

despacha habitualmente. He insistido en que las bibliotecas virtuales constituyen un excelente recurso virtual e icónico para proporcionar materias primas y también metodología a las HD, en particular he señalado cómo, tras el OCR de todo tipo de textos, las bibliotecas codifican el resultado mediante un estándar *de facto* internacional denominado ALTO. He tendido un puente a la otra orilla, donde se encuentran los humanistas digitales, señalando lo factible que resulta combinar ALTO y TEI, haciendo así un pleno al quince, es decir apostando por que TEI será el estándar, y además debería serlo, utilizado por la comunidad de las HD en España.

Por último, he hablado de la última gran revolución que ha experimentado la web, es decir la tecnología de datos abiertos enlazados, que va a permitir, entre otras cosas, enriquecer semánticamente textos codificados con TEI, mediante vocabularios de

valores, conjuntos de datos y conjuntos de elementos de metadatos que ya utilizan las bibliotecas y en especial las virtuales, fruto de las recomendaciones del *Informe Final del Grupo Incubador de Datos Abiertos Vinculados de Bibliotecas*. Creo que, finalmente, y con independencia del desarrollo propio de las HD, las bibliotecas virtuales estarán ahí para proporcionar la materia prima, como las proporcionaron en su día las bibliotecas tradicionales, que tanto juego dan y seguirán dando en años futuros. El humanista digital siempre ha contado con la tarea de los bibliotecarios, archiveros y museólogos para auxiliarle; no les debe de quedar ninguna duda a las HD de que las bibliotecas digitales estarán siempre a su disposición, proporcionando como siempre los materiales que la investigación requiere.

X. AGENJO /  
LAS BIBLIOTECAS  
VIRTUALES  
ESPAÑOLAS...

X. A.—FUNDACIÓN LARRAMENDI

## ANTONIO ROJAS CASTRO / DE LOS DATOS AL METAMEDIO

Pocos cambios han sucedido con tanta rapidez como los experimentados por nuestra sociedad desde mediados de los noventa, hasta el punto que, integrados completamente en nuestra vida diaria, hace falta señalarlos para que se vuelvan perceptibles. La tecnología WWW inventada por Tim Berners-Lee, la liberación de la gestión de internet puesta en manos de la comunidad de internautas —pero también de las empresas—, y, más recientemente, la generalización de la banda ancha, las redes inalámbricas o el móvil inteligente, han dado pie a que el sociólogo Manuel Castells se refiera con la expresión «sociedad-red» al modo en que nos comunicamos y relacionamos en la actualidad. No es mi objetivo determinar aquí qué fue antes —los cambios tecnológicos o los sociales— pero es innegable que nuestras relaciones se basan cada vez más en intereses, valores y proyectos personales. Ahora bien, esto no significa el fin de la comunidad ni el aislamiento; al contrario, en la última década la sociabilidad se ha reconstruido «en forma de individualismo y comunidad en red a través de la búsqueda de personas afines, en un proceso que combina interacción virtual (*online*) con interacción real (*offline*), ciberespacio con espacio físico y real» (Castells, 2013: 136).

A una escala menor, como si fueran un microcosmos, las HD tienen mucho de «sociedad-red». En su definición minimalista —uso de metodologías y herramientas informáticas—, las HD han sido practicadas desde hace varias décadas en el hispanismo con proyectos como ADMYTE, *PhiloBiblon*, la *Biblioteca Virtual Miguel de Cervantes*, el CORDE o los recursos digitales creados por el grupo SIELAE (Rojas Castro, 2013); en su definición maximalista —estudio de la cultura digital y exploración del medio electrónico—, las HD son un fenómeno nuevo que solo se entiende si tenemos en cuenta dos factores: por un lado, nuestro conocimiento práctico, casi íntimo, sobre las posibilidades de almacenamiento, multimedia e interacción del medio digital; en otras palabras, el nacimiento de nueva conciencia sobre el impacto de la tecnología e internet en nuestras vidas y, en concreto, en los objetos de estudio propios de las disciplinas humanísticas; por el

otro, la penetración de las redes de comunicación que conjugan el individualismo y la colaboración. De hecho, no es casual que la emergencia de las HD se haya producido en forma de redes, organizaciones, consorcios y asociaciones que buscan promover tanto proyectos individuales como conectar a investigadores con intereses comunes.

