

## Bibliografía consultada

- DIGITALHUMANITIES (2008). *A Digital Humanities Manifesto*. California: UCLA, 15 de diciembre. Disponible en: <<http://manifesto.humanities.ucla.edu/2008/12/15/digital-humanities-manifesto/>> [Consultado el 20 de julio de 2018].
- (2009). *A Digital Humanities Manifesto 2.0*. California: UCLA, 29 de mayo. Disponible en: <<http://manifesto.humanities.ucla.edu/2009/05/29/the-digital-humanities-manifesto-20/>> [Consultado el 20 de julio de 2018].
- AUTORES.UY (2018). «Breve reseña del panorama de la digitalización en Uruguay», 21 de mayo. Disponible en: <<http://blog.autores.uy/2018/05/breve-resena-del-panorama-de-la-digitalizacion-en-uruguay/>> [Consultado el 20 de julio de 2018].
- AGENCIA DE GOBIERNO ELECTRÓNICO Y SOCIEDAD DE LA INFORMACIÓN Y DEL CONOCIMIENTO (AGESIC) (2016). *Encuesta Específica de Acceso y Uso de TIC (EUTIC 2016)*. Disponible en: <[https://www.agesic.gub.uy/innovaportal/file/6308/1/eutic2016\\_final.pdf](https://www.agesic.gub.uy/innovaportal/file/6308/1/eutic2016_final.pdf)> [Consultado el 20 de julio de 2018].
- GAYOL, V. y MELO FLÓREZ, J. A. (2017). «Presente y perspectivas de las humanidades digitales en América Latina». *Mélanges de la Casa de Velázquez*, vol. 47, n.º 2. Disponible en: <<http://journals.openedition.org/mcv/7907>> [Consultado el 20 de julio de 2018].
- PONS, A. (2013). *El desorden digital. Guía para historiadores y humanistas*. Madrid: Siglo Veintiuno Editores.
- VALDÉS GÁZQUEZ, R. y BIA, A. (2013). «Una propuesta educativa para las HD: el nuevo máster en HD de la UAB», en *Actas de I Congreso internacional de la HDH: Humanidades Digitales Hispánicas*. A Coruña: Universidade da Coruña.

## Una excursión al mundo de las humanidades digitales en el Cono Sur: objetos y prácticas

LEONARDO FUNES<sup>1</sup>

### Resumen

Se ofrece una introducción general al ámbito multidisciplinar de las humanidades digitales en la región, comenzando con una reseña histórica de su «prehistoria», es decir, de las experiencias de trabajo en el campo de las humanidades con el auxilio de herramientas informáticas. Luego se centra en la discusión de las diferentes modalidades del libro digital en el ámbito académico: el libro impreso digitalizado, el libro digital propiamente dicho y el que resulta de la edición digital académica. El análisis de cada tipo permite trazar un panorama de los problemas, desafíos y perspectivas de las humanidades digitales en el Cono Sur, en tanto construcción local en el marco globalizante del universo digital.

### Abstract

This article offers a general introduction to Digital Humanities, beginning with an overview of its «prehistory», i. e. research activities in human sciences with the help of informatic tools. Then it focuses on different versions of the digital book in the academic milieu: digitized printed book, born-digital book and academic digital editing. The analysis of each one of them offers an opportunity to point out problems, challenges and perspectives that Digital Humanities face in our region, as a local construction in the global context of the digital world.

<sup>1</sup> Instituto de Investigaciones Bibliográficas y Crítica Textual (IIBICRIT), Consejo Nacional de Investigaciones Científicas y Técnicas (Conicet), Universidad de Buenos Aires (UBA).

## Consideraciones preliminares

Con la rapidez que ya es común en el universo digital, la etiqueta *humanidades digitales* (HD) se ha expandido en el ámbito académico y extraacadémico de la región. A poco que nos pongamos a pensar en esta expansión surgen de inmediato dos comprobaciones: la primera tiene que ver con las enormes dificultades para definir qué son exactamente las HD; la segunda se relaciona con el descubrimiento de cuántos de nosotros y por cuánto tiempo veníamos practicando las HD sin saberlo.

En efecto, debemos renunciar a la búsqueda de una definición exacta de las HD, porque es un término que permite abarcar un gran número de actividades. Pero puestos a dar alguna formulación que ayude a visualizar el fenómeno, podríamos decir que están constituidas por el conjunto de prácticas y de saberes que nos permiten operar y estudiar la cultura digital. Surge, por tanto, del cruce de métodos computacionales con las disciplinas humanísticas (arte, Historia, letras, filosofía). Desde un enfoque más académico, también podemos definir las HD como un conjunto de teorías, métodos y valores en el que convergen el objeto de estudio (humanidades) con el uso de procedimientos informáticos para recoger datos, procesarlos, estructurarlos y presentarlos de manera interactiva y multimodal en un contexto académico.

