejemplo, el atributo *xml:lang* ha sido sustituido por el atributo *lang*). Una conversión a un XML-TEI puro, en caso de necesitarse, se podría realizar fácilmente.

Un objetivo del proyecto era compaginar la fidelidad de la edición con una agradable experiencia de lectura.⁴⁵ Para ello se ha optado por presentar al principio un texto sin apenas marcaciones, dando la opción de mostrar y ocultar los contenidos propios de cualquier edición mediante la selección de casillas de verificación (checkboxes). También se ha optado por el uso de la familia de fuentes Libre Baskerville del repositorio de GoogleFonts, y se ha presentado el texto en un tamaño algo mayor al habitual para mejorar la experiencia de lectura, un aspecto a veces poco cuidado en las ediciones digitales.

La página de inicio del proyecto presenta la lista de sermones editados (Ilustración 1). Una vez seleccionado el sermón deseado, aparece en una página con la información mínima del sermón (localización y *thema*), el texto transcrito y las posibilidades de visualización de la información (ver Ilustración 2).

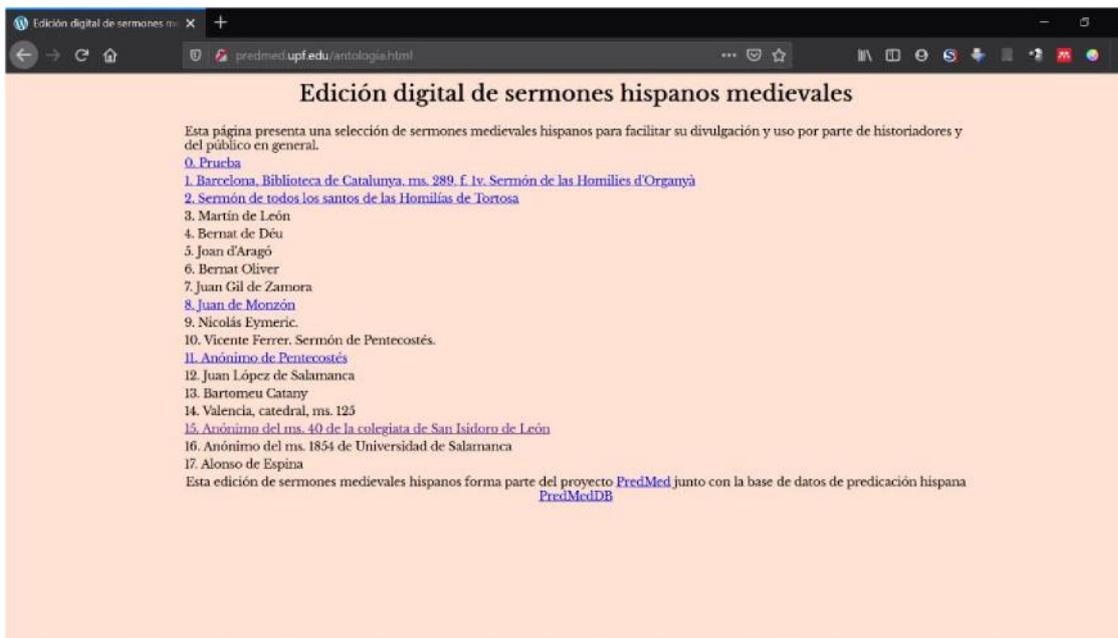


Ilustración 1. Borrador de la página de acceso a la edición.

⁴⁵ “Editors no longer have to choose between faithfully representing their sources and making their text readable” (Pierazzo 2016, 308)

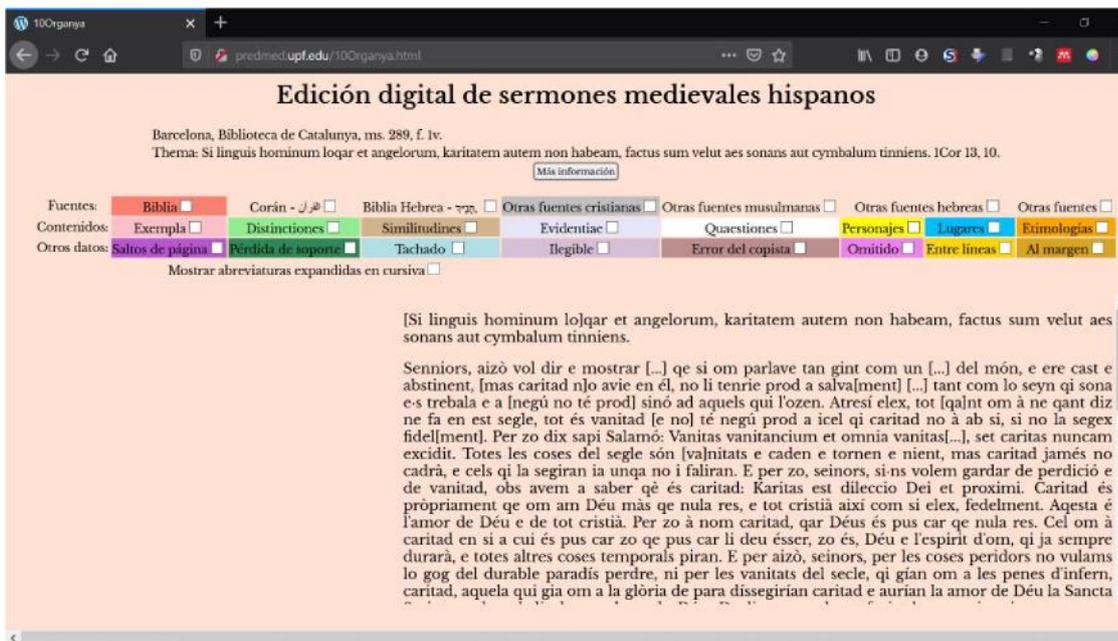


Ilustración 2. Borrador de la edición en línea a marzo de 2021 con la presentación inicial del texto.

A partir de ahí, se pueden visualizar contenidos y características del texto mediante la selección de las casillas de verificación (checkbox) entre las diversas opciones ofrecidas. La selección de las casillas colorea los resultados en el texto del sermón y, en algunos casos, se listan los resultados en pantalla (Ilustración 3). Tanto las listas como el coloreado de los datos se pueden ocultar deseleccionando la casilla correspondiente si se prefiere una lectura limpia del texto. La visualización de datos permite la selección de varias casillas a la vez, por lo que se pueden visualizar los contenidos deseados en cada momento. Esto permite, tal y como sugiere últimamente algunos autores (Rojas Castro 2016; González Blanco 2017, 254), ofrecer una edición edición dinámica o interactiva de los documentos.

Las categorías etiquetadas y por lo tanto visualizables son:

- Fuentes: Fuentes bíblicas, otras fuentes cristianas, Corán, otras fuentes musulmanas, Biblia hebrea, otras fuentes hebreas y otras fuentes. La inclusión de fuentes musulmanas y hebreas se ha hecho con el objetivo de editar también textos musulmanes y hebreos, para lo que se está negociando con expertos en estos campos para la inclusión de sermones no cristianos en el portal.
- Contenidos del sermón: , ejemplos, distinciones, similitudines, evidentiae, quaestiones, . personajes, lugares y etimologías.
- Características del texto: saltos de página, pérdida de soporte, texto tachado, texto ilegible, errores del copista, texto omitido, texto entre líneas, texto al margen y expansión de las abreviaturas (en cursiva).

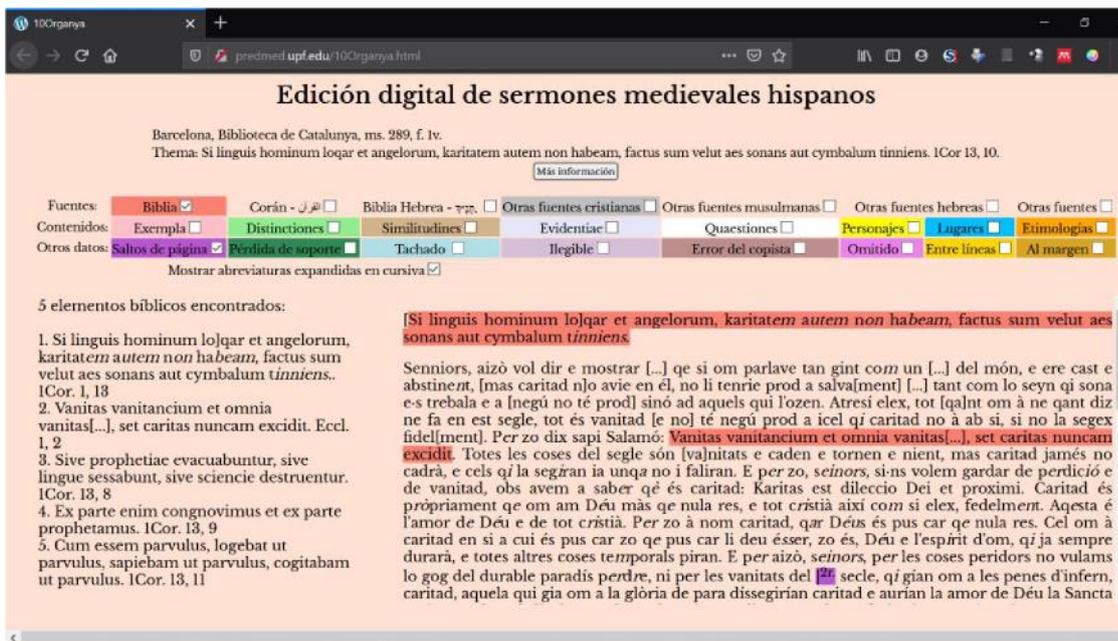


Ilustración 3. Borrador de la edición en línea a marzo de 2021 con las fuentes bíblicas coloreadas y en listado, los saltos de página coloreados y el desarrollo de las abreviaturas en cursiva.

Las etiquetas fundamentales utilizadas son:

- <cit> Citas, organizadas mediante los atributos *type* y *source*.
- <pb> Saltos de página
- <supplied> Texto omitido o perdido recuperable ya sea por pérdida de soporte o por otro motivo.
- <gap> Texto omitido o perdido no recuperable ya sea por pérdida de soporte o por otro motivo.
- <add> Texto añadido por el copista (sobre la línea, al margen...)
- Texto suprimido o tachado por el copista
- <surplus> Texto repetido
- <sic> Error del copista
- <corr> Correcciones al texto.
- <expan> Expansión de abreviaturas

El uso del language Javascript, compatible con los sistemas operativos y navegadores más habituales, permite la compatibilidad con la práctica totalidad de los ordenadores actuales.

Además de la información mínima del sermón (localización y thema) en la parte superior de la pantalla, se ha añadido más datos importantes sobre los sermones (autor, lugar y fecha de predicación o redacción, ediciones previas, y una pequeña introducción y comentario al sermón) que se pueden consultar apretando el botón "Más información" (Ilustración 4). Dicha información es ocultable para no entorpecer la lectura del texto y mantener limpia la pantalla.

Edición digital de sermones medievales hispanos

Barcelona, Biblioteca de Catalunya, ms. 289, f. iv.
 Thema: Si linguis hominum loqar et angelorum, karitatem autem non habeam, factus sum velut aes sonans aut cymbalum tinniens. 1Cor 13, 10.

Autor / predicador: Anónimo Lugar de predicación: Desconocido Fecha de predicación o redacción: Hacia 1200

Ediciones previas: Miret i Sans, Joaquim, "El més antig text literari escrit en català. Precedit per una colecció de documents dels segles XIè, XIIè, i XIIIè", *Revista de Bibliografia Catalana. Catalunya-Balears-Rosselló-València*, 7 (1904), pp. 5-47 i 215-220.

Introducción: Se trata de una de las Homilias de Organyà, uno de los más antiguos textos escritos en catalán (c. 1200), encontradas en 1904 por Miret i Sans en el pueblo pirenaico de Organyà. Relacionadas con las Homilias de Tortosa (ver sermón 2) y con la predicación al norte de los Pirineos, son una excepción en la escritura de sermones medievales, que no serán escritos en lengua vernácula hasta finales del siglo XIV.

Comentario: El sermón trata el tema de la caridad con varias referencias a la Carta de San Pablo a los Corintios.

Fuentes: Biblia Corán - القرآن Biblia Hebrea - התנ"ך Otras fuentes cristianas Otras fuentes musulmanas Otras fuentes hebreas Otras fuentes

Contenidos: Exempla Distinciones Similitudines Evidentiae Questiones Personajes Lugares Etimologías

Otros datos: Saltos de página Pérdida de soporte Tachado Illegible Error del copista Omitido Entre líneas Al margen

Mostrar abreviaturas expandidas en cursiva

5 elementos bíblicos encontrados:

- Si linguis hominum loqar et angelorum, karitatem autem non habeam, factus sum velut aes sonans aut cymbalum tinniens. 1Cor. 1, 13
- Vanitas vanitancium et omnia vanitas[...], set caritas nuncam excidit. Eccl. 1, 2
- Sive prophetiae evacuabuntur, sive lingue sessabunt, sive sciencie destruentur. 1Cor. 13, 8

[Si linguis hominum loqar et angelorum, karitatem autem non habeam, factus sum velut aes sonans aut cymbalum tinniens.

Senniors, aizò vol dir e mostrar [...] qe si om parlave tan gint com un [...] del món, e ere cast e abstinent, [mas caritat njo avie en el, no li tenie prod a salva]ment [...] tant com lo seyn qi sona e s trebala e a [negú no té prod] sinó ad aquels qui l'ozen. Atresi el ex. tot [qajnt om à ne qant diz ne fa en est segle, tot és vanitat [e no] té negú prod a icel qí caritat no à ab si, si no la segeg fidel]ment]. Per zo dix sapi Salamó: Vanitas vanitancium et omnia vanitas[...], set caritas nuncam excidit. Totes les coses del segle són [va]nitats e caden e tornen e nient, mas caritat jamés no cadrà, e cels qí la segiran ia unqa no i faliran. E per zo, seiors, si-ns volem gardar de perlició e de vanitat, obs avem a saber qé és caritat: Caritas est dileccio Dei et proximi. Caritat és

Ilustración 4. Borrador de la edición en línea a marzo de 2021 con la información del sermón desplegada.

¿Por qué una interfaz propia?

Aunque ciertamente las plataformas de edición digital se están multiplicando últimamente, las probadas para desarrollar el proyecto no convencieron por alguno de los siguientes motivos:

- No permitían visualizar la información que se deseaba presentar.
- Eran demasiado estáticas y no permitían cambiar los elementos visualizados.
- Eran de muy compleja instalación, gestión, o tenían una curva de aprendizaje muy lenta.
- Publican las imágenes de los documentos, de las que no tenemos los derechos de reproducción, y por lo tanto no puedo hacer públicas.
- Están diseñadas para textos con más de un ejemplar, mientras que la mayoría de sermones medievales se encuentran en manuscritos únicos.

Estos inconvenientes, junto a la evidencia que los documentos .xml serán utilizables más adelante si alguna plataforma se adapta mejor a los sermones aquí editados en un futuro, me decidió a crear una interfaz propia, que se adapta a las necesidades del proyecto.

Acceso a la edición

Además de la página de acceso ya comentada, los sermones serán accesibles desde el registro correspondiente de la base de datos [Predmed](#). Tanto el etiquetado XML como el código Javascript se subirán a Github y serán de libre acceso.

Precedentes y aprendizajes

Uno de los proyectos que más nos ha influido ha sido [Sermones.net](#). Sin embargo, se ha querido pulir algunas características de dicho proyecto, como el nombre de las páginas, algo más claras y fáciles de citar en nuestro proyecto. Otro proyecto fundamental ha sido el [Thesaurus Exemplorum Medii Aevi](#), recientemente modernizado y actualizado. Una vez ya en marcha el proyecto, he encontrado otros proyectos que han utilizado soluciones parecidas a las aquí tomadas, como la visualización por colores usada en [de Heresy](#), (proyecto que además incluye índices de autores, lugares, un diseño moderno e intuitivo y una agradable experiencia de lectura) o la opción de seleccionar diversas ediciones de un texto elegida en la edición digital de [La Entretenida](#) de Cervantes.

Bibliografía

TEI P5: Guidelines for Electronic Text Encoding and Interchange. Text Encoding Initiative Consortium, 2015, <<https://tei-c.org/guidelines/p5/>>

Driscoll, Matthew James; Pierazzo, Elena (eds.), *Digital Scholarly Editing: Theories and Practices*. Open Book Publishers, 2016.

González-Blanco, Elena, «La edición digital de textos literarios: planteamientos y perspectivas de futuro», *RILCE* 33.1 (2017), pp. 239-58

Rojas Castro, Antonio, «La edición crítica digital y la codificación TEI. Preliminares para una nueva edición de las Soledades de Luis de Góngora», *Revista de Humanidades Digitales* 1 (2016), p. 4-19. <<https://doi.org/10.5944/rhd.vol.1.2017.16379>>

Interfaces de ediciones digitales académicas: modelos y expectativas

María Luisa Alvite Díez

Palabras clave: Ediciones digitales académicas; Estándares; Interacción; Interfaces; Visualización

Resumen:

La definición de “Edición Digital Académica” (EDA), los desafíos para su desarrollo y los paralelismos entre los modelos de edición analógicos y digitales han sido ampliamente tratados en la bibliografía (Allés Torrent, 2020; González-Blanco, 2017; Pierazzo, 2015; Presotto, 2018; Sahle, 2016a, 2016b; Spence, 2014; Rojas Castro, 2018). Asimismo, la literatura hace referencia a un cambio de paradigma con respecto a la edición digital y corrobora la distinción firme de Sahle (2016) entre edición digital y edición digitalizada. Esto es, las ediciones digitales no son meras publicaciones en formato digital; son sistemas de información que adoptan una metodología determinada por un paradigma digital, mientras que las ediciones tradicionales siguen una metodología acorde a los paradigmas de la cultura impresa.

Como apostilla Spence (2014), una edición impresa conduce hacia una publicación concreta, mientras que editar en la era digital supone una mayor ambición, una perspectiva global dentro de la que se ha de poder incorporar una visión dinámica del texto dentro de la investigación. Señala Rojas Castro (2018) que, tras décadas de experimentación, se ha demostrado que la representación con lenguaje de marcado (XML-TEI) del establecimiento del texto permite superar lo que en el medio impreso daba lugar a un tipo u otro de edición (diplomática, modernizada o facsimilar), la EDA es capaz de conciliar distintas teorías del texto sobre una misma obra. Sin embargo, como el propio Rojas Castro subraya, de nada sirve etiquetar toda esta información si la publicación final consiste en un HTML sin interactividad o un simple PDF o ebook.

En esta línea, Allés Torrent (2020), sugiere que el modelo de edición crítica digital debe avanzar hacia la articulación de una metodología digital que refleje la fluidez del texto, atendiendo a la materialidad de las fuentes primarias, así como a la múltiple representación final destinada al lector o usuario.

En este contexto, nuestro trabajo pretende conocer el estadio actual de desarrollo de las interfaces web de ediciones digitales académicas. Para este fin, proponemos, en primer lugar, revisar los modelos o directrices establecidos para su evaluación (Galleron, et al, 2018; MLA, 2011, 2016; Sahle -Institut für Dokumentologie und Editorik- (2016a) en los ítems referidos a la publicación y presentación de EDAs. En segundo lugar, establecidos los criterios de calidad aplicables, se analizan una muestra representativa de EDAs de literatura hispánica.

El estudio pretende contribuir al debate sobre las expectativas de la edición digital académica relacionadas, de un lado, con las soluciones técnicas, funcionalidades, posibilidades de búsqueda e interacción, calidad y opciones de visualización, interoperabilidad, etc., y, de otro, con las posibles políticas sobre la sostenibilidad y preservación de este tipo de trabajos científicos que demandan infraestructuras académicas estables y, en consecuencia, acciones institucionales y colaborativas ineludibles.

Referencias

- Alles Torrent, S. (2020). Crítica textual y edición digital o ¿dónde está la crítica en las ediciones digitales?. *Studia Aurea*, 14, 63-98. <https://doi.org/10.5565/rev/studiaaurea.395>
- Allés Torrent, S. (2017). «Tiempos hay de acometer y tiempos de retirar»: literatura áurea y edición digital. *Studia Aurea: Revista de Literatura Española y Teoría Literaria Del Renacimiento y Siglo de Oro*, 11, 13–30. <https://doi.org/10.5565/rev/studiaaurea.261>
- Apollon, D., Bélisle, C. y Régnier, P. (Eds.) (2014). *Digital Critical Editions*. University of Illinois Press.
- Burgos Segarra, G. (2017). De la edición tradicional a la edición digital. Estudio y edición crítica digital compleja de La discreta enamorada de Lope de Vega. En C. Mata Induráin, y S. I. Santa Aguilar (Coords.), *“Posside sapientiam”*. *Actas del VI Congreso Internacional Jóvenes Investigadores Siglo de Oro, JISO 201638* (pp. 47–58). BIADIG. Biblioteca áurea digital del GRISO, 38. <https://hdl.handle.net/10171/43470>
- Fernández Riva, G., y Millet, V. (2019). La variación textual y su visualización en las ediciones digitales de textos medievales. En *Humanidades Digitales: Miradas hacia la Edad Media* (pp. 64–77). De Gruyter. <https://www.degruyter.com/view/title/535472>
- Fiormonte, D. (2018). *Per una critica del testo digitale: letteratura, filologia e rete*. Bulzoni editore.
- Galleron, I., Demonet, M.-L., Meynard, C., Fatiha, I., Pierazzo, E., Williams, G., Buard, P.-Y., Roger, J., y Ortega Deballon, P. (2018). *Las publicaciones digitales de corpus de autores - guía de trabajo, plantilla de análisis y recomendaciones* (Trad. al español por Paloma Ortega Deballon, rev. por Amelia del Rosario Sanz Cabrerizo y Fatiha Idmhand). Consortium Cahier. <https://halshs.archives-ouvertes.fr/halshs-02164065>
- González-Blanco García, E. (2017). La edición digital de textos literarios: planteamientos y perspectivas de futuro. *RILCE: Revista de Filología Hispánica*, 33(1), 239–258. <https://doi.org/10.15581/008.33.1.239-58>
- MLA - Modern Language Association. Committee on Scholarly Editions. (2011) *Guidelines for Editors of Scholarly Editions*. MLA. <https://www.mla.org/Resources/Research/SurveysReports-and-Other-Documents/Publishing-and-Scholarship/Reports-from-the-MLA-Committee-on-Scholarly-Editions/Guidelines-for-Editors-of-Scholarly-Editions>
- MLA - Modern Language Association. Committee on Scholarly Editions. (2016). *MLA Statement on the Scholarly Edition in the Digital Age*. MLA. <https://www.mla.org/content/download/52050/1810116/rptCSE16.pdf>
- Pierazzo, E. (2015). *Digital Scholarly Editing: Theories, Models and Methods*. <http://hal.univgrenoble-alpes.fr/hal-01182162>
- Presotto, M. (2018). La edición crítica digital del teatro clásico español: balance y perspectivas. *Cuadernos AISPI: Estudios de Lenguas y Literaturas Hispánicas*, 11, 31–46. <http://dx.doi.org/10.14672/O.2018.1426>
- Review journal for digital editions and resources (RIDE)*. <https://ride.i-d-e.de/>

Rojas Castro, A. (2018). La edición académica digital. De las teorías del texto a la visualización de la información. En I. Galina Rusell et al., *Humanidades digitales: edición, literatura y arte*. Bonilla Artigas.

Ruiz Soto, H., y Pezzini, S. (2017). Editar a Góngora con instrumentos digitales: bases metodológicas y perspectivas teóricas para un hipertexto de la polémica gongorina.

Revista de Humanidades Digitales, 1, 35–56. <https://doi.org/10.5944/rhd.vol.1.2017.16682>

Sahle, P. (2016a). *Criterios para la reseña de ediciones digitales académicas (EDA), versión 1.1*. (en colaboración con G. Vogeler y los miembros del Institut für Dokumentologie und Editorik; traducción al español por S. Allés Torrent, basada en las versiones 1.1 alemán e inglés de 2014). <https://www.i-d-e.de/publikationen/weitereschriften/criterios-version1-1/>

Sahle, P. (2016b). What is an Scholarly Digital Edition? En M. J. Driscoll, y E. Pizarro, *Digital Scholarly Editions: Theories and Practices* (pp. 19-41). Open Book Publishers. <https://www.openbookpublishers.com/product/483>

Spence, P. (2014). Edición académica en la era digital: modelos, difusión y proceso de investigación. *Anuario Lope de Vega: Texto, Literatura, Cultura*, 20, 47–83. <https://doi.org/10.5565/rev/anuariolopedevega.74>

Razones y reflexiones para una meta-historia digital de las literaturas ibéricas usando TEI y Open Linked Data

Santiago Pérez Isasi

Centro de Estudos Comparatistas, Faculdade de Letras, Universidade de Lisboa

Palabras clave: Humanidades Digitales, Historia literaria, Estudios ibéricos, TEI, Linked Open Data

El estudio y la producción de la historia literaria (disciplina tantas veces declarada “muerta”) sigue dependiendo, en gran medida, de presupuestos de origen romántico, y de modelos narrativos cronológicos, promovidos y fortalecidos por su relación con los sistemas educativos y editoriales nacionales. Ejemplo de estas dinámicas podrían ser las (por otro lado, de indudable valor y rigor) *Historia de la literatura española* en nueve volúmenes, publicado por la editorial Crítica, o la *História Crítica da Literatura Portuguesa*, también en nueve tomos, dirigida por Carlos Reis. Cabe mencionar también, por su ordenación genérica y no cronológica, la *Egungo Euskal Literaturaren Historia* (“Historia de la literatura vasca contemporánea”), coordinada por Jon Kortazar, en ocho volúmenes.

Alternativa o complementariamente a estos proyectos, que demuestran la vitalidad de la disciplina, han surgido, reciente y no tan recientemente, propuestas teóricas y metodológicas de renovación del campo de la historia literaria, ya sea a través del cuestionamiento de la narración, o del sujeto colectivo nacional, o bien al propio concepto de “literatura”. Surgen así iniciativas como la *Comparative history of literatures in the Iberian Peninsula* (Cabo et al, Domínguez et al) (por mencionar solo un ejemplo significativo), en el que se abandona el modelo cronológico/narrativo, se amplía el objeto de estudio más allá de los límites tradicionales de lo literario, y el concepto geoespacial de “Península Ibérica” viene a sustituir al sujeto nacional como base de estudio. Lo mismo sucede con el campo de los Estudios Ibéricos, que en los últimos años han venido a cuestionar los modelos académicos y disciplinares nacionales, para proponer reconsideración alternativa, relacional y dinámica, de las literaturas de la Península Ibérica (Resina, Pérez Isasi)

Este es el contexto teórico y disciplinar del proyecto que presento en esta comunicación, y que está en una fase inicial. Su objetivo fundamental consiste en la creación de un archivo de historias de las diversas literaturas ibéricas, compuestas y publicadas entre 1800 y 1930, editarlas y etiquetarlas empleando el estándar TEI, enriquecerlas con la introducción de *Open Linked Data* referentes a autores y obras, y ofrecerlas a los usuarios a través de una interfaz dinámica que permita la lectura y descarga de textos individuales, pero también búsquedas transversales al corpus por texto libre, por autor, obra, periodo o lengua.

En esta comunicación presentaré algunas reflexiones iniciales sobre cada uno de los pasos implicados en la consecución de este objetivo:

- Selección de las obras que compondrán el corpus, representativas no solo de todas las literaturas ibéricas, sino también de diversos periodos dentro del ámbito cronológico estudiado
- Localización de ejemplares digitalizados de las obras escogidas, o digitalización de aquellas que no existan; conversión en texto, revisión y corrección del resultado

- Etiquetado mediante estándar TEI y enriquecimiento con datos enlazados abiertos • Diseño y desarrollo de la interfaz de búsqueda y acceso a los textos

Se trata de un proyecto que aplica metodologías ya ampliamente testadas, tanto en un contexto internacional como en la Península Ibérica, en proyectos como la biblioteca digital Mnemosine del proyecto *La Otra Edad de Plata*, *BIDISO* o *Galiciana*, o el proyecto Andrés de Poza, dirigido por Carmen Isasi, en el que yo mismo participé. Con todo, estas metodologías se aplican en este caso a un corpus o tipo de texto poco explorado hasta ahora, como es el de las historias literarias, a pesar del evidente potencial de un proyecto de este tipo para la investigación, la divulgación o la docencia. Por otra parte, la introducción de *Linked Open Data*, por ahora poco explorada en el ámbito de la edición digital hispánica, puede ofrecer vías futuras de desarrollo para este y otros proyectos semejantes.

Además de estas posibles vías de reflexión y desarrollo metodológico, este proyecto puede contribuir para otros resultados o avances, contribuyendo, por ejemplo, para un mayor conocimiento de las interrelaciones entre literaturas ibéricas; disponibilizando un conjunto de textos hasta ahora poco accesibles o desconocidos; proporcionando herramientas didácticas para la aproximación al pasado literario; estableciendo vínculos e intercambios con otros proyectos semejantes; o contribuyendo para establecer diálogos entre los desarrollos en Humanidades Digitales en España y en Portugal.

Bibliografía

Andrés de Poza. URL <https://andresdepoza.com/>

BIDISO – Biblioteca Digital Siglo de Oro. URL www.bidiso.es/

Cabo, Fernando; Domínguez, César, y Abuín, Anxo (coords.) (2010). *A Comparative history of Literatures in the Iberian Peninsula*. John Benjamins, vol 1.

Domínguez, César; Abuín, Anxo y Sapega, Ellen (coords.) (2016). *A Comparative history of Literatures in the Iberian Peninsula*. John Benjamins, 2016, vol 2.

Galiciana. Biblioteca Dixital de Galicia. URL <https://biblioteca.galiciana.gal/>

La Otra Edad de Plata. URL <https://www.ucm.es/loep/>

Kortazar, Jon (coord.) (2007-2017): *Egungo Euskal literaturaren historia*. VitoriaGasteiz: UPV-EHU, 8 vols.

Mainer, Carlos (coord.) (2010-2013): *Historia de la literatura española*. Madrid: Crítica, 9 vols.

Pérez Isasi, Santiago (2020): “¿Hacia unos Estudios Ibéricos 2.0? Críticas, debates y caminos abiertos”, *Theory Now. Journal of Literature, Critique, and Thought*, 3(2), pp. 145-167. doi: [10.30827/tnj.v3i2.15542](https://doi.org/10.30827/tnj.v3i2.15542).

Reis, Carlos (coord.) (1992-2015): *História crítica da literatura portuguesa*. Lisboa: Verbo.

Resina, Joan Ramon (2009): *Del Hispanismo a los Estudios Ibéricos*. Madrid: Biblioteca Nueva.

ENTRIB Entremeses Ibéricos: hacia una edición sinóptica digital

José Pedro Sousa

Centro de Estudos de Teatro, Faculdade de Letras da Universidade de Lisboa

Uno de los objetivos del proyecto del Centro de Estudios de Teatro ENTRIB - entremeses ibéricos: inventario, edición y estudio (PTDC/LLT-LES/32366/2017) es publicar una edición académica digital de los entremeses ibéricos escritos en más de un idioma.

Tras realizar el inventario de los entremeses ibéricos, se seleccionaron los que son comunes a los distintos idiomas para su edición electrónica. La edición incluye (i) etiquetas (personaje; parlamento; verso, etc.), (ii) hiperlinks entre la transcripción textual y las imágenes digitales de los manuscritos o primeras ediciones utilizadas; y (iii) una comparación entre las distintas versiones de un mismo texto, identificando variantes, a las que se puede prestar especial atención para comprender los cambios y diferencias entre los dos o más textos.

En una edición crítica electrónica el investigador debe ser competente informáticamente. El conocimiento del “editor digital”, para utilizar la expresión de Gary Taylor (2000: 45), tiene que sobrepasar el universo del libro y el de la edición electrónica, ya que su trabajo se encuentra en la intersección de los dos al aplicar un medio distinto (el digital) al del objeto de estudio (el libro), abierto a un amplísimo abanico de posibilidades por descubrir. El modelo de organización es totalmente diferente y los modos de lectura son múltiples, lo cual incrementa, a la vez, la libertad y la responsabilidad del lector.

Hay que reconocer que en este tipo de literatura teatral las rescritas de los mismos temas en los diferentes idiomas pueden asumir diferentes formas, desde meras traducciones, hasta versiones, adaptaciones o reelaboraciones en las que el tema se puede reconocer como el mismo en los distintos testimonios, lo que permite presumir un arquetipo común, pero con diferencias textuales sustanciales. En estos casos, la edición sinóptica se vuelve más difícil.

¿Cómo demostrar los distintos tipos de diferencias textuales entre los textos? Tratándose de lenguas tan semejantes, ¿dónde están los límites para considerar la diferencia entre los textos? Y, sobre todo, ¿cómo mostrarlo en la pantalla al lector digital de las ediciones ENTRIB?

Bibliografía

CAMÕES, JOSÉ y SOUSA, JOSÉ PEDRO (2018), “Arte nuevo de editar comedias en este tiempo digital” Cuadernos AISPI, pp. 15-30.

MCGANN, JERRY (1997), “The rationale of hypertext”, *Electronic Text, Investigations in Methods and Theory*, ed. Kathryn Sutherland. Oxford: 19-46.

TAYLOR, GARY (2000), “c:\wp\file.txt 05:41 10-07-98”, *The Renaissance text*, ed. Andrew Murphy. Manchester/New York, Manchester University Press: 44-54.

Panel 3

Humanidades digitales y teatro

Modera: Javier J. González Martínez · *Asiste:* Brianda Otero Moreira

Panelistas: Álvaro Cuéllar; Claudia Demattè; Emma M^a Marcos Rodríguez; Hugo Sanjurjo-González; Olaia Andaluz-Pinedo

Humanidades digitales y teatro

Nombre y filiación del coordinador:

Javier J. González Martínez (Universidad Internacional de La Rioja)

Nombres y filiaciones de los participantes:

Olaia Andaluz-Pinedo (Universidad del País Vasco UPV/EHU) y Hugo Sanjurjo-González (Universidad de Deusto)

Álvaro Cuéllar (University of Kentucky)

Claudia Demattè (Università di Trento)

Emma M^a Marcos Rodríguez (Universidad de Valladolid)

Palabras clave: Corpus paralelo, estilística, estilometría, software, teatro.

Descripción del problema de investigación, metodología y resultados obtenidos:

Los estudios teatrales han encontrado en la tecnología el apoyo e impulso que requerían los métodos tradicionales. Y también es cierto que las humanidades digitales han hallado en el teatro un campo idóneo para la investigación desde muy diversas perspectivas: estudios textuales, localizaciones georreferenciales, creación de redes sociales de personajes y de escritores, recreaciones virtuales, etc. En este panel nos centraremos en los corpus y la estilometría para resolver cuestiones de traducciones, géneros dramáticos, dataciones y atribución de autorías.

El primer problema que será atendido es el de la escasez de corpus de textos dramáticos. Es especialmente acuciante en el caso de los corpus paralelos de obras teatrales y sus traducciones. Por medio de la construcción del corpus TEATRAD (Teatro traducido) y del desarrollo de un software específico (TAligner y ACM-theatre) para procesar este tipo de corpus, se ofrecen estadísticas, diferentes opciones de visualización, anotación, alineación y obtención de concordancias. Los resultados se concretan en un corpus y unas herramientas que facilitan el trabajo de análisis a todos los interesados en este tipo de textos.

Por otra parte, se analizan los usos de la estilometría a la hora de afrontar cuatro tipos de problemas: distinción de refundiciones, identificación de géneros literarios, datación y atribución de autorías. La combinación de la metodología filológica tradicional y el procesamiento informático de la frecuencia del uso léxico está abriendo nuevos horizontes en los estudios teatrales.

El sufrimiento premiado se encontraba entre las obras que Lope identificó como suyas, pero apareció en una colección de comedias de Juan Pérez de Montalbán. A través del estudio métricoestrófico, la ortología, la estilística y la estilometría se resuelve la cuestión de la autoría de esta pieza.

La producción dramática de Antonio Mira de Amescua es una de las más problemáticas en cuanto a autoría dentro del teatro del Siglo de Oro. La aplicación de ETSO (Estilometría Aplicada al Teatro del Siglo de Oro) ha permitido contrastar el corpus de este dramaturgo con más de dos mil obras de teatro del siglo XVII español. El resultado de este análisis estilométrico es la distinción de tres grupos dentro de las obras hasta ahora atribuidas a Mira de Amescua: las suyas, las dudosas y las de otros dramaturgos.

Descripción de las intervenciones de cada participante:

Humanidades digitales y teatro⁴⁶

Javier J. González Martínez (Universidad Internacional de La Rioja)

El carácter transdisciplinar del teatro y las humanidades digitales ha propiciado un fructuoso encuentro para ambos. En el campo de las humanidades digitales, la transdisciplinariedad permite a los investigadores trabajar en equipos que comparten información y técnicas de dos o más campos de especialización. Es un paso más allá de la multidisciplinariedad y la interdisciplinariedad, La multidisciplinariedad conlleva trabajar simultáneamente, pero en paralelo, a un grupo de investigadores de diferentes disciplinas. Sin embargo, la transdisciplinariedad aúna principios y métodos de distintas disciplinas y posibilita a los investigadores ir más allá de las disciplinas tradicionales.

En el teatro el enfoque transdisciplinar se da por su propia constitución de suma de artes: danza, música, interpretación actoral, literatura, pintura, escultura, vídeo, arquitectura, etc. Su estudio va a requerir también equipos de trabajo en los que participen de forma conjunta investigadores que ofrezcan diversidad de perspectivas. Además, el carácter efímero del teatro exige la aplicación de metodologías que permitan la grabación de las huellas (datos) dispersas por sus variados caminos. Esos rastros, que son manifestaciones de una presencia, están dispersos en cartelerías, ediciones, licencias de representación, etc., que podemos ir extrayendo de los más variados archivos y bibliotecas.

Las humanidades digitales aplicadas al teatro están desplegando un rico panorama de enfoques alternativos. Las visualizaciones de las redes sociales de escritores que escribían en colaboración nos están definiendo el contexto de las comunidades creativas de cada momento histórico. Los análisis de redes sociales de personajes dramáticos ayudan a definir el desarrollo dramático de cada trama y los usos de los escritores para desarrollar sus acciones dramáticas. La estilometría está aportando nuevos datos para la atribución de autorías literarias en un arte en el que los textos escritos han perdido protagonismo en favor de la escenificación.

La especificidad en la construcción y el análisis de un corpus paralelo de teatro: el caso de TEATRAD⁴⁷

Olaia Andaluz-Pinedo (Universidad del País Vasco UPV/EHU)

Hugo Sanjurjo-González (Universidad de Deusto) El marco multidisciplinar de las Humanidades Digitales integra una amplia gama de objetos de estudio y perspectivas. En los análisis basados en textos, los corpus destacan por su relevancia desde

diferentes disciplinas (Hoover, Culpeper y O'Halloran 2014; Doval Reixa and Sánchez Nieto 2019). No obstante, los corpus de teatro aún son escasos, frente al amplio número de corpus compuestos de otros géneros, como la narrativa.

Recientemente se han desarrollado casos de interés monolingües (como BETTE⁴⁸ o ESC⁴⁹); sin embargo, no tenemos

⁴⁶ Esta investigación es resultado de los proyectos “CLEMIT. Base de datos integrada del teatro clásico español. Segunda fase” (PID2019-104045GB-C52) e “ISTAE. Base de datos integrada del teatro clásico español” (PID2019-104045GA-C55), ambos financiados por el Plan Estatal de Investigación del Ministerio de Ciencia, Innovación y Universidades.

⁴⁷ Este trabajo se ha realizado en el marco de la Red de Excelencia CorpusNet (MINECO, FF12016-81934RED/AEI), el grupo de investigación consolidado TRALIMA/ITZULIK (Gobierno Vasco, IT1209/19; UPV/EHU GIU19/067) y la convocatoria de la UPV/EHU para la formación de personal investigador (PIF17/46).

⁴⁸ <https://github.com/GHEDI/BETTE>

⁴⁹ <http://corpora.lancs.ac.uk/esc-user-service/>

constancia de corpus paralelos de teatro (Andaluz-Pinedo, Sanjurjo-González y Merino-Álvarez en prensa). Este tipo de corpus, formado por la alineación de obras teatrales originales y sus traducciones (Xiao y Yue 2009: 240-241), permite estudiar el teatro traducido, un ámbito que apenas recibe atención en comparación con el teatro escrito originalmente en castellano, pese a constituir una parte fundamental de la cultura teatral en España (Merino-Álvarez 2016: 304). La ausencia de corpus paralelos de teatro podría deberse a la falta de software ajustado para procesar este tipo de corpus, ya que las herramientas condicionan en gran medida las posibilidades para la creación y el estudio de los corpus (Anthony 2013: 141-151).

Esta contribución se centra en las peculiaridades de la construcción y el análisis de un corpus paralelo de teatro denominado TEATRAD (TEAtrO TRADucido), compuesto por obras de teatro en lengua inglesa traducidas al español con una amplia difusión en el siglo XXI. Este corpus se ha utilizado para el desarrollo y testeo de herramientas ajustadas a la especificidad estructural de los textos teatrales: TAligner⁵⁰ y ACM-theatre (Sanjurjo-González 2017a, 2017b, 2018). Así, en esta ponencia se apuntan las necesidades encontradas respecto a anotación, alineación y la obtención de concordancias y estadísticas, junto con las características del corpus y de las herramientas. Cabe mencionar que, aunque se ponga el foco en un corpus paralelo, los aspectos señalados también pueden ayudar a dar respuesta a otro tipo de necesidades, ya que los textos que lo forman pueden consultarse de forma individual y los avances en las herramientas son aplicables asimismo para corpus monolingües.

En primer lugar, el proceso de construcción del corpus TEATRAD sirvió para pulir la adaptación a las particularidades de los corpus paralelos de teatro de la herramienta TAligner, desarrollada a partir de la aplicación TRACE Corpus Tagger/Aligner 1.0 (Gutiérrez-Lanza, Bandín, García-González y LobejónSantos 2015). La estructura singular de los textos teatrales requiere opciones específicas para la construcción de este tipo de corpus: la anotación estructural y la alineación a nivel de réplica (AndaluzPinedo y Sanjurjo González en prensa). Por un lado, en los corpus de teatro, ya sean paralelos o monolingües, es necesaria una anotación estructural que diferencie las unidades en las que se organizan las obras teatrales: unidades globales como actos o escenas y, fundamentalmente, réplicas, personajes, diálogos y acotaciones. De hecho, incluso existe un estándar TEI para la anotación de este tipo de textos (TEI Consortium 2019: 244-247). Por otra parte, una característica específica de los corpus paralelos de teatro consiste en la alineación a nivel de réplica, en lugar de a nivel de oración o párrafo como en los textos escritos en prosa (Merino 2007: 261; Bandín 2007: 29). El corpus TEATRAD también se ha empleado en el testeo de la adaptación para corpus teatrales de otra herramienta, ACM-theatre. ACMtheatre permite la anotación lingüística (gramatical y semántica), además de la anotación estructural de los textos y su alineación a nivel de réplica, en este caso automática. Dado que ofrece opciones complementarias a las de TAligner para la creación de corpus paralelos de teatro, se está construyendo una segunda versión de TEATRAD con las opciones de anotación lingüística disponibles en este software.

Asimismo, en el análisis de un corpus paralelo de teatro como TEATRAD es conveniente disponer de concordancias, estadísticas y visualizaciones basadas en la anotación estructural. Por un lado, TAligner se ha empleado para filtrar concordancias de acuerdo con la anotación de los diálogos y las acotaciones. Además, el desarrollo reciente de ACM-theatre amplía de manera significativa las opciones de análisis que utilizan la anotación estructural: se ofrecen concordancias, así como estadísticas y visualizaciones, en las que también es posible utilizar la anotación gramatical y semántica. El desarrollo

⁵⁰ <http://www.ehu.es/tralima/taligner.html>

de visualizaciones que interrelacionan la estructura del teatro con otros datos es un trabajo en curso que, no obstante, parece prometedor por las posibilidades que abre para el análisis de corpus paralelos, así como monolingües (Sanjurjo-González y Andaluz-Pinedo 2021).

En definitiva, a partir del proceso llevado a cabo para la construcción y explotación del corpus paralelo TEATRAD, en esta intervención se ilustran las particularidades de un tipo de corpus poco extendido y se refleja la interrelación entre este objeto de estudio digital y las herramientas software respecto a las posibilidades para su creación y análisis. Gracias a la colaboración interdisciplinar de investigadores en los campos de la traducción y la lingüística computacional, se han desarrollado herramientas que estarán disponibles para facilitar futuros trabajos de otros miembros de la comunidad científica.

Estilometría más allá de la atribución de autoría

Álvaro Cuéllar (University of Kentucky)

En los últimos años la estilometría se ha demostrado como un proceso informático muy útil para la atribución de autoría en el teatro del Siglo de Oro. La posibilidad de relacionar obras teatrales en función de la frecuencia de uso léxico, por ejemplo, parece conducirnos en muchas ocasiones hacia la ratificación o rectificación de propuestas de paternidad autorial. Más allá de este propósito, la estilometría puede también servir al estudio de multitud de aspectos distintos, que han sido menos trabajados desde las Humanidades Digitales aplicadas al teatro del Siglo de Oro. Esta presentación trata de explorar algunos de estos caminos de aplicación de la técnica informática no relacionados con la atribución de autoría. La estilometría puede servirnos, por ejemplo, para la detección automática de refundiciones, práctica tan habitual en ese tiempo, si la aplicamos a un amplio conjunto de textos sobre los que sería imposible trabajar de forma analógica. También parece prometedor el empleo de la estilometría para el estudio de los géneros literarios, categorías dúctiles que siempre pueden repensarse y reordenarse. Comedias, autos sacramentales, entremeses, etc., la estilometría puede ayudarnos a saber hasta qué punto estos subgrupos dentro del teatro del Siglo de Oro configuran un conjunto estilístico diferenciado del resto o si más bien esta distinción es una construcción arbitraria. Otra de las aplicaciones, todavía muy poco estudiada, es la datación de las obras teatrales. Si comprobamos que las obras de un autor determinado tienden a relacionarse estilométricamente por su momento de escritura, podremos inferir la datación de un texto de dudosa fecha de composición. Estas y otras vías de exploración serán presentadas con el objetivo de proporcionar al estudioso del Siglo de Oro, o de la literatura en general, un repaso a las posibilidades de análisis que ofrece esta técnica y que puede ser de utilidad para sus investigaciones.

Estilometría para la atribución de comedias del Siglo de Oro: el caso de *El sufrimiento premiado*

Claudia Demattè (Università di Trento)

En la primera edición de *El Peregrino en su Patria*, Lope de Vega incluye *El sufrimiento premiado* en la lista de sus comedias. Muchos años más tarde, en 1638, y sin casi ninguna otra noticia al respecto en este lapso temporal, una obra con el mismo título se publica en el *Segundo tomo de comedias* de Juan Pérez de Montalbán. El querido discípulo de Lope acababa de morir y su padre, el librero Alonso Pérez, se encargó de preparar el volumen con doce comedias, aunque posiblemente incluyendo algunas comedias que no eran de Montalbán. Entre estas, *El sufrimiento premiado* es la que ha

recibido alguna atención por haberse atribuido al Fénix gracias a los estudios de Dixon en los años Sesenta. La atribución de esta pieza será el objeto de nuestra intervención: a través del estudio métrico-estrófico, de la ortología, de la estilística y finalmente de la estilometría, intentaremos matizar los elementos que pueden llevar a apoyar la posible autoría de Lope de Vega.

Antonio Mira de Amescua en ETSO⁵¹

Emma M^a Marcos Rodríguez (Universidad de Valladolid) Antonio Mira de Amescua es uno de los dramaturgos más afectados por los problemas de atribución que envuelven a parte de la producción dramática de nuestro Siglo de Oro. Quizá, el motivo principal sea su retirada de la vida literaria para ocupar en Guadix, su tierra natal, el arcedianato de la catedral, pues supuso que jamás publicara las *partes* de sus comedias y que sus obras corrieran libres, representándose y publicándose a nombre de otros poetas dramáticos de su tiempo, con las variaciones textuales que ello conllevaría.

Dadas las circunstancias, los investigadores de los dos últimos siglos han procurado llevar a cabo una recuperación del teatro miramescuano lo más minuciosa posible y, sin embargo, los trabajos de catalogación más importantes realizados hasta la fecha difieren llamativamente en el número de títulos; así, pues, Ramón de Mesonero Romanos, en 1852, fue el primero en reunir en un minucioso catálogo treinta y nueve comedias y trece autos, muy pocas obras en comparación con las que recogió Cotarelo en 1931, las ochenta y cuatro recopiladas por Valladares en su *Bibliografía* del dramaturgo y las ochenta y dos editadas en el *Teatro completo* de Mira, coordinado por Agustín de la Granja y publicado por la Universidad de Granada entre el 2001 y el 2013. Asumimos que el último de estos trabajos, fruto de años de investigación por parte de un amplio grupo de estudiosos, reúne las piezas que con mayores motivos podemos pensar de Mira, pero basta una lectura de los trabajos introductorios a las ediciones para darnos cuenta de que en muchos casos carecemos de datos suficientes que nos permitan inclinar la balanza hacia la autoría del guadijeño.

El avance de la tecnología y, con ella, de las humanidades digitales, nos ha permitido revisar, desde una perspectiva innovadora, el trabajo de nuestros predecesores. La estilometría es el método digital más empleado por los investigadores actuales para determinar la autoría de los textos, puesto que se centra en la comparación del uso de las palabras más frecuentes en ellos para identificar a su autor dentro de un conjunto cerrado de posibilidades. El proyecto de Álvaro Cuéllar y Germán Vega GarcíaLuengos, ETSO (Estilometría Aplicada al Teatro del Siglo de Oro), cuenta en estos momentos con un corpus que ronda los dos mil textos teatrales, permitiéndonos obtener resultados bastante fiables de cara a una nueva investigación, ya que no podemos considerar definitivos los datos que nos ofrece, pues los textos, y sobre todo los destinados a los escenarios del siglo XVII, pudieron sufrir numerosos cambios respecto del original, e incluso podrían ser refundiciones o colaboraciones entre varios poetas. Todos estos problemas están reflejados en las piezas atribuidas a Mira de Amescua que se han analizado en ETSO, pues ha surgido todo lo contrario a lo que se esperaba, han aumentado todavía más las dudas que envuelven su corpus: un total de treinta y una obras de las publicadas en el *Teatro completo* ha dado problemas, siendo cuatro las que se acercan peligrosamente a otros dramaturgos y veintisiete las que no tienen una atribución clara. Asimismo, se ha demostrado la paternidad miramescuana de cuatro que andaban dudosas. Por otro lado, en contra de lo que sucede con algunas de las piezas atribuidas durante años a Mira de Amescua, que podrían pertenecer a la pluma de otros autores, solo tenemos el caso de una que se creía de Lope y podría ser de nuestro poeta.

⁵¹ Investigación cofinanciada por parte de la Consejería de Educación de la Junta de Castilla y León y el Fondo Social Europeo.

Por los motivos expuestos, dividiré mi intervención en cuatro apartados siguiendo los resultados del análisis estilométrico: las obras que se creían de Mira de Amescua y que podemos confirmar como suyas, las que podrían pertenecer a otros dramaturgos, aquella atribuida a Lope de Vega y que se acerca a Mira y, por último, las dudosas, que conforman el bloque más problemático. Por falta de tiempo, no podré detenerme en todas ellas, por lo que daré únicamente un apunte de cada una y me centraré brevemente en los casos más llamativos, sobre todo de los correspondientes al apartado de “dudosas”, que me servirán para reflejar la importancia de una herramienta como estilometría, para exponer los problemas con los que nos podemos encontrar a la hora de analizar textos antiguos y, además, pertenecientes al mundo escénico y, por supuesto, para visibilizar las nuevas investigaciones que se están llevando a cabo en el ámbito filológico-teatral con el fin de revisar e intentar completar los estudios que previamente realizaron nuestros maestros sin ningún tipo de medio digital a su alcance. Desde mi punto de vista, si aunamos los diferentes trabajos de investigación que se están realizando apoyados por las humanidades digitales y aquellos realizados de la manera más tradicional, podremos obtener mejores resultados, y esto es lo que, en mi caso, pretendo con el análisis autorial de las piezas atribuidas a Mira de Amescua: partir de los resultados de estilometría y de las investigaciones tradicionales previas para, así, llegar a nuevas hipótesis y conclusiones.

Bibliografía

- Andaluz-Pinedo, Olaia y Sanjurjo-González, Hugo. En prensa. “Corpus tools for parallel corpora of theatre plays: an introduction to TAligner and ACM-theatre”.
- Andaluz-Pinedo, Olaia, Sanjurjo-González, Hugo y Merino-Álvarez, Raquel. En prensa. Recursos de Humanidades Digitales para el estudio del teatro (traducido): bases de datos, corpus y herramientas desarrollados en TRALIMA/ITZULIK. *Talia. Revista de Estudios Teatrales*, 3.
- Anthony, Laurence. 2013. A critical look at software tools in corpus linguistics. *Linguistic Research*, 2: 141-161. <https://doi.org/10.17250/khisli.30.2.201308.001>
- Bandín, Elena. 2007. *Traducción, recepción y censura de teatro clásico inglés en la España de Franco. Estudio descriptivo-comparativo del Corpus TRACETci (1939-1985)*, Tesis. León: Universidad de León. <https://buleria.unileon.es/handle/10612/1885>
- Barrera y Leirado, Cayetano Alberto de la. 1860. *Catálogo bibliográfico y biográfico del teatro antiguo español: desde sus orígenes hasta mediados del siglo XVII*. Madrid: Imprenta y Estereotipia de M. Rivadeneyra.
- Calvo Tello, José. 2016. Entendiendo Delta desde las humanidades. *Caracteres: estudios culturales y críticos de la esfera digital*, vol. 5, no. 1: 140-76.
- Calvo Tello, José y Juan Cerezo Soler. 2018. *La conquista de Jerusalén ¿de Cervantes? Análisis estilométrico sobre autoría en el Teatro del Siglo de Oro español*. *Digital Humanities Quarterly*, vol. 12, 1: 1-10.
- Calvo Tello, José. 2019. Delta Inside Valle-Inclán: Stylometric Classification of Periods and Groups of His Novels. *Romanische Studien*: 151-64. <http://www.romanischestudien.de/index.php/rst/article/view/625>.
- Calvo Tello, José. 2019. Diseño de corpus literario para análisis cuantitativos. *Revista de Humanidades Digitales*, vol. 4 : 115-35, doi:10.5944/rhd.vol.4.2019.25187.
- Cerezo Soler, Juan y José Calvo Tello. 2019. Autoría y estilo. Una atribución cervantina desde las Humanidades Digitales. El caso de *La conquista de Jerusalén*. *Anales Cervantinos*, vol. 51: 231-50, <https://doi.org/10.3989/anacervantinos.2019.011>.
- Cotarelo y Mori, Emilio. 1931. *Mira de Amescua y su teatro. Estudio bibliográfico y crítico*. Madrid: Tipografía de la Revista de Archivos.
- Cuéllar, Álvaro y Germán Vega García-Luengos. *ETSO. Estilometría aplicada al Teatro del Siglo de Oro, 2017-2021*, Recurso web <etso.es>.
- Cuéllar, Álvaro. En Prensa. “Stylometry and Spanish Golden Age Theatre: An Evaluation of Authorship Attribution in a Control Group of Undisputed Plays”. Heidelberg University Press.
- Demattè, Claudia. 2019. ¿Una nueva comedia en colaboración entre ¿Calderón?, Rojas Zorrilla y Montalbán? *Empezar a ser amigos* a la luz del análisis estilométrico. *Rilce*, 35-3: 852-874. <<https://www.unav.edu/publicaciones/revistas/index.php/rilce/issue/view/633>> DOI: <https://doi.org/10.15581/008.35.3.852-74>
- Demattè, Claudia. En prensa. El *Segundo tomo de comedias* de Juan Pérez de Montalbán: hacia un mapa estilométrico de las comedias auténticas, dudosas y ajenas. *Talia. Revista de estudios teatrales*, 3.

- Di Pinto, Elena. 2011. "El arte de la refundición según Moreto (I): *El despertar a quien duerme* de Lope de Vega vs. *La misma conciencia acusa de Moreto*". En *Actas del VIII Congreso de la AISO*, vol. 6, no. 1: 107-14.
- Doval-Reixa, Irene, y M. Teresa Sánchez-Nieto (eds.). 2016. *Parallel corpora for contrastive and translation studies: New resources and applications*. Amsterdam/Philadelphia: John Benjamins.
- Eder, Maciej. 2010. Does Size Matter? Authorship Attribution, Small Samples, Big Problem. *Proceedings of Digital Humanities*: 132-35. Doi:10.1093/llc/fqt066.
- Eder, Maciej, y Jan Rybicki. 2013. Do birds of a feather really flock together, or how to choose training samples for authorship attribution. *Literary and Linguistic Computing*, vol. 28, no. 2: 229-36. Doi:10.1093/llc/fqs036.
- Eder, Maciej. 2013. Mind your corpus: Systematic errors in authorship attribution. *Literary and Linguistic Computing*, vol. 28, no. 4: 603-14. Doi:10.1093/llc/fqt039.
- Eder, Maciej. 2016. Rolling stylometry. *Digital Scholarship in the Humanities*, vol. 31, no. 3: 457-69. Doi:10.1093/llc/fqv010.
- Eder, Maciej, et al. 2016. Stylometry with R: a package for computational text analysis. *R Journal*, vol. 7, 1: 107-121.
- Eder, Maciej. 2017. Visualization in Stylometry: Cluster Analysis Using Networks. *Digital Scholarship in the Humanities*, vol. 32 nº 1: 50-64. <https://doi.org/10.1093/llc/fqv061>.
- Gutiérrez-Lanza, Camino, Elena Bandín-Fuertes, José Enrique García-González y Sergio Lobejón-Santos. 2015. "Desarrollo de software de etiquetado y alineación textual: TRACE Corpus Tagger/Aligner 1.0@". En *II Congreso Internacional de Humanidades Digitales Hispánicas: Innovación, globalización e impacto*. Madrid: UNED. <https://humanidadesdigitaleshispanicas.es/wp-content/uploads/2020/02/Humanidades-Digitales-Hisp%C3%A1nicas-https://humanidadesdigitaleshispanicas.es/wp-content/uploads/2020/02/Humanidades-Digitales-Hisp%C3%A1nicas-Innovaci%C3%B3n-Globalizaci%C3%B3n-e-Impacto.pdf>
- Hernández Lorenzo, Laura. 2019. Poesía áurea, estilometría y fiabilidad: métodos supervisados de atribución de autoría atendiendo al tamaño de las muestras. *Caracteres. Estudios Culturales y Críticos de La Esfera Digital*, vol. 8, no. 1: 189228. <http://revistacaracteres.net/wp-content/uploads/2019/06/Caracteresvol8n1mayo2019-estilometria.pdf>.
- Hoover, David L., Culpeper, Jonathan y O'Halloran, Kieran (eds.). 2014. *Digital Literary Studies. Corpus Approaches to Poetry, Prose and Drama*, 9-33. Nueva York: Routledge. <https://doi.org/10.4324/9780203698914>
- Merino-Álvarez, Raquel. 2007. "La homosexualidad censurada: estudio sobre corpus de teatro TRACETi inglés-español (desde 1960)". En Raquel Merino-Álvarez (ed.), *Traducción y censura en España (1939-1985). Estudios sobre corpus TRACE: cine, narrativa, teatro*, 243-313. Bilbao: UPV/EHU.
- Merino-Álvarez, Raquel. 2016. Censura, traducción e integración en el teatro de la época franquista: José López Rubio, hombre de teatro y traductor. *Hispanic Research Journal. Iberian and Latin American Studies*, 4: 303-321. <https://doi.org/10.1080/14682737.2016.1200854>
- Mesonero Romanos, Ramón. 1852. El teatro de Mirademesca. *Semanario Pintoresco Español*, 82-83.
- R Core Team. 2020. *R: A Language and Environment for Statistical Computing*, Vienna, Austria, <https://www.r-project.org/>.
- Sanjurjo-González, Hugo. 2017a. *ACTRES Corpus Manager* [software].
- Sanjurjo-González, Hugo. 2017b. *Creación de un framework para el tratamiento de corpus lingüísticos / Development of a Framework for Corpus Linguistic Analysis*. Tesis. León: Universidad de León. <https://buleria.unileon.es/handle/10612/6920>
- Sanjurjo-González, Hugo. 2018. *Creación de un framework para el tratamiento de corpus lingüísticos*. León: Universidad de León.
- Sanjurjo-González, Hugo y Andaluz-Pinedo, Olaia. 2021. "Statistics and visualisations of theatre corpora using corpus analysis software". *Poster presentado en el 38º Congreso Internacional de la Asociación Española de Lingüística Aplicada*. A Coruña: Universidade da Coruña. DOI: 10.13140/RG.2.2.15071.79529
- Underwood, Ted. 2020. Machine Learning and Human Perspective. *PMLA*, vol. 135, no. 1: 92-109. doi:10.1632/pmla.2020.135.1.92.
- Underwood, Ted. 2019. *Distant Horizons: Digital Evidence and Literary Change*. Chicago: University of Chicago Press.
- Vega García-Luengos, Germán. 1998. La reescritura permanente del teatro español del Siglo de Oro: nuevas evidencias. *Críticón*, no. 72: 11-34.
- Vega y Carpio, Lope de (atrib.). 1967. *El sufrimiento premiado*, ed. de Victor Dixon. Londres: Thamesis Books.

Sesión 4

La vía de lo femenino

Estudios culturales I

Modera: Amelia Sanz Cabrerizo · *Asiste:* Brianda Otero Moreira

- 8 HILAME. De la persona al dato, del dato a la persona
Agurtzane Paz, Teresa Jular
- 12 Discursos científicos sobre Arqueología Prehistórica desde una perspectiva de género: Un estudio de caso de inicios del siglo XX
Raquel Licerias Garrido, Alba Comino, Patricia Murrieta-Flores
- 29 Redes y comunidades de escritoras españolas de la primera Edad Moderna: análisis de datos y visualización
María Dolores Martos Pérez

HILAME. De la persona al dato, del dato a la persona

Agurtzane Paz Moro

Universidad del País Vasco (UPV-EHU)

Teresa Jular Pérez-Alfaro

XLI design+thinking

Palabras clave: prosopografía, Edad Media, base de datos, historia, transferencia

1. Presentación

HILAME, acrónimo de Hidalgos, Labradoras, Mercaderes, es un proyecto de Humanidades Digitales que busca recopilar los datos biográficos de todas aquellas personas que aparecen reflejadas en las fuentes documentales de época medieval e inicios de la moderna con el objeto de reconstruir, en la medida de lo posible, sus transcurso vitales.

El objetivo final es reconstruir las redes sociales configuradas en el periodo y espacio escogido, inicialmente los territorios de la cornisa cantábrica, al final de la Edad Media.

La metodología principal de HILAME es el análisis heurístico de la documentación medieval y su procesamiento informático. La clave del proceso es el concepto de “factoide”, una forma de categorizar las informaciones que aparecen en los testimonios medievales referidas a un individuo, normalizándolas para su tratamiento masivo y para la correlación de coincidencias (Tinti, 2003 y Smythe, 2007).

El equipo de HILAME está formado fundamentalmente por investigadores de la UPV-EHU en el marco del [Grupo de Investigación Consolidado](#) “Sociedad, Poder y Cultura en el País Vasco (siglos XIV-XVIII)”, codirigido por José Ramón Díaz de Durana y Arsenio Dacosta, y del cual forma parte la investigadora y docente Agurtzane Paz, experta en el tratamiento prosopográfico de las fuentes medievales trabajadas.

A ellos se sumó un pequeño grupo de diseñadores e informáticos externos, incorporados desde el comienzo para, todos juntos, cocrear una cultura de proyecto innovadora, acorde a los objetivos y principios del mismo. Entre estos profesionales, Teresa Jular, responsable de la coordinación de las tareas técnicas y la comunicación digital.

2. Cronología

HILAME nació en 2016 como proyecto orientado a la Ciencia Abierta, con unos principios de transparencia y conocimiento compartidos comunicados a través de múltiples canales, publicaciones y eventos. Y un proceso de ideación y desarrollo que podría resumirse en:

Nombre, la identidad gráfica, la arquitectura de la web y algunas primeras aplicaciones para la difusión del proyecto.

Una sede web, www.hilame.info, con licencia *creative commons* by-nc-sa, que detalla el proyecto, presenta al equipo investigador y muestra el trabajo en proceso en varias bases de datos complementarias. Se han concebido tutoriales para facilitar su uso a distintos perfiles de interesados, se incluyen referencias webgráficas y bibliográficas, con algunos libros completos accesibles para su descarga, así como audiovisuales de divulgación científica para mejor entendimiento del contexto histórico en el que se desenvolvían las personas identificadas a partir de las fuentes históricas tratadas. Una infografía actualizada trimestralmente, "HILAME en cifras", relata la evolución del proyecto en una sola hoja.

Una base de datos prosopográfica y un importador de datos tabulados «SMH Bypass» desarrollados a medida en PHP y protegidos, para la alimentación de las bases de datos públicas.

Un perfil en Academia.edu, que recoge artículos y eventos científicos relacionados con el proyecto.

Un perfil en [Twitter](https://twitter.com), desde el que se comunican avances en tiempo real y se dialoga sobre #HumanidadesDigitales en general y sobre #Historia en la #EdadMedia de manera específica.

Un canal en [YouTube](https://www.youtube.com), creado a finales de 2016 para hacer públicas varias entrevistas sobre el concepto de hidalguía y las guerras banderizas, y que luego se ha ido enriqueciendo con ponencias y talleres, disponibles para personas interesadas en la evolución del proyecto.

En 2017 tuvo lugar la presentación oficial de HILAME, en el Coloquio Internacional "Hidalgos e Hidalguía en la Península Ibérica al final de la Edad Media" celebrado en la Casa de Velázquez (Madrid). De allí partió el libro colectivo "Hidalgos e Hidalguía. Siglos XII-XV", publicado en Marcial Pons Historia con el artículo "HILAME en la escalera del diseño" (Dacosta y Jular, 2018).

La jornada "Terra Incognita?" organizada a finales de ese mismo año por la UPV-EHU en Vitoria para las personas especialistas inscritas, fue grabada en su totalidad, iniciando ya entonces lo que ahora se ha generalizado para todos los seminarios y congresos.

En 2018, el diseño gráfico de HILAME fue Selección en los Premios Internacionales CLAP y Finalista en la Bienal Iberoamericana de Diseño, en la categoría Investigación y Diseño.

Ese mismo año De Gruyter publicó un libro titulado "Humanidades Digitales. Miradas hacia la Edad Media" y editado por Déborah González y Elena Bermúdez Sabel, responsables del congreso homónimo celebrado en Santiago de Compostela, donde se publicó el artículo "HILAME: procesamiento y visualización de datos prosopográficos" (Jular y Dacosta, 2018).

Desde sus inicios, el proyecto aspira a la convergencia e interoperabilidad con otros proyectos (Charle, 1985; Millet, 1987; Boute, 2002), logro conseguido en 2020 con *Scripta manent*, un proyecto de investigación del Plan Nacional de I+D (HAR2016-77423-R), coordinado por Cristina Jular Pérez-Alfaro (IH-CCHS, CSIC). Esto ha permitido incorporar nuevas fuentes documentales y varios cientos de nuevos individuos a la base de datos principal, además de sumar una base de datos secundaria a las existentes: "la quinta bola" (Dacosta y Jular, 2020).

En el momento actual, HILAME ha incorporado al proyecto cinco estudiantes en prácticas, acogidos con una consulta preliminar y unas jornadas formativas online, disponibles en el canal de YouTube de HILAME.

3. La dimensión pública de HILAME

Durante los últimos meses de 2020 y primeros de 2021 hemos estado trabajando en mejoras para la visibilidad y la proyección pública (Baraibar, 2014), aunque sin olvidar el principal propósito del proyecto: facilitar la producción científica.

Esta labor ha venido precedida de un testeo del sitio web público desde distintos puntos de vista, con la intención de contar con la mirada de distintas audiencias, seguido de la implantación de mejoras destinadas a facilitar la consulta de sus contenidos, tanto en las bases de datos, principal y secundarias, donde hemos mejorado la precisión de las consultas y la organización de los resultados, como en las novedades relacionadas con la actividad del equipo.

Personas

Lugares

Factoides

Crono

"constanza lopez de ayala"

Buscar

CONSTANZA LÓPEZ DE AYALA

Mapa Satélite

Notas biográficas: Monja en Quejana

Lugares relacionados: Quejana, San Juan de Quejana, Orduña

Factoides asociados (10)

Participación en reunión de cabildo parroquial - monástico - catedralicio	Hermana	
Participación en reunión de cabildo parroquial - monástico - catedralicio		
Participación en reunión de cabildo parroquial - monástico - catedralicio	suprior / supriora	suprior / supriora
Participación en reunión de cabildo parroquial - monástico - catedralicio	prior / priora	
Participación en reunión de cabildo parroquial - monástico - catedralicio	prior / priora	

- Lugar: Quejana
- Fecha absoluta 1457/06/30
- Juana Díez de Ayala

Precisamente, es en este proceso de dimensionar HILAME -como paso previo y paralelo al mismo- donde se incorpora el nuevo proyecto que aquí presentamos: la transferencia. En este contexto, no se trata de una transferencia de conocimientos en sentido estricto, sino un ejercicio de transparencia y de comunicación de

nuestro trabajo, que se ha transmitido a un grupo de estudiantes que cursan el segundo ciclo del grado de Historia en la Universidad del País Vasco (UPV/EHU).

Las aportaciones de estos jóvenes, debidamente supervisadas, serán incorporadas a la base de datos principal con el resto de la información disponible

4. La transferencia: estudio de caso

Este grupo de estudiantes se ha integrado en el proyecto a través de un programa de prácticas que se ha desarrollado en tres fases. En primer lugar, una participación activa en la evaluación de la plataforma a través de un cuestionario preparado para la ocasión.



ANÁLISIS	RESUMEN	RESPUESTAS [6]
	¿Qué tal estás? Te pedimos feedback sobre la web hilame.info, que está siendo objeto...	21
*	1 ¿Antes que nada, ¿cómo conociste el PROYECTO HILAME?	16
*	2 ¿Podrías describir en qué consiste en pocas palabras?	12
*	3 ¿A quiénes va dirigido su sitio web?	12
*	4 ¿Cuál fue la primera impresión que te causó hilame.info?	12
*	5 ¿Podrías explicar por qué?	9
*	6 ¿Cuántas bases de datos nos ofrece en este momento?	9
*	7 ¿Podrías nombrarlas?	9
*	8 ¿Cuántos resultados encontramos para Montoya en la Base de Datos Principal?	9
*	9 ¿Qué hay que hacer para tenerlos todos en un solo listado?	9
*	10 ¿Cuántos Oñacinos y Gamboinos están documentados en HILAME este momento?	9

Esta iniciativa, de gran valor por aportar miradas nuevas, nos proporcionó una visión más imparcial de su funcionalidad y una perspectiva a nivel de usuario de las necesidades a cubrir. Por tanto, consideramos que ha sido una actividad esencial para implementar las mejoras en HILAME y avanzar en su dimensión pública.

En una segunda fase, buscamos dotar a los estudiantes de los instrumentos necesarios para el desarrollo de su trabajo. Para ello, elaboramos un [ciclo formativo](#), *Jornadas formativas sobre Humanidades Digitales: el proyecto HILAME* (online debido a las circunstancias actuales, marcadas por una pandemia por covid-19), que dividimos en dos sesiones. La primera sesión se destinó a presentar el proyecto, introduciendo algunos de sus signos identitarios. Primeramente, el “diseño ecoinformático” (Dacosta y Jular, 2018), que implica la integración de profesionales de diferentes especialidades como historiadores, antropólogos, diseñadores, técnicos informáticos, etc. en la génesis del proyecto. En segundo lugar, el fundamento teórico de la plataforma y su conceptualización: la prosopografía y el uso de factoides para la descripción de individuos. Y, finalmente, la metodología empleada para el volcado de datos, la denominada “orfebrería de datos”, fundamentada en la selección de un contexto que nos permite contextualizar -valga la redundancia- en su tiempo y lugar a cada uno de los sujetos históricos citados entre la documentación ya publicada, así como establecer sus vínculos sociales, completando su ficha de forma manual y personalizada. La segunda sesión, en cambio, se centró en la labor que iban a desarrollar, profundizando en la metodología de trabajo específica, las fuentes y sus características, así como los problemas y dificultades más habituales que podían encontrar.

The screenshot displays the HILAME web interface. At the top, there is a 'Categoría' dropdown menu. Below it, a list of categories is shown, each with a folder icon: 'Identificación del individuo', 'Relaciones de parentesco', 'Vínculos sociales', 'Posesiones y estado patrimonial', 'Eventos políticos', 'Eventos vitales', 'Eventos, actos y relaciones socioeconómicas', and 'Actos religiosos'. A blue instruction bar reads: 'Seleccione las entidades involucradas en el factoido y pulse en + para añadirla a la tabla'. Below this, there are three sections for adding entities: 'Entidades del factoido' with a dropdown set to 'Personas' and an 'Añadir persona' button; a dropdown set to 'Lugares' and an 'Añadir lugar' button; and a dropdown set to 'Cronos' and an 'Añadir crono' button. At the bottom, there is an 'Observaciones' field.

Orden	Id	Nombre	Tabla
-------	----	--------	-------

En este contexto, optamos por procesar un tipo documental concreto: las Actas Municipales de Vitoria-Gasteiz, que están digitalizadas y en acceso abierto a través de la web pública del [Ayuntamiento de Vitoria](#). Es una documentación de gran valor para el estudio de la sociedad y del funcionamiento urbano, que se conserva ininterrumpidamente desde septiembre de 1479 hasta la actualidad. Asimismo, presenta unas características formales bien definidas para la fecha: la periodicidad anual del curso político, con el nombramiento de oficiales el

29 de septiembre; la estructura de los acuerdos tomados en cada reunión concejil, que aportan la información necesaria y están precedidos de un resumen de su contenido; y el tipo gráfico, cortesana que está evolucionando hacia procesal, poco cuidado.

A pesar de ser unas fuentes inéditas, y a excepción de un documento previo, el que recoge los acuerdos municipales del año 1428 y parte de los del 29 (Díaz de Durana, 1984), su edición documental es un proyecto en ciernes de nuestro grupo de investigación. Por tanto, hemos podido facilitarles una transcripción provisional, aunque pulida, de los manuscritos a procesar. La metodología escogida para el procesamiento documental ha seguido un sistema indirecto, consistente en el volcado de datos en una plantilla de Excel que nos permite procesar un gran volumen de individuos de forma seriada con el empleo de pocos factoides. Posteriormente, los datos ya volcados son revisados por los documentalistas del equipo, labor imprescindible para poder vincular las personas entre sí y fijar, por ejemplo, sus lazos de parentesco (Dacosta y Jular, 2020: 98).

Ficha persona:									Ficha Crono	Ficha Crono	Ficha Crono	Ficha Lugares							
Apellido / Antonomasia	Nombre personal	Nombre 1	Apellido patronímico	Nombre 2	Apellido (activo o nombre de pila)	Subnombre	Oficio	Sexo	Año	Mes	Día	Lugares	Factoid1	Factoid2	Factoid3	Factoid4	Factoid5	Factoid6	Notas del editor
Juan			Martínez	de	Vergara		Masculino		1429			Vitoria	Miembro u oficial de concejo, regimiento, ayuntamiento/Alcalde						
Diego			Alonso	de	Luskano		Masculino		1429			Vitoria	Miembro u oficial de concejo, regimiento, ayuntamiento/Alcalde						
Juan			Pérez	de	Luzú		Masculino		1429			Vitoria	Miembro u oficial de concejo, regimiento, ayuntamiento/Regidor						
Juan			Iñiguez	de	Ullivarri		Masculino		1429			Vitoria	Miembro u oficial de concejo, regimiento, ayuntamiento/Regidor						
Juan			Martínez	de	Landa		Masculino		1429			Vitoria	Miembro u oficial de concejo, regimiento, ayuntamiento/Regidor						
Fernando				de	Cachu		Masculino		1429			Vitoria	Miembro u oficial de concejo, regimiento, ayuntamiento/Regidor						
Fernán			Ruiz	de	Arcadie		Masculino		1429			Vitoria	Miembro u Dedicación - ocupación principal/Escribano - notario						
Miguel			Pérez	de	Matauco		Masculino		1429			Vitoria	Miembro u oficial de concejo, regimiento, ayuntamiento/jurado						
Pedro			Sánchez	de	Urtarte		Masculino		1429			Vitoria	Miembro u o Artesanos/oficios especializados						
Juan			Pérez	de	Matauco		Masculino		1429			Vitoria	Miembro u oficial de concejo, regimiento, ayuntamiento/Escribano						
Juan			Martínez	de	Heali		Masculino		1429			Vitoria	Miembro u oficial de concejo, regimiento, ayuntamiento/Procurador - Representante legal o político						
Juan			Martínez	de	Andoñu		Masculino		1429			Vitoria	Miembro u oficial de concejo, regimiento, ayuntamiento/Procurador - Representante legal o político						
Pedro				de	Villafrañca		Masculino		1429			Vitoria	Miembro u oficial de concejo, regimiento, ayuntamiento/Carcelero						
Miguel			García	de	Escolla		Masculino		1429			Vitoria	Miembro u oficial de la Hermandad o Diputación						
Juan			Pérez	de	Oña el mozo		Masculino		1429			Vitoria	Miembro u oficial de la Hermandad o Diputación/Alcalde						
García				de	Mendoza		Masculino		1429			Vitoria	Merino Mayor						

Finalmente, los estudiantes han llevado a cabo el procesamiento documental, que ha sido particularmente enriquecedor desde dos perspectivas. Por un lado, desde la perspectiva de los alumnos, convirtiéndose en una herramienta idónea para profundizar en su formación medievalista, especialmente entre aquellos interesados en la historia bajomedieval (siglos XIVXV), pues favorecemos un primer contacto estrecho con las fuentes primarias y su tratamiento. De este modo, gracias a la transcripción que reciben y al acceso abierto a los documentos digitalizados, se familiarizan con la metodología de edición de fuentes, su estructura, el léxico propio del periodo, los tipos escriturarios de finales de la Edad Media, los usos y costumbres urbanos, el funcionamiento de una organización concejil de envergadura, las competencias que asumían sus oficiales, etc. En definitiva, reciben una inmersión en la metodología de trabajo para el análisis histórico en el periodo que transcurre entre la Edad Media y la Modernidad.

Por otro lado, desde el punto de vista del grupo de investigación y del proyecto, es una oportunidad inmejorable para ampliar los datos disponibles en la plataforma. Esta no es una cuestión baladí, pues el acrónimo HILAME (Hidalgos, Labradoras, Mercaderes) es un reflejo de la dicotomía de la sociedad medieval -campo, ciudad-, a la que se ha añadido una perspectiva de género. A grandes rasgos, la palabra Hidalgos representa el ámbito rural, estructurado en torno a los señores que lo gobernaban. De hecho, el primer aspecto que abordó HILAME fue el conflicto banderizo que enfrentó a los linajes vizcaínos bajomedievales (Guerra, 1930). Posteriormente, la información sobre estos linajes se ha ido ampliando con la inclusión de otras fuentes que completan otros aspectos de su *modus vivendi*: espacios de poder, fuentes de renta, patrimonio, entre otros. (Goicolea *et alii*, 2005; Dacosta *et alii*, 2010 y Munita *et alii*, 2014). Asimismo, la referencia a Labradoras representa la perspectiva de género. La elección del término tiene su origen en el análisis que Díaz de Durana, 2004, realizó sobre las labradoras alavesas; no obstante, la mayor aportación sobre las mujeres que integran la base de datos

procede de dos fuentes monásticas femeninas alavesas (Paz Moro, 2013 y 2017). Por último, los Mercaderes simbolizan el mundo urbano, el más infrarrepresentado hasta la fecha. Por tanto, las Actas Municipales de Vitoria contribuyen a diversificar aún más el contenido de HILAME, ilustrando las peculiaridades de la ciudad a través de los oficios concejiles y las familias que los ostentaban, sus profesiones y actividades económicas principales, los conflictos judiciales en los que se veían inmersos, la participación de las minorías étnicas en diversas actividades socioeconómicas, entre otras cuestiones.

5. Valoración y resultados

Aunque todavía no estamos en posición de realizar una valoración exhaustiva y global de la iniciativa, pues las prácticas aun están en ejecución, podemos adelantar algunas ideas. En primer lugar, las mayores dificultades que han tenido que afrontar los estudiantes. Estas se han manifestado esencialmente en el aspecto metodológico, siendo la selección de contextos y la asignación de factoides los mayores escollos. En este sentido, consideramos que estos son problemas inherentes a la metodología de trabajo, al desconocimiento de la plataforma y a la inexperiencia; a pesar de lo cual, hemos intentado ofrecer una orientación personalizada que les permitiera superar esas problemáticas iniciales.

Además, esperamos lograr el volcado de todos los individuos contenidos en las Actas Municipales de la villa/ciudad de Vitoria desde 1479 a 1484, así como los del manuscrito de los años 1428 y 1429. Por tanto, incorporaremos una relación completa de los oficiales concejiles de ese periodo y de otros vecinos pertenecientes a gremios o minorías étnicas. Esta circunstancia nos aportará una mirada de conjunto a esos datos facilitando el análisis de la organización socioeconómica de la Vitoria de finales del siglo XV, sus usos y costumbres, las familias preeminentes, la conflictividad política y social, entre otras muchas cuestiones. Este es un resultado notorio, pues creemos que sin la participación de los alumnos de la UPV-EHU no habríamos alcanzado un grado de procesamiento documental tan detallado.

Finalmente, nos gustaría analizar la experiencia desde el punto de vista de los ejecutores del proyecto, los alumnos de Historia de la UPV-EHU, a través de una valoración que les solicitaremos tras la conclusión del trabajo.

6. Bibliografía

Baraibar, Álvaro, *Visibilidad y divulgación de la investigación desde las Humanidades digitales: experiencias y proyectos*, Pamplona, Universidad de Navarra, 2014.

Boute, Bruno, “Towards more uniform database structures for prosopographical research: work in progress in University history – the example of the Lovanienses Database”, en K.S.B.

Keats-Rohan (ed.), *Resourcing Sources. The Use of Computers in Developing Prosopographical Methodology*, Oxford, University of Oxford, 2002, pp. 35-48.

Charle, Christophe, “Problèmes de traitement informatique d’une enquête sur trois élites en 1901”, en H. Millet (ed.), *Informatique et prosopographie*, París, CNRS, 1985, pp. 233-246.

Dacosta, Arsenio; Díaz de Durana, José Ramón, “Prosopografía y bases de datos. Desafíos teóricos y metodológicos para el estudio de la Edad Media”, en A. I. Carrasco Manchado, *El historiador frente a las palabras: lenguaje, poder y política en la sociedad medieval: nuevas herramientas y propuestas*, Lugo, Editorial Axac, 2017, pp. 191-217.

- Dacosta, Arsenio; Jular, Teresa, “Hidalgos en la escalera del diseño”, en A. Dacosta, C. Jular, J. R. Díaz de Durana, *Hidalgos e hidalguía en la Península Ibérica (siglos XII-XV)*, Madrid, Marcial Pons, 2018, pp. 429-446.
- Dacosta, Arsenio; Jular, Teresa, “HILAME ante los desafíos de las Humanidades Digitales: convergencia e interoperabilidad”, en S. de la Torre, E. Etxeberria, J. R. Díaz de Durana, *Valer más en la tierra, Poder, violencia y linaje en el País Vasco bajomedieval*, Madrid, Sílex, 2020.
- Dacosta, Arsenio; Lema, J. Ángel; Munita, J. Antonio; Díaz de Durana, J. Ramón, *Poder y privilegio. Nuevos textos para el estudio de la nobleza vizcaína al final de la Edad Media (1416-1527)*, Bilbao, EHU Press, 2010.
- Díaz de Durana, José Ramón, *Vitoria a finales de la Edad Media (1428-1476)*, Vitoria, Diputación Foral de Álava, 1984.
- Díaz de Durana, José Ramón, *La otra nobleza, escuderos e hidalgos sin nombre y sin historia. Hidalgos e hidalguía universal en el País Vasco al final de la Edad Media (1250-1525)*, Bilbao, EHU Press, 2004.
- Goicolea, Fco. Javier; Villanueva, Eider; Lema J. Ángel; Fernández de Larrea, Jon Andoni; Munita, J. Antonio; Díaz de Durana, J. Ramón, *Honra de hidalgos, yugo de labradores: nuevos textos para el estudio de la sociedad rural alavesa (1332-1521)*, Bilbao, EHU Press, 2005.
- Jular, Teresa; Dacosta, Arsenio, “HILAME (Hidalgos, Labradoras, Mercaderes): procesamiento y visualización de datos prosopográficos”, en *Humanidades Digitales. Miradas hacia la Edad Media*, Berlín, Boston, De Gruyter, 2018, pp. 79-89.
- Millet, Hélène, “From sources to data. The construction of a prosopographical data-bank”, en P. Denley, D. Hopkin (eds.), *History and computing*, Manchester, Manchester University Press, 1987, pp. 63-67.
- Munita, J. Antonio; Dacosta, Arsenio; Lema J. Ángel; Paz, Agurtzane; Díaz de Durana, J. Ramón, “*En tiempo de ruidos e bandos*”. *Nuevos textos para el estudio de los linajes vizcaínos: los Barroeta de la merindad de Marquina (1355-1547)*, Bilbao, EHU Press, 2014.
- Paz Moro, Agurtzane, *El monasterio de Barría. Historia y documentos (1232-1524)*, Bilbao, EHU Press, 2013.
- Paz Moro, Agurtzane, *Colección documental del monasterio de San Juan de Quejana (Álava, 1332-1525)*, Bilbao, EHU Press, 2017.
- Smythe, Dion C., “A whiter shade of Pale: issues and possibilities in prosopography”, en K. S. B. Keats-Rohan (ed.), *Prosopography approaches and applications: a handbook*, Oxford, University of Oxford, pp. 2007, 127-137.
- Tinti, Francesca, “Facing factoids”, paper presentado en el *King's College London, Second PASE Colloquium, 26 April 2003*, documento electrónico disponible en: http://rae2007.cch.kcl.ac.uk/pase/redist/pdf/col_001.pdf

Discursos científicos sobre Arqueología Prehistórica desde una perspectiva de género: Un estudio de caso de inicios del siglo XX

Raquel Licerias-Garrido (Universidad de Lancaster, Reino Unido)

Alba Comino (investigadora independiente)

Patricia Murrieta-Flores (Universidad de Lancaster, Reino Unido)

Palabras clave

Género, Arqueología Prehistórica, Discurso, Inteligencia Artificial, Análisis Lingüístico

Descripción del problema de investigación, metodología y resultados obtenidos

Los orígenes de la Arqueología española como disciplina científica se remontan al siglo XIX. El primer tercio del siglo XX es especialmente significativo, caracterizado por la diversidad de iniciativas investigadoras y la profesionalización de la disciplina. Los discursos arqueológicos creados durante ese periodo sentaron las bases de la interpretación actual, construyendo las narrativas históricas sobre lógicas actualistas y patriarcales que han llegado hasta nuestros días. Durante este periodo, las mujeres comenzaron a incorporarse progresivamente a las instituciones vinculadas con la gestión del patrimonio, especialmente a partir del *Real Decreto del 8 de marzo de 1910* que les permitió acceder a estudios universitarios (Díaz Andreu 1998).

En este contexto, se desarrolló el proyecto del *Catálogo Monumental de España* (CME), cuyos objetivos fundamentales fueron la documentación sistemática del patrimonio histórico-artístico disperso por la geografía española, para protegerlo de manera coherente y eficaz, y la publicación de los resultados obtenidos, con el fin de fomentar su investigación y transferir el conocimiento generado a la sociedad (Domingo y Cendón 2012). El CME constituye una obra imprescindible para la arqueología puesto que reúne descripciones, fotografías, planos y dibujos de numerosos bienes culturales -en ocasiones desaparecidos o transformados en la actualidad-, y, además, contiene interpretaciones teóricas sobre la cultura material descrita y relatos sobre la experiencia de los autores a la hora de elaborar el catálogo. Desde esta perspectiva, el CME sigue siendo una importante fuente documental para estudiar la evolución del patrimonio de España a lo largo del tiempo, así como para analizar el desarrollo de la propia disciplina.

Este trabajo analiza la representación de las mujeres en la investigación arqueológica de contextos prehistóricos en general, y de la Edad del Hierro en particular, tomando como base los volúmenes de las provincias de Ávila, Soria y Burgos. Además, se reivindica el papel de las mujeres en la investigación arqueológica y en la elaboración del propio CME. Para ello, se utilizan herramientas y metodologías desarrolladas en el campo de las Humanidades Digitales. En las últimas décadas, esta disciplina ha servido como puente entre los instrumentos que proporcionan las Ciencias de la Computación y los marcos teóricos y problemáticas de las Humanidades, desarrollando técnicas y métodos que permiten realizar análisis de datos a gran escala sobre la utilización del lenguaje en la construcción de las narrativas, entre otras múltiples posibilidades.

El CME es una obra que nunca llegó a publicarse en su totalidad, por lo que gran parte de los volúmenes pueden leerse en la versión manuscrita entregada por sus autores al Ministerio de Instrucción Pública y Bellas Artes.

Actualmente, se conservan y son accesibles, gracias a la digitalización llevada a cabo por el Instituto de Patrimonio Cultural de España a partir de 2008, a través del repositorio digital de la Biblioteca Tomás Navarro Tomás del CSIC. Por este motivo, para examinar los discursos historiográficos desde una perspectiva de género, el paso inicial de esta investigación ha sido obtener una versión de los catálogos a analizar en formato legible por computadora. En primer lugar, se ha recurrido a [Transkribus](#), una herramienta que emplea Inteligencia Artificial para producir transcripciones semiautomáticas de textos manuscritos a partir de una muestra generada manualmente. De este modo, se han creado modelos personalizados para transcribir cada una de las letras de los autores, alcanzando una precisión de en torno al 95% en cada una de las transcripciones automáticas. Posteriormente, estas transcripciones han sido revisadas para garantizar la correspondencia con la fuente original. En segundo lugar, se han utilizado técnicas de la Lingüística Computacional integradas en la aplicación [Voyant Tools](#) para ejecutar el análisis lingüístico de género sobre la Prehistoria y la Edad del Hierro, evaluando la frecuencia de vocablos, su contexto y las palabras clave, entre otros aspectos. Finalmente, además de explorar los contextos y el lenguaje que alude a las mujeres en los distintos catálogos examinados, este estudio ha llevado a cabo un análisis historiográfico de algunas de las mujeres coprotagonistas en su elaboración, así como también de aquellas mencionadas en el texto.

