

Historiografía digital



FICHA BIBLIOGRÁFICA

MIRELLA ROMERO RECIO, M^a JESÚS COLMENERO (Eds.), Madrid, Instituto de Historiografía Julio Caro Baroja. Universidad Carlos III de Madrid, 2016, 219 págs. ISBN: 978-84-16829-01-9.

María Martín de Vidales García **Universidad Carlos III de Madrid**

Historiografía digital. Proyectos para almacenar y construir la historia. Este título constituye el número cuatro de los Anejos de la Revista de Historiografía editado por el Instituto de Historiografía Julio Caro Baroja. Esta vez, una edición coordinada por las Profesoras Mirella Romero Recio y María Jesús Colmenero de la Universidad Carlos III de Madrid. Entre ellas, Romero Recio ha sido la investigadora principal del proyecto Almahisto¹ y ha conseguido,

1. Este proyecto de investigación financiado por el MINECO tiene como título “El Almacén de la Historia. Repositorio de Historiografía española (1700-1939)” (HAR2011-27540).

junto a un equipo colaborador de investigadores, el objetivo propuesto al crear un Repositorio de historiografía en red que facilite el trabajo de investigación.

Se estructura en doce capítulos redactados por varios autores y en los cuales se ejemplifica, a través de varios casos de estudio, una metodología adecuada para las humanidades digitales. Por otro lado, la introducción es tajante. Realizada por las Coordinadoras, se ponen sobre la mesa las pautas de mayor interés correspondientes al trabajo. Ante todo, el reconocimiento de las humanidades digitales. Hasta hace muy poco tiempo, las humanidades y el mundo digital parecían ocupar puestos integrados en ámbitos totalmente distanciados. Sin embargo, esta situación no es compatible en una sociedad que va de la mano del mundo digital. De esta forma, la incorporación digital se presenta como esencial a la hora de crear un repositorio en el cual se “almacenen” los documentos necesarios para el estudio historiográfico. Ha sido un proceso complejo, pero desde luego, presenta un carácter novedoso que mejora la investigación.

Los doce capítulos que conforman el libro son dispares pero reflejan con precisión el carácter del proyecto. De ellos, se puede extraer una idea sobre la metodología de aplicación del mundo digital al estudio historiográfico. En resumen, cada capítulo proyecta uno de los pasos que conforman dicha metodología. Los primeros son la digitalización de documentos y el volcado de datos teniendo en cuenta su diverso carácter y obteniendo, de esta forma, la separación por áreas. Una vez que se poseen los datos, la gestión de los mismos a partir de estructuras adecuadas y el enlace de estos a través, por ejemplo, de la investigación comparada, el uso de Linked Data y el tratamiento de textos. Por último, se conseguirá la creación de bases de datos con áreas diversas pero puntos en común que favorezcan el enriquecimiento de la investigación.

De forma más detallada y para empezar, se presenta la colaboración con la Biblioteca de la Universidad Carlos III a través de la cesión del Repositorio Institucional e-Archivo. No fue inminente pues la Biblioteca universitaria se encontraba en un proceso de renovación del mismo pero la espera merecía, pues el proyecto se aseguraba su mantenimiento en el futuro. Por otro lado, e-Archivo ofrece otras posibilidades al agilizar el proceso y facilitar la visualización de los investigadores. Se utiliza DSpace como Software y este asocia una URI a cada ítem asegurando la preservación digital de los archivos. Sabiendo que se recurriría a e-Archivo, era lógico establecer los parámetros jurídicos implicados. Santiago Mediano explica de forma clara en el segundo capítulo qué es Almahisto jurídicamente hablando². Hace hincapié en tener en cuenta al menos tres aspectos implícitos en el proyecto: la propiedad intelectual asociada a los contenidos utilizados, la relación entre los aportadores y el proyecto para conocer a quién había que solicitar la autorización de uso, y por último, las condiciones que poseía la plataforma en relación a la disposición de códigos.

2. El Proyecto sería “una plataforma compuesta por un conjunto de herramientas tecnológicas en código abierto, en las cuales se integra una base de datos que permite a los usuarios acceder a través de un sistema de búsqueda a los documentos historiográficos relativos al periodo de 1700 a 1939 inclusive, de la misma plataforma o permitiendo la visualización de documentos alojados en otras bases de datos a través de enlaces”, según señala Mediano en pág. 20.

El desarrollo de los procesos en proyectos como el que nos atañe suele ser dificultoso y evoluciona de forma paralela a las necesidades que van surgiendo. María Jesús Colmenero-Ruiz y Victoria Rasero explican el proceso que se siguió en este caso. Una vez delimitadas las fechas, que no son fruto del azar, sino que responden a un cambio en la Historia de la historiografía, se establece una metodología de trabajo dividida en dos fases. La primera de ellas ya se había cumplido, elegir un modelo de repositorio: e-Archivo. La segunda se ocuparía de la incorporación de los documentos al repositorio pero se desmembró en varias sub-fases a las que habría que hacer frente: diseño, desarrollo y adaptación del Software, incorporación del esquema de metadatos Almahisto y para finalizar, el diseño de plantilla de entrada de datos. El trabajo multidisciplinar que se proponía, sin embargo, potenció un refuerzo en las relaciones de los diferentes investigadores.