Asimismo, mucha gente realiza actividades humanísticas mediante el empleo sistemático de herramientas informáticas, y lo viene haciendo desde hace bastante tiempo, sin detenerse a pensar que lo que hace pertenece al campo de las HD. Como el burgués gentilhomme de la comedia de Molière, que descubre que habla en prosa, estos colegas descubren, no sin asombro, que están practicando HD. Esto nos ofrece un excelente ejemplo de la potencia referencial de una denominación atractiva, una cuestión sobre la que volveré brevemente más adelante; pero ahora solo quisiera subrayar que toda esa actividad que se viene desarrollando desde hace unas décadas podemos entenderla hoy como una suerte de «prehistoria» de las HD.

Esta prehistoria se remonta a los inicios de la década del ochenta; en aquel entonces, si se trataba de dar un nombre a esa actividad, se prefería hablar de *informática y humanidades*. De hecho, en el ámbito hispánico, este fue el título de un importante libro publicado por el lingüista y filólogo Francisco Marcos Marín (1994); en él se pasaba revista a los distintos campos en los que se venían desarrollando nuevas modalidades de trabajo con la lengua y con los textos: las novísimas «industrias de la lengua», que habían comenzado a tallar en los tiempos de la celebración del Quinto Centenario del descubrimiento de América (convenientemente actualizado como «encuentro de culturas»), los primeros repositorios y bases de datos (archivos digitales textuales, de texto e imagen, bibliográficos), los nuevos modos de editar textos (edición unificada, crítica textual informatizada) y la rápida ampliación de la conectividad (acceso remoto y comunicaciones). El

libro puede aprovecharse hoy como una excelente reseña histórica de los comienzos de este campo disciplinar, pero también nos permite apreciar de modo dramático cuán apropiado es hablar en este caso de «prehistoria»: es que la precariedad de los medios disponibles a comienzos de la década de 1990 se vuelve hoy tan evidente, dado el extraordinario progreso en el campo de la informática y, especialmente, la gran revolución que significó Internet para la comunicación y para el acceso a (y la circulación de) los textos, que el modo en que se trabajaba hace tan poco tiempo atrás —según leemos en sus páginas— nos resulta verdaderamente «prehistórico».

Por su naturaleza eminentemente práctica y experimental, los trabajos enfocados en la relación entre informática y humanidades fueron marginales en relación con las corrientes dominantes en los estudios culturales y literarios. Eso cambió con la revolución de Internet en los noventa; pero aun así, nunca fue impulsada desde el centro de la Academia, sino que se fue colando en la agenda desde los márgenes de diversas disciplinas.

La propia denominación se acuñó en la academia anglosajona y se registró por primera vez en 2001, cuando se creó un seminario permanente de *Digital Humanities* en la Universidad de Virginia, y su difusión quedó asegurada con el volumen colectivo *A Companion to Digital Humanities*, publicado en 2004 (Schreibman, Siemens y Unsworth, 2004). Desde hace 15 años, las *Digital Humanities* son un área en pleno desarrollo en Europa y el mundo anglosajón (Estados Unidos, Canadá, Australia).

La aparición en el ámbito hispanoamericano de la expresión *humanidades digitales* es mucho más reciente, pero tuvo una rápida aceptación y propagación. Constituye un caso claro de «acción discursiva», en la medida en que la denominación trabaja creativamente al menos en dos planos: como clave de identidad que viene poniendo en contacto a grupos muy diversos y como elemento constructivo del propio objeto que pretende designar.

En efecto, en el caso del ámbito rioplatense, el término empezó a circular en un contexto específico, en el que diferentes grupos, proyectos e investigadores aislados venían desarrollando trabajos en educación, archivística, ecdótica, bibliotecología, lingüística. Ante ese panorama, en el que las prácticas atravesaban disciplinas humanísticas tan diversas, podría decirse que, en la instalación de este nuevo campo, se operó de manera institucional: por una parte, proponiendo un espacio institucional (que terminó siendo en 2013 la Asociación Argentina de Humanidades Digitales, AAHD) como lugar de encuentro y puesta en común de experiencias (previas y en desarrollo) tan heterogéneas; por otra parte, absteniéndose de fijar límites (al menos demasiado estrictos) en el momento de definir lo que entra y lo que no entra en el marbete de las HD. Complementariamente, se trataba no tanto de *introducir* a los colegas en el nuevo campo como de *reconocer* que se estaba en él desde hacía un tiempo.

No hace mucho, en el capítulo introductorio de una *Historia de la literatura española*, publicada por la prestigiosa Cambridge University Press, el autor se explayaba en los motivos por los cuales era imposible escribir una historia de la literatura española, muy complacido con la brillante paradoja de su propuesta y absolutamente impermeable a la evidencia del par de miles de páginas que a continuación ponían en entredicho tal imposibilidad (Gies, 2004). En fin, lo paradójico posee un atractivo irresistible. Pero en nuestro caso, la propuesta de una nominación englobante de prácticas múltiples y heterogéneas no buscaba complacerse en la aporía, sino, muy por el contrario, proveer una reflexión que fuera más allá del mero uso instrumental de ciertas técnicas provistas por la informática.