Los principales resultados de este trabajo han sido:

- La creación de tres modelos para transcribir los textos manuscritos, disponible en la plataforma Transkribus.
- La transcripción de los volúmenes correspondientes a las provincias de Ávila, Soria y Burgos, disponible en los repositorios abiertos de [figshare](#) y [github](#) (Licerias-Garrido *et al.* 2020 a, b, c).
- Un estudio comparativo entre los resultados de los análisis lingüísticos y el discurso de género percibido por las investigadoras durante el proceso de creación de la muestra y de revisión de las transcripciones.
- Un análisis lingüístico de las narrativas de la Prehistoria y la Edad del Hierro en clave de género que permite identificar los patrones, sesgos y estereotipos en los discursos sobre el pasado construidos por autores españoles de principios del siglo XX.
- La reivindicación del papel de mujeres como Elena Rodríguez-Bolívar, Sagrario Gómez-Moreno y Antonia Herreros en el éxito de la empresa.

Bibliografía

- Díaz-Andreu, M. (1998): Spanish women in a changing world. Strategies in the search for self-fulfilment through antiquities. En M. Díaz-Andreu y M. L. Stig (eds.): *Excavating Women. A History of Women in European Archaeology*. Routledge, London y New York: 125–145.
- Domingo, M. y Cendón, O. (eds.): *El Catálogo Monumental de España (1900-1961). Investigación, restauración y difusión*. Ministerio de Cultura, Madrid.
- Licerias Garrido, R.; Comino, A. y Murrieta Flores, P. (2020a): *Transcripción del Catálogo Monumental de España: Provincia de Ávila por Manuel Gómez Moreno (1900-1901)*. Figshare. Dataset. <https://doi.org/10.6084/m9.figshare.12006318.v1>.
- Licerias Garrido, R.; Comino, A. y Murrieta Flores, P. (2020b): *Transcripción del Catálogo Monumental de la Provincia de Soria por Juan Cabré (1916-1917)*. Figshare. Dataset. <https://doi.org/10.6084/m9.figshare.12006273.v1>.
- Licerias Garrido, R.; Comino, A. y Murrieta Flores, P. (2020c): *Transcripción del Catálogo Monumental y Artístico de la Provincia de Burgos por Narciso Sentenach (1925)*. Figshare 2020. Dataset. <https://doi.org/10.6084/m9.figshare.12006327.v1>

Redes y comunidades de escritoras españolas de la primera Edad Moderna: análisis de datos y visualización

María D. Martos, Bieses, UNED, Madrid, mdmartos@flog.uned.es

Palabras clave: Redes sociales, visualización, gehi, escritoras españolas, primera Edad Moderna

Descripción del problema de investigación, metodología y resultados obtenidos

En la tarea de reconstrucción de los entornos de escritura de las mujeres en la primera Edad Moderna nos topamos frecuentemente con la escasez de datos y con una información bastante desestructurada. El proyecto BIESES viene trabajando en los últimos años en las redes de sociabilidad de las escritoras como instrumento de análisis para entender cuáles fueron los factores que influyeron en su producción literaria: qué tipo de apoyos recibían, si existieron grupos o personas que alentaron especialmente su actividad como escritoras, si esta se vio favorecida al gestarse en alguna tipología de red social más que en otras, etc.

La propuesta de Bieses parte de la valiosa información contenida en los paratextos, con el apoyo en otras fuentes secundarias, para analizar las relaciones de los agentes implicados en el proceso que lleva al producto impreso y a la intervención que tiene cada uno de ellos en este proceso, definiendo así tipos de relaciones, de redes y de estrategias de posicionamiento autorial de las escritoras. Para el diseño y visualización de redes se ha trabajado con distintas herramientas, en función de los objetivos de la investigación, de la tipología de red identificada y de las necesidades de formalización: Gephi, Visone y Tableau. De este proceso ha resultado un tejido relacional de individuos (nodos) y vínculos (aristas) que aporta una nueva visión sobre el entorno de las escritoras españolas hasta 1800. En esta comunicación se expondrá el análisis de datos que se ha llevado a cabo así como su visualización.

Figuras, tablas, mapas o gráficas



Bibliografía

- Burt, R. S., Kilduff, M., & Tasselli, S. (2013). Social network analysis: Foundations and frontiers on advantage. *Annual review of psychology*, 64, 527–547.
- Choudhury, A., Kaushik, S., & Dutt, V. (2017). Social-Network Analysis for Pain Medications: Influential physicians may not be high-volume prescribers (pp. 881–885). ACM.
- Fagan, J. (2017). Introduction to GEPHI.
- Gephi - The Open Graph Viz Platform. (s. f.). Recuperado a partir de <https://gephi.org/>
- Gualda, E. (2018). Big data y análisis de redes sociales.
- Guerra-Gomez, J. A., Wilson, A., Liu, J., Davies, D., Jarvis, P., & Bier, E. (2016). Network Explorer: Design, Implementation, and Real World Deployment of a Large Network Visualization Tool. En *Proceedings of the International Working Conference on Advanced Visual Interfaces* (pp. 108–111). ACM.
- Jacomy, M., Venturini, T., Heymann, S., & Bastian, M. (2014). ForceAtlas2, a continuous graph layout algorithm for handy network visualization designed for the Gephi software. *PloS one*, 9(6), e98679.
- Leifeld, P. (2017). Discourse network analysis. *The Oxford Handbook of Political Networks*, 301.
- Lodhi, P., Annapoorna, E., Mishra, O., & Sinha, A. (2018). Analysis of Real Life Networks Using Centrality Measure.
- Mackinlay, J., & Winslow, K. (2015). Designing Great Visualizations. *Tableau Software*.
- Maimon, O., & Browarnik, A. (2009). NHECD-Nano health and environmental commented database. En *Data mining and knowledge discovery handbook* (pp. 1221–1241). Springer.
- Marten Düring, "De la hermenéutica a las redes de datos: Extracción de datos y visualización de redes en fuentes históricas", traducido por Maria José Afanador-Llach, *The Programming Historian en español* 1 (2017), <https://doi.org/10.46430/phes0002>.
- Shah, P., & Mehta, R. (2017). Comparative Analysis of Social Network Analysis and Visualisation Tools.
- Stovel, K., & Shaw, L. (2012). Brokerage. *Annual Review of Sociology*, 38, 139–158.
- Tableau Public. (s. f.). Recuperado a partir de <https://public.tableau.com/en-us/s/>
- Visone. (s. f.). Recuperado, a partir de <https://visone.info/>
- Wellman, B. (1993). An egocentric network tale: comment on Bien et al. (1991). *Social Networks*, 15(4), 423–436.

Hacia la codificación del patrimonio musical hispánico en el contexto de las Humanidades Digitales: desafíos y repercusión historiográfica, económica y cultural

Modera: Fátima Díez Platas · *Asiste:* Alexandra Iglesias Cortés

Dos cambios que ha traído la revolución digital están produciendo un efecto determinante en el desarrollo de las Humanidades y en particular de la Musicología:

- Del acceso limitado a la información y su difusión hemos pasado al acceso a “big data” que afecta a la dimensión de los proyectos, impensables hace pocos años, facilitado también por los enlaces globales en acceso abierto (“linked open data”).
- De recursos documentales basados en microfilm, papel, archivos de imagen y PDF hemos pasado a contenidos codificados (de texto, sonido y notación musical) que pueden ser leídos por máquinas.

Este cambio de paradigma incrementa exponencialmente la capacidad de catalogación, estudio, análisis y difusión de las fuentes de interés musical y su interpretación, así como el tipo de preguntas y proyectos multidisciplinares a los que nos vemos abocados en un entorno de colaboración más estrecha con otras disciplinas. En esta conferencia se examinarán los desafíos y oportunidades que las Humanidades Digitales nos presentan en el ámbito del patrimonio musical hispánico, así como sus repercusiones historiográficas, económicas y culturales.

Estudios de corpus

Modera: Fátima Díez Platas · *Asiste:* César González-Pérez

- 72 Los proyectos HERES y CHARTA 3.0: retos y estrategias comunes para la creación, el alojamiento y el uso de diferentes corpus
Víctor Caballero Gómez, Ricardo Pichel
- 49 El corpus "Oralia diacrónica del español": características y opciones de explotación lingüística
Gael Vaamonde, Rocío Díaz-Bravo
- 47 Modelizando una literatura en el olvido: LDA aplicado a corpus españoles sobre Guinea Ecuatorial y Filipinas
Emilio Pedro Vivó Capdevila
- 101 Proyecto HDA Teatro UNIR, realidades y perspectivas para el estudio del teatro desde las HD
Teresa Santa María Fernández, Monika Dabrowska
- 54 Cervantes desde la Estilometría y el análisis de redes: estudio de su obra
Laura Hernández-Lorenzo

Los proyectos HERES y CHARTA 3.0: retos y estrategias comunes para la creación, el alojamiento y el uso de diferentes corpus

Víctor Caballero Gómez (UAH/UCM)

Ricardo Pichel (UAH)

Palabras clave: XML/TEI, TEITOK, lingüística de corpus, lingüística computacional

Uno de los mayores retos para quienes nos dedicamos a la edición digital de documentos antiguos es conseguir una plataforma o un marco de trabajo que sirva, al mismo tiempo, para alojar los textos anotados, para recuperar y explotar toda la información y para compartir ese trabajo con los investigadores interesados. En esa misma encrucijada se encontraron, en el ámbito de la Universidad de Alcalá y la Universidade de Santiago de Compostela, dos proyectos de orientación e intereses divergentes, pero igualmente convencidos de la necesidad de una herramienta como la que acabamos de describir. Dadas sus características, que ya empiezan a ser de sobra conocidas, y para optimizar esfuerzos y recursos, desde estos dos proyectos recurrimos a la plataforma TEITOK (desarrollada por Maarten Janssen en 2012, <http://teitok.org>), que permite su instalación en un subdominio compartido, reserva un espacio diferenciado para cada corpus y facilita la creación y explotación de los documentos (Janssen, 2016). En esta comunicación presentamos, de un modo muy breve, cada uno de estos proyectos y, con algo más de detalle, el plan y la metodología de trabajo seguidos para este propósito común, prestando especial atención a las distintas fases en las que estos se encuentran y, por ello, a sus diferentes necesidades técnicas.

El primero de estos proyectos (“CHARTA 3.0: De la edición digital a la web semántica”; ref. CM/JIN/2019-008) se propone la conversión al lenguaje XML/TEI y la integración en TEITOK de los más de dos mil documentos de los que consta el corpus CHARTA (Corpus Hispánico y Americano en la Red: Textos Antiguos; <http://corpuscharta.es>). El segundo (“HERES. Patrimonio textual ibérico y novohispano. Recuperación y memoria”; ref. CM/2018-T1/HUM-10230) persigue, entre otros objetivos, la puesta en marcha y desarrollo de varios corpus y repositorios digitales, principalmente: el Corpus Diacrónico del Español de Norteamérica (CORDINA), un corpus de referencia del español histórico norteamericano para los siglos XVI-XX; el Escritorio Galego-Portugués Antigo (EGPA), un repositorio compuesto por fuentes archivísticas y textos literarios gallegos y portugueses de entre los siglos XIII y XIX; y, por último, el Corpus de textos troyanos ibéricos (NOSTOI), conformado por todos los testimonios medievales vinculados con el ciclo clásico (materia de Troya) entre los siglos XIII y

XVI.

Esta notable diversidad se incrementa por las características del corpus CHARTA, concebido como agregador de los trabajos de una decena de grupos. Así, en esta iniciativa confluye una variada y nutrida cuenta de documentos, metodologías y objetivos, que van desde el estudio de la escritura de mujeres (Almeida Cabrejas, 2017) hasta la propia investigación de carácter ecdótico (SánchezPrieto Borja, 1998, 2011, 2012; Red CHARTA, 2013; Pichel, 2013; Díez del Corral y Martín Aizpuru, 2014), centrada también en la edición digital (Martín Aizpuru, 2016).

En cuanto a los aspectos técnicos y de utilidad para todos los proyectos implicados, ha sido necesario realizar una serie de ajustes de los criterios de edición digital propuestos para la Red CHARTA (Isasi Martínez et al., 2020), modificaciones siempre orientadas a permitir la explotación de todos los recursos ofrecidos por TEITOK y a asegurar una mejor adecuación a la arquitectura de la plataforma. También de interés general es el desarrollo de herramientas y procedimientos, complementarios a los integrados en TEITOK como el algoritmo que efectúa un primer etiquetado POS (Janssen, 2012), que automaticen la creación de la totalidad o, al menos, de partes sustanciales de los documentos XML/TEI a partir de bases de datos cuya información procede de los formatos tradicionalmente empleados en la Red CHARTA (HTML, DOCX, XLSX, etc.). Estas estrategias tendentes a la automatización permiten conservar el trabajo de edición realizado hasta la fecha y reducen las posibilidades de cometer errores en la conversión. Por su parte, para la edición de los nuevos documentos, aquellos que se crearán de manera nativa en XML/TEI, estamos preparando una serie de módulos formativos que faciliten el etiquetado, ya sea en editores de XML o, empleando un protomarcado diseñado a tal efecto, desde el conversor de documentos habilitado en TEITOK.

De cara a la explotación semántica de los distintos corpus, contamos con las potentes herramientas de búsqueda integradas en la propia plataforma. No obstante, para generar transformaciones que ofrezcan los textos en versiones cómodamente publicables en trabajos científicos o para realizar estudios de carácter cuantitativo basados en diferentes aspectos del mismo o de varios testimonios, estamos desarrollando y pondremos a disposición del investigador algunas hojas de estilos XSL, en las que se tendrán en cuenta las especificidades de nuestro etiquetado y las eventuales necesidades de recuperación de información que puedan surgir. Esto último, además, sirve para evidenciar la escasa dependencia respecto de la plataforma elegida. De este modo, y gracias al empleo de XML, un estándar idóneo para la interoperabilidad y la explotación de los datos, podremos seguir creando documentos y analizando los corpus con independencia de la mayor o menor durabilidad de una plataforma, de enorme valor y utilidad, pero no imprescindible, como la que centra nuestra comunicación.

Figuras



Fig. 1: Visualización parcial de uno de los documentos del Corpus CHARTA



Fig. 2: Opciones de búsqueda, todavía provisionales, en la plataforma

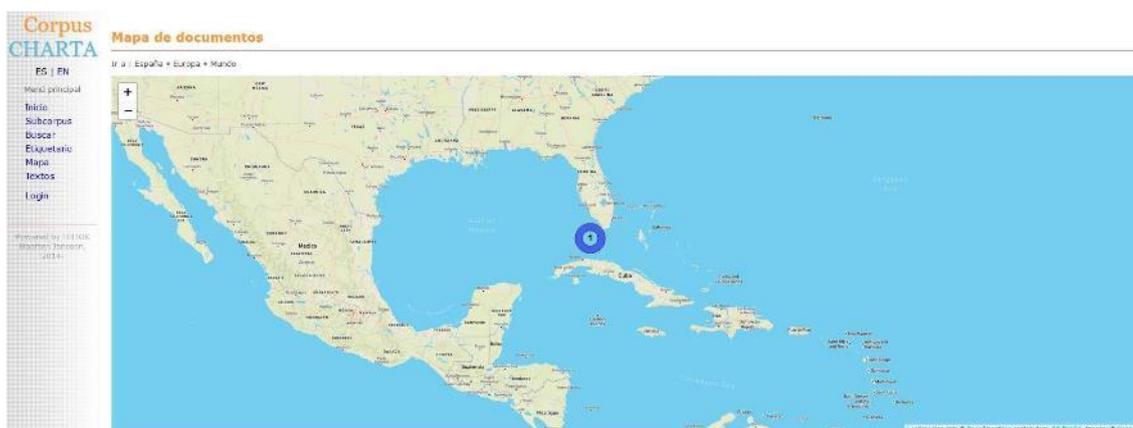


Fig. 3: Navegación en el mapa de documentos

Bibliografía

Almeida Cabrejas, Belén (2017): "Mujeres que escriben y sistema(s) gráfico(s) en el siglo XVI: acercamiento desde un corpus de cartas", en Belén Almeida Cabrejas, Rocío Díaz Moreno y María del Carmen Fernández López (eds.), "Cansada tendré a vuestra excelencia con tan larga carta": estudios sobre aprendizaje y práctica de la escritura por mujeres en el ámbito hispánico (1500-1900). Lugo: Axac, pp. 9-53.

Díez del Corral Areta, Elena; Leyre Martín Aizpuru (2014): "Sin corpus no hay historia: la Red CHARTA como un proyecto de edición común", Cuadernos de Lingüística de El Colegio de México 2, pp. 287-314.

Isasi Martínez, Carmen; Leyre Martín Aizpuru, Santiago Pérez Isasi, Elena Pierazzo y Paul Spence (2020): Edición digital de documentos antiguos: marcación XML-TEI basada en los criterios CHARTA. Sevilla: Universidad de Sevilla [DOI: <https://dx.doi.org/10.12795/9788447222049>].

Janssen, Maarten (2012): "NeoTag: a POS tagger for grammatical neologism detection", en Proceedings of the Language Resources and Evaluation Conference (LREC 2012), ELRA (Estambul, mayo de 2012), pp. 2118-2124.

Janssen, Maarten (2016): "TEITOK: Text-Faithful Annotated Corpora", en Proceedings of the Tenth International Conference on Language Resources and Evaluation (LREC 2016). Portorož: European Language Resources Association (ELRA), pp. 4037-4043.

Martín Aizpuru, Leyre (2016): "Algunos recursos informáticos al servicio de la edición de textos: la edición en XML-TEI", en Chiara Albertin y Santiago del Rey Quesada (coords.), Hispanica

Patavina. Estudios de historiografía e historia de la lengua española en homenaje a José Luis Rivarola. Padua: CLEUP, pp. 139-154.

Pichel, Ricardo (2013): "En torno a la presentación crítica de un texto medieval gallego-castellano (s. XIV). Algunos aspectos grafemáticos", en María Concepción Fernández López, Mercedes Suárez Fernández y Alexandre Veiga (eds.), *Oh lux Iberiae. En torno a las letras en la España medieval*. Lugo: Axac, pp. 135-162.

Red CHARTA (2013² [2011]): Criterios de edición de documentos hispánicos (Orígenes - siglo XIX) [[https:// www.redcharta.es/criterios-de-edicion/](https://www.redcharta.es/criterios-de-edicion/)].

Sánchez-Prieto Borja, Pedro (1998): *Cómo editar los textos medievales: criterios para su presentación gráfica*. Madrid: Arco/Libros.

Sánchez-Prieto Borja, Pedro (2011): *La edición de textos españoles medievales y clásicos. Criterios de presentación gráfica*. San Millán de La Cogolla: Cilengua.

Sánchez-Prieto Borja, Pedro (2012): "La red CHARTA: nuevas perspectivas en la edición y estudio de documentos hispánicos", en María Jesús Torrens y Pedro Sánchez-Prieto Borja (coords.), *Nuevas perspectivas para la edición y el estudio de documentos hispánicos antiguos*. Berna: Peter Lang, pp. 17-44.

El corpus *Oralia diacrónica del español*: características y opciones de explotación lingüística

Vaamonde, Gael (Universidad de Granada)

Díaz-Bravo, Rocío (Universidad de Granada)

Palabras clave: lingüística de corpus, edición digital, lingüística histórica, XML-TEI, anotación de corpus

1. Introducción

Es sabido que la compilación de corpus históricos está condicionada por ciertas limitaciones, que suelen ser más acusadas a medida que retrocedemos en el tiempo: conservación fragmentaria de textos, dificultad de datación, falta de contextualización, distribución errática de géneros, etc. (Kohnen, 2007; Claridge, 2008; Kytö, 2011). Entre estas limitaciones resulta especialmente reseñable, por su relevancia para la investigación lingüística, la carencia de fuentes directas de discurso hablado: la lingüística histórica necesita acceder a muestras de uso real del lenguaje, pero, simultáneamente, debe asumir la palabra escrita como fuente legítima de datos.

Afortunadamente, en los últimos años se han ido elaborando corpus históricos formados, total o parcialmente, por tipos textuales próximos a la oralidad —esto es, documentos que permiten reconstruir la dimensión de lo oral a través de lo escrito (Oesterreicher, 2004)—, buscando de este modo una representación de la lengua hablada en épocas pasadas. Nos referimos a corpus compuestos por cartas privadas (Raumolin-Brunberg y Nevalainen, 2007; Vaamonde, 2018), diálogos (Culpeper y Kytö, 2010), peticiones de ayuda a la beneficencia (Sánchez-Prieto Borja y Vázquez Balonga, 2017) o declaraciones de testigos (Calderón Campos, 2015; Huber et al., 2012), entre otros.

Siguiendo esta línea de investigación se sitúa el corpus ODE (*Oralia Diacrónica del Español*), que damos a conocer en esta comunicación (<http://corpora.ugr.es/ode/>). ODE es un recurso digital que recoge documentación manuscrita inédita del periodo comprendido entre 1492 y 1900, centrada en tres tipos textuales: inventarios de bienes, declaraciones de testigos en juicios criminales y certificaciones periciales de cirujanos vinculadas con esos juicios.

2. Tipos textuales

Los tres tipos textuales representados en *ODE* suponen el traslado al papel de las declaraciones orales de testigos, tasadores y cirujanos, por lo que no es de extrañar que se cuelean en los documentos rasgos lingüísticos del vernáculo de los declarantes o del escribano. En los tres casos la información se transmite de forma muy similar: el declarante habla y el escribano anota lo que oye, con una finalidad puramente práctica (registrar la verdad de los hechos), y sin una excesiva preocupación ni control por la forma, con el filtro normativo de su formación escrituraria, en pugna con sus propios hábitos dialectales.

Respecto a las declaraciones de testigos, la mayoría de los juicios transcritos tienen relación con delitos de robos, asesinatos, peleas, riñas de vecinos, injurias, estafas, violaciones o abusos de poder, procesos que se han seleccionado por reproducir situaciones en las que es fácil que los testigos narren acontecimientos con fuerte implicación emocional; además, con frecuencia recurren al estilo directo para dar mayor veracidad a sus testimonios. Las expresiones en estilo directo incluidas con frecuencia en las declaraciones de testigos conforman el subcorpus de ODE más próximo al discurso hablado. Los inventarios de bienes, entendiendo como tales «cualquier texto hecho con la finalidad de enumerar, de la forma más minuciosa posible, los bienes de una persona o una institución» (Morala, 2012: 200), contribuyen al conocimiento de las condiciones de vida del pasado y permiten profundizar en la realidad cotidiana de las sociedades antiguas. Al mismo tiempo, son una fuente complementaria de primer nivel para el estudio histórico del léxico, puesto que permiten documentar vocabulario dialectal no habitual en corpus generales.

Finalmente, los pleitos y probanzas criminales pueden incluir declaraciones médico-legales llevadas a cabo por cirujanos, barberos o sangradores. Son certificaciones médicas en las que estos facultativos dictaban al escribano de turno una descripción minuciosa de las heridas sufridas por las víctimas de la agresión que estaba siendo investigada. Se trata de un tipo textual especialmente interesante para estudiar el léxico médico que se empleaba en el día a día de la práctica terapéutica, porque recoge tanto las voces populares como los tecnicismos médicos que servían para nombrar o describir las partes del cuerpo o las dolencias de las víctimas.

Cada tipo textual facilita el análisis de algún nivel lingüístico: los inventarios de bienes sirven preferentemente para estudiar el léxico de la vida cotidiana, el que se suele encontrar en los atlas lingüísticos y etnográficos; las certificaciones médicas complementan a los inventarios proporcionando designaciones anatómicas populares y su contraste con la terminología técnica incipiente; por último, las declaraciones de testigo permiten analizar en mayor profundidad que los textos anteriores la evolución de las estructuras morfosintácticas del español.

3. Metodología

El corpus ODE es continuación del *Corpus diacrónico del español del reino de Granada* (CORDEREGRA). ODE supone, no obstante, una ampliación tanto desde el punto de vista cronológico como geográfico con respecto a su predecesor. Así, ODE incluye documentación de todo el siglo XIX y perteneciente al centro-norte peninsular y a todo el territorio andaluz (frente al límite temporal de 1833 y las actuales provincias de Granada, Málaga y Almería que caracterizaban al corpus

CORDEREGRA).

Otro aspecto destacable de ODE radica en la metodología utilizada, que incluye la transcripción de los manuscritos en lenguaje XML –siguiendo el estándar TEI–, la tokenización y estandarización ortográfica de los textos, la anotación morfosintáctica y lematización del corpus –basadas en el estándar EAGLES–, así como la visualización de los datos a través la plataforma TEITOK (Janssen, 2016). En último término, ODE sigue de cerca el modelo realizado con éxito en el proyecto *Post Scriptum* (Vaamonde, 2018) con el objeto de reunir documentación histórica y ofrecerla en dos formatos preparados para la búsqueda: el de la edición crítica digital y el del corpus lingüísticamente anotado (Calderón Campos y Díaz-Bravo, 2021).

3.1. Transcripción en XML-TEI

Las directrices TEI (*Text Encoding Initiative*), basadas en el lenguaje de marcado XML, ofrecen un estándar para la representación de textos en formato digital ampliamente consolidado en el ámbito de las Humanidades Digitales. Su adopción en ODE no solo permite contar con datos estructurados para facilitar su posterior recuperación, sino que garantiza su preservación, su reutilización o su posible integración en repositorios digitales, entre otras ventajas conocidas (Burnard, 2014).

Entre los aspectos textuales que se han marcado en ODE se incluyen cancelaciones (), adiciones (<add>), lagunas textuales (<gap/>), abreviaturas debidamente expandidas (<abbr> y <ex>), conjeturas (<supplied>), erratas (<sic>), expresiones en estilo directo (<quote>) y palabras en otras lenguas (<foreign>), además de cuestiones estructurales como los inicios de línea (<lb/>) y de folio (<pb/>). A modo de ilustración, recogemos a continuación un fragmento del facsímil de una declaración de testigo (Figura 1), la transcripción correspondiente en texto plano (Figura 2) y la transcripción con marcación XML-TEI que se ha aplicado en ODE (Figura 3):

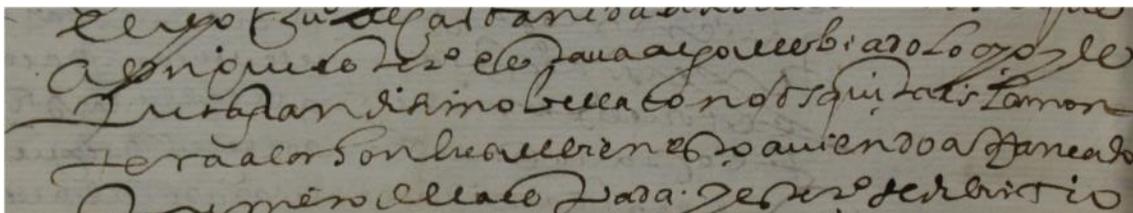


Figura 1. Fragmento de facsímil de una declaración de testigo de 1627

1 el dho Jua de Castaneda dixo a altas bozes, que
2 aunque este to estaba algo desbiado lo oyo: yde
3 puta, grandisimo bellaco, no os quitais la mon
4 tera a los hombres de uien; auiendo arrancado
5 primero de la espada; y este to se dibirtio

Figura 2. Transcripción

4 <lb/> el <abbr>d<ex>ic</ex>ho</abbr> <abbr>Jua<ex>n</ex></abbr> de Castaneda
5 dixo a altas bozes, que <lb/> aunque este <abbr>t<ex>estig</ex>o</abbr> estaba
6 algo desbiado lo oyo: <quote>yde<lb/>puta, grandisimo bellaco, no os quitais la
7 mon<lb/>tera a los hombres de uien; auiendo arrancado <lb/> primero de la espada;
8 </quote> y este se dibirtio

Figura 3. Marcación en XML-TEI

A esta información hay que sumar los metadatos recogidos para cada documento: año, lugar, título del documento, tipología textual o ubicación y signatura en el archivo histórico correspondiente. Cualquiera de estas

cuestiones puede ser utilizada como filtro de búsqueda en el corpus, ya sea mediante lenguaje XPath o mediante lenguaje CQP (*Corpus Query Protocol*) (Evert and Haride, 2011).

3.2. Tokenización

Una vez creada la edición digital, el siguiente paso es convertir dicha edición en un corpus anotado. Esta conversión se lleva a cabo a través de la plataforma TEITOK, que está especialmente diseñada para ver, crear y editar corpus que contienen tanto marcación en XML-TEI como anotación lingüística. Las transcripciones en XML-TEI son importadas a TEITOK y la propia plataforma asiste al compilador en diversas tareas relacionadas con el procesamiento automático de textos. La primera de estas tareas es la tokenización, esto es, la identificación y delimitación de cada palabra y signo de puntuación incluidos en el corpus. Tomando como punto de partida el texto anterior (Figuras 1, 2 y 3), esta tarea, que se aplica de forma automática, devuelve un resultado como el que se ofrece a continuación (la marcación se ha simplificado por razones de claridad):

```
<tok form="ydeputa" id="w-21">yde<lb/>puta</tok>
<tok id="w-22">,</tok>
<tok id="w-23">grandisimo</tok>
<tok id="w-24">bellaco</tok>
<tok id="w-25">,</tok>
<tok id="w-26">no</tok>
<tok id="w-27">os</tok>
<tok id="w-28">quitis</tok>
<tok id="w-29">la</tok>
<tok form="montera" id="w-30">mon<lb/>tera</tok>
<tok id="w-31">a</tok>
<tok id="w-32">los</tok>
<tok id="w-33">hombres</tok>
```

Figura 4. Ejemplo de tokenización en ODE

Como se puede apreciar, el sistema TEITOK añade un elemento <tok> y un identificador único a cada *token* dentro del archivo XML. Si el *token* en cuestión incluye elementos XML en su interior (por ejemplo, <tok>yde<lb/>puta</tok>), el sistema extrae la palabra correspondiente, esto es, el conjunto de caracteres excluyendo la marcación XML-TEI, y lo almacena en un atributo específico (@form) dentro del elemento <tok>, para facilitar su procesamiento o su recuperación posterior.

3.3. Estandarización ortográfica

Una vez tokenizado el texto, el siguiente paso aplicado en ODE es una estandarización de la ortografía de cada palabra. Este proceso resulta necesario, ya que la alta variación ortográfica de los documentos históricos constituye un escollo para obtener resultados precisos en el análisis de frecuencias, en la etiquetación morfosintáctica automática o en la recuperación de información. TEITOK incorpora un normalizador automático que asocia la forma original de cada *token* con su correspondiente forma con grafía normalizada según el estándar actual. El resultado de este proceso se almacena dentro de un atributo @nform (*normalized form*), tal como se refleja en el fragmento de la Figura 5:

```

<tok form="ydeputa" id="w-21" nform="hideputa">yde<lb/>puta</tok>
<tok id="w-22">,</tok>
<tok id="w-23" nform="grandísimo">grandisimo</tok>
<tok id="w-24">bellaco</tok>
<tok id="w-25">,</tok>
<tok id="w-26">no</tok>
<tok id="w-27">os</tok>
<tok id="w-28" nform="quitáis">quittais</tok>
<tok id="w-29">la</tok>
<tok form="montera" id="w-30">mon<lb/>tera</tok>
<tok id="w-31">a</tok>
<tok id="w-32">los</tok>
<tok id="w-33" nform="hombres">honbres</tok>

```

Figura 5. Ejemplo de tokenización y estandarización ortográfica en ODE

El resultado del normalizador es revisado manualmente por un equipo de lingüistas para depurar posibles cuestiones ortográficas que no hayan sido debidamente tratadas de manera automática.

3.4. Anotación morfosintáctica y lematización

Finalmente, una vez que el texto está tokenizado y estandarizado, el último paso es enriquecer el corpus con anotación lingüística. La plataforma *TEITOK* incorpora un etiquetador que, de manera automática, asocia una etiqueta morfosintáctica y un lema a cada *token* del corpus y, además, utiliza los propios datos etiquetados –una vez revisados manualmente– como corpus de entrenamiento, lo que significa que mejora a medida que el corpus aumenta de tamaño. De forma análoga al proceso de normalización ortográfica, el resultado de este proceso también se almacena en atributos dentro del elemento <tok>. La etiqueta morfosintáctica se guarda en un atributo @pos (siglas de *part of speech*) y el lema correspondiente a cada *token* se almacena en un atributo @lemma:

```

<tok form="ydeputa" id="w-21" nform="hideputa" lemma="hideputa" pos="NCMS000">yde<lb id="e-3"/>puta</tok>
<tok id="w-22" lemma="," pos="Fc">,</tok>
<tok id="w-23" nform="grandísimo" lemma="grande" pos="AQSMS0">grandisimo</tok>
<tok id="w-24" lemma="bellaco" pos="NCMS000">bellaco</tok>
<tok id="w-25" lemma="," pos="Fc">,</tok>
<tok id="w-26" lemma="no" pos="R">no</tok>
<tok id="w-27" lemma="os" pos="PP2CP000">os</tok>
<tok id="w-28" nform="quitáis" lemma="quitar" pos="VMIP2P0">quittais</tok>
<tok id="w-29" lemma="el" pos="DA0FS0">la</tok>
<tok form="montera" id="w-30" lemma="montera" pos="NCF5000">mon<lb id="e-4"/>tera</tok>
<tok id="w-31" lemma="a" pos="SPS00">a</tok>
<tok id="w-32" lemma="el" pos="DA0MP0">los</tok>
<tok id="w-33" nform="hombres" lemma="hombre" pos="NCMP000">honbres</tok>

```

Figura 6. Ejemplo de anotación morfosintáctica y lematización en ODE

El sistema de etiquetas utilizado para anotar el corpus *ODE* está basado en la propuesta del grupo *EAGLES* para la anotación morfosintáctica de lexicones y corpus para todas las lenguas europeas (Leech y Wilson, 1996). El conjunto de etiquetas *EAGLES* se rige por un sistema de posiciones: cada etiqueta consta de una secuencia de letras y números, donde cada letra o número representa un rasgo morfosintáctico determinado dependiendo de su posición dentro de la secuencia. Por ejemplo, la forma *hombres* lleva la etiqueta NCMPOOO, que representa nombre (N), común (C), masculino (M), plural (P).

El resultado del anotador morfosintáctico también es revisado manualmente y de forma colaborativa por un equipo de lingüistas. Asimismo, tras la toma de decisiones con respecto a formas que suscitan dudas (Artstein y Poesio, 2008) como *dicho* y su paradigma, o adjetivos deverbales derivados de participios de pasado (Calderón Campos y Díaz-Bravo 2021: 8-9), se aplica un proceso semiautomático de edición múltiple de *tokens* a través de búsquedas CQP (Corpus Query Processor), el cual permite identificar errores regulares y corregirlos de manera sistemática desde la interfaz.

4. Algunos ejemplos de búsquedas en el corpus

A través de la interfaz de búsqueda de *ODE* es posible obtener información de diferente tipo. Se pueden obtener, por ejemplo, las ocurrencias contextualizadas de una palabra determinada (por ejemplo, *vecino*), como se muestra en la Figura 7 (se han registrado un total de 326 ocurrencias de la palabra *vecino* en *ODE*)

contexto	a Jua de Salinas, vezino d esta dha çivdad;	1602 Málaga
contexto	de Herran y Guevara, vecino que fue d esta dha	1721 Almería
contexto	el ganado de Xptobal Rodrigues vo d esta villa; mas	1726 Almería
contexto	parezio Juan Garzia Morales, vezo d esta ua, viudo	1754 Almería
contexto	con Juan Garzia Morales, vezo de esta ua, viudo	1754 Almería
contexto	dn Diego Ramon Ategui, vezino de esta dha villa;	1775 Almería
contexto	de Pedro Sevilla Cortinas, vecino qe fue de la villa	1786 Almería
contexto	estado mancebo, natural y vezino de esta villa de Tabernas	1822 Almería
contexto	su ultima disposición testamentaria, vecino que fue de esta poblacion	1827 Almería
contexto	yo, Diego Morales, vo de la villa de Abla	1834 Almería

Figura 7. Muestra de resultado de concordancias en *ODE*

Posteriormente, se puede organizar el resultado por un criterio específico. Por ejemplo, es posible ordenar el resultado por la forma original transcrita, de tal modo que se obtiene la variación ortográfica de la palabra *vecino* en el corpus:

Group	Count ▼
vezo	83
vezino	69
vo	39
vecino	31
vzo	25
veçino	19
vso	10
vno	7
vesino	7
ueçino	6

Figura 8. Muestra de formas transcritas que fueron normalizadas como *vecino*

También es posible recuperar información a partir de la anotación lingüística del corpus. Por ejemplo, se puede obtener la lista de sustantivos más frecuentes en el corpus y ordenar el resultado por lema. El resultado, que se muestra en la Figura 8, permite hacerse una idea de cuáles son los asuntos tratados con mayor frecuencia en los documentos de ODE:

Group	Count ▼
real	5,686
don	1,564
número	1,420
testigo	1,334
ciudad	813
casa	778
lienzo	694
año	674
vara	649
día	607
villa	599
par	593
vecino	566
señor	566
bien	553
mujer	529
alcalde	477
plata	468

Figura 9. Muestra de lemas de sustantivos más frecuentes en ODE

Puesto que cada documento de ODE está asociado a unas coordenadas geográficas, también es posible rastrear aspectos de tipo dialectal mediante el mapeado de búsquedas en el corpus. Por ejemplo, es posible contrastar las formas *vecino* y *vesino* e identificar en qué zonas geográficas se registra la forma seseante (*vesino*)



Figura 10. Contraste entre *vecino* y *vesino* en un mapa

Sirvan estos resultados como una muestra del tipo de información que es posible obtener desde la interfaz de búsqueda de ODE.

5. Conclusiones

ODE es un recurso electrónico de acceso abierto que ofrece dos productos en uno. Por un lado, constituye un archivo digital compuesto por textos de inmediatez comunicativa (inventarios de bienes, declaraciones de testigo y certificaciones médicas) cuya transcripción puede ser consultada y visualizada junto a la imagen facsimilar de cada documento. Por otro lado, las ediciones digitales en lenguaje XML (estándar TEI) son transformadas en un corpus histórico estandarizado y lingüísticamente anotado que permite realizar búsquedas simples o complejas mediante lenguaje CQP.

Entre sus principales características se pueden destacar las siguientes:

- Usabilidad tanto para el equipo investigador como para los usuarios.
- Niveles personalizables de marcación textual y de anotación lingüística.
- Visualización de versiones diferentes de la misma fuente textual (imagen facsimilar del manuscrito, transcripción paleográfica y edición normalizada), mostrada a través de una interfaz que permite la comparación simultánea de dos versiones.
- Corpus lematizado y anotado morfosintácticamente del español clásico y moderno.
- Posibilidad de descarga de textos como archivos TXT y XML.
- Búsquedas en el corpus a través de interfaces amigables gracias a un *query builder* y a un *tag builder* (creador de búsquedas y de etiquetas)

- Numerosos tipos de búsquedas y de maneras de visualización de los resultados (incluidos índices, palabra clave en contexto o concordancias y mapas).
- Los tipos de búsquedas pueden aplicarse al texto o combinar información textual y metadatos como año, lugar y tipo textual, lo cual es útil para estudios de dialectología histórica.

En el momento de redactar estas líneas, ODE cuenta con un archivo digital compuesto por 238 documentos editados en XML-TEI, lo que se traduce en un corpus de más de 370.000 *tokens* debidamente estandarizados y anotados lingüísticamente. Al finalizar el proyecto, en 2023, se espera llegar a un corpus de un millón de *tokens*, aproximadamente.

6. Bibliografía

- ARSTEIN, R. Y M. POESIO (2008). «Inter-coder agreement for computational linguistics». *Computational Linguistics*, 34(4): 555-96. <https://www.aclweb.org/anthology/J08-4004.pdf>
- BURNARD, L. (2014): «Introduction», en *What is the Text Encoding Initiative? How to add intelligent markup to digital resources* [en línea]. Marseille: Open Edition Press. Disponible en Internet <<http://books.openedition.org/oep/679>>. ISBN: 9782821834606. DOI: <https://doi.org/10.4000/books.oep.679>.
- CALDERÓN CAMPOS, M. (2015): *El español del reino de Granada en sus documentos (1492-1833). Oralidad y escritura*. Bern: Peter Lang (*Fondo Hispánico de Lingüística y Filología*, 22).
- CALDERÓN CAMPOS, M. y R. DÍAZ-BRAVO (2021): «An online corpus for the study of historical dialectology: *Oralia diacrónica del español*», *Digital Scholarship in the Humanities* (1), 1-19.
- CLARIDGE, C. (2008): «Historical Corpora», en A. Lüdeling y M. Kytö (eds.): *Corpus Linguistics: An International Handbook (Vol.1)*, Walter de Gruyter, Berlin/New York, 242-259.
- CULPEPER, J. y M. KYTÖ (2010): *Early Modern English Dialogues: Spoken Interaction as Writing*. Cambridge: Cambridge University Press.
- EVERT, S. & A. HARDIE (2011): «Twenty-first century corpus workbench: Updating a query architecture for the new millennium». In *Proceedings of the Corpus Linguistics 2011 conference*, Birmingham, UK, July. University of Birmingham
- HUBER, M., M. NISEEL, P. MAIWALD & B. WIDLITZKI (2012): *The Old Bailey Corpus. Spoken English in the 18th and 19th centuries*. Disponible en línea: www.uni-giessen.de/oldbaileycorpus.
- JANSSEN, M. (2016): «TEITOK: Text-Faithful Annotated Corpora», en *Proceedings of the 10th Language Resources and Evaluation Conference (LREC 2016) ELRA*. Portoroz, Eslovenia, mayo de 2016, pp. 4037-4043.
- KOHENEN, T. (2007): «From Helsinki through the centuries: the design and development of English diachronic corpora», en Päivi Pahta, Irma Taavitsainen, Terttu Navelainen y Jukka Tyrkkö (eds.), *Studies in Variation, Contacts and Change in English. Volume 2: Towards Multimedia in Corpus Studies*. [28, 06, 2017] <<http://www.helsinki.fi/varieng/series/volumes/O2/kohenen/>>.
- KYTÖ, M. (2011): «Corpora and historical linguistics», *Revista Brasileira de Linguística Aplicada, Belo Horizonte*, 11/2, 417-457.
- LEECH, G. y A. WILSON (1996): *Recommendations for the Morphosyntactic Annotation of Corpora*.

EAGLES Document EAG-TCWG-MAC/R, marzo de 1996. Disponible en <http://www.ilc.cnr.it/EAGLES96/annotate/annotate.html> [Consulta: 26/4/2021]

MORALA, J. R. (2012): «Léxico e inventarios de bienes en los siglos de oro», en G. Clavería, M. Freixas, M. P. y J. Torruella (eds.), *Historia del léxico: perspectivas de investigación*, Madrid/Frankfurt am Main: Iberoamericana/Vervuert, pp. 199-218.

OESTERREICHER, W. (2004). «Textos entre inmediatez y distancia comunicativas. El problema de lo hablado escrito en el Siglo de Oro», en R. Cano Aguilar (coord.), *Historia de la Lengua Española*, cap. 29, pp. 729-769.

RAUMOLIN-BRUNBERG, H. y T. NEVALAINEN (2007): «Historical Sociolinguistics: The *Corpus of Early English Correspondence*», en J. C. Beal, K. Corrigan y H. Moisl (eds.), *Creating and Digitizing Language Corpora: Diachronic Databases*, vol.2, Houndmills: Palgrave, pp. 148-171.

SÁNCHEZ-PRIETO BORJA, P., y D. VÁZQUEZ BALONGA (2017): «Hacia un corpus de beneficencia en Madrid (siglos XVI-XIX)», *Scriptum digital*, 6, pp. 83-103.

VAAMONDE, G. (2018): «La multidisciplinariedad en la creación de corpus históricos. El caso de *Post Scriptum*», *Artnodes*, 22, pp. 118-127.

Modelizando una literatura en el olvido: LDA aplicado a corpus españoles sobre Guinea Ecuatorial y Filipinas

Emilio Vivó Capdevila, Universiteit Antwerpen, Bélgica, emilio.vivocapdevila@uantwerpen.be

Palabras clave: Filipinas en la literatura española; Guinea Ecuatorial en la literatura española; LDA; Topic modelling; Literatura del franquismo

Introducción y objetivos

Esta comunicación se aproximará a una serie de colecciones de textos publicados en España que tienen como eje temático Guinea Ecuatorial y Filipinas. En esta presentación intentaré demostrar dos cosas: 1) que en la representación de Filipinas en los textos publicados después de su independencia se comparten tópicos con el discurso franquista sobre Guinea Ecuatorial; 2) que una serie de tópicos decimonónicos sobre Filipinas desaparecen en el discurso poscolonial sobre este territorio, trasladándose, con el despegue de la colonización en Guinea Ecuatorial durante el franquismo, al discurso colonial sobre Guinea Ecuatorial. Para ello emplearé diferentes versiones de Latent Dirichlet Allocation (LDA) a diferentes corpus. Este análisis sirve de base para, primero, una etiquetado e interpretación de los resultados y, segundo, una comparación “distante” de la producción sobre ambos territorios entre sí y entre dos momentos diferentes, coincidentes con picos y bajadas del número de publicaciones y la intensificación/final de la presencia colonial (Figuras 1 & 2).

Corpus

El corpus principal de este trabajo (el corpus “estrecho”) se compone de un total de 969.656 palabras, subdividido en dos grupos temáticos (uno por país): 490.119 palabras para el corpus estrecho guineano, 479.537 para el corpus estrecho filipino. Estas han sido extraídas de una selección de obras de temáticas, volumen y autorías diversas preprocesadas en texto plano mediante el programa *Transkribus*. Siguiendo a Calvo Tello (2019, pp. 121-122), se trata de un “corpus oportunístico” (esta colección no está digitalizada) pero balanceado, ya que busca representar los diferentes géneros literarios dentro de esta producción:

- Filipinas:
Xavier, Adro. Ocaso del Imperio, 1940 (novela misional); Sanz y Díaz, José. Legazpi, conquistador de Filipinas, 1940 (biografía); Alcalá López, Antonio. Bajo el cielo Filipino, 1943 (novela); Bernia, Juan. Viaje a Nueva Castilla, 1947 (novela de viajes); Pérez de Olaguer, Antonio. Hospital de San Lázaro, 1953 (novela); Llovet, Enrique. Los últimos de Filipinas, 1954 (novela); Piñar, Blas. Filipinas, país hispánico, 1957 (ensayo); Pérez de Olaguer, Antonio. Mi padre, un hombre de bien, 1967 (novela).
- Guinea Ecuatorial:

Miranda, Agustín. Cartas de la Guinea, 1940 (novela híbrida). Bravo Carbonell, Juan. Anecdótico Pamúe, 1942 (novela etnográfica); Soler, Bartolomé. La selva humillada, Bartolomé, 1951 (novela de viajes); Bautista Velarde, Germán. Fang-Eyeyá, 1950 (novela); Masoliver, Liberata. Efún, 1954 (novela); Olangua, Augusto. Una cruz en la selva, 1955 (novela misional). Manfredi, Domingo. Tierra negra, 1957 (novela); Xavier, Adro. España en África: ayer y hoy, 1962 (novela de viaje).

Por otro lado, esta comunicación empleará como corpus de referencia (“corpus ancho”) una serie de colecciones creadas manualmente en la plataforma HuthiTrust, disponibles públicamente y creadas por Emilio Vivo y Rocío Ortuño. Por un lado, dos corpus/*worksets* con una muestra de literatura decimonónica (periodo 1850-1899): [Filipinas colonial](#); y [Decimonónico Guinea 1850-1899 15.4.21](#). Por otro, una serie de colecciones públicas con obras contenidas en la colección de Huthi Trust para el periodo del primer franquismo (1939-1959) pero sin libre acceso a la obra: [Literatura franquista sobre Filipinas](#); y [Franquismo Guinea 40-59 15.4.21](#). En el caso de ambos territorios, estos dos periodos (1850-1899) y (1939-1969) coinciden con picos y bajadas del número de publicaciones y de intensificación/final de la presencia colonial:

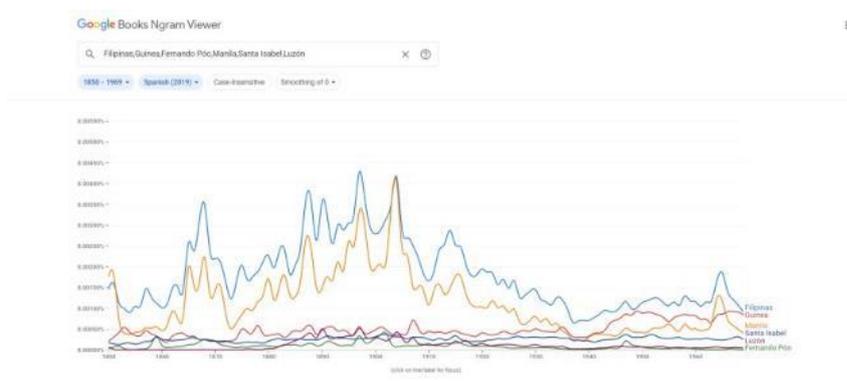


Figura 1 Google Ngrams de Filipinas; Guinea; Fernando Poo, Manila, Santa Isabel y Luzón, Spanish, years 1850-1969

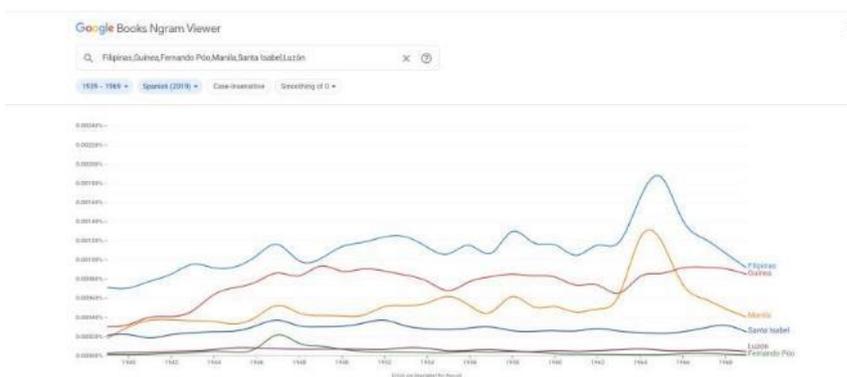


Figura 2 Google Ngrams de Filipinas; Guinea; Fernando Poo, Manila, Santa Isabel y Luzón, Spanish, years 1939-1969

El LDA incorporado a HuthiTrust Analytics permite incorporar grandes colecciones textuales a nuestra pregunta de investigación, mejorando la representatividad de la hipótesis y sirviendo de grupo de control y de comparación para un corpus más estrecho y más controlado.

Justificación bibliográfica

En los últimos años, tanto Filipinas como Guinea Ecuatorial han sido objeto de una explosión de trabajos centrados en la colonización española, especialmente en su momento de mayor intensidad (mediados del XIX y del XX, respectivamente). No obstante, estos o bien han sido llevados a cabo por historiadores (Aranzadi & Álvarez Chillida, 2020; Elizalde Pérez-Grueso, 2009) o, en el caso de los estudios culturales y literarios, se han centrado bien en el siglo XIX español (Guillén Arnaiz, 2021), bien en la producción filipina en español (Donoso Jiménez, 2013). Rocío Ortuño escribía recientemente uno de los únicos trabajos en revisar tanto las relaciones hispanofilipinas entre 1940 y 1960 como la producción literaria franquista sobre la excolonia y planteaba ahí precisamente la necesidad de un estudio cuantitativo sobre la continuidad en el franquismo de imaginarios coloniales del XIX (Ortuño Casanova, 2021). Respecto a Guinea, si bien aún escasos, sí que es posible encontrar diversos trabajos que han analizado tanto la literatura colonial franquista (Arbaiza, 2021; Djangany, 2014) como la literatura de la colonia (Lewis, 2007), pero ninguno de manera cuantitativa. Por otro lado, son muy escasos los trabajos que han yuxtapuesto ambos territorios (Lifshey, 2012). Esta comunicación pretende, así, proponer un primer acercamiento cuantitativo a este corpus y hacerlo de forma comparativa y basada en la idea de tropo. El estudio (cualitativo) de los tropos como elementos repetitivos, colectivos e históricos (Curtius, 1967, p. 89) y sus efectos performativos (Pritsch, 2011, pp. 104-106) tienen una larga historia, pero recientemente diversos autores han teorizado y aplicado LDA, el método estándar de topic modelling (incluido en el paquete MALLET), al análisis textual (NavarroColorado, 2018; Silge & Robinson, 2018; Jockers & Mimno, 2013; Tangherlini & Leonard, 2013; Blei, 2012; Newman et al., 2010, p. 215; Block, 2006). No obstante, nadie ha aplicado todavía este tipo de análisis ni a la literatura del franquismo, ni a literatura española sobre Filipinas y/o Guinea Ecuatorial, pese a que el uso de herramientas computacionales y su aplicación a la literatura mejora la representatividad de nuestras afirmaciones y facilita la recuperación de obras que, como estos corpus, están todavía por explorar (Jockers & Mimno, 2013, p. 751).

Metodología

Por lo que respecta a la selección del corpus “ancho” procedente de Hathi Trust, se han seguido los criterios de “full view”, búsqueda en el catálogo y referencia a Guinea/Filipinas. Los resultados se han filtrado luego manualmente de forma que todos los títulos contuvieran, como mínimo, entre 10 y 15 referencias a Filipinas/Guinea u otras palabras clave relacionadas, como (Fernando) Póo, Luzón,

Manila..., y luego he utilizado el código de Murdock y Allen integrado en Hathi Trust Analytics. (Murdock & Allen, 2015), aplicándolo, sin emplear palabras vacías, a cada una de las colecciones antes mencionadas. El algoritmo ha hecho 200 iteraciones y ha generado modelos de 20 tópicos. Para aplicar el LDA a mi corpus “estrecho” he empleado R Studio, adaptando el código de José Manuel Fradejas (2019, Capítulo 9). Respecto a la práctica común de usar palabras vacías, tras diversas pruebas, he optado por conservar los nombres propios (ya

que permiten orientar la interpretación y el etiquetado sin alterar el corpus) y he adaptado la lista de Fradejas (2019, Capítulo 9) (añadiendo verbos dicendi, formas de tratamiento y verbos comunes). Respecto al número de tópicos, se suele recomendar bien 100 (Doyle & Elkan, 2009, p. 7; Tangherlini & Leonard, 2013, p. 731), bien un número cuanto más amplio mejor (ya que solo un 10% suelen ser significativos) (Wallach et al., 2009, pp. 262-265). He realizado una serie de experimentos con diferentes números de tópicos múltiples de 5: 10, 25, 50 y 100, basándonos en los trabajos de Borja Navarro (Navarro-Colorado, 2018, p. 5). Para esta propuesta me centraré sobre todo en los modelos con 25 tópicos, ya que son más fáciles de visualizar y ya ofrecen resultados interpretables. Respecto al número de palabras por tópico, la literatura suele coincidir en el uso de más de 5 (Block, 2006; Navarro-Colorado, 2018; Silge, 2018). Nosotros hemos optado por 10 (top_n = 10). Además, he decidido dividir el corpus estrecho manualmente en dos sub-corpus, uno para Filipinas y otro para Guinea, en línea con Tangherlini et al. (2013, p. 732), confiando en la fase de curación cualitativa. El etiquetado e interpretación de los resultados se ha efectuado, finalmente, de modo manual y siguiendo criterios propios (Tangherlini & Leonard, 2013, p. 728).

Resultados

Los modelos de HTRC Analytics muestran un predominio en la producción decimonónica sobre Filipinas de la ficción sentimental, los textos sobre misioneros, y en general los textos sobre los españoles (colonos) en el archipiélago. Además, encontramos una preocupación por la administración colonial, la explotación agrícola (del tabaco) y la geografía y la historia natural. Temas como Cuba o la colonización del sur del archipiélago aparecen como particularismos relevantes del corpus. Mientras tanto, los modelos de HTRC Analytics de la literatura franquista sobre Filipinas mostrarían un predominio de la historia colonial y los héroes de la leyenda imperial, ofreciendo una visión positiva de la colonización, vinculándola a la ciencia y al progreso. También la historia misional aparece en repetidas ocasiones. Destacan además la figura del médico y del escritor Antonio Pérez de Olaguer, la preocupación por la masonería y su historia en Cuba y Filipinas (es decir, el 98) o los Austrias mayores, que encajan dentro de una idea de hispanidad franquista, esencialista y católica.

Respecto a Guinea, los modelos de HTRC Analytics destacarían en el XIX la evangelización, la plantación (también de tabaco), la administración colonial y la enseñanza de los "indígenas", de forma paralela al XIX en Filipinas, aunque aparecen las dificultades de los misioneros frente a la escasez de recursos y la presencia protestante: se destacan las deportaciones, y las dudas sobre el futuro de la colonia, y la imposibilidad de hacer valer los "derechos" de España. También destaca la preocupación por la gramática nativa, como forma de generar conocimiento sobre el colonizado. El continente parece solo como lugar de viaje y de expedición, y parece que el corpus se centra en Fernando Póo. Los modelos de HTRC Analytics de la literatura franquista sobre Guinea sugieren el avance de la colonización, y el conocimiento sobre la colonia con publicaciones sobre historia

natural, geografía y etnografía de la colonia , pero también la persistencia del tema misional; se ensalzan las expediciones militares y se tematiza la economía de explotación colonial en sus diferentes facetas. Aquí también destacan las publicaciones sobre las relaciones entre españoles y nativos en la colonia, especialmente centrándose en la historia misional y su impacto (positivo).



Figura 3 corpus estrecho Filipinas K = 25 , top_n = 10, con nombres propios

Cuadro 1 – Tópicos por tema – Corpus estrecho – Filipinas

- o Hospital: 1, 3, 10, 15, 19. o Militares: 2, 8, 11, 18, 22. o Biografía de los Pérez: 4, 5, 13, 16, 21. o Historia Colonial: 6, 9, 23, 23, 25.
- o Paralelismos Medicina-Religión-Colonización, sin quedar clara la diferencia entre padre misionero y padre progenitor 7, 13.
- o Comunidad española: 14, 15, 17, 20.

En el corpus “estrecho” sobre Filipinas destacan: la historia colonial (la “conquista” por Legazpi, las campañas de Mindanao y Joló de finales del XIX, y la guerra de independencia filipina, y, sobre todo, el episodio del Sitio de Baler); el ejército (asociado a la historia colonial); la masculinidad (el protagonismo masculino); la labor

misional (también en Mindanao/Davao) y médica (en la curación de la lepra), que el modelo, significativamente, confunde; y, finalmente, las memorias y recuerdos de los españoles que vivían en Filipinas entre la guerra de independencia filipina y la segunda guerra mundial. Esto es coherente con los resultados de HTRC Analytics y permitiría argumentar la continuidad del interés por la evangelización de Filipinas, el desinterés por temas administrativos y políticos (o incluso económicos), la centralidad de la vida “española” en el archipiélago (bien en forma de memorias o testimonios, bien en forma de novelas románticas) y, ante todo, la irrupción de las historias militares (en torno al sitio de Baler) en el imaginario español.

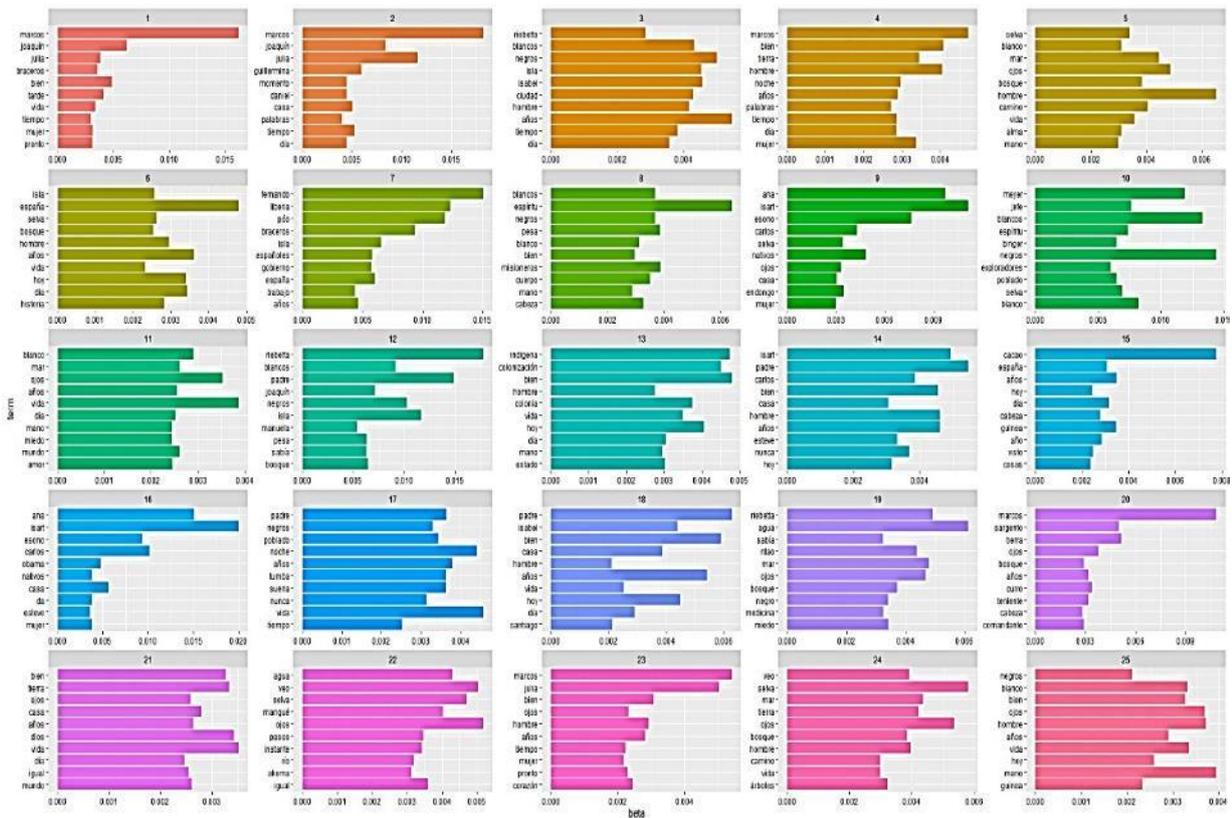


Figura 4 corpus estrecho Guinea k = 25 , top_n = 10, con nombres propios

Cuadro 2 – Tópicos por tema – Corpus estrecho – Guinea:

- o vida en la plantación: 1, 4, 9, 14, 16, 20, 23 o economía de plantación (braceros para el cacao): 1, 7, 15.
- o novela sentimental: 1, 2, 9, 14, 16, 23 o masculinidad colonial : 1, 3, 9, 23 o Interacciones colonizador-colonizado: 1, 3, 5, 8, 9, 10, 11, 12.,
- 13, 15, 16, 17, 21, 22, 24, 25 o Desarrollo de la colonización: 3, 6, 13, 14, 15, 20, 21, 25.
- o misioneros: 3, 8, 17, 18, 21 o naturaleza colonial (homogeneidad, miedo): 5, 6, 9, 10, 11, 12,
- 19, 20, 21, 22, 24 o cuerpo 9 (cabeza, mano...): 8, 15, 17 (baile y sensualidad), 20,
- 25 o 10 viaje (a la selva) : 10, 19, 22.

Por lo que respecta a los resultados del corpus "estrecho" guineano, el LDA muestra una preocupación por los españoles en la colonia (un paralelo con la literatura franquista sobre Filipinas), especialmente los misioneros pero también colonos y finqueros protagonistas de la novela romántica que (de forma paralela al XIX Filipino) aparece durante el franquismo. Hay un interés por la masculinidad, pero las mujeres, españolas e incluso nativas, son mucho más relevantes. Además, sí que aparecen la administración colonial ya efectiva o la ya rentable economía de plantación y deforestación, (de forma paralela al XIX Filipino). Se sugieren interacciones colonizador-colonizado ubicuas y multinivel, ahondando en el impacto (positivo) de la colonización en la vida y

organizaciones guineanas y el contraste entre el blanco y el negro. La mirada y los cuerpos también aparecen constantemente, reforzando las diferencias colonizadores-narradores y colonizados-narrados. El paso del tiempo es una referencia constante, pero no como memoria, sino como comparación: el antes serviría para compararlo con un presente de la colonización más productivo y más benigno para el nativo (haciendo con el presente en Guinea lo mismo que la historiografía franquista sobre Filipinas). La producción de conocimiento sobre la colonia (especialmente su naturaleza) se refleja aquí en una obsesión con la selva y una naturaleza amenazante y homogénea, a la que se integra el nativo animalizado y primitivo. Los militares tienen escasa presencia (son sobre todo expedicionarios), y el protagonismo lo tienen, como en el XIX filipino, los misioneros.

En resumen en esta comunicación pretendo mostrar que el LDA tanto del corpus “amplio” de Huthi Trust como del corpus “estrecho” mostraría que la producción cultural española sobre Filipinas y Guinea Ecuatorial se centra en la colonización de la espada y la cruz, militar y misional. Esta coexistiría con un interés por la explotación económica y los problemas particulares de la colonización. Además, esta se centra justamente en la producción sobre España y los españoles y, podríamos sugerir, sigue un principio nacionalista y dependiente de la posibilidad de una colonización efectiva, rentable y étnicoreligiosa, de la que dependerían la frecuencia, los temas e incluso los géneros de los textos producidos.

Referencias

- Aranzadi, J. R., & Álvarez Chillida, G. (2020). *Guinea Ecuatorial (des)conocida. (Lo que sabemos, ignoramos, inventamos y deformamos acerca de su pasado y su presente)*.
- Arbaiza, D. (2021). 'The Last Mission: Religion and Colonialism in the Spanish Cultural Production during Francoism'. *Revista Canadiense de Estudios Hispánicos*, 44(2).
- Blei, D. M. (2012). Topic modeling and digital humanities. *Journal of Digital Humanities*, 2(1), 8-11.
- Block, S. (2006). Doing More with Digitization: An introduction to topic modeling of early American sources. *Common-Place*, 6(2).
- Calvo Tello, J. (2019). Diseño de corpus literario para análisis cuantitativos. *Revista de Humanidades Digitales*, 4, 115-135.
- Curtius, E. R. (1967). *Europäische Literatur und lateinisches Mittelalter*. Francke.
- Djangany, J. S. (2014). Hilvanando derechos en la literatura colonial sobre Guinea Ecuatorial. *Debats*, 123, 12-27.
- Donoso Jiménez, I. (2013). *Historia cultural de la lengua española en Filipinas: Ayer y hoy*. Editorial Verbum.
- Doyle, G., & Elkan, C. (2009). *Accounting for burstiness in topic models*. 281-288.
- Elizalde Pérez-Gruoso, M. D. (2009). Sentido y rentabilidad. Filipinas en el marco del Imperio Español. *Repensar Filipinas. Política, identidad y religión en la construcción de la nación filipina*, 45-78.
- Finkel, J. R., Grenager, T., & Manning, C. (2005). Incorporating Non-local Information into Information Extraction Systems by Gibbs Sampling. *Proceedings of the 43rd Annual Meeting of the Association for Computational Linguistics*, 363-370. <http://nlp.stanford.edu/~manning/papers/gibbscrf3.pdf>

- Fradejas Rueda, J. M. (2019). *Cuentapalabras*. <http://www.aic.uva.es/cuentapalabras/>
- Guillén Arnaiz, C. (2021). *Reinventando las Filipinas: Imágenes y representaciones literarias de las islas a través de textos españoles de finales del siglo XIX (1872-1896)*. University of Antwerp.
- Jockers, M. L., & Mimno, D. (2013). Significant themes in 19th-century literature. *Poetics*, 41(6), 750-769.
- Lewis, M. A. (2007). *An Introduction to the Literature of Equatorial Guinea: Between Colonialism and Dictatorship*.
- Lifshey, A. (2012). *The Magellan fallacy: Globalization and the emergence of Asian and African literature in Spanish*. University of Michigan Press.
- Murdock, J., & Allen, C. (2015). Visualization Techniques for Topic Model Checking. *AAAI Conference on Artificial Intelligence 2015*. <https://www.hypershelf.org/>
- Navarro-Colorado, B. (2018). On poetic topic modeling: Extracting themes and motifs from a corpus of spanish poetry. *Frontiers in Digital Humanities*, 5, 15.
- Newman, D., Noh, Y., Talley, E., Karimi, S., & Baldwin, T. (2010). Evaluating topic models for digital libraries. *Proceedings of the 10th annual joint conference on Digital libraries*, 215-224. <https://doi.org/10.1145/1816123.1816156>
- Ortuño Casanova, R. (2021). Nostalgia del imperio: Literatura filipina y franquismo. En *Poéticas y cánones literarios bajo el Franquismo* (Diego Santos y Fernando Larraz, pp. 81-108). Vervuert.
- Pritsch, S. (2011). Rhetorik als Selbsttechnologie: Postmoderne Figuren des Subjekts zwischen ›Sagen‹ und ›Tun‹. *Rhetorik*, 30(2011), 96-108. <https://doi.org/10.1515/9783110236385.96>
- Silge, J. (2018). Training, evaluating, and interpreting topic models. URL <https://juliasilge.com/blog/evaluating-stm/> (accessed 12.13. 18).
- Silge, J., & Robinson, D. (2018). Topic modeling. *Text Mining with R: A Tidy Approach*.
- Tangherlini, T. R., & Leonard, P. (2013). Trawling in the Sea of the Great Unread: Sub-corpus topic modeling and Humanities research. *Poetics*, 41(6), 725-749.
- Wallach, H. M., Mimno, D., & McCallum, A. (2009). Rethinking LDA: Why priors matter. *Proceedings of the 22nd International Conference on Neural Information Processing Systems*, 19731981.

Proyecto HDA Teatro UNIR, realidades y perspectivas para el estudio del teatro desde las HD

Teresa Santa María y Monika Dabrowska

(Universidad Internacional de La Rioja-UNIR)

En este panel se quiere presentar la labor desarrollada dentro del proyecto HDATEATROUNIR en cuatro vertientes diferentes: descripción de la biblioteca BETTE, alojada también en DraCor, muestra del análisis de redes a partir de textos etiquetados, posibilidades didácticas de los textos dramáticos en TEI y los nuevos textos digitalizados en TEI. Con todas ellas se pretende describir la investigación llevada hasta la fecha, problemas y metodologías HD aplicadas en el estudio del teatro de la Edad de Plata y presentar las dos líneas nuevas hacia las que se encamina el grupo.

El Grupo de Humanidades y Edición Digital (GHEDI) de la Universidad Internacional de La Rioja (UNIR) comenzó en 2015 un proyecto de edición digital colaborativa en español. Este proyecto se ve hoy continuado por el grupo HDATEATROUNIR, centrando su atención en la creación de una Biblioteca Electrónica Textual de Teatro Español (1868-1939), BETTE, cuya finalidad es editar y publicar en lenguaje de marcado XML-TEI y analizar mediante métodos digitales las piezas dramáticas más relevantes de ese periodo (Jiménez y Calvo, 2020). El corpus total incluye obras de Clarín, Dicenta, Galdós, Lorca, Muñoz Seca, Unamuno y Valle-Inclán.

AUTORES	OBRAS
GARCÍA LORCA	Doña Rosita Yerma La zapatera prodigiosa Bodas de sangre La casa de Bernarda Alba
VALLE-INCLÁN	Divinas palabras Luces de bohemia Águila de Blasón Romance de Lobos Cara de Plata
MUÑOZ SECA	Los pergaminos El refugio La conferencia de Algeciras Usted es Ortiz
VALERA	Asclepigenia La venganza de Atahualpa
UNAMUNO	El otro Fedra
ECHEGARAY	Mancha A fuerza de arrastrarse
CLARÍN	Teresa
DICENTA	Juan José
GALDÓS	Electra Casandra Doña Perfecta

Figura 1. Especificación de obras en XML-TEI y sus autores, editadas por BETTE

Fuente: (Martínez y Santa María, 2019)

Estas obras se encuentran distribuidas de la siguiente manera:

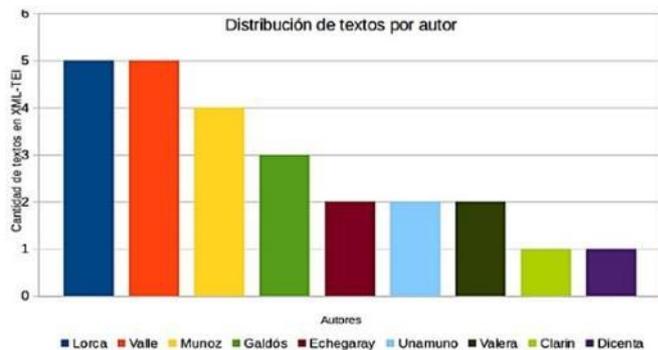


Figura 2: Distribución de textos teatrales según autores. Fuente: (Martínez y Santa María, 2019)

Las veinticinco obras procesadas están libres de derechos de autor decantándose por la variedad de temas, subgéneros teatrales o estilos que presenten un panorama completo de la dramaturgia española y sus cambios durante los casi setenta años que abarca la llamada Edad de Plata del teatro español. Este corpus de teatro ha sido sometido a una revisión filológica para su digitalización en TEI con miras a dar homogeneidad a todos los textos según unas directrices comunes.

Los dramaturgos seleccionados van desde autores del Realismo y Naturalismo, como Galdós, Valera y Clarín, hasta dramaturgos que renovarían la escena española a principios del siglo XX, como Valle-Inclán y Unamuno; pasando por autores que cultivaron más la comedia como es el caso de Muñoz Seca o aquellos que forman parte del llamado teatro de la Restauración y decimonónico como Echegaray y Dicenta (Arana, Santa María y Simón, 2015).

Al mismo tiempo, esas veinticinco obras digitalizadas se encuentran accesibles en el repositorio de GitHub del grupo (<https://github.com/GHEDI/BETTE>). Además, las veinticinco piezas aparecen bajo el nombre de "Spanish Drama Corpus" (<https://dracor.org/span>) en el Drama Corpora Project (DraCor). Este proyecto está liderado por el profesor Frank Fischer, que al mismo tiempo es director de DARIAH (Digital Research Infrastructure for the Arts and Humanities), uno de los proyectos de HD más importante a nivel europeo.

DraCor es un proyecto del Centre for Digital Humanities at HSE (Moscow) y de la Universität Potsdam (Alemania) y alberga otras colecciones de teatro en ruso, latín, griego antiguo, sueco, alemán, italiano, tártaro y alsaciano y de dos autores específicos: Shakespeare y Calderón de la Barca. Además de proporcionar los textos en XML-TEI, DraCor enlaza las obras y sus autores con Wikidata. Igualmente, esta plataforma de investigación digital sobre teatro europeo proporciona, como se ha mencionado, una serie de corpus de teatro etiquetado en TEI, pero, al mismo tiempo, permite que la investigación sea reproducible. En la actualidad contienen más de mil obras de teatro. La importancia de formar parte de este proyecto se justifica, ante todo, por las sinergias que brinda DraCor, es decir, la posibilidad de que diferentes investigadores, haciendo estudios similares, puedan llegar más lejos trabajando en equipo, de forma colectiva y en colaboración, tan necesaria en las HD. Así mismo, gracias a DraCor el proyecto BETTE gana en visibilidad y difusión a nivel internacional posibilitando la

comparación de datos entre corpus teatrales de distintos países. Del mismo modo, para la enseñanza en HD, DraCor aporta una herramienta interesante, novedosa e incluso divertida con la que se puede trabajar fácilmente (Fischer, Santa María, Calvo-Tello y Milling).

De esta manera, BETTE resulta un conjunto de textos y datos abierto a la comunidad investigadora con numerosas posibilidades de análisis comparativo de obras con las que corroborar o rechazar hipótesis.

Aplicaciones didácticas de las Humanidades Digitales en el área estudios teatrales.

Ante digitalización del mundo y de los estudios humanísticos cabe reflexionar sobre los nuevos modelos de enseñanza y los retos educativos que se plantean en la era del silicio. A saber: ¿Cómo formar a los futuros humanistas digitales y aplicar los medios computacionales en educación a filólogos, maestros, espectadores teatrales, actores y profesionales de escena, así como a estudiantes en general? El objetivo de nuestra intervención es presentar algunos ejemplos de los posibles usos didácticos de las herramientas de Humanidades Digitales e indagar sobre la importancia de los nuevos enfoques didácticos en las disciplinas humanísticas y la innovación educativa digital.

Con este estudio se pretende mostrar, en primer lugar, el horizonte actual de la vertiente educativa de los proyectos de Humanidades Digitales, es decir programas educativos y aplicaciones interactivas, como Simulated Environment for Theatre SET (Roberts-Smith 2013), The Magic Mirror Theater (Bozia, 2018), The New Variorum Shakespeare Digital Challenge, The Shakespeare Standard, o el software Luminary Digital Media para enseñar a Shakespeare en el siglo XXI (Revenga, 2014). En este contexto hablaremos de los posibles usos didácticos del proyecto BETTE/HDAUNIR que actualmente se está desarrollando en la Universidad Internacional de La Rioja (UNIR).

Los estudios de caso realizados por proyectos interdisciplinares en HD están explorando nuevos campos de estudio, en los que los estudiantes pueden desarrollarse dentro y fuera del aula. Asimismo, estudiantes universitarios involucrados en la aplicación de herramientas interactivas trabajan de forma cooperativa con el profesorado en proyectos de investigación expansivos, que ofrecen oportunidades hasta ahora inalcanzables en entornos de aprendizaje virtuales (Hannan, y Chatterjee, 2013; Faull y Jakacki, 2015).

El avance de las herramientas tecnológicas al servicio de las Humanidades Digitales, y de estas al servicio de la innovación educativa, ha sido enorme durante la última década. Las oportunidades que ofrece la realidad aumentada y entornos 3D, los contextos de inmersión y simulación facilitan el diseño de nuevas experiencias educativas. El uso de metodologías como Object Based Learning (OBL), Objects Interactive Learning, Virtual Reality Simulation-Based Learning, Open Educational Resources (OER) u OER Enabled Pedagogy permite potenciar y enriquecer los aprendizajes y responder a las nuevas exigencias educativas y pedagógicas de la sociedad digital.

La mejora de las aplicaciones digitales en áreas como la Antigüedad clásica, la arqueología, la lingüística y la literatura ha puesto de manifiesto que los entornos de aprendizaje en realidad virtual brindan acceso a modelos en 3D de los artefactos teatrales, así como a réplicas virtuales de lugares antiguos, resultando especialmente beneficiosos para los estudiantes (Bozia, 2018) y la percepción del espacio teatral, la actuación y la recreación del ambiente teatral. Importantes cuerpos académicos de educación superior, por ejemplo, los de la Universidad de

Pensilvania, la Universidad de Stanford, la Universidad Macquarie o el Instituto de Educación de Hong Kong (Barlow, 2017), están centrando sus esfuerzos en demostrar cómo los docentes pueden enseñar conocimientos específicos de distintas materias transfiriendo habilidades entre diferentes disciplinas y niveles académicos y cómo será posible transformar las prácticas educativas implementando las metodologías y aplicaciones didácticas digitales mencionadas más arriba, y que constituyen el objeto de este estudio.

Los ejemplos analizados dan muestra de las nuevas tendencias en los actuales entornos de aprendizaje y resultan relevantes para abrir nuevos horizontes en innovación educativa. Los resultados obtenidos, a su vez, redundarán en la investigación y análisis comparativo de numerosas aplicaciones didácticas, tanto en el ámbito de la lingüística y la literatura, como en otras disciplinas humanísticas.

Y la mujer irrumpe en BETTE. Dramaturgas españolas incorporadas en el corpus

Tras estos primeros años, BETTE cuenta con textos en abierto dentro de los repositorios mencionados y otras siete obras en TEI. Estas últimas no pueden ponerse en GitHub ni pasar al repositorio de DraCor porque no están aún en dominio público los textos de los autores. En concreto, contamos con las ediciones de: *Espejo de avaricia* de Max Aub, *Rosas de otoño* de Jacinto Benavente, *Tres sombreros de copa* de Miguel Mihura, *Usted tiene ojos de mujer fatal*, *Cuatro corazones con freno y marcha atrás* y *Un marido de ida y vuelta* de Enrique Jardiel Poncela y *La esfinge* de Unamuno. Este hecho – la necesidad de poder contribuir a la estructura de textos para el estudio y trabajo de otros investigadores (Fischer et al. 2019) -nos ha llevado a plantearnos como objetivos, para la tercera fase de edición de obras dramáticas de la Edad de Plata, completar, por un lado, el listado de obras de algunos dramaturgos ya incluidos en BETTE y Spanish Drama Corpus, así como otorgar mayor visibilidad a dramaturgas de calidad dentro de estos corpus. De este modo, sus obras dramáticas podrán consultarse y trabajar con ediciones en TEI, así como contar con las ventajas de su alojamiento dentro de DraCor.

Para incorporar nuevas obras y creadoras al corpus BETTE (1868-1939) que no tengan ya vigentes sus derechos de autoría, se ha partido del listado de autores que han pasado a dominio público, porque fallecieron antes de 1938 y que pone a disposición la Biblioteca Nacional de España. Se revisaron todas las entradas y quedó el siguiente listado, donde aparecen solo doce dramaturgas:

Acuña, Rosario (1850-1923)	Enlace al catálogo	Autor en BDH	Autor en Datos.BNE.es
Calé y Torres de Quintero, Emilia (1837-1908)	Enlace al catálogo	Autor en BDH	Autor en Datos.BNE.es
Contreras de Rodríguez, María del Pilar (1861-1930)	Enlace al catálogo	Autor en BDH	Autor en Datos.BNE.es
Coronado, Carolina (1820-1911)	Enlace al catálogo	Autor en BDH	Autor en Datos.BNE.es

Díaz de Mendoza y Aguado, Carmen, Condesa de San Luis (1864-1929)	Enlace al catálogo	Autor en BDH	Autor en Datos.BNE.es
Freixas, Narcisa (1859-1926)	Enlace al catálogo	Autor en BDH	Autor en Datos.BNE.es
García Balmaseda de González, Joaquina (1837-1911)	Enlace al catálogo	Autor en BDH	Autor en Datos.BNE.es
Monserdà de Macià, Dolors 1845-1919	Enlace al catálogo	Autor en BDH	Autor en Datos.BNE.es
Muñiz y Más, Adelaida m. 1906	Enlace al catálogo	Autor en BDH	Autor en Datos.BNE.es
Pardo Bazán, Emilia, Condesa (1851-1921)	Enlace al catálogo	Autor en BDH	Autor en Datos.BNE.es
Sáenz de Tejada, Victorina (1841-1909)	Enlace al catálogo	Autor en BDH	Autor en Datos.BNE.es
Velilla y Rodríguez, Mercedes (1852-1918)	Enlace al catálogo	Autor en BDH	Autor en Datos.BNE.es

Figura 5. Elaboración propia, a partir del catálogo de la BNE:

<https://datos.gob.es/es/catalogo/e00123904><https://datos.gob.es/es/catalogo/e00123904-autores-espanoles-en-dominio-publico-en-2019-fallecidos-en-1938>[autores-espanoles-en-dominio-publico-en-2019-fallecidos-en-1938](https://datos.gob.es/es/catalogo/e00123904-autores-espanoles-en-dominio-publico-en-2019-fallecidos-en-1938)

De entre todas ellas, se han elegido a dos autoras, Emilia Pardo Bazán y Rosario Acuña, por dos razones. En primer lugar, por cumplirse el centenario de ambas próximamente, lo que ayudaría a su mayor divulgación y difusión, aprovechando el eco que recibirían durante 2021 y 2023, centenario del fallecimiento de una y nacimiento de otra, respectivamente. En segundo, lugar, porque podremos contrastar la influencia en el futuro de dos autoras con un reconocimiento en los manuales de literatura y de teatro desigual. La primera de ellas sí aparece en uno de ellos (Pedraza Jiménez y Rodríguez Cáceres, 1983) como dramaturga (pp. 783-784), mientras que Rosario Acuña no aparece en ninguno de los volúmenes de dicho manual ni en los dos más específicos sobre teatro de Ruiz Ramón (1986 y 1996). Incidimos, de este modo, en las posibilidades que ofrecen las HD de dar visibilidad y alcance a autores no tan valorados hasta la fecha, como se constata en proyectos *Mnemosine*, vinculado por el periodo de la Edad de Plata al que se circunscribe; y, dentro del ámbito de la mujer, *Bieses*

(<https://www.bieses.net/>) y la iniciativa en Wikipedia de *Editatón* (Editatón. <https://es.wikipedia.org/wiki/Editat%C3%B3n>), por poner solo algunos ejemplos.

Las obras que tenemos o tendremos en breve en BETTE son las siguientes y cubren la dramaturgia conocida de dichas creadoras:

Rosario Acuña	<i>Rienzi el tribuno</i>
Rosario Acuña	<i>Amor a la patria</i>
Rosario Acuña	<i>El padre Juan</i>
Rosario Acuña	<i>Tribunales de venganza</i>
Emilia Pardo Bazán	<i>El becerro de metal</i>
Emilia Pardo Bazán	<i>El vestido de boda. Monólogo</i>
Emilia Pardo Bazán	<i>Verdad. Drama en cuatro actos, en prosa</i>
Emilia Pardo Bazán	<i>La suerte</i>
Emilia Pardo Bazán	<i>Cuesta abajo</i>
Emilia Pardo Bazán	<i>Juventud</i>
Emilia Pardo Bazán	<i>Las raíces</i>

Figura 6: Listado de obras para incluir en BETE (Fase 2020-2022)

Creemos que disponer de estas piezas puede contribuir a que resulten más conocidas – aunque se puedan consultar ya en PDF desde la Biblioteca Digital Hispánica – y a reclamar mayor atención por parte de los investigadores, al facilitarlas en el estándar TEI. Por otro lado, compensamos una deuda histórica con las creadoras, aún minoría o prácticamente invisibles, en muchas infraestructuras de las HD. También facilitamos su uso didáctico a las nuevas generaciones, a través de los grafos, distribuciones de personajes y de texto, etc. que podemos manejar, cuando dichas obras se incorporen al “Spanish Drama Corpus” de DraCor (<https://www.dracor.org/>).

- Bibliografía (en orden de presentación y comunicaciones)

Arana, R.; Santa María, T. y Simón, M. (2015). El teatro de la España contemporánea (1868-1936). *Texto Digital*, 11(2) (julio-diciembre), 111-127 <http://dx.doi.org/10.5007/1807-9288.2015v11n2p111>

Fischer, F., Santa María Fernández, T., Calvo-Tello, J. y Milling, C. Una perspectiva comparativa del teatro español: integración de la "Biblioteca Electrónica Textual del Teatro en Español" en la plataforma “DraCor” en *Actas del Congreso de Humanidades Digitales Hispánicas*. Universidad de Castilla-La Mancha. (en prensa).

Jiménez, C.M.; Santa María, T. y Calvo, J. (2017). *Biblioteca Electrónica Textual del Teatro en Español (BETTE)*, La Rioja: UNIR <https://bit.ly/2SGtwhB>

Jiménez Fernández, C. y Calvo Tello, J. (2020). Grafos de escenas y estudios literarios digitales: una propuesta computacional crítica. *452°F*, 23, 78-101 DOI: 10.1344/452f.2020.23.4

- Martínez Carro, E. y Santa María, T. (2019). Biblioteca electrónica textual del teatro español (1868-1936) e investigación con grafos. *Revista de Humanidades Digitales* 3, 23-45.
<http://revistas.uned.es/index.php/RHD/article/view/23144/>
- Bellido, P. (1983). La funcionalidad del personaje y la motivación de una trilogía (las *Comedias Bárbaras* de Valle-Inclán). En M. A. Garrido (Coord.) *Crítica semiológica de textos literarios hispánicos: Actas del Congreso Internacional sobre Semiótica e Hispanismo*. Vol. II, 843-850.
 Consejo Superior de Investigaciones Científicas.
- Calvo, J., Jiménez, C y Santa María, T. (2017). *Biblioteca Electrónica Textual del Teatro en Español (BETTE)*. Universidad Internacional de La Rioja. <https://github.com/GHEDI/BETTE>
- Couceiro, E. (1998). Troncalidad, fuero y sociedad estamental gallega en las *Comedias Bárbaras* de Valle-Inclán. Contextualización antropológica del texto literario. *Revista de Antropología Social*, 7, 121-147.
- Gagnon, A. (2015) *Lectura feminista de las "Comedias bárbaras" de Ramón del Valle-Inclán: Lilith y las mujeres en la decadencia del patriarcado*. Tesis, no publicada. Universidad de Montreal. Recuperado de:
<http://hdl.handle.net/1866/13703>
- Landín, P. (2015). *La expresión de lo trágico en el "Ciclo mítico" de Valle-Inclán. La presencia de Shakespeare y paralelismos con las tragedias de García Lorca*. Tesis, no publicada. Universidad Nacional de Educación a Distancia.
- Santos Zas, M. y Vilchez Ruiz, C. (2021). Todo Valle-Inclán en red. El archivo digital del GIVIUS. *Signa. Revista de la Asociación Española de Semiótica*. 30, 139-162.
- Vilchez Ruiz, C. E., Santos Zas, M., Mascato Rey, R., Serrano Alonso, J., Abalo Gómez, A. y Martínez Rodríguez, F. (2018). *Archivo Digital Valle-Inclán*. Universidad de Santiago de Compostela.
<https://www.archivodigitalvalleinclan.es/publica/principal.htm>
- Barlow, A. (2017). Beyond Object Lessons: Object-Based Learning in the Academic Library. *The Experiential Library*. Rhode Island College, Providence, RI, United States.
- Bozia, E. (2018). Reviving Classical Drama: virtual reality and experiential learning in a traditional classroom. *Digital Humanities Quarterly*. University of Florida.
- Faull, K. M. y Jakacki, D. K. (2015). Digital Learning in an Undergraduate Context: Promoting Long-Term Student-Faculty-Place-Based Collaboration. *Digital Scholarship in the Humanities*, 30 (1).
- Hannan, L., Chatterjee, H., y Duhs, R. (2013). Object Based Learning: A Powerful Pedagogy for Higher Education. En A. Boddington, J. Boys, y C. Speight (Eds.), *Museums and Higher Education Working Together: Challenges and Opportunities* (pp. 159-168). Ashgate Publishing.
- Revenga, N. (2014). La edición crítica digital de textos teatrales. La estrella de Sevilla como ejemplo. *Anuario Lope de Vega. Texto, literatura, cultura* 20, 99-121.
- Roberts-Smith et al. (2013). Visualizing Theatrical Text: From Watching the Script to the Simulated Environment for Theatre (SET). *Digital Humanities Quarterly*, 7(3), 1-17.
- Pedraza Jiménez, F. B. y Rodríguez Cáceres, M. (1983). *Manual de literatura española. 7: Época del realismo*. Cénlit.

Fischer, F., Börner, I., Göbel, M., Hecht, A., Kittel, C., Milling, C. y Trilcke, P. (2019).

Programmable Corpora: Introducing DraCor, an Infrastructure for the Research on European Drama. In *Proceedings of DH2019: "Complexities"*, Utrecht University, doi:10.5281/zenodo.4284002.

Mnemosine. <http://repositorios.fdi.ucm.es/mnemosine/queesmemosine.php>

Ruiz Ramón, F. (1986). *Historia del teatro español. Siglo XX*. Cátedra.

Ruiz Ramón, F. (1996). *Historia del teatro español (desde sus orígenes hasta 1900)*. Cátedra.