El uso de metodologías y herramientas informáticas en la universidad tampoco es una novedad. Las famosas Tecnologías de la Información y la Comunicación (TIC's) llevan entre nosotros muchos años y han afectado mayoritariamente al modo de enseñar los contenidos docentes. La propuesta de las HD va más allá de la adopción de recursos didácticos en formato digital porque no solo modifica la transmisión o comunicación de contenidos, sino —y ante todo— la producción del conocimiento. Al mismo tiempo las HD plantean una reflexión más profunda sobre la tecnología y los medios, es decir, son —o deberían ser— tan instrumentales como especulativas. En lo que sigue ilustraré algunas de estas ideas, esbozaré el proceso de trabajo de los humanistas digitales y justificaré el valor que añaden las HD a las disciplinas humanísticas tradicionales.

### Cómo hacer Humanidades Digitales

En la introducción a *Radiant Textuality*, Jerome McGann (2001: 5) sugiere que las HD cobraron fuerza en Estados Unidos tras el auge del New Criticism y las teorías postestructuralistas sobre el hipertexto: los académicos se dieron cuenta de que las ideas en torno a la intertextualidad tenían un fundamento real tras la invención de la WWW y podían llevarse a la práctica trabajando con documentos reales en bibliotecas y archivos. Así, las HD, en su definición minimalista, han sido percibidas como una disciplina que importa de las Ciencias unas metodologías cuantitativas y que las aplica al estudio de la literatura o del arte. Se trata, a mi juicio, de una percepción errónea porque las Humanidades siempre han





A. ROJAS CASTRO /  
DE LOS DATOS  
AL METAMEDIO

sido disciplinas con un componente empírico, además de crítico: la lectura, interpretación o análisis de una obra —literaria o artística— no puede sino basarse en la observación de evidencias encontradas en un texto oral, escrito, pictórico, semiótico o audiovisual.

Estas evidencias —o datos, como se llaman ahora— son la materia prima de los proyectos de HD, pues sin ellas el ordenador no puede procesar nada. Al iniciar una edición digital, una colección de imágenes o una base de datos, se necesitan —valga la redundancia— datos, es decir, información discreta y manipulable por el ordenador como nombres de personas, de personajes o de lugares, años o períodos temporales, títulos de obras, eventos en forma de enunciados, etc., pero también textos completos, imágenes o sonidos en formato digital. El modo en que se obtienen esos datos difiere en cada proyecto y depende del objeto de estudio. Es posible que hayan sido recogidos por alguna organización y por tanto estén disponibles en su web en algún formato abierto como CSV, RDF o SPARQL —nunca en PDF— y a la espera de que alguien los utilice y les dé sentido; esto ocurre sobre todo cuando se obtienen mediante encuestas o cuando las instituciones apuestan de manera deliberada por la transparencia y el acceso abierto en general y en concreto por los datos abiertos enlazados. También es posible que nuestro objeto de estudio ya haya sido digitalizado y que por tanto no sea necesario digitalizar los documentos, sino obtener el permiso para utilizarlos. En otras ocasiones será necesario, tras escanear el documento y obtener una imagen, captar la información contenida en él mediante un algoritmo OCR con el fin de generar un archivo de texto.

La etimología latina de la palabra «dato» —‘lo que se da’— no debería hacernos caer en la trampa de creer que los datos nos son «dados», es decir, que son inequívocos e independientes del investigador como si fueran unas partículas que flotan en el aire. Lo ha dicho en más de una ocasión Johanna Drucker (2011): «*All data is captaw*». En otras palabras, los datos son captados, percibidos y recogidos en función de los intereses del investigador y de la pregunta de investigación que fundamenta su proyecto; en este sentido, los datos son siempre una selección o reducción del fenómeno que se pretende estudiar (Schöch, 2013). Por ello, tener un conocimiento de la tradición humanística resulta indispensable. De lo contrario nuestra mirada no sabrá reconocer las particularidades del objeto de estudio y seguramente se plantearán preguntas irrelevantes o ya contestadas sin metodologías informáticas.

Una vez tenemos los datos en formato digital conviene estructurarlos para hacer explícitas las relaciones existentes entre ellos y así responder la pregunta que mueve la investigación. Conviene insistir: sin pregunta o hipótesis no hay HD. Existen distintas maneras de estructurar o «modelar» los datos: la codificación o etiquetaje de un texto con TEI es una manera de estructurar información comprensible por el ordenador y, además, intercambiable y reutilizable por otros investigadores; la definición de entidades, atributos y relaciones con MySQL es otra manera de traducir los datos al lenguaje de los ordenadores. En ocasiones la adición de metadatos —el título, las palabras clave, la descripción, la fuente, el tipo de recurso, las relaciones o la cobertura— es una operación suficiente para que el ordenador pueda manipularlos. Durante esta fase es imprescindible tener en cuenta dos variables: la interoperabilidad —el uso de algún estándar de modo que la información estructurada pueda intercambiarse y el proyecto se comunique con otros— y la preservación, esto es, que el proyecto tenga una vida

longeva e independiente del programa informático con el que fue creado, libre de los azares del mercado.