La acción aglutinante y difusora de la Asociación tuvo dos hitos importantes en la realización de un congreso nacional en 2014 y un congreso internacional en 2016.<sup>2</sup> En esos eventos académicos se pudo apreciar un panorama bastante completo del amplio abanico de experiencias y reflexiones sobre un fenómeno en plena efervescencia. Además de los debates sobre el marco epistemológico y los perfiles disciplinares e institucionales de las HD, se dieron a conocer trabajos en curso en las áreas más diversas: repositorios digitales, edición académica, construcción de bases de datos, proyectos de digitalización y preservación de archivos y documentación, publicación digital en sentido amplio, proyectos educativos, comunicación digital (texto e imagen en entornos digitales), plataformas para la investigación en ciencias sociales. También, en un segundo plano de recursividad, hubo espacio para la reflexión cultural sobre el impacto de las nuevas tecnologías en las formas de producción, circulación y recepción de los discursos (textuales y audiovisuales) y la reflexión lingüística sobre las nuevas prácticas comunicacionales.

Para no extender más lo que pretendía ser una breve introducción y reseña histórica del campo disciplinar, solamente agregaré que Gimena del Río, nuestra principal especialista en el tema, en una conferencia plenaria que diera en el TEI Victoria 2017,<sup>3</sup> llamaba la atención sobre las significativas diferencias que había entre las *Digital Humanities* y las humanidades digitales, pues no se trataba de una mera traducción terminológica, sino de una construcción disciplinar fundada en condiciones tecnológicas, sustrato lingüístico y antecedentes históricos muy diversos en relación con los vigentes en el Primer Mundo. Todo ello hace de las HD un fenómeno local en el contexto globalizante del universo digital. De un pequeño rincón de ese fenómeno voy a hablar en lo que resta del trabajo.

2 Las actas correspondientes pueden consultarse en la página web de la AAHD: <<http://aahd.net.ar/>>.

3 Me refiero a la Text Encoding Initiative 17<sup>th</sup> Annual Conference, evento celebrado en Victoria (Canadá), del 11 al 15 de noviembre de 2017.

## El objeto libro desde una óptica filológica digital

Si bien este campo disciplinar engloba iniciativas muy heterogéneas (artísticas, informales, educativas, académicas), mi práctica se circunscribe al ámbito académico, en tanto investigador del Conicet y director de un instituto de investigación dedicado a las ciencias humanas.

Por lo tanto, hablaré aquí de un objeto muy específico —el libro digital— desde esa perspectiva académica, que es la única que conozco con cierto detalle. Esta perspectiva es, en concreto, la de la filología digital, una expresión —en este caso sí— equivalente a la fórmula inglesa *Digital Philology*.

Todo esto necesita una serie de aclaraciones previas para evitar posibles malentendidos.

*Filología* es una denominación que cayó en desuso en nuestro medio hace bastante tiempo. Sobrevivió hasta hace unos pocos años en el Conicet en el nombre de la Comisión de Filología, Lingüística y Literatura; hasta que fue desterrada y la comisión pasó a llamarse de Literatura, Lingüística y Semiótica.

*Filología* tiene un penetrante olor a naftalina en nuestra región, pero es un término perfectamente vigente en gran parte del universo mundo, porque alude con precisión a la estrecha relación que hay entre lengua y literatura, que es el objeto de gran parte de los estudios literarios.

También debo advertir que los casos concretos de filología digital a los que me voy a referir están relacionados con la literatura de la Edad Media. De hecho, mi trabajo concreto está enfocado en el período medieval y consiste mayormente en la edición crítica de textos en castellano antiguo, labor que pertenece a una subdisciplina filológica que recibe el nombre de ecdótica o crítica textual.

A muy pocos se les ocurrirá conectar de primera intención y en forma espontánea los adelantos de la informática con los estudios medievales. Pero contra todo lo esperable, esta área de estudios ha sido una de las más relacionadas con los diferentes avances que ha habido en la cultura digital, desde los inicios de la computación hasta el presente.

En todos los trabajos introductorios a este campo disciplinar se encontrará una referencia insoslayable: el primer trabajo de HD fue encarado por el padre Roberto Busa, con la asistencia de técnicos informáticos de IBM, para generar el *Index Thomisticus*, una base de datos con el léxico y el paradigma conceptual de la obra completa del teólogo medieval Tomás de Aquino (Busa, 1980). De modo que los estudios medievales ya estuvieron involucrados en el mismo nacimiento de las HD. Por eso, el hecho de que hoy haya una revista académica virtual llamada *Digital Philology* y que se dedique exclusivamente a los estudios medievales no debería llamar la atención.

En el ámbito local, esta relación se dio ya desde los años 1980: a mediados de esa década un grupo de investigación en hispanomedievalismo obtuvo financiamiento del Conicet para desarrollar un proyecto de investigación sobre «Aplicación de los

medios informáticos al análisis y edición de textos»; esos textos eran crónicas del siglo XIV. El objetivo era encontrar algún modo de automatizar las operaciones de cotejo de diferentes versiones y de elaborar listados exhaustivos de variantes. Por supuesto, todavía no estaban dadas las condiciones tecnológicas para alcanzar resultados relevantes o duraderos (el conjunto de los discos rígidos de las máquinas de que disponíamos tenía una capacidad de almacenamiento inferior a la de un teléfono celular básico de nuestros días), pero al menos los objetivos marcaban una tendencia en cuanto a la búsqueda de una relación productiva entre la informática y las humanidades en términos estrictamente académicos.