El volcado de datos es una tarea fundamental en este proceso. Implica una división por temas que en historia suele corresponder a la cronología, ámbito geográfico, etc. Teodoro Crespo Mas analiza el proceso a través de los documentos de Historia Moderna que se han volcado como datos en el repositorio. La digitalización es esencial para dicho volcado y además, facilita el acceso a los datos y por tanto, el estudio. No es extraño que surjan problemas debido al carácter ambiguo de los documentos, se mezclan los contenidos de diversos apartados como advierte Jesús Salas Álvarez en relación al ámbito arqueológico. Todo este proceso es ejemplificado por el estudio llevado a cabo por las profesoras León Gómez y Gil Fernández al recopilar las fuentes en relación a los edificios lúdicos en el siglo XVIII en un ámbito geográfico concreto: Valencia, Andalucía y Extremadura.

Como estamos percibiendo el volcado de datos genera mucha información que debe gestionarse para “no morir ahogado por los datos”, como dice Sancho Caparrini³. Para ello, el Profesor propone la adopción de una estructura adecuada y con riqueza que sea capaz de tratar la información no estructurada para extraer todo su valor. Por el carácter estructural, semántico y contextual que poseen las humanidades, las estructuras suelen ser más débiles. Se propone, para combatir este estado, el uso de grafos multirrelacionales, que por las características que poseen podrán transformar la información almacenada en nuevo conocimiento. Las estructuras utilizadas pueden favorecer diferentes métodos de investigación. En el siguiente capítulo se destaca el sistema de investigación comparativa pues permite el enlace de datos mediante relaciones significativas proporcionando resultados novedosos. Como ejemplo de este método propone la iniciativa que supuso el proyecto Linked Open Data (LOD) al publicar datos enlazados en la web mediante URIs. Para Boer, Meroño-Peñuela y Ockeloen, el uso de Linked Data es muy beneficioso en la investigación histórica y lo justifican a través de tres ejemplos que se corresponden con tres proyectos de investigación histórica digital: CEDAR Project, The Dutch Ships and Sailors Project y The BiographyNet Project.

Pero toda esta intención de mejorar la metodología de investigación en Historia para ser capaces de obtener resultados más precisos y facilitar el trabajo al investigador no debe ceñirse únicamente a un ámbito universitario donde los investigadores ya hayan adquirido un alto nivel de madurez. Los Profesores Lasala Navarro y Gudín de la Lama proponen el

3. Sancho Caparrini en pág. 97.

trabajo con repositorios digitales para acercarse a la historia desde las primeras etapas formativas. Para ellos, lo importante del estudio de la historia es el uso de las fuentes, seguir un esquema de trabajo que podría corresponder al siguiente: recogida de información, hipótesis, trabajo con las fuentes y explicación histórica. A través de una propuesta de trabajo con los alumnos del Grado de Maestro de Educación Primaria llamada *La Guerra Civil en tu localidad* pueden observar la falta de metodología por parte de los alumnos al no darle la importancia necesaria al tratamiento de las fuentes. Por esta razón, proponen que se adquiera una mayor concienciación por parte de los profesores, sobre el estudio de la historia en los centros de formación a través de herramientas digitales.

Para gestionar los datos volcados, también hay que tener en cuenta el tratamiento de textos automatizado. A través de procesos de reconocimiento textual se permite un análisis de los diferentes niveles del lenguaje. Sin embargo, y en cuanto a lo que respecta a los Named Entity (NE) o nombres propios, el reconocimiento se convierte en una labor más compleja. Cada sistema de reconocimiento y clasificación de NEs hace uso de varias técnicas de identificación. Iglesias Moreno y Sánchez-Cuadrado aconsejan optar por un software libre- como Freeling y Python- para el procesamiento de reconocer y clasificar los datos. Aplican esta metodología al trabajo realizado por el grupo de investigación QUASTEIO, con el libro *Becerro de las Behetrías de Castilla* (1352).

Por último, se generan grandes bases de datos entre las que destacan los catálogos de piezas gestionados con herramientas digitales en relación con disciplinas como la arqueología o la historia del arte. Sin embargo, hay que tener cuidado con la gestión de estas bases de datos para no caer en el error de formar grandes e inabarcables catálogos estáticos. Es mucho más interesante que las herramientas digitales no formen simples almacenes, sino que sirvan de utilidad al investigador para el estudio. Otra dificultad que puede surgir debido a la gestión masiva de datos es la producción de errores, aunque podría solventarse con la mejora del diseño del software y el control de calidad por parte del investigador, para que los resultados no se distancien de la realidad. Pedro Luengo propone el estudio tridimensional como herramienta que permite la profundización tanto en la arquitectura como en la arqueología. Sin embargo, no considera esencial obtener datos específicos de cada objeto de estudio para rellenar una base de datos, sino poner en relación los diferentes resultados para abrir nuevas líneas de investigación.

De esta forma y desde mi punto de vista, la publicación por un lado facilita al lector, la adquisición de diversas herramientas digitales capaces de mejorar el estudio historiográfico, mientras que por otro lado, analiza el estado en el que se encuentran las humanidades digitales. Aunque estas, cada vez sean más frecuentes en los ámbitos de estudio, están experimentando aún en diferentes vías de aplicación. Una publicación muy acertada pues es esencial que la comunidad científica se familiarice con la gestión digital, apueste por el uso de herramientas digitales en los procesos de estudio, que se estrechen las relaciones y que, por supuesto, aporten nuevas posibilidades al mundo humanístico.