Por último, la publicación del recurso digital debería tener en cuenta los problemas de compatibilidad de los navegadores y cómo el usuario interactúa con la interfaz. No solo es determinante que la arquitectura de la información de la web sea coherente, consistente y clara, también lo son factores estéticos como los colores y las formas y las prácticas de la web social y de la «cultura participativa» como la colaboración y la apropiación (Jenkins, 2006). Usabilidad, estética y tipo de audiencia, pues, van de la mano. De hecho, cualquier interfaz puede entenderse como un dispositivo retórico que pretende persuadir al usuario para que haga una tarea: buscar, explorar o analizar el contenido, visualizar la información, aplicar filtros, descargarse un archivo, enriquecer, compartir o remezclar, entre otras. Si el usuario decide abandonar la web, ya porque la información no está actualizada, ya porque percibe errores ortográficos, ya porque el diseño gráfico no le parece atractivo, podemos concluir que el mensaje no ha sido recibido y que hemos fracasado. En definitiva, los proyectos de HD son modelos que pretenden que el usuario haga algo, lleve a cabo una *performance* —o una lectura performativa—. En este sentido las palabras de Wartofsky sobre las representaciones del conocimiento científico se ajustan como un guante a este tipo de contribuciones académicas: son —o deberían ser— una «llamada a la acción» (1979: 143).

### Por qué hacer Humanidades Digitales

Un convencimiento pocas veces declarado anima a las HD: hay tareas, funciones o procesos que el libro impreso, en tanto que tecnología, no puede realizar o automatizar. Los aspectos en que el ordenador supera al libro son varios: almacenaje, multimodalidad e interactividad. Un libro impreso, en principio, no tiene un límite de extensión pero la experiencia nos dice que una longitud excesiva se hace engorroso e incómodo. De sobra es sabido que el ordenador tiene una memoria superior. Un libro impreso, en principio, puede contener texto e imágenes, en blanco y negro y a color, pero el libro ilustrado no es la norma sino la excepción en la industria editorial y se produce a costa de aumentar su precio; nuestros navegadores, en cambio, no solo pueden visualizar imagen y texto, sino que también pueden reproducir vídeo y audio. Por último, un libro necesita ser abierto por el lector para que se ponga en marcha; este puede decidir seguir leyendo o cerrarlo, anotar en los márgenes e incluso romper sus páginas. El ordenador y la web permiten todo esto y además automatizar tareas como la búsqueda de palabras, manipular la información aplicando filtros (por ejemplo seleccionar una categoría concreta o desplegar una vista), compartir contenidos en redes sociales o anotar mediante etiquetas.

El progreso tecnológico no está exento de promesas: Marshall McLuhan (2001) habló en los sesenta de la tecnología como una «extensión» de la conciencia; Douglas Engelbart quería «potenciar» las capacidades cognitivas del ser humano (Barnett, 2013); y más recientemente Willard McCarty (2009) ha hablado del ordenador como si fuera un «telescopio». Todas estas metáforas son un tanto atrevidas, pero en los últimos años han surgido varios proyectos que exploran de manera creativa las posibilidades del ordenador y del medio digital. Las ideas que los fundamentan —hipertextualidad,

multimedia y participación— no son nuevas; de hecho, se encuentran en muchas publicaciones de mediados de los noventa como «The Rationale of Hypertext» de Jerome McGann (1996, 1997; cf. 2001); lo que es nuevo es que ahora disponemos de un medio económico de publicación —la web— cuyo nivel de penetración en la sociedad es muy amplio, y que los humanistas ya saben con qué lenguajes informáticos pueden realizar estos proyectos. Es, ante todo, un cambio de mentalidad y actitud hacia la informática.