No voy a hacer aquí la historia de los últimos treinta años, porque estaría fuera de lugar, pero sí querría llamar la atención sobre una característica que se ha mantenido constante y que no carece de su costado polémico: se trata de la conexión que siempre hubo entre estos proyectos locales y programas de investigación internacionales con sede central en el Primer Mundo. Aquellos proyectos de los ochenta y principios de los noventa estaban articulados con los programas (en el sentido de conglomerados de proyectos) DOSL [Dictionary of Old Spanish Language] y BOOST [Bibliography of Old Spanish Texts] dirigidos por el Hispanic Seminary of Medieval Studies, radicado en la Universidad de Wisconsin-Madison. Posteriormente, los resultados fueron exportados a la gigantesca base de datos del programa CORDE [Corpus Diacrónico del Español] gestionado por la Real Academia Española y que es hoy una herramienta fundamental para todos los especialistas interesados en Historia de la lengua española o lingüística histórica.

Esto quiere decir que los instrumentos de programación necesarios para la edición de los textos y la infraestructura a escala necesaria para el almacenamiento de estas grandes bases de datos estaban y provenían del Primer Mundo, y esto es algo que perdura en la práctica actual de la *Digital Philology*, con ciertos matices.

Salteándonos, entonces, los pormenores de treinta años de evolución en este campo, con las diversas formulaciones que trataron de identificarlo (informática y humanidades, HD, filología digital), voy a referirme a las experiencias actuales y las perspectivas inmediatas.

Dentro de esas coordenadas, puedo decir que el libro digital se está desarrollando en tres modalidades muy diferentes que conviene precisar aquí.

En primer lugar, tenemos el libro digitalizado. Es decir, el libro que nació impreso o manuscrito y que ahora es fotografiado, escaneado y digitalizado.

Todos conocemos el ambicioso proyecto de Google Books y su aspiración a digitalizar todos los libros existentes; pero hay centenares de proyectos con objetivos más modestos y realistas y las bibliotecas y repositorios digitales han proliferado por el mundo. El impacto de este fenómeno en el desenvolvimiento de las ciencias, tanto las físico-naturales como las sociales y humanas, ha sido enorme y decisivo: el acceso inmediato a la bibliografía crítica o a las fuentes documentales

permite trabajar con una velocidad y una riqueza de datos impensable hace apenas veinte años atrás.

La disponibilidad de herramientas informáticas básicas para esta tarea permitió una difusión exponencial de las iniciativas de digitalización. Y junto con la masa inabarcable de textos en formato digital llegaron también nuevos problemas y nuevos desafíos.

La multiplicación de iniciativas individuales, implementadas y coordinadas por personas de muy variado conocimiento sobre informática (desde expertos en el manejo de programas de OCR hasta personas que solo sabían pulsar el botón de un escáner, manejado como una fotocopidora), derivó en resultados muy dispares en cuanto a la consistencia y confiabilidad de las bibliotecas y de los repositorios.

Como ya alertara hace años Robert Darnton, el director de la biblioteca de la Universidad de Harvard, refiriéndose a Google Books, tenemos un grave problema de control de calidad —además de riesgos de todo tipo: la constitución de un monopolio imbatible, el hecho de que esta iniciativa digitalizadora sea llevada adelante no por un consorcio de instituciones educativas y culturales públicas, sino por una empresa privada, el siempre latente peligro de obsolescencia tecnológica, aun tratándose de una empresa fundada en la innovación permanente (Darnton, 2009).

Ante la inexistencia de un código de buenas prácticas consensuado internacionalmente, cada uno se ha puesto a escanear como le permitía su grado de conocimiento, o como se lo permitían los medios a su disposición, o como le dio la gana. Y así, junto a proyectos de primer nivel como el Proyecto Gutenberg o Gallica (de la Biblioteca Nacional de Francia) o la Biblioteca Digital Hispánica (de la Biblioteca Nacional de España), tenemos chapucerías megalómanas como Google Books (donde las quejas son constantes por falta de páginas que el pobre técnico contratado para escanear centenares de libros se saltó y no tuvo tiempo de controlar)<sup>4</sup> y modestísimos intentos locales de baja calidad de imagen, como, por ejemplo, el proyecto de revistas nacionales argentinas AHIRA [Archivo Histórico de Revistas Argentinas], muy útil a pesar de las deficiencias formales.

Desde el punto de vista del público lector en general, esto no parecería tan grave: después de todo, es mejor tener la posibilidad de leer un texto, aunque le falten un par de páginas o alguna esté escaneada en diagonal, que carecer completamente de esa posibilidad. Pero desde el punto de vista académico, el problema es más serio, sobre todo cuando hablamos de libros anteriores a la imprenta industrial (es decir, libros impresos antes de 1830, que es la fecha convencional que normalmente se acepta) o de manuscritos antiguos, medievales y modernos, casos en los cuales la alta definición de la imagen y el cuidado con que se la haya procesado hace toda la diferencia entre documentación útil o descartable.