Cervantes desde la estilometría y el análisis de redes: estudio de su obra

Laura Hernández-Lorenzo

Universidad de Sevilla

Palabras clave: Estilometría, análisis de redes, Cervantes, Estilística computacional, Género literario 1.

Introducción

Desde los orígenes de la Estilometría han predominado sus aplicaciones a problemas de atribución de autoría. Sin embargo, en los últimos años están aumentando los trabajos que exploran mediante métodos estilométricos otras cuestiones estilísticas, como las corrientes literarias o artísticas (Burrows, 2003), la influencia del sexo del autor (Rybicki, 2016) o el género literario (Jannidis & Lauer, 2014). Los buenos resultados obtenidos muestran el potencial de los métodos y medidas estilométricas para el análisis de corpus textuales desde estas perspectivas y suponen un aliciente para el estudio del estilo literario en general desde el punto de vista estilométrico.

En esta línea, se presenta un análisis mediante técnicas estilométricas y de análisis de redes de las obras completas de un único autor, Miguel de Cervantes (1547-1616). El objetivo final es explorar las relaciones entre sus obras, determinar qué papel juegan la periodización de estas, así como los diferentes géneros literarios que abarcan, y en última instancia, descubrir qué puede aportar esta perspectiva de estudio al conocimiento del autor español más reconocido y estudiado de todos los tiempos.

2. La Estilometría más allá de la autoría y sus posibles aplicaciones a la obra de Cervantes

La Estilometría más allá de la autoría tiene un importante precedente en los estudios de Lutostawski sobre la cronología de las obras de Platón (Lutostawski, 1897). Posteriormente, John Burrows, el creador de la popular medida Delta (Burrows, 2002), planteó cómo estos métodos podían ser útiles no solo para atribución de autoría, sino también para avanzar en otras cuestiones de la Historia de la Literatura (Burrows, 1987), realizando él mismo análisis estilométricos con Delta de distintos géneros y épocas literarias (Burrows, 2003, 2004). Siguiendo a Burrows, Jannidis y Lauer exploran estas cuestiones mediante la aplicación de Delta a un corpus de literatura alemana con el objetivo de evaluar si puede detectar similitudes genéricas y de épocas, con alentadores resultados (Jannidis & Lauer, 2014). Del mismo modo, Hoover extiende estos experimentos a otros narradores de la literatura inglesa (Hoover, 2014), y Rybicki aplica estas técnicas para detectar rasgos distintivos en la escritura de hombres y de mujeres (Rybicki, 2016). En los últimos años, destacan los trabajos llevados a cabo por los investigadores del proyecto CLiGS, con estudios sobre el género literario en el teatro clásico francés o la narrativa en lengua española de finales del XIX y principios del XX a través de herramientas estilométricas (Calvo Tello et al., 2017; Henny-Krahmer, 2018; Schöch, 2018). Todos estos trabajos coinciden en que, si bien la señal autorial es la predominante, cuando esta se neutraliza en la medida de lo posible en el corpus, es posible apreciar otras relaciones estilísticas y de realizar aportaciones de interés sobre obras y autores a través del análisis cuantitativo de estas.

A pesar del potencial que presenta esta línea de investigación, en el caso de la literatura española únicamente contamos con los trabajos de José Calvo para la detección y evaluación de géneros literarios con un corpus de novela española de la Edad de Plata (Calvo Tello, 2019, 2017). Con el objetivo de aislar estas señales y neutralizar la señal autorial, este investigador aplica las técnicas estilométricas sobre el corpus de un único autor

por análisis. Por ello, en este estudio nos centramos en la obra del que es, sin duda alguna, el autor más emblemático de la Literatura española, Miguel de Cervantes. La importancia y repercusión de su obra, así como su singularidad han motivado su elección como caso de estudio mediante esta metodología: ¿puede la Estilometría contribuir algo nuevo al conocimiento de la obra del autor más reconocido y estudiado de la Literatura española? La variedad de géneros cultivados y su carácter como escritor que desborda los moldes establecidos creando su propia teoría literaria (Riley, 1998), así como su particular concepción de la Literatura (Canavaggio, 1998; Close, 1998), refuerzan los atractivos de la aplicación de la Estilometría y los enfoques cuantitativos para explorar su obra.

3. Preparación del corpus

La obra completa de Cervantes ha sido obtenida del portal dedicado a este escritor en la Biblioteca Virtual Miguel de Cervantes⁵². Debido a la breve extensión de las composiciones líricas en verso, estas se mantienen en un único conjunto textual, que abarca, por tanto, distintas fechas. Del mismo modo, las *Novelas ejemplares* se mantienen separadas, ya que son obras independientes. El corpus resultante contiene, pues, 36 obras de diversos géneros y temas: entre otras, la novela pastoril de *La Galatea* (1585), piezas teatrales posiblemente escritas entre 1581 y 1615 (publicadas este último año), obras en verso como *El Viaje del Parnaso* (1614), los dos *Quijotes* (el de 1605 y el 1615), las *Novelas ejemplares*, publicadas en 1613 y que siguen el modelo italiano de la *novella*, y el *Persiles*, su novela bizantina, publicada de forma póstuma en 1617. La lista completa de obras con los géneros, subgéneros a los que pertenecen y su extensión puede consultarse en la Tabla 1.

Título	Género	Subgénero	Palabras
El Amante Liberal	Narrativo	Novela ejemplar	18561
El Casamiento Engañoso	Narrativo	Novela ejemplar	5163
El Celoso Extremeño	Narrativo	Novela ejemplar	13699
El Coloquio de los Perros	Narrativo	Novela ejemplar	23050
El Gallardo Español	Dramático	Comedia	18362
El Juez de los Divorcios	Dramático	Comedia	3081
El Laberinto de Amor	Dramático	Comedia	18796
El Licenciado Vidriera	Narrativo	Novela ejemplar	9482
El Retablo de las Maravillas	Dramático	Entremés	3451
El Rufián Dichoso	Dramático	Comedia	16547
El Rufián Viudo	Dramático	Entremés	2909
El Trato de Argel	Dramático	Comedia	15552
El Viejo Celoso	Dramático	Entremés	3980
El Vizcaín Fingido	Dramático	Entremés	4557
La Galatea	Narrativo	Novela pastoril	117969

⁵² Web: http://www.cervantesvirtual.com/portales/miguel_de_cervantes/ [Consulta: 03/05/2021].

La Casa de los Celos	Dramático	Comedia	16630
La Cueva de Salamanca	Dramático	Entremés	3721
La Elección de los Alcaldes de Daganzo	Dramático	Entremés	2541
La Entretenida	Dramático	Comedia	16972
La Española Inglesa	Narrativo	Novela ejemplar	16041
La Fuerza de la Sangre	Narrativo	Novela ejemplar	7696
La Gitanilla	Narrativo	Novela ejemplar	23472
La Guarda Cuidadosa	Dramático	Entremés	3878
La Gran Sultana	Dramático	Comedia	16657
La Ilustre Fregona	Narrativo	Novela ejemplar	20016
La Numancia	Dramático	Tragedia renacentista	16042
Las Dos Doncellas	Narrativo	Novela ejemplar	14971
La Señora Cornelia	Narrativo	Novela ejemplar	14352
Los Baños de Argel	Dramático	Comedia	18265
Pedro de Urdemalas	Dramático	Comedia	16970
El Persiles	Narrativo	Novela bizantina	142710
Poesía	Poesía	Poesía lírica	10404
Primera parte del Quijote	Narrativo	Novela	181178
Segunda parte del Quijote	Narrativo	Novela	191759
Rinconete y Cortadillo	Narrativo	Novela Ejemplar	13696
<u>El Viaje del Parnaso</u>	<u>Narrativo</u>	<u>Poema narrativo</u>	<u>24802</u>

Tabla 1. Corpus de obras de Cervantes utilizado.

4. Metodología

Este estudio utiliza la metodología de la Estilometría más allá de la autoría, siguiendo los trabajos citados en el estado de la cuestión (sección 2), e inscrita habitualmente dentro de los enfoques cuantitativos en Humanidades Digitales (Hoover, 2008; Jannidis et al., 2017). Se combinan métodos clásicos de esta disciplina, como los análisis de grupos y los árboles de consenso (Eder, 2013), con los análisis de redes estilométricas siguiendo la implementación de Eder (2017). Como explica este investigador, puesto que los análisis de redes estilométricas mantienen un mayor número de conexiones entre textos (no solo las más consistentes, sino también otras más débiles), resultan de especial utilidad para explorar cuestiones estilísticas diferentes a la detección autorial. El análisis de grupos y el árbol de consenso han sido realizados mediante el paquete stylo (Eder et al., 2016), y se ha utilizado Gephi (Bastian et al., 2009) para convertir el *output* del árbol de consenso en un grafo. En cuanto a los parámetros del análisis, se ha utilizado Cosine Delta (Evert et al., 2017) por ser la medida de distancia que generalmente obtiene mejores resultados, aplicada sobre las 100 palabras más frecuentes en el análisis de grupos

y sobre las 100 a 1000 palabras más frecuentes en el árbol de consenso. Este rango ha sido elegido siguiendo los parámetros habitualmente utilizados y con el objetivo de restringir el análisis en la medida de lo posible a las palabras función.

5. Resultados

Para comenzar, se muestra el resultado del análisis de grupos o *cluster analysis* en la Figura 1. Como puede observarse, las obras de Cervantes quedan divididas fundamentalmente en dos grandes grupos: un primero que incluye sus textos dramáticos y poéticos, y un segundo con sus textos narrativos. De este modo, en el segundo grupo encontramos las *Novelas ejemplares*, las dos partes del *Quijote*, *La Galatea* y el *Persiles*; en cambio, el primer grupo contiene todas las comedias y entremeses compuestos por Cervantes, junto a su poesía lírica y el *Viaje del Parnaso*, formando estas dos últimas un pequeño subgrupo.

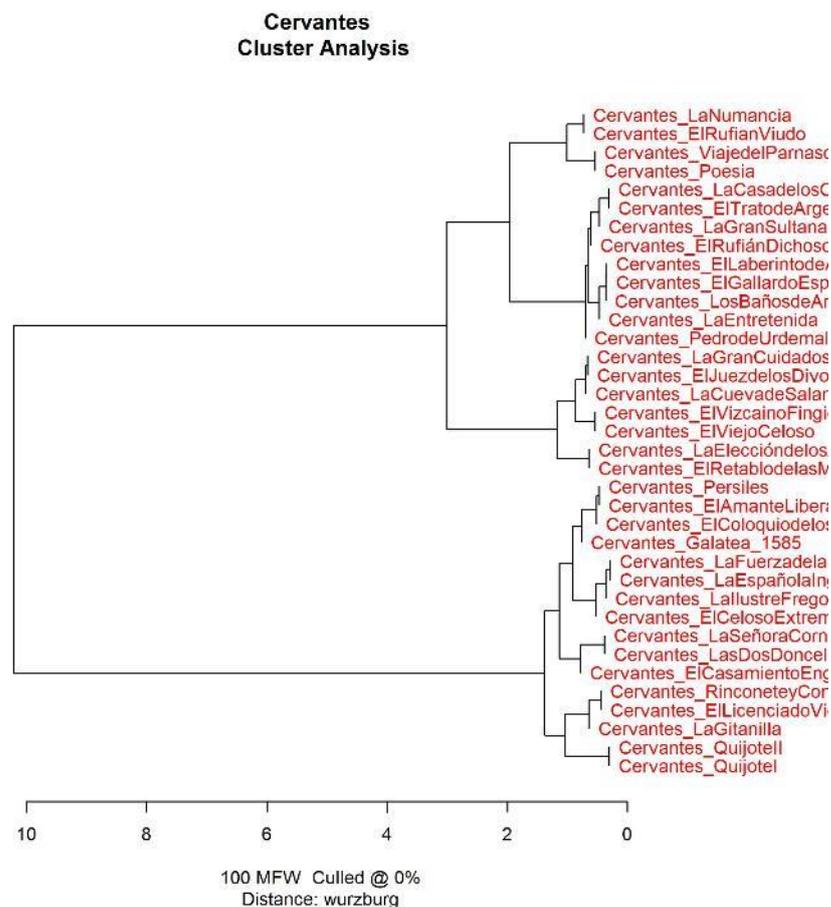


Figura 1. Análisis de grupos de la obra completa de Cervantes aplicando Cosine Delta a las 100 palabras más frecuentes.

Seguidamente, la Figura 2 presenta el árbol de consenso aplicando Cosine Delta sobre las 100 a 1000 palabras más frecuentes en la obra cervantina. Como es sabido, esta técnica recoge las conexiones más estables entre textos a largo de una serie de iteraciones. Como resultado, se observan menos relaciones entre textos y algunos de ellos cuelgan directamente de la raíz, como es el caso de varias *Novelas Ejemplares* (*El amante liberal*, *El casamiento engañoso*, *El celoso extremeño*, *La gitanilla*, *La ilustre fregona* y *Rinconete y Cortadillo*) y del *Persiles*. En este caso, se confirma que el dendograma es más interesante que el árbol de consenso para explorar cuestiones estilísticas no relacionadas con la detección de autoría, tal y como apuntan Jannidis y Lauer en su

estudio (Jannidis & Lauer, 2014). A pesar de ello, se mantiene la estrecha relación entre la poesía lírica y el *Viaje del Parnaso*, y encontramos tres ramas con las obras teatrales de Cervantes, una de las cuales incluye todos sus entremeses. Del mismo modo, los dos *Quijotes* aparecen juntos, así como algunas de las *Novelas ejemplares*.

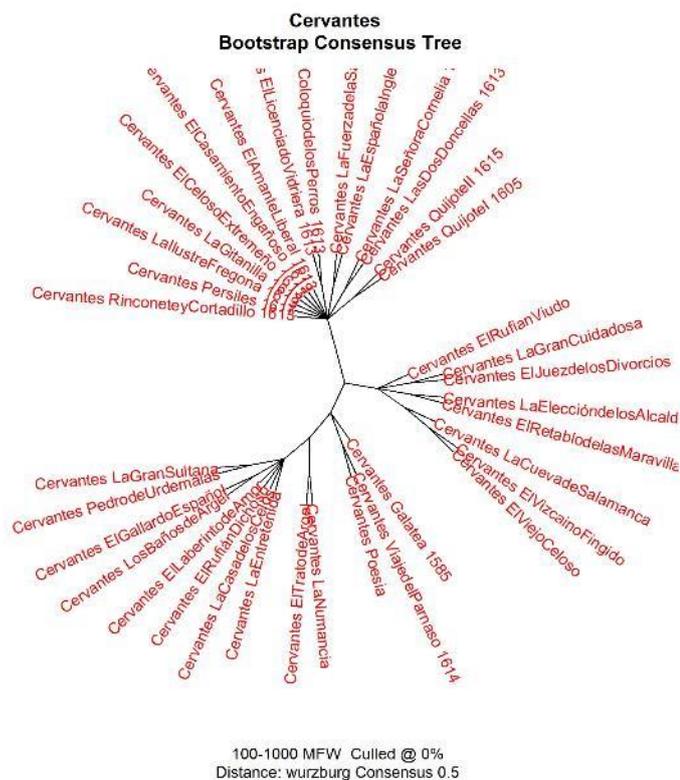


Figura 2. Árbol de consenso de la obra completa de Cervantes aplicando Cosine Delta sobre las 100 a 1000 palabras más frecuentes.

Por último, la Figura 3 muestra la red estilométrica obtenida con el *output* para Gephi generado por *stylo* del árbol de consenso anterior. Como puede apreciarse, las obras de Cervantes forman una red alargada, con el extremo izquierdo dominado por las obras teatrales y el derecho por las narrativas. El factor determinante en la formación de la red no parece ser la cronología, puesto que una obra temprana como *La Galatea* aparece junto a la novela bizantina póstuma de Cervantes, el *Persiles*. La aplicación del algoritmo de Modularidad ha permitido distinguir cuáles son las principales comunidades de la red. Con una resolución de 1, pueden distinguirse cinco comunidades: las dos de color verde claro y oscuro contienen las obras teatrales, mientras que la rosa y la naranja están constituidas por los textos narrativos. La quinta comunidad, en color celeste, se localiza en el centro de la red, y en ella se encuentra la poesía lírica y narrativa de Cervantes junto a dos de sus obras teatrales (*La Numancia* y *El trato de Argel*). Es interesante cómo estas dos últimas, a pesar de encontrarse cerca de las comunidades teatrales, parecen estar más cercanas a la poesía que el resto. Todas las obras teatrales del Siglo de Oro están escritas en verso, y Cervantes en este caso no es la excepción, por lo que nuestro resultado parece indicar un mayor aliento poético de estas dos piezas. Además, si se presta atención a las dos comunidades teatrales, se advierte que la comunidad de color verde claro agrupa los entremeses, mientras que la de color verde oscuro contiene las comedias. En este sentido, la única excepción es *El rufián viudo* que aparece en esta última comunidad a pesar de ser un entremés. En cuanto a las comunidades narrativas, cabe destacar cómo la de color naranja parece constituida por las obras más cómicas y grotescas, como la novela ejemplar picaresca de *Rinconete y Cortadillo* o los dos *Quijotes*, mientras que la de color rosa muestra obras con un tono más elevado o más

idealizadas como la novela pastoril de *La Galatea*, la novela bizantina del *Persiles* o la novela ejemplar de *La española inglesa*.

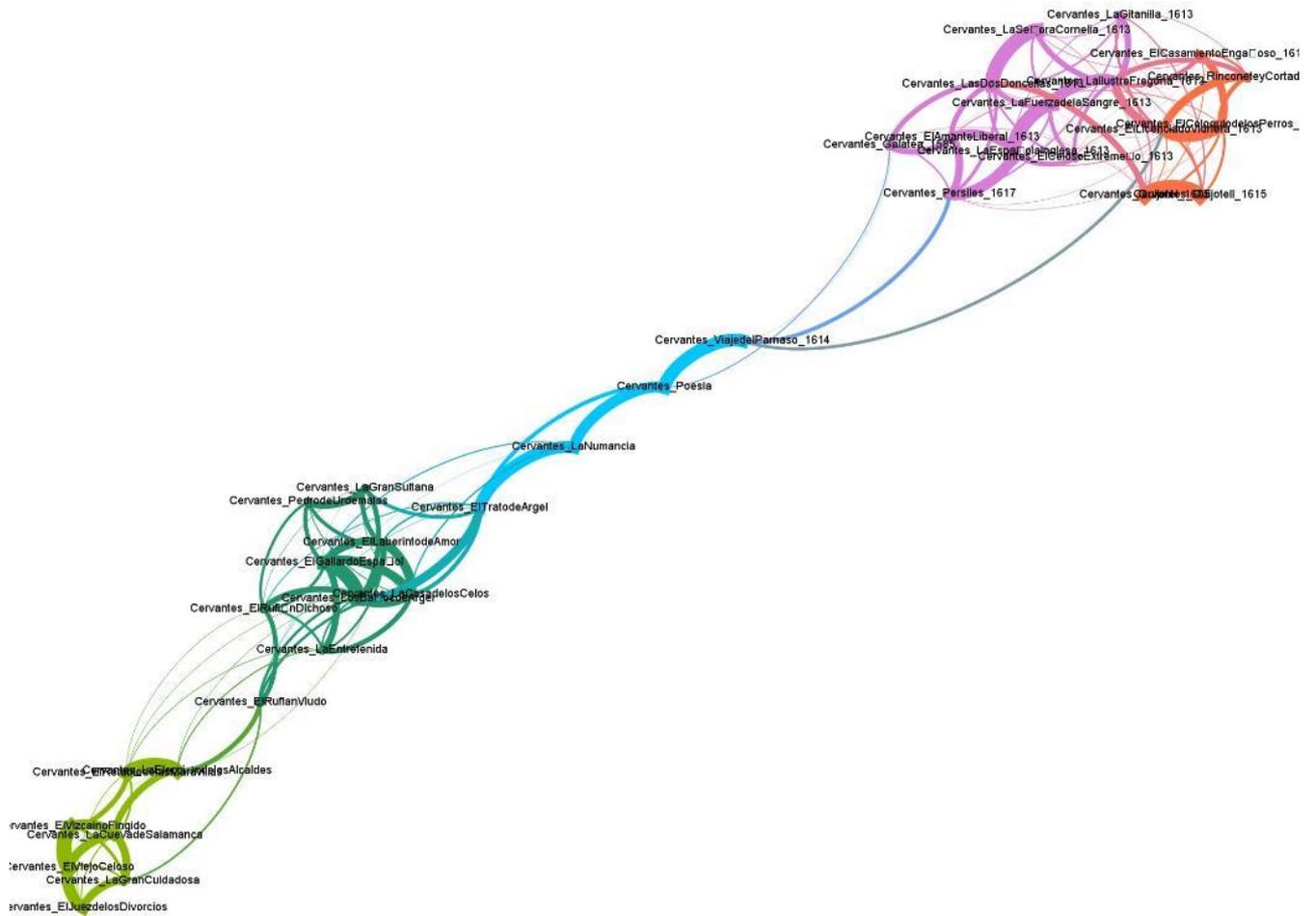


Figura 3. Análisis de redes de la obra completa de Cervantes coloreado por Modularidad, con una resolución de 1.

6. Conclusiones

Para concluir, el análisis estilométrico sobre la obra de Cervantes ha arrojado algunos resultados de gran interés. Los métodos estilométricos han conseguido detectar conexiones de género literario en el corpus, en línea con los resultados obtenidos en los estudios anteriores con otras obras, autores y literaturas. Esta señal de género literario parece ser la más fuerte en la obra cervantina, pues ha dominado tanto la formación del análisis de grupos como de la red estilométrica. Sin embargo, se aprecia una cierta influencia de la temática en la formación de las comunidades de textos narrativos detectadas mediante el cálculo de la Modularidad. Del mismo modo, una aportación interesante y que necesitará un estudio más extenso en el futuro es la estrecha relación entre la poesía cervantina y dos de sus obras teatrales, la tragedia renacentista *La Numancia* y la comedia *El trato de Argel*, así como la aparición del entremés *El rufián viudo* con las comedias, en lugar de con los demás entremeses.

Bibliografía

- Bastian, M., Heymann, S., & Jacomy, M. (2009). Gephi: An Open Source Software for Exploring and Manipulating Networks. *Third International AAAI Conference on Weblogs and Social Media*, 361-362.
- Burrows, J. (1987). *Computation into Criticism: A Study of Jane Austen's Novels and an Experiment in Method*. Oxford University Press.
- Burrows, J. (2002). 'Delta': A Measure of Stylistic Difference and a Guide to Likely Authorship. *Literary and Linguistic Computing*, 17(3), 267-287.
- Burrows, J. (2003). Questions of Authorship: Attribution and Beyond: A Lecture Delivered on the Occasion of the Roberto Busa Award ACH-ALLC 2001, New York. *Computers and the Humanities*, 37(1), 5-32.
- Burrows, J. (2004). Textual analysis. En R. Schreibman, R. Siemens, & J. Unsworth (Eds.), *A Companion to Digital Humanities* (pp. 323-347). Blackwell.
- Calvo Tello, J. (2019). Delta Inside Valle-Inclán: Stylometric Classification of Periods and Groups of His Novels. *Romanische Studien. Theorien von Autorschaft und Stil in Bewegung: Stilistik und Stilometrie in der Romania*, 6, 151-164.
- Calvo Tello, J. (2017). What does Delta see inside the Author?: Evaluating Stylometric Clusters with Literary Metadata. *III Congreso de la Sociedad Internacional Humanidades Digitales Hispánicas Sociedades, políticas, saberes*, 153-161. <http://hdh2017.es/wp-content/uploads/2017/10/Actas-HDH2017.pdf>
- Calvo Tello, J., Schlör, D., Henny, U., & Schöch, C. (2017). Neutralising the Authorial Signal in Delta by Penalization: Stylometric Clustering of Genre in Spanish Novels. *Digital Humanities 2017. Conference abstracts. McGill University & Université de Montréal. August 8-11, 2017*. <https://dh2017.adho.org/abstracts/O37/O37.pdf>
- Canavaggio, J. (1998). Vida y literatura: Cervantes en el «Quijote». En F. Rico (Ed.), *Don Quijote de la Mancha* (pp. 41-66). Instituto Cervantes.
- Close, A. (1998). Cervantes: Pensamiento, personalidad y cultura. En F. Rico (Ed.), *Don Quijote de la Mancha* (pp. 67-86). Instituto Cervantes.
- Eder, M. (2013). Computational stylistics and Biblical translation: How reliable can a dendrogram be? En T. Piotrowski & L. Grabowski (Eds.), *The Translator and the Computer* (pp. 155-170).
- Eder, M. (2017). Visualization in stylometry: Cluster analysis using networks. *Digital Scholarship in the Humanities*, 32(1), 50-64.
- Eder, M., Rybicki, J., & Kestemont, M. (2016). Stylometry with R: A Package for Computational Text Analysis. *The R Journal*, 8(1), 107-121.
- Evert, S., Proisl, T., Jannidis, F., Reger, I., Pielström, S., Schöch, C., & Vitt, T. (2017). Understanding and explaining Delta measures for authorship attribution. *Digital Scholarship in the Humanities*, 32(June 2017), ii4-ii16.
- Henny-Krahmer, U. (2018). Exploration of Sentiments and Genre in Spanish American Novels. *Digital Humanities 2018 Puentes-Bridges*, 399-403. https://web.archive.org/web/20190204111506/https://dh2018.adho.org/wp-content/uploads/2018/06/dh2018_abstracts.pdf

- Hoover, D. L. (2008). Quantitative Analysis and Literary Studies. En R. Siemens & S. Schreibman (Eds.), *A Companion to Digital Literary Studies*. Blackwell.
http://www.digitalhumanities.org/companion/view?docId=blackwell/9781405148641/9781405148641.xml&chunk.id=ss1-6-9&toc.depth=1&toc.id=ss1-6-9&brand=9781405148641_brand
- Hoover, D. L. (2014). A Conversation Among Himselfes: Change and the Styles of Henry James. En *Digital Literary Studies: Corpus Approaches to Poetry, Prose and Drama. The Corpus, the Computer and the Study of Literature* (pp. 90-119). Routledge Advances in Corpus Linguistics.
- Jannidis, F., Kohle, H., & Rehbein, M. (Eds.). (2017). *Digital Humanities: Eine Einführung*. J.B. Metzler Verlag.
- Jannidis, F., & Lauer, G. (2014). Burrows's Delta and Its Use in German Literary History. En M. Erlin & L. Tatlock (Eds.), *Distant Readings. Topologies of German Culture in the Long Nineteenth Century* (pp. 29-54). Camden House.
- Lutostawski, W. (1897). *The origin and growth of Plato's logic: With an account of Plato's style and of the chronology of his writings*. Longmans, Green & Co.
- Riley, E. C. (1998). Cervantes: Teoría literaria. En F. Rico (Ed.), *Don Quijote de la Mancha* (pp. 129-141). Instituto Cervantes.
- Rybicki, J. (2016). Vive la différence: Tracing the (authorial) gender signal by multivariate analysis of word frequencies. *Digital Scholarship in the Humanities*, 31(4), 746-761.
- Schöch, C. (2018). Zeta für die kontrastive Analyse literarischer Texte. Theorie, Implementierung, Fallstudie. En T. Bernhart, S. Richter, M. Lepper, M. Willand, & A. Albrecht (Eds.), *Quantitative Ansätze in den Literatur- und Geisteswissenschaften. Systematische und historische Perspektiven* (pp. 77-94). De Gruyter.

Sesión 6

La vía del testimonio

Estudios culturales II

Modera: María Luisa Alvite Díez · *Asiste:* Alexandra Iglesias Cortés

- 9 Poesía sobre el movimiento estudiantil mexicano de 1968, un acercamiento desde el testimonio y los imaginarios sociales
Jaime Ricardo Huesca Orozco
- 10 Clústeres y redes: una aproximación al ecosistema galerístico neoyorquino. Picasso como estudio de caso
María Ortiz Tello
- 24 Estrategias de recuperación de información mediante *pooling* para la construcción de colecciones de referencia desde redes sociales: caso de estudio durante las tensiones raciales de 2020
Patricia Martín-Rodilla, David Otero

Poesía sobre el movimiento estudiantil mexicano de 1968

Un acercamiento desde el testimonio y los imaginarios sociales

Jaime Ricardo Huesca Orozco

Maestría en Estudios de la Cultura y la Comunicación, Universidad Veracruzana

Palabras clave: poesía mexicana, movimientos sociales, *Voyant Tools*, testimonio, análisis de *corpus*.

La presente investigación emplea como objeto de estudio tres antologías que compilan (a lo largo de 40 años) gran parte de la poesía escrita en torno al Movimiento estudiantil mexicano de 1968. Más allá del contenido estético de los textos, el objetivo principal circunda en distinguir elementos discursivos y testimoniales referentes a la Masacre de Tlatelolco, acontecimiento donde, bajo órdenes del gobierno autoritario de México, fuerzas militares y judiciales agredieron con armas de fuego a jóvenes manifestantes en la Plaza de las Tres Culturas (Tlatelolco) en la tarde-noche del 2 de octubre; dichas acciones dieron como resultado un incontable número de muertos, desaparecidos, presos políticos y demás víctimas. Tal acontecimiento fue minimizado por el gobierno y medios de comunicación, quedando impune dentro de la historia contemporánea de México.

En tal ambiente de represión, violencia y control mediático, este género literario se consolida como espacio para la libre expresión, la crítica y denuncia hacia los agresores. A partir de la suma de estos escritos, se reconstruye un testimonio colectivo en torno a la masacre estudiantil, versión que resiste y a la vez apela los diversos relatos “oficiales” impuestos por el gobierno e insertos en la opinión pública gracias al apoyo de la prensa. Por lo anterior, es importante estudiar esta poesía debido a su contexto de producción y el papel fundamental que desempeñó para hacer evidente los crímenes de Estado.

Para comprender el fenómeno anterior y realizar una interpretación más acertada del mismo, se confeccionó un instrumento teórico-metodológico que incluyera tres conceptos relevantes: testimonio (discurso-testimonio), imaginarios sociales y hegemonía⁵³. A partir de un conjunto de teorías, se visualiza al testimonio como un tipo de escritura emergente utilizada por el intelectual como acto de solidaridad hacia grupos sociales marginados, exponiendo un relato vedado por fuerzas coaccionadoras. Así mismo, dicho género literario posee características discursivas ya que, desde la acción literaria, son vistos como documentos de urgencia que tratan de registrar un momento de crisis social, pensados para su publicación editorial y dirigidos contra “figuras antagonistas” (dictadores, presidentes, instituciones).

Desde Yúdice, el testimonio toma protagonismo en Latinoamérica a partir de 1968 con la Masacre de Tlatelolco y diversos movimientos en el centro y sur del continente, pues ilustran el fracaso de la hegemonía actual y proyectos de nación “modernos”, permitiendo la entrada de este género como alternativa para ejercer la democracia y la acción política desde la esfera cultural. Es preciso comprender estas movilizaciones desde el carácter hegemónico, pues la relación Movimiento-Estado guarda su sentido a partir del conflicto entre ambos.

En la poesía, es posible observar a un conjunto de voces vedadas por el gobierno y sus violentos actos de represión. En el lenguaje que emplean se observa la unificación de un imaginario social, entendido como “una forma específica de ordenamiento de un conjunto más amplio de representaciones que las sociedades dan para sí

⁵³ Para ello se empleó el trabajo de teóricos como Renato Prada, Beatriz Sarlo, George Yúdice, Bronislaw Baczko y Raymond Williams.

[...] una realidad específica que reside en su misma existencia, en su impacto variable sobre las mentalidades y los comportamientos colectivos, en las múltiples funciones que ejercen en la vida social (Baczko,

1984, p. 8)”. Tales ideas se legitiman desde el conflicto con otros imaginarios sociales: en los textos se observan vivencias compartidas, visiones del mundo, ideologías, emociones afines que originan la organización social desde abajo. En su suma, la poesía busca imponer un testimonio conformado por tal (o tales) imaginario social.

Desde un sentido teórico es posible aplicar el conocimiento anterior, pero al poseer un *corpus* de 155 poemas la práctica resulta complicada dada su extensión. Por ello, fue necesaria la elaboración de un método mixto donde se analice la información desde un enfoque cuantitativo y la interpretación se complementa con la teoría demostrada. Gracias a las facilidades de la “lectura lejana” que ofrece *Voyant Tools*, una plataforma web de código abierto especializada en el análisis de *corpus*, se compensó la carencia mencionada. Por otro lado, dicho ejercicio dio resultados que posiblemente no se percibirían desde la lectura cercana de quien investiga y, además, afianzó la comprobación de ciertas hipótesis.

Previo la digitalización del *corpus* y ciertos criterios de selección, el procesamiento en *Voyant Tools* permitió estudiar los datos a partir de diversas herramientas de conteo, visualización y correlación de términos. Una de las bases interpretativas consistió en el número de recurrencias presentes en cada término y cómo algunas palabras específicas se interconectaban con el lenguaje general del *corpus*. El primer paso fue la visualización de los términos más sobresalientes de esta poesía a partir de la función “Cirrus” o mayormente conocida como “nube de palabras”.

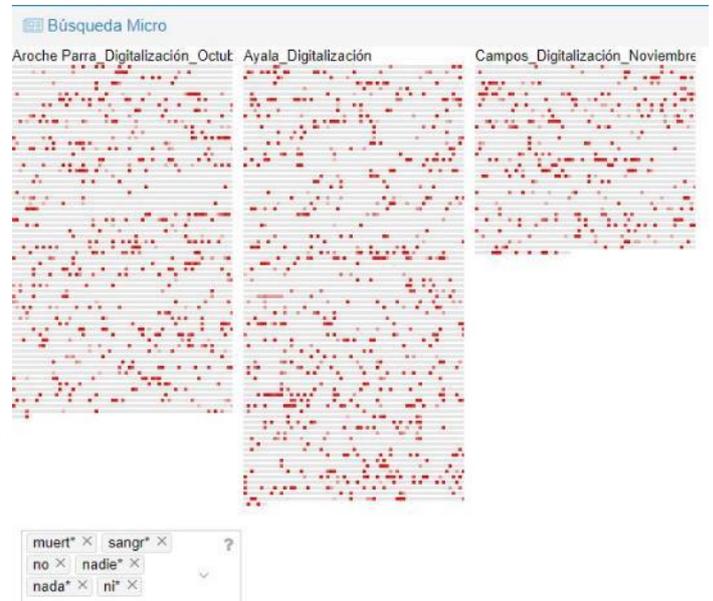
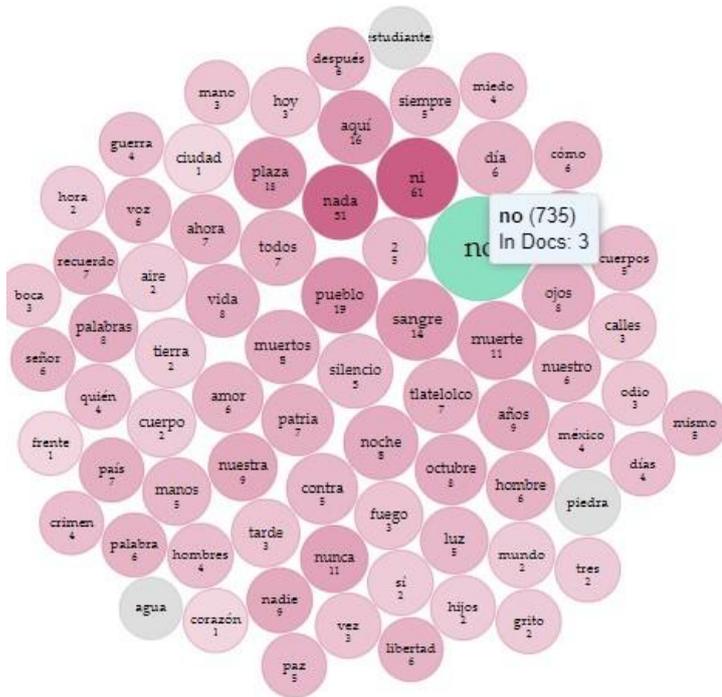


	Términos	Contar
1	no	735
2	sangre	167
3	ni	164
4	muerte	131
5	pueblo	109
6	nada	103
7	todos	101
8	silencio	100
9	día	97

El resultado de esta herramienta permitió la contabilización de los términos, a mi modo de ver, más importantes de esta poesía. La presencia dominante de adverbios de negación en la mayoría de los textos permite comprender que, desde un carácter discursivo, esta negación se enuncia hacia los acontecimientos registrados el 2 de octubre, un contexto social de imposibilidad y cuestionamiento político hacia las esferas gubernamentales de México. Por su parte, términos como “sangre”, “muerte”, “plaza”, “Tlatelolco”, entre otros, posicionarán a los textos por su carácter testimonial y la intención por reconstruir un relato a partir de la mirada en primera persona de sus autores o prestando su voz a las víctimas directas o indirectas del atentado. Además, desde esta impresión puede realizarse una primera interpretación de cómo se constituyen los imaginarios sociales desde este lenguaje común y cómo se utilizarán de forma discursiva.

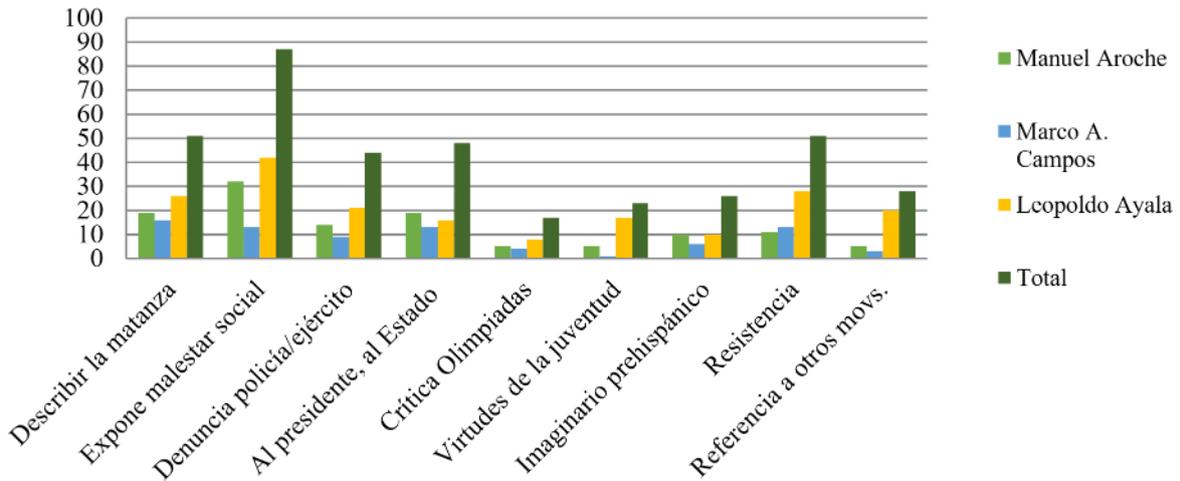
Dicho conteo, a partir de la herramienta “TermsBerry”, que muestra las correlaciones existentes entre un término con su respectivo *corpus* (por medio de cercanías dentro del texto) y “Microsearch”, que elabora una visualización abstracta de los documentos procesados y posiciona una palabra seleccionada en tal abstracción, se comprendió

que, para el estudio cercano de los poemas, hay palabras que por su excesiva presencia no pueden estudiarse salvo desde la mirada cuantitativa; por otro lado, éstas cumplen una función mayor en la significación de los poemas: no pueden estudiarse desde su particularidad pero sí como modificadores de los enunciados que las incluyan, a este tipo de elementos los denominó “términos base”, referentes a “no”, “ni”, “sangre”, “muertos” y sus diversas lematizaciones.



Ahora bien, para la óptima interpretación de los poemas, se propone un análisis en descenso donde, por medio de diversos pasos, se traslade desde los datos cuantitativos que la plataforma proporciona y, así, emplear tal información en una posición cualitativa. Para ello, gracias a la mediación entre la lectura cercana-distante se aprovechó la herramienta “Cirrus” para realizar clasificaciones. Con la lectura propia de las antologías se distinguieron diversos temas a analizar en la poesía, los cuales establecen rutas tópicas de interpretación a partir de campos semánticos establecidos por medio de los conteos de “Cirrus”. Para el desarrollo del análisis, se seleccionaron palabras pertenecientes a uno de los campos semánticos y que resultaran representativas de algún tema en concreto. Aunque en el siguiente gráfico se observen nueve opciones temáticas, la investigación se concentró en el estudio de los tres temas con mayor frecuencia: la descripción (generalmente testimonial) de la masacre estudiantil, la exposición del malestar social en el contexto sociopolítico en 1968 y, por último, aquellos textos que invitan a la resistencia social desde diversas estrategias; es necesario mencionar que tales poemas, en su mayoría, son complejos, extensos y contienen más de un referente tópico, donde pueden distinguirse subtemas que dependen directamente de las tres rutas señaladas. Como paréntesis, las frecuencias se distinguen a partir del nombre de sus compiladores.

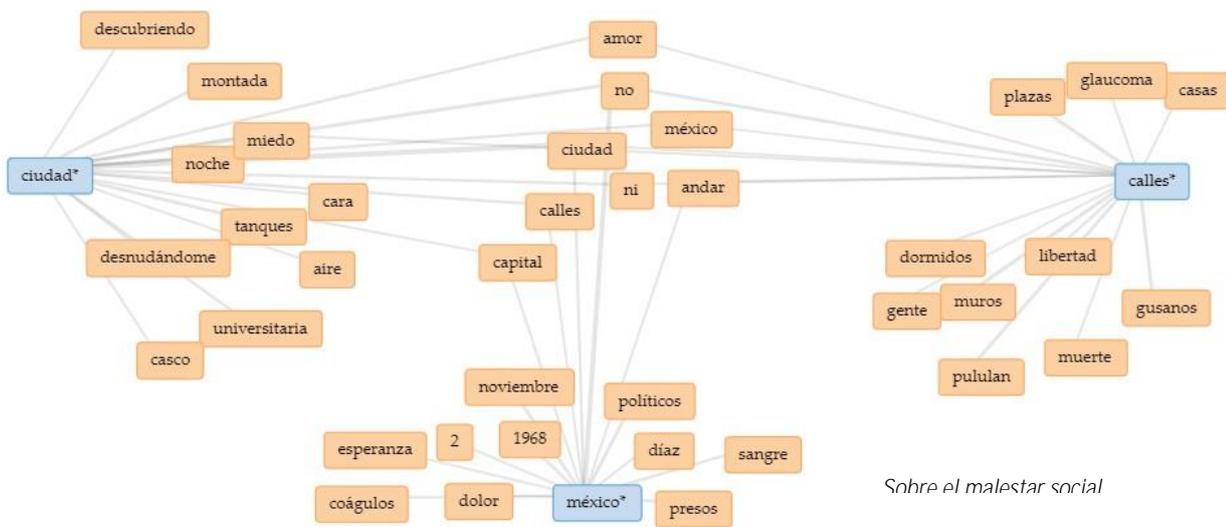
Sumario de tópicos por antología



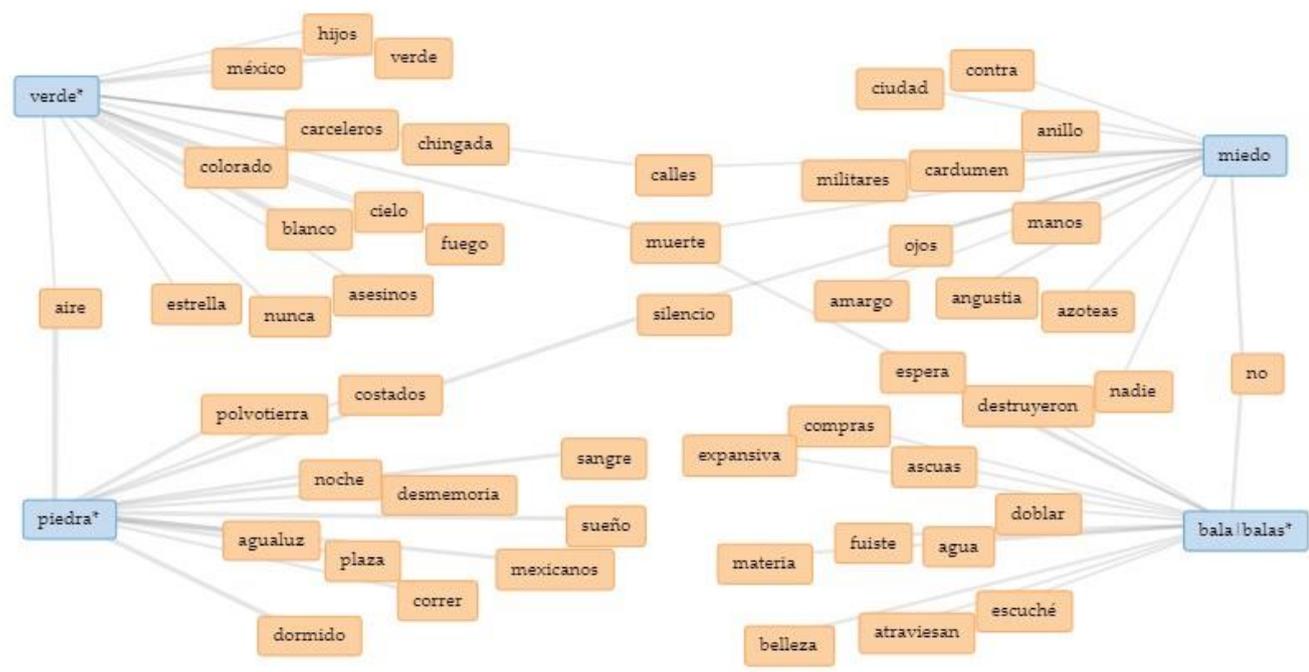
Para el análisis de cada tema, se seleccionó un conjunto de palabras en proporción con la relevancia de cada tema, después de ello, con la herramienta "Links" se correlacionan dichos términos (por tema) y observan las ramificaciones denotativas entre éstas. Tal resultado dibuja una interpretación previa a la muestra poética y deja entrever ciertas estrategias discursivas en cuanto a la enunciación y carga significativa de los textos.



Invitación a la resistencia.



Sobre el malestar social



Descripción de la matanza

Para cada palabra seleccionada se tomó una muestra textual de tres a cuatro fragmentos de poema y se observó cómo, desde el término seleccionado, se identifican diversos elementos que, en su conjunto, reiteran imágenes poéticas o estrategias discursivas. En su suma, los resultados muestran una compleja unidad de ideas y motivos que, desde el plano enunciativo, cuentan una historia alternativa que responde a las imposiciones de su contexto: testimonio, crítica, movilización, resistencia y la espera (pasiva) de un cambio positivo en el ejercicio político y social.

Por mencionar dos ejemplos, al seleccionar el término “país” (referente al malestar social en México de 1968) sobresale una crítica hacia dicha construcción social ya que, al sobrevivir en tal ambiente negativo, hay un cuestionamiento hacia la veracidad del concepto. Sobresalen adjetivos negativos, la añoranza o recuerdo de un “país” distinto antes de Tlatelolco. Por otro lado, persiste la consideración de este espacio como un infierno.

Hoy, pinto en una barda:
"La hora,
el día, mi país
es un gran horno,
no se ve más que luz." (Aroche, pp.168-169).

Ardiente, amado, hambriento, desolado,
bello como la duda, la sagrada blasfemia;
país de oro y limosna, país y paraíso,
país-infierno, país de policías (Aroche, p.145).|

[...] Y hasta el niño más anciano advertiría
Que el país que tuvimos ya no lo tenemos.

Un nuevo territorio
-En este siglo expansionista-
Al infierno fue anexado
En un dos de octubre mexicano (Aroche, p.141).

Por otra parte, respecto a la construcción testimonial del atentado en Tlatelolco, diversos referentes son reiterativos a lo largo de los poemas. Tal recurrencia legitima la veracidad de dicho acontecimiento y fortalece el testimonio general dentro de la poesía. En torno al empleo de la palabra "verde", que podría relacionarse comúnmente desde un aspecto positivo por representar la vida, la juventud, la esperanza, etcétera, dentro de este discurso será relacionado con la muerte, ya que la señal para que diera inicio el tiroteo consistió en el lanzamiento de tres bengalas de color verde; a su vez, dicho color se emplea como una crítica hacia los símbolos patrios por ser esta tonalidad parte importante de la bandera mexicana.

Nunca el verde fue más tétrico y odiado
que en esta noche que produce un rojo desolado,
caliente y borboteante,
con el viaje del plomo despiadado
que equivocó, de rumbo. (Aroche, p.39)

[...] sobre el cielo de México
la verde luz,
principio del horror y las tinieblas,
"Nosotros estuvimos allí"
quiere decir, oímos
escuchamos la sangre, el estertor, la muerte
rozando nuestra espalda [...] (Aroche, p. 86).

[...] de allá del cielo bajó,
fue del cielo que bajó
como un relámpago,
del cielo bajó el rayo verde,
los hombres se reunían para iniciar la danza,
luego llegaron los perros de la muerte,
babeaban,
sus pisadas sobre las escalinatas,
plam, plam, plam, [...]
sus pisadas,
plam, plam.
Del cielo cayó como un cometa,
un rayo verde, un relámpago verde, una estrella verde.
Nosotros somos testigos. (Ayala, p.229).

Aunque es necesario profundizar más respecto al análisis, los resultados obtenidos son valiosos para el estudio de este tipo de poesía, donde el ejercicio interpretativo desde *Voyant Tools* se aprecia como una alternativa viable para el estudio literario. Debo aclarar que su procesamiento está orientado en el análisis de contenido textual y, para acceder a un análisis enunciativo, es necesario el trabajo conjunto y mediado con la plataforma desde un enfoque metodológico mixto, más al tratarse de una investigación sobre poesía.

En términos generales, *Voyant Tools* facilitó el rango de observación de los textos, iluminando ciertas tendencias generales como el empleo de figuras retóricas básicas: símil, reiteración, sinécdoque, metáfora impura, que, desde la interpretación discursiva, comprenden una intención autoral hacia quienes lean los poemas, lo anterior con la finalidad de elaborar una comunicación clara y libre de oscuridades, ya que el lenguaje utilizado expresa un relato o crítica desde la alteridad, más social que estético. Dentro de estas tendencias, también es curioso observar la construcción de imágenes poéticas similares que podrían mostrar un modo específico de escritura o de pensamiento que validaría la existencia de imaginarios sociales concretos referentes al movimiento y masacre estudiantil. Debo insistir en la relevancia de este *corpus* como una unidad interconectada con cada uno de sus términos, donde no es posible excluir ninguno de ellos: en todo momento sobresale una diversidad de

términos, imágenes e ideas vinculadas con cualquier ruta. Toda palabra adquiere el carácter de ser principio y también fin.

Referencias

- Arcohe Parra, M. (1972). *53 poemas del 68 mexicano*. Editora y Distribuidora Nacional de Publicaciones.
- Ayala, L., Tlatletapas, J., & Ramírez Centeno, M. (2009). *EL libro rojo del 68. Poesía y gráfica social: Movimiento Estudiantil Mexicano 1968-2008*. Fundación de Estudios Sociopolíticos, Económicos, Autogestión y Poder Popular del Distrito Federal, AC.
- Baczko, B. (1984). *Los imaginarios sociales. Memorias y Esperanzas Colectivas* (2.ª ed.). Nueva Visión.
- Campos, M. A., & Toledo, A. (1996). *Poemas y narraciones sobre el movimiento estudiantil de 1968* (2.ª ed.). Universidad Nacional Autónoma de México.
- Prada Oropeza, R. (1989). *El discurso -testimonio* [Article]. Centro de Investigaciones Lingüístico Literarias. Universidad Veracruzana. <https://cdigital.uv.mx/>
- Sinclair, S. & G. Rockwell. (2021). Voyant Tools. Retrieved March 27, 2021, from <https://voyanttools.org/?stopList=keywords-1e8ec0c865a14cf12ffb60434b30ef78&panels=corpusterm%2Ctermsberry%2Ctrends%2C>
- Yúdice, G. (1992). Testimonio y concientización. *Revista de Crítica Literaria Latinoamericana*, 18(36), 211-232. <https://doi.org/10.2307/4530631>

Clústeres y redes: una aproximación al ecosistema galerístico neoyorquino. Picasso como estudio de caso

María Ortiz Tello
Universidad de Málaga

Palabras clave: Picasso; Nueva York; Redes; Clusterización; Galerías

1. Introducción

El objetivo de la presente comunicación es abordar las siguientes cuestiones: ¿Cuáles fueron y cómo se distribuyeron, tanto a nivel espacial como temporal, las distintas galerías neoyorquinas (ubicadas en Manhattan primordialmente) a lo largo del siglo XX y cuál fue el posicionamiento de Picasso dentro de esta distribución? Lo que nos lleva a plantearnos si, a medida que la fama de Picasso va aumentando y su obra revalorizándose, ¿su geoposicionamiento dentro de los clústeres de galerías neoyorquinas tiende a situarse de las periferias hacia los núcleos de los cúmulos galerísticos?

El análisis de los datos desde una perspectiva macroscópica permite analizar las dinámicas de origen, desarrollo y desaparición de los clústeres galerísticos neoyorquinos, revelando dónde y en qué períodos se sitúan. A continuación, hallados cuáles son estos clústeres y su disposición espaciotemporal, pueden ser analizadas las causas que los originan, examinándolos desde una perspectiva más cercana, microscópica, dando lugar a la tercera pregunta principal –y última– que rige este trabajo: ¿existe una red relacional entre las galerías aglomeradas? Si es así ¿cómo se producen tales relaciones?

De este modo, lo que esta prospección se propone es abordar una perspectiva de estudio del ecosistema galerístico-expositivo picassiano neoyorquino innovadora, atendiendo a las dinámicas espaciales, temporales e interpersonales de un amplio conjunto de galerías.

Este estudio se circunscribe, por lo tanto, dentro de los análisis de los llamados “fenómenos de clusterización”, iniciados a finales del siglo XIX por el economista Alfred Marshall. Marshall desarrolla la teoría de los distritos industriales, entendidos estos como las agrupaciones de empresas de un mismo sector que se benefician tanto de su proximidad geográfica como de las redes establecidas entre las mismas. Al concentrarse en una misma ubicación, se produce reciprocidad entre ellas, dando lugar a beneficios económicos y sociales.

A partir de Marshall, muchos otros autores comenzaron a preocuparse por estudiar y definir las características inherentes a las agrupaciones empresariales, destacando especialmente dos autores: Giacomo Becattini, que recupera y amplía el concepto de distrito industrial, y Michael E. Porter, responsable del concepto de clúster, definiéndolo como un territorio acotado en el que las empresas y otras entidades se interrelacionan, cooperando y compitiendo entre sí, fomentando, de este modo, el desarrollo de la innovación y la productividad entre las empresas conectadas. A partir de estas teorías, multitud de autores se han interesado tanto por el término y su definición como por sus características.

A pesar de que cada investigador desarrolla su propia teoría del clúster –o distrito– industrial, existen, no obstante, una serie de parámetros en los que todos coinciden a la hora de definir el fenómeno.

En primer lugar, el aspecto territorial es esencial. Existen, no obstante, diferentes dimensiones para el concepto de territorialidad, pudiendo entenderse desde un punto de vista local, regional, nacional o internacional. Se trata esta, pues, de una variable siempre presente, pero de carácter dinámico, determinada por las circunstancias de cada caso. Por otro lado, se encuentran las interrelaciones establecidas entre las empresas integrantes del clúster. Las entidades que conforman el clúster interactúan entre sí, construyendo entre todas una red social –que puede ser representada como un grafo– donde, por un lado, las empresas y demás agentes implicados representan los nodos, y por el otro, las aristas o lazos constituyen las relaciones o vínculos establecidos entre los mismos. Las entidades que representan los nodos de la red suelen estar especializadas en el mismo sector, lo que implica que se produzcan relaciones de competitividad entre las ellas, dando lugar a mayores exigencias de calidad, innovación y, finalmente, producción, lo que incentiva el crecimiento de las compañías incluidas en la red. Asimismo, estas empresas comparten conocimientos, cooperando y coordinándose entre ellas de manera habitual, generando una atmósfera de confianza indispensable para el correcto funcionamiento del clúster. De igual modo, también deben ser tenidos en cuenta para la conformación y continuidad del clúster los factores externos que lo posibilitan.

Aunque existen múltiples tipologías de clústeres, el que atañe a este estudio es el denominado clúster o distrito cultural. La característica del distrito cultural es que este concentra en un área delimitada un gran conjunto de agentes culturales, que bien pueden tratarse de instituciones u organizaciones artísticas, galerías, marchantes, artistas o casas de subasta, entre otros. Su principal fortaleza reside en la red social creada a partir de la interacción de este conjunto de agentes.

Atendiendo a todos estos aspectos, el presente estudio se compone por dos grandes bloques: la primera parte está elaborada desde una perspectiva macro, analizando los distintos distritos galerísticos que van surgiendo, estableciéndose y desapareciendo a lo largo del siglo, indicando en cuáles de ellos se va situando la obra de Picasso, elucidando si su posicionamiento dentro de la red se va centralizando a medida que es expuesto y difundido, es decir, comprobando si la fama de Picasso es medible geográficamente. El segundo bloque se concentra en un año concreto, desde una perspectiva micro, analizando las redes relacionales establecidas entre las diferentes galerías activas en el año 1939, lo que me permitirá dilucidar si las galerías se encuentran relacionadas entre sí, afectando a la obra de Picasso y su exposición. De esta manera, este estudio se fundamenta en la convergencia de tres ejes: el espacial, el temporal y el interpersonal, aportando nuevo conocimiento sobre las galerías neoyorquinas, los distritos industriales y la difusión, exposición y comercialización de la obra picassiana.

2. Metodología

Para poder responder a las preguntas de investigación planteada, ha sido empleada una metodología de carácter mixto, que combina las metodologías cuantitativa y cualitativa. En un primer estadio, el proceso metodológico es cuantitativo, buscando y recopilando datos en una base (en este caso, han sido registrados en hojas de cálculo). Posteriormente, estos datos han sido analizados a través de su procesamiento en ArcGIS, mediante el que he elaborado los mapas, y Gephi, con el que construí la red relacional.

Para localizar las galerías existentes en cada década, analicé, en primer término, diversas revistas y secciones culturales de periódicos neoyorquinos, anotando todas las galerías que aparecían en cada uno de ellos, sus direcciones y estilos artísticos. Dada la extensión del estudio, resolví acotarlo y escoger un año específico de cada década, sirviendo como ejemplo representativo de la misma. La decisión fue seleccionar el año «cinco»: 1905, 1915, 1925... se trata de un año intermedio y, de unos a otros, pueden observarse con claridad los cambios operados en el plano geográfico de la ciudad. En total, han sido recopiladas más de 1.200 galerías, todas ellas

geolocalizadas en los diferentes mapas establecidos. Asimismo, una vez que supe qué galerías se encontraban activas en 1939, comencé a buscar información sobre las mismas, anotando las relaciones que entre ellas iban apareciendo, de tal modo que reuní un *dataset* de relaciones surgidas entre las galerías implicadas. Estas relaciones fueron simplificadas, de tal modo que una entidad mantiene una conexión «x» con otra entidad. La categorización de esas relaciones surgió tras analizar el conjunto general y observar que muchas de ellas se repetían. De este modo, las relaciones fueron divididas en tipologías: ubicación, familiar, laboral, cooperación, competitividad, stock y amistad.

No obstante, existen diversas problemáticas que hay que tener en cuenta: en primer lugar, son muy pocos los estudios que se dediquen a analizar las redes de conexiones establecidas dentro del sistema artístico neoyorquino; en segundo lugar, subsiste un problema general en cuanto a la nomenclatura de las galerías, que aparecen indexadas con distintos nombres a lo largo de las fuentes; por otro lado, las galerías no son estáticas, cambian de ubicación cada cierto tiempo, pues a medida que su prestigio va aumentando, su localización se sitúa en espacios más exclusivos (aunque existen excepciones). Por lo tanto, averiguar la dirección de la galería en cada año supone una investigación concienzuda y extensa; finalmente, la pérdida de datos es inexorable, lo que implica que la base de datos galerística contiene vacíos, caracterizándose por su organicidad, abierta a futuras inclusiones.

3. Resultado del análisis

Dada la idiosincrasia de la capital del arte estadounidense, Nueva York es uno de los mejores ejemplos de distritos culturales que existen en la actualidad. Estos distritos han estado presentes a lo largo de todo el siglo XX bajo diversas formas. Aunque gran parte de los estudios sobre los fenómenos de clusterización galerística neoyorquina centran sus intereses en el SoHo, a partir de los años 70, y en la 57th St como clúster histórico por antonomasia, el análisis de los datos hallados y su posterior visualización en mapas permite refutar estas creencias generalizadas de que solo merecen atención estos dos espacios, especialmente el primero. De este modo, gran parte del estudio ha sido dedicado a dilucidar cuáles han sido los distritos culturales presentes en la metrópoli a lo largo del siglo y qué circunstancias las rodearon, así como qué distribución tuvo Picasso en ellas (fig. 1).

El primer clúster galerístico del siglo XX se situó en la Quinta Avenida, primero en su tramo sur y, con el paso de los años, en torno a la segunda década del siglo, desplazándose hacia Central Park, concomitante a la expansión del arte moderno europeo por la ciudad. En la década de los años 20, todavía la Quinta Avenida ostenta el primer puesto como lugar preferente para el asentamiento de las galerías artísticas, aunque la 57th St comienza a tener ya muchísima presencia, seguida por Madison Avenue. A partir de los años 30, la 57th arrebatará el puesto a la Quinta, que nunca más volverá a situarse como epicentro. El asentamiento de la 57 como núcleo galerístico neoyorquino coincide, asimismo, con el asentamiento del arte moderno europeo en sus programaciones y compraventas. La 57 se mantendrá como núcleo principal del distrito galerístico neoyorquino situado entre Midtown y el Upper East Side hasta prácticamente la llegada de los años sesenta, seguida por Madison Avenue que, en la década de los sesenta y los setenta, se convertirá en el núcleo principal del clúster, superando en galerías a la 57th St. Asimismo, en estos años comienzan a situarse en la zona sur de la isla algunas galerías. Es el germen de un nuevo clúster que, especialmente a partir de los años 80 y 90, adquirirá gran relevancia: es el SoHo, un distrito cultural que podríamos denominar como “programado”. En los ochenta, Madison pierde importancia, aumentando de nuevo el volumen de galerías en la 57th y en el SoHo que, en los años noventa, se convertirá en el distrito artístico-galerístico más importante de la ciudad, desplazando tras cincuenta años a la 57.

Dentro de este panorama, Picasso supone un paradigma idóneo para ejemplificar cómo se mueve el arte moderno europeo en el ecosistema galerístico neoyorquino del siglo XX. Su primera aparición expositiva se produce en 1911. En esta década, la obra de Picasso es expuesta de manera dispersa, tanto dentro como fuera de los núcleos principales. Sin embargo, a medida que avanza el siglo, el artista va centralizándose cada vez más progresivamente. Hacia la década de los años treinta y cuarenta, la exposición de su obra alcanza su cénit, llegando a ser exhibida en más de 150 ocasiones en cada década. Esto se refleja en su distribución, exhibiéndose de forma muy centralizada en los principales núcleos. La obra de Picasso es mostrada, por lo tanto, siempre en las zonas más relevantes hasta prácticamente la década de los setenta, cuando su exposición comience a convertirse en un asunto casi exclusivamente museístico.

Una vez analizado el panorama expositivo neoyorquino a lo largo del siglo XX y dilucidando cómo se desarrollaron en la ciudad los diversos distritos artístico-galerísticos (así como la presencia de Picasso en ellos), el estudio se dirigió al análisis micro de un determinado año, 1939, investigando las conexiones relacionales que posibilitan y determinan la existencia de tales distritos (el año elegido ha sido 1939 porque se trata del año con más presencia picassiana en el panorama expositivo, produciéndose ciertos hechos particulares, como la presencia de *Guernica*, la adquisición de *Les Femmes d'Alger* por el MoMA, que también celebrará la retrospectiva del artista más grande hasta la fecha en territorio norteamericano).

Estas relaciones se crean a partir de vínculos establecidos entre los propios agentes, pudiendo ser directos o indirectos. En el caso de este análisis, han sido tenidos en cuenta únicamente los vínculos directos, aquellos que se producen de manera inmediata entre una galería y otra. Es este el primer estadio de una futura investigación. Los lazos fuertes establecidos entre galerías obedecen a múltiples y muy heterogéneas circunstancias.

Para poder establecerlos, cada una de las galerías ha sido investigada individualmente, anotando sus vicisitudes (incluyendo las exposiciones celebradas en ellas durante todo el periodo anual).

Tras anotar los múltiples datos que en este período han sido hallados con respecto a las galerías que conforman el ecosistema neoyorquino del año 1939 (insisto en que este no es más que un prototipo de lo que será un estudio mayor), estos han sido simplificados y agrupados en categorías, según el tipo de relación que constituyesen: ubicación, amistad, familiar, stock, laboral y colaboración y competitividad. En esta primerísima fase del estudio, esta categorización se presenta como un instrumento útil para, una vez detectadas las conexiones, conocer a qué tipo de relación obedecen. No obstante, no se ha tratado de la cuestión primordial en este estadio. Lo importante en este primer estadio es dilucidar, de manera fehaciente y cuantitativa, que existen múltiples interrelaciones, tal y como afirman los autores incluidos en el estado de la cuestión.

Tras adecuar los datos a las hojas de cálculo (en csv) para su posterior procesamiento en Gephi, el programa de análisis de redes empleado en esta prospección, el resultado es el que se muestra (fig. 2).

Lo que el grafo permite visualizar en un simple vistazo es que realmente existe una red relacional entre las galerías que conforman los clústeres (aun sin tener en cuenta los lazos débiles). De esta forma, han sido halladas un total de 420 relaciones entre 134 galerías, a pesar de que aún quedan muchas relaciones por descubrir –estas son un primer paso en la investigación–; con este volumen, ya se establecen relaciones entre un amplio conjunto de las galerías (un 70 % aproximadamente). Así, el objetivo principal de ese apartado, verificar que existen relaciones es alcanzado. Una vez demostrado que se producen tales relaciones, ahora es el momento de analizarlas en profundidad.

Del mismo modo, extraje aquellas galerías que durante este período habían expuesto la obra de Picasso. Estas son, por lo tanto, las galerías implicadas en el ecosistema expositivo picassiano neoyorquino de 1939. Localizadas en emplazamientos muy cercanos unos con otros, se propicia así una relación territorial entre ellas; no obstante, en el grafo general también queda claro que estas forman parte, a su vez, de la red relacional

instaurada por el conjunto de decenas de galerías que conforman los clústeres. Las más sobresalientes, al igual que con el grafo anterior, han sido examinadas profundamente, advirtiendo su presencia a través de los orígenes de sus relaciones.

4. Conclusiones

El estudio de las galerías establecidas en la isla neoyorquina a lo largo del siglo XX desde una perspectiva macro permite comprender cuáles son las dinámicas y lógicas que configuraron el ecosistema galerístico neoyorquino, dilucidando cómo se establecieron los distritos galerísticos, en qué momento se originaron, cuánto perduraron y dónde se establecieron. El primer clúster no es la 57, previamente ya existía, hasta los años veinte, la Quinta Avenida. Madison Avenue es, durante los sesenta y los setenta, el principal núcleo del distrito cultural situado en el Upper East Side, por encima del SoHo, pese a que gran parte de la producción académica relativa a este fenómeno se centre en esta área casi exclusivamente (la idiosincrasia del SoHo reside en que es un clúster eminentemente dirigido). Asimismo, mientras que la característica principal de la 57 reside en que en ella se crean pequeños subclústeres localizados en edificios concretos, como el edificio Fuller, situado en el 41 E, en el SoHo, Madison y la Quinta, dada la amplitud de estas vías, las galerías se sitúan más dispersas, sobre todo en el SoHo, barrio compuesto por múltiples calles colindantes. Asimismo, el análisis de la situación dentro de estos clústeres de las exposiciones picassianas indica que el artista fue situándose a lo largo de los años en los lugares de mayor interés, quedando demostrado que, en su caso, sí que puede ser medido geográficamente –en un primer estadio– el reconocimiento artístico-social del artista.

Por otro lado, el análisis micro ha confirmado la existencia de una red relacional entre las galerías, lo que permite elucidar cuáles son las causas que motivan tales agrupaciones y qué relaciones subyacentes las fortalecen. Este aspecto es esencial para comprender en profundidad el sistema del arte neoyorquino y su idiosincrasia.

De esta forma, esta comunicación representa únicamente una pequeña introducción hacia un estudio de unas dimensiones muy superiores.

Figuras, tablas y gráficas



Figura 1. Distribución de las galerías neoyorquinas durante el siglo XX (en negro) y de la geolocalización de Picasso en ellas (en gris).

Bibliografía

BECATTINI, G., «Del distrito industrial marshalliano a la “teoría del distrito” contemporánea. Una breve reconstrucción crítica». *Investigaciones Regionales*, 1, 2002, págs. 9-32.

CAPÓ-VICEDO, J., EXPÓSITO-LANGA, M. y TOMÁS-MIQUEL, J., «Creación de Redes Interorganizativas en un Cluster Territorial». *Información Tecnológica*, 18 (4), 2007, págs. 3-10.

LAZZERETTI, L., «Distritos industriales, clusters y otros: un análisis trespassing entre la economía industrial y la gestión estratégica». *Economía industrial*, (359), 2006, págs. 59-72.

LÓPEZ POSADA, L. y CALDERÓN HERNÁNDEZ, G., «Análisis de las dinámicas culturales al interior de un clúster empresarial». *Estudios gerenciales*, 99, 2006, págs. 13-37.

PORTER, M. E., «Clusters and the New Economics of Competition». *Harvard Business Review*. 76 (6), 1998, págs. 77-90.

Estrategias de recuperación de información mediante *pooling* para la construcción de colecciones de referencia desde redes sociales: caso de estudio durante las tensiones raciales de 2020

PATRICIA MARTIN-RODILLA, *Information Retrieval Lab*, Centro de Investigación CITIC, Universidade da Coruña

DAVID OTERO, *Información Retrieval Lab*, Centro de Investigación CITIC, Universidade da Coruña

PALABRAS CLAVE: redes sociales, racismo, colección, *pooling*, recuperación de información

1. Introducción: las redes sociales como fuente de colecciones de referencia y archivos sociales

Determinados procesos sociales de especial relevancia en humanidades, como los procesos lingüísticos, de memoria e identidad, colaborativos o patrimoniales tienen su reflejo en las redes sociales, donde los usuarios expresan sus opiniones, debaten, comparten contenido e incluso planifican acciones, como manifestaciones o procesos participativos.

En estos contextos, numerosos investigadores estudian este contenido generado en redes sociales desde una perspectiva archivística, siendo posible encontrar numerosos ejemplos de archivos nativos digitales desde redes sociales (Pybus, 2013; Acker and Brubaker, 2014; Pybus, 2015) que buscan recuperar, preservar y poner a disposición de la comunidad partes de esa interacción en redes sociales para su posterior estudio.

Este nuevo enfoque, denominado en ocasiones “archivo social” (Ruiz Gómez and Vallès, 2020), permite “archivar las redes sociales” (Vallès, 2020), creando colecciones de referencia para el estudio de diversas temáticas y su reflejo e impacto en redes sociales (Gehl, 2009; Zubia Muxika, 2018; Perpinyà-Morera and Cid-Leal, 2020). Un ejemplo reciente de esto son los archivos sociales desde Twitter creados para estudiar aspectos de violencia de género (Vallès and Soler Jiménez, 2016; Ruiz Gómez and Vallès, 2020) o en contextos de enfermedad mental (Parapar, Martin-Rodilla, Losada and Crestani, 2021).

Existen protocolos que permiten la búsqueda, recuperación y archivo de páginas webs (Schebeck, Spiliotopoulos and Risse, 2012) o correspondencia electrónica desde un punto de vista archivístico. Sin embargo, en el caso de las redes sociales, los procesos de recuperación, búsqueda e indexación de información en dominios humanísticos siguen realizándose en su mayor parte de forma manual o semi-manual (Schebeck, Spiliotopoulos and Risse, 2012; Mordell, 2019), o bien usando técnicas computacionales no adaptadas a la naturaleza de la información que recuperan ni a su dominio (Winters and Prescott, 2019). Esto convierte a las tareas de RI en tareas muy costosas en tiempo y recursos. Además, la mayor parte de los protocolos de recuperación de información aplicados para crear colecciones desde redes sociales en el ámbito humanístico no abordan desde su creación aspectos de reproductibilidad (impidiendo replicar el proceso a otros investigadores) ni de privacidad y seguridad de la información, como los procesos de anonimización (McNealy, 2011; Obodoruku, 2016).

La recuperación de información (abreviada RI) como disciplina computacional, ha abordado la creación de colecciones de referencia desde sus inicios, desarrollando numerosas técnicas que permiten realizar estas

tareas de forma más eficiente, y evaluando la calidad de las colecciones resultantes. Combinando técnicas de RI con los protocolos emergentes de archivos sociales, algunos de los problemas identificados anteriormente pueden mejorarse sustantivamente. En concreto, para evaluar sistemas de recuperación y sus colecciones resultantes, así como para evitar que un asesor humano tenga que revisar todos los documentos de la colección creada (Losada, Parapar and Barreiro, 2017) surge la técnica denominada *pooling*, como método para extraer una muestra de un conjunto de documentos (posts o publicaciones en el caso de redes sociales), obteniendo el conjunto más completo e imparcial de documentos relevantes sobre un tema determinado. Esto es especialmente útil en las redes sociales, debido a la gran cantidad de posts existentes y temas tratados por el mismo usuario o en el mismo hilo o publicación.

Este artículo ilustra la metodología de aplicación de *pooling* como estrategia de recuperación de información desde redes sociales para la construcción de colecciones de referencia que aplicamos en nuestro grupo de investigación. Como caso real para una investigación en curso, se creó una colección de referencia desde la red social Reddit para el estudio del impacto en la comunidad hispanohablante en redes sociales del movimiento *Black Lives Matter* y las revueltas a raíz del asesinato de *George Floyd* en 2020.

2. Propuesta y metodología

El *pooling*, una metodología madura en el campo de la recuperación de información, consiste en seleccionar un subconjunto de información como un número manejable y finito de documentos recuperados (*pool*) para una temática dada (en este caso de *posts* desde redes sociales) (Sanderson, 2010). Una vez creado el *pool*, se presenta a asesores humanos que evalúan cada documento como relevante o no para la consulta enviada. Se han evaluado diferentes estrategias para ordenar este conjunto de documentos (Losada, Parapar and Barreiro, 2017), que sirven de base para la metodología de *pooling* aplicada en este trabajo.

Este número manejable de documentos (*pool*) permite posteriormente al sistema recuperar documentos similares a los juzgados como relevantes por asesores humanos, lo que hace asegura la calidad de la colección en cuanto a su adecuación a la temática deseada. La metodología que proponemos en este trabajo se detalla en la siguiente figura:

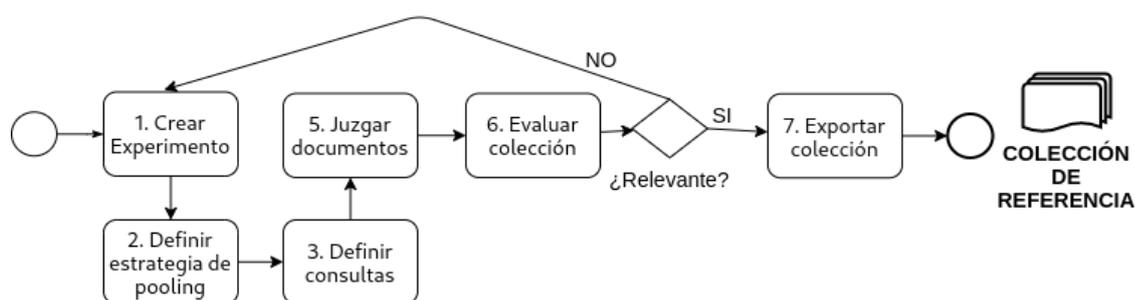


Figura 1: Metodología de pooling

En este contexto, un experimento es cada una de las instancias en las que desarrollamos el flujo de trabajo de la Figura 1, y abarca la selección de la estrategia de *pooling*, la definición de las necesidades de información y consultas, el juicio de los documentos y, finalmente, la creación de la colección final.

Dentro de cada experimento, el primer paso consiste en la selección de la estrategia de *pooling*. Ésta dicta el orden en el que los documentos serán presentados a los asesores expertos según diferentes criterios. A

continuación, el siguiente paso es definir las consultas, que se emplearán para la búsqueda de los documentos (*posts* en redes sociales). Estas consultas están normalmente formadas por varios términos clave que representan las necesidades de información de la colección, como se verá más adelante en el caso de estudio. Una vez seleccionada la estrategia de *pooling* y definidas las consultas que se utilizarán para recuperar documentos, el siguiente paso es establecer los juicios de los documentos seleccionados. Estos juicios pueden ser emitidos por una persona, o un grupo de personas (expertos en el dominio y temática en el que se encuadra la colección que se construirá finalmente). Los juicios consisten en que cada asesor experto determina si un documento es relevante o no para la temática abordada. Una vez los juicios se han completado, el último paso consiste en revisar si el contenido de la colección se adecúa a las necesidades de información que se establecieron al principio. Finalmente, los documentos que forman la colección en conjunto con los juicios establecidos por los expertos, la colección ya puede ser liberada y exportada para su uso en el futuro.

3. Caso de Estudio y Resultados

Durante el último año, nuestro grupo de investigación ha aplicado la metodología propuesta (ver Figura 1) para la construcción de colecciones de referencia en diversos ámbitos y proyectos de investigación en humanidades digitales. Entre ellos, una línea de interés era el seguimiento del impacto en redes sociales de las tensiones y protestas provocadas por la muerte de *George Floyd* el 25 de mayo de 2020. Para ello, y usando la red social *Reddit* (Reddit, 2020) como fuente, se aplicó en el pasado la metodología descrita para esta misma temática en inglés, creando una colección actualmente disponible (Otero, Martín-Rodilla and Parapar, 2020). La elección de *Reddit* viene dada por su alto número de usuarios en Estados Unidos (foco de las revueltas), su especialización por *subreddits* temáticos (haciendo más efectivo el *pooling*) y por su organización tipo blog, que permite a sus usuarios expresar opiniones largas sin limitación de caracteres. Presentamos aquí el caso de estudio y la colección resultante para español, explicando paso a paso la metodología aplicada para su creación.

3.1. Definición de términos de búsqueda

El proceso de diseño de consultas y selección terminológica constituye un proceso clave en la calidad posterior de la colección, y requiere de la colaboración interdisciplinar para su éxito. Algunas de sus implicaciones a nivel multilingüe y de calidad de la colección han sido discutidas previamente (Barreiro, Parapar and Otero, 2019).

Tabla 1: Términos en español seleccionados en la fase de definición de consultas.

	racismo	vandalizado
protesta		vandalismo esclavo
racista		
George Floyd		esclavitud
blm		policía
Black lives matter		violencia

estatua	monumento
busto	revuelta
placa	

La Tabla 1 muestra los términos empleados para la selección del *pooling* de documentos de la colección. La selección de términos se realizó como un proceso en paralelo a la colección en inglés (Otero, Martín-Rodilla and Parapar, 2020; Otero, Martín-Rodilla and Parapar, 2021) , así como guiado por diversas pruebas previas de cobertura (seleccionando, ante varias alternativas y sinónimos, el término que arrojaba mayores resultados en la red social). La facilidad del mecanismo de *pooling* para automatizar la posterior recuperación de información en base a las consultas diseñadas permite el rediseño continuo de las consultas a modo de ensayo error, la ampliación de esta lista de términos y la replicación de este proceso para ajustar cada colección a las necesidades de investigación, idiomáticas y regionales en cada caso.

3.2. Juicio sobre relevancia del *pooling*

Las consultas definidas dieron como resultado 219 publicaciones diferentes para juzgar. Los juicios fueron hechos por dos evaluadores a partes iguales, a los que se les presentó una publicación a la vez. Solo se anotaron como publicaciones relevantes los posts en español que contenían referencias a la muerte de *George Floyd* y sus consecuencias (revueltas, manifestaciones, detenciones y ataques a entidades patrimoniales, etc.). Este proceso duró en torno a una hora por evaluador. Una vez emitidos los juicios, siguiendo la metodología, se descargó el contenido de todos los hilos que contenían publicaciones con estas características.

3.3. Colección resultante

Después del proceso de definición de consultas y juicio experto, los posts juzgados como positivos trajeron consigo un total de 77 hilos diferentes, con un total de 19303 posts en español en la colección de referencia. Esto nos da una media de 250 posts incorporados a la colección por hilo, recuperando información de *subreddits* de toda Latinoamérica, España y comunidades hispanas en otros países, con al menos 15 países diferentes identificados.

Esta gran cantidad de hilos diferentes asegura que la metodología permite recuperar e incorporar a la colección publicaciones no solo de hilos explícitos sobre el movimiento *Black Lives Matter*, sino también de muchos hilos con diferentes temas centrales que hagan referencia a estos eventos en algún momento. Esto es importante para asegurar la cobertura del tema dentro de la red social, no solo recuperando aquellos hilos de conversación explícitos (que pueden reflejar puntos de vista que están excesivamente polarizados o dirigidos por asociaciones o personas directamente implicadas en los conflictos) sino también llegando a hilos con temáticas diferentes al tema principal, tratando las protestas en algún momento. Estas conversaciones pueden ser menos polarizadas e incluyen ciudadanos con perfiles y responsabilidades más diversos.

Así mismo, los posts integrantes fueron escritos entre el 25 de mayo de 2020 (día del fallecimiento de *George Floyd*) y el 10 de abril del 2021. Nótese que solo fue necesario juzgar los 219 posts del pool

inicial durante la fase de juicio (evitando la anotación completa de la colección) para recuperar casi 1 año de actividad en Reddit de la comunidad hispanohablante acerca de las protestas antirracismo. La colección está disponible para su descarga en este vínculo⁵⁴.

Planeamos en el futuro ampliar la colección presentada con la información de perfil de los usuarios cuyos posts forman parte de la colección, lo que permite una investigación más profunda acerca de perfiles demográficos y/o psico-sociales de usuarios dentro de la colección, pero teniendo en cuenta desde el diseño inicial de la colección protocolos específicos de anonimización y seguridad.

4. Discusión y Conclusiones

La metodología propuesta permite la construcción de colecciones de referencia desde redes sociales como archivos sociales relevantes en humanidades. Se presenta aquí una colección en español desde Reddit que recupera, monitoriza, documenta, conserva y permite el estudio evolutivo en esta red social del fenómeno de protesta antirracismo desencadenado por la muerte de *George Floyd* en mayo de 2020.

Como análisis crítico, cabe discutir varios aspectos de la propuesta (y de la propia colección). En primer lugar, es necesario un estudio en profundidad del nivel de acuerdo entre anotadores en la fase de *pooling* para garantizar cierta ausencia de sesgos en el diseño de consultas y en la fase de juicio posterior. Este aspecto está ampliamente desarrollado en la literatura (Cohen, 1960; Fleiss, Levin and Paik, 1981) y también desde el punto de vista de RI como disciplina (Hripcsak and Rothschild, 2005; Artstein and Poesio, 2008; Artstein, 2017) , por lo que planeamos ampliar el número de anotadores e incorporarlo en el futuro.

En segundo lugar, somos conscientes que el estudio de un fenómeno tan complejo como el racismo y/o el movimiento *Black Lives Matter* van mucho más allá de esta colección que presentamos. También de la necesidad de ampliar a más redes sociales el estudio, y de discutir la representatividad social en las mismas. Sin embargo, creemos que, debido a la generalización de las redes sociales como una fuente adicional en dichos estudios, esta colección (y la metodología que subyace), permite complementar estudios del fenómeno desde diferentes perspectivas (sociológicas, antropológicas, históricas, patrimoniales, de medios de comunicación etc.) con otras fuentes y/o técnicas que los investigadores usen, así como posibilita el uso de técnicas computacionales sobre ella (procesado de lenguaje natural, *text mining* o *Social Network Analysis* (Kelly and Kittle Autry, 2011), entre otros). Un estudio comparativo tanto a nivel metodológico como entre las colecciones resultantes en inglés (Otero,

Martin-Rodilla and Parapar, 2020) y en español (el que aquí se presenta) arrojará en el futuro datos sobre las implicaciones de la lengua empleada para la creación de estas colecciones desde redes sociales.

Destacamos así mismo que la metodología presentada permite replicar el proceso de creación de la colección con las mismas consultas y fuentes en *Reddit*, pudiendo, por ejemplo, ampliar la colección en rangos de fechas para estudiar el fenómeno de manera diacrónica. En el caso específico de *Black Lives Matter*, ciertos eventos como la celebración del juicio por la muerte de *George Floyd* en abril de 2021 pueden hacer resurgir el movimiento y la actividad en redes sociales, por lo que su ampliación puede ser interesante desde el punto de vista investigador.

Es importante también reseñar que, aunque la técnica de *pooling* y la metodología empleada resuelve algunos de los problemas detectados en cuanto a la aplicación de técnicas de recuperación de información en este

⁵⁴ <https://pages.irilab.org/david.otero.freijeiro/hdh2021>

tipo de trabajos, no se pretende con ello automatizar la tarea de crear colecciones de referencia (como archivos nativos digitales) a partir de las redes sociales. Somos conscientes de que las mejoras en el enfoque tecnológico que ofrece deben ser aplicadas desde una perspectiva transdisciplinar, donde sea la colaboración de expertos en el dominio de las colecciones, en archivística y documentación y en recuperación de información la que permita la mejora de este tipo de colecciones.

Bibliografía

- Acker, A. and Brubaker, J. R. (2014). "Death, memorialization, and social media: A platform perspective for personal archives." *Archivaria* 77: 1-23.
- Artstein, R. (2017). "Inter-annotator agreement". In *Handbook of linguistic annotation*, Springer: 297-313.
- Artstein, R. and Poesio, M. (2008). "Inter-coder agreement for computational linguistics." *Computational Linguistics* 34(4): 555-596.
- Barreiro, Á., Parapar, J. and Otero, D. (2019). *Exploiting Pooling Methods for Building Datasets for Novel Tasks*. FDIA@ ESSIR.
- Cohen, J. (1960). "A coefficient of agreement for nominal scales." *Educational and psychological measurement* 20(1): 37-46.
- Fleiss, J. L., Levin, B. and Paik, M. C. (1981). "The measurement of interrater agreement." *Statistical methods for rates and proportions* 2(212-236): 22-23.
- Gehl, R. (2009). "YouTube as archive: Who will curate this digital Wunderkammer?" *International journal of cultural studies* 12(1): 43-60.
- Hripcsak, G. and Rothschild, A. S. (2005). "Agreement, the f-measure, and reliability in information retrieval." *Journal of the American medical informatics association* 12(3): 296-298.
- Kelly, A. R. and Kittle Autry, M. (2011). *A humanistic approach to the study of social media: Combining social network analysis & case study research*. Proceedings of the 29th ACM international conference on Design of communication.
- Losada, D. E., Parapar, J. and Barreiro, A. (2017). "Multi-armed bandits for adjudicating documents in pooling-based evaluation of information retrieval systems." *Information Processing & Management* 53(5): 1005-1025.
- McNealy, J. (2011). "The privacy implications of digital preservation: social media archives and the social networks theory of privacy." *Elon L. Rev.* 3: 133.
- Mordell, D. (2019). "Critical questions for archives as (big) data." *Archivaria* 87(87): 140161.
- Obodoruku, B. (2016). *Social Networking: Information Sharing, Archiving and Privacy*. 24th BOBCATSSS Conference Proceedings & Abstracts.
- Otero, D., Martín-Rodilla, P. and Parapar, J. (2020). "Building Cultural Heritage Reference Collections from Social Media through Pooling Strategies: The case of 2020's tensions over race and heritage. Reference Collection" from <https://pages.irlab.org/david.otero.freijeiro/heritage/>.
- Otero, D., Martín-Rodilla, P. and Parapar, J. (2021). "Building Cultural Heritage Reference Collections from Social Media through Pooling Strategies: The case of 2020's tensions over race and heritage." *Journal on Computing and Cultural Heritage*. In review.

- Parapar, J., Martin-Rodilla, P., Losada, D. E. and Crestani, F. (2021). *eRisk 2021: Pathological Gambling, Self-Harm and Depression Challenges*. 43rd EUROPEAN CONFERENCE ON INFORMATION RETRIEVAL ECIR, Italia.
- Perpinyà-Morera, R. and Cid-Leal, P. (2020). "Memoria de la ciudadanía: preservación de archivos de movimientos sociales y movilizaciones ciudadanas." *El Profesional de la Información* 29(6).
- Pybus, J. (2013). "Social networks and cultural workers: Towards an archive for the prosumer." *Journal of Cultural Economy* 6(2): 137-152.
- Pybus, J. (2015). "15 Accumulating Affect: Social Networks and Their Archives of Feelings." *Networked affect*: 235.
- Reddit. (2020). Retrieved 04/11/2020, from <https://www.reddit.com>.
- Ruiz Gómez, V. and Vallès, A. (2020). "# Cuéntalo: the path between archival activism and the social archive (s)." *Archives and Manuscripts* 48(3): 271-290.
- Sanderson, M. (2010). *Test collection based evaluation of information retrieval systems*, Now Publishers Inc.
- Schefbeck, G., Spiliotopoulos, D. and Risse, T. (2012). "The recent challenge in web archiving: archiving the social web." *context* 7: 9.
- Vallès, A. and Soler Jiménez, J. (2016). "El tractament arxivístic de Twitter."
- Vallès, A. M. (2020, 13 Junio 2020). "Archivando las redes sociales. Una charla sobre datos abiertos y archivos sociales." *Internation Archives Week 2020 IAW2020*, from <https://www.ica.org/es/iaw2020-webinar-archivando-las-redes-sociales-una-charla><https://www.ica.org/es/iaw2020-webinar-archivando-las-redes-sociales-una-charla-sobre-datos-abiertos-y-archivos-sociales>.
- Winters, J. and Prescott, A. (2019). "Negotiating the born-digital: a problem of search." *Archives and Manuscripts* 47(3): 391-403.
- Zubia Muxika, J. J. (2018). *Preservación de información de redes sociales virtuales*. Grado en Información y Documentación, Universitat Oberta de Catalunya.