Las herramientas Prism e Ivanhoe, desarrolladas por Scholars' Lab, plantean la lectura como un juego interpretativo y fomentan la anotación colaborativa entre los usuarios. La plataforma de publicación Scalar creada por The Alliance for Networking Visual Culture tiene por objetivo reinventar el libro en el siglo XXI permitiendo visualizar cómo se estructura la información en forma de red o índice, crear múltiples senderos o rutas de lectura, anotar imagen, vídeo y texto, y etiquetar con palabras clave el contenido creado por el usuario. *Woolf Online* contiene múltiples ediciones y versiones de *Al faro*, es decir, es un libro que contiene muchos libros, en función de los intereses del lector. Además, se puede reconstruir el contexto de producción mediante un repositorio con ensayos escritos por la célebre escritora, reseñas contemporáneas, fotografías, borradores y notas. Por último, *Pelagios* es una colección de topónimos del Imperio romano mencionados en documentos antiguos que permite no solo visualizar en un mapa dónde se encuentra tal topónimo y descubrir los documentos que lo referencian, sino conectar los datos en otros contextos mediante un Identificador de Recursos Uniforme (URI). Estos proyectos y herramientas son solo cinco ejemplos recientes a la vanguardia de la experimentación en HD pero la tendencia en pensar en términos sociales, visuales e interactivos la literatura, el arte o simplemente los datos contenidos en libros crece año tras año.

La traducción de signos materiales a signos digitales, lo que se conoce como «remediación» (Bolter, 1999), oscila entre la imitación del medio original y el desafío de lo nuevo que plantea el medio de destino —el ordenador, la web—. Para que los proyectos de HD sean relevantes, para que amplíen los límites de las disciplinas humanísticas, es necesario que se dejen de lado algunas convenciones del libro impreso y que se trate al ordenador como un medio nuevo. Ya lo anunciaron Alan Kay y Adele Goldberg en 1977: el ordenador en realidad es un «metamedio» porque simula todos los medios precedentes y a la vez añade propiedades nuevas (Kay y Goldberg, 2003; Manovich, 2013) como buscar, contar y clasificar palabras, cortar y pegar textos e imágenes, ver la información por capas estructuradas, etiquetar contenido, interrelacionar documentos web, etc. Los humanistas digitales deberían, pues, llevar a la práctica las ideas de estos pioneros y no limitarse a simular las funciones del libro en papel. Con esto no pretendo declarar la enésima muerte del libro impreso sino señalar, por un lado, su especificidad —extensión limitada, texto con pocas imágenes y sin

enlaces externos— y, por el otro, la existencia de unos géneros discursivos académicos que hasta ahora no tenían cabida —por su escala y por su naturaleza multimedia, interconectada e interactiva— en las tapas y márgenes del libro impreso. Es este tipo de conocimiento científico, que si se imprime en papel pierde cualidades, el que las HD aspiran a producir y difundir.

A. ROJAS CASTRO /  
DE LOS DATOS  
AL METAMEDIO

### Conclusiones

El impacto del ordenador e internet en nuestra sociedad y, de manera más específica, en los estudios humanísticos está lejos de ser analizado y conocido en todas sus dimensiones. No obstante, las HD pueden entenderse como la reacción académica a los cambios tecnológicos recientes que pretende responder a preguntas como: ¿qué metodologías existen para digitalizar, estructurar y presentar recursos digitales? ¿Qué nuevas posibilidades nos brinda el ordenador y la web ajenas al código? ¿Cuáles son las características de la cultura digital y de las formas de expresión artística nacidas en y para ser experimentadas en la pantalla? En la intersección entre la tradición y la innovación, el humanista digital es el investigador más preparado para plantear nuevas preguntas sobre la cultura, pasada y presente, y para dotar de sentido a los fenómenos digitales gracias a la unión del juicio y la crítica con las herramientas digitales.

Desde un punto de vista institucional en los últimos años las HD se han organizado en redes y organizaciones. En nuestro país contamos con numerosos proyectos de HD ligados al hispanismo pero, aunque el interés es creciente, sigue siendo necesario consolidar estos principios metodológicos, valores y prácticas mediante la creación de centros de HD (Rodríguez Yunta, 2014 y Romero Frías-del Barrio García, 2014). Concebidos como un espacio físico en el que colaboran humanistas, informáticos, bibliotecarios y diseñadores gráficos, los centros y laboratorios de HD han jugado un papel activo en la expansión de la disciplina en el ámbito anglosajón y se adivina que serán fundamentales en suelo hispánico para que se desarrollen proyectos, se asesore a investigadores interesados y se garantice la preservación de los recursos creados en el pasado. En otras palabras: la investigación y la innovación deberían ser las prioridades. Una vez se haya investigado de manera intensiva durante algunos años y se haya creado el ambiente adecuado para experimentar con nuevas tecnologías será posible crear programas formativos de HD.

A. R. C.—UNIVERSITAT POMPEU FABRA



 Logotipo de la  
plataforma Scalar.

INSULA 801



EL AFORISMO  
ESPAÑOL DEL  
SIGLO XX

INSULA 822  
JUNIO 2015

17