4 Ver al respecto Galina Russell (2017).

¿Cuál es el futuro de esta modalidad de libro digital? Si logramos consensuar códigos de buenas prácticas y cumplir con procesos de digitalización de alta calidad, el futuro es altamente positivo, pero antes de llegar a él es necesario superar este desafío previo.

Los equipos de escaneo y procesamiento de imágenes de alta resolución con reconocimiento de caracteres son previsiblemente caros para los magros presupuestos de las instituciones que patrocinan la formación de repositorios digitales, de modo que hay que aguzar el ingenio y la inventiva para acceder al instrumental adecuado.

Un ejemplo es el escáner desarrollado por dos emprendedores jóvenes con el que el Instituto de Investigaciones Bibliográficas y Crítica Textual, que actualmente dirijo, está trabajando en estos momentos. Se trata de un sistema de digitalización integral de bajo costo desarrollado con tecnología abierta. Matías Butelman y Juan Pablo Suárez desarrollaron un escáner «en V» para libros y documentos encuadernados. Consiste en un par de cámaras digitales de gama media con *software* modificado para que puedan ser operadas remotamente desde una computadora, un estativo compuesto por una cuna en V para apoyar el libro, una placa móvil, también en V, que se apoya sobre el documento y corrige la curvatura de las páginas, y un soporte para las cámaras que las mantiene en posición. Las cámaras digitales son configuradas de un modo especial que permite controlarlas desde una computadora y realizar disparos sincronizados, descargas inmediatas y un procesamiento semiautomatizado de las imágenes. El resultado final del proceso (para el que se usan programas como Scan Tailor) es un documento en formato PDF para libros, revistas y documentos encuadernados o en formato JPG para fotografías, ilustraciones, planos, etc., que se puede someter a un tratamiento de OCR y finalmente ser publicado en una plataforma web para su consulta en línea. Me detengo en estos pormenores técnicos para ilustrar también de qué modo el trabajo en HD, cualquiera sea su dimensión y modalidad, debe ser forzosamente grupal y colaborativo: no se espera ni sería deseable que el humanista digital sea al mismo tiempo un técnico informático, por lo que es imprescindible la presencia de alguien que cumpla esa función dentro del equipo de trabajo. Conviene aclarar también que los detalles técnicos que acabo de describir se inscriben en una perspectiva humanística digital de apoyo al *software* libre y los programas de código abierto, lo que nos remite a la contienda ideológica que hoy se sostiene en el universo digital y sobre la que volveré en el cierre de este trabajo.

La segunda modalidad es la del libro que nace digital. En el caso del libro digital académico, justamente por razones de costos y por el tipo de circulación restringida que suele tener este tipo de publicaciones, se trata de una modalidad que no deja de crecer. Ajena a los problemas que enfrenta el *e-book* en términos de mercado, la publicación académica se va adaptando con bastante rapidez a los formatos digitales; un fenómeno que no se da sin sus problemas y conflictos.

El conflicto académico se manifiesta en las diferencias de valoración entre la publicación en papel y la publicación virtual. Todavía subsiste en ciertos organismos de evaluación científica una sospecha sobre la calidad de los trabajos difundidos por medios digitales. Es probable que este prejuicio derive del impacto que ha tenido el aumento exponencial de información disponible en la red no sometida a validación según estándares satisfactorios para la comunidad científica. La falta de discriminación de los motores de búsqueda en cuanto a la calidad de los resultados (una página de un instituto de ciencia aparece al lado del blog de un aficionado) rara vez puede ser superada por el usuario no especializado, puesto que se privilegia la rapidez en la accesibilidad. De todos modos, en la medida en que el libro digital académico está dirigido a un público específico, hay herramientas en la web para sortear estas dificultades: el conflicto en el ámbito científico es indicio de que en el *staff senior* de ciertas instituciones hay un problema de desconocimiento con respecto a las herramientas informáticas.

También hay un problema económico: ¿las publicaciones científicas deben ser pagas o gratuitas? La política mundial orientada a lo que se denomina *Open Access* busca la libre circulación del conocimiento científico, lo cual entra en conflicto con editoriales y organismos de difusión. Tenemos en la Argentina una ley al respecto y las universidades y el Conicet están poniendo en acceso libre y gratuito gran parte de la producción científica. Esta política está siendo adoptada por muchos países, pero no por todos, lo cual está generando una situación muy injusta: el mundo anglosajón se mantiene a pie firme en un modelo comercial y hay que pagar —y mucho— por las publicaciones científicas estadounidenses y británicas. Por suerte los *hackers* rusos están cumpliendo un invalorable servicio a la comunidad científica internacional y esta es una lucha en la que el FBI lleva todas las de perder.

Pero más allá de estas dificultades, es fácil prever que en un futuro próximo esta modalidad del libro digital será absolutamente dominante en el ámbito académico, algo ya muy visible en el caso de las revistas, donde el uso de la tecnología OJS [Open Journal System] es obligatoria en la mayoría de las universidades argentinas.