Digitalización de archivos, bibliotecas y museos

Modera: Nieves Pena Sueiro · *Asiste:* Brianda Otero Moreira

- 58 Bibliotecas digitales de textos griegos y latinos antiguos. Luces y sombras
Cecilia Criado
- 50 El DEMel: Una base de datos para el léxico del español medieval
Caroline Müller
- 13 Inteligencia Artificial y modelos de reconocimiento automático de texto manuscrito: experimentando con textos sobre Patrimonio Cultural en español de inicios del siglo XX
Raquel Licerias Garrido, Alba Comino, Patricia Murrieta-Flores
- 32 La biblioteca digital TECER
Montserrat Ribao Pereira
- 5 PMSierra, una Colección Digital de la Pintura Mural en la comarca de la Sierra de Huelva
Antonio Manuel Cuaresma
- 39 Catalogar el patrimonio histórico-artístico en tiempos de pandemia: Google Maps como herramienta investigadora y docente
Antonio Ledesma, Sergio Sánchez Jiménez

Bibliotecas digitales de textos griegos y latinos antiguos. Luces y sombras

Cecilia Criado cecilia.criado@usc.es

Departamento de Filología Clásica, Francesa e Italiana de la Universidad de Santiago de Compostela

Palabras clave: Filología Clásica; Bibliotecas digitales; Edición electrónica de textos; Codificación TEI XML; Mitotebano

La Filología Clásica fue líder en la digitalización del ingente corpus textual de autores griegos y latinos antiguos, tarea que fue arrojada en los años setenta y ochenta del siglo pasado. Así, el TLG (*Thesaurus Linguae Graecae*) ofreció en soporte magnético los autores latinos en 1975 y, entre 1985 y 1992, en soporte óptico; y PHI (*Packard Humanities Institute*), sacó a la luz el CD-ROM de los autores latinos en 1987. Los contenidos de los CD-ROM de TLG y PHI vertida a soporte electrónico por la [Perseus Digital Library](#) en 1995. La finalidad era divulgar entre un público amplio el acervo literario de la Grecia y Roma antiguas. Puede decirse que, a día de hoy, tal objetivo se ha cumplido con éxito.

Ahora bien, los avances tecnológicos que en los últimos veinte años han revolucionado la edición digital están poniendo en evidencia el precio de esta precocidad. Una de las más importantes limitaciones en la oferta de bibliotecas de textos griegos y latinos existentes es, sin duda, la ausencia de compatibilidad mundial. En cambio, la digitalización de los corpus de las lenguas modernas, por tratarse mayoritariamente de iniciativas más recientes, se está llevando a cabo con lenguajes estándares e incorporando funcionalidades no imaginables cuando los contenidos del TLG y y del PHI fueron editadas en soporte electrónico.

Los primeros pasos en la creación de bibliotecas digitales de autores clásicos Sin afán de exhaustividad, éstas son los más sobresalientes, y también más antiguas, bibliotecas digitales de autores griegos y latinos:

- 1) El [Thesaurus Linguae Graecae](#) (TLG) es un centro de investigación de la Universidad de California, Irvine, dirigido en la actualidad por la prof. M. Pantelia que combina filología y tecnologías de la información. Desde el momento de su fundación en 1972 hasta la actualidad ha digitalizado gran parte de los textos escritos en griego desde el Homero hasta la caída de Bizancio (diez mil obras de cuatro mil autores), siendo su objetivo llegar a la actualidad. En 2001 TLG dejó de comercializar su CD-ROM, siendo hoy día posible acceder a sus contenidos sólo de forma electrónica, aunque no libre. Existe, no obstante, una [versión online abreviada gratuita](#) que, si bien es de gran utilidad para los estudiantes, es de todo punto de vista insuficiente para un investigador de Filología Griega.
- 2) El *Packard Humanities Institute* (PHI) fue creado en 1987 con el objetivo de crear herramientas que facilitasen el acceso a contenidos y a búsquedas temáticas y filológicas en el amplio campo de las Humanidades (historia, literatura y música). Tres de los ámbitos privilegiados fueron la literatura latina, y los papiros e inscripciones griegas. Una parte importantísima de estos textos fueron publicados en soporte magnético. Dada la envergadura de la empresa, desde un principio el proyecto fue concebido como colaborativo, participando en él universidades de prestigio internacional (Yale, Princeton, Duke, Cornell y Michigan). Igual que el TLG, desde 2004 PHI no comercializa el CD-ROM, aunque existe [acceso electrónico a la base de datos](#) de forma gratuita.

- 3) La [Library of Latin Texts](#) (LLT) es un proyecto del *Traditio Litterarum Occidentium* (CTLO), dirigido por P. Tombeur. Dicho centro continúa las actividades que el Cetodoc (Université Catholique de Louvain) había realizado en el campo de los estudios latinos. Constituye el mayor repertorio de textos latinos existente pues contiene la *Bibliotheca Teubneriana Latina* y la *CETEDOC Library of Christian Texts* (CLCLT) Ofrece dos series de distinta entidad. La finalidad de la Serie B es poner a disposición de investigadores y estudiantes el mayor número de textos latinos en formato electrónico. Por mor de una rápida consecución de este objetivo, esta serie extrae su corpus de ediciones ya existentes; esto es, las ediciones no son siempre de referencia y, por supuesto, no tienen aparato crítico (como no lo tienen ninguna de las bibliotecas electrónicas). La verdadera joya del centro es la Serie A, que es una monumental base de datos que desde 1991 ha digitalizado 3.200 obras de 950 autores latinos. En un principio, su corpus se limitaba a la literatura latina cristiana y medieval pero, desde 2002, se están editando los autores antiguos.

La aportación trascendental de la Serie A de la LLT fue que, por primera vez en la red, se realizaron ediciones críticas de los textos latinos (*Bibliotheca scriptorum Romanorum Teubneriana*). Con todo, la biblioteca no ofrece acceso libre.

- 4) Frente a LLT y al TLG online, [Perseus Digital Library](#) es de acceso gratuito. Dirigido desde 1987 por el prof. G.R. Crane (Tufts University –Boston–), el *Perseus Project* ha creado una biblioteca online que ha ido digitalizando una ingente cantidad de obras [de distintos ámbitos de las Humanidades](#), cuyos derechos de autor habían expirado. En lo que se refiere estrictamente a su colección de [Greek and Roman Materials](#), los textos han sido vertidos desde los antiguos soportes magnéticos del TLG y del PHI. Son muchas las funcionalidades que la plataforma ofrece: lectura lineal y radial, traducciones en paralelo, búsquedas y enlaces a múltiples recursos. No obstante, adolece de importantes defectos entre los que cabe destacar las numerosas erratas derivadas de una deficiente corrección tras practicar la OCR a los textos escaneados; las ediciones y traducciones son muy antiguas, esto es, no son estándares con lo cual su empleo con fines investigadores es desaconsejable; su motor de búsquedas es precario y su funcionalidad, muy deficiente cuando se realizan búsquedas complejas o, concretamente, de nombres propios, búsqueda imprescindible para el estudioso de literatura; y, por último, no son ediciones críticas.

Desde finales de la primera década de este siglo, se han creado diversas plataformas, aún en desarrollo, que corrigen estas deficiencias y que, además, incorporan el lenguaje de marcado TEI XML. Me limitaré a citar [Perseus Digital Library](#). [Scaife Viewer](#) (que mejora sustancialmente el entorno de lectura y las búsquedas de *Perseus* aunque el etiquetado de los textos, al menos a día de hoy, es muy somero, limitándose a título, autor, número de libros y números de versos) y [Perseus under PhiloLogic](#) (también filial del *Perseus Project*, está sometiendo las bases de datos documentales a un complejo sistema de codificación y lematización desarrollado por el proyecto [ARTFL](#)). Una y otra biblioteca, sin embargo, se nutren de los textos de *Perseus Project* con lo cual persiste su principal defecto, la baja calidad de las ediciones.

Un paso más: bibliotecas con codificación TEI XML que no utilizan los textos del Perseus Project

La comunidad internacional de investigación en Filología Clásica es muy consciente de la necesidad de que la oferta libre de textos antiguos esté en consonancia con lo que lo que la ecdótica y la alta investigación científica en Filología requieren. En lo que toca estrictamente a textos literarios, son dos los proyectos que están dando importantes pasos en este sentido:

- 1) [Musisque Deoque](#) inició su andadura de forma modesta en 2005. Su finalidad era y sigue siendo crear un archivo digital de poesía latina y neolatina al que paulatinamente se va incorporando el aparato crítico. Las ediciones, marcadas en TEI XML, son de referencia; así, por citar un ejemplo, el texto de la *Eneida* es el de Geymonat 1973/2008, frente al texto de J. B. Greenough 1900 que ofrece el *Perseus Project* y la práctica totalidad de bibliotecas electrónicas de acceso abierto. El proyecto es colaborativo y en él participan las universidades italianas de Venecia, Calabria, Padua, Parma, Perugia y Federico II de

Nápoles. Su desarrollo en los últimos cinco años ha sido extraordinario y a él se han incorporado otros proyectos interdisciplinarios que permiten nuevas funciones: Epigraphica, Metrical typology y Hellenica. Es una herramienta esencial para la docencia universitaria. Como herramienta de investigación, sin embargo, adolece de mucha irregularidad en la calidad de las ediciones críticas. Asimismo, al haber utilizado un set reducido de etiquetas TEI, la posibilidad de visibilizar toda la “cultura” de los textos es limitada.

- 2) [Digital Latin Library](#) (DLL) es un proyecto ambicioso dirigido por S. Huskey (University of Oklahoma) con el patrocinio y participación de la *Society for Classical Studies* (antigua *American Philological Classical Association*), la *Medieval Academy of America* y la *Renaissance Society of America*. Dio sus primeros pasos en 2012 y hasta 2018 desarrolló un gran trabajo en su [Catalog](#). Su [Library of Digital Latin Texts](#) está en progreso; por ello, sólo hay una demostración de su método editorial, las *Bucólicas* de Calpurnio Sículo (ed. C. Giarratano 1910), codificada por S. Huskey y H. Cayless. No obstante, esta edición piloto sirve para constatar la competencia en la codificación TEI XML alcanzada y la investigación tecnológica que llevaron a cabo entre los años 2012 y 2018.

La aportación de la biblioteca digital *Thebarum fabula*

La biblioteca [Thebarum fabula](#) está siendo desarrollada por el equipo de investigación *Oedipodioniae Thebae. Bibliotheca grecorromana sobre la guerra civil* (FFI2015-68599R MINECO/FEDER), dirigido por Cecilia Criado (Universidad de Santiago de Compostela), con Sandra Romano (Universidad Autónoma de Madrid) como directora técnica y Neil Coffee (University at Buffalo) como asesor consultivo externo. Es una biblioteca temática que incluye las obras griegas y latinas que versan sobre el mito tebano. Es una biblioteca comprensiva y un proyecto de largo recorrido que, en fases posteriores, incorporará las obras del resto de las sagas míticas.

El proyecto surgió de la convicción de que, una vez conseguido el objetivo de que una gran parte de los textos clásicos son de libre acceso en la red, la Filología Clásica debe asumir la responsabilidad de que la presencia de estos textos en Internet sea menos puramente informativa o divulgativa, y adopte una perspectiva más académica, dirigida a usuarios especializados.

Esto exige la mejora en la calidad de las ediciones y en el sistema de codificación de los textos, que en *Thebarum fabula* realizamos enteramente con TEI-XML y una base de datos documental. Las ediciones están publicadas en un aplicación de software libre y pueden ser descargadas en formato TEI-XML (en un futuro, también será posible su descarga en formato PDF e EPUB). Son ediciones críticas a las que se les están incorporando notas, comentarios y traducciones a distintas lenguas modernas alineadas respecto al texto original.

Los aparatos críticos de cada obra están basados en varias ediciones estándares de reconocida solvencia que el miembro del equipo de investigación modifica en virtud de sus criterios editoriales. La metodología, por tanto, coincide con la de los ya citados *Musique deoque* y *DLL*, así como con la de otros proyectos de similares características que están siendo desarrollados en la actualidad, aunque su objetivo no es la creación de una biblioteca sino la edición original de un único autor u obra.

En ámbito español, *Thebarum fabula* es una iniciativa pionera en el mundo de la Filología Clásica. Según creo que podré mostrar en la comunicación, el set de etiquetas TEI con el que trabajamos es mucho más extenso que el utilizado por las bibliotecas que han sido mencionadas *supra* como las más solventes. Ello queda patente, precisamente, en la atención que nuestras ediciones prestan a las etiquetas de los aparatos críticos. Al no codificarlos con la etiqueta <note>, práctica habitual en la oferta existente, los aparatos pueden dar cuenta cumplidamente de toda la “cultura” que a ellos subyace, aparte de posibilitar su óptima visualización web.

El DEMel: Una base de datos para el léxico del español medieval

Caroline Müller (Universidad de Rostock)

Palabras clave: Digitalización / Base de datos / Español medieval / Lexicografía

El proyecto Diccionario del Español Medieval electrónico (DEMel), apoyado por la DFG (Fundación Alemana para la Investigación Científica), tiene como objetivo digitalizar el extenso archivo del Diccionario del Español Medieval y publicarlo en la web. Con este propósito, el proyecto DEMel pretende transformar este archivo en papel no solo en un fichero digitalizado, sino también en una base de datos lematizada y cronológicamente estructurada que se podrá consultar a través de una aplicación web.

Después de haber publicado 26 fascículos (a-almohatac), el proyecto Diccionario del Español Medieval (DEM) terminó en el año 2007 por razones financieras. Aunque una gran parte del fichero haya quedado inacabada debido al repentino final del proyecto, el archivo constituye una valiosa fuente para el estudio del español medieval: se basa en un corpus de más de 600 obras o colecciones de documentos tanto literarias como jurídicas, religiosas o científicas del siglo X hasta inicios del siglo XV. El fichero llamado 'nuclear' consta de unos 700 000 fichas que muestran documentaciones de los más de 31 000 lemas. Además, el archivo dispone de unos 161 000 fichas con información adicional, por ejemplo, etimológica, bibliográfica o toponímica. Después de la digitalización de todas estas fichas y su publicación en la web gracias al proyecto DEMel, la comunidad científica internacional tendrá acceso libre a este extenso archivo⁵⁵.

En la presente comunicación describiremos las diferentes etapas del proyecto, la mayoría de las cuales ya han terminado. Después del escaneo de todas las fichas por una empresa especializada en grandes proyectos de digitalización, se usó el sistema de digitalización de la Biblioteca Universitaria de Rostock para clasificar las imágenes de las fichas según sus lemas. A continuación, las imágenes se archivaron junto con sus metadatos en el repositorio propio de la biblioteca y se implementó un modelo relacional en un sistema gestor de bases de datos MySQL para registrar los demás datos de las fichas. A causa de la heterogeneidad del material, no resultó posible descifrar los datos automáticamente, por lo que se elaboró un editor con una interfaz de entrada, con el cual el equipo filológico está registrando los datos más importantes de las fichas, como la forma documentada de la palabra medieval y la fuente de la documentación, en la base de datos. Con el registro de los datos ya casi terminado, el desarrollo de la aplicación web está en pleno proceso. Al describir estas etapas, también entraremos en los retos que hemos afrontado durante el proyecto.

A continuación, nos centraremos en la aplicación web. Explicaremos las opciones de consulta que se ofrecerán a los usuarios y que permitirán investigaciones exhaustivas y precisas en el abundante material lexicográfico⁵⁶. De este modo, se pondrá a disposición del público el valioso archivo del DEM facilitando un nuevo entorno de investigación no solo para los hispanistas, sino, gracias a la gran diversidad de textos, también para investigadores de otras disciplinas como los estudios históricos, árabes o teológicos. Aunque el DEMel no proporcionará artículos lexicográficos en sentido estricto, la riqueza documental ofrecida por la base de datos contribuirá al conocimiento más profundo de la historia del léxico español.

Para terminar mostraremos las perspectivas de una posible evolución del DEMel en el futuro. De acuerdo con la demanda a la interoperabilidad en las Humanidades Digitales, deseamos, por ejemplo, conectar el DEMel con

⁵⁵ Para más información acerca del archivo y de los orígenes del proyecto, véase Arnold et al. (2016 y 2018), Franke / Serafin (2018) y Arnold (2020).

⁵⁶ Para más información acerca de la aplicación web, véase Müller / Reiter (2021).

otras herramientas de la lexicografía histórica con el fin de facilitar el estudio del español medieval a través de la web.

La aplicación web (versión preliminar)

Bibliografía

Arnold, Rafael (2020): "La digitalización del fichero del "Diccionario del español medieval" (DEM)", en: *Scriptum digital. Revista de corpus diacrònics i edició digital en Llengües iberoromàniques*, [en línia], Núm. 9, 191–207, <https://www.raco.cat/index.php/scriptumdigital/article/view/377294> [Consulta: 13-04-2021].

Arnold, Rafael / Langenbacher-Liebgtott, Jutta / Serafin, Stefan / Álvarez Vives, Vicente (2016): "El Diccionario del Español Medieval electrónico", en: Domínguez-Rodríguez, María Victoria et al. (eds.): *Words across History: Advances in Historical Lexicography and Lexicology*. Las Palmas de Gran Canaria: Servicio de Publicaciones y Difusión Científica, 30–39.

Arnold, Rafael / Langenbacher-Liebgtott, Jutta / Franke, Anna-Susan / Serafin, Stefan (2018): "Una nueva fuente para la historia del léxico español: el DEMel", en: Corbella, Dolores / Fajardo, Alejandro / Langenbacher-Liebgtott, Jutta (eds.): *Historia del léxico español y Humanidades digitales*. Berlin: Lang, 165–188.

Franke, Anna-Susan / Serafin, Stefan (2018): "Historisches Sprachmaterial modern aufbereitet: Von den Ursprüngen der historischen Lexikographie des Spanischen zu den Herausforderungen im digitalen Zeitalter", en: Franke, Anna-Susan / Álvarez Vives, Vicente (eds.): *Romania Pontes. Beiträge zur Sprache in der Gallo- und Iberoromania*. Berlin: Lang, 111–142.

Müller, Bodo (1987–2005): *Diccionario del español medieval*, vol. 1, fascículos 1–10, vol. 2, fascículos 11–20, vol. 3, fascículos 21–26. Heidelberg: Winter.

Müller, Caroline / Reiter, Martin (2021): "El DEMel – Del fichero en papel a la publicación de acceso abierto en la web", en: Schøsler, Lene / Härmä, Juhani / Lindschouw, Jan: *Actes du XXIXe Congrès international de*

linguistique et de philologie romanes (Copenhague, 1-6 juillet 2019). Strasbourg: Éditions de linguistique et de philologie, 829–840.

Inteligencia Artificial y modelos de reconocimiento automático de texto manuscrito: experimentando con textos sobre Patrimonio Cultural en español de inicios del siglo XX

Raquel Licerias-Garrido (Universidad de Lancaster, Reino Unido)

Alba Comino (investigadora independiente)

Patricia Murrieta-Flores (Universidad de Lancaster, Reino Unido)

Palabras clave

Texto Manuscrito, Español de inicios del siglo XX, Inteligencia Artificial, Transcripción Automática, Transkribus

Descripción del problema de investigación, metodología y resultados obtenidos

La Arqueología tiene una trayectoria de más de un siglo como disciplina científica y, además, ha encabezado la colaboración entre las Humanidades y las Ciencias de la Computación. Sin embargo, la aplicación potencial de técnicas digitales a los recursos textuales arqueológicos producidos durante este tiempo permanece en gran parte inexplorada. Del mismo modo, los métodos y análisis de la Lingüística Computacional, o aquellos derivados de la Inteligencia Artificial como son el Procesamiento del Lenguaje Natural o la Minería de Textos siguen siendo extremadamente escasos en su aplicación a la Arqueología (Amrani *et al.* 2008; Murrieta-Flores y Gregory 2015; Richards *et al.* 2015; Licerias-Garrido *et al.* 2019). Para implementar estos análisis es imprescindible contar con texto en un formato legible por computadora. Afortunadamente, en las últimas décadas, se han llevado a cabo importantes proyectos de digitalización, especialmente en bibliotecas, archivos y museos, que han escaneado los documentos originales y han compartido abiertamente en formato de imagen una enorme cantidad de datos que posibilitan los enfoques con acercamientos de tipo Big Data.

En el campo de la Inteligencia Artificial, el aprendizaje profundo (*Deep Learning*) ha evolucionado como una categoría avanzada del aprendizaje automático (*Machine Learning*) impulsando mejoras sustanciales mediante el desarrollo de métodos innovadores para la identificación y clasificación de imágenes. Las redes neuronales son un componente fundamental de cualquier arquitectura de aprendizaje profundo. Basándose en redes neuronales recurrentes y convolucionales, el reconocimiento de texto manuscrito (*Handwritten Text Recognition*) es una tecnología que permite a la computadora interpretar entradas escritas a mano a partir de imágenes y producir texto codificado de alta calidad.

En los últimos años, los programas de reconocimiento de texto manuscrito han proliferado tanto en formato abierto como propietario. Una de las herramientas en línea más utilizada en investigación es Transkribus, desarrollada por los proyectos europeos tranScriptorium (2013-2015) y READ (2016-2019). Transkribus utiliza una combinación de redes neuronales recurrentes y convolucionales para reconocer la escritura a mano usando una cantidad limitada de datos de entrenamiento (Kahle *et al.*, 2017; Muehlberger *et al.*, 2019). Este trabajo utiliza la tecnología de reconocimiento de texto manuscrito integrada en Transkribus para explorar las

posibilidades de producir transcripciones automáticas de tres volúmenes manuscritos del Catálogo Monumental de España (CME), creado a principios del siglo XX.

El CME es una fuente fundamental en lengua española para disciplinas como la Arqueología, la Historia, el Patrimonio, las Artes y la Arquitectura, entre otras, ya que enumera y describe los principales edificios, monumentos y antigüedades de España a inicios del siglo XX, e incluye datos únicos e inéditos para la investigación arqueológica. Este proyecto experimenta con los volúmenes de la provincia de Ávila (3 tomos), escritos por Manuel Gómez-Moreno (1900-1901), de Soria (8 tomos), elaborados por Juan Cabré (1911-1917), y de Burgos (7 tomos), recopilados por Narciso Sentenach (1921-1924). El corpus inicial es una versión en imagen digital de los manuscritos, disponible en la Biblioteca Virtual Tomás Navarro Tomás del Consejo Superior de Investigaciones Científicas y libre de derechos de autor.

Para llevar a cabo el proceso de transcripción automática, se diseñaron dos estrategias principales. En primer lugar, se crearon dos modelos personalizados de reconocimiento de texto manuscrito basados en los mayores conjuntos de datos: Cabré y Sentenach. En segundo lugar, a partir estos modelos, se produjo otro que combinaba la caligrafía de Cabré y de Sentenach para realizar transcripciones automáticas de los libros de Gómez-Moreno. Esta segunda estrategia tenía como objetivo probar modelos creados con datos de entrenamiento de diferentes caligrafías para evaluar su rendimiento utilizando datos que no estuviesen incluidos en el entrenamiento del modelo.

Los principales resultados de este trabajo han sido:

- La creación de tres modelos para transcribir los textos manuscritos, disponibles en la plataforma Transkribus, con una precisión de en torno al 95% en cada uno de los conjuntos de datos.
- La transcripción de los 18 tomos correspondientes a las provincias de Ávila, Soria y Burgos, disponibles en los repositorios abiertos de [figshare](#) y [github](#) (LiceriasGarrido *et al.* 2020 a, b, c).
- Un estudio historiográfico que ha permitido comprobar que la letra del catálogo de Ávila elaborado por Manuel Gómez-Moreno correspondía en realidad a su hermana Sagrario Gómez-Moreno, encargada de pasar a limpio la información.
- Un análisis lingüístico computacional de las narrativas de la Prehistoria y la Edad del Hierro en clave de género que permite identificar los patrones, sesgos y estereotipos en los discursos sobre el pasado construidos por autores españoles de principios del siglo XX.

Bibliografía

Amrani, A., Abajian, V., Kodratoff, Y. and Matte-Tailliez, O. "A Chain of Text-mining to Extract Information in Archaeology", *3rd International Conference on Information and Communication Technologies: From Theory to Applications*, Institute of Electrical and Electronics Engineers Computer Society, Piscataway (2008), pp. 1-5 <https://doi.org/10.1109/ICTTA.2008.4529905>.

Kahle, P. Colutto, S., Hackl and G., Mühlberger, G. "Transkribus - A Service Platform for Transcription, Recognition and Retrieval of Historical Documents", *14th IAPR International Conference on Document Analysis and Recognition 4*, Institute of

Electrical and Electronics Engineers Computer Society, Los Alamitos (2017), pp. 19-

24. <https://doi.org/10.1109/ICDAR.2017.307>

Licerias Garrido, R.; Comino, A. y Murrieta Flores, P. (2020a): *Transcripción del Catálogo Monumental de España: Provincia de Ávila por Manuel Gómez Moreno (1900-1901)*. Figshare. Dataset. <https://doi.org/10.6084/m9.figshare.12006318.v1>.

- Liceras Garrido, R.; Comino, A. y Murrieta Flores, P. (2020b): *Transcripción del Catálogo Monumental de la Provincia de Soria por Juan Cabré (1916-1917)*. Figshare. Dataset. <https://doi.org/10.6084/m9.figshare.12006273.v1>.
- Liceras Garrido, R.; Comino, A. y Murrieta Flores, P. (2020c): *Transcripción del Catálogo Monumental y Artístico de la Provincia de Burgos por Narciso Sentenach (1925)*. Figshare 2020. Dataset. <https://doi.org/10.6084/m9.figshare.12006327.v1>
- Müehlberger, G., Seaward, L., Terras, M., Ares Oliveira, S., Bosch, V., Bryan, M., Colutto, S., Déjean, H., Diem, M., Fiel, S., Gatos, B., Greinoecker, A., Grüning, T., Hackl, G., Haukkovaara, V., Heyer, G., Hirvonen, L., Hodel, T., Jokinen, M., Kahle, P., Kallio, M., Kaplan, F., Kleber, F., Labahn, R., Lang, E. M., Laube, S., Leifert, G., Louloudis, G., McNicholl, R., Meunier, J., Michael, J., Mühlbauer, E., Philipp, N., Pratikakis, I., Puigcerver Pérez, J., Putz, H., Retsinas, G., Romero, V., Sablatnig, R., Sánchez, J.A., Schofield, P., Sfikas, G., Sieber, C., Stamatopoulos, N., Strauß, T., Terbul, T., Toselli, A. H., Ulreich, B., Villegas, M., Vidal, E., Walcher, J.,
- Weidemann, M., Wurster, H. and Zagoris, K. "Transforming scholarship in the archives through handwritten text recognition: Transkribus as a case study", *Journal of Documentation*, 75.5 (2019): 954-976. <https://doi.org/10.1108/JD-07-2018-0114>
- Murrieta-Flores, P. and Gregory, I. "Further frontiers in GIS: Extending Spatial Analysis to Textual Sources in Archaeology", *Open Archaeology*, 1 (2015): 166-175. <https://doi.org/10.1515/opar-2015-0010>
- Liceras-Garrido, R., Favila-Vázquez, M. Bellamy, K., Murrieta-Flores, P., Jiménez Badillo D. and Martins, B. "Digital Approaches to Historical Archaeology: Exploring the Geographies of 16th Century New Spain", *Open Access Journal of Archaeology and Anthropology*, 2.1 (2019), <http://dx.doi.org/10.33552/OAJAA.2019.02.000526> Richards, J., Tudhope, D. and Vlachidis, A. "Text Mining in Archaeology: Extracting Information from Archaeological Reports". In J. Barceló and I. Bogdanovic (eds), *Mathematics and Archaeology*. CRC Press-Taylor & Francis Group, Boca Raton (2015), pp. 240-254.

La Biblioteca Digital TECER

Montserrat Ribao Pereira

Universidad de Vigo

Palabras clave: base de datos, biblioteca digital, literatura medieval, literatura contemporánea.

Tecer es un verbo transitivo gallego (“*enlazar unha serie de fíos paralelos con outra serie colocada perpendicularmente, formando un tecido ou unha tea*”; *Diccionario da Real Academia Galega*).

TECER (TExtos, Contextos, Ecos y Relecturas) es la Biblioteca Digital creada por el grupo de investigación NOCELBA para abordar las relecturas y reescrituras (literarias, artísticas, audiovisuales) de que han sido objeto los textos del *Cancionero de Baena* y su contexto, desde la composición del mismo hasta la actualidad, con el fin de favorecer el estudio del delicado y complejo tejido que conforma la primera poesía cancioneril castellana desde una perspectiva diacrónica.

Si TECER es el acrónimo de esta Biblioteca Digital, su icono es una “S” inicial del manuscrito del *Cancionero de Baena*, en concreto la que encabeza los *Proverbios* de Santillana, originalmente con siete líneas de altura y un trazo ornamental que separa los dos elementos de la banda roja” (Faulhaber y Perea 2018: 32-33)-

El grupo de investigación NOCELBA está dirigido por el profesor Antonio Chas Aguión (Universidad de Vigo). Desde 2009 este grupo estudia, por una parte, la poesía de Cancionero en el tiempo de los primeros Trastámara castellanos y, en paralelo, los ecos y revisiones a que dicha poesía y su entorno cortesano, en sentido amplio, da lugar a partir del siglo XVI. Esta segunda vertiente de la investigación, al amparo de diferentes subvenciones de la Universidad de Vigo y avalada por diversos proyectos competitivos de la Xunta de Galicia y del Plan Nacional de Investigación, es el germen de [TECER Biblioteca Digital](#).

NOCELBA está integrado en la actualidad por nueve investigadores/as de tres universidades españolas (Universidad de Vigo, CEU, Universidad de León) y cinco europeas (Universidades de Udine, Verona, Rennes II, Lisboa y Northumbria), además de personal en formación de la Universidad de Vigo. Este es el núcleo académico que trabaja en los contenidos de [la Biblioteca Digital TECER](#), coordinada por Montserrat Ribao Pereira (Universidad de Vigo). La empresa de diseño y programación IntroVisual (Oviedo) ha creado el soporte informático y la apariencia del sitio de acuerdo con las necesidades específicas del proyecto. La tecnología de UVigo-TV (Universidad de Vigo) hace posibles los recursos polimedia propios que alberga la biblioteca; dos traductoras de lengua española de signos trabajan con nosotros para hacer accesibles dichos recursos a los colectivos con dificultades auditivas.

El sitio [TECER](#) nace en 2020 y se encuentra ahora mismo en su primera fase de expansión. Permitirá en el futuro acceder a una base de datos de las reescrituras y relecturas (literarias, iconográficas y audiovisuales) de que son objeto los textos y el contexto del tiempo de Enrique II, Juan I, Enrique III y Juan II, desde 1368 (instauración de la dinastía Trastámara en Castilla) hasta 1454 (muerte de Juan II). Actualmente [TECER](#) se centra en la literatura de los siglos XVIII y XIX que reescribe el tiempo de Juan II y Enrique III; en fases posteriores ampliaremos el marco cronológico hasta completar el panorama de la reescritura moderna y contemporánea de los cuatro monarcas.

El diseño de las páginas es funcional. En todas figura el logo y un menú superior con los accesos informativos habituales (quiénes somos, objetivos, cómo buscar), contacto, enlace a la web del grupo de investigación [NOCELBA](#) y a redes sociales, opción de idioma y acceso a la base de datos.

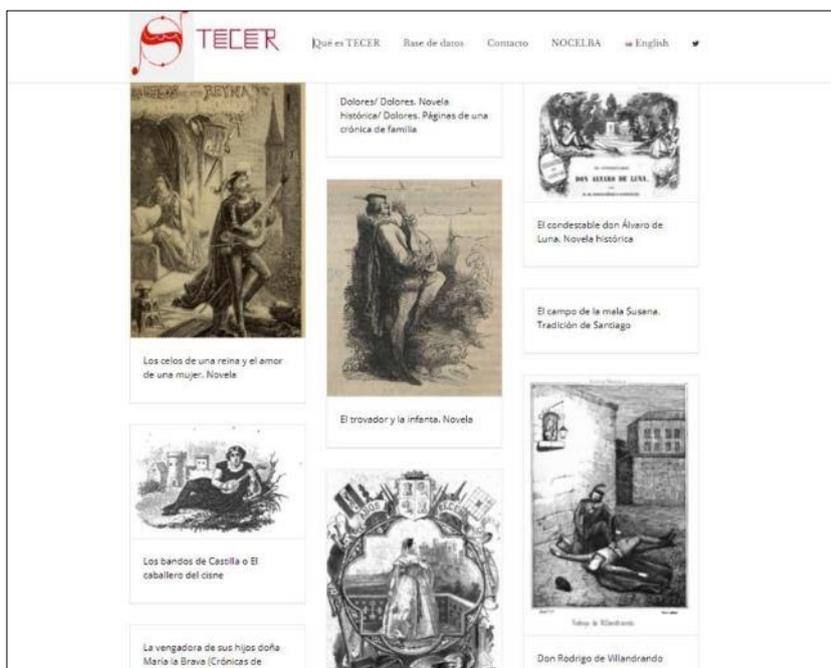
Los contenidos que alberga [TECER](#) son de diferente naturaleza. Desde la página principal es posible acceder a los cuatro bloques principales:

- a) Autores/as.
- b) Títulos.
- c) Recursos:
 - c.1. Enlaces a estudios específicos. Son producto, mayoritariamente, de la investigación de los miembros del grupo.
 - c.2. Enlaces a bibliografía complementaria.
 - c.3. Ilustraciones literarias de los títulos: galería de las imágenes que acompañan la edición de los textos en sus diversos formatos: libro, prensa, pliego...
- d) Mediateca. Píldoras polimedia creadas por las investigadoras del grupo con la tecnología de UVigo-TV, en español y en lengua española de signos.



En estas páginas el scroll es el mínimo necesario para visualizar los logos de las entidades financiadoras (Ministerio y Universidad), los enlaces al grupo y a la Universidad, y el pie de página. En él aparece el aviso legal y una referencia a la empresa de diseño, así como el ISSN en este momento en trámite.

La aportación más relevante de nuestra Biblioteca Digital es su base de datos, a la que puede accederse de dos modos desde el menú superior. Un buscador simple a partir de un mínimo de tres caracteres permite localizar cualquier contenido de la base que albergue esa secuencia, al margen de mayúsculas/minúsculas o tildes. El resultado de esta búsqueda es un panel de obras en cuya ficha aparece la secuencia en cuestión; el título o la imagen son hipervínculos para acceder a la ficha de informe.



La búsqueda avanzada posibilita cruzar datos de cuatro parámetros: autor, género (novela, relato, poema, teatro, otros), época (Edad Media, Edad Moderna, Edad Contemporánea) y título. El resultado, en esta ocasión, es un listado alfabético de títulos que se ajustan a los parámetros seleccionados, que funcionan como hipervínculos a sus correspondientes fichas.

The screenshot shows the 'Buscador avanzado de obras' (Advanced search for works) form. It features a header 'BUSQUEDA AVANZADA' and four input fields: 'Autor', 'Género', 'Época', and 'Título'. Below the fields are 'Buscar' and 'Resetear' buttons. A prominent red button labeled 'VER LISTADO DE ILUSTRACIONES' is positioned below the search form.

Las fichas-informe de cada obra contienen, además de los datos básicos (autor, género, época) una descripción de su soporte de difusión (edición o ediciones en libro o prensa), breve sinopsis o valoración crítica, personajes principales y recursos. Este último apartado ofrece:

- a) El enlace al texto original, tanto a la primera como a las sucesivas ediciones si las hubiese. Cuando los textos están digitalizados en sitios estables (Biblioteca/Hemeroteca Nacional de España, bibliotecas

virtuales de las diferentes comunidades autónomas, Biblioteca Histórica Municipal de Madrid, Biblioteca Virtual Miguel de Cervantes...), o en Google Books, el link remite a ellos. También hemos digitalizado fondos propios y, en otros casos, presentamos nuestra propia edición de los textos, derivada de la investigación del grupo.

- b) Estudios: enlace a una selección de trabajos específicos sobre el título o el autor; son, fundamentalmente, estudios propios de NOCELBA. Estos trabajos se almacenan en el apartado correspondiente de "Recursos", accesible desde la página principal.
- c) Media: píldoras polimedia de creación propia sobre el autor, la obra o su contexto.
- d) Enlaces externos: documentos, estudios, vídeos, material gráfico... a propósito del autor, la obra o su contexto.
- e) Galería de imágenes: texto icónico que acompaña la publicación (libro, prensa, pliego). Es posible acceder a las galerías de imágenes, ordenadas alfabéticamente, desde el apartado correspondiente en "Recursos".

Doce años de regencia (crónica del siglo XV)

Autor: Blanch e Illa, Narciso

Título: Doce años de regencia (crónica del siglo XV)

Género: novela

Época: Edad Contemporánea

Descripción. Publicado con grabados en Barcelona, Corominas/Espasa; Madrid, Librería Española; Gerona, Librería Gerundense, 1868. Es una novela histórica protagonizada por Catalina de Lancaster, regente de Castilla (1406-1418), y por su válida Leonor López de Córdoba. Se trata de una obra muy bien documentada, que sigue al pie de la letra los episodios de la regencia referidos por la historiografía medieval disponible en el siglo XIX, a la que se suma una intriga romancesca atractiva y bien trabada con la trama histórica. El resultado es un texto de contenido político que, como su autor, busca en el ideario carlista respuestas a la convulsa situación de España en vísperas de la revolución del 68. Se publica con ocho láminas de Planas y Vázquez.

Personajes clave: Catalina de Lancaster. Leonor López de Córdoba. Inés de Torres. Ruy López de Álalos. Álvaro de Luna. Fernando de Antequera.

Recursos.



- **Texto** (Google Libros).
- **Estudios:** Ribao Pereira 2019c.
- **Media:**
 - Ribao Pereira, Montserrat. *Novela popular, folletín y entrega.*
 - Ribao Pereira, Montserrat. *Los escritores románticos y el tiempo de Juan II.*
 - Ribao Pereira, Montserrat. *El interés romántico por los primeros Trastámara castellanos.*
- **Enlaces externos:**
 - Fundación Tatiana Pérez de Guzmán el Bueno. Ciclo de conferencias de Historia -Reinas, Damas e Infantas durante la dinastía Trastámara-. Oscar Perea Rodríguez: «**Leonor López de Córdoba e Inés de Torres**».
 - Revista de Historia. *Catalina de Lancaster, reina de Castilla.*
 - RTVE. Mujeres en la Historia. *Catalina de Lancaster.*
 - RTVE. Mujeres en la Historia. *Leonor López de Córdoba.*






Desde el punto de vista técnico el desarrollo programativo se ha basado en el lenguaje PHP y maquetación mediante HTML5 y hojas de estilo CSS3. Mediante los mismos se consigue una buena visualización en todo tipo de resolución de pantalla, así como en sus versiones responsivas. Se ha tenido en cuenta los protocolos estándares de usabilidad y accesibilidad web según W3C de tipo AA. Se ha buscado que cualquier usuario, con independencia de cualquier tipo de problemas cognitivos, así como los dispositivos y/o conexiones de que pueda disponer, acceda y utilice la aplicación con la mayor facilidad posible. Se ha puesto el foco también en la optimización de velocidad y de gestión del servidor y de la propia navegación con énfasis en aspectos de

posicionamiento orgánico o natural SEO. El CMS utilizado es portable, se trata de tecnología wordpress con un diseño y una programación a medida. Existe un nivel de usuario administrador para la gestión de las ediciones de contenidos que facilita cualquier gestión desde su panel privado. Las conexiones en servidor para recogida de información en base de datos MySQL se efectúan con llamadas en el mismo servidor y una gestión PHPMyAdmin.

Formalmente la web es sencilla, limpia e intuitiva, pensada para facilitar el acceso a sus materiales a diversos perfiles de usuario. Su diseño responsivo permite la correcta visualización en todo tipo de dispositivos. El signado de las píldoras polimedia, por su parte, acerca los contenidos orales de creación propia a los colectivos con dificultades auditivas.

La Biblioteca Digital da visibilidad a contenidos poco estudiados, dispersos y de difícil acceso. Ofrece al usuario general un abanico de recursos asociados a los textos que favorecen su comprensión y una red de relaciones literarias e históricas que redimensionan la pertinencia, en el momento de su creación, de títulos y autores/as hoy al margen del canon. Quien accede desde un punto de partida no especializado, elige qué autor, qué grupo de leyendas, qué período cronológico, que temática novelesca o qué recurrencias históricas quiere rastrear. En palabras de Mainer Blanco (2019: 202), “acercamos los contenidos a un público de masas mediante información personalizada”. Para los estudiosos y los investigadores TECER es, además, un instrumento que simplifica la búsqueda de bibliografía primaria y secundaria y ofrece información propia al respecto, rigurosa, exhaustiva y accesible en línea, sobre la reescritura de un período muy concreto, que puede servir a su vez como punto de partida a nuevas investigaciones en diferentes líneas de trabajo, tal y como demuestra la investigación predoctoral que nuestro personal en formación comienza a realizar, pese a la juventud de la Biblioteca.

La Biblioteca Digital TECER pone al alcance de los interesados en la literatura medieval castellana, o moderna y contemporánea de contenido histórico, una parte importante de los resultados de investigación del grupo que la nutre. NOCELBA devuelve así a la sociedad los frutos de su trabajo en la investigación pública. Y lo hace contribuyendo a redefinir la metodología de acercamiento a las disciplinas humanísticas, que debe aunar el rigor científico, crítico y epistemológico con las herramientas tecnológicas que permitan síntesis pertinentes y productivas de conocimiento, un saber, en definitiva, que sin el soporte tecnológico tardaríamos en conseguir o conseguiríamos parcialmente. La perspectiva tecnológica, sin embargo, no debe desdibujar, a nuestro juicio, el peso de los conocimientos literarios que sustentan la herramienta, “y que esto sea un valor y no un demérito” (Santos Zas 2021: 159).

TECER no es un repositorio de recursos, sino una herramienta digital de tipo relacional, en proceso de creación, que interconecta las ediciones de textos y títulos poco conocidos y menos estudiados, en sus diferentes soportes, con sus ilustraciones, su contexto de creación y difusión, y los materiales bibliográficos y audiovisuales imprescindibles para su comprensión. Este planteamiento abre vías nuevas, libres y accesibles, al estudio de la reescritura moderna y contemporánea de la Edad Media, al tiempo que favorece el acercamiento diacrónico a sus textos, a sus contextos y a sus diversas reelaboraciones.

Bibliografía

Faulhaber, Charles B. y Perea Rodríguez, Óscar (2018), “El manuscrito del *Cancionero de Baena* (PN1): descripción codicológica y evolución histórica”, *Magnificat Cultura i Literatura Medievals*, 5, 19-51. <https://doi.org/10.7203/MCLM.5.10664>

Mainer Blanco, Belén (2019), "Continuidad del legendario hispánico en la cultura emergente actual", en Pilar Vega Rodríguez y Belén Mainer Blanco, coords., *Lecturas del pasado. Política y usos culturales de la leyenda literaria*, Madrid, Iberoamericana-Vervuert, 201-231.

Santos Zas, Margarita y Carmen E. Vílchez Ruiz (2021), "Todo Valle-Inclán en red: el archivo digital GIVIUS", *Signa*, 30, 139-162.

PMSierra, una Colección Digital de la Pintura Mural en la comarca de la Sierra de Huelva

Antonio Manuel Cuaresma Maestre. Universidad Pablo de Olavide.

Palabras clave: Colección Digital, Exposición Virtual, Pintura Mural, Sierra de Huelva, Difusión.

Con más de ochenta restos de pintura mural repartidos por la comarca de la Sierra de Huelva, este apartado de la Historia del Arte serrano se nos presenta como uno de los más interesantes a la hora de profundizar en él, ya que nadie antes lo ha hecho en el sentido que estamos haciéndolo en esta investigación. Realizaremos un acercamiento, proponiendo una metodología de investigación que nos va a llevar al mundo de las humanidades digitales, utilizando recursos multimedia para mostrar una propuesta de difusión.

Debemos tener en cuenta que es un aspecto artístico no tratado anteriormente de una manera tan amplia. Cronológicamente podemos encontrar pinturas que van desde la primera mitad del siglo XV, hasta el siglo XX, siendo este último siglo bastante fructífero y pudiéndose observar las pinturas, obviamente, en mejor estado de conservación.

En cuanto a su estilo, los modelos responden, desde los más antiguos, a influencias venidas directamente de los grandes maestros muralistas italianos, hasta la utilización del trampantojo en época moderna, o el eclecticismo predominante en el siglo XX. La técnica utilizada para realizar estas pinturas va acorde a los siglos en que se realizan, podemos encontrar frescos, mezcla de fresco con óleo, técnica mixta, o pinturas al temple, que llenaban de color, hoy muchos de ellos perdidos por el mal estado de conservación u ocultos bajo la cal, las parroquias y ermitas de cada localidad.

Interesantes son también sus motivos iconográficos donde principalmente se representan escenas o personajes del Nuevo Testamento que servían como elementos catequizantes para poder comprender mejor lo que se leía en las Sagradas Escrituras. Últimas Cenas, Anunciaciones, el Apóstol Santiago a caballo, Evangelistas o Apóstoles, son los motivos más representados, también la decoración de rocallas barrocas simulando yeserías, el marmoleado o motivos geométricos, de carácter mudéjar, muestran la belleza de las mismas.

La falta de un estudio más exhaustivo, nos hace no poder hablar de autores de las mismas, pero si podemos referirnos a influencias directas, que como se indicaba anteriormente, vienen, en sus inicios, directamente de las grandes escuelas italianas que entran hasta Castilla, sirviendo de modelos a los artistas españoles que se concentran en "escuelas", sirviendo de foco para llegar a los lugares más cercanos. Podemos hablar de un importante foco pictórico en Sevilla, alrededor de su Catedral, Real Alcázar o el Monasterio de San Isidoro del Campo, en Santiponce, o un fuerte foco alrededor del Monasterio de Santa Clara de Moguer, hoy en la provincia de Huelva. Recordemos que la comarca serrana pertenecía a la Diócesis Hispalense, que controlaba todo el antiguo Reino de Sevilla, y los artistas iban de un lado a otro allá donde estaban sus encargos. No obstante muchas de estas pinturas responden a modelos populares en su factura, no siendo esto algo que le reste importancia a las mismas.

Localidades como Aracena, con numerosas iglesias con restos pictóricos, Aroche, donde también podemos encontrar pinturas en varios edificios, Cala, con ese magnífico retablo pictórico dedicado a la Magdalena,

Puerto Moral o Jabugo, por citar donde más cantidad y más antigüedad tienen, o Santa Ana la Real, Valdalarco o Alájar, que por ser más recientes se conservan en mejor estado, llenan, la geografía del norte de Huelva, de pinturas las paredes de sus iglesias.

También es interesante saber que no solo en edificios religiosos hay restos pictóricos, ya que sobre todo en el siglo XVIII, por ejemplo en el Castaño del Robledo, Alájar o Aroche, entre otros municipios, hay también pinturas en muchas casas, sobre todo nobiliarias, tanto en el exterior de las mismas como decorando las bóvedas de algunas de sus estancias interiores.

Por todo ello, ponemos de manifiesto la importancia de estas pinturas y para, sobre todo, en aquellos templos en los que se están perdiendo, dejar constancia de lo que hay antes de que definitivamente desaparezcan por su mal estado de conservación. Todavía queda mucho por hacer, por eso la puesta en valor mediante trabajos de investigación como el propuesto, estudios científicos y multidisciplinarios, la difusión a través de jornadas, congresos, publicaciones o artículos, juegan un gran papel en su puesta en valor.

También la gran cantidad de ejemplos pictóricos que se reparten en diferentes municipios, en distintos edificios y de una cronología dispar carecen de un estudio unificado, que englobe a todos y cada uno de dichos ejemplos. Han sido diferentes autores los que han tratado el tema, pero de una manera estanca y no viendo este fenómeno como un todo importante para la Historia del Arte de la Sierra de Huelva, donde se marca una relación directa con la producción sevillana dentro de la limitación geográfica de su antiguo Reino, tratándose de una región periférica del mismo, pero que además aglutina influencias que se relacionan estrechamente con la comunidad autónoma limítrofe, Extremadura, y con el país vecino, Portugal.

Al realizar un recorrido historiográfico sobre las pinturas murales de la comarca de la sierra onubense, podemos afirmar que hay numerosas publicaciones, más de una veintena, pero, como remarcamos, todas ellas, a excepción de las últimas publicadas, hacen referencia aislada a cada ejemplo pictórico. Se carece de un estudio que recoja todas las muestras que podemos observar hoy día.

El objetivo principal de nuestro proyecto es poner en valor la pintura mural de la Sierra de Huelva. Para llevar a cabo este objetivo se plantea, en primer lugar documentar todas las representaciones; y en segundo lugar, construir una herramienta digital que permita la consulta de esta información a los investigadores interesados y al público en general.

Siguiendo estas premisas, pretendemos recoger todas las muestras que hay y ordenarlas con la plataforma virtual seleccionada para ello, catalogándolas y creando diferentes colecciones dependiendo de nuestros intereses.

Es de considerar que la implementación en este proyecto de recursos digitales aportará un mayor rigor científico, académico y divulgativo al poderse organizar toda la información que los murales muestran, siguiendo una metodología académica, expuesta en nuestro trabajo de investigación.

Para conseguir nuestro objetivo, el de dar a conocer este aspecto de la historia artística de la Sierra de Huelva a todos los niveles, utilizamos la exposición virtual, teniendo como referencia algunas propuestas digitales de los mejores museos nacionales e internacionales. Y en torno a ello gira nuestra metodología para conseguirlo.

Para llegar a unas conclusiones concretas y que hablen de la importancia del tema investigado, hemos recopilado, en primer lugar, todo lo publicado sobre la pintura mural en la Sierra de Huelva.

A estas lecturas, le ha seguido un trabajo de campo de búsqueda y comprobación del estado de conservación de las pinturas murales halladas.

A las muchas preguntas de la investigación, fueron surgiendo respuestas con la preparación de diferentes participaciones en las jornadas de investigadores de la comarca de la Sierra. No obstante hasta la elaboración de una robusta base de datos no fue disipándose la problemática planteada para investigación.

Junto a esta base de datos principal, sobre las pinturas murales, se implementa otra base de datos sobre la bibliografía utilizada, haciendo uso para ello del programa de software libre Zotero, gestor de referencias bibliográficas, una herramienta esencial para recopilar y unificar en una colección todo lo referente a nuestro tema de investigación.

Tener claro las entidades y el conjunto de atributos de la base de datos principal, aportó una visión más general de la pintura mural. A hacer más fuerte la investigación contribuyó la aportación de datos por parte de Jesús Mendoza e Inmaculada Garrido, restauradores desde finales del siglo XX de muchas de las pinturas a estudiar. Este conjunto de entidades o pinturas que serán descritos y representados en la base de datos conforman, según Luis Codina y M^a. Eulalia Fuentes, el llamado *“Sistema de entidades registrables (SER)”*.

Siguiendo las indicaciones de Luis Codina procedemos al diseño de un diccionario de datos, teniendo como referente dicha base de datos documental, que para su creación hemos seguido su recomendación al indicarnos que *“el proceso de creación de una base de datos debe ir siempre desde los aspectos lógicos hacia los aspectos físicos”*.

Una vez identificadas las pinturas, procedemos a su organización por imágenes siguiendo el Id que ocupan en la base de datos, pudiéndolas tener organizadas para su uso en la exposición virtual.

La conclusión de este proyecto es la definitiva exposición virtual, donde se verá el uso de los datos en la creación a través de la base empírica que tiene la investigación. Siendo esta exposición virtual, como adelantábamos, el principal objetivo de nuestro proyecto, ya que nos va a ayudar a la puesta en valor de las pinturas murales mostradas en él.

Incidimos en que es de destacar en nuestro apartado metodológico, aspectos fundamentales como la documentación, donde se hace referencia del objeto en sí a exponer, y el producto, donde mostramos el uso que le daremos a los datos extraídos de los objetos.

Una de las labores más innovadoras de este proyecto es la de enumerar uno a uno los diferentes ejemplos de pintura mural en la Sierra de Huelva y que de manera inédita se muestran en este trabajo con la ya citada finalidad de aportar calidad y originalidad a su exposición virtual.

La documentación, constituye una investigación que por sí misma goza de suficiente entidad, realizándose un recorrido, en primer lugar, de movimientos y lugares que influyen directamente la pintura de la zona geográfica que nos ocupa, encontrando relaciones internacionales, nacionales y por último provinciales, siendo una aportación interesante, por qué no, para la Historia del Arte andaluz.

Como veníamos avanzando no existe ningún estudio que aglutine la cantidad de ejemplos pictóricos repartidos por la geografía serrana onubense, encontrándose testimonios en veinticinco de los veintinueve

municipios que componen la comarca de la Sierra de Huelva, así pues, para dotar de un importante y riguroso contenido a nuestro proyecto, que posteriormente se verá reflejado en la plantilla de recursos de la exposición virtual, se pretende realizar un somero recorrido cronológico, a la vez que iconográfico y, por supuesto, geográfico, de las pinturas a exponer.

Para desarrollar nuestra colección digital y sus posteriores exposiciones virtuales, utilizaremos el programa Omeka, aplicación de software libre, flexible de código abierto, pensada para mostrar colecciones y exposiciones, adaptándose perfectamente a nuestras necesidades, al permitir la creación de un repositorio mediante la organización de objetos digitales de diferentes colecciones, los cuales pueden integrarse en tantas exposiciones como se quiera.

Funciones como las de exponer, difundir y conservar (digitalmente) la obra, nos ha llevado a querer mostrar el gran repertorio de pintura mural de la Sierra de Huelva a través de su exposición virtual, con dos principales motivos, el comunicacional, abriendo nuestra exposición a la sociedad según los principios de democracia cultural y democratización de la cultura, y el económico al poder ser mostrada tal exhibición a través de un mundo virtual que no suponga un sobre coste para los visitantes, ni para la persona que lleva a cabo la acción creativa del proceso de exponer.

Nuestra creación gana en capacidad al estar conectada a otras fuentes digitales de información digital, formando parte así de una estrategia para transformar la organización y liberar el valor organizacional de la información, el movimiento *Open data*, que a día de hoy está revolucionando el mundo del acceso a los datos.

El resultado esperado es que, una vez concluido el proyecto y dada una buena publicidad al producto creado, la pintura mural serrana llegue a través de un dispositivo electrónico a todos y cada uno de los visitantes que quieran conocerla, animándoles a viajar hasta donde se encuentran.

La aplicación, además, debe permitir a los visitantes compartir y difundir las imágenes de las pinturas, y dar a conocer el estudio de este campo histórico artístico que se realiza en edificios tanto religiosos como civiles, por ello la colaboración de los usuarios, es importante para aportar opiniones y nuevas referencias a la colección, y que así vaya creciendo en importancia este área, pudiendo tener efectos beneficiosos, desde un punto de vista del turismo cultural, para toda la comarca.

Cada vez hay más información al respecto de la cultura digital, y las exposiciones virtuales concretamente, es un medio que está en auge, relacionado con la educación y el mundo del arte, vinculándolo directamente a la realidad virtual y la realidad aumentada, o que nos hace referencia a como la utilización de la tecnología puede ser un puente para vivir la experiencia, que es concretamente lo que pretendemos con nuestra exposición virtual, y así conseguir que la tecnología no sea el fin sino el medio para generar la experiencia.

La forma más gráfica de evaluar los resultados será la de acudir a la propia plataforma web del repositorio PMSierra una vez finalizada. Creación con el software Omeka, que nos proporciona todos los requisitos

necesarios para alcanzar nuestros objetivos. Así pues, nuestra propuesta se ha encaminado por esta visualización digital, para poder acceder a las pinturas de una manera sencilla, y sobre todo visual.

Dicha propuesta puede ser un referente a nivel autonómico, realizándose a partir de ella un prototipo exportable a otros campos, e incluso podría formar parte del sitio *Google Arts and Culture*, dentro de su apartado Maravillas de Andalucía, o por qué no, en colaboración con equipos multidisciplinares trabajar en un proyecto conjunto donde con el uso de unas gafas de realidad virtual podamos transportarnos a los edificios donde están las pinturas, e incluso marcar rutas por siglos pictóricos o motivos iconográficos.

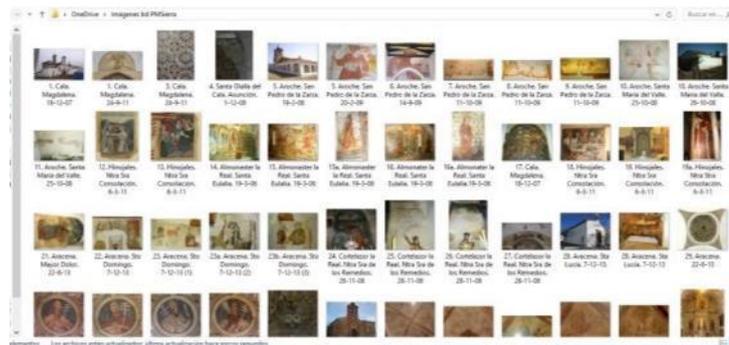
Figuras, tablas, mapas o gráficas

1. Implementación de los datos a PMSierra

ID	Lugar	Edificio	Nombre Edif.	Cronología	Periodo_Crono_E	Coordenadas	Zona	Cronología	Periodo_Crono_Pir	Iconografía	Ma
1	Cala	Iglesia	Santa María Magdalena	Siglo XV	Medios del siglo	37.972043, -6.308045	Ábside	Siglo XV	Primera mitad del sig	Ángel	Angel de San Mateo
2	Cala	Iglesia	Santa María Magdalena	Siglo XV	Medios del siglo	37.972043, -6.308045	Ábside	Siglo XV	Primera mitad del sig	Santo	San Cristóbal
3	Cala	Iglesia	Santa María Magdalena	Siglo XV	Medios del siglo	37.972043, -6.308045	Retablo	Siglo XV	Primera mitad del sig	Decoración de base	Decoración de base
4	Santa Chabalgüena	Ermita	Nuestra Señora de la Asunción	Siglo XV	Primera mitad del sig	37.962115, -6.234786	Ábside	Siglo XV	Siglo XV	Crucif.	Restos de un rostro
5	Arcena	Ermita	San Pedro de la Zarza	Siglo XVI	Siglo XVI, ampliación	37.889904, -6.948598	Muro norte	Siglo XV	Primera mitad del sig	Santo	San Cristóbal
6	Arcena	Ermita	San Pedro de la Zarza	Siglo XVI	Siglo XVI, ampliación	37.889904, -6.948598	Muro norte	Siglo XV	Segunda mitad del sig	Maria	Anunciación a la V
7	Arcena	Ermita	San Pedro de la Zarza	Siglo XVI	Siglo XVI, ampliación	37.889904, -6.948598	Muro sur	Siglo XV	Segunda mitad del sig	Santo	Escena con Santiago
8	Arcena	Ermita	San Pedro de la Zarza	Siglo XVI	Siglo XVI, ampliación	37.889904, -6.948598	Muro norte	Siglo XV	Segunda mitad del sig	Maria	Anunciación a la V
9	Arcena	Ermita	San Pedro de la Zarza	Siglo XVI	Siglo XVI, ampliación	37.889904, -6.948598	Muro norte	Siglo XV	Segunda mitad del sig	Crucif.	Ultima Cena
10	Arcena	Ermita	San Pedro de la Zarza	Siglo XVI	Siglo XVI, ampliación	37.889904, -6.948598	Muro norte	Siglo XV	Segunda mitad del sig	Maria	Virgen con el niño,
11	Arcena	Ermita	Santa María del Valle	Siglo XVI	Finales del siglo XI	37.967541, -6.947905	Ábside	Siglo XV	Segunda mitad del sig	Decoración geomé	Decoración geomé
12	Hinojales	Iglesia	Nuestra Señora de la Cor	Siglo XV	Finales del siglo XV	38.008495, -6.587550	Naves de la Epistola	Siglo XV	Finales del siglo XV	Santa	Maria Magdalena t
13	Hinojales	Iglesia	Nuestra Señora de la Cor	Siglo XV	Finales del siglo XV	38.008495, -6.587550	Naves de la Epistola	Siglo XV	Finales del siglo XV	Maria	Presentación de la V
14	Almonaster	Ermita	Santa Eulalia	Siglo XV	Medios del siglo	37.809000, -6.607780	Ábside	Siglo XV	Finales del siglo XV	Maria	Santa Eulalia, fianc
15	Almonaster	Ermita	Santa Eulalia	Siglo XV	Medios del siglo	37.809000, -6.607780	Ábside, lateral derecho	Siglo XV	Finales del siglo XV	Santo	Escena con Santiago
16	Almonaster	Ermita	Santa Eulalia	Siglo XV	Medios del siglo	37.809000, -6.607780	Ábside, lateral izquierdo	Siglo XV	Finales del siglo XV	Escena	Escena con Santiago
17	Cala	Iglesia	Santa María Magdalena	Siglo XV	Medios del siglo	37.972043, -6.308045	Ábside	Siglo XVI	18 de marzo de 1521	Santa	Retablo dedicado a
18	Hinojales	Iglesia	Nuestra Señora de la Cor	Siglo XV	Finales del siglo XV	38.008495, -6.587550	Ábside, lateral izquierdo	Siglo XVI	Primera mitad del sig	Crucif.	Ultima Cena
19	Hinojales	Iglesia	Nuestra Señora de la Cor	Siglo XV	Finales del siglo XV	38.008495, -6.587550	Naves del evangelio	Siglo XVI	Primera mitad del sig	Santos	Santos, Sant
20	Hinojales	Iglesia	Nuestra Señora de la Cor	Siglo XV	Finales del siglo XV	38.008495, -6.587550	Naves del evangelio	Siglo XVI	Primera mitad del sig	Santos	Inscripción gótica
21	Arcena	Iglesia	Nuestra Señora del Mayor	Siglo XV	Medios del siglo	37.809442, -6.561592	Isocrista	Siglo XVI	Siglo XVI	Maria	Virgen de la Anunci
22	Arcena	Convento	Santo Domingo	Siglo XVI	Medios del siglo	37.889486, -6.558298	Ábside	Siglo XVI	Siglo XVI	Maria	Anunciación a la V
23	Arcena	Convento	Santo Domingo	Siglo XVI	Medios del siglo	37.889486, -6.558298	Ábside	Siglo XVI	Siglo XVI	Varios	Diferentes escenas
24	Cortelazar	Iglesia	Nuestra Señora de los Reyes	Siglo XVI	Segunda mitad del sig	37.934772, -6.624189	Ábside	Siglo XVI	Último cuarto del sig	Crucif.	Caballo, con Jesús
25	Cortelazar	Iglesia	Nuestra Señora de los Reyes	Siglo XVI	Segunda mitad del sig	37.934772, -6.624189	Ábside, lateral derecho	Siglo XVI	Último cuarto del sig	Santa	Santa sin atributos
26	Cortelazar	Iglesia	Nuestra Señora de los Reyes	Siglo XVI	Segunda mitad del sig	37.934772, -6.624189	Ábside, lateral izquierdo	Siglo XVI	Último cuarto del sig	Crucif.	Bautismo de Cristo
27	Cortelazar	Iglesia	Nuestra Señora de los Reyes	Siglo XVI	Segunda mitad del sig	37.934772, -6.624189	Ábside, lateral izquierdo	Siglo XVI	Último cuarto del sig	Crucif.	Bautismo de Cristo
28	Cortelazar	Iglesia	Nuestra Señora de los Reyes	Siglo XVI	Segunda mitad del sig	37.934772, -6.624189	Ábside, lateral izquierdo	Siglo XVI	Último cuarto del sig	Crucif.	Bautismo de Cristo

Primeros campos con registros

2. Banco de imágenes de testimonios de pinturas murales en la Sierra de Huelva



Algunos ejemplo de imágenes donde se muestra el ID, el motivo iconográfico y la fecha de visita a las mismas

3. Campos en el diccionario de datos de PMSierra

Campos de representación de atributos de las pinturas	Campos de control administrativo
(01) Lugar (02) Edificio (03) Nombre del edificio (04) Cronología del edificio (05) Periodo cronológico del edificio (06) Coordenadas (07) Zona (08) Cronología de la pintura (09) Periodo cronológico de la pintura (10) Iconografía (11) Motivo iconográfico (12) Técnica (13) Estilo (14) Dimensiones (15) Autoría (16) Similitudes cercanas (17) Año restauración (18) Restaurador (19) Más información (20) Visitada	(21) Fecha de alta (22) Fecha de modificación (23) Operador (24) Número de registro

Campos en el diccionario de datos

4. Diccionario de datos de PMSierra

01

Campo Lugar	
Etiqueta	Lugar
Dominio	Nombre del municipio donde se encuentra la pintura mural
Tipo de dato	Alfanumérico
Indexación	Sí
Tratamiento documental	Diccionario de autoridades para nombres geográficos
Web	http://tesauros.mecd.es/tesauros/geografico/1109471.html
Lengua	Castellano
Otros controles	No puede quedar vacío. Valor repetible
Observaciones	Puede incluirse municipios o pedanías
Ejemplos válidos	Aroche

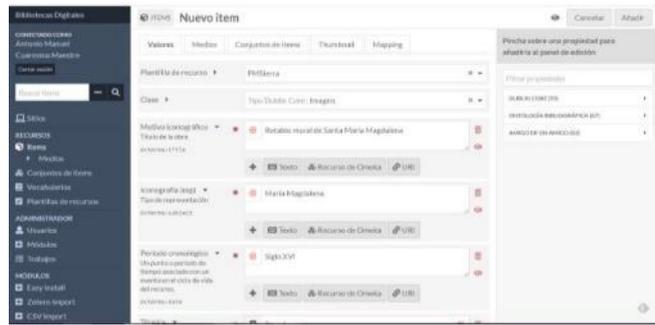
Ejemplo de ficha del diccionario de datos

5. Plantilla de recursos en Omeka (Proceso de creación de la colección digital)

Etiqueta original	Tipo de datos	Etiqueta alternativa	Comentario alternativo	¿Requerido?	¿Privado?
Título	Por defecto	Motivo iconográfico	Título de la obra	Sí	No
Nombre	Por defecto	Iconografía (enpl)	Tipo de representación	Sí	No
Fecha	Por defecto	Periodo cronológico		Sí	No
Medio	Recurso: Diccionario de Técnicas	Estilo	Técnica artística de realización de la obra	Sí	No
Creador	Por defecto	Autor		No	No
Cobertura Espacial	Recurso: Diccionario de Bienes Culturales	Edificio	Tipo de edificio	Sí	No
Cobertura	Recurso: Diccionario Geográfico	Lugar	Localidad	Sí	No
Descripción	Por defecto	Descripción	Descripción iconográfica	No	No
Relación	Por defecto	Similitudes cercanas	Obras de similar iconografía que pueden influir	No	No

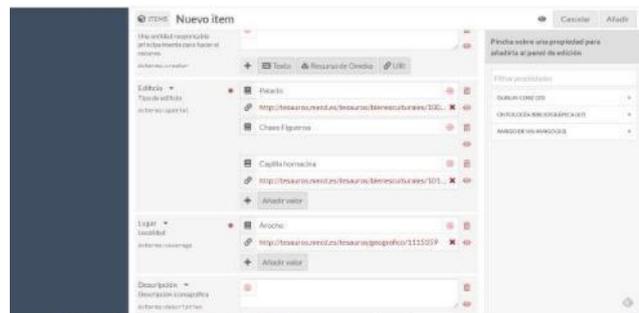
Imagen donde se muestran los primeros recursos

6. Creación de ítem



Campos a rellenar para la creación de un ítem

7. Metadatos



Utilización de tesauros en el ítem

Bibliografía

Abadal, Ernest, y Luís Codina. «Sistemas de gestión documental y bases de datos documentales». UOC. Accedido 26 de mayo de 2020.

Agénjo-Bullón, Xavier, y Francisca Hernández-Carrascal. «Registros de autoridades, enriquecimiento semántico y Wikidata». *Anuario ThinkEPI* 12, nº 0 (26 de abril de 2018): 361-72.

— — —. «Wikipedia, Wikidata y Mix'n'match». *Anuario ThinkEPI* 14, nº 0 (31 de marzo de 2020).

Arrieta Urtizberea, Iñaki. *El desafío de exponer: procesos y retos museográficos*. Servicio Editorial de la Universidad del País Vasco/Euskal Herriko Unibertsitateko Argitalpen Zerbitzua, 2015.

Ciurea, Cristian, y Florin Gheorghe Filip. «New Researches on the Role of Virtual Exhibitions in Digitization, Preservation and Valorization of Cultural Heritage». *Informatica Economica* 20, nº 4/2016 (30 de diciembre de 2016): 26-33.

Codd, E F. «A Relational Model of Data for Large Shared Data Banks» 13, nº 6 (1970): 11.

Codina, Lluís. «Sistemas de Gestión de Bases de Datos Documentales». *Departamento de Comunicació. Universitat Pompeu Fabra. Barcelona*, 2015, 24.

Codina, Lluís, y M Eulàlia Fuentes. «Documentación periodística y bases de datos: elemento para su fundamentación como disciplina y propuesta de conjunto nuclear de diccionario de datos», 1999, 20.

- Cuaresma Maestre, Antonio Manuel. «La pintura mural en Huelva. La comarca de la Sierra». En *Pintura mural en la Edad Moderna. Entre Andalucía e Iberoamérica*, Vol. II. Universo Barroco Iberoamericano. Sevilla, 2018.
- DCMI. «Dublin Core Metadata Initiative». Dublin Core Metadata Initiative, 2020 de 1995.
- Googel Arts & Culture. «Maravillas de Andalucía». Google Arts & Culture, 2020.
- «Guía para la creación de metadatos octubre pub 2009.pdf». Accedido 26 de mayo de 2020.
- Laguens García, José Luis. «Tesauros y lenguajes controlados en Internet». *CEOE. Anales de documentación* 9 (2006): 105-21.
- López Ochoa, Alberto. «Creación de un repositorio web con el software Omeka». En *Actas de las Jornadas de Estudiantes de Ciencias de la Documentación. Compartiendo conocimiento*, Facultad de Ciencias de la Documentación, Universidad Complutense de Madrid., 61-68. Madrid, 2018.
- Martini, Ricardo G., Mónica Guimarães, Giovani R. Librelotto, y Pedro Rangel Henriques. «Creating Virtual Exhibition Rooms from Emigration Digital Archives». *Universal Access in the Information Society* 16, nº 4 (1 de noviembre de 2017): 823-33.
- Mendoza Ponce, Jesús. «Pinturas murales en la Sierra de Huelva: el caso de Cortelazor la Real». En *Actas de las XIII Jornadas del Patrimonio de la Comarca de la Sierra. Cortelazor (Huelva)*, 199-210. Cortelazor la Real, Huelva: Diputación Provincial de Huelva, 1999.
- Oliver, Carlos Alberto, Alfonso Hernández Pleguezuelo, y José M. Sánchez. *Guía histórico-artística de la Sierra de Aracena y Picos de Aroche*. Aracena, Huelva: Iniciativas Leader Sierra de Aracena y Picos de Aroche, 2004.
- Open Archives. «Iniciativa de archivos abiertos». Open Archives, 2017.
- Peñalver Gómez, Eduardo, y Jesús Bocanegra Linares. «OMeka S para exposiciones digitales», octubre de 2018.
- Respaldiza Lama, Pedro José, Juan Luis Ravé Prieto, Inmaculada Garrido Márquez, y Timoteo Rivera Jiménez. *Restauración de los retablos mayores de la iglesia parroquial de Santa Magdalena de Cala (Huelva)*. Huelva: Delegación Provincial de Cultura, 2003.
- Rodríguez Guillén, Antonio, y Antonio Manuel Cuaresma Maestre. «Pintura Mural en la Sierra de Huelva. Pintura mural en Hinojales». En *Actas de la XXVI Jornadas del Patrimonio de la Sierra de Huelva*, 223-44. Huelva: Diputación de Huelva, 2011.
- Saorín Pérez, Tomás. «Exposiciones digitales y reutilización: aplicación del software libre Omeka para la publicación estructurada». *Métodos de información* 2, nº 2 (29 de julio de 2011): 29-46.

Catalogar el patrimonio histórico-artístico en tiempos de pandemia: *Google Maps* como herramienta investigadora y docente.

Antonio Ledesma

Dr. en Historia del Arte (investigador independiente)
<https://independent.academia.edu/AntonioLedesma21>

Sergio Sánchez Jiménez

Profesor de Geografía e Historia en el I.E.S. José Luis L. Aranguren de Ávila.

Palabras clave

Google Maps, Historia del Arte, investigación, docencia y pandemia

Descripción del problema

En el año 2020 la Fundación Santa María la Real del Patrimonio Histórico encomendó a un grupo de investigadores la elaboración de un estudio monográfico sobre los registros materiales pertenecientes al arte románico que se conservan en la provincia de Salamanca (con especial celo en los casos inadvertidos hasta la fecha por parte de la historiografía)⁵⁷. Se trataba de un trabajo de por sí complejo, arduo y costoso que requería de una buena organización y disponibilidad. Si a esto se sumaba una fuente de financiación limitada y en un contexto de pandemia que por el estado de alarma impedía o reducía el trabajo de campo, así como la consulta de bibliotecas, fototecas y otros centros similares para el reconocimiento y el análisis de los bienes del estudio, resultaba imprescindible la búsqueda de herramientas auxiliares poco comunes que facilitasen la identificación con la máxima precisión, limitando los traslados y reduciendo los costes (tanto materiales como humanos). Ante esta coyuntura *Google Maps*, junto con *Google Earth*, se ha revelado como una fuente idónea y asequible con la que obtener información rigurosa y actualizada para alcanzar los objetivos del estudio.

Metodología

Partiendo de los grandes corpus desarrollados a lo largo del siglo XX, surgía la necesidad de examinar la naturaleza de los casos identificados hasta la fecha y la posibilidad de ampliar su número tras el reconocimiento individualizado de cada núcleo de población. La provincia de Salamanca es una de las más extensas en superficie en Castilla y León y se divide en 362 municipios (sin incluir en esta agrupación alquerías, despoblados, etcétera). Y cada una de estas localidades cuenta en todos los casos con una iglesia, construcción que de un modo u otro podría conservar algún vestigio de arte románico. Ante esta coyuntura, se procedió en primer lugar a contar con un listado de todas las poblaciones y en segundo lugar se pasó al análisis de cada uno de los templos con el empleo prioritario de *Google Maps*, junto con otras fuentes. Las presentaciones *Street View* y *360º* permitían en la casi totalidad de los casos rodear el edificio y ver la obra en su conjunto, pero también con cierto detalle gracias a la ampliación de pormenores, llegando incluso a captar detalles escultóricos o marcas de cantería distribuidas por los muros. Asimismo esta herramienta incorporaba la posibilidad de visualizar imágenes del interior de estas

⁵⁷ <https://www.santamarialareal.org> [fecha de consulta 14-IV-2021].

iglesias gracias a la participación ciudadana⁵⁸. De igual modo, resulta necesario referirse a la opción de ver la evolución en el tiempo y los cambios que se han generado tanto en el entorno como en el propio edificio. También se ha recurrido a *Google Earth* para el estudio de las cubiertas y otros pormenores. Tras esta primera selección, en los casos más dudosos o que resultaba imprescindible la visita, se procedió a su reconocimiento *in situ*. Solo en algunos registros se ha observado que la valoración obtenida con las aplicaciones podría ser confusa o imprecisa, siendo contrastada con el uso de otras fuentes, mientras que para el resto resultaba ser un medio rápido y eficaz.

Resultados obtenidos

Gracias al trabajo con *Google Maps*, además de auxiliar en la confirmación o en el descarte de los testimonios que la historiografía ofrecía como obras románicas, se han incorporado más de cincuenta nuevos registros al listado existente, lo que supone más de un tercio del total. Además del número, su catalogación ha ofrecido información imprescindible para seguir ampliando los conocimientos existentes sobre el arte románico en la provincia de Salamanca. De este modo, se han obtenido otros datos de interés para la investigación sobre cuestiones relativas al estado de conservación de los edificios, intervenciones, señalización, piezas descontextualizadas en el entorno y otros pormenores que demuestran la utilidad de la herramienta en su nivel más básico, sin necesidad de formación previa especializada y con carácter altamente intuitivo. Si a esto se suma su condición gratuita, actualidad relativa y alto valor participativo, al permitir al usuario señalar tanto errores como añadir fotografías, las ventajas de su empleo son palpables más allá de la COVID-19. De igual manera *Google Maps* podría aplicarse en otras áreas, no solo en arte románico y en estudios de catalogación, ya sea en cuestiones vinculadas con la gestión cultural, la restauración, el urbanismo (*Timelapse* de *Google Earth*), la arquitectura (vistas en 3D), etcétera. Por último, su uso en la enseñanza superior es asumible y recomendable, al permitir al docente asignar labores muy diversas a los alumnos, desde la catalogación y la divulgación, a la gestión cultural y la restauración, puesto que “with easy access to spatial and cultural information, Google Maps/Earth has provided users with the means to understand their world and their communities of interest” (Dodsworth y Nicholson, 2012). En resumen, *Google Maps* es una herramienta cargada de futuro en el ámbito de las Humanidades Digitales.

⁵⁸ *Street View* ya está incorporando algunos interiores con la misma tecnología que para los exteriores, pero los ejemplos siguen siendo minoritarios en el cómputo total.

Figuras

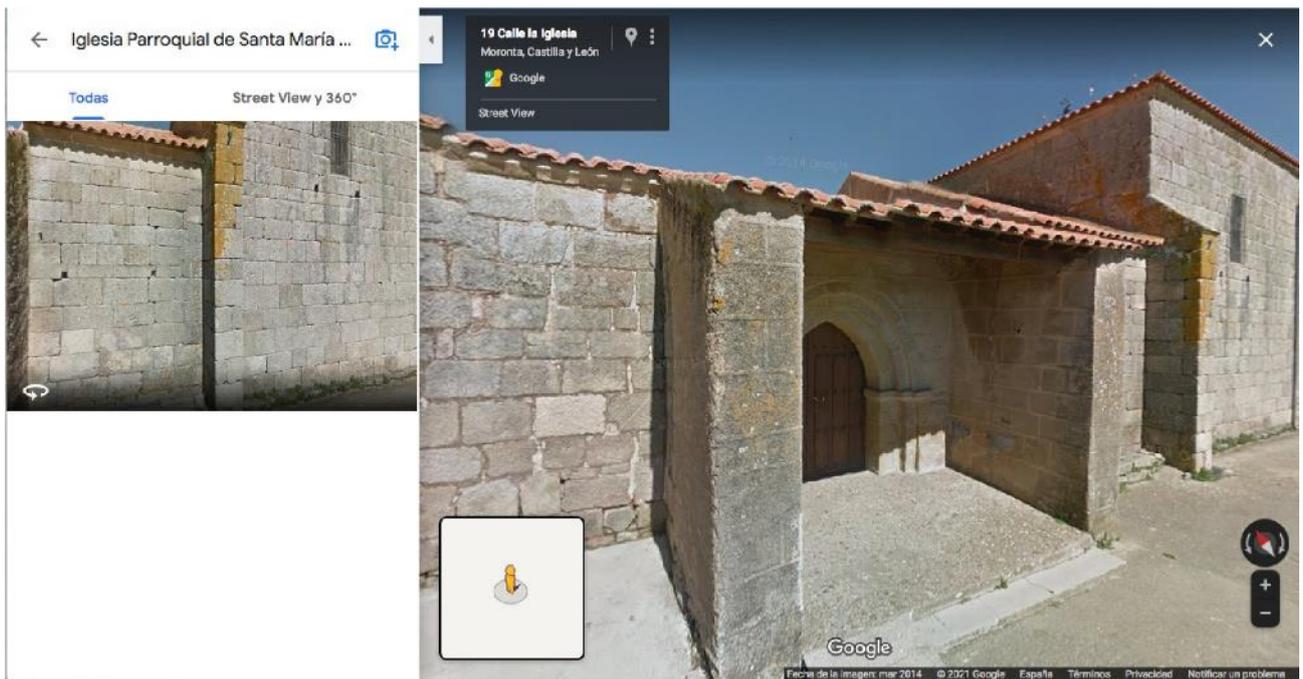


Figura 1. Ejemplo de templo románico identificado gracias a la tecnología *Google Maps*.



Figura 2. Iglesia en una finca privada no visitable.

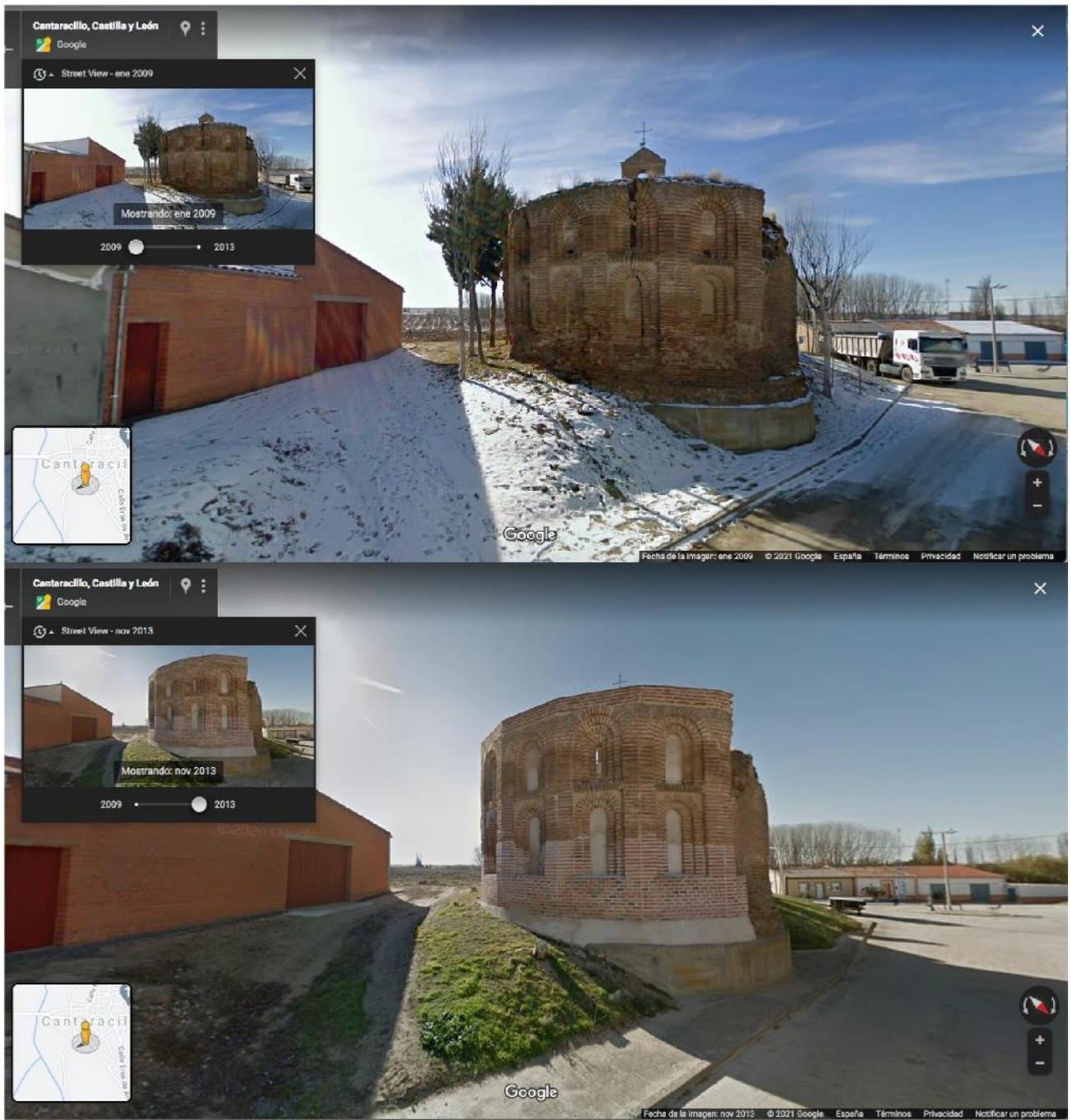


Figura 3. Estado de uno de los casos en 2009 y en 2013.



Figura 4. Presencia y ausencia de un sarcófago medieval.

Bibliografía:

MORÁIS MORÁN, J.A.; TEIJEIRA PABLOS, M.^a D. (coords.). *Academia, educación e innovación en la historia del Arte Medieval*. Granada: Editorial Universidad de Granada, 2020.

BOLLETTA, V. "El lugar de la investigación en tiempos de pandemia. Reflexividad y perplejidad".

Anuario Digital de Investigación Educativa, 2021, n.º. 4, pp. 71-77.

<http://revistas.bibdigital.uccor.edu.ar/index.php/adv/article/view/4944/3247>

DODSWORTH, E.; NICHOLSON, A. "Academic Uses of Google Earth and Google Maps in a Library Setting". *Information Technology and Libraries*, 2012, n.º 31 (2), pp. 102-117.

https://www.researchgate.net/publication/265002770_Academic_Uses_of_Google_Earth_and_Google_Maps_in_a_Library_Setting

LAMB, A.; JOHNSON, L. "Virtual Expeditions: Google Earth, GIS, and Geovisualization Technologies in Teaching and Learning". *Teacher Librarian*, 2010, n° 37 (3), pp. 81-85.

<https://scholarworks.iupui.edu/handle/1805/8651>

Cognición y materialidad

Modera: Fátima Díez Platas · *Asiste:* Alexandra Iglesias Cortés

¿Es posible estudiar si los objetos culturales influyen nuestros procesos cognitivos a través de los estímulos perceptuales que producen? ¿Influyen en particular la capacidad de predicción (o anticipación) de nuestra mente? Este es el objetivo de un proyecto reciente que financia el ERC (*European Research Council*). El proyecto aúna neurociencia fundamental y ciencias cognitivas con la arqueología y la antropología y plantea atacar esas preguntas de investigación desde una perspectiva histórica y transcultural. Es un ejemplo de la potencia que puede tener relacionar las Humanidades con otras disciplinas científicas. Pero también permite prever los efectos en los procesos cognitivos y la atención de la generalización de las tecnologías digitales.

Panel 4

Aracne Nodus. Red de humanidades digitales y letras hispánicas

Modera: Nieves Pena Sueiro · *Asiste:* César González-Pérez

Panelistas: Ángeles Saavedra Places; Anna Bognolo; Delfina Ramos Vidal; Stefano Bazzaco; Stefano Neri

Aracne Nodus. Red de humanidades digitales y letras hispánicas

Coordinadora: Nieves Pena Sueiro(UDC)

Comunicación 1: Anna Bognolo, Stefano Neri y Stefano Bazzaco (Universidad de Verona)

Comunicación 2: Delfina Ramos Vidal, Ángeles Saavedra Places, Alejandro Cortiñas, Miguel R. Luaces, Óscar Pedreira(Universidade da Coruña)

[Aracne, red de Humanidades Digitales y Letras Hispánicas](#), nació en el año 2012 por la agrupación de seis grupos de investigación de Letras con proyectos activos y que habían desarrollado bases de datos, bibliotecas digitales y otros recursos ([Biblioteca Saavedra Fajardo](#), [BIDISO](#), [BIESES](#), [CLARISEL](#), [DIALOGYCA](#) y [PHEBO](#)). El principal objetivo de la Red Aracne fue ofrecer información de los citados grupos de investigación y sus proyectos, así como los recursos que creados por estos (16 bases de datos o/y bibliotecas digitales) a través de un único sitio web o punto de acceso; otro de los objetivos fue desarrollar un [metabusgador](#) que permitiese la interoperabilidad entre los recursos de cada grupo, que sirviese de punto de acceso común para acceder a sus bases de datos, lo que, sin duda, resultó de gran utilidad y economizó tiempo a los investigadores interesados. La adaptación de cada uno de los recursos al protocolo OAI-PMH, necesaria para conseguir la interoperabilidad, redundó, además, en un aumento exponencial de la visibilidad de los resultados de investigación de los grupos, al ser recolectados por agregadores como HISPANA y EUROPEANA.

Sin embargo, el paso del tiempo puso en peligro la sostenibilidad de la Red Aracne por la obsolescencia de los formatos de interoperabilidad. El modelo utilizado por la Red Aracne era el modelo ESE (Europeana Semantic Elements), que ha sido sustituido recientemente por el modelo EDM, más flexible y expresivo. Los principios subyacentes de diseños de EDM se basan los principios fundamentales y buenas prácticas de la Web Semántica y Linked Data y actúa como una ontología común de alto nivel que permite la interoperabilidad manteniendo los modelos de datos originales y sus perspectivas de información. Este fue el motivo principal por el que en el año 2019 se solicitó financiación al Ministerio de Economía y Competitividad con el fin de actualizar la red Aracne, tanto en lo relativo a la tecnología como en los contenidos y, se aprovechó, además, para incorporar nuevos grupos de investigación.

Aracne Nodus surgió, en 2019, como continuación de la primigenia Red Aracne, con el objetivo principal de actualizar los datos y recursos existentes, agregar recursos nuevos creados por esos grupos, y, además, incorporar cuatro grupos más, de referencia en desarrollos de recursos digitales para la investigación: tres grupos humanísticos ([BECLaR](#), [PARNASEO](#) y [Progetto Mambrino](#)) y uno tecnológico, el LBD (Laboratorio de bases de datos) de la UDC. Así, la red Aracne Nodus está conformada por nueve grupos o nodos de información, todos ellos referentes en sus líneas de investigación y reconocidos en su trabajo en Humanidades Digitales (8 grupos de Literatura Española de la Edad Media y Edad Moderna, y un grupo experto en bases de datos y bibliotecas digitales).

El panel pretende, en la primera comunicación, presentar esta red, sus objetivos, sus logros y retos y ofrecer, además, en dos comunicaciones, los objetivos y recursos de dos de los grupos nuevos: el Progetto Mambrino, de la Universidad de Verona, que recientemente creó una Digital Library, y el Laboratorio de Bases de Datos, de la Universidade da Coruña, que está desarrollando una herramienta de creación automática de *software* con el fin de facilitar la creación de bibliotecas digitales con los requisitos definidos para integrarse en la Red Aracne.

Comunicación 1

Progetto Mambrino

Autores: Anna Bognolo (Università di Verona), Stefano Neri (Università di Verona), Stefano Bazzaco (Università di Verona)

Las tecnologías digitales han incidido en todos los procesos implicados en la investigación literaria: la planificación del estudio, la producción y difusión de resultados, la búsqueda de la información y la interacción con bases de datos semánticas. A partir de esta premisa, el [Progetto Mambrino](#) (Università di Verona) avió en los últimos años un proceso de integración de los resultados científicos ya obtenidos con el entorno digital para la creación de una Digital Library de los libros de caballerías que se publicaron en Italia durante el Renacimiento, traducciones y continuaciones originales de las exitosas obras españolas.

En la presente comunicación se ilustrará el proceso de ideación, creación y publicación de la Digital Library. En concreto se expondrán los objetivos principales del proyecto, las distintas fases de digitalización y modelización de los contenidos (transcripción semi-automática de las obras del corpus con la plataforma de HTR [Transkribus](#); ajuste de las imágenes facsimilares al estándar IIIF; maquetación XML TEI; creación de bases de datos de nombre, lugares, motivos caballerescos) y los resultados experimentales de visualización e interconexión en *standoff markup* con las bases de datos desde el nivel de la interfaz web generada con la herramienta TEI Publisher.

A la luz de las finalidades del proyecto, durante la exposición se discutirán los problemas que se encontraron en la planificación, y se aclararán las metodologías y soluciones adoptadas de cara a la interoperabilidad y a la búsqueda de recursos y estándares ya largamente difundidos en la comunidad científica de referencia.

Bibliografía:

Bazzaco, Stefano, «El Progetto Mambrino y las tecnologías OCR: estado de la cuestión», *Historias*

Fingidas 6 (2018), 257-272
<<http://historiasfingidas.dlts.univr.it/index.php/hf/article/view/89/135>>.

Bazzaco, Stefano, «El reconocimiento automático de textos en letra gótica del Siglo de Oro: creación de un modelo HTR basado en libros de caballerías del siglo XVI en la plataforma Transkribus», *Janus*, 9 (2020), 534-561 <<https://www.janusdigital.es/articulo.htm?id=160>>.

Bognolo, Anna, «Reti di libri, libri in rete. Libros de caballerías tra Italia e Spagna», *Trayectorias literarias hispánicas: redes, irradiaciones y confluencias*, AISPI Edizioni, 2018, 91-106.

Bognolo, Anna, «Didattica della letteratura e Informatica Umanistica/Humanidades Digitales», *Hispanismo y didáctica universitaria: cuestiones y perspectivas*, AISPI Edizioni, 2020, 119-133.

Bognolo, Anna y Stefano Bazzaco, «Tra Spagna e Italia: per un'edizione digitale del Progetto Mambrino», *eHumanista/IVITRA*, 16 (2019), 20-36.

Fiormonte, Domenico, *Per una critica del testo digitale*, Roma: Bulzoni, 2018.

Mancinelli, Tiziana, «Early printed edition and OCR techniques: what is the state-of-art? Strategies to be developed from the working-progress Mambrino project work», *Historias Fingidas* 4 (2016), 255-260.

Mühlberger, Günter *et al.*, «Transforming scholarship in the archives through Handwritten Text Recognition. Transkribus as a case study», *Journal of Documentation - Emerald Publishing*, vol. 75, n. 5 (2019), pp. 954-976 < <https://www.emerald.com/insight/content/doi/10.1108/JD-07https://www.emerald.com/insight/content/doi/10.1108/JD-07-2018-0114/full/html2018-0114/full/html> >.

Pierazzo, Elena, *Digital Scholarly Editing. Theories, Models and Methods*, Londres-Nueva York: Routledge, 2020.

Schillingsburg, Peter, *From Gutenberg to Google: Electronic Representations of Literary Texts*, Cambridge: CUP, 2009.

Comunicación 2: Reduciendo complejidad y coste de la creación de Bibliotecas Digitales con técnicas de Ingeniería Software

Autores: Delfina Ramos Vidal, Alejandro Cortiñas, Miguel R. Luaces, Óscar Pedreira, Ángeles Saavedra Places

Los organismos que necesitan construir una Biblioteca, Catálogo o Archivo digital, muchos de ellos grupos de investigación en Humanidades Digitales, se enfrentan desde el principio al problema de conseguir la financiación suficiente que requiere el desarrollo del *software* que le de soporte. En particular, las partidas de desarrollo de *software* de las subvenciones a las que pueden acceder estos grupos de investigación son verdaderamente insuficientes para abordar el trabajo completo de una sola vez, por lo que han de usar diferentes convocatorias para completar el presupuesto necesario, dilatándose demasiado en el tiempo la implantación de la Biblioteca en construcción.

Sin embargo, si se analizan desde el punto de vista de la Ingeniería del *Software*, las Bibliotecas Digitales tienen unas propiedades que las hacen encajar en dos paradigmas de desarrollo cuyo objetivo es automatizar la generación de código a partir de especificaciones de análisis: las Líneas de Producto Software (LPS) y el Desarrollo Dirigido por Modelos (DDM).

Para conseguir reducir la complejidad y el coste de la creación de Bibliotecas Digitales hemos desarrollado una plataforma que permite generar su *software* de forma automática [4,5,6] a partir de la especificación de su modelo de datos y sus características, de forma que el presupuesto y tiempo requerido para su construcción disminuye significativamente. De esta manera, los grupos de Humanidades Digitales pueden centrarse en su investigación en vez de preocuparse por el desarrollo del *software*.

Durante el desarrollo del proyecto hemos seguido una metodología, basada en un enfoque mixto LPS [1,2] y DDM [3], que ha sido creada por los autores de esta comunicación y probada con éxito también en otros dominios.

Este proyecto se ha desarrollado en el marco de la Red Aracne-Nodus financiada por el Ministerio de Ciencia, Innovación y Universidades (ref. RED2018-102755-T).

Referencias

- [1] Alejandro Cortiñas, Miguel R. Luaces, Oscar Pedreira, and Ángeles S. Places. "Scaffolding and inbrowser generation of web-based gis applications in a spl too"l. In *Procs. of the 21st International Systems and Software Product Line Conference-Volume B*, pages 46–49. ACM, 2017.

- [2] Alejandro Cortiñas, Miguel R. Luaces, Oscar Pedreira, Ángeles S. Places, and Jennifer Pérez. "Web-based geographic information systems sple: Domain analysis and experience report". In *Procs. of the 21st International Systems and Software Product Line Conference-Volume A*, pages 190–194. ACM, 2017.
- [3] Brisaboa, N. R.; Cortiñas, A.; Luaces, M. R.; Pedreira, O.: "Creating web-based GIS applications using automatic code generation techniques", en *Proc.15th Int. Symposium on Web and Wireless Geographical Information System (W2GIS 2017)-LNCS 10181*, Springer, Shanghai (China), 2017, pp. 19-34.
- [4] Delfina Ramos-Vidal. *Tool for the semi-automatic generation of software for Digital Libraries*. Trabajo fin de grado. Departamento de Ciencias de la Computación y Tecnologías de la Información. Universidade da Coruña, A Coruña, 2020.
- [5] Delfina Ramos-Vidal, Alejandro Cortiñas, Miguel R. Luaces, Oscar Pedreira, and Ángeles Saavedra Places. "A software product line for digital libraries", In Massimo Marchiori, Francisco José Domínguez Mayo, and Joaquim Filipe, editors, *Proceedings of the 16th International Conference on Web Information Systems and Technologies, WEBIST 2020*, Budapest, Hungary, November 3-5, 2020, pages 381–389. SCITEPRESS,2020.
- [6] Delfina Ramos-Vidal, Alejandro Cortiñas, Miguel R. Luaces, Oscar Pedreira and Ángeles Saavedra Places, "Development of digital libraries with software product line engineering". *Special Issue on Advanced Practices in Web Engineering 2021 of the Journal of Web Engineering* (ISSN: 1540-9589). JWE, 2021. Pendiente de publicación.

Bibliografía sobre la Red Aracne

Actas del Seminario Internacional sobre Bibliotecas digitales y Bases de datos especializadas en Literaturas Hispánicas

(*BIDESLITE*), Madrid, 4-5 julio de 2011, dir. Mercedes Fernández Valladares, ed. de María Casas del Álamo, Germán Redondo Pérez y Sara Sánchez Bellido, Madrid, IUMP (UCM), 2013. E-prints UCM. Accesible desde 7 de mayo 2013.

Alvite Díez, M^a Luisa y Nieves Pena Sueiro, "Colecciones digitales patrimoniales. Estudio de la Red Aracne", en *Actas del IV Congreso ISKO España-Portugal 2019, XIV Congreso ISKO España / coord. por Jesús Tramullas Saz, Piedad Garrido Picazo, Gonzalo Marco Cuenca*, 2020, pp. 185-195, ISBN 978-84-09-20065-8,

Arrigoni, Eleonora; Rodríguez-López, Eduardo, «La red de investigación de Humanidades Digitales y Letras Hispánicas: avance de Red ARACNE», en *Visibilidad y divulgación de la investigación desde las Humanidades Digitales. Experiencias y proyectos*, ed. Álvaro Baraibar, Pamplona, Servicio de Publicaciones de la Universidad de Navarra, 2014, pp. 243-251.

Baranda Leturio, Consolación y Eduardo Rodríguez López, "Red ARACNE: retos y objetivos de un proyecto de coordinación en letras hispánicas digitales", *Janus, (Anexo I)*, Universidade da Coruña, España, 2014, pp. 101-109.

Collantes Sánchez, Carlos M., "Red Aracne Nodus: consolidación y nuevos horizontes en las Humanidades Digitales", comunicación presentada en el *XII Congreso de la AISO*, 6 de noviembre de 2020.

Pena Sueiro, Nieves y Ángeles Saavedra Places, "Aracne. Red de Humanidades Digitales y Letras Hispánicas", *Historias Fingidas 7* (2019), pp. 409-214. DOI: <http://dx.doi.org/10.13136/22842667/149> . I

Panel 5

El proyecto «Mapping Chivalry. Spanish Romances of Chivalry from Renaissance to 21st Century: A Digital Approach»

Moderador: Daniele Crivellari · *Asiste:* César González-Pérez

Panelistas: Anna Bognolo; Elisabetta Sarmati; Giulia Tomasi

El proyecto «Mapping Chivalry. Spanish Romances of Chivalry from Renaissance to 21st Century: A Digital approach»

NOMBRE Y FILIACIÓN DEL COORDINADOR.

- Daniele Crivellari (Università di Salerno) – dcrivellari@unisa.it

NOMBRES Y FILIACIONES DE LOS PARTICIPANTES.

- Anna Bognolo (Università di Verona) – anna.bognolo@univr.it
- Elisabetta Sarmati (La Sapienza Università di Roma) – elisabetta.sarmati@uniroma1.it
- Giulia Tomasi (Università di Trento) – giulia.tomasi.2@unitn.it

PALABRAS CLAVE.

Bases de datos; libros de caballerías en España e Italia; motivos; teatro del Siglo de Oro; narrativa de los siglos XIX-XXI.

DESCRIPCIÓN DEL PROBLEMA DE INVESTIGACIÓN, METODOLOGÍA Y RESULTADOS OBTENIDOS.

- La propuesta de panel titulado «El proyecto “Mapping Chivalry. Spanish Romances of chivalry from Renaissance to 21st Century: a Digital approach”» surge de la voluntad de presentar a la comunidad científica el trabajo que están llevando a cabo varios investigadores de distintas universidades italianas (Roma, Salerno, Trento y Verona) en el ámbito de un proyecto nacional financiado en 2020 por el MIUR (Ministero dell’Istruzione, dell’Università e della Ricerca) y que finalizará en 2023. Partiendo de la consideración de la importancia que los libros de caballerías españoles tuvieron a lo largo del Renacimiento y de su constante impacto también en épocas sucesivas (desde el Barroco hasta la contemporaneidad), tanto en España como en otros países de Europa, el grupo de investigadores ha decidido abordar y analizar el fenómeno desde cuatro puntos de vista distintos y entrelazados:
- la unidad de Verona se ocupa del censo y el estudio de los libros de caballerías italianos que se publicaron en Venecia durante los siglos XVI y XVII, traducciones y continuaciones originales de las famosas obras españolas;
- la unidad de Trento enfoca su atención hacia la presencia de motivos comunes en los libros de caballerías españoles, ya que a través de este enfoque narratológico puede resaltarse la evolución del género durante el largo período de su auge en la escena editorial, además de proporcionar respuestas complejas sobre sus ricos contenidos narrativos;
- la unidad de Salerno se ocupa del impacto que los libros de caballerías tuvieron en la gestación y la evolución de la Comedia nueva barroca, rastreando las piezas de teatro áureo que guardan relación, a distintos niveles, con el corpus caballeresco;
- finalmente, la unidad de Roma centra su trabajo en las reescrituras modernas de los libros de caballerías, incluyendo las reescrituras de *Don Quijote* con autores, obras y ediciones digitalizadas, comparando las nuevas estructuras narrativas, los motivos y los personajes de las escrituras contemporáneas con los paradigmas caballerescos del siglo XVI y estudiando los diferentes grados de reescritura del nuevo corpus.

DESCRIPCIÓN DE LAS INTERVENCIONES DE CADA PARTICIPANTE.

- DANIELE CRIVELLARI (UNIVERSITÀ DI SALERNO): en esta intervención se presentará el proyecto titulado «Teatro caballeresco» que están llevando a cabo, desde la Universidad de Salerno (Italia), Daniele Crivellari y Paula Casariego Castiñeira. El punto de partida del proyecto es la observación de la fortuna editorial que los libros de caballerías tuvieron a lo largo del siglo XVI, pero también de su importancia a lo largo de las épocas sucesivas; particularmente, la atención del grupo de investigación de Salerno se ha centrado en la relación que el corpus caballeresco del Renacimiento mantiene con la producción teatral del Siglo de Oro. El proyecto se vertebra en torno a dos ejes fundamentales: por un lado, la creación de una base de datos en la que se recoja el corpus de todas las piezas de teatro áureo que están relacionadas, a varios niveles, con libros de caballerías. A partir de trabajos específicos sobre el asunto (cfr. por ejemplo Demattè, 2005) y siguiendo el modelo de proyectos digitales análogos (Artelope – artelope.uv.es/, Calderón digital – calderondigital.unibo.it/, etc.), para cada comedia se está creando una ficha digital en la que se incluye toda la información pertinente sobre el texto, para dar cuenta cabal no solo de los datos bibliográficos y temáticos (año de primera publicación, ediciones antiguas y modernas, sinopsis argumental, etc.), sino también de los enlaces que se establecen con el corpus caballeresco, esto es, qué relación mantiene cada pieza con posibles hipotextos caballerescos. Por otro lado, otro objetivo del proyecto «Teatro Caballeresco» es la publicación de ediciones de los textos dramáticos incluidos en la base de datos que no cuentan actualmente con una edición crítica fiable. Para este fin, se ha empezado un proyecto de colaboración con la editorial Reichenberger y con la Biblioteca Virtual Miguel de Cervantes, que publicará en acceso abierto dichas ediciones. El primer resultado de esta parte del proyecto es la publicación de la edición de la pieza *Amadís y Niquea*, del dramaturgo Francisco Leiva Ramírez de Arellano, a cargo de Claudia Demattè (en prensa). En la comunicación se presentarán tanto las características de la base de datos como la colección de ediciones críticas.
- ANNA BOGNOLO (UNIVERSITÀ DI VERONA): El Progetto Mambrino (Università di Verona) es un grupo de investigación que se ocupa del censo y el estudio de los libros de caballerías italianos que se publicaron en Venecia durante los siglos XVI y XVII, traducciones y continuaciones originales de las famosas obras españolas. Los libros de los dos ciclos más exitosos de Amadís y Palmerín, traducidos y continuados en su mayoría por Mambrino Roseo da Fabriano, tuvieron una difusión europea y acabaron por ser verdaderos *best-sellers* de esa época. Estos libros no cuentan con ediciones críticas modernas que permitan su rescate y su estudio por parte de los investigadores (Bognolo-Bazzaco, 2019). Por estas razones, a la luz de las posibilidades que ofrece el medio computacional (Tomasì, 2008; Pierazzo, 2020), tomó forma la idea de crear una Biblioteca Digital de libros de caballerías italianos de inspiración española que pudiese integrar varios recursos ya disponibles — *Repertorio delle continuazioni italiane ai romanzi cavallereschi spagnoli. Ciclo di Amadis di Gaula* (2013) y el *Repertorio ... Ciclo del Palmerino d'Oliva* (en prensa 2021)— con otros digitales de nueva creación para el fichaje bibliográfico, la edición, el marcado y la representación del corpus, dando como resultado una Biblioteca Digital interactiva. Dentro de un entorno digital los datos pueden ser relacionados y estructurados por medio de bases de datos de carácter bibliográfico y temático y el etiquetado de los textos transcritos. Los objetivos principales del proyecto son los siguientes: a) poner a disposición de los especialistas una base de datos de obras de libre consultación, fácilmente navegables y consultables a partir de distintas perspectivas de investigación; b) consentir la visualización de las ediciones digitales de los textos en modalidad *split-screen* con una conexión entre imágenes facsímil y transcripción, modelada según los criterios XML-TEI; c) contar con varios recursos (repertorios bibliográficos, resúmenes, índices de nombres de personajes y lugares, índices de motivos – véase: Luna Mariscal, 2020) vinculados a la edición de cada texto e integrados dentro de la Biblioteca Digital. En esta comunicación se presentarán los resultados actuales del proyecto, principalmente con respecto a la migración de las bases de datos bibliográficas al contexto digital, la transcripción automatizada de las fuentes y la modelización de los libros italianos del ciclo de Amadís de Gaula. La Biblioteca Digital del Proyecto Mambrino, como todo el Proyecto Mapping Chivalry, se propone conectar con otros grupos de investigación en España y vincularse con los repositorios internacionales (Red Aracne).

- ELISABETTA SARMATI (LA SAPIENZA UNIVERSITÀ DI ROMA): La narrativa caballeresca española, tras el extraordinario florecimiento del siglo XVI, vivió un largo período de crisis, al menos hasta el umbral del Romanticismo cuando, con el redescubrimiento de la Edad Media como cronotopo literario, el mundo y el mito de la caballería volvieron a ponerse de moda. Posteriormente, el modelo narrativo de los libros de caballerías vivirá otras épocas de renovada fortuna tanto en el contexto de la denominada «literatura falangista», que recuperó el modelo medieval del caballero adaptándolo a la ideología del nacionalcatolicismo, como en época contemporánea, con ejemplos muy significativos en la novela española e hispanoamericana de los siglos XIX-XX. Con el proyecto Caballeresca y reescrituras en los siglos XIX-XXI, parte del más amplio “Mapping Chivalry. Spanish Romances of chivalry from Renaissance to 21st Century: a Digital approach”, nos proponemos a) crear una base de datos titulada AMADÍSSIGLOXX (ya en vía de construcción) de las reescrituras modernas de los libros de caballerías, incluyendo las reescrituras de Don Quijote con autores, obras y ediciones digitalizadas; b) comparar las nuevas estructuras narrativas, los motivos y los personajes de las escrituras contemporáneas con los paradigmas caballerescos del siglo XVI; c) estudiar los diferentes grados de reescritura del nuevo corpus: imitación, actualización del género medieval, parodia. En la presente comunicación se detallará el Proyecto con la presentación de los criterios de formación del corpus y del trabajo realizado.

- GIULIA TOMASI (UNIVERSITÀ DI TRENTO): Los libros de caballerías son un género de gran éxito del Renacimiento español que durante los siglos XVI y XVII se difunde en España para llegar luego a toda Europa y hasta a América. Este género editorial predomina en las imprentas de la época y sus características exteriores hacen que los textos sean reconocibles, especialmente por su gran tamaño. Los estudios críticos más recientes han puesto de relieve el importante papel que desempeña la reiteración de unos motivos comunes en sus extensas tramas, ya que a través de este enfoque narratológico puede resaltarse la evolución del género de los libros de caballerías durante el largo período de su auge en la escena editorial, además de proporcionar respuestas complejas sobre sus ricos contenidos narrativos. Así pues, a través de las herramientas que brindan las Humanidades Digitales, el proyecto MeMoRam tiene por objeto crear una plataforma que recoja los principales motivos de libros de caballerías castellanos. Este enfoque permite marcar la trayectoria de los más sobresalientes rasgos narrativos de uno entre los géneros que sientan las bases de la literatura europea moderna. La colección de los datos se apoya asimismo en instrumentos computacionales que aseguran resultados cuantitativamente rentables; entre estos son especialmente útiles las técnicas de *text mining*, como por ejemplo el *topic modeling*. A través de la creación de nubes de palabras, estas metodologías permiten aislar las secuencias relacionadas con motivos específicos para obtener una definición precisa de los mismos. El proyecto se está llevando a cabo con el soporte técnico de una compañía italiana líder en las Humanidades Digitales y la plataforma cuenta con una sección de fichas bibliográficas donde se reúnen los datos sobre las obras del corpus; otra sección está dedicada a los motivos que se desarrollan en dichas obras y en la última se organizan los personajes secundarios que intervienen en las narraciones vinculándose con los motivos detectados. Las búsquedas pueden efectuarse entrecruzando entre sí los distintos niveles. En la comunicación se van a detallar los pasos que nos han llevado a concebir la plataforma desde su diseño conceptual hasta su construcción y utilización; se presentarán las numerosas posibilidades de investigación que otorga una base de datos de este tipo en el campo de la literatura de caballerías y, en fin, se dará muestra de algunos ejemplos concretos de búsqueda de motivos recurrentes en el interior del corpus de libros de caballerías a nuestro alcance.

BIBLIOGRAFÍA

- Bellomi, Paola, Anna Bognolo, Stefano Neri y Federica Zoppi, *Repertorio delle continuazioni italiane ai romanzi cavallereschi spagnoli. Ciclo di Palmerino di Oliva*, Roma, Bulzoni, en prensa 2021.
- Bognolo, Anna, dir., *Ciclo italiano di "Amadis di Gaula."* Collezione della Biblioteca Civica di Verona, en Anna Bognolo, Paola Bellomi, Federica Colombini y Stefano Neri, eds., Verona, QuiEdit, 2012.
- Bognolo, Anna, Giovanni Cara y Stefano Neri, *Repertorio delle continuazioni italiane ai romanzi cavallereschi spagnoli. Ciclo di Amadis di Gaula*, Roma, Bulzoni, 2013.
- Bognolo, Anna y Stefano Bazzaco, «Tra Spagna e Italia: per un'edizione digitale del Progetto Mambrino», *eHumanista/IVITRA*, 16, 2019, pp. 20-36.
- Bueno Serrano, Ana Carmen, *Índice y estudio de motivos en los libros de caballerías castellanos (15081516)*, tesis doctoral dir. por Juan Manuel Cacho Bleuca, Universidad de Zaragoza, Filología Hispánica, (Literaturas Española e Hispánicas), 4 vols., 2007.
- Cacho Bleuca, Juan Manuel, «Introducción al estudio de los motivos en los libros de caballerías: la memoria de Román Ramírez», en *Libros de caballerías, (de «Amadís» al «Quijote»)*. Poética, lectura e identidad, Eva Belén Carro Carbajal, Laura Puerto Moro, María Sánchez Pérez (eds.), Salamanca, Seminario de Estudios Medievales y Renacentistas, Sociedad de Estudios Medievales y Renacentistas, 2002, pp. 27-57.
- Campos García Rojas, Axayácatl, y Yordi Enrique Gutiérrez Barreto, *En línea caballescica. Lecciones del Seminario de Estudios sobre Narrativa Caballescica*, UNAM, 2020.
- Crescini, Alice, «Progetto Mambrino. "Spagnole romanzerie": esemplari censiti nel 20182020», *Historias Fingidas*, 8, 2020, pp. 321-352.
- Crivellari, Daniele, «Luis Vélez de Guevara y los libros de caballerías: *El caballero del Sol*», en C. Demattè (ed.), *Il mondo cavalleresco tra immagine e testo*, Trento, Università degli Studi di Trento, 2010, pp. 199-214.
- Cruz Casado, Antonio, «El mundo cavalleresco medieval en algunas novelas españolas del siglo XX», *1616: Anuario de la Sociedad Española de Literatura General y comparada*, XII, 2006, pp. 97-106.
- Demattè, Claudia, *Repertorio bibliografico e studio interpretativo del teatro cavalleresco spagnolo del sec. XVII*, Trento, Università degli Studi di Trento, 2005.
- Gil González, Antonio Jesús, «El mito caballescico en la novela falangista: el *Amadís* de Angel María Pascual», en *Mitos: actas del VII Congreso Internacional de la Asociación Española de Semiótica (Investigaciones Semióticas VII) celebrado en la Universidad de Zaragoza del 4 al 9 de noviembre de 1996*, coord. por Alberto Navarro González, Juan Carlos Pueo Domínguez, Alfredo Saldaña Sagredo, Túa Blesa, vol. 2, Zaragoza, Asociación española de semiótica, 1998, pp. 472-475.
- Karsdorp, Folgert, Van der Bosch, Antal, «Identifying Motifs in Folktales using Topic Models», en *Proceedings of the 22 Annual Belgian-Dutch Conference on Machine Learning*, Nijmegen, 2013, pp. 41-49. [https://www.researchgate.net/publication/237049600_Identifying_Motifs_in_Folktales_using_Topic_Models]
- Leiva Ramírez de Arellano, Francisco, *Amadís y Niquea*, ed. crítica de Claudia Demattè, Kassel-Alicante, Reichenberger-Biblioteca Virtual Miguel de Cervantes, en prensa.

- Luna Mariscal, Karla Xiomara, «De Stith Thompson a las plataformas digitales: algunas reflexiones (con un Índice de motivos de la Demanda del Santo Grial, Toledo, 1515)», *Historias Fingidas*, 8, 2020, pp. 55-128.
- Mainer, José Carlos, *Falange y literatura*, Barcelona, RBA, 2013.
- Neri, Stefano, «Cuadro de la difusión europea del ciclo del Amadís de Gaula», en José Manuel Lucía Megías y María Carmen Marín Pina (eds.), *Amadís de Gaula, quinientos años después (estudios en homenaje a Juan Manuel Cacho Blecua)*, Alcalá de Henares, Centro de Estudios Cervantinos, 2008, pp. 565-591.
- Neri, Stefano, «Cuadro de la difusión europea del ciclo palmeriniano (siglos XVI-XVII)», en Aurelio González, Axayácatl Campos García Rojas, Karla Xiomara Luna Mariscal y Carlos Rubio Pacho (eds.), *Palmerín y sus libros: 500 años*, México D.F., El Colegio de México-Centro de Estudios Lingüísticos y Literarios, 2013, pp. 285-314.
- Neri, Stefano, «Progetto Mambrino. 'Spagnole romanzerie': esemplari censiti nel 2014-2015», *Historias Fingidas*, 3, 2015, pp. 211-222.
- Neri, Stefano, «Progetto Mambrino. 'Spagnole romanzerie': esemplari censiti nel 2016-2017», *Historias Fingidas*, 5, 2017, pp. 185-206.
- Pierazzo, Elena, *Digital Scholarly Editing. Theories, Models and Methods*, Londres-Nueva York, Routledge, 2020.
- Sarmati, Elisabetta, «Motivos caballerescos y reescritura paródica: el *Florestán del Palier* (1959) de Wenceslao Fernández Flórez», *Historias Fingidas*, 8, 2020, pp. 257-273.
- Thompson, Stith, *Motif-Index of Folk Literature*, Bloomington-London, Indiana University Press, 1955-58.
- Tomasi, Francesca, *Metodologie informatiche e discipline umanistiche*, Roma, Carocci, 2008.
- Tomasi, Giulia, «Las humanidades digitales y la base de datos MEMORAM: Para un enfoque sistemático hacia los motivos en los libros de caballerías», *Historias Fingidas*, 8, 2020, pp. 129-156.
[<https://historiasfingidas.dlcs.univr.it/article/view/155>]
- Voigt, Vilmos, Preminger, Michael, Ládi, Lászlá, Darányi, Sándor, «Automated Motif Identification in Folklore Text Corpora», *Folklore*, 12, 1999, pp. 126-141.
[https://www.researchgate.net/publication/26428179_Automated_Motif_Identification_in_Folklore_Text_Corpora]

Panel 6

Redes Sociales del Pasado. Cartografía de la Modernidad Literaria Hispánica, 1898-1959

Moderadora: Diana Roig Sanz - *Asiste:* César González-Pérez

Panelistas: Ainamar Clariana; Aurea Mota; Elisabet Carbó; Laura Fólica; Marc Moreno Galimany; Marvin Matías Agüero Torales; Pablo Suárez-Mansilla; Ventsi Ikoff

Redes Sociales del Pasado. Cartografía de la Modernidad Literaria Hispánica, 1898-1959

Coordinadora: Diana Roig-Sanz, Investigadora ICREA; IP del proyecto ERC StG *Social Networks of the Past. Mapping Hispanic and Lusophone Literary Modernity, 1898-1959*, IN3UOC.

Palabras clave: Iberoamérica, Instituto Internacional de Cooperación Intelectual, traducción literaria, crítica cinematográfica, mujeres mediadoras.

Descripción del problema de investigación, metodología y resultados

La minería de datos y las perspectivas de *big data* están cambiando el modo en el que generamos conocimiento, accedemos a la información y preservamos nuestro patrimonio histórico y cultural. Esta propuesta de panel se propone presentar la investigación que estamos llevando a cabo en el proyecto ERC *Social Networks of the Past. Mapping Hispanic and Lusophone Literary Modernity, 1898-1959*, que aplica herramientas digitales y análisis a gran escala para estudiar el papel de los(as) mediadores(as) culturales iberoamericanos y las redes literarias hispánicas y lusófonas en la modernidad internacional entre 1898 y 1959, marco cronológico que refiere la pérdida de Cuba por parte del imperio español y la Revolución Cubana. Como es sabido, la bibliografía sobre las redes literarias y los sistemas mundiales se ha inspirado en distintas aproximaciones (Bourdieu, Latour), pero la mayoría de análisis presentan Europa como el centro de los procesos literarios mundiales y el papel del sur de Europa o de América Latina se ha visto eclipsado. En los estudios literarios, es de sobras conocido que Moretti fue pionero en una aproximación cuantitativa basada en el análisis de redes, que otros han seguido. Sin embargo, todavía predominan los análisis cualitativos y la utilización de la minería de datos o *machine learning* no se han aplicado suficientemente en el contexto iberoamericano para redefinir redes, reevaluar actores e influencias, y poner a prueba las ideas recibidas sobre el valor literario, las instituciones o la posición de los productores culturales en el campo literario.

Con este contexto, las presentaciones de este panel se proponen contribuir a tres objetivos principales: 1) recuperar la historia perdida de un grupo numeroso de mediadores(as) culturales iberoamericanos(as) que participaron activamente en las redes interculturales y multilingües del modernismo (en el sentido anglosajón) para poder visibilizar y reevaluar su papel; 2) reducir la brecha de conocimiento en términos de acceso y producción, en el espacio iberoamericano, para generar datos nuevos y fiables que puedan contribuir a llenar los vacíos de nuestro patrimonio histórico y cultural, no siempre bien documentado; y 3) ofrecer un modelo innovador y reproducible que pueda ser aplicado a distintos periodos, lenguas, y disciplinas para analizar fenómenos culturales transfronterizos, mediadores culturales invisibilizados o poco conocidos, y redes y escalas geográficas poco atendidas en sus relaciones con el espacio global.

Para ello, los investigadores de este panel planteamos una doble metodología que se basa, por un lado, en 1) la construcción de una base de datos relacional para el análisis cuantitativo y cualitativo de la actividad, circulación e impacto de estos mediadores y mediadoras y, por otro, en 2) el estudio más detenido de cuatro procesos de transformación cultural que fueron distintivos de las sociedades modernas de la época. En concreto, la internacionalización de las culturas iberoamericanas a través del papel de los mediadores(as) iberoamericanos(as)

en las redes y actividades que estableció el Instituto Internacional de Cooperación Intelectual (un precedente de la UNESCO); la circulación de literatura traducida en una amplia selección de revistas iberoamericanas preservadas en las bibliotecas nacionales o en otros archivos relevantes como el Instituto Iberoamericano de Berlín; el papel de los mediadores(as) iberoamericanos(as) en la institucionalización de un campo de crítica cinematográfica, y el rol de mujeres mediadoras iberoamericanas en el ámbito literario, artístico, cinematográfico o científico a través del estudio de sus relaciones (sociales y epistolares) y de su movilidad. Hasta el momento, las investigaciones que proponemos se están desarrollando a partir de distintas fases que incluyen: identificación de todas las fuentes objeto de nuestro interés y catalogación preliminar; descarga de documentos y ocerización; búsqueda y extracción; análisis de los datos, y visualización. En este sentido, cabe señalar que hasta ahora hemos utilizado herramientas de SNA como Gephi o nodegoat.

Con respecto a nuestros resultados, algunos preliminares, otros en un estadio más avanzado, este panel mostrará avances relativos al uso de OCR y herramientas digitales para automatizar la indexación de la correspondencia preservada en el Instituto Internacional de Cooperación Intelectual y, en particular, la correspondencia relativa a las distintas comisiones iberoamericanas. También presentará importantes avances sobre la literatura traducida en las revistas iberoamericanas de la época, así como las soluciones planteadas a nivel técnico y legal para avanzar con los procesos de extracción, limpieza y análisis de datos. En una línea similar, apuntaremos algunas reflexiones sobre la constitución del campo de la crítica cinematográfica iberoamericana a través del caso de estudio de los objetos de cine en las revistas culturales de Uruguay antes de 1940. En concreto, abordaremos cuestiones técnicas relativas a la segmentación de la página durante el proceso de reconocimiento óptico de caracteres, pero también sobre reconocimiento de entidades nombradas, en particular, de nombres propios de autores o críticos. Cabe señalar que la creación de una base de datos única (en arquitectura relacional, por ejemplo, el sistema de base de datos PostgreSQL) nos permitirá hacer consultas cruzadas de todas las hemerotecas, tanto para la investigación sobre traducciones como para el trabajo sobre crítica, permitiendo identificar obras, autores(as), traductores(as), lenguas, críticos(as) y figuras relativas a la industria cinematográfica. Finalmente, la última presentación de este panel, a partir de la extracción de datos sobre las relaciones y contribuciones de un grupo de mujeres intelectuales iberoamericanas, mostrará algunos datos preliminares sobre sus desplazamientos y sobre las relaciones sociales y epistolares que establecieron y que contribuyeron a conformar los cánones de la modernidad iberoamericana.

Todos los resultados, en su conjunto, permitirán matizar la idea, demasiado repetida en la bibliografía que se ha publicado, sobre todo, fuera de nuestro entorno, de “centros” innovadores y “periferias” imitativas, y ofrecer un método original, reproducible y empíricamente probado para el estudio de la interacción social humana con un enfoque global, cultural y descentrado. La comparación entre España y América Latina nos permite evaluar redes intra- e internacionales y poner a prueba conceptos geográficos que muestran patrones menos habituales en la circulación de textos y agentes, así como distintos canales de transferencia. Por ejemplo, hoy podemos visualizar a gran escala la diversidad de autores traducidos en las revistas hispánicas y lusófonas y demostrar cómo algunas muy significativas traducciones españolas se publicaron, en algunos casos, con anterioridad a las traducciones francesas o inglesas. A partir de esos datos podemos, por supuesto, extraer gran cantidad de información sobre los mediadores implicados (traductores, editores, críticos, etc.).

Las 4 presentaciones que se proponen combinan métodos computacionales, historia literaria y cultural, sociología histórica, literaria y de la traducción, estudios literarios globales, estudios de género, análisis sobre la transferencia cultural y la mediación, estudios fílmicos, y el análisis de redes sociales, y son una nueva oportunidad para pensar en la *big history*, y promover un modo distinto de escribir la historia que modifique las

grandes narrativas y la historia *mainstream* y, en nuestro caso, que nos permita reevaluar el rol de Iberoamérica en los estudios sobre el modernismo, en el sentido anglosajón, con una amplia y rica diversidad de materiales. En este sentido, todas las investigaciones se basan en un trabajo coral que quiere ser descentrado, a gran escala, basado en una amplia y rica diversidad de fuentes y fundamentalmente empírico, así como en un análisis transnacional de los mediadores culturales y de sus múltiples roles para concluir que las mal llamadas “periferias iberoamericanas” fueron innovadores culturales clave por derecho propio.

Bibliografía seleccionada

Hitchcock, Tim. 2013. “Confronting the digital: or how academic history writing lost the plot,” *Cultural and Social History*, 10 (1), 9-23.

Moretti, Franco. 2013. *Distant Reading*. London: Verso.

Roig-Sanz, Diana, and Meylaerts, Reine (Eds.). 2018. *Literary Translation and Cultural Mediators in ‘Peripheral’ Cultures. Customs Officers or Smugglers*. London/ New York: Palgrave MacMillan.

1. “Cooperación intelectual y diplomacia cultural a través de *big data*: el caso del Instituto Internacional de Cooperación Intelectual”

Panelistas: Elisabet Carbó (IN3-UOC/ KU Leuven, Bélgica) y Marc Moreno (IN3- UOC)

Esta presentación se centra en el estudio de la historia del Instituto Internacional de Cooperación Intelectual (IICI, París, 1925-1946), antecesor de la UNESCO, cuyas actividades se desarrollaron en campos como el de la educación, la museología, el cine o la literatura y la traducción e implicaron a actores provenientes de los cinco continentes, desde Japón a Perú y desde Rusia a Sudáfrica (Renoliet 1999).

El archivo del IICI, conservado actualmente como parte del archivo de la UNESCO en París, constituye una fuente muy rica de información susceptible de interesar a investigadores de varios campos como son la historia intelectual, la historia global o los estudios literarios globales. El fondo incluye documentación (actas de reuniones, resoluciones, etc.), publicaciones (volúmenes monográficos, boletín interno, etc.) y la correspondencia. Sin embargo, su extensión -el archivo incluye más de 650 cajas y más de 800.000 páginas de publicaciones- hace que toda aproximación cualitativa sea inevitablemente parcial, característica que apunta al interés de explorarlo en el marco de las humanidades digitales. En este sentido, cabe señalar que se trata de un fondo parcialmente digitalizado, dado que la misma UNESCO inició dicha labor en 2018, siendo hoy posible acceder a una parte importante de los documentos a través de su página [web](#). Por tanto, el objetivo de nuestra investigación es contribuir a la digitalización de este fondo y, sobre todo, a su ocerización para explorarlo mediante herramientas más adecuadas para trabajar con *big data*, como son las herramientas de visualización de datos, el *topic modelling* o el análisis de redes históricas. Dichas metodologías se movilizan con el objetivo de identificar patrones hasta ahora desconocidos relacionados con la participación y el rol de actores considerados periféricos en el sistema de la Sociedad de Naciones, ya sean países o individuos.

En este sentido esta presentación se sitúa en clara filiación con el proyecto desarrollado por Grandjean (2018) en relación con el Comité Internacional de Cooperación Intelectual (CICI, Ginebra 1922-1946), pero se distingue de esta investigación en su propósito de profundizar en el uso de herramientas digitales, automatizando también la indexación de la correspondencia. Cabe señalar, no obstante, que las características del mismo fondo dificultan su procesamiento automático, ya que no presenta unidad en términos lingüísticos, siendo el francés la lengua más usada en un conjunto documental multilingüe. Por otro lado, incluye también documentos manuscritos y mecanoscritos en diferente estado de conservación. En este sentido, esta presentación mostrará, desde una perspectiva metodológica, los retos que estamos abordando en la fase de preprocesamiento, así como las soluciones planteadas. Encontrar un software OCR lo suficientemente dinámico como para proporcionar buenos resultados en este corpus histórico tan poco homogéneo ha sido sin duda un desafío. Por tanto, teniendo en cuenta que entrenar el OCR implica realizar una agrupación de las imágenes en subconjuntos según sus características principales, y que esto puede conducir a errores de reconocimiento, se ha optado por utilizar el modelo configurando aquellos parámetros más generales y escogiendo el resultado que mejor se adapte al conjunto.

Bibliografía seleccionada

Renoliet, Jean-Jacques. 1999. *L'UNESCO oubliée. La Société des Nations et la coopération intellectuelle (1919-1946)*. Paris: Publications de la Sorbonne.

Grandjean, Martin. 2018. "Les réseaux de la coopération intellectuelle. La Société des Nations comme actrice des échanges scientifiques et culturels dans l'entre-deux-guerres." Lausanne: Université de Lausanne, tesis de doctorado.

-----, 2015. "Introduction à la visualisation de données: l'analyse de réseau en histoire". *Geschichte und Informatik*, 18/19, 109-128.

Liebl, Bernhard and Burghardt, Manuel. 2020. "From Historical Newspapers to MachineReadable Data: The Origami OCR Pipeline." Leipzig: Leipzig University, Computational Humanities Group.

Roig-Sanz, Diana and Jaume Subirana (Eds). 2020. *Cultural Organizations, Networks and Cultural Mediators in Contemporary Iberoamerica*. London/New York: Routledge. Archivos de la UNESCO: <https://digital.archives.unesco.org/fr/>

2. "La traducción literaria en revistas digitalizadas iberoamericanas (1898-1959): ventajas e inconvenientes del trabajo con herramientas de las HD"

Panelistas: Laura Fólica (IN3-UOC) y Marvin M. Agüero-Torales (IN3-UOC)

Esta presentación se propone mostrar la investigación que estamos llevando a cabo para la detección y análisis a gran escala de traducciones en revistas literarias y culturales del campo iberoamericano (1898-1959). Por un lado, las revistas han sido un objeto de estudio dilecto a partir del giro cultural de la década de los ochenta en los

estudios literarios, sobre todo, en el campo latinoamericano. Estos trabajos se basan en un análisis de *close reading* de un corpus reducido de revistas, siempre dentro del ámbito nacional; se trata del análisis de un corpus abordable por el ojo humano en proyectos de investigaciones a cargo de un/a investigador/a. Por otro lado, la traducción publicada en revistas ha sido tomada en cuenta -aunque no de forma privilegiada- en los estudios de historia de la traducción como parte de la obra de un autor/a o traductor/a, pero desatendiendo la especificidad del medio en el que se publicaba. Nuestra investigación se propone integrar los aportes de ambas perspectivas, es decir, el estudio de la traducción presente en las revistas, a partir de un enfoque transnacional, a gran escala (temporal y espacial) a través de herramientas de las HD.

En la presentación explicaremos cómo hemos estructurado nuestro *workflow*, así como las decisiones metodológicas que hemos seguido en cada instancia y qué acciones prevemos para el futuro. Para componer un corpus amplio de revistas y gracias a los proyectos de digitalización de muchas bibliotecas, nos hemos centrado en las hemerotecas digitales de las distintas bibliotecas nacionales, que tienen gran cantidad de obras, aunque con calidad de dato menor. Para compensar este aspecto, hemos sumado repositorios digitales de grupos de investigación de revistas que cuentan con menos cantidad de obras, pero mejor etiquetadas o analizadas. En este punto, explicaremos los pros y contras encontrados tanto a nivel legal como técnico relativos a la descarga y tratamiento de los metadatos de las hemerotecas que hemos limpiado hasta el momento (con técnicas de web scraping, API o descarga directa y persistiendo las fuentes en bases de datos individuales del tipo SQLite). El objetivo final, una vez estandarizados todos los metadatos, es agregarlos a una base de datos única (en arquitectura relacional, por ejemplo, el sistema de base de datos PostgreSQL) que nos permita hacer consultas cruzadas de todas las hemerotecas sobre obras, autores, traductores, lenguas. Por otro lado, una vez acabado el vaciado de las hemerotecas y la descarga y estandarización de metadatos, contemplamos trabajar con la ocerización (en particular, con herramientas preparadas para lidiar con OCR histórico) de parte del material textual para estudiar algunos casos de traducción (en arquitectura de búsqueda de texto, por ejemplo, PostgreSQL Full Text Searching o Lucene-based). Con parte del corpus que disponemos hasta el presente, mostraremos el tipo de análisis que podemos realizar, así como las técnicas NLP utilizadas para su pre-procesamiento, a través de visualizaciones geográficas y relacionales. El objetivo final es que gracias al análisis masivo de los datos, desde una perspectiva transnacional y a gran escala, emerjan redes y actores hasta ahora poco estudiados por la historia literaria y que nos permitan descentrar relatos canónicos sobre circulación de traducción literaria entre los llamados “centros” y “periferias”, es decir, buscamos discutir la división entre lenguas centrales, que traducen poco y son traducidas, y lenguas periféricas, traductoras y poco traducidas.

A nivel teórico, esta presentación se inscribe en lo que hemos denominado Big Translation History (BTH) y que definimos como un análisis a gran escala (geográfico y cronológico); el uso de datos masivos que son entendidos como *big data*, pero también como *little data*, y con una amplia variedad de fuentes a menudo heterogéneas y poco estructuradas; y el uso de técnicas computacionales como parte del proceso de investigación y para la producción de conocimiento y no la mera presentación (Roig Sanz y Fóllica, *forth.*).

Bibliografía seleccionada

Borgman, Ch. L. 2015. *Big Data, Little Data, No Data. Scholarship in the Networked World.*

Cambridge, MA: MIT University Press.

Guzmán, Constanza, ed. 2019. *Translation and/in Periodical Publications. Special issue of*

Translation and Interpreting Studies 14 (2). <https://doi.org/10.1075/tis.00036.int>

Fólica, Laura, Diana Roig-Sanz, and Stefania Caristia, eds. 2020. *Literary Translation in Periodicals. Methodological challenges for a transnational approach*, Amsterdam/Philadelphia, John Benjamins, col. Benjamins Translation Library 155.

Phillips, J. P., & Tabrizi, N. 2020. Historical Document Processing: Historical Document Processing: A Survey of Techniques, Tools, and Trends. arXiv preprint arXiv:2002.06300.

Roig-Sanz y Fólica, L. (forth.). "Big Translation History. Data Science Applied to the Analysis of Translated Literature: the Spanish-speaking Case (1900-1945)".

SAGE Publishing. 2015. Cultural Sociology - Volume 9, Number 3, Sep 01, 2015. SAGE Journals. <https://journals.sagepub.com/toc/cus/9/3>.

Special Issue: Sociology of Literature in the Early 21st Century: Away From the Centre.

3. "Aproximación a los macroelementos para una historia de la crítica de cine a través de la lectura distante. El caso trasatlántico de Uruguay (1900 - 1940)"

Panelistas: Pablo Suárez-Mansilla (IN3-UOC) y Ventsislav Iloff (IN3-UOC)

Esta presentación mostrará los resultados preliminares de una lectura (semi)automatizada y distante de la crítica cinematográfica en revistas culturales uruguayas publicadas en la primera mitad del siglo xx. Este ensayo forma parte de una investigación de mayor calado cuyo propósito es revisar el campo de la crítica cinematográfica en Iberoamérica a partir de los mediadores culturales y la circulación internacional de las ideas. Las revistas culturales del pasado fueron testigo de las dinámicas del panorama cultural e intelectual y fueron los instrumentos de los procesos de modernización cultural. En este sentido, se convierten en el objeto idóneo para rastrear la historia y la crítica del cine como máxima expresión de la modernidad. No obstante, el volumen de revistas culturales que circularon y la duración variable de las mismas —algunas llegando a publicarse durante más de 30 años de vida— imposibilita una lectura cualitativa. Por eso, esta investigación se apropia de las capacidades de procesamiento de los sistemas de computación para ofrecer nuevas soluciones metodológicas y facilitar una lectura amplia y masiva de la crítica cinematográfica de la época. En concreto, procesaremos una selección de revistas uruguayas disponibles en formato digital en el portal *Anáforas*, creado por el Seminario Fundamentos Lingüísticos de la Comunicación de la Universidad de la República Uruguay. En el flujo de trabajo abordaremos los siguientes problemas o retos:

- Identificación de los artículos cuyo objeto es el cine. Este problema está asociado a la problemática de la segmentación de la página durante el proceso de reconocimiento óptico de caracteres, pero también a temas relativos al análisis textual, ya que éste nos permitirá clasificar los distintos fragmentos de una o más revistas.
- Reconocimiento de entidades nombradas, y particularmente de nombres propios de autores o críticos, para nombrar algunos, con el objetivo de identificar los críticos de cine y las figuras del mundo cinematográfico que estos citan.
- Experimentar con modelos estadísticos como *topic modelling* para determinar las nociones o los temas tratados por los críticos de cine en estas revistas.

La presentación mostrará los resultados preliminares de este trabajo, así como los desafíos y lecciones aprendidas del proceso de investigación.

En la dimensión teórica, esta investigación se inscribe en la posibilidad de estudiar las relaciones entre revistas, ideas y mediadores culturales, permitiendo el estudio de temas y conceptos acerca del cine que contribuyen a su estructuración y a la configuración de su historia industrial, intelectual y transnacional, así como a la forma del género de la crítica cinematográfica. Por otra parte, esta investigación da lugar a debates acerca del potencial de una aproximación historiográfica digital y de la lectura distante para la historia de la institución cinematográfica.

Bibliografía seleccionada

- Bourdieu, Pierre. 2002. "Les conditions sociales de la circulation internationale des idées." *Actes de la recherche en sciences sociales* 5 (145): 3–8. <https://doi.org/10.3917/arss.145.0003>.
- Duval, Julien. 2016. *Le cinéma au xxe siècle. Entre loi du marché et règles de l'art*. París: CNRS editions.
- — —. 2020. "Une république mondiale du film." *CONTEXTES* 28, 2020. [Disponible en: <https://journals.openedition.org/contextes/9222#quotation>].
- Moretti, Franco. (coord.) 2018. *Literatura en el laboratorio. Canon, archivo y crítica literaria en la era digital*. Barcelona: Gedisa.
- Sapiro, Gisele. 2017. "Champ." *Politika* [Disponible en: <https://www.politika.io/fr/notice/champ>].

4. "Mediadoras culturales en la modernidad: los cuerpos y espacios del arte, de la literatura y de la ciencia en el mundo Iberoamericano"

Panelistas: Aurea Mota (IN3-UOC) y Ainamar Clariana Rodagut (IN3-UOC/ Universidad de Marburg).

Partiendo del análisis de las contribuciones de la colombiana Soledad Acosta de Samper (1833-1913), la mexicana Lola Álvarez Bravo (1903-1993), la española María Luz Morales Godoy (1889-1980) y la brasileña Maria Augusta Generoso Estrela (1860-1946), este trabajo avanzará algunos resultados sobre los procesos de mediación cultural modernos que llevaron a cabo algunas mujeres intelectuales iberoamericanas desde distintas disciplinas: en el caso que nos ocupa, la literatura, el cine, el periodismo, o la medicina. El foco de esta presentación es estudiar, por un lado, los procesos de mediación cultural que se establecen por medio del desplazamiento físico o a nivel relacional de estas mujeres y los distintos lugares del mundo a los que viajaron, o las numerosas relaciones que establecieron personalmente o de forma epistolar con otras personalidades de su tiempo; y, por otro, arrojar luz, desde una perspectiva de género, sobre el estatus diferenciado que se ha atribuido a los personajes masculinos y femeninos por lo que respecta a los procesos de conexión que se originaron durante la modernidad. Mientras la relevancia de un personaje histórico se ha medido, a menudo, por el impacto que ha tenido este en la formación de los cánones hegemónicos dentro del dominio artístico/literario y científico, las contribuciones de los personajes de géneros otros (no-masculinos) han sido invisibilizadas, en muchos casos, hasta fechas muy recientes, en la historia de la modernidad. En consecuencia, la aportación intelectual de estos personajes no-masculinos no ha sido reconocida como parte estructural de los cánones que han constituido los campos científico, literario y artístico de la primera mitad del siglo XX.

Esta presentación se propone avanzar en la extracción de datos sobre las relaciones y contribuciones intelectuales de estas mujeres para la conformación de los cánones de la modernidad en Iberoamérica. En concreto, utilizaremos una de las bases de datos relacional del proyecto (*nodegoat*) para analizar las relaciones sociales de algunas de ellas y los patrones que pueden observarse en sus desplazamientos.

Desde una perspectiva teórica, la investigación discutirá brevemente la relación que existe entre la formación de la modernidad global y las experiencias de movilidad de los sujetos sociales que actuaron en los dominios científico y artístico. Frente a la idea de que la modernidad surgió en un espacio aislado y luego se extendió por todo el mundo, nosotras argumentamos que la experiencia moderna se constituyó a partir de la síntesis de prácticas / culturas que no se pueden reducir fácilmente al mundo occidental o, más concretamente, a Europa y EEUU.

Los procesos de desplazamiento definen el rol que las y los mediadoras tuvieron en la formación de las interconexiones entre los (mal)llamados centros y las (mal)llamadas periferias de la modernidad, y así nos ayudan a entender de manera más profunda cómo el fenómeno de la experiencia espacial moderna se ha formado y transformado. Se sostiene que los desplazamientos, desconexiones y discontinuidades, en el tiempo y el espacio, realizados por el ser humano forman parte de una dinámica efectiva y constituyen un vector fundamental en el proceso de transformación del mundo moderno. Como consecuencia de ello, es posible concluir que el emprendimiento epistémico-espacial moderno es más dinámico y descentralizado de lo que tradicionalmente se ha supuesto. Revisitar el rol central de las mujeres mediadoras culturales en los procesos de formación de la modernidad se considera una herramienta muy útil para el desarrollo de este trabajo de crítica histórica y para el establecimiento de nuevas posibilidades analíticas y metodológicas.

Bibliografía seleccionada

Derrida, J. 1978. *Writing and Difference*, London: Routledge.

Monteiro, R. H. 1997. *Inventing women science, technology and gender*. *Cadernos Pagu* (10), 445-448.

Méndez Rodens, A. 2014. *Transatlantic Travels in Nineteenth-Century Latin America: European Women Pilgrims*. Lewisburg: Bucknell University Press.

Mota, Aurea. 2018. 'On Spaces and Experiences: modern displacements, interpretations and universal claims', in *Annual of European and Global Studies*, Vol. 4, Edinburgh: Edinburgh University Press, pp. 161-182.

Sesión 8

La vía de la expresión

Estudios culturales III

Modera: Patricia Martín-Rodilla · *Asiste:* Alexandra Iglesias Cortés

- 53 Estéticas de los 80's. Análisis comparativo de los modos texto, imagen y sonido en 100 videos musicales
Alexis Castellanos
- 75 El videoblog dentro del ciberdiscurso juvenil
José Luis De Piero
- 76 Usted está acá. Proyecto de investigación y realización intermedial con soporte en software para dispositivos móviles
Esteban Prado, Mauro Carusso, Pablo Spennato

ESTÉTICAS DE LOS 80's. ANÁLISIS COMPARATIVO DE LOS MODOS TEXTO, IMAGEN Y SONIDO EN 100 VIDEOS MUSICALES

Alexis Castellanos-Escobar¹

Palabras clave: Videos musicales, estéticas, análisis textual, análisis visual, análisis sonoro.

Desde su aparición los videos musicales han tenido abordajes analíticos muy diversos, entre ellos se destacan las aproximaciones semióticas, narrativas y etnográficas a estos productos audiovisuales. El mejor ejemplo de ello lo revela el recientemente publicado "*The Bloomsbury Handbook of Popular Music Video Analysis*" (2019), un voluminoso manual con una variedad de perspectivas teórico-metodológicas para analizar las transformaciones creativas y cambios en la producción de *videoclips* que circulan hoy en diferentes plataformas mediáticas. Empero, el giro computacional en la investigación y las Humanidades Digitales han planteado nuevos enfoques que potencializan la lectura/visualización/escucha distante, fragmentada y colaborativa frente a las prácticas investigativas que privilegian la lectura/visualización/escucha cercana, lineal y solitaria.

Asimismo, la llegada de tecnologías como la computación visual, la automatización y la analítica de macrodatos han contribuido a novedosas aproximaciones en el estudio de productos audiovisuales. Sin embargo, como bien señalan los editores del monográfico *AudioVisual Data in DH* publicado este año "la cultura del texto y la palabra ha disfrutado de una posición dominante en las Humanidades Digitales" (Arnold, Scagliola, Tilton & Van Gorp, 2021). Lo que suscita el interés en este trabajo sin desconocer el valor de una serie de esfuerzos históricos, entre ellos el análisis multimodal (recurriendo a la codificación manual a través de un software), la cinemétrica y recientemente la analítica cultural², abordajes que han desarrollado métodos y herramientas de estudio para materiales audiovisuales que intentan superar el acérrimo logocentrismo.

Descripción del problema de investigación

En esta comunicación se presentan los resultados de tres procedimientos realizados con un conjunto de datos extraídos del medio/ambiente YouTube pertenecientes a reconocidos artistas de los años ochenta del siglo pasado incluidos por MTV en su conteo de los 100 mejores videos musicales de la década transmitido el 24 de noviembre de 1989³. A partir de dichos datos se desarrolló un análisis comparativo de textos con las letras de las canciones usando Voyant Tools⁴, un análisis comparativo de imágenes fijas o fotogramas de los videos con ImageSorter⁴ y un análisis comparativo de sonidos y audios a través de Sonic Lineup⁵. Usamos la visualización como metodología para explorar grandes colecciones de datos, ver patrones a diferentes escalas, confrontar sesgos/estereotipos, hacer descubrimientos y aplicar operaciones digitales/computacionales.

¹ Profesor Asistente del Departamento de Comunicación, Pontificia Universidad Javeriana. Contacto: alexis.castellanos@javeriana.edu.co

² Según Manovich (2020) la analítica cultural es una caja teórica-metodológica que permite el estudio de contenidos en línea, producción de la cultura a escala global, ver cambios/flujo de artefactos, estilos, técnicas y usos de los medios en el tiempo, así como trabajar las características de diferentes objetos comunicativos.

³ Para el análisis exploratorio se tomó como punto de partida el MTV Top 100 Videos of the 1980s. Disponible en: <https://www.imdb.com/list/ls022380614/> (consultado: 14 de febrero de 2021). ⁴ La aplicación Voyant Tools fue desarrollada por los investigadores canadienses Stéfan Sinclair y Geoffrey Rockwell para el procesamiento de grandes corpus de datos textuales. Disponible en: <https://voyant-tools.org/>

⁴ El software libre llamado ImageSorter es un desarrollo construido por el laboratorio de Computación Visual HTW en Berlín-Alemania. Disponible en: <https://visual-computing.com/project/imagesorter/>

⁵ La aplicación Sonic Lineup es un software construido por el Centro de Música Digital de la Universidad Queen Mary de Londres, para musicólogos, investigadores y cualquier otra persona que esté interesado en explorar y visualizar archivos de audio. Disponible en: <https://sonicvisualiser.org/sonic-lineup/>

Para este caso, nos preguntamos sobre lo raro, lo recurrente y lo relevante en los 100 vídeos musicales a partir de un acercamiento a los datos textuales derivados de las letras de las canciones, los datos visuales de los fotogramas de los videos musicales y los datos sonoros extraídos del vídeo musical. Acudimos a YouTube para el acceso al conjunto de datos, reconociendo que la digitalización del material afectó los colores, la calidad de la imagen y el sonido, así como la resolución de los archivos descargados que fueron posteriormente procesados los software de análisis.

De acuerdo al ecólogo de los medios Sergio Roncallo-Dow (2012) los abordajes frecuentes para encarar un fenómeno como YouTube en tanto objeto de estudio, están directamente relacionados con sus *contenidos*, pero también como *medio*, especialmente atractivo por sus características en lo audiovisual y lo narrativo. De modo que los videos musicales en tanto artefacto estético combinan textos, imágenes y sonidos siendo un producto multimodal y un dispositivo narrativo influyente en la vida cotidiana a través de su ubicuidad en diferentes pantallas y tecnologías móviles. Desde su aparición masiva en canales dedicados a la promoción de la música como MTV, VH1 o Much Music, la velocidad, el montaje y la fragmentación formarán parte de las estéticas de nuestro tiempo (Roncallo-Dow & Uribe-Jongbloed, 2017) en la producción de *videoclips* narrativos, no narrativos y videoregistros. Asimismo, el *remix*⁶ o reciclaje cultural (Navas, 2018) es el patrón estético en la confección de propuestas audiovisuales en el campo de la música donde se evidencian prácticas de apropiación, combinación, resignificación de formas preexistentes en la edición de textos, imágenes y sonidos en los procesos de creación.

Por último, uno de los desafíos que presentó esta pesquisa consistió en dividir por modos comunicativos los videos musicales en tanto formas de naturaleza multimodal, teniendo en cuenta que los recursos semióticos se yuxtaponen en la experiencia estética y perceptiva del espectador. Desde la perspectiva teórica de la multimodalidad Kress (2010) y Kress & Van Leeuwen (2006) hacen referencia al uso de varios modos semióticos en el diseño de productos simbólicos, en él desempeñan un papel determinante las combinaciones modales y refuerzos mutuos cumpliendo roles complementarios. Los modos son elementos semióticos que permiten la realización simultánea de discursos y tipos de (inter)acción de acuerdo con los intereses de una situación comunicativa particular y pueden ser lenguajes, imágenes, músicas, sonidos, gestos, entre otros⁷.

Metodología

El abordaje metodológico que se implementó en esta investigación apostó por la integración de métodos computacionales para la consecución del corpus, visualización y su posterior análisis comparativo por cada uno de los modos: textos, imágenes, sonidos. Se tomó en cuenta la visualización de medios planteada por Manovich (2020) en tanto paradigma analítico que permite nuevos descubrimientos, evitando la reducción extrema y la hegemonía de las variables espaciales. En palabras de la investigadora Maria Giulia Dondero (2019) la visualización de medios implementada en los proyectos del Cultural Analytics Lab es “una operación metavisual, donde las imágenes son visualizadas y analizadas como imágenes” (p.1).

Si bien descubrir patrones, documentar la diversidad/singularidad y dinámica de los videos musicales, así como el mapeo a escala de influencias estéticas no ha sido una prioridad a través de los métodos clásicos, en los últimos años se han incorporado el uso de dispositivos computacionales para poder automatizar grandes conjuntos de datos, trabajar con metadatos y extraer características numéricas de los mismos. Ejemplos de ello son los esfuerzos del Multimodal Analysis Lab liderado por la profesora Kay O'Halloran en donde desarrollan y

⁶ Los Estudios de remezcla (Remix Studies) analizan la acción creativa y la práctica cultural siendo un campo de estudio útil para entender cómo y por qué producimos cosas nuevas a través del reciclaje o la reutilización de ideas de producciones materiales/simbólicas.

⁷ Para Kress & Van Leeuwen (2001) la comunicación humana es por naturaleza multimodal siendo fundamental los modos: verbales, visuales, sonoros, olfativos, táctiles y gustativos en la creación/producción/interpretación de significado y en los conocimientos socialmente construidos de la realidad (discursos).

comercializan software para el análisis de imágenes, videos, sitios web y medios interactivos. Empero, una de las apuestas académico-políticas de las Humanidades Digitales está relacionada con el acceso abierto y gratuito, la generación de comunidades de práctica, el uso de herramientas y software libre permitiendo un mayor alcance e impacto no solo en espacios académicos, sino también fuera de ellos.

Por tanto, las estrategias de visualización para el análisis de los 100 videos musicales están inspiradas en Manovich (2020), Arnold & Tilton (2019), Niederer & Colombo (2019) quienes trabajan en sus laboratorios ejercicios con microvistas, mesovistas y macrovistas bajo la apuesta del acceso abierto y con herramientas generativas no privativas⁸. El trabajo con este tipo de procedimientos analíticos permite visualizar las influencias de los productores audiovisuales, los cambios, similitudes y diferencias en la música, los temas y las transformaciones visuales y estéticas en los vídeos, pero a su vez advertir detalles y singularidades de cada uno de los *videoclips*.

Para esta investigación el conjunto de datos culturales seleccionado fueron 100 *videoclips* destacados por el famoso conteo realizado por la cadena televisiva MTV de los 100 mejores *videoclips* de los 80's. La decisión de tomar este período está fuertemente vinculada con el éxito del formato en el mundo audiovisual, su rápido avance, masificación y posicionamiento en las audiencias juveniles a través canales televisivos dedicados exclusivamente a la programación de contenidos musicales. De acuerdo con la *Figura 1. Búsqueda en Google Ngram de la entrada "Music video"* se puede advertir el crecimiento de este tipo de artefacto cultural y comunicativo, así como su alta penetración en el consumo audiovisual gracias a la televisión y a medios digitales con un alcance global.

Resultados obtenidos

Luego de la recolección de los datos en el sitio web Internet Movie Database - IMDb, ver *Anexo 1. Ficha técnica MTV Top 100 Videos of the 1980s* se desarrolló un ejercicio exploratorio a través de la herramienta Tableau Public en las que se usaron visualizaciones estándar como diagramas de barras verticales/horizontales, burbujas agrupadas, entre otras. Pese a los sesgos que pueda tener este conteo realizado por el canal MTV nos brinda una idea de la popularidad de los videos, las preferencias y gustos de los públicos en el consumo musical audiovisual. En una primera lectura se advierte la presencia de Madonna 5 videoclips, U2, Michael Jackson, Def Leppard con 4 videos musicales siendo los artistas con mayor número de frecuencias en el conteo y constituyendo el 17% de la muestra. Dos de ellos dedicados al género del pop y dos bandas de rock provenientes de las islas británicas. *Figura 2. Número de videos musicales por artista.*

Entre los directores con mayor número de *videoclips* se encuentran Marty Callner, Wayne Isham, Steve Barron y Nigel Dick. Todos ellos con una gran experiencia en la industria audiovisual, cine y televisión en el ámbito anglosajón. Llama la atención que la única directora es Mary Lambert reconocida por su trabajo con Madonna, Janet Jackson, Whitney Houston, artistas femeninas que fueron exitosas en la década de los 80. *Figura 3. Número de videos musicales por director(a).* En cuanto a la distribución de los videos musicales dentro del periodo de estudio, se pueden identificar dos picos, el primero en 1983 y el segundo en 1987, siendo este último el más alto con 19 *videoclips* del listado. *Figura 4. Número de videos musicales por año.*

Dentro de los géneros se destacan la versatilidad del rey del pop Michael Jackson con videos musicales que mezclan terror, crimen y cierta dosis documental. También sobresalen los *videoclips* animados de los artistas musicales Peter Gabriel, Dire Straits y la agrupación noruega A-Ha. El primero, empleaba técnicas de plastimación, pixilación y *stop motion* y fue dirigido por el experimentado animador y pintor Stephen R. Johnson,

⁸ El orden en el flujo de trabajo que implementamos comenzó con la extracción de los datos, la limpieza y estandarización, la construcción de una ficha técnica para luego generar visualizaciones macro con exploraciones de datos estructurados y no estructurados, visualizaciones micro y meso del *corpus*.

contando con un equipo de trabajo entre los que se encontraban animadores, escultores y titiriteros. Por su parte Money For Nothing es un video con rotoscopia, personajes animados en 2D y efectos especiales que se entrelazan con las imágenes de una actuación en vivo. Mientras tanto, el grupo A-Ha en 1985 buscando la proyección de la banda a nivel internacional decide trabajar con Michael Patterson en un nuevo video musical con rotoscopia que integra perfectamente imágenes y narrativa de una historia con los músicos.

En duración los videos de Michael Jackson y las bandas Guns N' Roses y Yes tienen la mayor extensión con 7 minutos sosteniendo un estilo cinematográfico que tiene preludios e intermedios que marcan una nueva narrativa del *videoclip* mezclando historias de amor, múltiples locaciones con una gran inversión en producción y postproducción. Entre los menos extensos están el famoso dueto de

David Bowie & Mick Jagger con Dancing In The Street y el trabajo del pianista Herbie Hancock en un video musical con marionetas y efectos especiales que curiosamente es la canción con el texto más corto con tan solo 16 palabras.

Posteriormente, se elaboraron macrovistas para advertir patrones textuales, visuales y sonoros, tendencias, relaciones, variaciones y excepciones con los datos no estructurados (textos, imágenes, sonidos). La visualización distante fue profundamente valiosa para poder comprender las estéticas de los años 80 a partir de los videos musicales. El conjunto de datos textuales tiene 100 documentos con 29,999 total de palabras y 2,394 formulario de palabra única. Las canciones con líricas más extensas son: Parents Just Don't Understand (920); Red Red Wine (606); Man In The Mirror (547) destacando que la primera es un rap americano y por su constitución estructural tendrá un mayor volumen de texto. Por su parte, los videos musicales con menor número de palabras en sus canciones fueron: Rockit (16); Wild Thing (109); Sharp Dressed Man (121). *Figura 5. Macrovista de los datos textuales con Voyant Tools.*

La apuesta cromática predominante está compuesta por colores oscuros (especialmente negro como patrón estético y técnico por los fundidos), rojos y azules muchos de ellos provenientes de las luces y ambientación de las actuaciones en vivo de los artistas, que a modo de videorregistro se utilizan en la edición para el ensamblaje del *videoclip*. *Figura 6. Macrovista de los datos visuales en Image Sorter.* Desde la perspectiva del sonido se presentan sintetizadores, *riffs* de guitarras eléctricas y baterías electrónicas con un repertorio de efectos, *loops* y *samples*. El amor es el gran tema de las letras de las canciones y sonoramente esto se ve musicalizado a partir del uso de acordes menores en una estructura de estrofa 1, pre-coro, coro, estrofa 2, pre-coro, coro, puente, coro. *Figura 7. Macrovista de los datos sonoros en Sonic Lineup.*

Por su parte en las mesovistas, se desarrollaron clústeres (conjuntos, grupos, series) de datos que se organizaron y clasificaron por tipologías. En los datos textuales provenientes de las letras de las canciones la palabra más frecuente es "love" con 339 recuentos en el corpus, considerando que en el idioma inglés "love" es sustantivo y verbo. Los sustantivos frecuentes son "baby", "heart" y "time" mientras que en verbos aparecen en los primeros lugares "just", "got", "like", "know" "want" y "make". Las conjugaciones como "I'm", "It's" y "you're" son las más empleadas en las líricas de las canciones de los años 80. Ocupan un lugar privilegiado las inflexiones y expresiones vocales propias de la lengua anglófona "oh" y "yeah", empleándose como *ad-libs* para acompañar solos de guitarras, espacios entre el puente de la canción, o el remate de la misma. *Figura 8. Mesovista de los datos textuales en Voyant Tools.*

En las visualizaciones realizadas a partir de las mesovistas y acercándose a los fotogramas de los videoclips, se evidencia la estrategia "banda" en la cual se dedican encuadres y tomas a cada uno de los integrantes de las agrupaciones musicales: vocalista, guitarrista, bajista, baterista. Ejemplo de ello es el caso Bon Jovi, Def Leppard, Aerosmith y Guns N' Roses. *Figura 9. Mesovista de los datos visuales en Image Sorter.* En el sonido hay particularidades como efectos con teclados y reverberación de instrumentos, armonización de voces con

distancias de tercera y quinta, diálogos al inicio de los videos derivados de la narrativa ampliada que realizan los productores audiovisuales en los *videoclips*. *Figura 10. Mesovista de los datos sonoros en Sonic Lineup.*

Las microvistas permitieron estudiar la diversidad cultural, detectar características específicas, rasgos particulares y singularidades de cada *videoclip* en una lectura cercana. Desde allí se pueden apreciar la variedad de géneros de los videos musicales a través de videoclips narrativos, no narrativos, videoregistros, animación y actuaciones en vivo. En lo textual, las letras en su mayoría se vinculan con el discurso del amor y el desamor en el pop y el rock como estilos musicales. *Figura 11. Microvista de los datos textuales en Voyant Tools.*

En la imagen los cambios creativos, estéticos y culturales sin duda están relacionados directamente con las innovaciones tecnológicas. Las técnicas audiovisuales utilizadas son cámaras de vídeo que permiten a los artistas producir su material promocional de mejor calidad y en menor tiempo. Visualmente, se privilegian las estéticas de las bandas de rock con sus cabelleras voluminosas, largas melenas en hombres y mujeres con chaquetas de cuero, tenis y moda epocal. También se encontraron en los videos musicales el (re)uso de transferencias estéticas y culturales de otros medios, entre ellos la televisión, el cine, los cómics y los dibujos animados. Se recuperan lenguajes y elementos, como por ejemplo secuencias de baile de los clásicos musicales de Hollywood que aparecen en los videos musicales de estrellas pop (Madonna y Michael Jackson). El *remix* como un elemento estético primordial tiene un marcado protagonismo en el desarrollo de las propuestas audiovisuales de la década de los años 80 y se presenta en las paletas de color, uso de la iluminación, planos/encuadres en las imágenes, y las fuentes de imaginación audio(visual). *Figura 12. Microvista de los datos visuales en Image Sorter.*

Los programas de postproducción toman mayor protagonismo, no obstante esto tendrá vínculos con las tecnologías de grabación y edición, el software, los efectos especiales y gráficos animados con la inclusión del ordenador en la confección de los videos musicales. Revisando los archivos de sonido existen patrones visuales y auditivos repetitivos obvios, por ejemplo el uso de efectos de audio en guitarras eléctricas, sintetizadores y voces. Asimismo, la velocidad son canciones de más de 90 BPM, los ritmos dominantes son pop y rock con mezclas de reggae, R&B, house, hip hop e incluso jazz. *Figura 13. Microvista de los datos sonoros en Sonic Lineup.*

Por supuesto, analizar vídeos musicales, plantea retos para los humanistas digitales, en tanto el acercamiento a los modos comunicativos. De esta manera, al trabajar con diferentes capas expresivas se tuvo cuidado en que los recursos imagen, texto, sonido fueran analizados como contenidos y a su vez como datos. Esta es una entrada bífida al conjunto de 100 videos musicales, dado que por un lado al revisarlos en su contenido, se pueden notar de forma individual características estilísticas y estéticas en lo textual, lo visual y lo sonoro, pero al cotejarlos como datos, era posible advertir patrones, agrupar por series y tipologías, reconocer las similitudes/diferencias visuales (color, textura, luminosidad, composición), examinar las frecuencias/infrecuencias, continuidades/discontinuidades y flujos/ritmos.

Desde el punto de vista técnico el uso de herramientas digitales permitió la revisión del corpus en un menor tiempo de lo que nos hubiera tomado codificando manualmente cada uno de los vídeos. Las limitaciones en el uso de aplicaciones digitales como Voyant Tools corresponden con la ausencia del contexto histórico, la dificultad de las variaciones idiomáticas, lingüísticas y semánticas de los datos. Adicionalmente, en el procesamiento de los datos audiovisuales por su tamaño, especialmente con los fotogramas, generó conflicto con la aplicación ImageSorter por su capacidad de procesamiento. Sobre el uso de Sonic Lineup, las reflexiones están relacionadas con el análisis y representación del sonido como sonido (sonificación en lugar de visualización). Este es un punto clave de discusión para las Humanidades Digitales formulado por Everardo Reyes (2014) y que sigue teniendo vigencia, dado el auge de la visualización en tanto técnica y paradigma investigativo.

Si bien la relación imagen-texto-sonido tiene una larga tradición en las humanidades, existe una brecha semántica al dividir por modos comunicativos. Para intentar mitigar este fenómeno, la propuesta tuvo la posibilidad de

triangulación del análisis textual (lectura distante, fragmentada y colaborativa), el análisis visual (visualización distante, fragmentada y colaborativa) y el análisis sonoro (escucha distante, fragmentada y colaborativa). Tradicionalmente, la lectura humana interpreta textos, distingue géneros, capta estilos; mientras la lectura computarizada reconoce una secuencia de letras separadas por espacios, la densidad léxica de las letras de las canciones y visualiza en nubes de palabras las frecuencias en los datos textuales. Por su parte la visión humana reconoce características y relación figura-fondo, mientras la visión computarizada explora patrones de color, elementos formales de similitud y el número de bits que componen la imagen. Por último, la escucha humana reconoce ritmos, estilos y se emociona, mientras la escucha computarizada identifica patrones rítmicos y estructuras.

Referencias

- Arnold, T., Scagliola, S., Tilton, L. & Van Gorp, J. (2021). *Introduction: Special Issue on AudioVisual Data in DH*. DHQ: Digital Humanities Quarterly. Volume 15, Number 1.
- Arnold, T. & Tilton, L. (2019). *Distant Viewing: Analysing large visual corpora*. Digital Scholarship in the Humanities, 34, 1 (2019): pp. 3-16.
- Burns, L. & Hawkins, S. (2019). *The Bloomsbury Handbook of Popular Music Video Analysis*. New York: Bloomsbury.
- Dondero, M. (2019). *Visual semiotics and automatic analysis of images from the Cultural Analytics Lab: How can quantitative and qualitative analysis be combined?*. Semiótica, 2019 (230), pp. 121-142. <https://doi.org/10.1515/sem-2018-0104>.
- Kress, G. (2010). *Multimodality: A Social Semiotic Approach to Contemporary Communication*. New York: Routledge.
- Kress, G. & Van Leeuwen, T. (2006). *Reading Images: The Grammar of Visual Design*. New York: Routledge.
- Kress, G. & Van Leeuwen, T. (2001). *Multimodal Discourse. The modes and media of contemporary communication*. London: Arnold.
- Manovich, L. (2020). *Cultural Analytics*. Cambridge: MIT Press.
- Navas, E. (2018). *Art, Media Design, and Postproduction: Open Guidelines on Appropriation and Remix*. New York: Routledge.
- Niederer, S. & Colombo, G. (2019). *Metodologías visuales para imágenes conectadas en red: diseño de visualizaciones para la investigación colaborativa, el análisis multiplataforma y la participación pública*. Diseña, (14), pp. 40-67. <https://doi.org/10.7764/disena.14.40-67>
- Reyes, E. (2014). *Explorations in Media Visualization*. HT'14: Santiago. Disponible en: http://ceur-ws.org/Vol-1210/datawiz2014_11.pdf
- Roncallo-Dow, S. (2012). *YouTube como medio aurático*. En: Codificar/Decodificar. Prácticas, espacios y temporalidades del audiovisual en Internet. Bogotá: Pontificia Universidad Javeriana.
- Roncallo Dow, S., & Uribe-Jongbloed, E. (2017). *La estética de los videoclips: propuesta metodológica para la caracterización de los productos audiovisuales musicales*. Cuadernos de Música, Artes Visuales y Artes Escénicas, 12 (1).

<https://doi.org/10.11144/Javeriana.mavae12-1.evpm>

Figuras

<https://docs.google.com/presentation/d/1jIElftpSnei8VkJGGtPU4xcb04PkKcxL7ElmBIUrs/edit?usp=sharing> Fuente: Elaboración propia.

Tabla 1. Ficha técnica MTV Top 100 Videos of the 1980s

<https://docs.google.com/spreadsheets/d/1ZpXa4p7mT9AnfLoteYINtGtzy4Rfab5ZTc16tYhtTOU/edit?usp=sharing>

Fuente: <https://www.imdb.com/list/ls022380614/>

El videoblog dentro del ciberdiscurso juvenil

José Luis De Piero

INVELEC – CONICET – UNT

Introducción

Cuando hablamos del lenguaje empleado en la web, de las formas de interacción y de las formas de comunicación, surge un problema sobre las maneras de denominar a los fenómenos que allí ocurren. Muchas veces, se apela a fenómenos conocidos del mundo analógico o no mediado por la tecnología computacional y se le agrega el adjetivo “digital” para diferenciarlo de su contraparte. Sin embargo, no siempre se emplean estos términos para designar a un mismo fenómeno con la salvedad del entorno en el que ocurre.

Con el discurso digital y su clasificación ocurren cosas similares. ¿El periodismo digital es “periodismo” que ocurre “sólo en el mundo digital”? ¿Comparte rasgos con el periodismo no digital? ¿A qué conjunto de características se refiere ese “digital”: sólo al medio o también a rasgos discursivos y lingüísticos? ¿Qué pasa con los géneros discursivos que nacieron y sólo tienen existencia en el mundo digital?

Con estas preguntas en mente nos proponemos este repaso por las distintas maneras de nombrar al fenómeno de la comunicación en la web, revisando sus alcances y limitaciones y luego tratar de acotar la definición para el abordaje de un fenómeno que sólo ocurre en el mundo digital: el videoblog.

Primero que nada, es necesario indicar que estamos hablando de discursos en la web, de cierto tipo de fenómenos, que identificamos como fenómenos discursivos, y que tienen lugar en el ciberespacio.

Entendemos por discurso, siguiendo a Fairclough (2008, 48) a “un modo de acción, una forma por la cual la gente puede actuar sobre el mundo y especialmente sobre otros”. Y también que el discurso “está formado y restringido por la estructura social (...) [al mismo tiempo que] es socialmente constitutivo”

Es decir, una práctica de uso del lenguaje socialmente constituida y constituyente, que es un modo de acción en sí misma.

Cuando ingresamos en el ámbito del ciberespacio, la terminología aplicada para referirse a los usos del lenguaje se vuelve muy variada, pero intentamos reducirla a tres grandes grupos:

1. Términos referidos a la comunicación
2. Términos referidos al discurso
3. Términos referidos al lenguaje.

Términos referidos a la comunicación

En el primer grupo encontramos un conjunto de posicionamientos más tradicionales, vinculado a los orígenes de la web y de las formas discursivas que aparecieron en la llamada Web 1.0. El término más empleado es el de “Comunicación Mediada por Computadoras” o CMC, aunque también se encuentran variedades como “Comunicación mediada por la Tecnología” (CMT) o “Comunicación Digital Interactiva”.

Estas definiciones aluden, en general, a fenómenos comunicativos entre dos usuarios. Se observa una etapa más primitiva, como la que recuperan Thurlow et. al. (2004, 14) de Santoro (1995,11), como una comunicación entre dos computadoras, sin que necesariamente se vea involucrada la persona humana.

Las otras definiciones ya implican la existencia de dos sujetos humanos que se involucran mediante la mediación de una computadora o equivalente (como son los teléfonos celulares).

Lo que tienen en común estas definiciones son el hecho de pensar en un contexto amplio en el cual se produce un fenómeno comunicativo que tiene rasgos particulares por estar mediado tecnológicamente.

Términos referidos al discurso

Este conjunto agrupa definiciones que parten de una postura discursiva, con las herramientas propias que brinda el Análisis del Discurso, para analizar las formas de uso de la palabra mediada por la tecnología. No ya a la comunicación como un proceso general o un entorno necesario, sino a la toma de la palabra posibilitada por los nuevos medios.

Aquí encontramos dos conceptos contrapuestos a los que nos referiremos en detalle más adelante que son el de "Discurso Digital" (también bautizado como Discurso Mediado por Computadoras, CMD) y "Ciberdiscurso".

Términos referidos al lenguaje

Este conjunto de términos proviene de estudios vinculados a la sociolingüística y hace referencia a los modos en los que el contacto de lenguas en el ciberespacio está produciendo variaciones y cambios lingüísticos en los distintos idiomas.

Aquí los conceptos centrales son los de "ciberlenguaje" y "lenguaje en-línea"

En la siguiente tabla presentamos los autores más importantes con el término acuñado y el año de aparición de una publicación de referencia, según los grupos que indicamos anteriormente.

Grupo	Término	Autor	Año
Comunicación	Comunicación Mediada por Computadoras	Thurlow, Lengel & Tomic	2004
	Comunicación Mediada por Computadoras	Santoro en Th et al	1995
	CMT	Parini y Giammatteo	2014
	Comunicación Digital	Parini y Giammatteo	2016
	Comunicación Digital Interactiva	Scolari	2008
Discurso	Discurso Digital	Cantamutto & Vela Delfa	2016
	Discurso Mediado por Computadora	Herring	2004
	Ciberdiscurso (juvenil)	Palazzo	2010
	Discurso de los nuevos medios	Aitchinson & Lewis	2003
Lenguaje	Ciberlenguaje	Vaqueiro Romero	2012
	Online language variation and change	Paolillo	1999

	Ciberlanguage	Androutsopoulos	2013
--	---------------	-----------------	------

Tabla 1. Conceptos vinculados al discurso en la web. Fuente: Elaboración Propia

Podemos ver cómo los términos vinculados a tradiciones más textuales son más antiguos que los que tienden a incluir la variación y el cambio lingüístico o el análisis de la lengua oral, más tardíos no sólo por las posibilidades técnicas de la web 2.0 que ahondó en la cultura de la participación con fragmentos discursivos audiovisuales llamados multimodales, sino también por el renovado interés en estos fenómenos y las nuevas posibilidades de análisis de estos discursos.

Los posicionamientos, también, se deben a factores como las áreas de la que provengan los estudios. Es normal encontrar el concepto de “discurso digital” en áreas como el periodismo o incluso la ingeniería, mientras que CMC es más frecuente en áreas vinculadas a los estudios del lenguaje.

El ciberdiscurso vs el discurso digital

Herrin (2004) Define al discurso digital como “la comunicación producida cuando los seres humanos interactúan entre sí transmitiendo mensajes a través de computadoras conectadas.” Se distingue además la importancia de observar y focalizar el lenguaje y el uso lingüístico en los medios digitales.

Esta definición pone el foco a la comunicación mediada por computadoras, enfocándose en el lenguaje, pero poniendo a la red sólo como un espacio de comunicación.

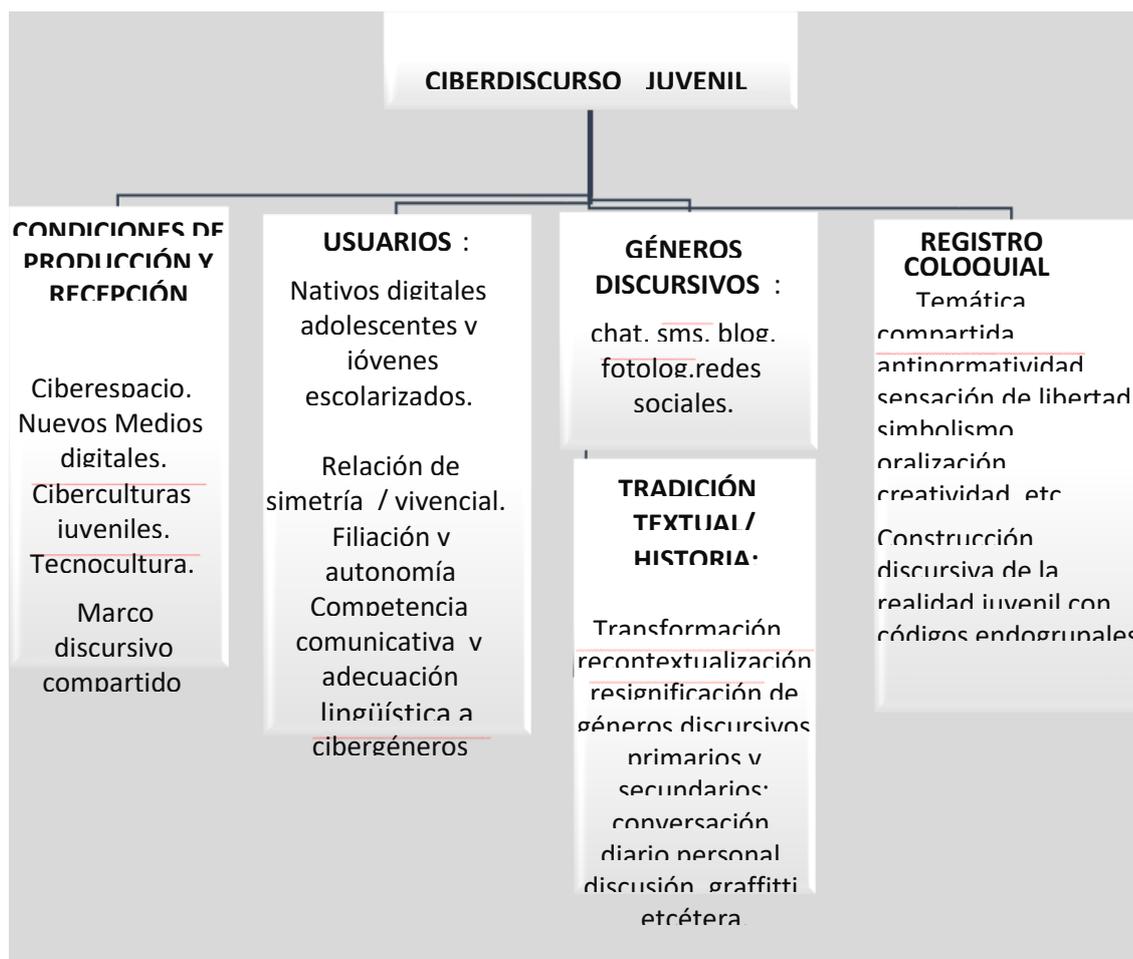
Una perspectiva complementaria es la que ofrecen quienes apelan al concepto de “ciberdiscurso”. Mayans i Planells (2001) quien, siguiendo a de Certau (1988) propone a internet como un espacio practicado, sin materialidad física, pero socialmente significativo para sus usuarios. Palazzo (2010) retoma estos conceptos y describe los rasgos principales de lo que ella denomina “ciberdiscurso juvenil”. Cuando los jóvenes acceden al uso de los llamados “nuevos medios” acceden a “formas de construcción sociocultural y de afirmación de sus identidades” (2010b, 4).

El ciberdiscurso pone el foco en las formas de interacción entre los usuarios en un entorno que es socialmente significativo, como es la web.

El ciberdiscurso juvenil

Estas notas hacen que los discursos de jóvenes y adolescentes montados en el ciberespacio sean auténticas “formas de acción social, determinado por el registro coloquial escrito como producto de las variedades de la lengua.” (Palazzo, 2010b, 9).

Predomina en estos discursos el objetivo de una comunicación efectiva, que desborda los límites del código escrito y fuerza a sus usuarios a buscar otras formas de expresividad. Las decisiones que toman los usuarios para aumentar la efectividad son adecuadas a los medios que emplean y forman parte de la competencia comunicativa en sus distintos aspectos. La autora elabora, a partir de estas consideraciones, un esquema que permite visualizar los rasgos principales a tener en consideración al momento de estudiar estos discursos. Estos son, a saber:



CUADRO X: RASGOS DEL CIBERDISCURSO JUVENIL. FUENTE: PALAZZO, 2010B, 11

De estos rasgos, la autora profundiza el registro coloquial, ya que considera que es allí donde se materializan las marcas propias de las construcciones discursivas de los usuarios.

Particularmente, como ya mencionamos, estos discursos permiten vehicular de una manera especialmente eficaz las identidades, en general, y las juveniles en particular.

Vlog, flog y blog (y sus versiones micro)

Estos términos hacen referencia a distintas formas o especies o géneros discursivos dentro de la web.

Un blog es un diario personal montado en la web y los términos derivados de él se refieren a espacios donde el contenido subido no es sólo texto: flog, vlog, microblog, refieren a espacios donde el contenido predominante son fotos, videos o micro-textos, respectivamente.

Los términos están asociados al origen de estas palabras, pero con el tiempo fueron diferenciándose los espacios y algunos ecosistemas, como los de los flogs, han perdido vigencia en la actualidad.

A partir de esta distinción, desarrollamos un cuadro comparativo, originalmente elaborado por Palazzo de los blogs y de los fotologs al que oponemos nuestra propia caracterización de los videoblogs para poder contrastar y

observar las regularidades y diferencias que aparecen según el medio vinculado con el concepto de “ciberdiscurso juvenil”

Blog	Flog	Vlog
Representación discursiva del “nosotros generacional”	Representación discursiva del “nosotros generacional	Representación discursiva del “nosotros generacional
Usuario joven escolarizado. Visión juvenil de temáticas variadas	Perfil de usuario adolescente. Temática juvenil.	Usuario joven, escolarizado o adolescente. Visiones juveniles de temas.
Construcción discursiva de la imagen personal. Mayor reflexión ideológica y estética	Construcción discursiva de la imagen juvenil. Estética visual	Imagen juvenil al lado que imagen personal. Con distintos grados de reflexión pero alto nivel de estética visual.
Reglas regulativas del género: texto/imagen, comentarios, historial.	Reglas regulativas: registración, <i>posteo</i> (foto, comentario, firma)	Elementos del canal, video y descripción, título y comentarios.
Reglas constitutivas: estrategias de filiación y autonomía, antinormatividad	Reglas constitutivas: estrategias corteses de filiación y autonomía, registro coloquial, antinormatividad	Constitutivas: filiación, pertenencia a comunidad (de youtubers y de líder de sus seguidores). Antinormatividad (encauzada por reglas de YouTube)
Registro coloquial escrito. Creatividad y juegos lingüísticos.	RCE. Juegos tipográficos.	Registro coloquial oral y escrito. Oralidad escrita en ocasiones (como juegos teatrales) Juegos lingüísticos.
Mayor presencia de registro oralizado y		Jerga juvenil tanto en video como en comentarios.

jerga juvenil en los comentarios.		Atenuación de variedades regionales para mayor eficacia de alcance.
(+) contenidos concretos (+) interés intelectual (-) personas	(+) contenidos concretos (-) interés intelectual (+) personas	(+) contenidos concretos (+) interés intelectual (+) personas

TABLA 2: FUENTE: PALAZZO (2010B) Y ELABORACIÓN PERSONAL EN LA TERCERA COLUMNA.