Pero quisiera referirme a otras cuestiones de índole cultural, más cercanas a mi formación y mi campo de especialidad, que es la literatura medieval. Las dos modalidades de las que hablé hasta ahora —el libro digitalizado y el libro digital— tienen algo en común sobre lo que quisiera llamar la atención: los editores, maquetadores y diseñadores del mundo digital trabajan con la misma imagen mental de lo que es un libro que los impresores desde Gutenberg hasta el presente. Es más, el diseño de página y la puesta en página de cualquier texto todavía sigue los parámetros inventados en el siglo XII. Para muchos historiadores del libro, entre los que me incluyo, ese fue un momento revolucionario que en términos cualitativos superó a la invención de la imprenta, ya que, en última instancia, la

impresión solo mecanizó y multiplicó un diseño que se venía perfeccionando de manera manuscrita desde tres siglos antes.

Esto demuestra una vez más —como ya sabemos al estudiar otro momento importante de transición: el del manuscrito al impreso— que se trata de un proceso lento y gradual, aunque tengamos la sensación de que todo está cambiando de modo radical y a gran velocidad.

Pero al mismo tiempo debo advertir que el cambio se está produciendo y que, a pesar de la pervivencia de muchos elementos formales, el libro digital posee una especificidad que lo vuelve irreductible al libro impreso. Aunque los programadores han logrado desarrollar efectos como la reproducción del sonido del papel al dar vuelta la hoja o el efecto visual de una página que se curva como si le pasáramos el dedo, la naturaleza virtual de la imagen en una pantalla hace del texto digital algo inconmensurable con un libro físico que sostenemos en la mano y cuyo peso y consistencia forma parte de nuestra percepción.

Pero es en la tercera modalidad del libro digital donde sí vemos ya un grado de transformación mucho más notable: se trata de la edición digital académica de textos. Esta es una clase de libro digital tan diferente que la propia denominación «libro digital» resulta completamente inapropiada.

Las prácticas y objetivos concretos de la filología digital consisten en la edición de textos y la compilación de bases de datos. En cuanto a la edición, se trata de poner a disposición de la comunidad de estudiosos versiones confiables y seguras de textos premodernos (desde la Antigüedad clásica hasta el Renacimiento). Esto se hace de dos maneras básicas: una es la transcripción paleográfica de los testimonios conservados, es decir, la copia más fiel que las actuales posibilidades tipográficas permiten de los registros de escritura manuscritos o impresos que constituyen los testimonios de las obras (literarias o no) que nos han llegado desde el pasado; la otra manera es la edición crítica, es decir, la reconstrucción, a partir de los testimonios conservados, de la versión más cercana posible a la redacción original de la obra en cuestión.

Para elaborar una edición digital tenemos que transcribir el texto en un lenguaje específico, el Extensible Markup Language (XML), y según un sistema de codificación conocido como Text Encoding Initiative (TEI). Para el trabajo editorial no sirven los procesadores de texto habituales, como el Word de Windows, sino que habitualmente se trabaja con programas de edición *ad-hoc* «tuneados» por programadores, sobre la base de matrices de código abierto, aunque existe algún programa de edición en TEI de carácter comercial, como Oxygen. El resultado del trabajo de transcripción y de edición de un original no es un archivo que uno pueda grabar en el disco rígido de su computadora o subir a la nube: requiere ser alojado en plataformas muy amplias y muy complejas. Lamentablemente, no hay ninguna de estas plataformas en el Cono Sur, y seguimos batallando para que las autoridades del Conicet, en el caso argentino, vean la importancia de contar con esta herramienta y, de ese modo, revertir —en ese campo muy acotado de la

edición digital— la tendencia al crecimiento de la brecha digital con los países desarrollados.

Me referiré a un par de casos concretos para ilustrar esta modalidad del libro digital. Participé en un proyecto radicado en la Universidad de Birmingham y con fondos de la Modern Humanities Research Association del Reino Unido, con la asociación de investigadores y centros de investigación de Francia, España y Argentina: The Estoria de Espanna Digital Project.

El proyecto consistió en la elaboración de una edición digital de la gran crónica general de España que mandó a redactar el rey Alfonso el Sabio en la segunda mitad del siglo XIII. Es fácil imaginar que se trata de una obra de gran extensión y de ambición enciclopédica, de modo que su redacción siguió una modalidad similar a la de la construcción de una catedral (o de una obra pública en mi país): un trabajo de décadas con infinidad de cambios, arreglos, ajustes, para quedar finalmente inacabado en sentido estricto. El problema adicional es que de cada etapa se fueron haciendo copias, con lo cual circularon y se multiplicaron versiones más o menos parciales, más o menos *aggiornadas* de la obra. Hoy se conservan alrededor de cien copias manuscritas de los siglos XIV y XV y, obviamente, ninguna es idéntica a otra. Hace ya más de cien años, el gran filólogo Ramón Menéndez Pidal publicó en dos tomos lo que creyó que era la versión oficial final de la obra (Menéndez Pidal, 1906). Hoy sabemos que no es así, que existieron al menos tres redacciones en tiempos diferentes (primitiva, crítica y amplificada) entre 1270 y 1284. Con lo cual los historiadores vienen usando desde hace un siglo un libro que no responde a la realidad textual de la obra y que no considera las especiales condiciones de una obra que se produjo y circuló de modo manuscrito durante los dos siglos previos a la invención de la imprenta.

Los participantes en el proyecto, como muchos otros medievalistas, llamaron la atención sobre el hecho de que el modo de producción y circulación de los textos en la web tenía sorprendentes puntos en común con las formas de la cultura manuscrita medieval. Si uno se refiere a un texto en cuya escritura mucha gente mete mano, sin que se sepa al final qué cosa fue poniendo o cambiando cada uno, que deja la sensación de que el texto nunca está terminado del todo y que es bastante inestable porque puede cambiar en cualquier momento, un medievalista estaría de acuerdo con que se está describiendo el modo en que trabajaban los escribas en la Edad Media, cuando en realidad se estaba hablando de un grupo trabajando con documentos compartidos en la web.

En consecuencia, la fundamentación de este proyecto insiste en que la forma más históricamente fiel de publicar este tipo de obras enciclopédicas medievales no es el formato libro impreso, sino el formato digital en internet.

La primera fase del proyecto fue la transcripción y comparación de cinco manuscritos. Las transcripciones están vinculadas a imágenes de los códices.

Y aquí ya afloran los primeros desafíos: por ejemplo, la transcripción, operación fundamental de la edición digital, sustituye un sistema semiótico por otro y

corre el riesgo de convertirse en una nueva autoridad, ya que su ubicación debajo de la imagen del código disfraza su condición de hipótesis. Porque las operaciones editoriales implican no solo la vinculación de transcripción e imagen, sino también la organización del texto. Estos desafíos se enfrentan mediante la modalidad de la colaboración abierta, que deja al descubierto todas las operaciones y permite una reorganización en cualquier momento del diseño global.

Las operaciones primarias o elementales son dos: la codificación de los datos básicos y la construcción de la edición propiamente dicha. Para ello se utiliza el sistema *Comunidades textuales* (*Textual communities*) que provee el *software* para transcribir —elaborado en la Universidad de Saskatchewan, Canadá— llamado Application Programmers Interface (API), que permite también la presentación de una edición crítica. Pero *Comunidades textuales* no es en sí la edición, sino la manera de compilar los metadatos. Permite, por un lado, la colaboración colectiva abierta (*crowdsourcing* en inglés), y por otro, la utilización de los metadatos por cualquier otro editor. Para poder trabajar en ese entorno hay que aceptar una serie de principios:

1. Las ediciones digitales son fruto de la colaboración científica.
2. Debe reconocerse públicamente la participación y el trabajo de todos.
3. Todo el trabajo realizado dentro de una comunidad virtual debe ser accesible a todos, sin restricciones.
4. Todos los materiales de cualquier edición digital deben conservarse en un almacenamiento de datos de larga duración y estar disponibles para la comunidad científica de manera independiente de su interfaz con el público. (Este es el principio fundamental: *Textual Communities* facilita la creación de archivos XML, el programa API interpreta esos archivos y permite la creación de recursos web que pueden servir para fines muy diversos).

En el caso de *Estoria de Espanna*, los materiales y los metadatos estarán almacenados en el Archivo de Investigación de la Universidad de Birmingham (UBIRA, por sus siglas en inglés).

Otro proyecto que estamos llevando adelante consiste en el estudio, edición y etiquetado de la poesía castellana dialogada de los siglos XII al XV.

El objetivo general es la elaboración de una base de datos de la poesía castellana medieval dialogada. Para ello, partimos de una selección del corpus, que abarca textos desde fines del siglo XII hasta el siglo XV a partir de una aproximación pragmática, esto es, desde una macroestructura basada en el elemento *diálogo* o *esquema secuencial dialogado*.

Tenemos como objetivo específico para la primera fase de trabajo —y anclaje inicial para el desarrollo de la investigación— el estudio y la edición digital paleográfica y crítica de la más antigua poesía castellana medieval en pareados (siglos XII-XIII): *Disputa del alma y el cuerpo*, *Razón de amor*, *Elena y María*, *Auto de los*

*Reyes Magos*, y la de otro grupo de textos de formato más breve, esparcidos principalmente en cancioneros.

Como primera medida, estamos elaborando una serie de criterios de transcripción paleográfica comunes a todos los textos en lo tocante a: desarrollo de las abreviaturas, representación de los signos generales de abreviatura, grafías y alógrafos.

Los tres pasos básicos del trabajo editorial son: la digitalización de las imágenes de los testimonios manuscritos, la transcripción paleográfica y finalmente la codificación, es decir, el etiquetado en lenguaje XML, para lo cual usamos el programa Oxygen. Para avanzar en esto y disponer del asesoramiento informático adecuado, nos asociamos con el Laboratorio de Innovación en Humanidades Digitales (LINHD), dependiente de la Universidad Nacional de Educación a Distancia de España. Los resultados terminarán incorporándose también (según una lógica similar a la de Textual Communities) a una gran base de datos gestionada por este laboratorio llamada Diálogo Medieval.

La mayor novedad desde mi punto de vista —y a la que no suele prestarse suficiente atención— es que esta modalidad del libro digital está dirigida a un nuevo público lector: los robots. Por primera vez generamos textos para ser leídos no por seres humanos, sino por máquinas. La intervención humana ahora consiste en interpretar las lecturas que los robots hacen de los textos digitales. Esto está estrechamente relacionado con lo que el teórico Franco Moretti (2013) denomina *lectura distante* (*distant reading*).