En la tabla se observa como la representación discursiva del nosotros generacional es una constante que atraviesa a todos los géneros. Esto se evidencia en la selección de las temáticas y en el empleo de la primera persona del plural al referirse a los miembros de esta comunidad de la que se siente parte. El usuario es un sujeto joven, escolarizado, también puede ser adolescente, pero de clase social media, con acceso a tecnologías que le permiten crear productos audiovisuales de cierta calidad. En este sentido pensamos que los videoblogs son algo más excluyentes que los otros dos porque requieren mayor inversión económica para lograr productos de calidad. Los temas que tratan, no obstante, son siempre relacionados a su momento y a su edad social. Además, hay videos que se inscriben dentro de prácticas habituales realizadas por adolescentes dentro de la comunidad de youtubers: el "Draw My Life" o el "Condon Challenge" que fueron formatos difundidos y viralizados en internet y practicados por muchísimos usuarios.

Existe, sin dudas, una construcción discursiva de la imagen personal, esta se da no solo en el lenguaje verbal: existe un altísimo nivel de cuidado con la estética visual del sujeto.

Las reglas del género tienen que ver con los requisitos técnicos del sitio: el video es el centro, pero están acompañados de un título y una descripción. Además, los comentarios son espacios propicios para la evaluación de alguna medida de la recepción de estos videos.

En relación con la constitución de comunidades, vemos que existen usuarios que nombran a sus seguidores utilizando alguna forma particular. En Argentina, por ejemplo, Julián Serrano nombra a sus seguidoras como "Serranistas".

Vlog como género discursivo

En tanto red social, YouTube es un espacio que alberga usuarios quienes comparten, al menos, el conocimiento de la red y el conocimiento del manejo de dicha red, es decir, el saber qué y el saber cómo de la página. Además, en tanto medio tecnológico, recordamos lo propuesto por Santos et al. (2007: 8) donde afirma que YouTube es "una red tecnológica cuya topología y conexiones están fuertemente influenciadas por el comportamiento social humano." Por otro lado, es un espacio virtual donde los diversos usuarios se reconocen como parte de una comunidad por los lazos que pueden forjar allí. Podemos pensar entonces que es un tipo de comunidad virtual. Es sabido que en este tipo de comunidades Siguiendo a Bergman y Luckmann (1995), la principal necesidad comunicativa de esta comunidad es la expresión de la identidad. Los autores hablan, además, de géneros reconstructivos, que son géneros comunicativos que sirven para reconstruir eventos pasados y que proveen soluciones para el registro material de dichos eventos. Prefieren hablar de géneros reconstructivos en lugar de

narración, historias o relatos porque consideran que estas son formas demasiado restrictivas y no contemplan aspectos como las múltiples posibilidades de trama. Los géneros reconstructivos pueden ser argumentativos y no (sólo) narrativos. (Bergman Luckmann 1995, 292). Sobre la necesidad específica para el desarrollo de estas reconstrucciones Griffith y Papacharizi detallan de una manera que consideramos suficientemente clara y precisa:

It is not easy to express oneself in any medium. Vlogging incorporates multiple platforms, making the medium more complex. But the complexity of a given set of tools has never stopped anyone from using them for self-expression. This study investigated vlogging as a new medium for self-presentation. Vlogs support self-presentation in multiple ways.

Presentation of self occurs in interaction — individuals present themselves in specific ways based on the impressions they want to give (Goffman, 1959). Some vloggers present themselves as if their vlog were a personal diary, using the site to express or reflect on oneself, store memories, or experiment with video technology. However, they are not intimate because these videos are highly controlled presentations.

Vlogs enable self-presentation by facilitating an ongoing negotiation of identities, not necessarily in a diary-oriented way. Other vloggers present themselves, whether intentionally or not, in narcissistic ways through

unwarranted self-promotion or self-absorption. Certainly this primary focus on the self gives rise to a narcissistic tone. Many of the vlogs in this study exhibit all of these themes throughout their sites.

Through various means of self-presentation, many vloggers share their lives, making personal video blogs fascinating gateways into individuals. Vloggers create and negotiate a self-online as well as a shape a new and more rich kind of cyberspace.

Las propuestas analizadas nos ofrecen un marco sistemático y coherente para la consideración del videoblog como un género discursivo. A la propuesta de Heinemann y Viehweger (1991) es necesario presentar con una cautela mayor algunos rasgos y realizar el análisis de un corpus más extensivo de videos. A su propuesta, consideramos pertinente agregar, además, una revisión de los fenómenos pragmáticos que, consideramos, son de vital importancia para el estudio de la comunicación en los nuevos medios. Un enfoque pragmático permitiría, además de considerar necesidades y objetivos de la comunicación, evaluar estructuras que normalizan de manera especial ciertos hábitos o costumbres propias de la comunidad. A modo de ejemplo mencionamos el análisis de actos de habla realizados en los videos: no es raro encontrarnos con “pedidos de disculpa” al principio de algunos videos. Estos pedidos tienen la finalidad de generar lazos más estrechos con los interlocutores, al mismo tiempo que contribuyen a crear la apariencia de transparencia en la construcción de identidades que se realizan en estos videos.

Por su parte, el análisis de la comunidad discursiva es un paso previo, necesario para poder considerar la existencia de géneros surgidos de ciertas necesidades particulares. Para ello, reconocemos necesario abarcar comunidades más homogéneas en una primera instancia, para poder realizar una aproximación descriptiva y luego una comparación entre diferentes comunidades (considerando, por ejemplo, la variable idioma como un

componente distintivo de cada comunidad), para comprender la universalidad de ciertos fenómenos que contribuirían al éxito comunicativo de estas producciones.

Finalmente, es necesario recalcar la necesidad de un estudio extensivo, cuantitativo, sobre la presencia de los rasgos aquí presentados en un corpus numeroso de producciones que confirmen o refuten nuestras hipótesis.

Referencias

- Androustopoulos, J. K., & Georgakopoulou, A. (2003). *Discourse Constructions of Youth Identities*. Amsterdam: John Benjamins B. V.
- Bergman, J. R., & Luckmann, T. (1995). Reconstructive Genres. En U. Quasthoff, *Aspects of Oral Communication* (págs. 289-304). Berlin: W. de Gruyter.
- Cantamutto, L., & Vela Delfa, C. (2016). El discurso digital como objeto de estudio: de la descripción de interfaces a la definición de propiedades. *Aposta. Revista de Ciencias Sociales*(69), 296-323. Recuperado el 12 de diciembre de 2019, de <http://apostadigital.com/revistav3/hemeroteca/cvela2.pdf>
- De Certau, M. (1988). *The Practice of Everyday life*. Berkeley: University of California Press.
- Fairclough, N. (2008). El análisis crítico del discurso y la mercantilización del discurso público: las universidades. *Discurso y Sociedad*, 2(1), 170-185.
- Griffith, M., & Papacharissi, Z. (2010). Looking for you: An analysis of video blogs. *First Monday*, 15(1). doi:<http://dx.doi.org/10.5210/fm.v15i1.2769>
- Heinemann, W., & Viehweger, D. (1991). *Textlinguistik. Eine Einführung*. Tübingen: Niemeyer.
- Herring, S. (2008). Language and the Internet. En W. Donsbach, *International Encyclopedia of Communication*. Blackwell Publishers.
- Mayans i Planells, J. (2002). *Género chat o cómo la etnografía puso un pie en el ciberespacio*. Barcelona: Gedisa.
- Palazzo, M. G. (2010). *La juventud en el discurso: Representaciones sociales, prensa y chat*. San Miguel de Tucumán: Facultad de Filosofía y Letras - Universidad Nacional de Tucumán.
- Santos, R. L., Rocha, B. P., Rezende, C. G., & Loureiro, A. A. (2007). Characterizing the youtube video-sharing community. *Universidade Federal do Minas Gerais (UFMG) Belo Horizonte Brazil Tech. Rep.*
- Thurlow, C., Lengel, L., & Tomic, A. (2004). *Computer Mediated Communication: Social Interaction and the Internet*. London: Sage.

“Usted está acá”

Proyecto de investigación y realización intermedial con soporte en software para dispositivos móviles

Dr. Esteban Prado; Centro de Letras Hispánicas, Facultad de Humanidades, Universidad Nacional de Mar del Plata.

Téc. Mauro Carusso; Facultad de Arquitectura, Urbanismo y Diseño, Universidad Nacional de Mar del Plata.

Ing. Pablo Spennato; Facultad de Ingeniería, Universidad Nacional de Mar del Plata.

INTERMEDIALIDAD; CIUDAD; SOFTWARE; DISPOSITIVOS MÓVILES; ARTE y LITERATURA

El presente trabajo es la primera comunicación que surge de "Cultura, Arte y Comunicación Transmedia Geolocalizadas", un proyecto de carácter interdisciplinario que reúne a investigadores de la Facultad de Humanidades, de Arquitectura, Urbanismo y Diseño y de Ingeniería de la Universidad Nacional de Mar del Plata. Se trata de un tipo de proyecto poco frecuente en nuestra casa de estudios dado que se otorgaron por vez primera en 2019 subsidios a la investigación con estas características.

Nuestro proyecto tiene dos perfiles: por un lado, busca generar un espacio de estudios enfocados en la exploración de los emergentes artístico-culturales del siglo XXI en Argentina. Hemos comenzado haciendo un recorte en función de las nociones de "transmedialidad" (Jenkins, 2008), "intermedialidad" (Rajewsky, 2005) y "geolocalización". Por otro lado, entendemos que en el paso de lo mecánico y analógico a lo digital, del *tejido sensitivo* al *conectivo* (Berardi, 2017), se ha posibilitado una reorganización de la cultura y el arte que fue guiada por la lógica de las corporaciones que generan los dispositivos y las plataformas de interacción con los "objetos digitales" (Hui, 2017). Ante esto nos preguntamos qué posibilidades han quedado vedadas por esa lógica y desde nuestro espacio como investigadores de una institución pública nos proponemos explorar un camino particular a partir de una intervención tecno-cultural en nuestra ciudad, Mar del Plata.

Problema

Nuestro objeto se constituye como una intervención en el tejido tecnocultural de nuestra ciudad. Vamos a enfocar la investigación en el desarrollo de una experiencia intermedial basada en una aplicación para dispositivos móviles desarrollada *ad hoc*. Esta experiencia intermedial estará centrada en la interrelación entre "objetos digitales" y espacios urbanos concretos. El problema de investigación que nos hemos planteado se puede enunciar como una pregunta: ¿cuáles son los modos de apropiación posibles de las herramientas tecnológicas para intervenir el espacio urbano y generar experiencias artísticas e instancias de difusión cultural? Como toda pregunta densa, no tiene una respuesta completa. Nuestra investigación trabajará uno de esos modos posibles y buscará demostrar que existe un campo mucho más amplio de posibilidades.

Asimismo, la investigación encuentra uno de sus desafíos centrales en la articulación del deseo universalizante de la tecnología y la ingeniería, los modos de ejecución de proyectos del diseño y la comunicación y la ciencia de casos particulares de la crítica literaria y cultural. Es decir, más allá de que no sea un problema de investigación en

sentido estricto, la articulación de lógicas disciplinares distintas no deja de ser un centro gravitacional oculto de todo nuestro trabajo. A esto se suma una cuarta dimensión vinculada ya no con la crítica cultural, el diseño o la ingeniería como disciplinas académicas sino con el estatuto de la investigación artística y su lugar en la construcción de conocimiento y de intervenciones críticas que nos proponemos con este proyecto. Un porcentaje considerable de los integrantes del proyecto se definen como investigadores académicos y, al mismo tiempo, como editores, curadores, escritores, realizadores audiovisuales y artistas transdisciplinares. Uno de los problemas que nos interesa abordar es el de los diálogos entre “ciencia” y “arte” y, más acá o más allá de estos diálogos, el del desarme de los límites que las mantienen separadas.

El problema que enfocamos con el desarrollo de nuestra experiencia intermedial, más allá de la exploración de posibilidades, consiste en un intento de interrelacionar los resultados obtenidos en investigaciones vinculadas a los estudios del mercado editorial argentino con la noción de *cosmotécnica* de Yuk Hui. Por un lado, retomamos el modelo de “negocios” y/o de “subsistencia” de las pequeñas editoriales argentinas surgidas luego del proceso de asimilación de la diversidad del mercado editorial por parte de los grandes grupos editoriales, en la década de los ‘90 (Botto, 2006; De Diego, 219). Estas pequeñas editoriales se caracterizaron por trabajar en la brecha que estos grupos descuidaban, la de los libros que van desde la producción artesanal hasta la industrial de escala acotada. Destacamos que este proceso garantizó una interesante bibliodiversidad en la producción editorial argentina durante los últimos veinte años. Entonces, una pregunta que se nos plantea como problema es si habría un modo de intervención en el ámbito artístico, técnico y cultural que pudiera analogarse con este modo de existencia de las editoriales *interdependientes* (Moscardi, 2017). Es una pregunta abierta que oficiará de norte para establecer algunos de los principios que guíen nuestra intervención. Desde otro ámbito nos interpela la propuesta de Yuk Hui en torno a la “cosmotécnica”, entendida esta como una noción que se contrapone al deseo totalizante de la tecnología y que entiende que pueden existir modos de apropiación particulares en los que naturaleza, cultura y técnica se interrelacionan tienen alcances ontológicos.⁶⁷ Con estas dos cuestiones presentes, nuestro proyecto busca intervenir en un espacio que las corporaciones descuidan, con una serie de condicionamientos y propuestas que difícilmente sean “monetizables” y/o “replicables” a gran escala. Asimismo, hipotetizamos que atender y desarrollar un proyecto en ese ámbito podría dar lugar a una “vernacularización” de los modos de uso de las herramientas disponibles para construir experiencias a través de aplicaciones para dispositivos móviles. Esta vernacularización -de la lengua franca, de la “tecnología universal”- podría significar un modo de apropiación, los prolegómenos de una cosmotécnica propia.

Ahora bien, si como afirman los estudios mediales, especialmente Müller (2012), un medio se constituye entre un dispositivo tecnológico y una configuración sociocultural determinada, podríamos plantear que para el caso de los dispositivos móviles y de la informática en general se trataría de un complejo conformado por 1. hardware, 2. software -aplicación (código; UX, UI + objetos digitales) 3. configuración sociocultural. De manera que lo que desarrollaremos será, en el sentido de las artes mediales, un medio de comunicación con una configuración sociocultural abierta y en la que se suspenderá la búsqueda compulsiva de “masividad”. Esta concepción la tomamos de Arlindo Machado quien señala que el arte de los medios constituye un ‘metalenguaje’, “un espacio de reflexión sobre las nuevas tecnologías capaz de alejarse de los sentidos atribuidos a ellas por la lógica del trabajo y el entretenimiento. ‘Desprogramar’ sería una consigna, reapropiarse de los dispositivos, reinventarles nuevos usos y contextos”. (Kozak, 2015: 193)

⁶⁷ La discusión de Hui se enmarca en una revisión del giro ontológico propio de la antropología (Ver: Hoolbrad y Perdersen, 2017).

Metodología

Nuestro planteo sostiene que el pensamiento crítico, la literatura, el arte y la tecnología construyen caminos “poco probables en el programa” (Flusser, 2015) cuando trabajan de modo conjunto: la apertura y el reconocimiento de posibilidades vedadas surgen de configuraciones interdisciplinarias poco transitadas. Consideramos que la investigación crítico-creativa en la que los investigadores se involucran en los procesos constructivos permite una perspectiva distinta tanto para el pensamiento crítico y la construcción de conocimiento como para la apropiación del *know how*.

En este sentido, trabajamos el estudio de casos concretos en correlación con el desarrollo de nuestra intervención, de manera que una y otra perspectiva se retroalimentan. De los posicionamientos teóricos de Josefina Ludmer rescatamos el planteo de que desde Argentina y Latinoamérica tenemos una posición descentrada y que nuestros proyectos teórico-críticos siempre adquieren un emplazamiento territorial y contingente. Desde un punto de vista más acotado, nos concebimos como “mediadores culturales” en la medida en que buscamos convertirnos “en un actor central como articulador entre el ámbito de la producción cultural y el de la producción de conocimiento” (Del Valle, 2018: 12). Esto implica que no buscamos la “innovación” en las distintas disciplinas involucradas; nuestro objetivo es el desarrollo de propuestas vinculadas a nuestro ámbito de inserción y la apertura de un espacio de apropiación, reflexión y usos críticos de la tecnología.

Hemos planteado un corpus amplio del que podemos destacar tres casos concretos que dan cuenta del estado de la cuestión de este tipo de experiencias en la ciudad: “Silencio. Dispositivo Sonoro”, “Nueva Cartas para Diana”, “Las campanas” de Virgili.⁶⁸

Como dijimos, hemos definido como objetivo el desarrollo de experiencias intermediales que combinan “objetos digitales” con espacios urbanos concretos. A partir de esta decisión inicial, se conformó un equipo con un miembro de cada unidad académica involucrada. En sucesivas reuniones de discusión se evaluaron posibilidades técnicas en relación con la exploración crítico-creativa que buscamos, llegando a una primera conclusión que define el carácter intermedial de nuestra intervención. La misma buscará que el “usuario” experimente la relación de un “objeto digital” con un espacio urbano de Mar del Plata -arquitectura, paisaje, tránsito- y, simultáneamente, con un territorio -suelo, atmósfera, condiciones climáticas-. Para lograr esta multiplicidad es que definimos nuestros “objetos digitales” como eminentemente sonoro-verbales, para que el resto de los sentidos y el cuerpo puedan vincularse libremente con el emplazamiento. Este sería uno de los desafíos estructurantes de la intervención. Como señala Van Dijk (2016), los usos de los dispositivos tecnológicos tienden hacia una cada vez mayor inmersión “online”, en la que la cultura de la conectividad es una captura de la atención. En nuestro caso, la intermedialidad que buscamos cancelaría esa ecuación, ya no se podría pensar el tiempo de uso de la APP como un momento de inmersión en el dispositivo sino como una experiencia heterogénea, no totalmente inmersiva. En términos metodológicos, iremos haciendo un registro del proceso constructivo y de su impacto en el entramado cultural de la ciudad, generando reflexiones críticas sobre los mismos.

⁶⁸ “Silencio. Dispositivo Sonoro” (2017-) es una vidriera a la calle barrial con *plugins* para auriculares, lo que implica una intervención del espacio público con obras que están disponibles las 24 hs. para ver y escuchar. “Nueve Cartas para Diana” (2020) fue una pieza que combinó monólogos teatrales con video digital y fue transmitida por redes sociales. Su particularidad radicó en que cada “carta” era proyectada en un emplazamiento distinto, de manera que la transmisión por celulares recuperaba la contingencia de los espacios de proyección. Por último, “Campana-Acantilado” (2020) de Martín Virgili consiste en un documental sonoro sobre la zona sur de Mar del Plata. Se distribuyó a través de la plataforma Soundcloud. Rescatamos el carácter “presencial” y el deseo de estar ahí que generan las voces y su paisaje sonoro.

En cuanto al marco desde el que pensamos la ciudad y nuestro usuario, en primer lugar, nos posicionamos desde el *derecho a la ciudad* (Lefebvre, 1967). Concebimos al usuario como un sujeto que al tiempo que accede a los bienes culturales disponibilizados ejerce este derecho. En ese sentido, adherimos al manifiesto de Raymond Williams, “Culture is ordinary” (1958), que sin dudas ha constituido las bases de las propuestas que vinculan la mediación cultural como un modo de garantizar un derecho ciudadano (Del Valle, 2018). En cuanto a los movimientos del usuario pensamos en una experiencia del caminar que alineamos con la propuesta de Francesco Careri

(2002) en torno al *andar como práctica estética*.

A partir de estas primeras definiciones, el equipo de ingenieros comenzó el desarrollo de la aplicación para dispositivos móviles y de aquí surgieron nuevas cuestiones que fueron dando otra cariz a la experiencia. La APP está siendo desarrollada a partir de Toolkit Flutter y se utilizará Firebase para el almacenamiento de los “objetos digitales”. Para establecer la relación entre estos y la geolocalización del usuario se utilizará otro servicio de Firebase llamado Firestore, una base de datos que, en tiempo real, permite establecer vínculos entre geolocalización y “eventos” que suceden en la APP para disponibilizar los contenidos digitales. Más precisamente, se trabaja con “geovalas” que definen áreas en las que se habilitan funciones y contenidos en la APP según la ubicación territorial del dispositivo móvil. Fue en el diálogo con los desarrolladores que desde el diseño se adoptó una configuración vinculada con la noción de *futuridad* (Gatto, 2018), lo que implica trabajar ya no desde una idea de futuro cerrado, sino recuperando la plasticidad del desarrollo histórico de todo proyecto. Así fue que definimos una idea de “Estación” como el anclaje en el que la aplicación disponibilizará un objeto digital en un espacio/territorio determinado.

Desde las primeras ideas al trabajo concreto, sobre todo en el diseño de la experiencia de usuario y el relevamiento de producciones culturales susceptibles de ser convertidas en nuestro repertorio de “objetos digitales”, surgió la evidencia de que la ciudad tiene una amplísima cantidad de textos literarios que fueron escritos aquí, suceden aquí o la tematizan⁶⁹ por lo que pasamos de concebir sólo recorridos que unieran distintas estaciones a pensar en dos opciones de propuestas por parte de la APP: por un lado, una muestra permanente que funcione como archivo de las producciones que se vinculan con Mar del Plata, curada por miembros del equipo y con una lógica aditiva y heterogénea en cuanto al conjunto de objetos culturales y aleatoria en cuanto a los modos de transitarla; por otro lado, muestras eventuales con producciones construidas *ad hoc* para la APP y con recorridos secuenciales. En esta segunda línea, cada “evento” definirá una dinámica propia.

Es en estas dos líneas que imaginamos la posibilidad de experimentar con la herramienta, evaluar sus límites y construir formas y modos posibles de articulación. De modo que la tomaremos como plataforma o laboratorio de exploración de arte y tecnología. Asimismo, el proyecto genera dos campos de interés susceptibles de ser estudiados en sí mismos: por un lado, el archivo de producciones vinculadas a la ciudad de la muestra permanente y de las muestras temporales; por otro lado, una diversidad de datos registrados por la APP respecto del comportamiento de los usuarios.

⁶⁹ Nuestra ciudad tiene varios perfiles: por un lado, la ciudad de la *Belle Époque*, construida como un espacio de esparcimiento para las élites a comienzos de siglo XX. Por otro, una ciudad de veraneo de las clases populares a partir de mediados de siglo. Finalmente, una ciudad con medio millón de habitantes que en verano llegan a multiplicarse por cuatro. Esto ha generado un amplísimo y heterogéneo caudal de producciones culturales de ilustres visitantes y existe una profusa producción de artistas y escritores locales que ha proliferado con la llegada del siglo XXI.

Primeras conclusiones: no una restauración del aura, sí una apertura a la experiencia

Las conclusiones que hemos obtenido al momento, antes de haber cumplido un año de desarrollo del proyecto, nos permiten afirmar que la interrelación disciplinar y el desarme de los órdenes que separan investigadores de artistas ha sido sumamente productivo para abrir un campo de exploración en el ámbito de los estudios y las artes mediales de la ciudad de Mar del Plata, Argentina, y para comenzar a trabajar en una intervención concreta de carácter socio-tecno-cultural. Asimismo, como resultado inesperado ha aparecido la posibilidad de construir un archivo de literatura y otras artes vinculado a la ciudad que no imaginábamos al inicio del proyecto. A su vez, esto ha modificado sustancialmente el modo de trabajo que habíamos propuesto dado que en un primer momento los objetos digitales serían construidos completamente *ad hoc*. Ahora, el descubrimiento de la posibilidad del archivo nos ha llevado a plantear las dos propuestas que caracterizan nuestra experiencia: la muestra permanente con el archivo de la ciudad y las muestras eventuales. Asimismo, nos hemos visto en la necesidad de adoptar la noción de *diseño postutópico* (Gatto, 2018) para abrir nuestro modo de relacionarnos con el proyecto y posibilitar una plasticidad abierta a lo que surja en el proceso constructivo. Es en este sentido que la investigación científico-creativa que estamos llevando a cabo permite algunas conclusiones vinculadas al problema que nos enmarca: este proceso de trabajo, en su desarrollo temporal-histórico, significa la apropiación tecnológica que nos proponemos y comienza a evidenciar algunas cuestiones vinculadas a las preguntas fundantes de los ensayos sobre *cosmotécnica* de Yuk Hui (2020). Estas respuestas no las encontramos -como sugiere este filósofo- en el marco de las culturas precolombinas que nos anteceden, sino que las encontramos en un modo particular y criollo del “hágalo usted mismo”, en el que una intensa capacidad de aprendizaje, un permanente diálogo y un modo de atar *las cosas con alambre* nos han traído hasta acá.

Bibliografía

- BERARDI, F., *Fenomenología del Fin*, Buenos Aires, Caja Negra, 2017.
- BOTTO, M., “La concentración y la polarización de la industria editorial”, *Editores y políticas editoriales en Argentina, 1880-2000*, José Luis de Diego (Dir.), Buenos Aires, Fondo de Cultura Económica, 2006, 209-248.
- CARERI, F. *Walkscapes. El andar como práctica estética*, Barcelona: GG, 2002.
- DE DIEGO, J. L., *Los autores no escriben libros*, Buenos Aires, Ampersand, 2019.
- DEL VALLE, D., “La mediación cultural y la cultura como derecho ciudadano”, ABOUDRAR, B. N. y MAIRESSE, F., *La mediación cultural*, Buenos Aires, Libros UNA, 2018, 7-13.
- FLUSSER, V. *El universo de las imágenes técnicas*, Buenos Aires, Caja Negra, 2015.
- GATTO, E. *Futuridades. Ensayos sobre política posutópica*, Rosario, Casagrade, 2018. HOLBRAAD, M., y PEDERSEN, M. A., *The ontological turn: An Anthropological Exposition*, Cambridge, Cambridge University Press, 2017.
- HUI, Y. “¿Qué es un objeto digital?”, *Virtualis*, Vol. 7, núm. 15, enero/junio 2017, 81-96, HUI, Y. *Fragmentar el futuro. Ensayos sobre tecnodiversidad*, Buenos Aires, Caja Negra, 2020.

- JENKINS, H. *Convergence culture: la cultura de la convergencia de los medios de comunicación*. Barcelona: Paidós, 2008.
- KOZAK, C. (Ed.) *Tecnopoéticas argentinas. Archivo blando de arte y tecnología*. Buenos Aires, Caja Negra, 2015.
- LEFEBVRE, H. *El derecho a la ciudad*, Madrid: Capitán Swing, 2017.
- LUDMER, J. *Clases 1986. Algunos problemas de teoría literaria*, Buenos Aires, Paidós, 2015.
- MACHADO, A. *Arte e mídia*, Rio de Janeiro, Zahar, 2007.
- MOSCARDI, M., *La máquina de hacer libritos*, Mar del Plata, Puente Aéreo, 2017.
- MÜLLER, J., "Intermedialidade revisitada: algumas reflexões sobre os princípios básicos desse conceito", en NOGUEIRA F., y VIEIRA, A.S. (org.), *Intermedialidade e estudos interartes: desafios da arte contemporânea*. V. 2, Belo Horizonte, Rona Editora, 2012, pp. 75-98.
- PRADO, E. "Intermedialidad. Una reflexión teórica sobre su potencia crítica", en *ASRI - Arte y Sociedad. Revista de Investigación*, 15, 2018, 65-79. Recuperado de <http://asri.eumed.net/15/intermedialidad-reflexiones-teoricas.pdf> [22/8/19].
- RAJEWSKY, I., "Intermediality, Intertextuality, and Remediation: A Literary Perspective on Intermediality", *Intermedialités*, 6, 2005, 43-64. Recuperado de http://cri.histart.umontreal.ca/cri/fr/intermedialites/p6/pdfs/p6_rajewsky_text.pdf [22/8/19].
- VAN DIJKC, J., *La cultura de la conectividad*, Buenos Aires, Siglo XXI, 2016.
- WILLIAMS, R., "Culture Is Ordinary", *Resources of Hope: Culture, Democracy, Socialism*, Londres, Verso, 1989, 3-14.

Sesión 9

La vía de la civilización

Educación, enseñanza y participación ciudadana

Modera: Nuria Rodríguez Ortega · *Asiste:* Brianda Otero Moreira

- 42 Ciencia ciudadana e Historia: procesos participativos y compromiso público
Lidia Bocanegra Barbecho
- 73 Patrimonio Herido: un trabajo-proyecto transdisciplinar para la innovación educativa
María Marcos Cobaleda, Antonio Jesús Santana Guzmán, Julia García González, Antonio Cruces Rodríguez, Nuria Rodríguez Ortega
- 65 La Edición Digital de Fernando Pessoa: consideraciones sobre las ediciones digitales como archivos abiertos
Ulrike Henny-Krahmer, Jorge Uribe, Pedro Sepúlveda, Inês Rebelo Do Carmo
- 25 El área de Ciencias y Técnicas Historiográficas como ejemplo de implantación de las Humanidades Digitales en el progreso de la docencia
Lorena Barco, Alicia Marchant

Ciencia ciudadana e Historia: procesos participativos y compromiso público

Lidia Bocanegra Barbecho, Universidad de Granada

Palabras clave: Ciencia ciudadana, Co-creación, Humanidades Digitales, Compromiso público, Investigación historia.

Descripción del problema de investigación

La presente ponencia pretende arrojar luz a aquellos procesos participativos, así como el compromiso público, llevados a cabo en la investigación histórica y analizados a través del proyecto Co-Historia: *análisis de la Participación Pública en la investigación histórica desde el ámbito de la ciencia ciudadana* (Ref. E-HUM-507-UGR18)⁷⁰.

La iniciativa Co-Historia se basa en detectar, analizar, testear y estandarizar aquellas prácticas metodológicas de participación pública en proyectos de investigación desde la disciplina de la historia, nacionales e internacionales, y ejecutados desde la esfera universitaria, o por científicos profesionales. El objetivo final es entender de qué manera se involucra a la ciudadanía, incluyendo partes interesadas, en la propia investigación en historia (compromiso); así como saber qué enfoques participativos se están llevando a cabo. Co-historia pretende responder cuestiones, tales como: ¿se utiliza más la participación a nivel co-creativo o participativo? ¿Prima más la participación a nivel digital o *face-to face* mediante los laboratorios ciudadanos? También pretende demostrar que, a diferencia de lo que se cree, cada vez más se ejecutan proyectos que implementan esta estrategia participativa durante el desarrollo de la investigación, propias de otras disciplinas y que, además, se implementa de manera improvisada, sin sacar el máximo partido.

Otro de los objetivos generales del proyecto es la de fomentar la unión, y el uso, de la ciencia histórica con la inteligencia colectiva, incentivando el compromiso público en la investigación histórica y situando, esta disciplina, en la esfera de la innovación social. Dicho de otro modo, Co-Historia incentivará la apertura de la investigación histórica al público para que entre todos la entretengan, la co-creen, con la coordinación de los profesionales científicos en esta materia y cuyos resultados, en parte, tengan un beneficio social.

Casar ciencia ciudadana e investigación histórica sitúan esta iniciativa en la frontera del conocimiento, a nivel metodológico acerca de la manera de investigar y comprometerse con la propia historia. Dicho de otro modo, antes había unas formas tradicionales de investigar esta disciplina, analizada por la historiografía y basada fuertemente en un tipo de investigación básica; mientras que ahora el movimiento de ciencia ciudadana abre un espacio nuevo de actuación favoreciendo, esta vez, una investigación multidisciplinar, e interdisciplinar, y con unos resultados cada vez más prácticos. Así pues, los resultados de la presente investigación supondrán un avance claramente significativo a nivel científico-metodológico, pero especialmente dará un nuevo uso del paradigma de la ciencia ciudadana aplicada en la disciplina histórica, ofreciendo nuevos enfoques de la misma y dando lugar a nuevos planteamientos.

⁷⁰ La presente ponencia se ha realizado en el marco del proyecto de investigación: *Análisis de la Participación Pública en la investigación histórica desde el ámbito de la ciencia ciudadana* (Co-Historia), Ref. E-HUM-507-UGR18, financiada en la convocatoria de Proyectos I+D+i del Programa Operativo FEDER 2018 de la Universidad de Granada, siendo la Dra. Lidia Bocanegra Barbecho la investigadora principal.

INVESTIGACIÓN EN HISTORIA

+

CIENCIA CIUDADANA

=

INVESTIGACIÓN DEMOCRÁTICA E INNOVADORA

¿Qué entendemos por ciencia ciudadana? Se trata de un método que está basado en la inclusión de miembros del público en algún aspecto de la investigación científica (Eitzel, 2017); nos referimos a cualquier forma de participación activa en la ciencia que va más allá de la investigación profesional. Es decir, se trataría de una investigación científica que cuenta con la implicación activa del público, que no tiene porqué ser especializado, junto con científicos y profesionales que son quienes dirigen la iniciativa y que, dependiendo del grado de involucración del público, trabajarán de una forma más horizontal o no. De la misma manera que existen varias definiciones acerca de las Humanidades Digitales, lo mismo sucede con la ciencia ciudadana, aportando, cada una de esas definiciones, una serie de rasgos que caracterizan la misma. Una de ellas define la ciencia ciudadana en relación con la educación cívica, como el trabajo realizado con las comunidades ciudadanas para avanzar en la ciencia, fomentar una mentalidad científica amplia y el compromiso democrático, ayudando a la sociedad a abordar problemas modernos complejos (Ceccaroni, Bowser & Brenton, 2017). Asimismo, algunos autores enfatizan la propia responsabilidad de la ciencia para con la sociedad y que ellos denominan ciencia ciudadana “democrática” (Irwin, 1995). Destacar, también, la importancia de los voluntarios y no profesionales en dicho método, quienes contribuyen participando colectivamente en una amplia gama de proyectos científicos para responder preguntas del mundo real. Tanto las contribuciones de los ciudadanos, como las actitudes de los investigadores con respecto a este enfoque, abarcan un amplio conjunto de actividades a múltiples escalas; de esta manera, encontramos interacciones masivas ocasionales a escala global, de manera virtual, pero también una participación proactiva regular en entornos locales (muchas veces con participación física y no solo virtual) que identifican nuevas preguntas de investigación (VV.AA, 2013). Para una mayor comprensión del enfoque de ciencia ciudadana, aconsejamos leer los diez principios clave que subyacen a las buenas prácticas de este método participativo y que han sido establecidos por la European Citizen Science Association (ECSA) (VV.AA, s.f). Algunos de estos principios ya se han ido nombrando en esta sección (fomentar la mentalidad científica y el compromiso público), pero nos gustaría destacar también aquel que expone que los datos, y metadatos, de los proyectos en ciencia ciudadana deberían ser públicos y, a ser posible, los resultados deberían publicarse en formato de acceso abierto, siempre y cuando no haya problemas de seguridad y/o privacidad que lo impidan. Esto es importante ya que este hecho, junto con el propio compromiso público mencionado anteriormente, son dos aspectos transversales que se enmarcan también en otras disciplinas afines, tales como las Humanidades Digitales y la Historia Pública Digital (Bocanegra Barbecho, 2020).

Metodología

Destacamos dos tipos de metodología implementadas en Co-Historia: una macro, dedicada a la ejecución completa del proyecto; y otra micro, basada en la estructuración de la base de datos para una correcta clasificación.

A nivel macro, el proyecto Co-Historia conlleva un marco metodológico desarrollado en cuatro fases y ejecutados en cinco paquetes de trabajo, con una temporalidad total de tres años (2020-2023). Son las siguientes:

1. Fase 1: *detección de proyectos de investigación en historia que utilizan la participación pública desde la ciencia ciudadana*. En esta fase se identificará y mapeará, a nivel nacional e internacional, aquellos

proyectos de investigación en historia que utilizan la participación pública a través de la ciencia ciudadana en ámbito digital y físico.

2. Fase 2: *metodología de estandarización sobre participación pública en la investigación histórica desde la ciencia ciudadana*. En esta fase se pretende relacionar los proyectos de investigación en historia que utilizan la ciencia ciudadana con la literatura científica, artículos e informes, que traten de procesos participativos en general y de la ciencia ciudadana en sí. En este punto, la literatura científica se evalúa para conectar la definición y los objetivos de las diferentes prácticas de la ciencia ciudadana en historia (co-creación, contribución y colaboración) y el compromiso público. Esta fase es muy importante porque llevará a cabo los primeros análisis y, como resultado, dará lugar a una primera catalogación de los procesos participativos (de involucración ciudadana); así como de las actividades de *engagement* (compromiso público).
3. Fase 3: *Testeo y mejora de la metodología de participación pública en la investigación histórica desde la ciencia ciudadana a través de los casos piloto*. Inicialmente, esta fase conllevaba la realización de una serie de laboratorios ciudadanos (*Living labs*) en donde se analizaba un tema en concreto; debido a la crisis sanitaria actual, se sustituirán los laboratorios ciudadanos por actividades participativas virtuales, como, por ejemplo: la transcripción de textos, etiquetado y enriquecimiento (entrada) de datos, entre otros. En la medida de lo posible, se involucrará al estudiantado de la Universidad de Granada en dichas actividades, con el fin de proyectar esta estrategia más allá de la generación de nuevo conocimiento; si no, también, con la idea fomentar una serie de competencias digitales a los alumnos, a la par que aprenden y se comprometen con determinados temas históricos.
4. Fase 4: *evaluación de la metodología de participación pública y compromiso público*. Tras el testeo previo de determinadas actividades participativas y de compromiso público, identificadas en la fase precedente, se desarrollará unas guías definitivas de buenas prácticas acerca de la participación pública, a nivel digital y físico (laboratorios ciudadanos), y compromiso público en la investigación en historia.

Actualmente estamos finalizando de clasificar los proyectos participativos nutriendo nuestra base de datos; a la par que estamos indagando la literatura científica acerca de este tema.

Resultados preliminares

Un primer resultado ha sido la de definir los parámetros de clasificación de los proyectos identificados; para ello hemos utilizado bibliografía científica sobre el tema; así como hemos hecho uso de nuestra propia experiencia en la ideación y gestión de proyectos participativos y su correspondiente *public engagement*. Destacar que tanto la IP del proyecto Co-Historia, como los miembros del equipo, tienen una amplia experiencia en ejecutar y participar en proyectos que incluyen la participación ciudadana.

Hemos definido la estructura de clasificación de la siguiente manera:

- un apartado descriptivo en donde se explica brevemente el objetivo principal del proyecto, el área a la que pertenece (disciplina), palabras claves que lo describen, institución, audiencia involucrada, idioma y URL.
- un apartado analítico en donde se clasifica los enfoques participativos (Contributivo, Colaborativo, Co-creado, Alojado), las actividades participativas (Enriquecimiento/entrada de datos, Etiquetado/Clasificación, Geolocalización, Recolección de fuentes, Recreación histórica, Transcripción/corrección textos, Transferencia, etc.), las estrategias de acercamiento al público; así como el modelo de estrategias de involucración del público.

La base de datos permite, además de parámetros de clasificación con listas cerradas, campos de texto libre en donde se añade información explicativa acerca de las estrategias participativas y de acercamiento al público.

Una vez ideada la estructura, hemos implementó en una base de datos web a través del cual se ha ido clasificando más de un centenar de iniciativas. Actualmente estamos finalizando dicha catalogación y hemos empezado a ejecutar los primeros análisis, los cuales se mostrarán en el congreso HDH2021 *Scire Vias. Humanidades Digitales y Conocimiento*.

Durante estas primeras fases de implementación, Co-Historia también ha ido identificando bases de datos a través del cual se ha ido nutriendo de proyectos que están en marcha, o bien ya han finalizado.



Proyectos

91 Elementos

1 2 3 4 5 6 7 8 9 » Last »



Project Querença



Preserving the Mayan-Achi Food System



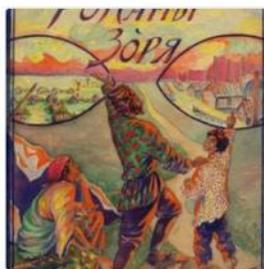
Preservation of the bobbin lace tradition in Balatonendré



Popular Musicscapes



Plugg
Plug into Cultural Her



Detalle de la sección de Proyectos identificados de la web de la iniciativa Co-Historia: <http://www.cohistoria.es/>

Bibliografía

Bocanegra Barbecho, L (2020). Ciencia ciudadana y memoria histórica: nuevas perspectivas historiográficas desde las Humanidades Digitales y la Historia Pública Digital. In *Terra Incognita: Libro blanco sobre transdisciplinariedad y nuevas formas de investigación en el Sistema Español de Ciencia y Tecnología*. Burgos: PressBooks. <http://doi.org/10.5281/zenodo.4046754>

Eitzel, M.V., et al. (2017). Citizen Science Terminology Matters: Exploring Key Terms. *Citizen Science: Theory and Practice*, 2(1): 1, pp. 1–20, <https://doi.org/10.5334/cstp.96>

Ceccaroni, L., Bowser, A. and Brenton, P. (2017). Civic Education and Citizen Science:

Definitions, Categories, Knowledge Representation In: Ceccaroni, L. and Piera, J. eds. *Analyzing the Role of Citizen Science in Modern Research*. Hershey, PA: IGI Global, pp. 1–23, <https://doi.org/10.4018/978-1-5225-0962-2.ch001>

Irwin, A. (1995). *Citizen Science: A study of people, expertise and sustainable development*. London: Routledge.136.

VV.AA. (2013). Green Paper on Citizen Science. Citizen Science for Europe: Towards a better society of empowered citizens and enhanced research, European Commission, p.21,

https://www.researchgate.net/publication/271130919_Green_Paper_on_Citizen_Science_Citizen_Science_for_Europe_Towards_a_better_society_of_empowered_citizens_and_enhanced_research/related

VV.AA. (s.f). Diez principios de ciencia ciudadana. European Citizen Science Association,

https://ecsa.citizenscience.net/sites/default/files/ecsa_ten_principles_of_cs_spanish_0.pdf

Patrimonio Herido: un trabajo-proyecto transdisciplinar para la innovación educativa

María MARCOS COBALEDA, Antonio Jesús SANTANA GUZMÁN, Julia GARCÍA GONZÁLEZ, Antonio CRUCES RODRÍGUEZ, Nuria RODRÍGUEZORTEGA

Universidad de Málaga (UMA)

Palabras clave: Patrimonio, Peligro, Tutela, Protección, Proyecto de Innovación Educativa (PIE)

El objetivo principal de esta contribución es presentar una línea de indagación docente basada en la hibridación entre metodologías de prototipado, estrategias ABP (Aprendizaje Basado en Problemas) e implementaciones tecnológicas, que forma parte de las acciones de innovación educativa llevadas a cabo por TransUMA (Laboratorio de Competencias Transdisciplinares de la Universidad de Málaga). TransUMA es un ecosistema de aprendizaje orientado a facilitar una formación transdisciplinar, entendida como integración de saberes heterogéneos donde se combinan acción, pensamiento divergente y discurso crítico⁷¹. En concreto, esta comunicación se centra en un estudio de caso específico: el proyecto *Patrimonio Herido*, en el que ha participado el alumnado del grado en Historia del Arte y del grado en Turismo de la Universidad de Málaga. El objetivo central de este proyecto ha sido (y es) crear mecanismos a través de los cuales los estudiantes (y la ciudadanía en general) puedan contribuir de manera efectiva y directa al cuidado y preservación del patrimonio cultural próximo. Como se describirá a continuación, los trabajos llevados a cabo por los estudiantes en el marco de la infraestructura tecnológica *Patrimonio Herido*, desarrollada *ad hoc* para este proyecto, les ha permitido adquirir competencias muy diversas, entre las que se incluyen las propiamente tecnológicas.

Teniendo en cuenta lo anterior, durante el curso académico 2020-2021, se ha escogido esta metodología de actividad formativa para la realización de un trabajo-proyecto colaborativo entre el alumnado de varias asignaturas de distintos grados de la Universidad de Málaga. Por un lado, han participado estudiantes de la asignatura de “Gestión del Patrimonio Cultural” del grado en Turismo (primer curso) y, por otro lado, estudiantes de las asignaturas “Arte Islámico, Andaluz y Mudéjar” (primer curso), “Historia y Conceptos Fundamentales del Patrimonio” (segundo curso) y “Arquitectura Contemporánea” (cuarto curso) del grado en Historia del Arte.

Como objetivos específicos para la realización de esta actividad se han planteado los siguientes:

1. Propiciar la transdisciplinariedad entre el grado en Historia del Arte y el grado en Turismo de la Universidad de Málaga.
2. Involucrar al estudiantado en el contacto directo con el patrimonio cercano y fomentar su conservación.
3. Promover acciones que permitan sanar el patrimonio cultural herido.
4. Concienciar de la importancia del turismo cultural como factor de preservación del patrimonio.
5. Familiarizar al alumnado con herramientas digitales para la difusión, protección y conservación del Patrimonio, y proveerles de capacitación técnica a través de talleres de formación específicos.

En virtud de los objetivos anteriores, el trabajo-proyecto del *Patrimonio Herido* se ha estructurado en tres fases. La primera fase ha consistido en la selección por parte del alumnado matriculado en las asignaturas de “Gestión

⁷¹ TransUMA es resultado del proyecto de innovación educativa de la Universidad de Málaga PIE17-123 que continúa en el marco PIE19-178 *Imaginando futuros posibles*. Véase en <http://transuma.uma.es/>

del Patrimonio Cultural”, “Arte Islámico, Andaluz y Mudéjar” y “Arquitectura Contemporánea” de un bien patrimonial (preferiblemente localizado en la provincia de Málaga) que, por alguna razón, se encontrara en riesgo de conservación o que hubiera sufrido algún deterioro o destrucción. Una vez elegido el bien, los estudiantes elaboraron una ficha básica de diagnóstico, en la que se incluyeron los datos para su identificación, su cronología, su ubicación exacta, el tipo de patrimonio en el que se inscribe y una breve descripción histórico-artística. Junto a ello, en la ficha diagnóstico, los estudiantes debían señalar los daños sufridos por el bien seleccionado. Asimismo, se solicitó al alumnado la inclusión de una bibliografía y documentación de referencia para familiarizarse con los diferentes sistemas de citación; una serie de imágenes que permitieran la identificación del bien y la valoración de los daños sufridos; un mapa básico de localización; y una reflexión sobre las posibles medidas que posibilitarían la curación del bien.

Una vez confeccionadas las fichas diagnóstico, el profesorado de las asignaturas procedió a una selección de los bienes más representativos, atendiendo a varios criterios:

1. Calidad de los trabajos.
2. Viabilidad de la realización de un proyecto de conservación sobre el bien.
3. Necesidades sociales en el entorno donde se emplaza el bien y las posibilidades de regeneración que su recuperación ofrece.
4. Inclusión de ejemplos de otros lugares de Andalucía.
5. Grado del daño.
6. Viabilidad de acciones de mejora.
7. Amplitud tipológica.

Así, el número total de fichas diagnóstico seleccionadas fueron sesenta y siete.

La segunda fase del proyecto fue realizada por los estudiantes matriculados en la asignatura “Historia y Conceptos Fundamentales del Patrimonio Cultural” del grado en Historia del Arte, quienes trabajaron sobre las fichas diagnóstico seleccionadas por el equipo docente, recopilando los datos esenciales del inmueble, entre ellos sobre su catalogación, así como realizando un análisis de la situación en la que se encuentra el bien, en la que se incluye una propuesta de uso, ideas para frenar el daño mediante unos “primeros auxilios”, y una intervención más específica y definitiva para recuperar el elemento. El alumnado tuvo que acompañar esta documentación con imágenes históricas, actuales y de la futura propuesta, a la vez que ampliar el número de referencias y bibliografía utilizada, haciendo una búsqueda de documentos más específicos de su disciplina.

Cabe aquí citar al alumnado de la asignatura “Arquitectura contemporánea”, que, haciendo directamente la selección del patrimonio en peligro para este proyecto, realizó la ficha descrita en el párrafo anterior. Esta decisión viene planificada debido a diversas cuestiones. Por un lado, a que dicha asignatura se imparte en el cuarto curso, año del grado en el cual su formación es más amplia, a la vez que durante el curso 2018-2019 los estudiantes ya habían participado en la asignatura “Historia y Conceptos Fundamentales del Patrimonio Cultural”, por lo que se encontraban familiarizados con el sistema. Por otro lado, ante la posibilidad de que no fuesen seleccionados bienes inmuebles de este periodo cronológico por parte del grado en Turismo, se permitió la elección directa, lo que además ha ampliado el abanico de tipologías a incluir en el proyecto.

Como ejemplo de bienes patrimoniales en riesgo en la ciudad de Málaga que han sido seleccionados para el proyecto, podemos mencionar el caso de la muralla islámica, donde en sus tramos visibles en la Plazuela Virgen de las Penas y en la C/ Arcos de la Cabeza ha sufrido varios daños causados por el vandalismo y por la presión urbanística, pues sobre ella se han construido varias viviendas, así como por el abandono, que ha provocado la aparición de vegetación en los muros (Figura 1). Otro caso a reseñar es el del *Molino de la Bóveda*, arquitectura vernácula con destacados valores históricos y etnológicos de diversos momentos del desarrollo de Torremolinos

(Málaga) y que se encuentra en un gravísimo estado de deterioro (Figura 2). Igualmente, el antiguo *Cine Lope de Vega*, inmueble que ha tenido diversas funciones hasta su abandono en la década de 1980 y que, tras su recuperación, podría paliar las carencias de usos existentes en su sector urbano (Figura 3). Por último, de época romana, destacamos el caso de la *Villa romana de los Molinillos* en Benalmádena (Málaga) que, a pesar del buen estado de conservación en el que se encontraba durante su excavación, dando lugar al conocimiento tanto de la *pars urbana* como de la *pars fructuaria* de la villa, en la actualidad adolece de medidas de conservación y adecuación que permitan la comprensión y el disfrute de la sociedad siendo objeto de actos vandálicos tales como pintadas en las estructuras (Figura 4).



Figura 1. Muralla islámica de Málaga, tramo de la Plazuela Virgen de las Penas. Fuente: ©Google Maps.



Figura 2. Antiguo *Molino de la Bóveda*, en la Cuesta del Tajo, Torremolinos (Málaga, 2016). Fuente: Miguel Vila, arqueólogo.



Figura 3.

Antiguo *Cine Lope de Vega*, en la Avenida Juan Sebastián Elcano, 65, Málaga (marzo de 2021).

Fuente: Cayetano Villodres.



Figura 4. Vista del estado actual del yacimiento *Villa romana de Los Molinillos*, Benalmádena (Málaga, mayo de 2021).

Fuente: ©Google Maps.

Una vez culminados los trabajos por parte del alumnado, en la tercera fase del proyecto toda la información recopilada ha sido incorporada por parte de los estudiantes a la plataforma *Patrimonio Herido* (<https://patrimonioherido.iarthislabs.eu>), desarrollada por Antonio Cruces Rodríguez. Esta plataforma ha sido creada con la finalidad de hacer visibles los daños en el patrimonio por parte de cualquier ciudadano. Para facilitar la supervisión del profesorado sobre las contribuciones al trabajo-proyecto realizadas por el alumnado, y, al mismo tiempo, permitir una publicación rápida, de alcance global, con un método de trabajo cuya curva de aprendizaje fuese muy leve, sin dejar por ello de ofrecer todas las garantías de seguridad y privacidad, *Patrimonio herido* (PH) se ha implementado mediante un mecanismo internacionalmente reconocido y utilizado con éxito en

diversos proyectos del grupo de investigación iArtHis_LAB y su rama tecnológica iASD (iArtHis_LAB Software Development), cuyo despliegue ha dado lugar a una tesis doctoral, y que ha servido como herramienta clave en dos más.

Se trata de una combinación de temas y complementos –plugins– desarrollados sobre la aplicación de fuente abierta WordPress (WP), cimentada en un conjunto de código PHP y JavaScript, más MySQL/MariaDB como base de datos. Haciendo uso de su versatilidad, los estudiantes, previamente identificados y una vez aceptados los términos legales de privacidad y las restricciones derivadas del uso de imágenes, envían a la plataforma sus hallazgos y colaboraciones. Tales envíos son inmediatamente notificados al profesorado destinatario, que, tras realizar la correspondiente evaluación, puede publicar lo recibido, o bien reclamar ulteriores revisiones a los remitentes.

El sistema, que puede consultarse en español e inglés, es tan adaptable y configurable que permite, en realidad, la participación de cualquier persona que desee aportar su ayuda en la difícil pero necesaria tarea de la defensa del patrimonio artístico, histórico, arqueológico y arquitectónico, bajo la supervisión de investigadores y docentes pertenecientes al Departamento de Historia del Arte de la Universidad de Málaga. Adicionalmente, PH ofrece no solo una interfaz de publicación electrónica con acceso directo desde Internet para cualquier internauta, sino como repositorio OAI2-PMH, estándar abierto que posibilita su incorporación directa a depósitos de información de alcance internacional tales como *Hispana*, dependiente del Ministerio de Cultura, y *Europeana* (auspiciado por la Unión Europea), ampliando así de forma muy notable el ámbito de difusión de sus contenidos (Figura 5).



Figura 5. Vista de la portada de la página de inicio del portal *Patrimonio Herido* (2021). Fuente: ©patrimonioherido.iarthislab.eu.

En el caso concreto de la plataforma *Patrimonio Herido*, para la localización de los bienes seleccionados se ha utilizado un sistema cartográfico open source (fundamentalmente OpenStreetMap) y un mecanismo de gestión Leaflet, basado en JavaScript. Gracias a ello, la plataforma puede mostrar mapas globales en los que se visualiza la localización de todos los bienes culturales incluidos en el sistema. Asimismo, gracias a la georreferenciación de los bienes, puede aportar operaciones SIG, como la generación de heatmaps, clusterización, operaciones geométricas, interpolación, etc.

Este trabajo concluye, al final del mes de mayo, con la realización de un seminario que lleva por título *Ética del cuidado y patrimonio cultural. Estrategias y participación ciudadana*. En él se reunirá el alumnado adscrito al proyecto de los grados en Turismo e Historia del Arte con profesionales en la materia y la ciudadanía para dar a

conocer el proyecto y posibilitar que éste salga del aula y adquiera una función social que permita la retroalimentación de todos los agentes implicados.

Como resultados de la realización del trabajo-proyecto sobre el *Patrimonio Herido*, hemos observado que este tipo de actividades formativas contribuyen a familiarizar al alumnado, en especial al de primer curso, con las problemáticas actuales relativas al patrimonio cultural, haciéndolo a su vez más próximo por acercarlo a su realidad cotidiana. Por su carácter transversal, así como por partir de la metodología ABP, este trabajo-proyecto ha acercado además al estudiante a la resolución de problemas reales que se encontrará en su actividad profesional, fomentando la originalidad en los planteamientos resolutivos y en el proceso de la investigación, que le serán de gran utilidad para afrontar desafíos posteriores en su formación, como el TFG, el TFM o una Tesis Doctoral. Asimismo, este trabajo-proyecto ha contribuido al desarrollo de una metodología de trabajo basada en fichas de diagnóstico sobre bienes patrimoniales, que le será de gran utilidad para su futura práctica profesional. Igualmente, al incorporar en las fichas la descripción histórico-artística del bien, ésta le ayuda, por un lado, a manejar la terminología específica de esta disciplina y a su mejor asimilación, y, por otro, a estructurar correctamente y de forma sintética los comentarios de los bienes culturales, tarea a la que tanto el alumnado de Turismo como el de Historia del Arte deberán enfrentarse a lo largo de su formación académica y en el desarrollo posterior de su actividad profesional.

Por otro lado, gracias al uso de las TICs por parte del alumnado, los estudiantes han sido partícipes de la elaboración de sus propios materiales didácticos, lo que ha redundado en una mayor motivación. Esto puede verse, por ejemplo, en el caso concreto de la elaboración de los mapas de localización de los bienes patrimoniales seleccionados como complemento a las fichas diagnóstico, para los que los estudiantes han utilizado las herramientas de Google, en concreto Google Earth y Google Maps. El uso de estos programas ha supuesto una primera aproximación del alumnado al uso de herramientas de carácter geográfico, que permiten posicionar un determinado objeto patrimonial en el espacio.

Para facilitar el uso de herramientas SIG por parte del estudiantado, se ha llevado a cabo, en la asignatura "Historia y Conceptos Fundamentales del Patrimonio Cultural" a

través del *I seminario de Patrimonio Cultural. El patrimonio arquitectónico: la importancia de su estudio y puesta en valor* una actividad impartida por el consultor territorial Anton Ozomek Fernández, bajo el título "Territorio y Patrimonio: GIS online". A través de la misma, el alumnado se ha familiarizado con portales como: Google Maps, la Oficina Virtual del Catastro, la Fototeca Digital del Instituto Geográfico Nacional, la Guía Digital del Patrimonio Cultural de Andalucía, el de la Gerencia Municipal de Urbanismo, Obras e Infraestructuras de Málaga.

Bibliografía

CAMBIL HERNÁNDEZ, María Encarnación; ROMERO SÁNCHEZ, Guadalupe. "Metodología por proyectos: un modelo innovador para la enseñanza y el aprendizaje del patrimonio cultural". En: CAMBIL HERNÁNDEZ, María Encarnación; TUDELA SANCHO, Antonio (coords.). *Educación y patrimonio cultural. Fundamentos, contextos y estrategias didácticas*. Madrid: Pirámide, 2017, pp. 61-80.

GUTIÉRREZ CARRILLO, M^a Lourdes; ROMERO SÁNCHEZ, Guadalupe. "Arquitectura defensiva de Granada. Propuesta metodológica fundamentada en itinerarios didácticos y ABP para su conocimiento y conservación". En: BERMÚDEZ SÁNCHEZ, M^a Paz (coord.). *Actas del congreso: Evaluación de la calidad de la investigación y de la Educación Superior. XVI FECIES*. Granada: Asociación Española de Psicología Conductual (AEPC), 2019, p. 527.

MARCOS COBALEDA, María; CALDERÓN ROCA, Belén. "Los trabajos-proyecto en Educación Infantil: propuesta metodológica para la enseñanza del Patrimonio cultural andaluz". En: CALDEVILLA DOMÍNGUEZ, David; ALARCÓN OROZCO, Elena; ALARCÓN MARTÍNEZ, Virginia (coords.). *Reformulando la docencia actual*. Barcelona: Gedisa, 2018, pp. 203-2018.

RESTREPO GÓMEZ, Bernardo. "Aprendizaje Basado en Problemas (ABP). Una innovación didáctica para la docencia universitaria". *Educación y educadores*, nº 8 (2005): 9-20.

SANTANA GUZMÁN, Antonio Jesús. "Experiencias sobre el acercamiento del patrimonio malagueño a la ciudadanía". *Universitas*, n. 33 (2020): 103-124. doi: <https://doi.org/10.17163/uni.n33.2020.05>

La Edición Digital de Fernando Pessoa: consideraciones sobre las ediciones digitales como archivos abiertos

Ulrike Henny-Krahmer – Universidad de Colonia

Jorge Uribe – Universidad EAFIT

Pedro Sepúlveda – Universidade Nova de Lisboa

Inês Rebelo do Carmo – Universidade Nova de Lisboa

Palabras clave: edición digital, archivo digital, TEI, literatura portuguesa, Fernando Pessoa

Introducción y objetivos

Esta comunicación busca dar a conocer entre el público hispanohablante la *Edición Digital de Fernando Pessoa: Proyectos y Publicaciones* (Sepúlveda, Henny-Krahmer, Uribe 2017ff., <http://www.pessoadigital.pt>) que ofrece a los lectores un espacio de consulta digital en el que las fuentes primarias, las publicaciones en revistas y los manuscritos autógrafos del autor portugués y modernista Fernando Pessoa (1888-1935) pueden ser leídos junto a un texto preparado por los editores. Las transcripciones y las codificaciones de todos los documentos siguen la norma de la *Text Encoding Initiative* (TEI) y se ofrecen en acceso abierto.

El uso de la *Edición Digital de Fernando Pessoa*, por parte de los lectores, en el ámbito de clases y eventos académicos promueve un acceso más directo a las fuentes, llevando la mediación de las herramientas digitales a una mayor proximidad con la materialidad de los textos que componen la obra, sin descuidar la importancia del trabajo editorial de selección y tratamiento interpretativo de estos materiales. La edición, que se ha desarrollado en el marco de una colaboración internacional entre investigadores residentes en Alemania, Portugal y Colombia utiliza el acceso digital para ofrecer una lectura a distancia de materiales de archivo pertenecientes a diferentes fondos de bibliotecas y colecciones privadas, permitiendo así el acceso global a un patrimonio cultural más allá del determinismo geográfico. El proyecto también ha avanzado y se ha ampliado en fases sucesivas, empezando por la edición de dos conjuntos: 1) listas editoriales y 2) poesía publicada en vida, y más recientemente continuando con la inclusión de un tercero, 3) la prosa publicada por el autor en revistas y periódicos. Así, la edición digital constituye un espacio que se desarrolla continuamente y se completa en dirección a un archivo, lo que requiere repensar su concepción en términos de versiones específicas.

Partiendo de la concepción de esta edición específica y tomando en consideración tendencias generales, otro objetivo de esta comunicación es reflexionar acerca de aspectos metodológicos de las ediciones digitales como “archivos abiertos”: la digitalización de los materiales, su tratamiento computacional y su presentación en la web pueden contribuir a aproximar las ediciones digitales a archivos que se abren para los lectores e investigadores, es decir, que son accesibles y también extensibles. Sin embargo, esa tendencia no se aplica a todo tipo de edición digital sino que depende de la selección de los materiales, del nivel de “digitalidad” así como de las posibilidades y estrategias de organización de los proyectos de edición. Además, una edición digital abierta en este sentido conlleva algunos desafíos metodológicos concernientes a su sostenibilidad. Sobre esas características de ediciones digitales como archivos abiertos reflexiona la presente propuesta.

La obra de Fernando Pessoa y su edición en [pessoadigital.pt](http://www.pessoadigital.pt)

Alrededor de Fernando Pessoa, cuya obra se considera la principal contribución en lengua portuguesa al modernismo literario europeo, se creó el mito de un poeta que rechazó la publicación y se refugió en el

anonimato, y que sólo alcanzó un amplio reconocimiento tras su muerte, a través de las publicaciones póstumas de su obra (cf. Simões 1989, 4 y Tabucchi 2019, 3). Como suele ocurrir con los mitos, este alberga una verdad parcial e inexacta, que esconde una compleja relación del autor con la exposición pública de la obra y su reconocimiento en vida. Si bien es cierto que Pessoa solamente publicó un libro de poesía en portugués, *Mensagem*, en 1934, y cuatro folletos de poesía en inglés, entre 1918 y 1921, no lo es menos que fue un colaborador habitual de las más diversas publicaciones periódicas de la época, incluyendo revistas literarias y periódicos generalistas.

Estas publicaciones son, por otra parte, solo una pequeña parte de una obra contenida principalmente en el famoso baúl de papeles que dejó a su muerte, hoy custodiado, principalmente, por la Biblioteca Nacional de Portugal y que suma unos 30.000 documentos (parcialmente accesibles en <http://purl.pt/1000/1/>). Entre estos documentos se encuentran numerosos proyectos y planes editoriales, presentes en notas, en la correspondencia con compañeros y críticos, y en varias listas, que permiten vislumbrar la dimensión potencial de la obra tal como fue concebida en diferentes momentos. La *Edición Digital de Fernando Pessoa: Proyectos y Publicaciones* ofrece nuevas formas de acceder a los escritos del autor, permitiendo conocer el contraste entre el carácter potencial de las listas de proyectos editoriales contenidas en su archivo y la concreción que implican las publicaciones, en periódicos y revistas, que abarcan el período de la vida adulta del autor, entre 1912 y 1935. Como se indica en la portada de la *Edición Digital*, se dispone de toda la poesía publicada en vida en periódicos y revistas, que comprende 72 publicaciones, en su totalidad. La prosa publicada por el autor en publicaciones periódicas y las listas de proyectos editoriales, que comprenden un conjunto de cerca de 120 publicaciones y 248 listas, ya están disponibles en su mayoría en la edición.

Es notable que, siendo Pessoa un autor tan sistemáticamente caracterizado como póstumo, el corpus de sus publicaciones en vida siga creciendo y no esté claro que esté ya definitivamente fijado, aunque haya sido objeto de sucesivos esfuerzos editoriales¹. De hecho, en los últimos veinte años, cerca de una veintena de publicaciones, la mayoría en periódicos, algunas de ellas firmadas y otras anónimas, y algunas que pueden haber sido escritas en colaboración con otros escritores, le han sido atribuidas por varios investigadores, a través de la colación de los documentos de su archivo, para ser integradas en el canon de la obra de Pessoa publicada en vida (cf. Uribe 2020). En este procedimiento, fue de importancia decisiva conocer y revisar las listas de proyectos editoriales en las que Pessoa dejó constancia de su propia actividad editorial, integradas en la *Edición Digital*, y

¹ En este sentido pueden citarse las siguientes ediciones recientes de la obra de Fernando Pessoa: *Mensagem e outros poemas publicados em vida* (Pessoa 2018); *Prosa publicada em vida* (Pessoa 2017a); *Ficções do Interlúdio* (Pessoa 2017b); *Proses – Volume I 1912-1922* y *Proses – Volume II 1923-1935* (Pessoa 2013a y 2013b); *Crítica, ensaios, artigos e entrevistas* (Pessoa 1999).

que en comparación con las publicaciones permiten una visión más completa de los procedimientos de escritura, planificación y publicación de Pessoa a lo largo de los años.

Considerando el carácter disperso de las fuentes que contienen la obra de Fernando Pessoa y también de los documentos que son testimonios de su planificación, y teniendo en cuenta el estado de una investigación que todavía está en desarrollo, una edición digital ofrece la posibilidad de integrar nuevos documentos sucesivamente. Así, por una parte, la selección de materiales para la *Edición Digital* fue predeterminada por un programa específico de investigación y edición: el de comparar la obra publicada en vida con fuentes autógrafas que muestran su planificación editorial. El objetivo de la *Edición Digital* es editar estos documentos críticamente, en el sentido de ofrecer:

- transcripciones diplomáticas en sinopsis con facsímiles digitales
- la indexación de nombres de personas, revistas y títulos mencionados en los textos

- la indicación de referencias entre varios documentos con *links*
- la elaboración de comentarios editoriales que explican el contexto a la historia editorial de cada documento

En el caso de los manuscritos, la transcripción diplomática incluye la codificación de la microgénesis de los documentos, registrando añadidos, sustituciones y borrados por parte del autor. Partiendo de esta transcripción diplomática, se ofrecen dos versiones editadas de los manuscritos, una versión primera y una última versión (véase un ejemplo en la fig. 1).

The screenshot displays a digital manuscript interface. On the left, there is a table of contents for 'Orpheu 3 e 4'. The table lists authors and their corresponding page numbers, with some entries marked with a '#' symbol. Below the table is a 'Note' button. In the center, there are three buttons: 'diplomatic transcription' (highlighted in red), 'first version', and 'last version'. On the right, there is a facsimile of a handwritten manuscript page, labeled 'Facsimile' and 'Indexes'. The facsimile shows a list of authors and page numbers, matching the table of contents, with a total of 128 pages indicated at the bottom. The manuscript page is dated 7/196 and has the number 87-26 in the top right corner.

Author	Page Number
Mario de Sá-Carneiro:	16
Albino de Menzes: #	6
José Coelho Pacheco:	8
Mário da Cunha e Sá: #	8
AN Fernando:	8
José de Almada-Negreiros:	18
Augusto Ferreira Gomes:	10
M. Fernando de Carvalho:	10
Barradas Telles D'Aviz: #	6
Fernando Pessoa:	8
Castello de Moraes:	12
Alvaro de Campos:	12
Indice do 1º Volume:	6
	<u>128 pp.</u>

Fig. 1: Sinopsis de la transcripción diplomática y del facsimil del manuscrito BNP 87-26r.

Las publicaciones de poesía y prosa también se transcribieron diplomáticamente, pero en los casos en que existen varias versiones publicadas durante la vida del autor, se estableció un aparato crítico indicando dónde hay variantes léxicas o de puntuación (véase un ejemplo en la fig. 2).

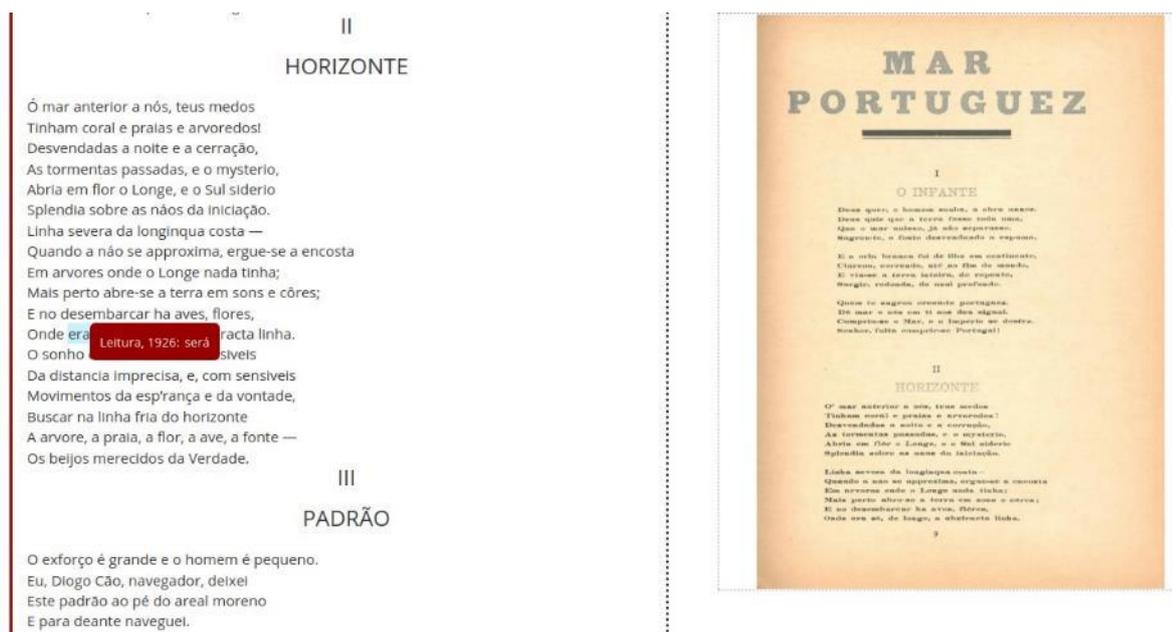


Fig. 2: Sinopsis de la transcripción y del facsímil del poema “Mar Português”, incluyendo una variante léxica.

Los elementos mencionados (selección programática de documentos, transcripciones diplomáticas, versiones editadas de los textos, indexaciones, comentarios) señalan claramente que la *Edición Digital* es una edición elaborada críticamente. Por otra parte, en la publicación de la *Edición Digital* el corpus no está fijado de antemano sino que se va completando siguiendo el estado de trabajo así como de investigación. En este sentido, la *Edición Digital* tiene el carácter de un archivo en construcción, y se publica con la fecha “2017ff.”, indicando que el trabajo alcanzó una primera etapa de finalización ya en el año 2017, pero que todavía se avanza. En la página web, la edición todavía se presenta bajo la indicación “Beta”, señalando que el corpus editado se completa y que también la aplicación web se desarrolla siguiendo la evolución del punto de vista de contenido. Este procedimiento caracteriza la edición como un proyecto abierto. Desde el principio, el trabajo colaborativo en los textos, en sus transcripciones y codificaciones en TEI, se ha realizado públicamente en GitHub (<https://github.com/cceh/pessoa>), con las primeras contribuciones en el año 2015. La frecuencia de las intervenciones muestra un punto culminante en 2017 cuando se lanzó la página web del proyecto (véase la fig. 3).



Fig. 3: Frecuencia de adiciones y eliminaciones en el código de la Edición Digital de Fernando Pessoa, 2015-2021.

Si bien el desarrollo del código y del trabajo en la edición se puede seguir detalladamente en GitHub, lo que no se hizo fue marcar etapas significativas en la evolución de los datos y de la aplicación con versiones semánticas (por ejemplo una versión 0.5, 1.0, etc.), lo que se manifiesta en la tensión entre la fecha de publicación y el adjunto “Beta”. Desde que se inició la *Edición Digital*, la conciencia sobre la importancia de definir estados sucesivos de ediciones y archivos digitales ha crecido (Broyles 2020) y actualmente se elabora una estrategia en el proyecto para definir versiones seleccionadas y citables de la edición.

Pero el carácter de archivo de la *Edición Digital* no solo se basa en su meta de completar un subcorpus específico de la obra de Fernando Pessoa sino también en la proximidad a la materialidad de los textos que componen esta obra, ofreciendo a lectores, estudiantes e investigadores un espacio abierto. Esta proximidad se establece mediante la ya mencionada visualización paralela de las fuentes primarias –de las publicaciones originales o documentos del archivo del autor– y el texto editado. Además, por el uso de *links*, algunos de estos textos, incluidas las listas, se convierten en ejemplos de edición hipervinculada, en la que un documento concreto es una puerta de entrada a otros con los que mantiene una relación intratextual. Esta condición recuerda uno de los rasgos distintivos de la edición digital destacados por la crítica especializada:

In an electronic critical edition the critical text will be the locus of a set of data connected to it by various kinds of links, some established specifically by the editor, others established automatically by software tools. The critical text will not exist as a self-sufficient isolate but rather as part of a rich environment ... (Faulhaber, *apud*. Vanhoutte 2010, par. 23).

El “ambiente rico” se configura explícitamente en esta edición a través de un vínculo entre las listas y las publicaciones, y de estas últimas entre sí, que es el resultado de un amplio trabajo de críticos y editores a lo largo de los años, identificando y traduciendo los vínculos pertinentes entre la aparición de un título o la referencia en una lista y un texto que formaría parte de otro archivo físicamente alejado del de Pessoa, de modo que ahora ambas fuentes conviven en el mismo espacio de lectura y configuran una misma legibilidad. En este sentido, la

Edición Digital se inscribe en el marco de la discusión en las humanidades digitales en torno a los conceptos de "archivo" y "edición", y reivindica esta doble condición como instancias que tienden a una integración orgánica:

Electronic editions and electronic archives, in Tanselle's view, are therefore synonymous.

At the same time, it is true that the meaning of the word 'archive' in connection with electronic textual editing has changed over the course of time. Originally denoting a mere repository of digital surrogates of material artefacts and processed data, the concept has come to include scholarly and critical material such as edited texts, annotations, scholarly essays and the like, alongside the digital resources. This transition has happened organically. (Vanhoutte 2010, par. 32-33).

La lectura de los textos de la *Edición Digital* también se ve facilitada por las diferentes formas de acceder a las mismas fuentes, a través de su indización por autor, tipo de documento, género, cronología y referencias a nombres y títulos de obras o publicaciones periódicas, inmediatamente visibles en la barra central del portal.

Ubicación metodológica de la edición digital como archivo abierto

Como se discutió ya en los últimos apartados, las ediciones digitales tienen una tendencia de acercarse a archivos, en varios niveles (siguiendo McGann 2001; Sahle 2013; Pierazzo 2015):

- constituyen un ambiente rico, hipervinculado, conectando datos de diversa índole (datos bibliográficos, textuales, referenciales, analíticos, etc.)
- pueden ser completados sucesivamente, explotando los archivos análogos y siguiendo el estado de investigación
- tienden a la exhaustividad (incluir todos los materiales accesibles, añadir informaciones contextuales, presentar una obra completa, ofrecer documentos previamente desconocidos o inaccesibles)
- integran una labor crítica con una proximidad a la materialidad de los documentos archivísticos (combinación de texto editado, contexto bibliográfico, facsímiles)

Al mismo tiempo se liberan de la ubicación y delimitación geográfica y organizacional de los archivos análogos, ofreciendo un espacio abierto a los usuarios. A diferencia de archivos en sentido estricto, la edición digital archivística mantiene el elemento crítico con su orientación hacia una obra, una época, un género, un estilo, i.e. siguiendo un criterio literario, histórico o lingüístico que limita la selección de las fuentes que se editan. También, en la edición archivística, todavía persiste el esfuerzo de constituir textos de base, sean de carácter diplomático o crítico, y de ofrecer comentarios filológicos o históricos para elaborar el material.

Sin embargo, la tendencia hacia un archivo abierto no se puede postular para todo tipo de edición digital. Algunos factores determinantes son la extensión de los contenidos de la edición (¿se procura, por ejemplo, crear una edición de una obra completa – o de una obra determinada, seleccionada?), la medida de "digitalidad" de una edición (¿se reproduce una edición impresa o se crea una edición "de origen digital"?) y también el marco del proyecto de edición (¿es un proyecto estrictamente limitado temporalmente, de capacidad del personal, y de financiación, o es un proyecto extensible?). Ya en el contexto de la obra de Fernando Pessoa se han realizado varios tipos de ediciones digitales. Aparte de la *Edición Digital* presentada en esta propuesta, existen una edición del "Libro del desasociado" (Portela e Rito Silva 2017, <https://ldod.uc.pt/>), una de la obra de teatro "Fausto" (Centro de Estudos de Teatro 2015, <http://www.faustodigital.com>) y un archivo de los textos publicados y editados anteriormente del autor (Areal 2008 <http://arquivopessoa.net>). La primera también se acerca a un archivo abierto porque es extensible y colaborativa y permite crear múltiples ediciones sobre la base de fragmentos de texto de la obra seleccionada. La segunda ofrece facsímiles y texto, pero se concentra en una obra específica del autor y en ofrecer un aparato crítico para el texto. La última procura presentar la obra entera del autor, pero reutiliza ediciones impresas.

Dependiendo del carácter y de la medida “archivística” de las ediciones digitales, se imponen ciertos desafíos metodológicos. Especialmente una edición abierta en el sentido de un trabajo en fases, como en el caso de la *Edición Digital de Fernando Pessoa*, significa que el modelo de datos, la estructura del módulo de presentación (el portal web) y las funcionalidades de la edición se adaptan en conformidad con los nuevos materiales añadidos y el trabajo crítico en progresión. Para asegurar la sostenibilidad de una edición de este tipo es necesario tomar medidas apropiadas como la generación de diferentes versiones. La discusión sobre el planeamiento, la realización, la publicación, y el almacenamiento de ediciones digitales del tipo “archivo abierto” en los niveles de contenido, tecnología y organización se encuentra todavía en sus albores.

Conclusiones

Es posible constatar que la obra del modernista portugués Fernando Pessoa se adecua con una edición digital del tipo “archivo abierto”, por ser una obra caracterizada por la tensión entre un planeamiento meticuloso y documentado en numerosos manuscritos autógrafos, y una publicación delimitada y cuidadosamente considerada por el autor, pero también dispersa en muchas revistas y periódicos. Al tratarse de una obra en la que la relación entre las partes y el todo es fundamental, la lectura invita a reflexionar sobre la forma particular en que las partes (los textos y sus diversos elementos) interactúan con el todo (la obra). Las ediciones digitales son herramientas versátiles: por un lado, permiten integrar diferentes tipos de fuentes y seguirlas de cerca, como en el caso de las variantes intratextuales de las listas editoriales y las diferentes versiones de las publicaciones de Pessoa. Por otra parte, todos los datos recogidos y creados en el proceso de representación digital de las fuentes constituyen también un archivo que puede ser interrogado y analizado desde diferentes perspectivas, para crear, por ejemplo, índices, redes, líneas de tiempo, representaciones gráficas y otras formas que exploran y sintetizan los datos de la edición, con vistas a producir conocimiento. De este modo, se crea una nueva versión de la obra del poeta, que invita al usuario de la *Edición Digital* a comparar productivamente lo que se le ofrece y a plantear sus propias preguntas a los materiales presentados, según la visión de Moretti (2008) de que el uso de los archivos digitales (y necesariamente también de las ediciones que se acercan a ellos) da lugar a una experimentación y a una interrogación de la cultura: “‘questions put to nature’ is how experiments are often described, and what I’m imagining here are questions put to culture” (114). La reflexión sobre las características de las ediciones digitales como archivos y los diferentes tipos de ediciones digitales de Fernando Pessoa mostraron que la metodología editorial se vincula estrechamente con el objeto de la edición digital, con consecuencias específicas para su producción y mantenimiento.

Bibliografía

Broyles, Paul A. (2020): “Digital Editions and Version Numbering.” *Digital Humanities Quarterly* 14 (2). <http://www.digitalhumanities.org/dhq/vol/14/2/000455/000455.html>.

McGann, Jerome J. (2001). *Radiant Textuality*. Literature after the World Wide Web. Palgrave. Moretti, Franco (2008). “The Novel: History and Theory”. *New Left Review* 42, 111-124.

Pessoa, Fernando (2018). *Mensagem e Poemas Publicados em Vida*. Edição de Luiz Fagundes Duarte, INCM.

---- (2017a). *Prosa publicada em vida*. Edição de Richard Zenith, Assírio & Alvim.

---- (2017b). *Ficções do Interlúdio*. Edição de Fernando Cabral Martins, Biblioteca de editores independentes.

---- (2013a). *Proses – Volume I 1912-1922*. Édition de José Blanco, Éditions de La Différence.

---- (2013b). *Proses – Volume II 1923-1935*. Édition de José Blanco, Éditions de La Différence.

---- (1999). *Crítica, ensaios, artigos e entrevistas*. Edição de Fernando Cabral Martins, Assírio & Alvim.

Pierazzo, Elena (2015). *Digital Scholarly Editing. Theories, Models and Methods*, Ashgate.

Sahle, Patrick (2013). *Digitale Editionsformen. Zum Umgang mit der Überlieferung unter den Bedingungen des Medienwandels. Teil 2: Befunde, Theorie und Methodik*. Schriften des Instituts für Dokumentologie und Editorik 8. BoD.

Sepúlveda, Pedro, Ulrike Henny-Krahmer e Jorge Uribe (eds.) (2017ff.). *Edição Digital de Fernando Pessoa: Projetos e Publicações*. Coordenação editorial por Pedro Sepúlveda, coordenação técnica por Ulrike Henny-Krahmer, IELT, Universidade Nova de Lisboa e CCEH, Universidade de Colónia, <http://www.pessoadigital.pt>. DOI: 10.18716/cceh/pessoa.

Simões, João Gaspar (1989). *Fernando Pessoa: retrato memória*, Universidade Católica Portuguesa.

Tabucchi, Antonio (2019). *Un baule pieno di gente. Scritti su Fernando Pessoa*, Feltrinelli.

Uribe, Jorge (2020). "Em vida de Fernando Pessoa - Lista de Publicações", *Revista Estranhar Pessoa* 7, editado por Jorge Uribe e Pedro Sepúlveda, IELT, FCSH, Universidade Nova de Lisboa, 13-52. <http://estranharpessoa.com/revista>.

Vanhoutte, Edward (2010). "5. Defining Electronic Editions: A Historical and Functional Perspective". *Text and Genre in Reconstruction: Effects of Digitalization on Ideas, Behaviours, Products and Institutions*, Open Book Publishers. <http://books.openedition.org/obp/654>.

El área de Ciencias y Técnicas Historiográficas como ejemplo de implantación de las Humanidades Digitales en el progreso de la docencia

Lorena C. Barco Cebrián y Alicia Marchant Rivera

Universidad de Málaga

Palabras clave: Humanidades Digitales, Historia de la Cultura Escrita, Paleografía, Docencia, Universidad.

Mediante nuestro trabajo pretendemos dar a conocer el progreso que se está llevando a término dentro del área de Ciencias y Técnicas Historiográficas de las Universidad de Málaga, exponiendo cómo ha cambiado la docencia de las diferentes asignaturas impartidas en el área, donde un papel primordial ha tenido la implantación de las Humanidades Digitales en el proceso de enseñanza-aprendizaje.

En el área de Ciencias y Técnicas Historiográficas de la Facultad de Filosofía y Letras de la Universidad de Málaga se dan diferentes asignaturas, todas ellas muy relacionadas también con las Humanidades Digitales. La docencia impartida desde nuestra área se centra principalmente en el Grado de Historia, al cual está adscrita. Así impartimos las asignaturas obligatorias en dicho Grado de *Paleografía General I* y *Paleografía General II*, donde se imparte la enseñanza de las diferentes grafías que se desarrollan en la Península Ibérica desde la conquista romana hasta el siglo XVIII, asignatura que se imparte en el 2º curso. Dentro del mismo Grado, en esta ocasión en el tercer curso del mismo, se imparte la asignatura optativa de *Acceso a las Fuentes Historiográficas*, la cual se imparte en su integridad en aula TIC y tiene un carácter eminentemente práctico y se desarrolla a través de las diferentes herramientas que ofrece internet para el progreso del trabajo de cualquier historiador, acercando al alumno a un gran número de recursos digitales que hoy en día se ofrecen para desarrollar la labor del historiador, por lo que es una asignatura que se integra dentro de las denominadas como Humanidades Digitales. En cuarto curso del dicho Grado de Historia la oferta es mayor y se imparten desde el área tres asignaturas optativas: *Archivística*, *El documento y el libro en la Edad Media*, y *El documento y el libro en la Edad Moderna*, las dos últimas dentro de lo que se denominan como Diplomática y Codicología. Las tres asignaturas utilizan y desarrollan en el proceso de enseñanza y aprendizaje desde los últimos años hasta la actualidad métodos y herramientas integradas en las denominadas como Humanidades Digitales, acercando al alumno otras herramientas diferentes a las utilizadas tradicionalmente. Por otro lado, también impartimos dos asignaturas dentro del *Máster Oficial en Gestión del patrimonio literario y lingüístico español* que se desarrolla en la Facultad de Filosofía y Letras de la Universidad de Málaga. A través de las asignaturas optativas de *Recuperación del patrimonio literario y lingüístico español* y *Edición Crítica de Textos*, se acerca al alumno las herramientas indispensables para poder hacer frente a la lectura de un documento histórico y se dan las herramientas web necesarias para la búsqueda de fuentes primarias útiles para su desarrollo como investigador del español. Además, en el curso actual, 2018-2019, se ha puesto en marcha por primera vez un Máster Propio, desarrollado íntegramente por el área de Ciencias y Técnicas Historiográficas en el que participan especialistas en las diferentes disciplinas que la integran y en Humanidades Digitales del ámbito nacional, cuyo título es *Máster Propio en Ciencias y Técnicas Historiográficas y Humanidades Digitales*, el cual se imparte íntegramente online, facilitando así el acceso a estas ciencias desde cualquier parte del mundo e integrando desde su génesis las Humanidades Digitales. Por último, desde el año 2016 se vienen realizando una serie de cursos online auspiciados por el área de Ciencias y Técnicas Historiográficas y la Fundación General de la Universidad de Málaga, cuyo objeto es acercar la Paleografía a cualquier persona interesada en ella que la necesite por cualquier

motivo personal o profesional. Estos cursos nacieron porque desde que se implantaron los Grados, la asignatura optativa o de libre configuración de *Paleografía* que se ofertaba a otras licenciaturas tales como Filología Clásica, Filología Hispánica o Historia del Arte, dejaron de ofertarse, y desde entonces los alumnos de dichas carreras, incluso algunos ya egresados, denunciaban una falta de formación y conocimiento en esta disciplina tan necesaria para poder leer y entender las fuentes originales y primarias a las que por necesidad deben enfrentarse a lo largo de sus carreras académicas y profesionales. Por ello, y a instancias de esta demanda creciente por parte de dichos alumnos, se decidió impartir los mencionados cursos online, los cuales acercan al alumno las herramientas necesarias para poder enfrentarse de forma satisfactoria a la documentación original. A través de dichos cursos, y por su propia naturaleza, fue necesario incluir en el proceso de enseñanza de los mismos diferentes herramientas que quedan incluidas en las Humanidades Digitales, para poder facilitarle al alumnado la mayor y mejor comprensión de los textos.

BIBLIOGRAFÍA

ANGULO MORALES, Alberto (2006), "Algunas reflexiones sobre los recursos de archivos históricos en Internet y la enseñanza de la Historia", *Hispania* (Madrid) 66, págs. 31-58.

CARTELLI, A. y PALMA, M. (2009). "Digistylus-An Online Information System for Palaeography Teaching and Research". En *Codicology and Palaeography in the Digital Age*, 2, F. Fischer, C. Fritze, G. Vogeler (eds.), B. Assmann, M. Rehbein, P. Sahle (colabs.), págs. 123-134.

CASTRO CORREA, A. (2014). "Paleografía Latina: recursos para docentes y estudiantes sobre cómo no perderse en la Red". *Espacio, Tiempo y Forma. Serie III. Historia Medieval*, 27, págs. 211-228.

CIULA, A. (2005). "Digital Palaeography: Using the Digital Representation of Medieval Script to Support Palaeographic Analysis". *Digital Medievalist*, 1.2, 1-31.

FERROSOTO, Carlos; MARTÍNEZ SENRA, Ana Isabel; OTERO NEIRA, Ma Carmen (2009), "Ventajas del uso de las TICs en el proceso de enseñanza-aprendizaje desde la óptica de los docentes universitarios españoles", *EduTec-e. Revista electrónica de Tecnología Educativa* (Palma de Mallorca) 29 <<http://www.edutec.es/revista/index.php/edutec-e/article/view/451>>

FISCHER, F., FRITZE, C. y VOGELER, G. (eds.); ASSMANN, B., REHBEIN, M. y SAHLE, P. (colabs.) (2010). *Codicology and Palaeography in the Digital Age*, 2.

HASSNER, T., SABLATNIG, R., STUTZMANN, D. y TARTE, S. (eds.) (2014). *Digital Palaeography: New Machines and Old Texts* [Dagstuhl Reports], 4.7, págs. 112-134.

PONS, Anacleto (2013), *El desorden digital. Guía para historiadores y humanistas*, Madrid, Siglo XX.

RAMÍREZ-SÁNCHEZ, Manuel (2003), "Las TIC como apoyo en la docencia universitaria de las Ciencias y Técnicas Historiográficas", *Actas de las I Jornadas canarias sobre las Tecnologías de la Información y la Comunicación en la docencia universitaria*, Universidad de Las Palmas de Gran Canaria, La Laguna-Las Palmas de Gran Canaria, págs. 123-131.

RAMÍREZ-SÁNCHEZ, Manuel (2016), "Las ciencias y técnicas historiográficas en el contexto de las humanidades digitales: oportunidades para su desarrollo", en "*Dicebamus hesterna die...*": estudios en homenaje

a los profesores Pedro J. Arroyal Espigares y M^a Teresa Martín Palma, Alicia Marchant Rivera, Lorena C. Barco Cebrián (coords.), págs. 365-392.

ROJAS-CASTRO, Antonio (2013), "Las Humanidades Digitales: principios, valores y prácticas", *Janus. Estudios sobre el Siglo de Oro*, 2, págs. 74-99.

STOKES, P. (2008). "Palaeography and Image Processing: Some Solutions and Problems". *Digital Medievalist*, 3. University of Lethbridge. Recuperado de <http://www.digitalmedievalist.org/journal/3/stokes/> el 20/04/2021.

____ (2009). "Computer-Aided Palaeography, Present and Future". En *Codicology and Palaeography in the Digital Age*, págs. 309-338.

____ (2010). "Teaching Manuscripts in the Digital Age". En *Codicology and Palaeography in the Digital Age*, 2, F. Fischer, C. Fritze y G. Vogeler (eds.); B. Assman, M. Rehbein y P. Sahle (colabs.), BoD, págs. 229-425.