Lentamente nos estamos habituando a comunicarnos con las máquinas, porque no otra cosa hacemos cuando tecleamos un *hashtag* en las redes sociales: eso va dirigido a los robots, de cuya lectura podemos derivar luego la interpretación de los *trending topics* en la web. Allí está quizás lo más desafiante del futuro del libro.

## Breves consideraciones finales

Para concluir, quisiera referirme a los aspectos todavía polémicos de estos proyectos y del propio campo multidisciplinar de las HD: en principio, comprobamos que la estructura de proyecto tradicional, cuyo diseño comienza y termina en el ámbito nacional suponiendo un «trabajar con lo nuestro», ya no es viable. En todos los casos es necesario asociarse con centros del exterior y aprovechar desarrollos de *software* por lo general elaborados en el Primer Mundo; hacerlo de otro modo implica o redescubrir la pólvora o anotarse en una carrera de Fórmula 1 subidos arriba de un *karting*.

De todas maneras, aquí interviene el nuevo factor que es la globalidad de la red de redes y que, al menos en el campo de las humanidades, sigue jugando a favor de la libre circulación del conocimiento académico. Queda, entonces, como

desafío y perspectiva a la vez, promover una mayor «conectividad» en el ámbito nacional, para ello nuestra humilde propuesta de un ámbito de diálogo y difusión como es la AAHD. Y también atraer a los informáticos al campo de las humanidades e incentivarlos para el diseño de programas *ad hoc* que nos permitan dialogar con nuestros colegas de Europa y Estados Unidos con un mayor equilibrio en cuanto a las herramientas de la filología digital. En fin, todo está por hacerse y eso es algo que mantiene en alto el entusiasmo.

## Bibliografía

- BUSA, R. (1980). «The Annals of Humanities Computing: The Index Thomisticus». *Computers and the Humanities*, vol. 14 (2), pp. 83-90. Disponible en: <<https://link.springer.com/article/10.1007%2FBF02403798>> [Consultado el 6 de agosto de 2018].
- DARNTON, R. (2009). *The case for books: past, present, and future*. Nueva York: Public Affairs.
- GALINA RUSSELL, I. (2017). «¿Hacia dónde van las ediciones digitales académicas?», en GODINAS, L. y otros (eds.). *De Pérgamo a la nube. Nuevos acercamientos y perspectivas a las edades del libro*. Ciudad de México: UNAM.
- GIES, D. T. (2004). «The Funes effect: making literary history». *The Cambridge History of Spanish Literature*. Cambridge: Cambridge University Press.
- MARCOS MARÍN, F. (1994). *Informática y Humanidades*. Madrid: Gredos.
- MENÉNDEZ PIDAL, R. (1906). *Primera crónica general de España*. Madrid: Bailly Baillière.
- MORETTI, F. (2013). *Distant Reading*. Londres: Verso.
- SCHREIBMAN, S.; SIEMENS, R. y UNSWORTH, J. (eds.) (2004). *A Companion to Digital Humanities*. Oxford: Blackwell Publishing.

## Anáforas: propósitos, procedimientos y prácticas de un programa

MAXIMILIANO BASILE

RODRIGO ECHÁNIZ

MARIANA NOGUERA

ARTURO RODRÍGUEZ PEIXOTO<sup>1</sup>

### Resumen

En 2004, en el Seminario de Análisis de la Comunicación, en la actualidad Fundamentos Lingüísticos de la Comunicación de la Facultad de Información y Comunicación, a cargo de Lisa Block de Behar, se propuso, entre otras actividades curriculares, docentes y de investigación, la construcción de sitios web destinados a conservar y difundir, en forma digital, obras impresas del acervo cultural nacional, así como otros materiales de carácter documental. Actualmente cuatro grandes colecciones (Biblioteca Digital de Autores Uruguayos, Publicaciones Periódicas del Uruguay, Figuras y Sobre la Prensa) están unificadas en el repositorio Anáforas ([anaforas.fic.edu.uy](http://anaforas.fic.edu.uy)) y funcionan mediante el software libre Dspace. El trabajo conjunto y entusiasta de un equipo de docentes, egresados y estudiantes hace posible el continuo crecimiento del portal. La labor involucra tareas de investigación, docencia y extensión, por lo que cumple con los principales cometidos que la ley orgánica asigna a la Universidad de la República.

Ya no se requiere fundamentar la construcción colectiva y acumulativa de archivos digitales de impresos que habilite el más amplio acceso público a ellos, tanto por su valor didáctico como por lo que podrían aportar a la difusión de aspectos del pasado cultural de una comunidad, tal como quedó registrado en periódicos, libros y otros documentos. En este caso se hace desde la Facultad de Información y Comunicación (FIC) de la Universidad de la República, inicialmente a partir de un seminario de grado en comunicación y ahora en el marco institucional de la Facultad. Todo lo digitalizado está disponible, en forma gratuita, en el portal web Anáforas: en total hay unos 38.000 documentos, pero su número aumenta todos los días.

<sup>1</sup> Por Anáforas, <<http://anaforas.fic.edu.uy/jspui/>>.