STUTZMANN, D. (2010). "Paléographie statistique pour décrire, identifier, dater... Normaliser pour coopérer et aller plus loin?". *Codicology and Palaeography in the Digital Age*, 2, F. Fischer, C. Fritze y G. Vogeler (eds.), B. Assmann, M. Rehbein y P. Sahle (colabs.), BoD, págs. 247-277.

WOLF, L., POTIKHA, L., DERSHOWITZ, N. (2011). "Computerized Paleography: Tools for Historical Manuscripts". En *Proceedings of International Conference on Image Processing (ICIP 2011)*, págs. 3545-3548.

ZOZAYA-MONTES, L. (2009). "El desarrollo de competencias en Historia y Ciencias y Técnicas Historiográficas a través de un espacio virtual en Internet". *Relada (Revista Electrónica de ADA)*, 3.3, págs. 240-248.

____ (2012). "La aplicación de técnicas informáticas con fines docentes para un manual de paleografía". En *La escritura de la memoria. Libros para la administración. IX Jornadas de la Sociedad Española de Ciencias y Técnicas Historiográficas*, J.Á. Munita Loinaz y J.A. Lema Pueyo (eds.), págs. 365-370.

____ (2014), "Cursos online de paleografía: he- rencias, limitaciones, logros y propuestas", *El profesional de la información* (Barcelona), 23/5, págs. 475-484.

____ (2017), "NTIC e innovación docente: juegos en soporte digital para complementar el aprendizaje paleografía", *Revista de Humanidades Digitales*, 1, págs. 150-180.

La decolonización de lo imaginario: hacia la inclusividad y la consideración de lo subalterno en las Humanidades Digitales

Moderador: César González-Pérez · *Asiste:* Alexandra Iglesias Cortés

El campo de las Humanidades Digitales ha emergido en los últimos años como un área no solamente emocionante por su carácter altamente interdisciplinario, sino también necesaria por su mirada innovadora hacia temas considerados como tradicionales dentro de las Humanidades. Tecnologías derivadas de campos como la Inteligencia Artificial incluyendo subdisciplinas como el Procesamiento Natural de Lenguaje y Visión por Computadora nos permiten ahora, por ejemplo, el minado automático de colecciones que pueden contener cientos o miles de textos y el reconocimiento automatizado de elementos en imágenes de interés histórico y arqueológico. Hace solo unas décadas, estos avances estaban todavía fuera del alcance de la gente no especializada y particularmente fuera de las ciencias computacionales. El avance tecnológico, la democratización de la información, así como la adopción de una posición más interdisciplinaria, ha permitido a las Humanidades dar un giro y aprovechar los beneficios que estos desarrollos representan. A pesar de esto, la mayoría de las tecnologías adoptadas, las metodologías, e incluso las ideas y acercamientos teóricos que derivan de ellas están profundamente enraizadas en la epistemología moderna occidental. Tomando como punto de partida la historia del encuentro entre las sociedades Indígenas mesoamericanas y las europeas, en esta plática recorreremos la espinosa historia de la opresión de conocimientos y culturas, analizaremos las implicaciones de la utilización de estas tecnologías para su estudio y su papel en la reproducción del pensamiento colonial, y reflexionaremos acerca de la importancia de la consideración de los marcos ideológicos y conceptuales tanto de las sociedades premodernas como Indígenas.

Panel 7

INTELE: Red estratégica para integrarse en las infraestructuras digitales europeas de investigación en Humanidades

Modera: Mikel Iruskieta · *Asiste:* Brianda Otero Moreira

Panelistas: Amelia Sanz Cabrerizo; José Calvo Tello

INTELE: Red estratégica para integrarse en las infraestructuras digitales europeas de investigación en Humanidades

Coordinador

Germán Rigau

(Universidad del País Vasco)

Participantes

German Rigau

Núria Bel

Amelia Sanz

Mikel Iruskieta y José Calvo Tello

Palabras clave

Infraestructuras europeas, Humanidades Digitales, Centro K, DARIAH, CLARIN

Descripción del problema de investigación, metodología y resultados

El comité *European Science Future Research Infrastructures* (ESFRI) fundamentó la importancia de las infraestructuras para obtener resultados de investigación de una forma eficaz y con los medios disponibles en la actualidad. El comité ESFRI recomendó la creación de dos infraestructuras de investigación para los ámbitos de humanidades: CLARIN (www.clarin.eu) y DARIAH (www.dariah.eu) que se consolidaron como ERICs (European Research Infrastructure Consortium) en 2014.

El propósito de este panel es dar a conocer la red INTELE a la comunidad de humanistas digitales. INTELE es una red de investigación estratégica constituida inicialmente por investigadores españoles que están relacionados, por su participación anterior y su interés actual, con las actuales infraestructuras europeas de investigación, ya constituidas como European Research Infrastructure Consortium o ERIC, para las humanidades y ciencias sociales: CLARIN y DARIAH. El propósito de CLARIN (*Common Language Resources and Technology Infrastructure*) es fomentar y difundir conocimiento, herramientas, datos y métodos de trabajo basados en el procesamiento y explotación automáticos de textos (escritos u orales) para las áreas de eHumanidades y ciencias sociales. Desde hace más de 12 años, CLARIN funciona como una red de centros que comparten misión, tecnología y recursos para ponerse a disposición de los investigadores de los Estados que son miembros de la infraestructura. Actualmente CLARIN brinda servicios a investigadores en psicología, lingüística, filología, ciencias políticas, o sociología entre otros. El propósito de DARIAH (*Digital Research Infrastructure for the Arts and the Humanities*) es facilitar el acceso y uso de la información de investigación digital sobre el Arte y las Humanidades europeas, constituyendo una red de conexión entre personas, información, herramientas y metodologías para investigar, explorar y apoyar el trabajo a través del amplio espectro

de las eHumanidades. Ambas infraestructuras se consideran complementarias y han participado junto con la infraestructura para arqueología ARIADNE en PARTHENOS: *Pooling Activities, Resources and Tools for Heritage E-research Networking, Optimization and Synergies*, iniciativa en la que ponen en común herramientas, recursos y actividades, para conservar el patrimonio. Además, en Holanda ambas infraestructuras se han juntado y han creado CLARIAH (www.clariah.nl), fusionando los objetivos de ambas infraestructuras.

La Red INTELE está financiada por el Ministerio Ciencia, Innovación y Universidades de España (RED2018-102797-E), en el marco de las acciones de dinamización «Redes de Investigación» correspondientes al Programa Estatal de Generación de Conocimiento y Fortalecimiento Científico y Tecnológico del Sistema de I+D+i, (2019-2021).

Grupos de investigación en INTELE

- i) **Grupo IXA - UPV/EHU (Universidad del País Vasco)**
- ii) **Grupo IULATERM-TRL-UPF (Universitat Pompeu Fabra, Barcelona)**
- iii) **LINHD - UNED (Universidad Nacional de Educación a Distancia en Madrid)**
- iv) **Fundación Biblioteca Virtual Miguel de Cervantes (Universidad de Alicante)**
- v) **Grupo TALG-UVIGO (Universidad de Vigo)**
- vi) **Grupo LOEP (Universidad Complutense de Madrid)**
- vii) **Grupo UJAEN (Universidad de Jaén)**

Descripción de las intervenciones de cada participante

"INTELE: sus objetivos" German Rigau (Universidad del País Vasco)

El objetivo general de la red INTELE (RED2018-102797-E) es impulsar actividades de promoción de las dos infraestructuras de investigación para los ámbitos de humanidades: CLARIN (www.clarin.eu) y DARIAH (www.dariah.eu) y lograr la participación oficial de España en las mismas. La participación de España ha de contribuir al avance de la investigación española en humanidades y ciencias sociales, así como a su posicionamiento estratégico en proyectos y programas internacionales, fundamentalmente en el contexto del Espacio Europeo de Investigación.

INTELE ha reunido en la red a grupos que están interesados en estas infraestructuras europeas de investigación que tienen el objetivo de reducir la brecha digital e impulsar nuevas líneas de investigación multidisciplinares en humanidades y ciencias sociales, participando en una transformación digital de las mismas con la ayuda de las tecnologías del lenguaje.

En este panel presentaremos los resultados obtenidos por la red hasta el momento y los objetivos que se plantean para un futuro próximo.

“Presentación del CLARIN Centro-K: algunos proyectos” *Núria Bel*

La infraestructura de investigación CLARIN está formada por una red de centros que incluye grupos de investigación de universidades, centros, academias nacionales de ciencia y academias nacionales de la lengua de cada Estado miembro. Además, el consorcio CLARIN ha invitado a incorporarse, como centro de conocimiento o CentreK, a grupos, de estados asociados o no – como es el caso de España–, que presten servicios estables y ofrezcan sus conocimientos y experiencia para la investigación en textos de una lengua (o tema) particular. En 2015, CLARIN concedió el reconocimiento de Centre-K especializado en las lenguas oficiales en España y desde entonces ofrece herramientas, aplicaciones web, y asesoría para el desarrollo de proyectos. El CLARIN Centro-K-español está formado por el grupo IULATERM de la Universidad Pompeu Fabra, el Laboratorio de Innovación en Humanidades Digitales de la Universidad Nacional de Educación a Distancia, el Grupo IXA de la Universidad del País Vasco/Euskal Herriko Unibertsitatea y el Grupo de Investigación en Tecnoloxías e Aplicacións da Lingua Galega de la Universidad de Vigo. El objetivo de esta intervención es presentar el Centre-K español y dar ejemplos de los servicios que ha prestado con el fin de darlo a conocer entre los investigadores de humanidades digitales para que puedan beneficiarse.

“La hoja de ruta de DARIAH: sus propuestas, sus aportaciones” *Amelia Sanz* (Universidad Complutense de Madrid)

Proponemos presentar las grandes líneas de los planes estratégicos de DARIAH para 2022-2026 y, a partir de ahí, sus aportaciones en lo que se refiere, primero, al

desarrollo de grandes proyectos sobre infraestructuras (PARTHENOS, SSHOC, TRIPLE) y a la formación en Humanidades Digitales (DARIAH-TEACH y DARIAHCAMPUS) y, segundo, a los servicios que ofrecen a la comunidad y las herramientas que comparten.

En este sentido, nos detendremos particularmente en presentar el *SSH Open Market*, un *marketplace* que servirá de vitrina a las herramientas disponibles para las *Digital Humanities*, como alternativa europea en abierto y al día a repositorios ya abandonados como DIRT, no actualizados como DH Toychest o la canadiense TAPOR.

“Facilitando la investigación en la nube del corpus de novelas de la Edad de Plata con las infraestructuras europeas CLARIN y DARIAH” *Mikel Iruskieta* (Universidad del País Vasco) y *José Calvo Tello* (Universidad de Goettingen)

En la presentación se utiliza el Corpus de novelas de la Edad de Plata, una colección de más de 358 novelas publicadas por autores españoles entre 1880 y 1939. La selección de textos sigue criterios fijados por manuales de literatura. Los textos han sido codificados en XML-TEI junto con los metadatos integrados de manera manual y las anotaciones lingüísticas. El conjunto de datos permite realizar descripciones estadísticas, evaluar hipótesis propuestas por otros investigadores o explorar patrones estadísticos. En la intervención en el panel se demostrará cómo podemos extraer información de los textos sobre aspectos literarios relacionados con los lugares que aparecen en estas novelas y se demostrará la utilidad que aportan las

infraestructuras de investigación de Ciencias Sociales y Humanidades CLARIN y DARIAH para texto escrito. Las aplicaciones ofrecen desde cómo depositar el corpus de forma permanente, hasta cómo analizar el corpus en la nube. El objetivo es realizar nuevas preguntas de investigación y explorar las primeras respuestas basadas en corpus de forma dinámica y eficaz.

Enlaces

<https://www.dariah.eu/> <https://www.clarin.eu/>

<http://clarin-es.org/>

<http://ixa2.si.ehu.es/intele/?q=home>

<http://clarin-es-lab.org/index-en.html>

www.evihinhd.com <http://hitz.eus>

<http://sli.uvigo.gal>

<https://www.ucm.es/loep>

Panel 8

Historia del Arte Digital: conocimientos renovados, nuevos conocimientos

Modera: Nuria Rodríguez Ortega · *Asiste:* Brianda Otero Moreira

Panelistas: Fátima Díez Platas; Jorge Sebastián Lozano; Pedro Luengo Gutiérrez

Historia del Arte Digital: conocimientos renovados, nuevos conocimientos

Nuria Rodríguez Ortega (UMA) (coord.)

Fátima Díez Platas (USC)

Pedro Luengo Gutiérrez (US)

Jorge Sebastián Lozano (UV)

Palabras clave: Historia del Arte Digital, fotogrametría, bibliotecas digitales, web semántica, IA

El ámbito de investigación denominado «Historia del Arte Digital» cuenta ya con un amplio recorrido en la escena internacional (Baca, Helmreich and Rodríguez-Ortega, 2013; Baca, Helmreich and Gill, 2019; Brown, 2020). En España, investigadores, doctorandos y grupos de investigación se van incorporando progresivamente a este campo mediante la inclusión de tecnologías digitales y metodologías computacionales como parte de sus estudios y proyectos. A diferencia, sin embargo, de lo que acontece en otros contextos, no existe todavía en nuestro país un proceso de institucionalización fuerte o un reconocimiento amplio de las posibilidades epistemológicas y metodológicas que el desarrollado de la Historia del Arte Digital comporta. Pervive, todavía, una cierta percepción que tiende a inscribir este ámbito en el marco de las meras aplicaciones tecnológicas, aquellas que tienen capacidad para transformar los usos y las prácticas pero no las formas del pensamiento, las categorías de análisis o los modelos interpretativos.

El objetivo de este panel, es, justamente, presentar distintas líneas de indagación históricoartística en las que la incorporación de tecnologías digitales y metodologías computacionales está dando lugar a una revisión de las categorías teóricas hasta ahora utilizadas en el análisis e interpretación de los procesos y fenómenos artístico-culturales. A través de cuatro vertientes de trabajo, que lideran los participantes de este panel, se mostrará y demostrará que la Historia del Arte Digital implica una transformación no solo técnica sino también epistémica y, en consecuencia, la emergencia de nuevas y renovadas formas de conocimiento. Situada en un espacio interseccional y transfronterizo, la Historia del Arte Digital se desenvuelve a través de continuidades y discontinuidades, dando lugar, en este proceso, a transformaciones y rupturas (conceptuales, epistemológicas y metodológicas) sobre las que conviene reflexionar más allá del simple aparataje técnico.

Las cuatro vertientes de trabajo que se presentan en este panel no agotan, naturalmente, los actuales desarrollos de la Historia del Arte Digital, pero sí ofrecen una panorámica lo suficientemente demostrativa para los propósitos de este panel: agregar complejidad a la tecnología y a su empleo en la investigación histórico-artística y, por tanto, relativizar la idea que le confiere valor por su simple función de herramienta auxiliar.

Este panel forma parte de las iniciativas emprendidas por la Red de Estudios Digitales sobre la Cultura Artística (ReArte.Dix).

Intervenciones de los panelistas

1. Nuria Rodríguez Ortega. Visión por computadora y estudios visuales digitales

Aunque se trata de un campo aún muy experimental y de reciente desarrollo, la aplicación de tecnologías de inteligencia artificial para el procesamiento de extensos corpus de imágenes de la cultura artística -especialmente las tecnologías de *deep learning* asociadas a la visión por computadora- está abriendo nuevas vías de exploración en el ámbito de la Historia del Arte y de los estudios visuales en general, así como nuevos espacios para la reflexión crítica (RodríguezOrtega, 2020). En líneas generales, este nuevo paradigma de análisis cultural viene definido por dos factores: en primer lugar, la base cuantitativo-matemática propia del procesamiento computacional; y en segundo lugar, el cambio de escala, es decir, la posibilidad que se nos brinda ahora de analizar cientos de miles de imágenes en un arco temporal y espacial teóricamente global.

Un rápido recorrido por las capacidades analíticas de estas tecnologías aventura interesantes oportunidades de investigación. Así, por ejemplo, las redes neuronales artificiales pueden identificar de manera automática objetos, figuras y detalles presentes en grandes colecciones de imágenes y agrupar (o clusterizar) estas imágenes según los elementos representados. También pueden detectar y medir grados de similitud y diferencia entre ellas; calcular umbrales de complejidad visual; estimar la probabilidad estadística de que un conjunto de imágenes pertenezca a un determinado estilo, género, periodo o autor; o descubrir patrones formales y compositivos que se mantienen o se transforman a lo largo del tiempo.

La incorporación de estas tecnologías a los estudios visuales está constituyendo, por tanto, un auténtico revulsivo para replantear desde un nuevo orden de escala y complejidad algunas de las preguntas que han configurado el núcleo de las preocupaciones intelectuales de la Historia del Arte y de los estudios visuales desde sus orígenes⁷². Así, por ejemplo, la tensión entre unidad y variedad formal, la dialéctica entre persistencia y transformación visual o la continuidad y discontinuidad de los elementos semántico-simbólicos que constituyen las imágenes pueden afrontarse ahora desde nuevos parámetros. Del mismo modo, determinadas líneas de investigación pueden expandirse a partir de estos análisis computacionales, como los estudios sobre la movilidad y circulación de motivos, formas y esquemas compositivos (Joyeux-Prunel, 2019); o los procesos de reapropiación y resemantización cultural de determinadas imágenes.

Ahora bien, junto a las oportunidades de investigación que se nos brindan, la introducción de tecnologías de inteligencia artificial en el campo de la Historia del Arte y de los estudios visuales entraña importantes retos intelectuales y metodológicos que también es necesario abordar. Si nos ceñimos al ámbito de la visión por computadora, donde el concepto de visión ha de entenderse como cálculo matemático de información numérica, hemos de preguntarnos cómo se van a transformar los procesos del ver y los conceptos de visión y mirada en la era de la inteligencia artificial y cuál será el nuevo papel que deberán desempeñar los historiadores del arte en cuanto analistas e intérpretes de la cultura visual en este nuevo contexto. Téngase en cuenta que los dispositivos de visión por computadora no supervisados generan unidades, patrones, correlaciones y regiones visuales que exceden las nociones tradicionales de formas o motivos que hasta ahora habían conformado la base de la lectura e interpretación de las imágenes. Estas otras formas del ver artificial provocan, sin duda, una desestabilización de los regímenes visuales tradicionales, pero también conllevan un indudable valor heurístico, pues la manera en la que las redes neuronales artificiales «ven» puede dirigir nuestra atención hacia aspectos hasta ahora no advertidos por nuestra visión y por nuestra mirada (cultural y disciplinariamente construida).

⁷² Véase DHNord2020 colloquim *The Measurement of Images: Computational Approaches in the History and Theory of Arts*. Maison Européenne des Sciences de l'Homme et de la Société in Lill, 18-20 noviembre, 2020. En: <https://www.meshs.fr/page/dhnord2020> [consulta: 9-5-2021].

Dado este contexto, en esta primera intervención se mostrarán las líneas fundamentales por las que discurren actualmente los estudios visuales digitales y, especialmente, se incidirá en las nuevas (o renovadas) categorías de análisis e interpretación que se están proponiendo con vistas a configurar un nuevo marco epistémico, con un especial énfasis en los conceptos de «similitud» y «diferencia» visual. Para ello, se tomarán como base los proyectos que al respecto se están desarrollando en el grupo de investigación iArtHis_Lab⁷³.

2. Fátima Díez Platas. La Biblioteca Digital Ovidiana y el conocimiento iconográfico

Dentro de la Historia del Arte, en el campo de trabajo sobre la imagen que supone la iconografía, el papel de la trasposición tecnológica ha sido siempre fundamental. Para el planteamiento de las tareas con imágenes interconectadas, la reproducción de las piezas de estudio, pinturas, esculturas o grabados, ha sido el instrumento básico para abordar el trabajo con conjuntos de objetos que se convierten en imágenes y se pueden manejar aprovechando la ventaja que ofrece el estatuto de la imagen para el estudio, que ya reclamara Aristóteles en la *Poética*.

De manera concreta, para trabajar con las imágenes desde el punto de vista iconográfico en relación con temas o personajes, la composición de repertorios que reúnen, categorizan y clasifican las representaciones con criterios temáticos ha supuesto la herramienta fundamental para el trabajo con las imágenes desde el punto de vista de los contenidos, y su utilización constituye el punto de partida fundamental para establecer panoramas y estados de la cuestión que aportan el rigor necesario a los estudios iconográficos.

La digitalización y las bases de datos de imágenes han permitido la transferencia de estos valiosos repertorios al entorno virtual, y con ello se han multiplicado las posibilidades en el trabajo iconográfico. A la evidente ventaja de trabajar de manera casi simultánea con grandes cantidades de imágenes, se ha unido la posibilidad de establecer nuevas relaciones entre ellas, facilitando la comparación, la ordenación y la categorización, además de su interconexión, que es una de las vías esenciales para producir conocimiento en los estudios de Historia del Arte. La construcción de corpus de imágenes categorizadas y etiquetadas constituye, por tanto, una infraestructura crucial para el estudio de la imagen dentro del ámbito de la iconografía.

Un tipo peculiar de corpus de imágenes con categorías y etiquetas se encuentra dentro de la

Biblioteca Digital Ovidiana desarrollada en el marco del proyecto *Ovidius Pictus*⁷⁴. La Biblioteca Digital Ovidiana es una biblioteca digital de ejemplares de ediciones ilustradas sobre Ovidio, pero, funciona como una suerte de «museo» virtual de imágenes constituido por los grabados que ilustran los ejemplares, cuyas imágenes están etiquetadas por temas y personajes. De este modo, todas las imágenes que forman parte de los libros ilustrados son piezas que se pueden combinar y con las que se puede operar para crear relaciones relevantes para abordar determinadas investigaciones iconográficas. La búsqueda de imágenes que representan un determinado tema o un determinado personaje nos da como resultado un corpus de imágenes relacionadas que se presentan ordenadas cronológicamente dentro del rango de tiempo de producción de las ediciones ilustradas, de modo que de la visualización de las relaciones que se establecen entre las distintas imágenes se pueden extraer conclusiones sobre la evolución de los temas, la composición, los mecanismos de representación, los modelos de la tradición o las influencias y las divergencias en la concepción de las escenas.

Estas operaciones propias de la Historia del Arte y de la iconografía afectan a la imagen desde un punto de vista intrínseco y producen por sí mismas un conocimiento sobre los mecanismos de figuración y sobre los procesos de creación de los contenidos visuales. Sin embargo, el conocimiento que puede aportar el establecimiento de estas

⁷³ <http://iarthislab.es> [9-5-2021].

⁷⁴ <http://www.ovidiuspictus.es/> [9-5-2021].

relaciones entre las distintas imágenes de un personaje o de un mismo tema necesita enriquecerse con otras relaciones que hay que establecer fuera de la propia imagen, sea con el contexto histórico, el ámbito estilístico, el medio en el que la imagen se inserta o se produce o, incluso, en el caso concreto de las ilustraciones, el propio texto que las justifica. De este modo, el corpus de imágenes de la *Biblioteca Digital Ovidiana* se beneficia de la propia contextualización que supone el hecho de que las imágenes están individualizadas pero no están desgajadas de su soporte, el libro ilustrado, y de manera específica el ejemplar concreto de una determinada edición ilustrada, con sus características y su materialidad, que se encuentran no tanto reproducidas de modo digital, sino sobre todo detalladas y en gran medida estudiadas dentro de la propia biblioteca digital.

A lo largo de la exposición se expondrán ejemplos tomados de la *Biblioteca Digital Ovidiana* que permiten mostrar de qué modo este tipo de corpus son susceptibles de impulsar una reorientación de la investigación iconográfica.

3. Pedro Luengo. La historia de la arquitectura y del espacio arquitectónico desde nuevos parámetros

La historia de la arquitectura se ha desarrollado tradicionalmente hacia la decoración, la funcionalidad tipológica o diversos aspectos constructivos, información habitual entre la documentación de archivo, que puede complementarse fácilmente con determinados análisis formales. En los últimos años, las Humanidades Digitales han proporcionado una nueva perspectiva que permite abordar las particularidades espaciales, un aspecto que, aunque había despertado el interés de los investigadores, difícilmente se podía estudiar desde la solidez de los datos, al carecerse de las herramientas adecuadas.

El primer campo en desarrollarse de una forma destacable estas nuevas metodologías de análisis fue el de la incidencia lumínica y su utilización por parte de los arquitectos. La publicación en 2017 del manual sobre el tema por Oxford University Press, coordinado por Costas Papadopoulos y Holley Moyes, significó la consagración de estas técnicas como campo de exploración. La gestión de la luz natural, especialmente desde el ámbito de la arqueología, ha ofrecido resultados de gran interés, que progresivamente se han visto replicados en investigaciones de arquitectos e historiadores del arte. Ya más recientemente, han sido los estudios sobre dispositivos lumínicos de época antigua los que han ofrecido un paso más en el estudio de los interiores.

En paralelo se han desarrollado notablemente los estudios de simulación acústica, o lo que también se conoce como «arqueoacústica». Inicialmente, fueron los investigadores del campo de la Física los que estudiaron a través de mediciones *in situ* diferentes espacios históricos. Por su parte, los historiadores de la música también dirigieron su interés por los escenarios performativos a través de la historia. Han sido los trabajos de Dorothea Baumann (2011) los que han puesto en relación ambas líneas. Sin duda, las nuevas herramientas digitales en desarrollo permitirán un acercamiento más fluido de musicólogos e historiadores de la arquitectura al hecho acústico, facilitando un diálogo técnico interdisciplinar más profundo.

El campo de la simulación arquitectónica por medios digitales está lejos de agotarse con estas dos líneas. Diversos programas informáticos están apostando por analizar el comportamiento térmico de los interiores, en lo que supone una variable crucial para la comprensión de los edificios históricos y la experiencia de sus usuarios. Algo similar puede decirse de otros aspectos, como el análisis de estructuras, o incluso los dedicados a mecánicas de fluidos, que pueden resultar de interés para los trabajos sobre jardines o la arquitectura hidráulica, solo por citar algunos ejemplos.

No obstante, estas tecnologías tienen sus limitaciones, las más relevantes asociadas a su precaria adaptación a las circunstancias históricas. Así, por ejemplo, los materiales previstos por el *software* se limitan a los utilizados en la

actualidad, muy alejados de las realidades constructivas pasadas. Por ello, resulta clave establecer vínculos entre los desarrolladores y los humanistas digitales, promoviendo la incorporación de mediciones de materiales históricos, como ya se ha hecho con la iluminación artificial en época antigua. Este reto informático debe unirse al que propone la limitada capacidad de análisis de muchos de estos programas. Previstos para abordar espacios poligonalmente poco complejos, los modelos ofrecidos por escáneres 3D, o incluso por procedimientos fotogramétricos, resultan excesivamente ricos. Parece probable que esta limitación será solventada progresivamente con el desarrollo general de los lenguajes informáticos y la capacidad de procesamiento de los equipos, pero también mediante la incorporación progresiva de los historiadores del arte a estas investigaciones.

Por último, cabe señalar que estas limitaciones técnicas no son superiores a las dificultades de asimilación de estas metodologías de análisis por la práctica investigadora tradicional. Parece obvio que se trata de herramientas útiles para complementar tanto las discusiones ya planteadas por la historiografía como para ofrecer nuevas vías de interpretación. En cambio, su desarrollo parece haber despertado un mayor interés entre los humanistas digitales que entre los historiadores en general. Probablemente, la generalización de publicaciones que incorporan estas técnicas, sumado a la extensión de actividades formativas que acerquen los programas tanto a estudiantes como a investigadores, llevará a su asimilación progresiva en las diferentes disciplinas.

4. Jorge Sebastián Lozano. Terminología artística revisitada

La conservación y estudio del patrimonio histórico artístico requiere de la normalización terminológica. Objetos o prácticas sin una terminología asentada son más difíciles de identificar, analizar y conocer. Esto, que es válido también para entornos tradicionales o analógicos, es un principio básico en el medio digital. Sin embargo, su aplicación eficaz en nuestra realidad dista de estar tan extendida como debería. La normalización terminológica no es una tarea fácil ni entretenida, que rara vez recibe el reconocimiento que merece.

En el medio español, el Ministerio de Cultura ha desempeñado una labor encomiable en las últimas décadas que, sin estar terminada, proporciona una base más que suficiente. Los tesauros del patrimonio cultural de España están disponibles en línea, también como *Linked Open Data*⁷⁵. De esta forma, repositorios independientes pueden conectarse entre sí a través del uso de vocabularios controlados y jerarquizados. Yendo más allá de nuestro ámbito estatal, otro elemento muy positivo en este campo es la traducción española del *Art and Architecture Thesaurus* de la Getty, el estándar internacional de facto en la normalización terminológica para la historia del arte. Dicha traducción fue obra del Centro de Documentación de Bienes Patrimoniales de Chile⁵. Este magno esfuerzo nos ofrece una enorme ventaja, la posibilidad de conectar con otros idiomas cualquier descripción normalizada que emplee este vocabulario, dado que no solo existe en inglés y castellano sino en un número creciente de otras lenguas, entre las que se cuenta también el catalán. Valdría la pena plantear una revisión, aunque pequeña, de la citada traducción, que en algunos casos no incluye el castellano utilizado en la Península Ibérica. En todo caso, su disponibilidad supone un gran paso adelante, puesto que los vocabularios especializados tienen múltiples aplicaciones, entre ellas su uso para el análisis y modelizado de textos mediante algoritmos de inteligencia artificial.

Junto con la normalización, otro campo en mejora parcial pero aún insuficiente es la propia disponibilidad de datos. Los museos, y en general las instituciones del patrimonio cultural, generan una cantidad masiva de información. Sin embargo, parte muy importante de esa información se queda en la institución, no se publica ni se pone en valor en una sociedad que, como la contemporánea, está aprendiendo que la gestión de los datos tiene una posición cada vez más central. Por desgracia, muchos de nuestros museos van muy rezagados en esta área: los catálogos en línea son todavía más una excepción que la regla, aunque muchos sí que tienen sistemas digitales

⁷⁵ <http://tesauros.mecd.es/tesauros> [9-5-2021]. ⁵ <https://www.aatespanol.cl/> [9-5-2021].

de gestión interna de la colección. Y no ha sido por falta de ejemplos: una institución como el Museo Lázaro Galdiano tuvo toda su colección disponible en línea en fecha tan temprana como 2002, lo cual demuestra que no necesariamente es una cuestión de disponibilidad de muchos medios tecnológicos o humanos.

Con todo, la apuesta por un sistema de documentación compartido, como es Domus, ha permitido crear un recurso tan excepcional como CER.ES, Colecciones en Red, que sigue creciendo en número de museos participantes y de registros incorporados. No debíamos olvidar aquí otros esfuerzos pioneros, como el sistema Odiseus de Andalucía (SIPHA), o la base de datos *Museus en línia* de la Generalitat de Cataluña, construida sobre el programa DAC (Documentació Assistida de Col·leccions). Pero todavía no hay ningún gran museo español que haya apostado seriamente por los datos en abierto, algo que no es solo una cuestión tecnológica (imágenes de alta calidad, puntos SPARQL, APIs...), sino museológica, de concepción del museo. El Prado ha digitalizado y publicado todo su archivo histórico, cosa muy loable; ¿pero cuándo va a publicar en abierto los datos de su colección para que los historiadores del arte podamos aprovecharlos para la generación de conocimiento y valor?

La web semántica no solo requiere de vocabularios normalizados y de datos, sino de ontologías y modelos de datos que permitan enlazar los muchos sistemas descriptivos que empleamos alrededor de nuestros objetos de estudio. Tradicionalmente, los esfuerzos normalizadores tampoco han tenido gran implantación entre nosotros, como en general dentro de las humanidades. Ya contamos con estándares internacionales, como CIDOC-CRM⁷⁶, que proporcionan el «pegamento semántico» para conectar y relacionar entre sí repositorios o bases de datos. En este caso, el Prado sí está mostrando un liderazgo muy de agradecer con su modelo ontológico, cuya explotación les está permitiendo publicar recursos experimentales tan interesantes como la línea del tiempo de la colección. Otro fructífero campo de aplicación es la conexión de nuestros repositorios digitales semantizados con sistemas de información geográfica, como está haciendo ahora mismo el proyecto SeMap, a partir de los registros disponibles en CERES. Un recurso ya disponible, como la infraestructura Expofinder-Pathfinder⁷⁷, que compila y publica datos masivos sobre exposiciones artísticas celebradas en España, desarrollada en el marco de los proyectos Exhibitium y Andalex, ilustra los diversos usos a que podemos someter los datos digitales, por ejemplo, para la planificación de políticas culturales públicas.

Referencias bibliográficas

Baca, Murtha, Helmreich, Anne & Rodríguez-Ortega, Nuria (2013). *Digital Art History*. Special double issue of *Visual Resources. An international Journal of Documentation*, XXIX (1-2), 2013.

Baca, Murtha, Helmreich, Anne, and Gill, Melissa (2019). *Digital Art History*. Special issue of *Visual Resources*. vol. 35 (2), 2019.

Baumann, Dorothea (2011). *Music and Space: A Systematic and Historical Investigation Into the Impact of Architectural Acoustics on Performance Practice Followed by a Study of Handel's Messiah*. Berna: Peter Lang.

Brown, Kathryn (ed.) (2020). *The Routledge Companion to Digital Humanities and Art History*. New York: Routledge, 2020.

Díez Platas, Fátima (2020). «La *Biblioteca Digital Ovidiana*: un proyecto patrimonial de vocación iconográfica sobre el "Ovidio ilustrado», *Norba. Revista de Arte*, XL, pp. 21-50.

Joyeux-Prunel, Béatrice (2019). «Visual Contagions, the Art Historian, and the Digital Strategies to Work on Them», *Artl@s Bulletin*, vol. 8, n.º 3, pp. 128-144.

⁷⁶ <http://www.cidoc-crm.org/> [9-5-2021].

⁷⁷ <https://hdplus.es/pathfinder/> [9-5-2021].

Rodríguez-Ortega, Nuria (2020). «Image Processing and Computer Vision in the Field of Art History». En: Brown, K. (ed.). *The Routledge Companion to Digital Humanities and Art History*. New York: Routledge, pp. 338-357.

Papadopoulos, Costas, and Moyes, Holley (2017). *The Oxford Handbook of light in archeology*. Oxford: Oxford University Press.

Panel 9

Eventos festivos históricos y Humanidades digitales: reconstrucción virtual y estudio interdisciplinar de la entrada de Isabel de Valois en Toledo (1560)

Modera: Sagrario López Poza · *Asiste:* César González-Pérez

Panelistas: Estefanía López Salas; José Julio García Arranz; Manuel Alberto Santos Miñambres

Eventos festivos históricos y Humanidades digitales: reconstrucción virtual y estudio interdisciplinar de la entrada de Isabel de Valois en Toledo (1560)

Coordinadora: Sagrario López Poza (Universidade da Coruña. España)

Participantes:

Estefanía López Salas (Universidade da Coruña. España)

José Julio García Arranz (Universidad de Extremadura. España)

Manuel Alberto Santos Miñambres (Universidad de Extremadura. España)

Palabras clave: Arquitectura efímera, crónica festiva, recreación arquitectónica, modelado 3D, Iconografía

Descripción del problema de investigación, metodología y resultados obtenidos

Los estudios sobre las solemnidades de los siglos XVI y XVII se han desarrollado mucho en los últimos años por varias razones. Por un lado, en varios países de Europa, se han contemplado las recreaciones festivas históricas como espectáculos que proporcionan muchos beneficios económicos a las ciudades que los organizan, lo que ha impulsado la existencia de un área de investigación vinculada a ese fenómeno. Independientemente de esta tendencia, la investigación humanística se ha visto favorecida por la posibilidad de crear grupos de investigación interdisciplinarios que permiten enfoques de estudio sobre la diversidad de expresiones artísticas involucradas en tales acontecimientos festivos de forma que podamos comprender mejor tales fenómenos culturales complejos. El desarrollo de la tecnología digital ha sido otro factor importante, al permitir recrear en tres dimensiones arquitecturas efímeras, simular recorridos de una comitiva, recrear la música, los fuegos artificiales, etc. haciendo patente lo que hasta hace poco había de ser imaginado.

Un tipo de solemnidad muy especial la constituían las entradas solemnes de reinas que, casadas por poderes en su país con los reyes de la casa de Austria, llegaban a España para ratificar el matrimonio y ser presentadas y recibidas por el pueblo. Las ciudades competían para poder organizar la entrada, de lo que esperaban beneficios de diverso tipo, pero el rey solo autorizaba a algunas ese privilegio, que por otra parte suponía un gasto descomunal y un esfuerzo gigantesco de muchos ciudadanos. Para que perdurara en la memoria ese esfuerzo, los ayuntamientos solían encargar a un cronista competente la exposición por escrito de todos los detalles del acontecimiento. Esas relaciones del suceso festivo son fuentes históricas y literarias de capital importancia, pero en raras ocasiones ofrecen imágenes de lo descrito. El lector ha de imaginarse, gracias a la detallada descripción verbal, los arcos triunfales, soportales, escenarios, estatuas, fuentes y carros triunfales, así como las procesiones y atavíos de los asistentes.

Entre las maquinarias efímeras, los arcos de triunfo, aunque eran estructuras temporales de madera, lienzo y yeso, simulaban el mármol, el oro y la piedra de los monumentos permanentes. Estas decoraciones llevaban una profusión de pinturas, inscripciones, estatuas y emblemas según un programa iconográfico diseñado por un entendido en Letras Humanas, con mensajes destinados a glorificar al monarca y su estirpe, así como a la esposa; servían para transmitir un mensaje cargado de significaciones políticas y propagandísticas. Su contenido iconográfico procedía de lecturas de obras que la imprenta había conseguido divulgar y también de la observación “in situ” de solemnidades celebradas en Italia y Flandes de tantos caballeros y miembros del séquito de Carlos V o del príncipe Felipe, su hijo, en el largo viaje que este realizó entre 1648 y 1552 por los estados de su padre, el emperador, con quien debía reunirse en Bruselas para viajar luego por los Países Bajos. Juan Cristóbal Calvete de

Estrella dejó extraordinario testimonio escrito en *El felicísimo viaje...* (Amberes, 1552) y su libro, con detalladas descripciones, fue inspiración posterior para quienes tenían que idear programas iconográficos.

En 1559, Felipe II se casaba con su tercera esposa, Isabel de Valois, hija del rey Enrique II de Francia y Catalina de Médicis. Con esta unión se afianzaba la paz entre ambas casas, por lo que Isabel de Valois también fue conocida como Isabel de la Paz. La boda se había realizado por poderes en París, el 22 de junio de 1559, representando al rey Felipe el duque de Alba, y se ratificó en Guadalajara, el 31 de enero de 1560, en el Palacio del Infantado. Se celebraron varias entradas de la reina en ciudades, pero la más solemne tuvo lugar en Toledo, a donde la nueva reina llegó el 12 de febrero de 1560. La recibió una gran fiesta pública que se extendía por toda la ciudad y transformaba sus puertas de entrada, calles y plazas en espacios de un complejo espectáculo multi-sensorial, que giraba en torno a su figura protagonista: la monarquía.

Uno de los recursos utilizados para modificar el aspecto habitual de la ciudad y crear una imagen ideal de la misma, en las entradas reales del siglo XVI, fue el diseño y construcción de arquitecturas efímeras. Estas adquirirían principalmente la forma de arcos de triunfo que se disponían en determinados puntos de un recorrido urbano planificado. Los monarcas avanzaban por la ciudad, a la manera de emperadores romanos, y su paso y parada en los espacios distinguidos por las arquitecturas efímeras, simbolizaba, entre otras cuestiones, la renovación de la lealtad de la ciudad que los acogía, sus instituciones y habitantes a la Corona (Ferrer Valls, 2003: 31).

En las comunicaciones que integran este panel se plantea mostrar el trabajo desarrollado para la reconstrucción digital de los arcos triunfales diseñados por el arquitecto Nicolás de Vergara, el Mozo, para la entrada real de Isabel de Valois en Toledo, tomando como punto de partida la relación de sucesos más importante y completa que se conserva de este acontecimiento, la de Álvaro Gómez de Castro, publicada en 1561 (Fernández Travieso, 2007; Torre Fazio, 2012)⁷⁸. Se trata de un conjunto de cinco aparatos cuyos nombres estaban directamente vinculados a su posición dentro de la trama urbana y que, desde el punto de entrada a la ciudad hasta el fin del itinerario real, eran: el arco levantado en las inmediaciones de la puerta de la Bisagra, el arco a la puerta de la segunda muralla y junto a la Herrería, el arco localizado por encima de la segunda muralla de la ciudad, desde donde se veía el palacio de Galiana y la huerta del rey, el arco a la puerta del Perdón de la catedral y el arco situado a la entrada de la plaza Zocodover.

No conservamos representaciones gráficas que nos permitan comprender en su totalidad el ambiente de la fiesta que tuvo lugar como consecuencia de esta entrada real. Tampoco han sido localizadas, hasta la fecha, ilustraciones, dibujos, pinturas o grabados que hoy pudieran ser memoria visual de las construcciones efímeras entonces levantadas. No obstante, un rasgo distintivo de la relación de sucesos de Álvaro Gómez de Castro, dentro de la literatura vinculada a la fiesta pública desarrollada durante el Siglo de Oro, es que ofrece descripciones minuciosas de cada aparato que, por el tipo y número de datos recolectados, en algunos casos son casi memorias técnicas, dibujadas con palabras, de las arquitecturas efímeras ejecutadas para la entrada real. Como consecuencia, estas descripciones pormenorizadas permiten una reconstrucción gráfica de las trazas arquitectónicas de cada aparato festivo, siempre que estas se combinen con el conocimiento de: las leyes propias de la gramática clásica de la arquitectura, la realidad de la trama urbana histórica que las acogió y los modelos de su misma tipología hechos antes o en los años posteriores inmediatos a la entrada de Isabel de Valois en Toledo.

En otras palabras, a través de un determinado flujo de trabajo y método de visualización digital computarizada es posible recuperar, con rigor académico en el proceso y en los resultados, la imagen perdida de la arquitectura de los arcos de triunfo citados. Este es un primer paso para descifrar el mensaje, eminentemente visual, de la

⁷⁸ Este trabajo se enmarca dentro del Proyecto I+D+i 'Biblioteca Digital Siglo de Oro 6 (BIDISO 6), PID2019105673GB-I00, concedido por el Ministerio de Ciencia e Innovación para el periodo 2020-2023, IP: Nieves Pena Sueiro.

compleja red de espectáculos de la que formaron parte que, al mismo tiempo, abre la puerta hacia nuevas interpretaciones de la fiesta pública del siglo XVI, como acontecimiento histórico, artístico y cultural.

Aparte de la reconstrucción arquitectónica de los arcos, ofrecemos también la identificación de las imágenes que decoran los arcos –fundamentalmente personificaciones, alegorías, emblemas y empresas, retratos y elementos heráldicos– de las que no se conserva documentación visual ni de su estructura ni de sus detalles. A través de las descripciones textuales de Álvaro Gómez de Castro y teniendo muy en cuenta algunas construcciones coetáneas de estas características cuyo aspecto sí nos es conocido, se ha llevado a cabo una labor de rastreo de aquellas figuras o composiciones que mejor se pudieran adaptar a los detalles ornamentales y simbólicos descritos en la referida crónica, considerando al mismo tiempo los aspectos temáticos –identificación del personaje o de la composición–, tipológicos –si se trata de esculturas o pinturas–, significativos y cronológicos, y buscando la mayor proximidad con el elemento requerido a fin de obtener una recreación lo más rigurosa posible. Se expondrán los principales manuales y fuentes utilizados para esta labor, los recursos disponibles en red, así como las principales dificultades que hemos tenido que afrontar a la hora de adaptar las figuras encontradas a la temática requerida –sobre todo en el caso de composiciones con varios personajes o elementos–, a las características formales, o al diseño del lugar –soportes, enjutas, frisos, coronamientos...– en el que se insertaron originalmente. Aplicando a estas imágenes los adecuados tratamientos digitales para su más precisa adaptación y ajuste, el objetivo final es el de recrear de una manera lo más fidedigna y detallada posible, no solo las estructuras arquitectónicas, sino también los programas ornamentales y simbólicos que las complementaban, en aras de obtener una nueva percepción y comprensión global del elemento festivo que propicie nuevas aproximaciones e interpretaciones de este fenómeno.

Descripción de las intervenciones de cada participante

Sagrario López Poza, como coordinadora del panel, presentará un estado de los estudios sobre cultura festiva histórica y los avances conseguidos en los últimos años, propiciados por el empleo de herramientas y métodos derivados de la tecnología digital, en que se incardina el trabajo interdisciplinar del grupo SIELAE-Hispania para el proyecto *Biblioteca Digital Siglo de Oro 6*.

Estefanía López Salas realizará la presentación: La entrada en Toledo de la reina Isabel de Valois: (re)construcción digital de arcos triunfales a partir de una relación de sucesos, en que expondrá el trabajo realizado como arquitecta para el proyecto *Biblioteca Digital Siglo de Oro*

6.

José Julio García Arranz y Manuel Alberto Santos Miñambres se ocuparán de la presentación: Recreando imaginarios festivos: la reconstrucción icónica de maquinarias efímeras en la entrada en Toledo de la reina Isabel de Valois (1560), que tratará de los procedimientos metodológicos aplicados al rastreo, localización y manipulación de las diversas figuras e imágenes que adornaron las arquitecturas efímeras analizadas.

Bibliografía

Gómez de Castro, Alvar. 1561. *Recebimiento que la imperial ciudad de Toledo hizo a la majestad de la reina nuestra señora doña Isabel, hija del rey Enrique II de Francia, cuando nuevamente entró en ella a celebrar las fiestas de sus felicísimas bodas con el rey don Filipe N. Sor II deste nombre (Toledo, 1561)*. Estudio y edición de Carlota Fernández Travieso. A Coruña: SIELAE & Sociedad de Cultura Valle-Inclán, 2007.

Gómez de Castro, Alvar. 1561. *Recebimiento que la imperial ciudad de Toledo hizo a la Magestad de la Reyna Nuestra Señora Doña Ysabel*. Edición facsímil de Julia de la Torre Fazio. Málaga: Servicio de Publicaciones e Intercambio Científico de la Universidad de Málaga, 2012.

- Ferrer Valls, Teresa. 2003. "La fiesta en el Siglo de Oro: en los márgenes de la ilusión teatral". En: *Teatro y fiesta del Siglo de Oro en tierras europeas de los Austrias*, 27-37. Madrid: SEACEX.
- Redondo, Augustin. 1997. "Fiesta, realeza y ciudad: Las relaciones de las fiestas toledanas de 1559-1560 vinculadas al casamiento de Felipe II con Isabel de Valois". En: S. López Poza y N. Pena Sueiro (eds.). *La fiesta. Actas del II Seminario de Relaciones de Sucesos*, 303313. Ferrol: Sociedad de Cultura Valle Inclán.
- Fernández Travieso, Carlota. 2006. "La cultura emblemática en la entrada en Toledo de Isabel de Valois en 1560". En: Anthony Close (ed.), con la colaboración de Sandra M.^a Fernández Vales, *Edad de Oro Cantabrigense. Actas del VII Congreso de la Asociación Internacional del Siglo de Oro (AISO). Robinson College, Cambridge, 18-22 de junio, 2005*, pp. 245-254. Madrid / Frankfurt, Iberoamericana / Asociación Internacional del Siglo de Oro (AISO).
- Fernández Travieso, Carlota. 2011. "Reutilización de elementos del programa iconográfico creado para el recibimiento a Isabel de Valois en Toledo en 1560". En: Rafael Zafra y José Javier Azanza, *Emblemática trascendente. Hermenéutica de la imagen, iconología del texto*, pp. 301-310. Pamplona, SEE-Universidad de Navarra.
- Torre Fazio, Julia de la. 2015. "The Grand Entry of Elizabeth of Valois into Spain (1559)". En *Ceremonial Entries in Early Modern Europe. The Iconography of Power*, J.R. Mulryne, M^a Inés Aliverti y Anna M^a Testaverde (eds.), pp. 149-166. Ashgate Publishing, Farnham Surrey.
- "Una relación de la entrada de Isabel de Valois en tierras españolas en la Colección Granvela", *Avisos. Noticias de la Real Biblioteca*, Año XI, núm. 47 (octubre - diciembre, 2006):
<<http://avisos.realbiblioteca.es/?p=article&aviso=55&art=912>>
- Horozco, Sebastián. 1981. *Relaciones históricas toledanas*, introducción y transcripción de Jack Weiner, Toledo, Instituto Provincial de Investigaciones y Estudios Toledanos.
- López Poza, Sagrario. 1996. "Emblemas españoles manuscritos en Toledo en 1562". En: Sagrario López Poza, (ed.), *Literatura emblemática hispánica. Actas del I Simposio Internacional (La Coruña, 14-17 de septiembre, 1994)*, pp. 129-174. La Coruña, Universidade da Coruña.

Sesión 10

La vía de la definición

Modelos, ontologías y simulaciones

Modera: María Luisa Díez Platas · *Asiste:* Alexandra Iglesias Cortés

- 17 De alfares y alfareros: Gestión de fuentes múltiples y modelado de datos para el análisis territorial de la producción de cerámica
Esther Travé Allepuz, Sonia Medina Gordo, Joan Vicens Tarré, Alfred Mauri Martí

- 20 Identificando el sujeto del paisaje: Modelado e integración de la dimensión social para una (re)lectura de la geografía altomedieval
Sonia Medina Gordo, Esther Travé Allepuz

De alfares y alfareros: gestión de fuentes múltiples y modelado de datos para el análisis territorial de la producción de cerámica

Esther Travé Allepuz^{1,2}, Sonia Medina Gordo^{1,2}, Joan Vicens Tarré^{2,3}, Alfred Mauri Martí^{2,4}

1. Universitat de Barcelona
2. GRAMP.UB – Grup de Recerca d'Arqueologia Medieval i Postmedieval
3. Universitat de Girona
4. Centre d'Estudis Martorellencs

Palabras clave: Modelado de datos, Sistemas de Gestión, Cerámica, Unidad Topográfica, Actor

Descripción del problema de investigación, metodología y resultados obtenidos

El estudio arqueológico de la cerámica medieval y moderna ha experimentado un avance significativo en las últimas décadas. Las razones para ello debemos buscarlas en la consolidación de la Arqueología Medieval como disciplina y el interés por la materialidad de los periodos más recientes, que cristaliza en el desarrollo de proyectos arqueológicos de época postmedieval, moderna y contemporánea. A ello se añade la consolidación de la arqueometría como metodología habitual también en estos periodos más tardíos.

En la actualidad desarrollamos un proyecto de estudio de las producciones alfareras de cocción reductora, las denominadas cerámicas grises u obra negra, centrado en el análisis de los procesos de transformación formal, tecnológica y comercial que operan durante el tránsito entre la Edad Media tardía y el periodo moderno. En este contexto, contamos con una gran multiplicidad de fuentes, de índole diversa, que permiten abordar el estudio de los procesos de transferencia tecnológica y de producción y distribución de cerámicas utilitarias desde la perspectiva que ofrece el análisis del territorio y profundizar por esta vía en cuestiones relativas a la selección de materias primas, distribución de productos, o transferencia de prácticas tecnológicas.

Para el desarrollo de la investigación, por tanto, resulta imprescindible definir la estructura de la información que manejamos en clave ontológica y generar un sistema de gestión de la información en base a un refinado gradual de modelos de datos (González-Pérez 2018). Las Humanidades Digitales ponen a nuestro alcance las herramientas necesarias para interpretar la construcción de los paisajes alfareros medievales y modernos desde una perspectiva territorial e interdisciplinar, en donde las fuentes de materia prima, el producto, los circuitos de distribución o los actores que forman parte de todo el proceso sean los protagonistas del relato histórico. Esta comunicación pretende abordar las bases para hacer posible este estudio de los materiales arqueológicos y su caracterización arqueométrica, a la luz de las fuentes documentales y bibliográficas; y presentar la estructura inicial de un sistema de gestión de la información diseñado con el objetivo de hacer posible la interdisciplinariedad y trazabilidad del proceso de investigación.

J. Huggett (2020) ha advertido en un trabajo reciente acerca del surgimiento de un nuevo paradigma de gestión de la información en el cual la interpretación de la información desde la perspectiva hipotético-deductiva parece quedar sustituida por un análisis en donde el dato, la información, cobran relevancia por sí mismos en un énfasis de la cuantificación y la fe ciega en la precisión de los conjuntos de datos que a menudo comprometen la

necesaria validación o valoración –imprescindibles en arqueología– de la fiabilidad o completitud de nuestras fuentes de información.

En un esfuerzo por buscar la transparencia en la gestión de la información, la integración de los registros que emanan de las distintas disciplinas vinculadas al análisis ceramológico en arqueología y normalización metodológica, presentamos un sistema de gestión de la información que aspira a definir una estructura de carácter ontológico. En los últimos años, nuestro equipo ha desarrollado una aproximación de carácter semántico al conocimiento histórico y diseñado modelos de datos que hemos aplicado a distintos proyectos de investigación, deviniendo así el fundamento de sistemas en transformación adaptables a las necesidades de los usuarios y, aun así, interoperables.

El sistema de gestión GREYWARE que presentamos ha sido diseñado para la ejecución del proyecto de investigación homónimo, con la aspiración de constituir una plataforma válida para el análisis de las producciones cerámicas desde una perspectiva interdisciplinar, con independencia de la época o funcionalidad, que permita interpretar de manera integrada la información procedente de múltiples fuentes. Nuestra propuesta inicial se inspira en proyectos de estandarización de datos, como PARTHENOS (Krawer 2019) y su aplicación posterior ARIADNE (Meghini et al. 2017), y tomó forma en trabajos posteriores a los que nos remitimos (Travé et al. 2020; 2021).

La digitalización de la información relativa a los conjuntos cerámicos para su explotación en clave histórica o arqueológica plantea múltiples problemas derivados de la necesidad de formalización del registro, habitualmente generado en múltiples soportes. Pese a ello, ha contado con algunas experiencias notables en el desarrollo de las Humanidades Digitales en nuestro campo de estudio. Nos referimos por ejemplo al lanzamiento de una base de datos en línea como PETRODATABASE (Quinn et al. 2011) para compartir material relativo a los estudios de cerámica en el campo de la petrografía y, más recientemente, a una aplicación móvil para la clasificación automática de cerámicas arqueológicas como ARCHAIDE (Gattiglia 2018; Wright & Gattiglia, 2018). En cualquier caso, carecemos de un sistema de información integral que permita formalizar todo el proceso de análisis y gestión del registro específico en este campo de estudio que garantice la trazabilidad del proceso y la incorporación de fuentes de información de origen múltiple y naturaleza diversa.

El sistema GREYWARE que proponemos se fundamenta en los principios de la archivística y la gestión documental y el cuadro de clasificación (Figura 1) del mismo da respuesta –en el campo específico de la ceramología– a algunas de las propuestas metodológicas del International Research on Permanent Authentic Records in Electronic Systems (InterPARES) (Gänser & Michetti, 2018). El desarrollo del proyecto contempla la definición de una ontología para el estudio de la cerámica arqueológica que permita superar los estreñimientos propios de los sistemas de información definidos en base a proyectos de investigación concretos, es decir, su caducidad al finalizar el proyecto en cuestión y la imposibilidad de integrar dicha información y estructura de datos en proyectos posteriores (Doerr 2009).

GREYWARE ha sido diseñado para la gestión de la información generada en el estudio acerca de la evolución tecnológica y funcional de las cerámicas de cocción reductora en el tránsito entre la Edad Media y la Moderna, pero su diseño y funcionalidad es extensible a cualquier tipo de estudio cerámico con independencia de la datación o asignación de los materiales que son objeto de investigación. Entre sus componentes principales, GREYWARE utiliza distintas aplicaciones de software comercial y/o libre, y cuenta con una base de datos creada para el desarrollo del proyecto con distintas interfaces de usuario en función del estadio de gestión –primario o secundario– de la información.

La definición de los procesos metodológicos que precisa el desarrollo de la investigación sirve de guía a través del sistema que se estructura en actividades, series y unidades documentales compuestas o simples, siguiendo el

Estándar Internacional de Descripción Archivística (Matas & Rufi 2001) e identifica las fases, los procesos, y la documentación generada o derivada de cada uno de ellos. Con independencia del vocabulario específico en función de la disciplina concreta que lo articula y del tipo de análisis –macroscópico o microscópico–, el sistema permite un cierto grado de interoperabilidad e interdisciplinariedad en base a la utilización de unas categorías mínimas comunes y del vocabulario propio de la ceramología a través de tesauros normalizados generados a partir de consensos mínimos dentro de la disciplina (Rice 1987, Quinn 2013, Whitbread 1989 & 2001).

Esta forma de gestionar la información facilita la investigación, agiliza los procesos de análisis y –lo que es más importante– asegura la trazabilidad del proceso de gestión de datos en cada una de las fases de investigación. El trabajo colaborativo e interdisciplinar deviene posible en una estructura organizativa compartida a partir del cuadro de clasificación, del que participan los distintos investigadores implicados en el proyecto. En la medida en que la introducción y manejo de la información se ajusta a dicha estructura, cada uno de los documentos generados, con independencia del momento de creación y su finalidad, es fácilmente localizable y accesible. Los archivos son clasificados mediante la introducción de metadatos normalizados que incluyen información genérica como la identificación del yacimiento, de la muestra en cuestión, la autoría de la documentación generada, el origen y formato de la misma, el contexto arqueológico y el contexto histórico interpretativo (ambos expresados como Unidades Topográficas), así como otras informaciones más precisas según el tipo de fuente, como por ejemplo la magnificación o iluminación en microfotografías. La información relativa a la gestión secundaria de la información –como la asignación a grupos cerámicos concretos generados a partir de distintas variables de clasificación– también forma parte de este cuerpo de metadatos.

Un elemento clave en el desarrollo del sistema ha sido la creación de una base de datos acorde a las necesidades del proyecto, pero fundamentada en un modelo de datos preexistente (Mauri 2006, Travé et al. 2021) que utiliza los conceptos de Unidad Topográfica, Unidad Estratigráfica y Actor como categorías elementales en la gestión del registro y delimita los requisitos mínimos para que se ajuste a dicho modelo. En su versión actual, la base de datos de GREYWARE define una estructura multinivel de introducción y explotación de datos con un modelo de datos (Figura 2) que se ajusta a una escala variable en función del grado de precisión requerido en el análisis o a la fase en ejecución dentro del proyecto. Permite el registro y explotación del inventario de muestras investigadas, su documentación gráfica y fotográfica, su descripción y contexto, así como la información obtenida de las fuentes escritas y bibliográficas o de la experimentación y etnografía en relación con las áreas y centros de producción del territorio en estudio.

Esta base de datos la conforman diez tablas relacionadas, cada una con su interfaz propia. Algunas interfaces de visualización adicionales permiten mostrar el resultado de búsquedas automatizadas a través de la ejecución de guiones y plantillas de informe o impresión. El gráfico UML (Figura 3) que presentamos resume los elementos fundamentales de su arquitectura. Esta aproximación multinivel se ajusta a dos planteamientos ontológicos distintos: por un lado, una estructura de carácter relacional basada en el modelo de datos definido por Mauri (2006) y las tablas relacionadas que derivan de dicho modelo, tal como hemos sugerido en trabajos anteriores (Travé et al. 2021: 52); y, por otro lado, una estructura centrada en el objeto (cerámico en nuestro caso). Estos elementos constituyen los requisitos mínimos para posibilitar la gestión de la información desde una perspectiva interdisciplinar e integrada para la generación del conocimiento histórico. En el primer nivel de modelado, que relaciona entidades y acontecimientos, las categorías de Unidad Topográfica y Actor son las unidades mínimas de información que articulan el sistema, mientras que en el segundo nivel la categoría Muestra de Cerámica es la unidad mínima de información que relaciona todos los componentes y que articula los elementos descriptivos, gráficos y fotográficos del estudio de materiales cerámicos.

Una categoría importante en este contexto es la Unidad Estratigráfica, que, como tal, define el contexto arqueológico de las muestras en la mayoría de los casos y constituye un concepto de uso habitual en arqueología.

La dialéctica entre Unidad Topográfica y Unidad Estratigráfica ha sido analizada y discutida en trabajos anteriores (Travé et al. 2020: 13-14) y permite una comparación eficaz entre las acciones del pasado en función de su materialidad o la ausencia de ella. En este sentido, el contexto de las muestras que no tienen un origen arqueológico (materiales de origen etnográfico, experimental, o procedentes de colecciones de museo) se formaliza a través de su asignación a una Unidad Topográfica. La información contextual de las muestras procedentes de excavaciones arqueológicas está poco desarrollada en esta base de datos, porque GREYWARE como sistema de información se nutre de otros sistemas específicos como los diseñados para la gestión del registro arqueológico (representado en rojo en el modelo UML de la Figura 3). Por ello, resulta redundante incorporar duplicidades en la base de datos más allá de consignar la Unidad Estratigráfica a la que se asignan los materiales y una descripción breve.

Precisamente porque un fragmento de cerámica aislado no explica nada por sí mismo, agrupar las muestras en función de distintas variables (tipología, composición de la pasta, grupos de proveniencia, etc.) es un paso imprescindible en el análisis de materiales y su interpretación en términos antrópicos. La consistencia y significación de la clasificación final debe ser evaluada (Whittaker et al. 1998), a fin de definir el carácter explicativo de los grupos en términos sociales, funcionales, cronológicos o tecnológicos relativos a la producción y utilización de cerámica en el pasado (Chang 1967: 68). Los grupos verdaderamente significativos son entonces considerados como producciones, e integrados o no en áreas de producción más amplias, prácticas tecnológicas o circuitos de distribución e intercambio entre distintas comunidades. Todos estos elementos se corresponden con la definición de Unidad Topográfica y son introducidos como tales en la base de datos.

En paralelo, la información relativa a los talleres e instalaciones, los alfareros conocidos, las comunidades que adquieren y utilizan estos productos, la estructura gremial o el comercio y múltiples datos que podemos conocer a través de las fuentes escritas, iconográficas o etnográficas, constituyen un elemento valioso para la interpretación del registro material. Por todo ello, las categorías de Unidad Topográfica y Actor permiten una explotación conjunta a este nivel, por lo que las posibilidades de gestionar la información desde una perspectiva verdaderamente interdisciplinar se multiplican. Identificar los principales elementos de la producción cerámica y las transformaciones que operan en esta actividad económica y considerarlos de manera normalizada como Unidades Topográficas permite asignarles una localización precisa en términos espaciotemporales, por lo que facilita una explotación desde la óptica territorial o de la construcción del paisaje (Figura 4), elemento que aún hoy ha sido poco desarrollado en el ámbito de la ceramología.

Desarrollar estrategias para garantizar un registro documental adecuado es un elemento crucial en arqueología (Del Fresno 2016; Bordoni 2013), puesto que el método de análisis es destructivo; apreciación que también aplica a las analíticas de materiales que pueden destruirlos o alterar su materialidad en parte. En este sentido, GREYWARE es un proyecto actualmente en su fase inicial, pero con expectativas de futuro que aspira a normalizar la gestión de la información relativa a los materiales arqueológicos y su interpretación histórica, incluyendo para ello las propuestas existentes.

Figuras, tablas, mapas o gráficos

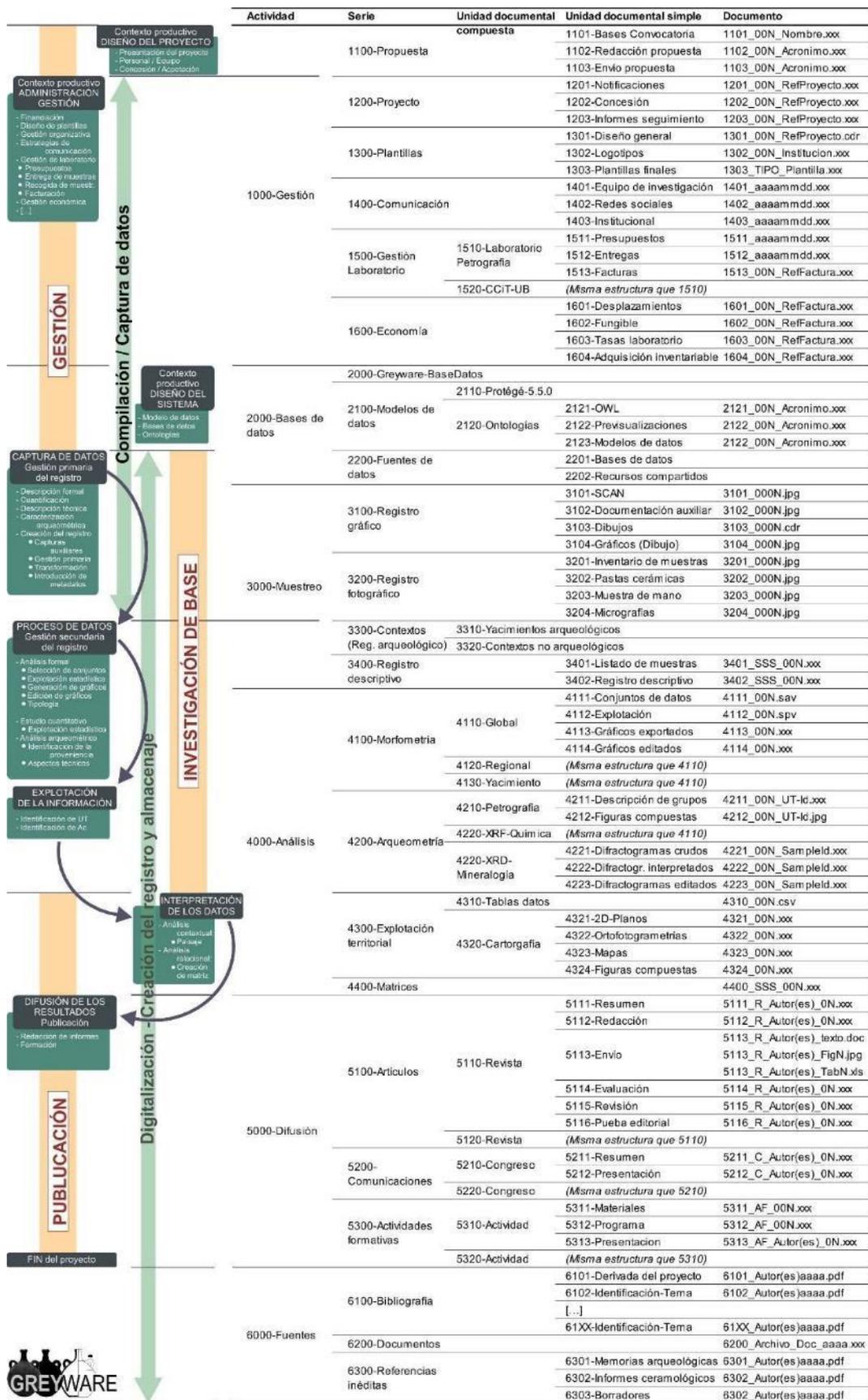


Figura 1: Cuadro de clasificación del sistema

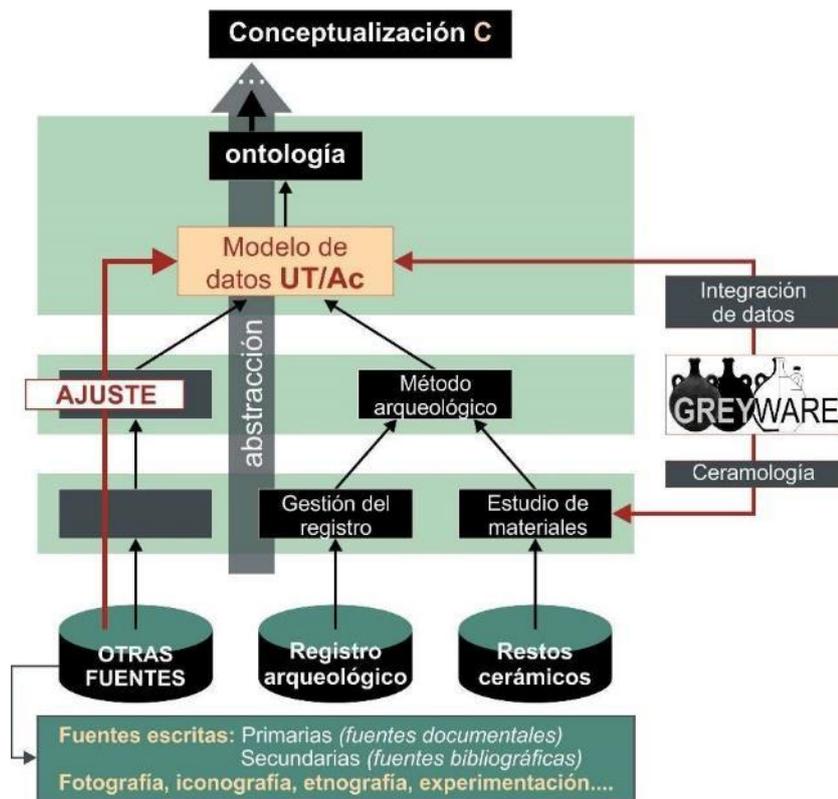


Figura 2: Ajuste multinivel del sistema a distintos modelos de datos.

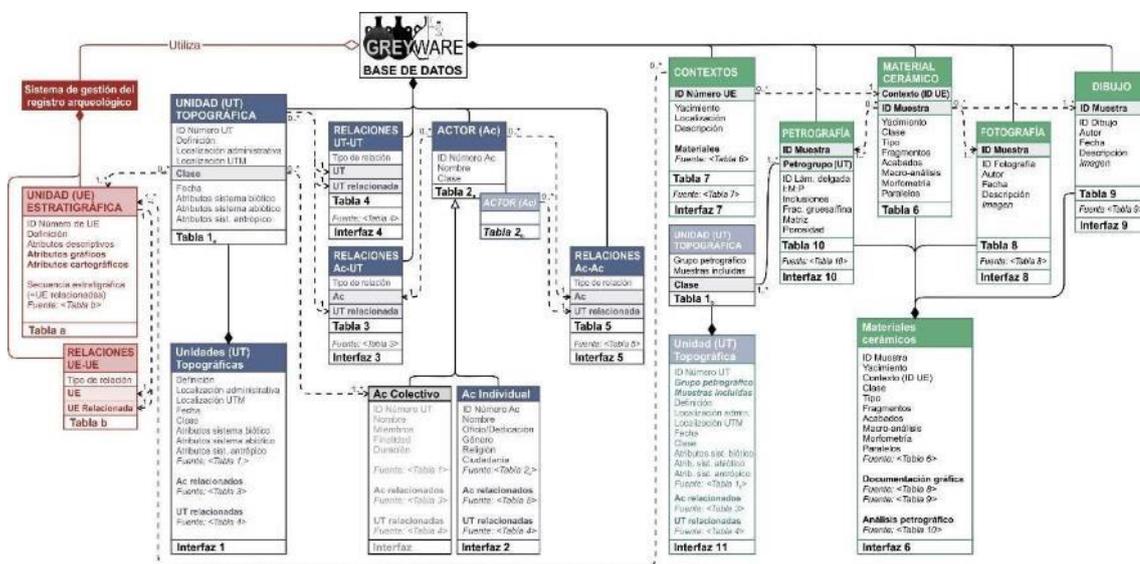


Figura 3: Diagrama UML del modelo de datos de la base de datos GREYWARE

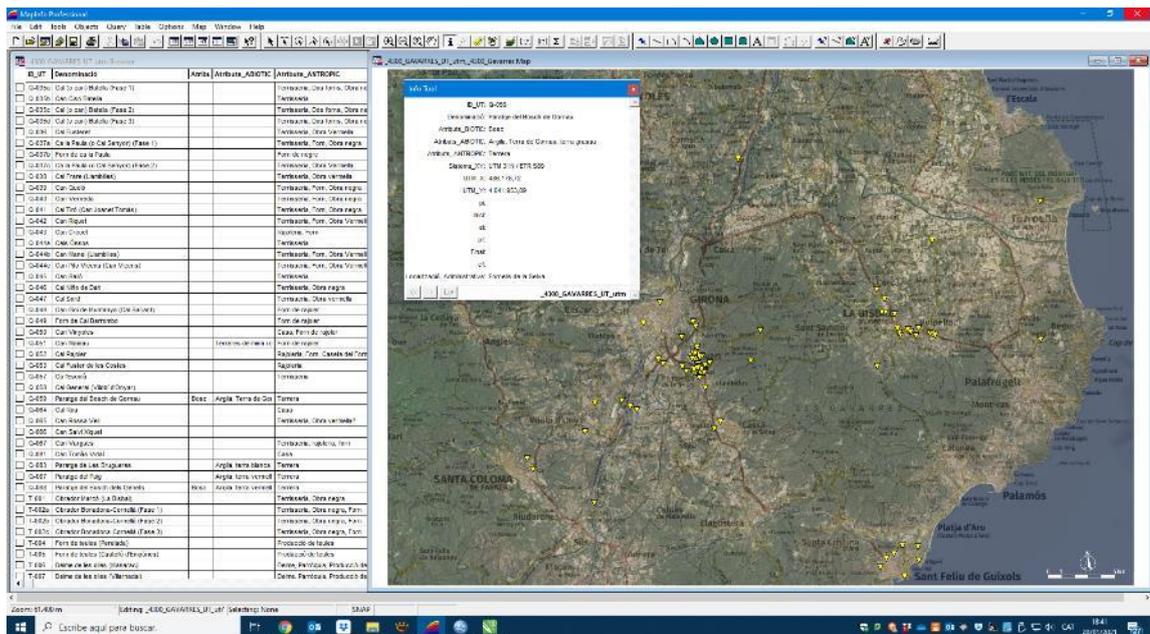


Figura 4: Captura de pantalla de la explotación territorial de la información mediante sistemas de información geográfica.

Bibliografía

Bordoni, L.; Costantini, M.; Curcio, A. (2013): A case study in archaeological documentation with ontological modelling. In CHNT 18 Proceedings. International Conference on Cultural Heritage and New Technologies, Viena.

Chang, K.C. (1967): Rethinking Archaeology. Random House: New York.

Del Fresno, P. (2016) Sistema de Información Arqueológica: Propuesta de Normalización, Desarrollo Conceptual e Informático. Tesis doctoral, Universidad del País Vasco, Vitoria.

Doerr, M. (2009): Ontologies for Cultural Heritage. In Handbook on Ontologies, 1st ed.; International Handbooks on Information Systems Book Series (INFOSYS); Springer: Berlin / Heidelberg; pp. 463–486.

Gänsler, G.; Michetti, G. (2018): InterPARES Trust Project. Ontology of Functional Activities for Archival Systems. TR05-Final Report-20180526. 2018. Disponible: <https://interparestrust.org>.

Gattiglia, G. (2018) : Classificare le ceramiche: dai metodi tradizionali all'intelligenza artificiale. L'esperienza del progetto ArchAIDE. In Archeologia, Quo vadis? Riessoni metodologiche sul futuro di una disciplina. Atti del workshop internazionale, Catania 1819 Gennaio 2018. 1st ed.; Monografie IBAM 14, CNR: Catania, Italy.

González-Pérez, C. (2018): Information Modelling for Archaeology and Anthropology. Software Engineering Principles for Cultural Heritage, Springer International.

Hugget, J. (2020): Is Big Digital Data Different? Towards a New Archaeological Paradigm. Journal of Field Archaeology, 45 (S1), S8–S17.

Krauer, S. (2019): The Parthenos Project. In Proceedings of the Parthenos CEE Workshop, Sofia, Bulgaria, 7–9. October 2019.

Matas, J.; Ruff, J. (trad.) (2001): Norma Internacional General de Descripción Arxivística. [International Council on Archives. ISAD(G) General International Standard Archival Description]. Associació d'Arxivers de Catalunya / Generalitat de Catalunya Departament de Cultura: Barcelona.

Mauri, A. (2006): La Configuració del Paisatge Medieval: El Comtat de Barcelona Fins al Segle XI. Tesis doctoral. Universitat de Barcelona, Barcelona.

Meghini, C.; Scopigno, R.; Richards, J.; Wright, H.; Geser, G.; Cuy, S.; Fihn, J.; Fanini, B.; Hollander, H.; Niccolucci, F.; et al. (2017): ARIADNE: A research infrastructure for archaeology. *ACM J. Comput. Cult. Herit.*, 10, 1–27.

Musen, M.A. (2015) The Protégé project: A look back and a look forward. *AI Matters*, Association of Computing Machinery Specific Interest Group in Artificial Intelligence, 1 (4), DOI: 10.1145/2557001.2575003.

Quinn, P.S. (2013): *Ceramic Petrography: The Interpretation of Archaeological Pottery and Related Artefacts in Thin Section*, Archaeopress: Oxford.

Quinn, P.S.; Rout, D.; Stringer, L.; Alexander, T.; Armstrong, A.; Olmstead, S. (2011): Petrodatabase: An on-line database for thin section ceramic petrography. *Journal of Archaeological Science*, 38 (9), 2491–2496.

Rice, P.M. (1987): *Pottery analysis: a sourcebook*. University of Chicago Press: Chicago, 1987.

Travé, E.; Del Fresno, P.; Mauri, A. (2020): Ontology-mediated Historical Data Modelling: Theoretical and Practical Tools for an Integrated Construction of the Past, *Digital Humanities Special Issue. Information*, 11, 182.

Travé, E.; Del Fresno, P.; Mauri, A.; Medina, S. (2021): The Semantics of History. *Interdisciplinary Categories and Methods for Digital Historical Research*, *International Journal of Interactive Multimedia and Artificial Intelligence*, 6 (5), 47–56.

Whitbread, I.K. (1989) A proposal for the systematic description of thin sections towards the study of ancient ceramic technology. In *Archaeometry. Proceedings of the 25th international symposium*, 1st ed.; Maniatis, Y. Ed.; Amsterdam, 1989, pp. 127-138.

Whitbread, I. K. (2001): *Ceramic Petrology, Clay Geochemistry and Ceramic Production. From Technology to the Mind of the Potter*. In *Handbook of Archaeological Sciences*; Brothwell, D.R.; Pollard, A.M.; Eds.; Michigan, pp. 449-458.

Whittaker, J.C.; Caulkins, D.; Kamp, K.A. (1998): Evaluating Consistency in Typology and Classification. *Journal of Archaeological Method and Theory*, 5 (2), 129–164.

Wright, H.; Gattiglia, G. (2018): ArchAIDE: Archaeological Automatic Interpretation and Documentation of cEramics. In *Proceedings of the Workshop on Cultural Informatics research and Applications co-located with the International Conference on Digital Heritage*. 1st ed.; Euromed: Nicosia, Chipre, 2018; pp. 60–65.

Identificando el sujeto del paisaje: modelado e integración de la dimensión social para una (re)lectura de la geografía altomedieval

Sonia Medina Gordo, Esther Travé Allepuz

Universitat de Barcelona

GRAMP.UB – Grup de Recerca d'Arqueologia Medieval i Postmedieval

Palabras clave: Modelo de datos semántico, Alta Edad Media, Actor, Rol, Unidad Topográfica

Los estudios de la Tardoantigüedad y la Alta Edad Media han puesto de relieve todo un conjunto de elementos que caracterizan el paisaje y articulan el territorio bajo unas dinámicas de poblamiento diferentes de etapas precedentes. A fin de procurar un conocimiento más ajustado y preciso del periodo, suele ser habitual (y necesario) un acercamiento de las fuentes escritas, poco abundantes en comparación con otros periodos, con el registro material, altamente fragmentado y recuperado de contextos territoriales heterogéneos. En esta aproximación, sin embargo, rara vez se lleva a cabo una reflexión acerca de los conceptos y definiciones que permiten una estrategia de comunicación entre datos verdaderamente cruzada y compartida.

La caracterización de estos nuevos espacios en formación puede ser examinada desde perspectivas de estudio diversas, como las distintas formas ocupación, la administración de bienes, o a través de los sujetos protagonistas de toda la actividad que se está desarrollando. Aquí, nuestra labor como investigadores debería capacitar su identificación e interrelación dentro de un entorno de trabajo colaborativo y común, posibilitando una gestión fluida de la información. En este sentido, por tanto, y dada la naturaleza de cada uno de los registros a disposición, conviene llevar a cabo una reflexión en torno al modelo de datos que nos debe permitir una administración integrada y multinivel de los mismos.

Atendiendo tales necesidades, nuestra comunicación pretende presentar algunas propuestas respecto a la construcción de este modelo, incidiendo de manera particular en la dimensión social que participa del paisaje y la manera en que ésta puede ser conceptualizada e integrada. Como estudio de caso, presentamos el marco de trabajo de un proyecto en curso que considera la incorporación del territorio del Alto Arlanza (Burgos, España) a las estructuras del sistema feudal castellano a través de una amplia variedad de fuentes.

Uno de nuestros propósitos a nivel interpretativo es prestar atención a los cambios que propicia y experimenta la sociedad por medio de relaciones de todo tipo (sociales, políticas o de poder, económicas, jurídicas), que a su vez han quedado reflejadas sobre el territorio de diferentes maneras. Para su estudio, es necesario concretar qué datos (y, por tanto, qué variables) son los que nos deben ayudar en la resolución y consolidación de esta finalidad dentro de la investigación, sirviéndonos para ello de herramientas digitales que puedan mejorar y facilitar el proceso. De hecho, recientemente se ha comenzado a incorporar la lectura de esta etapa temprana de la Edad Media por medio del registro, cuantificación y análisis de una gran cantidad de datos (Martín Viso, 2020; Carvajal y Escalona, 2020). No obstante, todavía no se ha conseguido integrar un modo de trabajar toda esta información desde acercamientos verdaderamente intercambiables y normalizados, ni menos aun reflexionado

sobre la forma de recopilar dicha información, especialmente cuando la colección de datos es resultado del vaciado de múltiples fuentes.

Por todo ello, creemos que nuestro trabajo debe enfocarse hacia una metodología más cercana a la gestión *inteligente* (Schöch, 2013), a la par que crítica (Huggett, 2020), de los datos. Un buen punto de partida para ello sería comenzar definiendo un modelado ontológico en el que queden reflejadas las entidades implicadas y las asociaciones posibles, atendiendo, al mismo tiempo, toda implicación semántica que pudiéramos derivar durante el desarrollo. En la experiencia ofrecida durante el diseño de otros modelos que acercan la dimensión social, centrados por ejemplo en información relevante sobre la vida de las personas, como es el caso de Bio CRM, podemos observar una mayor tendencia a la elaboración de modelos de datos semánticos interoperables que pretenden enriquecer los repositorios nutriendolos de datos (biográficos) procedentes de una variada tipología de fuentes (Tuominen et al., 2018; Doerr, 2003).

De igual manera nuestro sistema de registro queda definido en base a un modelo de datos semántico enriquecido por fuentes varias que, en definitiva, busca un equilibrio entre una concepción del pasado consensuada por una mayoría considerable de investigadores –y, por tanto, interoperable– y las particularidades propias del mismo estudio; en nuestro caso, recordemos, debe dar respuesta global a la asimilación y desarrollo del feudalismo castellano por parte de un espacio periférico y de montaña del norte peninsular (Álvaro et al., 2018). Trabajamos particularmente sobre un modelo conceptual diseñado por Mauri (2006) cuya aplicación práctica cuenta ya con un cierto recorrido (Mauri et al., 2012; Travé et al. 2021) y se vertebra sobre los conceptos *Unidad Topográfica* (UT), *Actor* (Ac) y *Unidad Estratigráfica* (UE), formalmente también definidos en otros trabajos (Mauri, 2006; del Fresno, 2016; Travé et al., 2020), y enmarcado en la plataforma de gestión de la información patrimonial HORAI, actualmente en construcción, que será objeto de otras contribuciones en este congreso.

La forma en la que interpretamos y recuperamos la información de una realidad pasada para la cual es preciso incluir el papel del factor humano determina el modo en que estructuramos nuestro modelo (Tuominen et al., 2018; Kozaki et al. 2006; Nuninger et al., 2020). En el marco de trabajo de esta investigación, ello considera una interrelación entre cada uno de los conceptos propuestos por Mauri, cuya identificación queda supeditada a los atributos definitorios de cada uno. Para el caso específico de los Actores, categoría fundamental para el registro de los sujetos que interactúan con el paisaje, reducimos su determinación a una primera tipología general seguida de otra más detallada. Para el primer acercamiento, un Actor puede ser identificado como individual o corporativo, dependiendo del número de participantes reconocidos que integran la categoría. Sobre ello, unos y otros son clasificados como definidos o no definidos, en función de aparecer o no determinados nominalmente.

De esta forma, si el nombre de un Actor designa a una persona particular que actúa a título personal, debe ser identificado como un *Actor individual definido*; en cambio, si la unidad se incluye en la fuente designando un colectivo general, sin la explicitación de sus integrantes – siendo *visigodos* un buen ejemplo de ello –, ésta queda registrada como un *Actor individual no definido*. En lo relativo a los Actores corporativos, podemos encontrar enunciados como “los santos Julián y Basilia”, y podríamos entonces presumir a Julián y Basilia dentro de un único *Ac corporativo definido* por el término *santos*, independientemente de si uno u otro actúa a título personal en otros contextos; por su parte, podría darse el caso de tener que lidiar con los testigos de una donación que podrían no ser reconocidos como un grupo *per se*, ya que su denominación como *Ac corporativo no definido* depende de nuestra interpretación, y para ello es suficiente con que cada individuo firme en calidad de testigo en ese momento y para esa donación.

Otro aspecto de gran relevancia dentro del modelo es la interacción entre unidades (con otras UT o Ac), que en el caso particular del mismo entendemos que se trata de una relación de la cual podemos derivar el *rol*. En toda modelización que considera dicho concepto, de hecho, conviene señalar que generalmente se habitúa a diferenciar

entre el rol que es entendido como un atributo producto de una relación binaria entre dos objetos (A, B), o bien el que podría recoger una característica que le es inherente a cada objeto con independencia de cualquier relación existente dentro del sistema (Guarino, 1992; Steinman, 2000). En este sentido, y a pesar de la dificultad que supone su aplicación práctica en la creación de bases de datos (Jaekel, 2013), contamos con experiencias previas perfectamente funcionales cuyos diseños ya parten de una noción diferenciada en la definición y caracterización del concepto rol, bien para la elaboración de un sistema conceptual propio (Kozaki et al. 2006), o bien aprovechando modelos previos (Tuominen et al., 2018).

La noción que nosotros ofrecemos y acercamos, y que nos servirá de base para la declaración de las asociaciones posibles entre cada uno de los elementos, asume también explícita y separadamente el rol como un atributo derivado de una relación, diferenciándose de la otra forma de representación, que para nuestro estudio queda perfectamente recogido como los atributos propios que definen al Actor y que han sido ya señalados. Partiendo de estos preceptos, en su vinculación con las UT, a cada Actor se le puede establecer un papel activo o pasivo sobre la primera. De esta manera, en la donación de bienes de un determinado territorio no solamente puede ser posible la ubicación más o menos exacta del conjunto propiedades, sino que al mismo tiempo nos permite establecer y vincular al (Ac) donatario y al (Ac) beneficiario.

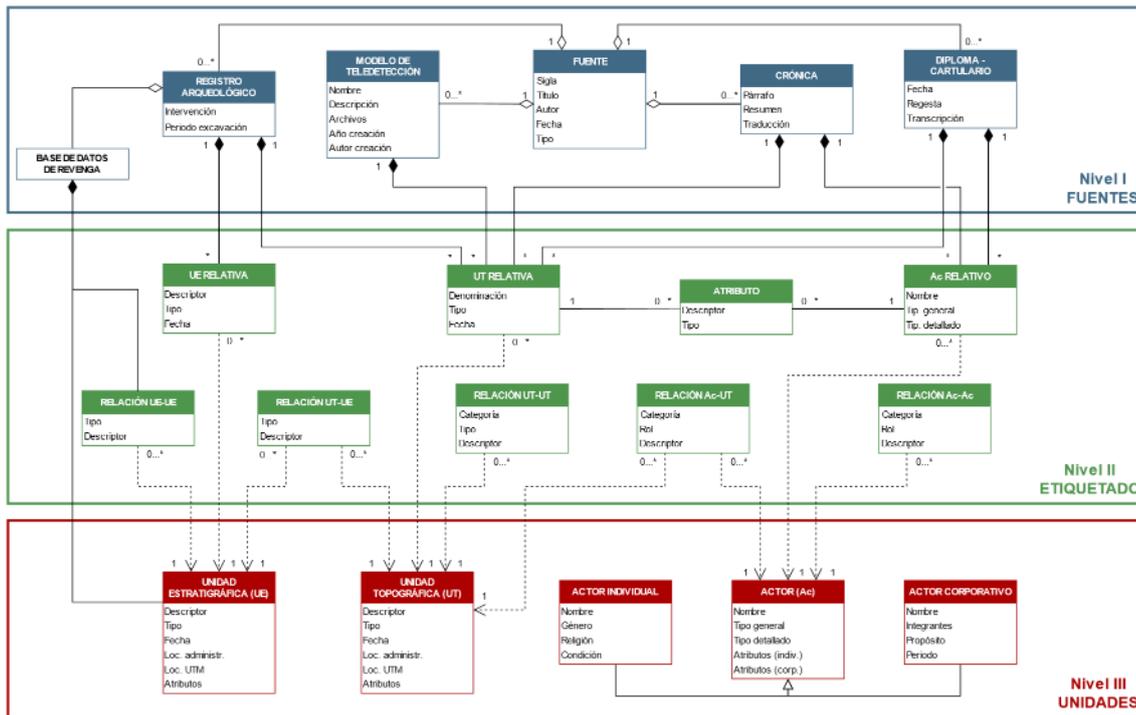
Paralelamente, las asociaciones entre Actores y el rol resultante también es un aspecto que varía en función del tipo de relación establecida. Por ejemplo, las relaciones familiares pueden determinar la paternidad/maternidad de un Actor con respecto a otro, al mismo tiempo que determina la filiación del segundo con respecto al primero. De igual manera, en aquellas de carácter político o social, nos encontramos con otro amplio abanico de posibilidades sistematizadas según el contexto del que son producidas. Así, es más sencillo de registrar el papel de ciertos personajes dentro del entramado que conformará el paisaje altomedieval, como pudiera ser el caso de los *boni homine* con respecto a las comunidades campesinas o la élite local (Luis Corral, 2015; Álvarez, 2013), o del monarca en relación con otros personajes notables del periodo (Carvajal, 2017).

Todas y cada una de estas ideas son recogidas y materializadas dentro de una base de datos (Figura 1) para la cual se han habilitado tres niveles de trabajo que responden a las necesidades propias del estudio. Un primer nivel se corresponde con la caracterización de las fuentes, para lo cual cinco tablas registran y definen los elementos que nos permiten localizar, así como individualizar, los trozos de información manejados. El siguiente nivel pretende relativizar cada unidad identificada, dando una respuesta al trabajo específico de vinculación del dato con cada uno de los registros y permitiéndonos su contextualización dentro de las mismas fuentes. Esto es posible gracias al tercer nivel, que nos permite describir de manera separada la unidad, individualizando su información en un único registro sin importar el número de veces que pueda aparecer para cada fuente.

De este modo, el flujo de trabajo resultante deviene un proceso relacionado entre sí, pero con un cierto grado de autonomía y bajo un grado de control mayor que el que ofrece una vinculación directa de la unidad a la fuente; un acercamiento que nos forzaría la esencial memorización o revisión constante de todos y cada uno de los datos que se van identificando e incrementando en número conforme avanza la investigación.

Con todo, y hasta el momento, nuestra investigación ha estado caracterizando un considerable número de datos en el estudio del proceso de feudalización en la cuenca alta del Arlanza. Aquí el estudio no solamente nos está permitiendo ubicar en espacio y tiempo la actuación materializada en el paisaje a través de elementos, denominaciones o dinámicas varias, sino que, a su vez, subraya la importancia y ofrece el testimonio de todos aquellos sujetos que han realizado o sufrido dicha interacción en consonancia con las diferentes etapas del proceso en su totalidad, diferenciándose así de aquellos otros estudios difícilmente reusables y escasamente integrados y compartidos.

Figura 1. Modelo de datos al que se ajusta nuestro proyecto de investigación.



Bibliografía

- ÁLVAREZ, I. (coord.). *Comunidades locales y poderes feudales en la Edad Media*. Universidad de La Rioja, Servicio de Publicaciones, 2013.
- ÁLVARO, K.; TRAVÉ, E.; LÓPEZ, M. D. «La acción monástica y eremítica en los procesos de poblamiento altomedievales (ss. VI-XI): Los conjuntos singulares del Alto Arlanza y sus problemas interpretativos». *Arqueología y Territorio Medieval*, 2018, 25, 145-167.
- CARVAJAL, Á. *Bajo la máscara del Regnum. La monarquía asturleonese en León (854-1037)*. Madrid: Consejo Superior de Investigaciones Científicas, 2017.
- CARVAJAL, Á.; ESCALONA, J. «The value of status: monetary penalties in the charters from León (854–1037)». *Journal of Medieval History*, 2020, 46:1, 23-49. DOI:10.1080/03044181.2019.1697726.
- DEL FRESNO, P. *Sistema de Información Arqueológica: Propuesta de Normalización, Desarrollo Conceptual e Informático*. Tesis doctoral. Vitoria: Universidad del País Vasco, 2016.
- DOERR, M. «The CIDOC Conceptual Reference Module. An Ontological Approach to Semantic Interoperability of Metadata». *AI Magazine*, 2003, 24, 3, 75-92.
- GUARINO, N. «Concepts, Attributes, and Arbitrary Relations Some Linguistic and Ontological Criteria for Structuring Knowledge Bases». *Data & Knowledge Engineering*, 1992, 8, 249-261.
- JAEKEL, T. «The Role Concept for Relational Database Management Systems», en J. Parsons y D. Chiu (eds). *Advances in Conceptual Modeling. ER 2013 Workshops, LSAWM, MoBiD, RIGiM, SeCoGIS, WISM, DaSeM, SCME, and PhD Symposium, Hong Kong, China, November 11-13, 2013, Revised Selected Papers*. Springer, 2014, 277-286.

- KOZAKI, K.; SUNAGAWA, E.; KITAMURA, Y.; MIZOGUCHI, R. «Fundamental Consideration of Role Concepts for Ontology Evaluation», en D. Vrandečić, M. C. SuárezFiguroa, A. Gangemi y Y. Sure (eds.). *Proceedings of EON2006. Evaluation of Ontologies for Web. Edinburg, UK, May 22, 2006*. CEUR, 2006. Disponible en: <http://ceur-ws.org/Vol-179/eon2006kozakietal.pdf>.
- LUIS CORRAL, F. «Lugares de reunión, boni homines y presbíteros en Valdevimbre y Ardón en la Alta Edad Media». *Medievalista*, 2015, 18, ISSN-e 1646-740X.
- MAURI, A. *La configuració del paisatge medieval: El comtat de Barcelona fins al segle XI*. Tesis doctoral. Barcelona: Universitat de Barcelona, 2006.
- MARTÍN VISO, I. «Commons and the construction of power in the early Middle Ages: tenth-century León and Castile». *Journal of Medieval History*, 2020, 46:4, 373-395. DOI:10.1080/03044181.2020.1784777.
- NUNINGER, L. ET AL. «Developing FAIR Ontological Pathways: Linking Evidence of Movement in Lidar to Models of Human Behaviour». *Journal of Computer Applications in Archaeology*, 2020, 3(1), pp. 63–75. DOI: 10.5334/jcaa.46.
- SCHÖCH, C. «Big? Smart? Clean? Messy? Data in the Humanities». *Journal of Digital Humanities*, 2013, 2, 3. Disponible en: <http://journalofdigitalhumanities.org/2-3/big><http://journalofdigitalhumanities.org/2-3/big-smart-clean-messy-data-in-the-humanities/smart-clean-messy-data-in-the-humanities/>.
- STEINMAN, F. «On the representation of roles in object-oriented and conceptual modelling». *Data & Knowledge Engineering*, 2000, 35, 83-106.
- TRAVÉ, E.; DEL FRESNO, P.; MAURI, A. «Ontology-mediated Historical Data Modelling: Theoretical and Practical Tools for an Integrated Construction of the Past». *Information*, 2020, 11, 182. DOI:10.3390/info11040182.
- TRAVÉ, E.; DEL FRESNO, P.; MAURI, A.; MEDINA, S. «The Semantics of History. Interdisciplinary Categories and Methods for Digital Historical Research». *International Journal of Interactive Multimedia and Artificial Intelligence*, 2021, 6 (5), 47-56.
- TUOMINEN, J.; HYVÖNEN, E.; LESKINEN, P. «Bio CRM: A Data Model for Representing Biographical Data for Prosopographical Research», en A. Fokkens, S. ter Braake, R. Sluijter, P. Arthur y E. Wandl-Vogt (eds.). *Proceedings of the Second Conference on Biographical Data in a Digital World 2017*, 2018, 59-66.

Representaciones 3D, realidad virtual y realidad aumentada

Modera: José Julio García Arranz · *Asiste:* Alexandra Iglesias Cortés

- 19 3D: más allá de la reconstrucción virtual. Una herramienta al servicio de la gestión integral del patrimonio.
Ermita de Sant Llorenç de la Senabra
Pablo del Fresno, Isidre Ot Padilla, José Manuel Martínez Torrecilla
- 4 Realidad virtual para el registro de fenómenos intangibles. Instalaciones y ambientes lumínicos contemporáneos como estudio de caso
Leticia Crespillo Marí
- 2 Hacia el modelado digital de una arquitectura desaparecida: el asilo de las Mercedes de Lugo
Estefanía López Salas
- 62 Teatro digital: en busca de una herramienta de investigación de reconstrucción escénica
Elena Truan
- 30 Patrimonio Cultural y conocimiento: Propuestas para el diseño de una experiencia inmersiva basada en Realidad Mixta para el sitio arqueológico Cueva de las Manos (Argentina)
Marina Gutiérrez De Angelis, Gorka López de Munain

3D: más allá de la reconstrucción virtual. Una herramienta al servicio de la gestión integral del patrimonio. Ermita de Sant Llorenç de la Senabra

Pablo del Fresno Bernal^{1,2}, Isidre Ot Padilla³, José Manuel Martínez Torrecilla⁴.

1. Sistemes de Gestió de Patrimoni SCCL
2. GRAMP.UB – Grup de Recerca d’Arqueologia Medieval i Postmedieval
3. Virtual Past
4. Qark Arqueología SL

Palabras clave: Reconstrucción virtual 3D, Gestión Integral, Patrimonio construido, Documentación virtual, Sant Llorenç de la Senabra

Descripción del problema de investigación, metodología y resultados obtenidos

Presentamos un proyecto de gestión integral del Patrimonio que tiene como objetivo reflexionar acerca de las razones por las que trabajamos en entornos virtuales en contextos de gestión patrimonial. Ciertamente, los procesos de reconstrucción virtual suscitan un interés creciente en el campo de las humanidades digitales, pero resulta imprescindible a nuestro juicio una reflexión serena acerca de las posibilidades que este nuevo escenario nos ofrece.

Nuestro caso de estudio, la ermita de Sant Llorenç de la Senabra, (Santa Margarida i els Monjos, Parc del Foix, Barcelona) está siendo objeto en la actualidad de un proyecto de revalorización patrimonial iniciado en el año 2011 y financiado por la Diputación de Barcelona y el Consorcio del Parc del Foix (Figura 1).

Enclavada dentro del paraje de la Senabra, en el Parc del Foix (red de parques naturales de la Diputación de Barcelona), presentaba un estado total de ruina y abandono. Las sucesivas campañas arqueológicas han puesto de manifiesto una ocupación del lugar que podemos sintetizar en dos grandes fases históricas (Del Fresno & Socorregut 2018).

1. Construcción y uso de una iglesia románica (s. XII-XIV)
2. Construcción y uso de una explotación agrícola ganadera que reaprovecha la iglesia ya en ruinas (s. XVIII-XIX)

En el año 2018 se realizó la última campaña de excavación y una intervención de conservación preventiva de los abundantes y deteriorados restos materiales documentados correspondientes a ambas fases constructivas.

En este punto, nuestra ermita presenta una problemática común para muchos otros elementos patrimoniales, la ausencia de estrategias de gestión integral del sitio que permitan su revalorización patrimonial y en ocasiones, su propia integridad estructural.

Por este motivo, en el año 2020, la dirección del parque encargó la redacción de un Plan Director que estructurase y coordinase la secuencia de actuaciones que permitiesen tanto la salvaguarda del conjunto, como su socialización y gestión como equipamiento cultural integrado dentro de la estructura del parque y la red patrimonial municipal.

Sistemes de Gestió de Patrimoni SCCL, empresa dedicada a la gestión del patrimonio cultural, ha sido la encargada de coordinar la redacción de dicho Plan Director, para ello, ha configurado un equipo de trabajo integrado por diferentes Ciencias del Patrimonio: empresas especializadas en la conservación y restauración

(ÀBAC SL), en la investigación y gestión del patrimonio construido (QARK Arqueología SL), en la recreación virtual (Virtual Past) y en la ilustración de evocaciones históricas (dbòlit).

Los objetivos concretos de esta primera fase del Plan Director, organizados según la Cadena de Valor del Patrimonio Cultural son los que detallamos a continuación, sin que ello menoscabe la posibilidad de potenciar las estrategias de socialización en un futuro inmediato (Figura 2).

Significación:

- Interpretar, entender y contextualizar cada una de las fases de ocupación y abandono del conjunto.

Evaluación:

- Valoración del estado de conservación.
- Valoración de los riesgos estructurales, ambientales y sociales.

Conservación:

- Redacción del proyecto de conservación preventiva.
- Redacción del proyecto de conservación curativa.
- Redacción del proyecto de restauración.

Circulación

- Redacción del proyecto de adecuación.
- Redacción del proyecto de narrativa.

Toda esta información ahora presentada en el Plan Director, entre otras cuestiones, debe servir de referencia para la redacción de un proyecto arquitectónico que permita la ejecución del equipamiento cultural planteado.

El aspecto metodológico de esta investigación es el que nos lleva a la necesidad de modelar en entornos virtuales, en dos momentos bien diferentes de la secuencia de actuaciones que forman el Plan Director.

En primer lugar, tras la búsqueda de paralelos documentados (Figura 3) que nos permitieran comprender las diversas soluciones constructivas desaparecidas y adoptadas en diferentes épocas, pero superpuestas en el mismo espacio, vimos necesario reconstruir las diferentes arquitecturas para comprobar la viabilidad de las teorías adoptadas. Es decir, reconstruimos virtualmente para entender la construcción física ya desaparecida, y su relación con aquello que aún hemos podido documentar in situ.

En la actualidad son infinitos los ejemplos de aplicación de reproducciones virtuales, realidad virtual y realidad aumentada a los proyectos patrimoniales. Es evidente que con la democratización de la tecnología necesaria, obtener evocaciones virtuales de los volúmenes construidos del pasado reciente o remoto, están mucho más al alcance de lo que estaba solo hace unos años. Un proceso similar lo vivimos con la posibilidad de desarrollar estudios de modelado del terreno (MDT), en poco tiempo, nos encontramos ante una avalancha de publicaciones que presentaban modelos digitales del espacio estudiado.

Pero para ambos casos, nos surge la misma duda. ¿Por qué lo hacemos? Es decir, ¿hay siempre una pregunta detrás de cada reconstrucción virtual o de cada MDT? Llegamos a lo virtual, ¿por qué podemos, o por qué debemos?

En el marco del proyecto que ahora presentamos, hemos necesitado generar reproducciones 3D que nos permitieran documentar, entender, interpretar y por último representar aquello que está siendo objeto de estudio y gestión. Y a la vez, hemos identificado necesidades metodológicas, conceptuales e instrumentales que de

momento no hemos podido integrar en nuestra dinámica de trabajo, pero que implican la necesidad de desarrollar líneas de investigación inmediatas sobre las que podemos desarrollar algunas ideas en nuestra comunicación.

En definitiva, presentamos un planteamiento de acercamiento a lo virtual. Como historiadores, como arqueólogos, creemos que las infinitas posibilidades del mundo virtual deben ir más allá de la simple evocación de una construcción en un momento determinado, sin quitarle un ápice al inmenso poder divulgativo que esta posibilidad representa. Pero además, consideramos que deben ayudar a entender y modelar tanto el propio elemento objeto de estudio, como el contexto histórico en el que esas construcciones se han erigido.

En este sentido, teniendo en cuenta la dimensión de realidad cambiante atribuible al Patrimonio Cultural y en extensión al arqueológico, el registro –cualquier registro, pero más el que se puede obtener a partir del método arqueológico– nos permite entender y explicar las diferentes realidades documentadas, las aún existentes y las ya desaparecidas de un objeto patrimonial y *es precisamente esta evolución constante la que origina que la información herede el carácter patrimonial de los elementos que describe dado que, desde el momento en que éstos ya no existen tal y como se documentaron, la información capturada queda como única fuente para su estudio. En consecuencia, se puede afirmar que la información sobre los elementos patrimoniales es, en sí misma, patrimonio y que, por este motivo, debe ser conservada* (Rodríguez 2014: 1).

Por lo tanto, afirmamos que el Patrimonio Arqueológico, aquel que se configura a través de la mirada arqueológica de las fuentes históricas, no solo está formado por los elementos materiales existentes, sino también por las diferentes evidencias identificadas, por los diferentes indicadores informativos, y por todas aquellas recreaciones documentadas, virtuales o no, que podamos construir a partir del Registro Arqueológico recogido de manera normalizada y sistemática (Del Fresno 2016).

Y a su vez, la sistematización con la que abordemos la construcción de reproducciones virtuales, es decir, la capacidad que podamos tener de transformar los diferentes elementos reproducidos, en unidades informativas que podamos sintetizar junto con aquellas evidencias procedentes de la lectura arqueológica de la ruina material, determinará si estamos ante un ejercicio meramente evocador del pasado, o construyendo herramientas de investigación histórica. En definitiva, creemos que cuando reproducimos virtualmente, también debemos documentar históricamente.

De momento no podemos desarrollar mucho más esta idea, solo señalar que no es una preocupación particular. Como casi todas las cuestiones de nuestro campo de procedencia, la Arqueología, las problemáticas son comunes, pero las soluciones no lo son tanto y nunca están exentas de un cierto debate.

En este sentido, acudiremos a desarrollos ya existentes que encajan perfectamente con nuestro actual entorno conceptual e instrumental (la plataforma de gestión integral Horai). Se trata de la Extended Matrix (EM), lenguaje formal con el que realizar un seguimiento de los procesos de reconstrucción virtual y que está destinado a ser utilizado por arqueólogos y especialistas en patrimonio para documentar de manera robusta sus actividades científicas (Demetrescu 2015).

El EM se basa en el enfoque de lectura estratigráfica y tiene como objetivo crear un marco común que conecte la documentación arqueológica y la reconstrucción virtual en las primeras etapas de la excavación de un sitio o un estudio 4D de un monumento (Demetrescu & Fanini 2017).

Los trabajos que hemos realizado en los últimos años a partir de volúmenes reales (virtuales, pero reales al fin y al cabo) y como ello nos ha permitido contrastar y corregir algunas de las ideas iniciales respecto a la apariencia de la construcción en sus diferentes fases, nos han permitido conseguir algunos resultados al respecto en Sant Llorenç de la Senabra. Así, por ejemplo, el trabajo de modelado virtual ha permitido descartar la existencia de algunos elementos estructurales arquitectónicos propios de la ermita románica pero ausentes ya en la masía,

como la cubierta de la nave central o la misma bóveda del ábside. Conclusiones, estas, a las que sólo hemos podido llegar tras una restitución virtual aproximada de los volúmenes de ambas construcciones en las que se han tenido como referencia, especialmente en el caso de la masía, los escasos recursos gráficos que han llegado hasta nuestros días, en los cuales podía apreciarse la altura original de las diversas naves (Figura 4).

Por otro lado, el uso de las herramientas digitales no sólo ha permitido la reconstrucción del modelo virtual en sus fases pretéritas, si no también reflejar de manera visual y detallada la propuesta para este monumento como equipamiento cultural (Figura 5) en el futuro.

Y es este, precisamente, uno de los puntos fuertes del proyecto, en el que utilizamos las potencialidades y posibilidades de las herramientas digitales para conjugar, a partir de un modelo 3D, pasado, presente y futuro de Sant Llorenç de la Senabra, articulando a través de él todo el proceso de gestión integral que incorpora las conclusiones aportadas desde los distintos campos profesionales que han intervenido en el proyecto (arqueología; arqueología virtual; conservación; arquitectura; etc). Acercando nuestro proyecto a la metodología de trabajo colaborativa *Building Information Modeling (BIM)* para la creación y gestión de un proyecto de construcción. No en vano, su objetivo es centralizar toda la información del proyecto en un modelo de información digital creado por todos sus agentes.

Y todo ello sin causar un impacto real e irreversible sobre el propio monumento, pues, el trabajo en entornos virtuales nos permite determinar e integrar algunas de las soluciones constructivas ya empleadas en el pasado, valorando la idoneidad de las mismas y, de cara a posibles actuaciones futuras, reducir el impacto visual y estructural de las intervenciones, manteniendo una coherencia arquitectónica e histórica que ayude a la comprensión de la diacronía existente.

Figuras, tablas, mapas o gráficas



Figura 1: Estado inicial y actual del conjunto

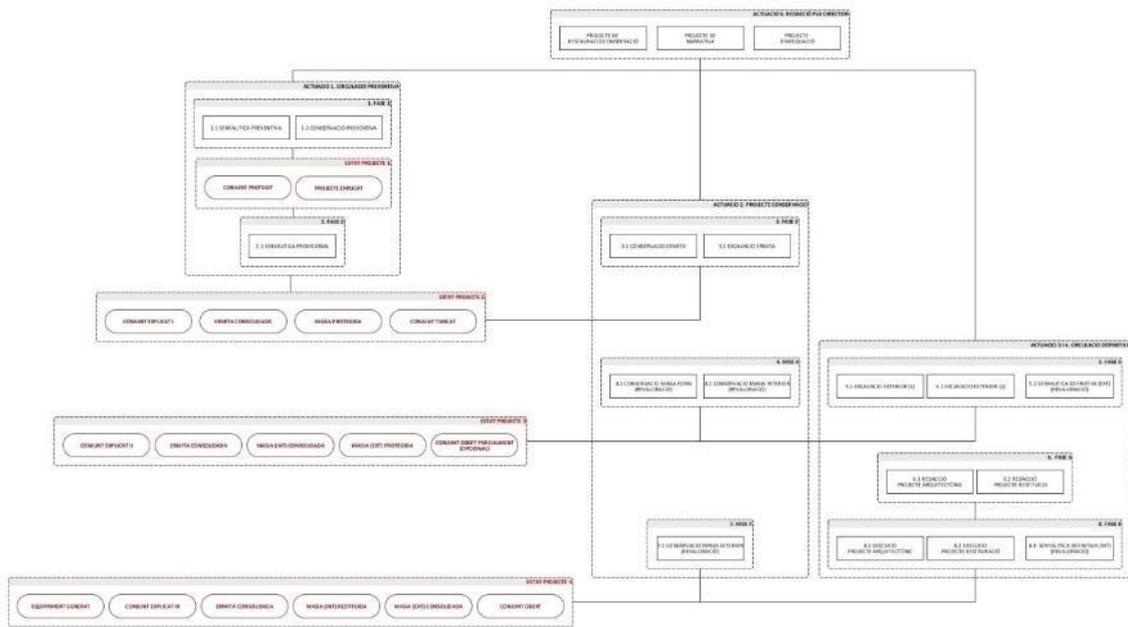


Figura 2: Cronograma actuaciones Plan Director

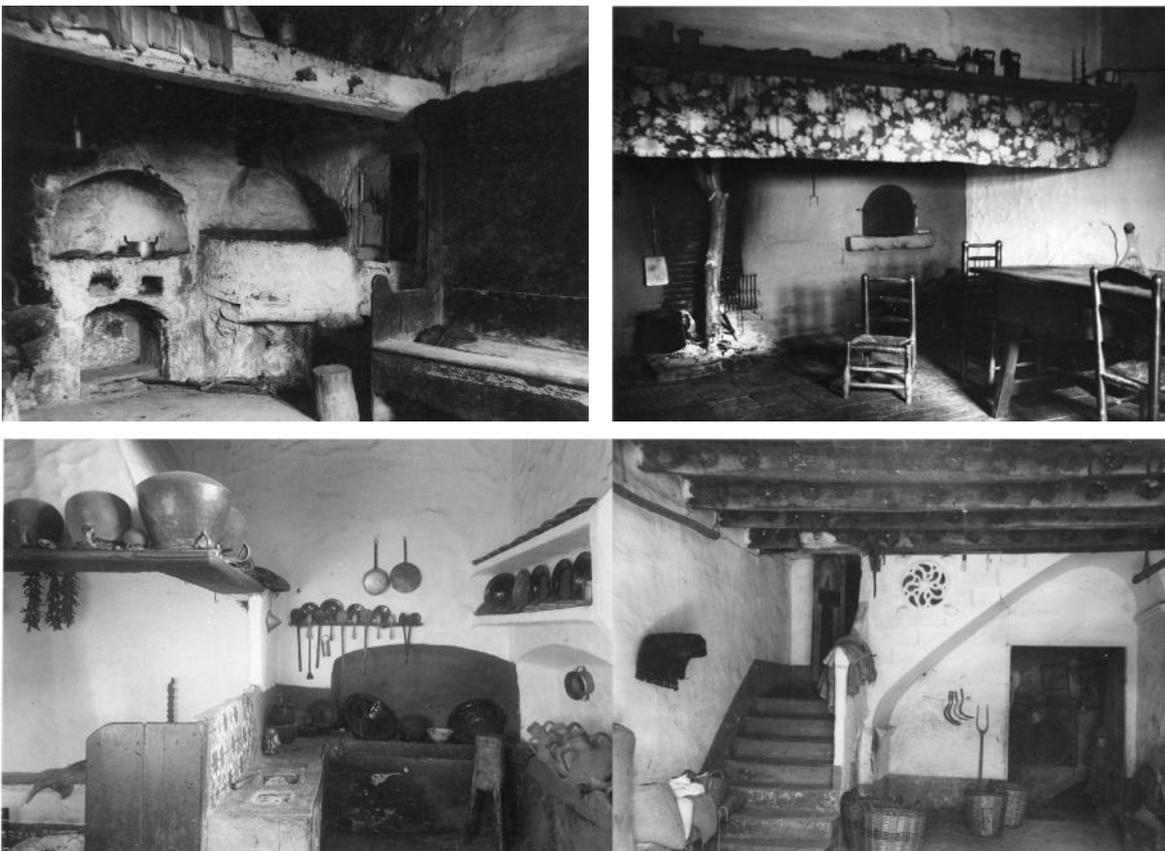


Figura 3: Documentación gráfica de paralelos. Fase masía

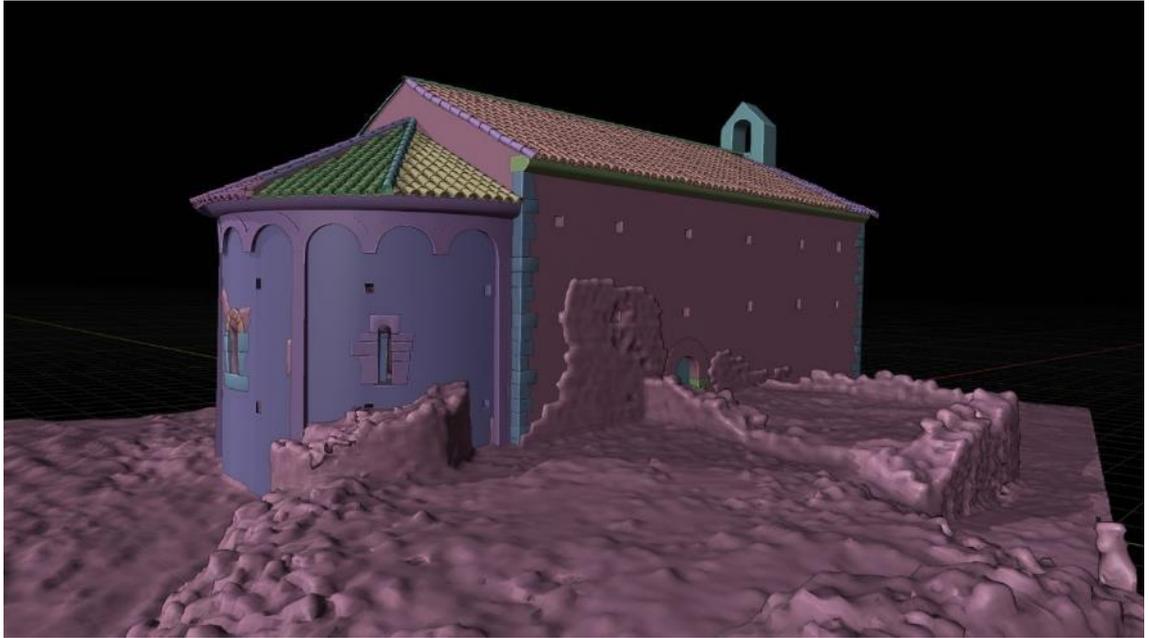


Figura 4: Reconstrucción virtual. Fase ermita (en elaboración)



Figura 5: Reconstrucción virtual. Equipamiento cultural (en elaboración)

Bibliografía

Del Fresno, P. (2016): *Sistema de Información Arqueológica: Propuesta de Normalización, Desarrollo Conceptual e Informático*. Tesis doctoral, Universidad del País Vasco, Vitoria.

Del Fresno, P. Socorregut, J. (2018): *Memòria de la intervenció arqueològica preventiva a Sant Llorenç de la Senabre d'acord amb la resolució d'autorització 470 K121 N-661 2018-1/22745 SantaMargarida i els Monjos (Alt Penedès)*. Memoria técnica inédita. Disponible en <https://www.academia.edu/40432667/>.

Demetrescu, E. (2015): Archaeological stratigraphy as a formal language for virtual reconstruction. Theory and practice. *Journal of Archaeological Science* 57, 42-55.

Demetrescu, E., Fanini, B. (2017): A White-Box Framework to Oversee Archaeological Virtual Reconstructions in Space and Time: Methods and Tools. *Journal of Archaeological Science: Reports* 14, 500–514.

Mauri, A. (2006): *La Configuració del Paisatge Medieval: El Comtat de Barcelona Fins al Segle XI*. Tesis doctoral. Universitat de Barcelona, Barcelona.

Rodríguez Miranda, A. (2014): *Documentación espacial del patrimonio: preservación de la información. Necesidades, posibilidades, estrategias y estándares*. Tesis doctoral, Universidad del País Vasco, Vitoria.

Travé, E.; Del Fresno, P.; Mauri, A. (2020): Ontology-mediated Historical Data Modelling: Theoretical and Practical Tools for an Integrated Construction of the Past, *Digital Humanities Special Issue. Information*, 11, 182.

Travé, E.; Del Fresno, P.; Mauri, A.; Medina, S. (2021): The Semantics of History. Interdisciplinary Categories and Methods for Digital Historical Research, *International Journal of Interactive Multimedia and Artificial Intelligence*, 6 (5), 47–56.

Realidad virtual para el registro de fenómenos intangibles. Instalaciones y ambientes lumínicos contemporáneos como estudio de caso

Leticia Crespillo Marí (Universidad de Málaga)

Palabras clave: Realidad Virtual, Light Art Contemporáneo, registro catalográfico, estética, museografía

- Descripción del problema de investigación

El catálogo de exposición ha sido siempre una herramienta muy importante en cuanto al registro y difusión de contenidos culturales e histórico-artísticos en entornos museísticos. Su formato impreso ha trascendido las barreras del tiempo adaptándose en cada momento a las necesidades comunicativas que derivan de los diferentes contextos sociales. Paulatinamente hemos visto esta herramienta museográfica evolucionar gracias a la digitalización tecnológica permitiendo un acceso y democratización de contenidos, hasta entonces, nunca visto. Sin embargo, hay que reflexionar profundamente sobre cómo se muestran determinados enunciados artísticos, valorando así, si el catálogo tradicional (libro) o el digitalizado (en una pantalla) son capaces de solucionar problemas derivados del nuevo contexto tecno-social. En este sentido, como historiadora del arte, me hago preguntas sobre el enorme y rápido desarrollo de las tecnologías de análisis computacional, *softwares*, *hardwares*, dispositivos o prótesis etc., que pueden sernos útiles en nuestra línea de investigación ligada, tanto a la museografía de intangibles, como al registro mismo del componente estético y experiencial derivado de la participación directa del sujeto-espectador en la intervención artística. Si algo necesita la museografía, en la actualidad, es una renovación en las técnicas de registro, catalogación y puesta en valor de la obra de arte, sobre todo, de cara a un usuario cada vez más tecno-inteligente e hiperestimulado, cuyos dispositivos son parte de su cotidianidad. Con ello, quiero incidir en la necesidad de un proceso comunicativo adaptado a las nuevas relaciones e interpretaciones significativas que surgen de la experimentación dentro de un museo, una galería, un aula o cualquier institución cultural.

Mi objeto de estudio es la instalación/environment de luz contemporáneo (abarcando propuestas que van desde los años 60 y 70 del siglo XX hasta los años 20 de este siglo XXI)⁷⁹. Estas intervenciones son complejas por sus cualidades estéticas; el propio uso de la luz como material intangible o el entorno espacial tridimensional inmersivo como lugar de experiencia corporal no son registrables en dos dimensiones. Esto supone un problema pues la inmersión también caracteriza a estas manifestaciones lumínicas y, para poder ser experimentadas en su totalidad, es necesario estar allí. Desde una perspectiva fenomenológica será necesario transitar el espacio e involucrarse en el mismo, tanto performativa como procesualmente. El sujeto-espectador ya no es un elemento ajeno, si no que se convierte en el centro de la obra, completándola subjetivamente a través de la propia experiencia en primera persona que obtiene interactuando. Es una casuística compleja, pues hablamos de obras con un carácter abierto (Eco, 1962) y rizomático (Deleuze y Guattari, 1980) de las que se pueden extraer múltiples lecturas; un artefacto que necesita del sujeto-espectador, su sensibilidad y un *feedback* continuo para existir como realidad en un tiempoespacio determinado en el que, además, se ve involucrado el constructo fisiológico único e intransferible del mismo individuo. Todos estos ítems quedan fuera del soporte editorial 2D propio, tanto del catálogo tradicional como del digitalizado, cuyas fórmulas se antojan insuficientes, por cuanto son limitadas.

⁷⁹ Dan Flavin, James Turrell, Ann Veronica Janssens, Matthew Schreiber, Carlos Cruz-Diez, Olafur Eliasson, Bruce Nauman, Anthony McCall, entre muchos otros artistas analizados.

Las nuevas tecnologías han evolucionado rápido y esto es una ventaja para este estudio: dispositivos mediales, *hardwares* o *softwares* gráficos han superado con creces todas nuestras expectativas de futuro. Con ello, se abre un campo enorme a la experimentación dentro de las humanidades digitales en general y la historia del arte en particular (desde la perspectiva museográfica). La Realidad Virtual se convierte en una herramienta útil para estos casos, pues la interfaz ofrece inmersión, interacción, participación corporal, performatividad en un espacio construido y procesualidad gestual y situacional donde, en resumen, *el valor estético va a depender de la actitud que tome el espectador frente al artefacto* (Hernández, 2011). Por tanto, su investigación, documentación y preservación van a tener un impacto enorme en su revalorización como producción cultural.

Se trata, pues, de una investigación con un carácter eminentemente transversal que aúna la investigación histórico-artística cualitativa y cuantitativa (análisis de datos), así como el desarrollo tecnológico derivado de la propuesta de prototipado virtual (RV)⁸⁰ realizado junto a la empresa *Opossum Studios Virtual Builders*⁸¹.

- Metodología y resultados obtenidos

Se trata de una investigación de tipo mixto (cualitativo y cuantitativo). El primer paso (cualitativo) ha sido indagar, desde la perspectiva de la filosofía del arte y la estética, cuestiones relacionadas con el sujeto de la experiencia y su relación con la luz como elemento plástico y estético intangible. Se realiza una síntesis bibliográfica (algo extensa) que se adhiere a otros aspectos propios del campo de la percepción visual y corporal. Además, se atiende a cuestiones sobre la recepción y las nuevas tecnologías (estética de la simulación, estética de la participación (Sánchez Vázquez, 2005), estética de la habitabilidad (Hernández, 2003), estética virtual) así como sobre la fisiología del cuerpo humano (Huston et al, 2015) y la fenomenología de la percepción (Husserl, 2002), la neurofenomenología de la percepción (Laughlin, McManus y D'Aquili, 1990), la somaestética (Shusterman, 2013), la psicología perceptiva (Evans, 1948), la kinestesia, la somática o la neuroestética (Zahavi y Gallagher, 2014). Todo ello nos puede ayudar a entender esa creación de vínculos sensibles por parte del individuo receptor con la intervención (Maturana 1996; 2004; 2006)⁸².

Realizado esto, se procede a la recopilación de textos de estudio a analizar. Se crea una base de datos relacional relativa a catálogos tradicionales que se han adquirido o consultado (catálogos relativos a exposiciones de los artistas individuales seleccionados y relacionados con el objeto de este estudio). La selección de artistas se hace con base a la relación de influencias mutuas⁵ que existe entre los periodos anteriormente indicados (años 60-70 siglo XX- actualidad). Esta relación se obtiene del cribado y análisis (cuantitativo) de los datos obtenidos de los textos. El corpus analizado lo forman 386 textos extraídos de 170 catálogos y, a su vez, de 40 artistas. Estos textos están formados por ensayos, comentarios de obra, reflexiones de artistas y entrevistas categorizadas en la base de datos relacional. Lo que nos interesa de los textos son las concurrencias de conceptos (estéticos y de experiencia) así como la correlación de los mismos respecto de los artistas seleccionados (frecuencia de repetición). Estos términos, a los que he denominado “descriptores de experiencia” nos devuelve al análisis cualitativo de nuevo, pues son la base para entender cómo funcionan las intervenciones y que elementos no registrables en 2D pueden ser desplazados al entorno tridimensional reconstruido. Los catálogos tradicionales, sin

⁸⁰ Siempre, respetando que esto es un complemento y no un sustituto de la obra original.

⁸¹ Opossum Studios Virtual Builders

⁸² A tener en cuenta que esto es una mínima muestra de bibliografía consultada y trabajada para la tesis. ⁵ Imagen de ecosistema que se mostrará en la presentación y que irá en la tesis doctoral.

otros y que tienen que ver con la condición o la cualidad experiencial de estas intervenciones lumínicas: *space, time, experience, perception, meaning, viewer, process* o *presence*. Todos estos elementos nos van a marcar la ruta de acción para desarrollar un prototipo que resuelva estos ítems, permitiendo una comprensión real del fenómeno completo. La inmersión va a solucionar el problema del espacio tridimensional transitable y la experiencia temporal individual, así como la necesidad de presencia y la percepción en proceso (relación situacional). Incluye al espectador (*viewer*) y su cuerpo (*body*) que en el entorno (*place*) experimentará en presente la intervención (con ello, resolvemos la participación y la interacción).

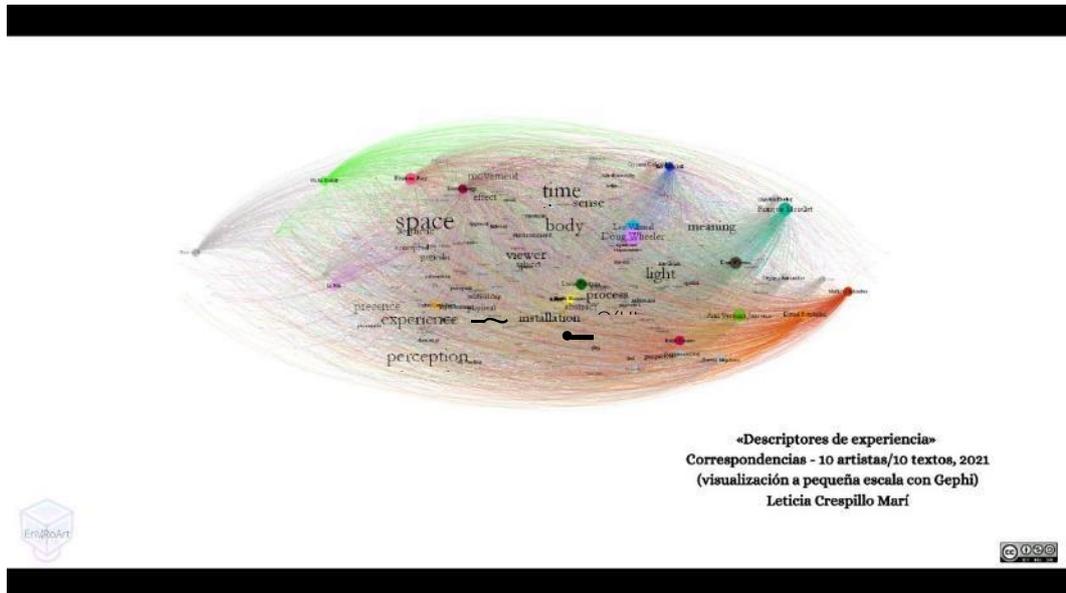


Fig. 3 Relación de correspondencia de "descriptores de experiencia" entre 10 artistas y 10 textos del corpus. Autora: Leticia Crespillo Mari (2021)

No obstante, y para que se pueda ver un pequeño ejemplo de concurrencias, podemos hacer pequeños mapas conceptuales con distintos artistas. En este caso, incluyo tres de los artistas más significativos relacionados con el ámbito del *light art* (en sus respectivas formas de trabajo).

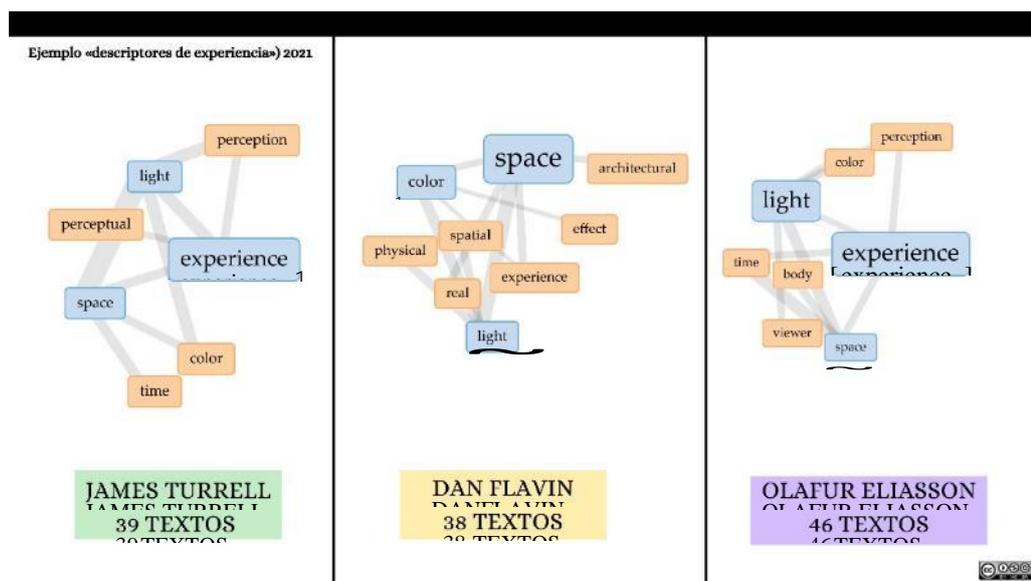


Fig. 4 Pequeña muestra de correspondencia "descriptores de experiencia" James Turrell / Olafur Eliasson / Dan Flavin.

Autora: Leticia Crespillo Mari (2021)

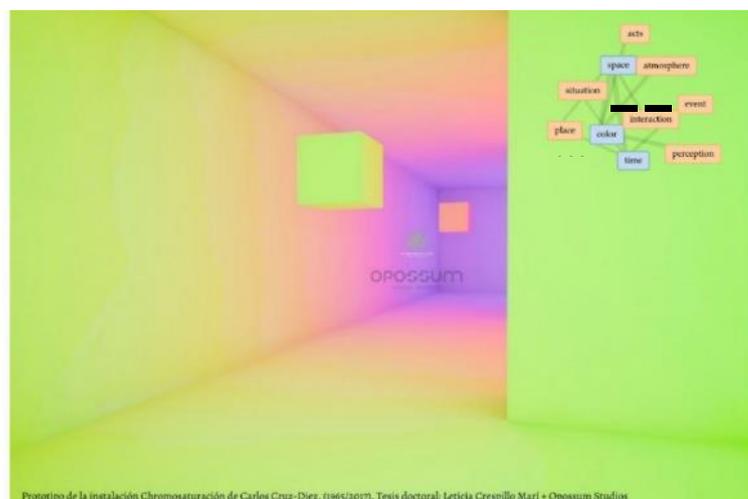
Con ello, se puede cartografiar completamente el comportamiento de estas intervenciones y los requerimientos básicos necesarios para realizar el prototipo correspondiente. También se le suma una plataforma de conocimiento (PosDoc) para la difusión de todo este conocimiento relativo a los catálogos seleccionados, los textos extraídos, el análisis conceptual los artistas y las 800 obras que lo complementan.

Previo a todo este proceso de análisis se dejaron listos tres prototipos vacíos en realidad virtual (en fase beta):

- Dan Flavin, *Untitled (to Helga and Carlo with respect and affection)* 1974; extraído del catálogo *Series and progressions* de la Galería David Zwirner de Nueva York⁸³
- Ann Verónica Janssens, *Yellowbluepink*, 2015; extraído del catálogo *Experienced* de la Galería Toni Tàpies de Barcelona para su reconstrucción.
- Carlos Cruz Díez, *Chromosaturación*, 1965/2017; extraído de la exposición realizada en el Museo Würth de La Rioja y el catálogo *Informed by color*.



Fig. 5 Prototipo (fase beta) en realidad virtual instalación Dan Flavin, *Untitled (to Helga and Carlo with respect and affection)* 1967. Autores: Leticia Crespillo y Gabriel González (2020)



⁸³ Crespillo, L & Ruiz, J.G (2020) demo *Prototipo de Dan Flavin, Untitled (to Helga and Carlo with respect and affection) 1977* [Online] Disponible en: <https://youtu.be/99v7-L6a-U>. (consultado: 27 Feb 2021)

Fig. 6 Prototipo (fase beta) en realidad virtual instalación Carlos Cruz-Diez, Chromosaturation (1965/2017) Autores: Leticia Crespillo y Gabriel González (2021)

La fase final del proyecto de tesis pretende ofrecer la alternativa de registro catalográfico ya testada y probada. Estos prototipos han sido realizados en Unreal Engine 4 y pueden visualizarse, en un principio, a través de gafas HTC Vive Pro. Han sido producidos junto a *Opossum Studios Virtual Builders*. A continuación, en la siguiente fase, se implementará la información de modo que la experiencia sea intuitiva y sencilla de manejar. Dos de estos prototipos, más concretamente, los primeros se testaron en 2019 en el New Tech Observatory de Málaga (Polo Digital de contenidos) con resultados bastante interesantes y ligados a la experiencia personal de los usuarios. La tesis será defendida en el primer semestre de 2022.

- Bibliografía

Arheim, R. (2013). *Arte y percepción visual*. Madrid: Alianza Editorial S.A.

Brea, J.L. (2010). *Cambio de régimen escópico: del inconsciente óptico a la e-imagen*. Madrid: AKAL.

Deleuze, G. y Guattari, F. (1980). "Rizoma" en Deleuze, G. y Guattari, F, *Mil mesetas: Capitalismo y esquizofrenia*. Valencia: Editorial Pre-Textos.

Evans, R. M. (1948). *An Introduction to color study of physical, psychophysical and psychological aspects*. EEUU: John Wiley & Sons.

Giannetti, C. (2012). *Estética digital. Sintopía del arte, la ciencia y la tecnología*. Barcelona. ACC L'Angelot.

Hernández García, I. (2003). *Estética de la habitabilidad y nuevas tecnologías*. Colombia: Universidad Javeriana

Hernández, Hernández, Francisca (2011). *El museo como espacio de comunicación*. Gijón: Editorial Trea.

Hudson, J.P; Nadal, M; Mora, F; Agnati, L.F. y Cela Conde, C.J. (2015). *Art, Aesthetics, and the Brain*. Oxford: UOP Oxford.

Husserl, E. (2002). *Lecciones de fenomenología de la conciencia interna del tiempo*. Madrid: Editorial Trotta.

Laughlin, C.D; McManus, J y D'Aquili, E.G. (1990). *Brain, Symbol & Experience. Toward a Neurophenomenology of perception*. Columbia: Columbia University Press.

Manovich, L. (2005). *El lenguaje de los nuevos medios de comunicación. La imagen en la era digital*. Barcelona: Paidós Comunicación.

Martín Prada, J. (2012). *Otro tiempo para el arte. Cuestiones y comentarios sobre el arte actual*. Valencia: Sedema.

Maturana, U y Pörksen, B. (2004). *Del ser al hacer. Los orígenes de la biología del conocer*. Chile: Editorial Comunicaciones Noreste LTDA.

Maturana, U. (1996). *El sentido de lo humano*. Chile: Dolmen Ediciones S.A.

Maturana, U. (2006). *Desde la biología a la psicología*. Chile: Editorial Universitaria SA.

Moreno Sánchez, I. y Moreno Gant, M. L. (2013). *Arte y museo del siglo XXI. Entre los nuevos ámbitos y las intersecciones tecnológicas*. Barcelona: Ediciones UOC.

Pireddu, M y Serra, M. (Eds.). *Mediología: cultura, tecnología y comunicación*. Barcelona: Editorial Gedisa.

Sánchez Vázquez, A. (2005). *De la estética de la Recepción a una estética de la participación*. México: UNAM.

Shusterman, R. (2013). *Conciencia del cuerpo. La filosofía como arte di vivere e la somaestetica*. Milano: Christian Marinotti Edizioni.

Soto Solier, P. M. (2012). *Metáfora de flujos lumínicos*. Granada: Universidad de Granada.

Eco, U. (1962). *La obra abierta*. Barcelona: Planeta DeAgostini.

Zahavi, D. y Gallagher, S. (2014). *La mente fenomenológica*. Madrid: Alianza Editorial.

Hacia el modelado digital de una arquitectura desaparecida: el asilo de las Mercedes de Lugo

Estefanía López Salas, Universidade da Coruña

Palabras clave: historia de la arquitectura, reconstrucción virtual, CAD, modelado 3D, asilo de las Mercedes

El primer propósito de la reconstrucción digital aplicada al campo de la investigación y divulgación de la arquitectura histórica es el de crear un modelo virtual espacial, temporal o semántico que responda a la pregunta de *cómo* era una entidad patrimonial en un determinado momento de su pasado. Tener acceso, desde la contemporaneidad, a la percepción de un espacio histórico, total o parcialmente no conservado, abre las puertas a la recuperación de su memoria y significado, y a la posibilidad de llevar a cabo nuevos análisis de esa arquitectura histórica. La cuestión del *cómo* era es así pareja a la segunda pregunta más importante del *por qué* abordar su representación con herramientas digitales.

Cuando se trata de bienes completamente desaparecidos, la reconstrucción virtual se realiza principalmente por medio de sistemas de modelado computarizado manual 2D y 3D, que requieren la búsqueda e interpretación humana de fuentes de diferente naturaleza, con información sobre la apariencia física de la entidad histórica. La producción de un modelo digital riguroso a partir de estas conduce a un proceso complejo y diferente del de la investigación exclusivamente realizada a través de textos y/o imágenes para producir resultados eminentemente textuales. Cuando intentamos *re*-construir de modo visual un espacio perdido nos enfrentamos a múltiples cuestiones sobre su forma, dimensiones y materialidad, entre otras, claves para su materialización física en el mundo real, pero igualmente en el virtual. Al intentar darles respuesta de modo visual adquirimos un mejor conocimiento del edificio histórico, que se produce durante el proceso de *hacer* el modelado. El modelado digital se convierte así en un método de investigación en sí mismo, pues revela numerosas cuestiones espaciales que pasan desapercibidas con el uso de otros procedimientos tradicionales.

En la presente comunicación proponemos mostrar el conocimiento adquirido sobre una arquitectura histórica concreta durante las primeras fases del flujo de trabajo llevado a cabo para su reconstrucción virtual. Nuestro caso de estudio es el edificio proyectado y construido como asilo de las Mercedes en Lugo, a finales del siglo XIX, y demolido en 1964.⁸⁴ Pionero en su tipología en Galicia, no llegó nunca a utilizarse con su función original y sufrió múltiples transformaciones en el tiempo, principalmente para su uso como colegio militar y cuartel, hasta su completa desaparición. De sus diferentes fases evolutivas, la que abordamos en este trabajo es la primera y más antigua, la del asilo, por ser esta la que definió su fábrica original y condicionó las posteriores obras de adaptación a otros programas funcionales. Con la reconstrucción digital de este espacio histórico no sólo buscamos recuperar la imagen perdida del complejo asistencial, sino también sentar las bases para, en una fase futura, analizar las intenciones del arquitecto en la creación del espacio y la importancia de la propuesta en su época.

A partir de una recopilación de fotografías y planos históricos del proyecto original y de otros de reformas sucesivas localizados en archivos y prensa histórica, elaboramos las trazas principales del complejo, primeramente como dibujos bidimensionales en CAD con medidas reales para cada espacio componente, extraídas de datos contenidos en fuentes textuales o del escalado de la planimetría histórica. Las trazas resultantes sirvieron de base para el modelado tridimensional posterior, asegurando una total consistencia entre

⁸⁴ Este trabajo se enmarca dentro del Proyecto I+D+i 'Memoria del patrimonio arquitectónico desaparecido en Galicia. El siglo XX', PID2019-105009GB-I00, concedido por el Ministerio de Ciencia e Innovación para el periodo 2020-2022.

los dos productos virtuales. En esta fase del flujo de trabajo se produjo un constante proceso comparativo entre las diferentes fuentes disponibles, para poder afrontar la representación del espacio no conservado. Este proceso reveló cuáles son los datos más relevantes para la comprensión del espacio histórico desaparecido, cuáles están ausentes en las fuentes, qué información espacial presenta incongruencias entre unas evidencias documentales y otras, y en qué partes son necesarias, posibles o limitadas, las hipótesis para producir un modelo final completo, coherente y riguroso desde un punto de vista histórico-artístico. Documentar y reflexionar sobre esta parte del flujo de trabajo nos permitirá mostrar, por un lado, el cambio de lente que los métodos de reconstrucción virtual conllevan con respecto a prácticas previas en el estudio del espacio de la arquitectura histórica y, por otro, los problemas asociados a la investigación en la que los métodos y los resultados son eminentemente visuales, frente a la textual, todavía dominante en el mundo académico.

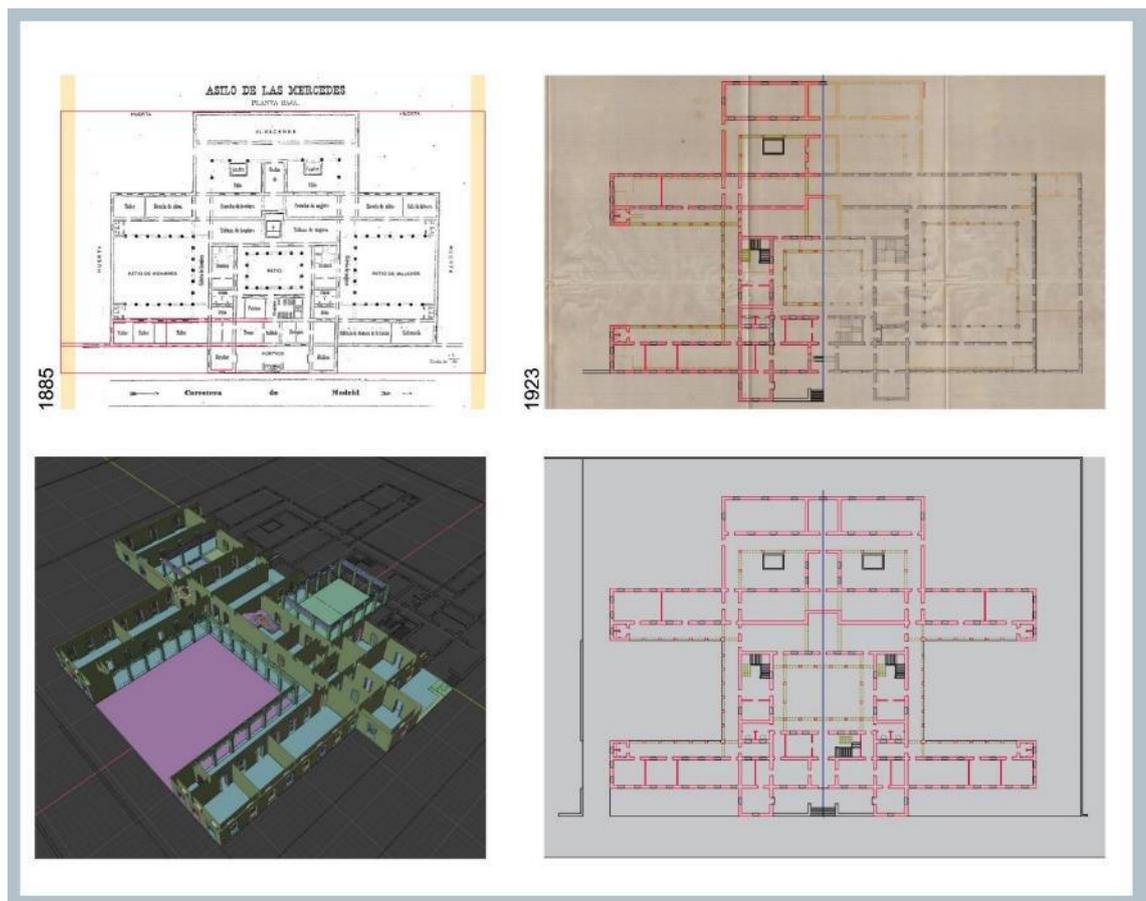


Figura 01. Del documento histórico al modelo digital del desaparecido asilo de las Mercedes de Lugo.

Bibliografía

Abel Vilela, Adolfo. *Urbanismo y Arquitectura en Lugo. Arquitectura Isabelina y de la Restauración*. A Coruña: Edicións do Castro, 1996, pp. 175-181.

Castro Freire, Salvador. *Lugo y sus hombres: ensayo de síntesis histórica*. Lugo: Ediciones Celta Lugo, 1951, pp. 122-123, 140-150.

De Francesco, Giuliana, y D'Andrea, Andrea, "Standards and guidelines for quality digital cultural three-dimensional content creation". En: Marinos Ioannides et al. (ed.). *Digital Heritage: Proceedings of the 14th International Conference on Virtual Systems and Multimedia. Project Papers*, pp. 229-233. Budapest: Archaeolingua.

Fernández Fernández, Carlos M., Sánchez García, Jesús Ángel. "Espacios para el refugio. La asistencia a la vejez y los asilos en Galicia", *Sémata, Ciencias Sociais e Humanidades*, vol. 18, 2006, pp. 167-203.

ICOMOS. *Principles of Seville. International Principles of Virtual Archaeology*. New Delhi: 2017. Recuperado de: <https://icomos.es/wp-content/uploads/2020/06/Seville-Principles-IN-ESFR.pdf>.

La Carta de Londres para la visualización computarizada del patrimonio cultural, versión 2.1. Febrero 2009. Recuperado de: <http://www.londoncharter.org/index.html/>.

Münster, Sander, Friedrichs Kristina y Hegel, Wolfgang, "3D Reconstruction Techniques as a Cultural Shift in Art History?", *International Journal for Digital Art History. Digital Space and Architecture*, nº 3, 2018, pp. 38-60, <https://doi.org/10.11588/dah.2018.3.32473>.

Teatro digital: en busca de una herramienta de investigación de reconstrucción escénica.

Elena Truan Aguirre

Universidad de Salamanca

Palabras clave: teatro, escenografía, reconstrucción, espacios escénicos, digital

El objetivo de esta comunicación consiste en proponer una nueva vía de investigación en los estudios teatrales a través del uso de la tecnología, exponiendo los obstáculos y necesidades que han llevado a ella, y descubriendo los orígenes de un proyecto de humanidades digitales interdisciplinar que pueda facilitar la investigación de los estudios teatrales, particularmente los estudios de teatro del Siglo de Oro y medieval.

Uno de los principales escollos a la hora de investigar obras de teatro del Siglo de Oro se halla en la interpretación de las acotaciones, a menudo vagas y escasas; no obstante, éstas pueden proveer al investigador de valiosa información con respecto a la posición de los actores, la dinámica habitual del movimiento escénico de la época, e incluso la configuración arquitectónica del teatro donde se representó la obra. Las acotaciones son a menudo una parte del texto teatral que, si bien siempre presentes en el género y fundamentales para la interpretación, a menudo caen en una tierra de nadie situada entre los estudios puramente literarios y los estudios teatrales de una obra. Estos dos enfoques, no obstante, deben de ir unidos para una exhaustiva y correcta investigación en torno a una obra teatral. El carácter literario, poético, de una obra de teatro del Siglo de Oro coexiste con su técnica teatral, con su necesidad de ser interpretada y vista por una audiencia. Es por ello por lo que, a la hora de investigar los aspectos fundamentales de una obra, la reconstrucción escenográfica debe ser tomada en cuenta como algo de vital interés, y hoy en día no existen maneras efectivas de llevar a cabo dicha reconstrucción de manera teórica.

Durante el curso de mi investigación para mi tesis doctoral, que se centra en la edición y estudio contextual de la obra de teatro *El Saco de Amberes*, una obra histórica atribuida a Calderón de la Barca, ha sido de gran interés el análisis no únicamente del texto que conforma el diálogo de la obra, sino también de las acotaciones. Esto se debe a que las acotaciones proporcionan detalles de interés con respecto a la interpretación de cada personaje, e instrucciones que indicarían al actor cómo moverse, reaccionar, o gesticular. Dichas instrucciones resultan altamente relevantes a la hora de analizar cómo se planteaban personajes históricos que estarían destinados a generar simpatía o rechazo en la audiencia, contribuyendo al carácter propagandístico de esta comedia flamenca de contexto bélico que representa los momentos previos al saqueo de Amberes de 1576 tras sofocar la actuación de tropas flamencas sublevadas contra el dominio español. Igualmente, las acotaciones han resultado iluminadoras frente a la calidad e inexactitud de las copias impresas que existen de la obra, y que no son cuidadosas en su recuento de personajes inicial; permiten llevar a cabo un recuento de los personajes que entran y salen, calculando la cantidad total de actores que hubieron de participar en la representación. Este recuento, las frecuentes menciones de las acotaciones de armas y objetos de atrezzo, una vez más proporcionando valiosos detalles, y los aspectos pragmáticos de las salidas y entradas de dichos personajes, es lo que lleva a la investigación por los derroteros de la reconstrucción escenográfica y del movimiento escénico de la obra.

Llegados a este punto, se genera la necesidad de visualizar el movimiento escénico de manera eficiente, que permita un análisis detallado del movimiento que marca el texto, y cómo se ve afectado por la configuración del

teatro en que la representación tuvo lugar. El carácter dramático del texto, pensado para su interpretación y no para su lectura, faltando así descripciones detalladas de los espacios o de los movimientos, deja en evidencia la importancia de estudiar los textos teatrales desde un punto de vista pragmático y no puramente literario. Sin embargo, los estudios teatrales que se enfocan en la representación están a menudo separados de los estudios literarios. Cabe preguntarse, pues, ¿cómo visualizar una reconstrucción teórica del movimiento escénico que ayude a la investigación? La propuesta se basa en encontrar un término medio entre un estudio literario del texto teatral, que se centraría meramente en el texto como pieza literaria, ignorando los aspectos pragmáticos de la representación; y un estudio interpretativo, llevado a cabo en la práctica sobre un escenario con unos actores. No se trata de la investigación a través de la práctica que se encuentra, por ejemplo, en el Globe Theatre de Londres; se trata de permitir al investigador una visualización accesible y práctica de la obra en la que se trabaje.

Lo primero que se descarta en el desarrollo de la rama de investigación en torno a la reconstrucción del movimiento escénico es el uso del papel. Resulta evidente enseguida que es imposible visualizar el proceso dinámico del transcurso de una obra teatral, reflejando el movimiento de los actores y cualquier cambio escénico, basándose en esquemas o diagramas a mano alzada. El uso de las tres dimensiones es fundamental en el teatro, puesto que es necesario contar con la posición sobre el escenario, incluyendo movimientos a proscenio, apariciones en galerías superiores como las que se encuentran en corrales de comedias, y posicionamientos de actores que no están por completo en escena y se mantienen junto al foro o pegados al paño, como aquellos que se esconden o los que están a punto de entrar en escena, pero aún no intervienen. El espacio escénico demanda tres dimensiones. Una reconstrucción física tridimensional, si bien útil para escenógrafos en muchos casos, resulta ineficiente en cuestiones de tiempo y esfuerzo en una investigación en torno a la obra teatral. La reconstrucción de un corral de comedias no es el motivo por el cual reconstruir y visualizar el movimiento escénico de una obra resulta un asunto complicado. Actualmente se dispone de las ilustraciones del Corral del Príncipe de Manuel Canseco publicadas en *La puesta en escena de los teatros comerciales del Siglo de Oro* de Ruano de la Haza, lo más fieles posibles a cómo se vería en realidad el corral, y es precisamente esta reproducción tridimensional la que podría resultar enormemente útil para desarrollar una herramienta digital de reconstrucción escénica para obras del teatro del Siglo de Oro. La rápida conclusión es el uso de una herramienta digital para visualizar los aspectos prácticos de la representación sin la necesidad de asistir a una producción teatral. Por el momento, tal herramienta, adaptada a las necesidades mencionadas, no parece existir. Sólo dos proyectos destacan por su potencial uso como herramienta aplicable a las humanidades digitales:

Stage Block

La herramienta Stage Block está diseñada con el objetivo de ayudar a dramaturgos a visualizar sus escenas, y a directores a comunicar a los actores su movimiento escénico. No obstante, Stage Block es una idea poco desarrollada, diseñada por un grupo de estudiantes de la universidad de Berkeley: Helen Xu, Julianna White, Eugene Lee, y BunJin Kim. El proyecto, parece, formó parte de un trabajo universitario y no parece que haya habido posterior desarrollo; los intentos de contactar con los creadores han sido en vano, por el momento. El principal usuario para el que está pensada Stage Block es el director de producciones teatrales a pequeña escala; y la necesidad a la que responde principalmente es la de coordinar el guion con el movimiento escénico en sus apuntes. Su diseño, planteado con una previsualización de su funcionamiento en la página donde se expone, es rudimentario y no provee una visualización tridimensional del espacio escénico (Figs. 1 y 2). No obstante, sus creadores sí comienzan a reconocer útiles funciones, como la posibilidad de visualizar la reconstrucción a vista de pájaro además de desde la perspectiva del patio de butacas, o la adición de una imagen con el decorado planeado para visualizar mejor el movimiento de los personajes. Stage Block permite también identificar a los personajes gracias a un código de colores y arrastrar sus iconos moviéndolos por el escenario, incluyendo bocadillos de texto

para sus intervenciones. Ofrece, finalmente, un botón que reproduce el movimiento registrado hasta ahora para una vista general de la obra, en el que aparecerán cuadro por cuadro las posiciones indicadas, junto con el texto en la parte superior, algo tal vez simplista y que soportaría únicamente textos sencillos

Stage Write

Al contrario que Stage Block, Stage Write es una herramienta completamente desarrollada que se ofrece bajo suscripción en su página web. Stage Write, sin embargo, no está orientada a directores de teatro en general ni a investigadores que deseen una ligera reconstrucción del movimiento escénico, sino a coreógrafos y directores de escena de musicales de gran escala. Mientras que comparte con Stage Block la codificación de colores de los iconos de los personajes, Stage Write únicamente ofrece una vista de pájaro del escenario, y tampoco funciona con animaciones. De nuevo, el usuario debería organizar un esquema que visualizaría la posición de los actores cuadro por cuadro (Figura 3). En lugar de reproducir la obra en la aplicación, el usuario puede descargarse un documento PDF con cada cuadro que conforma la obra, de manera que al moverse por el documento se puede visualizar el movimiento fluidamente. Además de poder incluir las dimensiones reales del escenario, Stage Write permite incluir en el plano elementos como escaleras, mobiliario, y atrezzo con el que contar en escena, aunque al sólo ofrecer vista de pájaro, no cuenta con dobles alturas, como la posición de un actor en la galería de un corral de comedias. Stage Write resulta útil a la hora de organizar la posición de los actores y sus entradas y salidas, pero no ofrece una buena reconstrucción visual del movimiento escénico para obras del Siglo de Oro.

Ambas aplicaciones se hallan en el camino de una buena y completa reconstrucción que le ayude al investigador a visualizar la obra estudiada, pero fallan en diversos aspectos, como un planteamiento más flexible de la posición de los actores y de la configuración del espacio escénico, una manera dinámica de asociar cada cuadro con su parte correspondiente del texto, una visualización más fluida del movimiento en los cuadros en conjunto, o una mayor personalización de los iconos de personaje.

En definitiva, esta comunicación expone la necesidad de una herramienta digital que ayude en la investigación de textos teatrales con la visualización y reconstrucción del movimiento escénico, arrojando luz sobre la valiosa información que a veces proveen las acotaciones sobre la interpretación y posición de los personajes, y permitiendo una mayor familiarización con la pieza estudiada sin la necesidad de depender de las producciones teatrales, sujetas además a las decisiones artísticas subjetivas de cada director, para hacerse una idea de la representación sobre el escenario. Respondiendo a las necesidades del investigador sobre los aspectos a reflejar en una herramienta así, como la visión tridimensional, la consciencia de los espacios escénicos del Siglo de Oro, o la posibilidad de visualizar la obra desde distintos ángulos, cabe la posibilidad de desarrollar una aplicación realmente útil para la investigación, enmarcada en el ámbito de las humanidades digitales. En torno a la investigación realizada en *El Saco de Amberes*, se está valorando y trabajando la creación de esta herramienta, utilizando la obra investigada como base para identificar las posibilidades y principales exigencias que el estudio de una obra presenta para una visualización útil y eficiente.

Apéndice: Imágenes



Figura 1

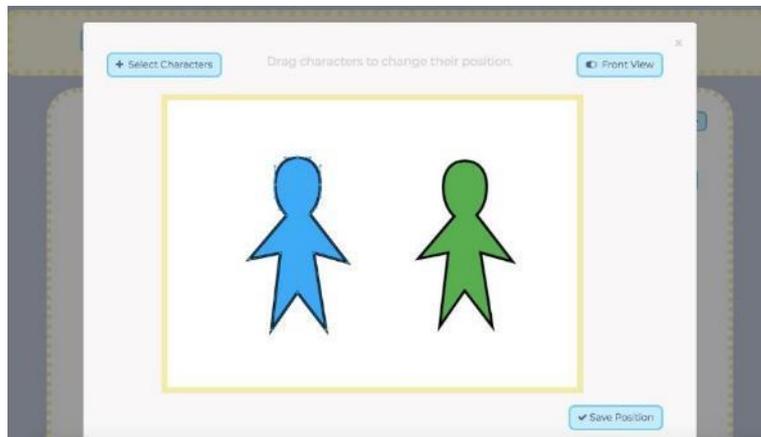


Figura 2

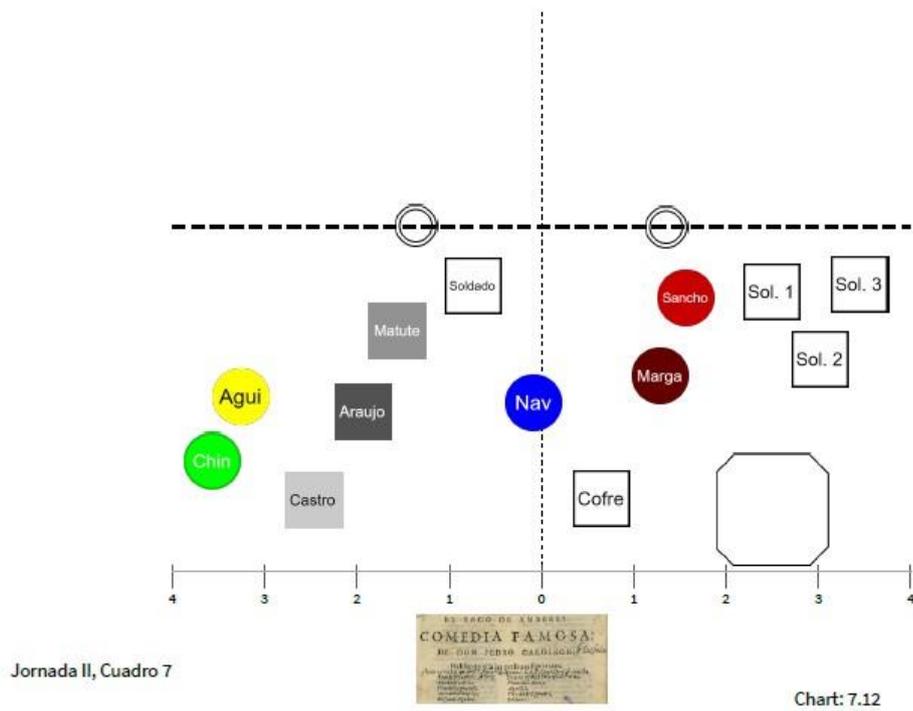


Figura 3

Bibliografía

de la Haza, J. M. R. La puesta en escena en los teatros comerciales del Siglo de Oro. Editorial Castalia, 2000.

Xu, H. Stage Block — A stage visualization and blocking tool for playwrights. [www.medium.com](https://medium.com/@helenxu_91186/stage-block-a-stage-visualization-andblocking-tool-for-playwrights-63151e7de76b) (2019).

Recuperado de: https://medium.com/@helenxu_91186/stage-block-a-stage-visualization-andblocking-tool-for-playwrights-63151e7de76b [20/04/2021]

Stage Write Software. <https://www.stagewritesoftware.com/> [20/04/2021]

Patrimonio Cultural y conocimiento: Propuestas para el diseño de una experiencia inmersiva basada en Realidad Mixta para el sitio arqueológico Cueva de las Manos (Argentina)

Marina Gutiérrez De Angelis - Universidad de Buenos Aires-ICA-AAV

Gorka López de Munain, UBA, Álava Medieval / Erdi Aroko Araba - UNED (Vitoria-Gasteiz)

Palabras clave: Patrimonio cultural digital, Realidad Mixta, Conocimiento, Divulgación científica, arte rupestre.

El proyecto Cueva de las Manos

Partiendo de la pregunta por la relación entre las imágenes del pasado y su conexión con el presente, desde el equipo de investigación del Área de Antropología Visual de la Universidad de Buenos Aires iniciamos en 2018 un trabajo de registro en el sitio Cueva de las Manos con vídeo 360°, con el objetivo de crear una experiencia inmersiva. En el año 2020 comenzó una segunda fase de trabajo que se vio inevitablemente interrumpida por la crisis sanitaria del Covid-19, pero que pronto podrá ser retomada con nuevos registros *in situ*⁸⁵. El proyecto tiene en este sentido un objetivo doble; por un lado, explorar las posibilidades que ofrece la Realidad Mixta (RM) para crear un entorno de aprendizaje basado en la exploración corporal y, por otro lado, profundizar en el desarrollo de un marco teórico interdisciplinario basado en el cuestionamiento de una serie de ideas acerca del concepto de arte rupestre.

Cueva de las Manos es uno de los sitios arqueológicos más importantes de Argentina y fue declarado Patrimonio de la Humanidad por la UNESCO en 1999. Está ubicado en el cañadón del Río Pinturas, entre las ciudades de Perito Moreno y Bajo Caracoles, en la provincia de Santa Cruz, y contiene rastros de hace diez mil años de grupos de cazadores-recolectores, entre los que se destacan manifestaciones de grafismo parietal con figuras de guanacos, escenas de caza, figuras antropomorfas y los famosos negativos de manos, que superan en total la cifra de 2.000. Las imágenes al aire libre se localizan en abrigos rocosos vulnerables no solo a la erosión del viento o el agua sino también a la erosión antrópica (fig.1-2). Aunque fue declarado Patrimonio de la Humanidad, cabe destacar que la gestión del patrimonio arqueológico en ocasiones establece un fuerte vínculo con la comunidad, reforzando el sentido de pertenencia regional (Guráieb y Frere 2008, p. 17). Pero en otras ocasiones, como sucede con muchos sitios arqueológicos en Argentina, existe una ruptura en la percepción de ese pasado que se traduce en la falta de creación de políticas adecuadas para su conservación.

Considerados estos aspectos, el proyecto propone diseñar una experiencia inmersiva para confrontar al visitante con las imágenes encontradas en la cueva y los imaginarios sobre la prehistoria. La introducción de la RV dentro del ámbito arqueológico hace necesario reflexionar sobre qué pueden aportar estas tecnologías al aprendizaje sobre el pasado. La RV puede

⁸⁵ El proyecto "Patrimonio arqueológico y realidad virtual en el sitio Cueva de las Manos, Santa Cruz, Argentina" comenzó en 2018 a partir del subsidio otorgado por el Fondo Nacional de las Artes. En 2020 fue otorgado un nuevo financiamiento con el proyecto UBACyT 20020190200046BA (2020-2022).

transformar un entorno de aprendizaje en un espacio de experimentación, ofreciendo múltiples puntos de vista a través de la interacción con objetos integrados. Los visitantes ya no se encuentran frente a una pantalla sino dentro de las imágenes. Esto transforma por completo los principios e ideas de base del diseño de nuevas experiencias en el campo del Patrimonio Cultural digital y exige también el desarrollo de bases teóricas aplicadas a este medio específico. Entendemos este proyecto como una contribución a la discusión en curso sobre el rol de las tecnologías digitales en el ámbito del patrimonio cultural.

Marco teórico-metodológico

Durante la primera fase del trabajo de campo, surgieron una serie de hipótesis que orientaron la investigación hacia la discusión del concepto de arte rupestre, optando por la expresión de grafismo parietal para evitar toda alusión a concepciones modernas aplicadas a configuraciones visuales del pasado. El problema del arte rupestre no es simplemente conceptual, sino que también expresa el papel que juegan, en el campo de la arqueología, las imágenes técnicas (Bredenkamp et al., 2015b). Como señala Palacio Pérez, la diversidad de ideas sobre la Prehistoria en el último tercio del siglo XIX generó una gran variedad de discursos sobre el arte paleolítico, unificados bajo el concepto de arte (Palacio Pérez, 2017, p. 35). Frente a las imágenes realizadas en las cuevas, por lo general se busca descifrar lo que representan, ilustran o simbolizan. Como plantea W.J.T Mitchell (2017), pero también autores como Alfred Gell (2016), Conkey (1978), White (2003) o Moro Abadía y González Morales (2007), el estudio de lo pictorial no debe sustentarse primariamente en la decodificación de significados de las imágenes o en la comprensión de una intención detrás de ellas. La Cueva Virtual es pensada como un espacio de pensamiento que desencadena procesos de reflexión, puesto que mirar no es una competencia sino una experiencia (Gutiérrez De Angelis et al, 2019).

A partir de esta premisa, retomando el concepto de *Bildakt* (Bredenkamp, 2015) y los aportes desde las *experimental aesthetics* (Gallese, 2015), como base de una teoría sobre la percepción y sobre nuestra relación con diferentes medios de la imagen, el proyecto propone explorar la capacidad de la RM para estimular la experiencia de la sensorialidad espacial y háptica de la imagen, así como una experiencia de conocimiento en primera persona. Actualmente la tendencia se dirige hacia el uso de formas más atractivas y naturales de interactuar con los contenidos digitales (Gonizzi et. al, 2015). En ese sentido, las tecnologías hápticas permiten crear entornos que no alejan al usuario/a del sitio, sino que lo invitan a explorar un nuevo concepto de interfaz donde la telepresencia se articula para crear una experiencia de aprendizaje y conocimiento. El proyecto incorpora el uso de tecnologías de bajo costo como Oculus Quest 2, combinadas con sensores de movimiento para manipulación de objetos modelados en 3D, vídeo y sonido 360°, junto con software especializado para arqueología como *DStretch*. El uso de estas tecnologías permite registrar aspectos fenomenológicamente significativos de la cultura visual que no pueden ser transmitidos por otros medios de imágenes, involucrando la participación física del público.

Tomando en consideración los resultados de diversos estudios sobre recepción de RV en el ámbito museográfico (Pujol, 2007; Forte, 2006; Belaën, 2003), este proyecto reflexiona sobre el modo en que el uso de la RM ofrece nuevas herramientas para promover el aprendizaje a través de entornos inmersivos. Uno de los principales elementos que surge de estos estudios es aquel relacionado con el aprendizaje y la centralidad que se otorga al lenguaje en ese proceso. Si, por un lado, la RV es valorada como medio capaz de generar una sensación de presencia a través de la inmersión, no se la considera relevante a la hora de la riqueza de la información que ofrece en términos discursivos. Este trabajo sostiene la hipótesis de que la RM permite superar las dicotomías

entre físico-virtual, visual-textual, control-interacción, creando espacios capaces de integrar todos esos niveles a través de una experiencia en primera persona.

La perspectiva de la *experimental aesthetics*, propone un modelo relativo a la percepción basado en el concepto de encarnación simulada (Gallese y Guerra, 2015, p. 72). La percepción contiene un componente sensorio motor que nos permite sentir la imagen de una forma corporizada, en virtud de la representación motora que produce la imagen mediante la simulación, ya que la simulación encarnada posibilita su vivencia (Freedberg y Gallese, 2007). Esta simulación puede ser resaltada por la RM, puesto que se produce una implicación corporal que incluye nuestra acción física en relación con las imágenes (Pennisi et al., 2019). La intervención de la habilidad motora y la mano capaz de interactuar con la pantalla construyen una experiencia completa donde la imagen es resultado de un acto motor intencional (Gallese y Guerra, 2015, p. 263). Los entornos digitales son así un extraordinario amplificador tecnológico de la multimodalidad natural de la visión que define una nueva forma de *Bildakt*, en la que el papel activo de las imágenes es ofrecerse como objeto de un deseo físicamente tangible (Gallese y Guerra, 2015, p. 266). La relación cuerpo-cerebro entendida como un sistema, ha demostrado cómo la corporalidad juega un papel decisivo en los procesos cognitivos, a diferencia de las perspectivas vinculadas al cognitivismo clásico, según las cuales nuestras sensaciones y emociones adquieren sentido sólo cuando se convierten en objeto de una representación en términos lingüísticos (Gallese y Guerra, 2017, p. 74). Las imágenes son enactivas en tanto son capaces de *embody information* y nuestro cuerpo es un complejo sensorio-motor perceptivo (Krois 2011, p.17). Las imágenes no son ilustraciones, sino cosmos que ofrecen una semántica creada de acuerdo a sus propias leyes (Bredenkamp, 2017, p. 17). Desde este punto de partida, el trabajo en el sitio Cueva de las Manos nos llevó a reflexionar acerca del modo en que estas imágenes del pasado pueden conectarse con las imágenes de nuestro presente y nos obligan a revisar una serie de ideas sobre ese pasado y el carácter de lo visual.

La Cueva Virtual

La necesidad de crear una Cueva Virtual surge por varios motivos. La excepcionalidad del sitio arqueológico de la Cueva de las Manos lo convierte en un caso de estudio ideal para implementar soluciones tecnológicas que promuevan un acceso global a su conocimiento. Hasta la fecha, el desarrollo de propuestas innovadoras para favorecer el conocimiento del patrimonio por medios digitales ha sido escaso, parcial y se ha concentrado en grandes instituciones como museos (Jung, 2016). A pesar de la incipiente implantación de aplicaciones de realidad virtual en estos espacios museísticos, se ha demostrado el creciente interés de los visitantes.

En cuanto al desarrollo de la Cueva Virtual, son múltiples los retos a los que nos enfrentamos. En primer lugar, es necesario tener en cuenta la ubicación de las figuras en el paisaje. El sitio se encuentra en un cañadón, resultado de la erosión de las aguas del deshielo de las glaciaciones. El sitio se encuentra a 88 m de altura sobre el cauce del río mientras que la visita tiene un recorrido total de aproximadamente 1500 m. Estos elementos lo convierten en un sitio arqueológico de difícil acceso (fig. 3).

El primer elemento que contempla el diseño es el de la experiencia espacial, a partir del concepto de telepresencia, entendido como la posibilidad de encontrarse presente en un espacio distante a través del acceso a espacios virtuales. El diseño permite al usuario/a elegir por sí mismo, en qué dirección y en qué tipo de perspectiva, quiere ver el espacio, ya que el espectador ya no está simplemente frente a la imagen sino que es parte de ella (Grau 2003, 234). Esto crea una orientación inmediata, activando la participación a partir del

compromiso físico del cuerpo. Con ese objetivo, el diseño ubica al visitante en medio del cañadón en una posición flotante entre ambos frentes, lo que permite acceder a un punto de vista físicamente inaccesible. Esto busca ofrecer una comprensión de las dimensiones del territorio y el sitio así como de la forma, posición, distribución y tamaño de cada sector, confrontando al visitante con la ubicación de las figuras en relación a su propio cuerpo y posición.

El segundo elemento es la interacción física. Dentro de la experiencia con RM, el usuario no se encuentra inmóvil frente a una pantalla, sino inmerso en un entorno que solicita sus movimientos. La imagen estimula mucho más que un dato visual, estimula una *haptic visuality* que nos permite establecer una relación carnal con la trama de la imagen (Gallese y Guerra, 2015, p. 233). El movimiento con la mano con los sensores de Oculus Quest 2 convierten el movimiento del usuario en un movimiento efectivo. La experiencia táctil puede ofrecer al usuario la posibilidad de interactuar con las imágenes al diferenciarlas, recortarlas, explorarlas y animarlas. El diseño de la interfaz explora la cualidad táctil de la percepción. Con ese objetivo las imágenes fueron tratadas a partir del uso *DStretch*, para realzar pinturas y obtener dobles de las figuras a partir de operaciones de mejoramiento de contraste (figs. 4). Las imágenes obtenidas a través de estos procesos, son la base para crear modelos tridimensionales que permiten “arrancar” de la roca estas figuras, así como recrear la experiencia de su superposición a lo largo de los siglos, para que el usuario pueda ver detalles, rotarlas y recorrerlas en su emplazamiento físico (Gutiérrez De Angelis, 2021). El modelado 3D permite profundizar en esta interacción con la mano, guiando la mirada sobre detalles a partir del movimiento de la mano y la exploración de su forma y dimensión. La experiencia motriz permite comprender la importancia de la relación entre las figuras, los materiales y técnicas con las que se realizaron y las actividades sociales, accidentes geográficos y circuitos de movilidad de los que no pueden ser separadas.

Conclusiones

El proyecto aún se encuentra en la primera etapa de desarrollo. Nuestro principal objetivo en esta primera etapa fue profundizar en la comprensión del modo en que los medios virtuales permiten diseñar experiencias que apelan a la comprensión física de los entornos, a la exploración detallada de objetos, a la manipulación de objetos y a las sensaciones hápticas y auditivas. En ese sentido, retomando a O’Sullivan (2005), en lugar de aislar las imágenes y ofrecer interpretaciones, la Cueva Virtual las considera como *objects of encounter* antes que *objects of recognition*, puesto que el encuentro produce un corte que nos hace reflexionar más allá de lo que vemos. Estos entornos pueden potenciar la construcción del propio conocimiento a través de la experiencia sensorio-motora y afectiva de la realidad.

Con este trabajo hemos presentado una serie de reflexiones que guían el diseño de un modelo para experiencias inmersivas basado en la exploración de las capacidades de la RM para crear entornos de conocimiento a través de la participación de nuestro cuerpo. La interacción gestual permite crear una experiencia basada en el modo en que la imagen, la espacialidad y lo háptico se articulan a través de la experiencia virtual. Esto hace necesario el desarrollo de bases teóricas aplicadas a este medio específico, puesto que los entornos virtuales no pueden ser pensados a partir de las lógicas propias de otros medios. Estas tecnologías plantean nuevas estrategias metodológicas y expresivas para la protección, difusión y defensa del patrimonio cultural.

Figuras



Figura 1. Cañadón del Río Pinturas. Fotografía Digital. Santa Cruz, Argentina. 2018. © Área de Antropología Visual.



Figura 2. Detalle de los aleros cubiertos de manos y animales, Sitio II, Sector b, Cueva de las Manos, Santa Cruz, Argentina. 2018. © Área de Antropología Visual.



Figura 3 Detalle acceso a la cueva, Sitio II, Cueva de las Manos. 2018. © Greta Winckler, Área de Antropología Visual.



Figura 4 Cueva de las Manos, Sitio II, Sector c. 2018. Fotografía Digital y fotografía procesada con DStretch. Santa Cruz. Argentina. © Área de Antropología Visual.

Bibliografía

- Belaën, F. (2003). "L'immersion Au Service Des Musées De Sciences", *Ichim 03: Les institutions culturelles et le numérique*, París, Ecole du Louvre.
- Bredenkamp, Horst. (2015). *Immagini che ci guardano. Teoria dell'atto iconico*. Milán: Raffaello Cortina.
- Bredenkamp, Horst et al. (2015b). *The Technical Image: A History of Styles in Scientific Imagery*. Chicago: University of Chicago Press.
- Conkey, Margaret. (2010). "Images without Words: The Construction of Prehistoric Imaginaries for Definitions of 'Us'". *Journal of Visual Culture*, 9 (3), 272-83.
- Forte, M; Pescarin, S; Pujol Tost, L (2006). "2VR applications, new devices and museums: visitors's feedback and learning. A preliminary report", *The 7th International Symposium on Virtual Reality, Archaeology and Cultural Heritage VAST*.
- Gallese, Vittorio. (2017). *Visions of the Body: Embodied Simulation and Aesthetic Experience. Aisthesis. Pratiche, Linguaggi E Saperi dell'estetico*, 10(1), 41-50.
- Gallese, Vittorio, y Michele Guerra. (2015). *Lo schermo empatico. Cinema e neuroscienze*. Milán: Raffaello Cortina.
- Gell, Alfred. (2016). *Arte y agencia. Una teoría antropológica*. Buenos Aires: Sb.
- Gonizzi Barsanti, S; Caruso, G.; Micoli, L.L.; Covarrubias Rodríguez, M.; Guidi, G. (2015). "3D Visualization of Cultural Heritage Artefacts with Virtual Reality devices", *ISPRS International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, XL-5/W7, 2015, 165-72.
- Gutiérrez De Angelis, M. (2021) Antropología visual expandida: grafismo parietal y realidad mixta en el sitio arqueológico Cueva de las Manos (Argentina)", *Revista de Antropología Visual*, Número 29 – Santiago, Chile, pp. 1-20 pp
- Gutiérrez De Angelis, M; Winckler, G; Bruno, P; Guarini, C (2019). "Rethinking Paleolithic Visual Culture", *View. Theories and Practices of Visual Culture* 25.
- Grau, Oliver. (2003). *Virtual Art. From Illusion to Immersion*. MIT Press Leonardo Book Series.
- Guráieb, Ana Gabriela; Frere, María Magdalena. (2008). *Caminos y encrucijadas en la gestión del patrimonio arqueológico argentino*. Buenos Aires: Facultad de Filosofía y Letras de la Universidad de Buenos Aires.
- Jung, Timothy; tom Dieck, M. Claudia; Lee, Hyunae; Chung, Namho. (2016). "Effects of Virtual Reality and Augmented Reality on Visitor Experiences in Museum", *Information and Communication Technologies in Tourism*, Inversini A., Schegg R. (eds), Springer, Cham.
- Krois, J.M. (2011) "Enactivism and Embodiment in Picture Acts. The Chirality of images", *Sehen und Handeln*, H. Bredenkamp y J.M. Krois (eds.), Berlín: Akademie Verlag.
- Mitchell, W. J. T. (2005). *What Do Pictures Want?: The Lives and Loves of Images*. Chicago: University of Chicago Press.

- Moro Abadía, Oscar, y Manuel R González Morales. (2013). "Paleolithic Art: A Cultural History". *Journal of Archaeological Research*, 21, 269-306.
- O'Sullivan, S. (2005). *Art Encounters. Deleuze and Guattari: Thought beyond Representation*. Basingstoke, UK: Pallgrave Macmillan.
- Palacio Pérez, Eduardo. (2017). *El arte paleolítico. Historia de una idea*. Santander: Ediciones Nadir.
- Pennisi, Antonio et al. (2019). "The Extended Theory of Cognitive Creativity: Interdisciplinary Approaches to Performativity", *Perspectives in Pragmatics, Philosophy & Psychology*, 23, Springer Nature.
- Pérez, J. M. (2015). "e-Art realidad aumentada y arte rupestre", *XXXIV Reunión de Asociaciones y Entidades para la Defensa del Patrimonio Cultural y su Entorno*, 1-4.
- Pujol-Tost, Laia. (2007). "Exploring the suitability of virtual reality interactivity for exhibitions through an integrated evaluation: the case of the Enane Museum". *Museology*, 4, 81-97.
- White, Randall. (1992). "Beyond Art: Toward an Understanding of the Origins of Material Representation in Europe", *Annual Review of Anthropology*, 21, 537-64.

Plenaria 5

Amparo Alonso Betanzos

Las Humanidades en la era de la Inteligencia Artificial: cooperando para modelar sostenibilidad

Moderador: César González-Pérez · *Asiste:* Alexandra Iglesias Cortés

La Inteligencia Artificial no solamente es el cambio tecnológico más importante desde la invención de la electricidad, también ocasiona un cambio social. Nuestro objetivo no es solamente desarrollar buenas herramientas, sino desarrollar herramientas buenas para las personas, en esta IA que la Unión Europea preconiza confiable, segura, sostenible y al servicio de las personas. Y para poder conseguirlo es necesario contar con profesionales de las diferentes ramas de las Humanidades. Como ejemplo, comentaremos el desarrollo de un modelo basado en agentes inteligentes que, con las contribuciones de sociólogos y psicólogos, puede ayudarnos en el diseño de mejores políticas medioambientales, que faciliten la adhesión de la ciudadanía y contribuyan a dar confianza en la tecnología.