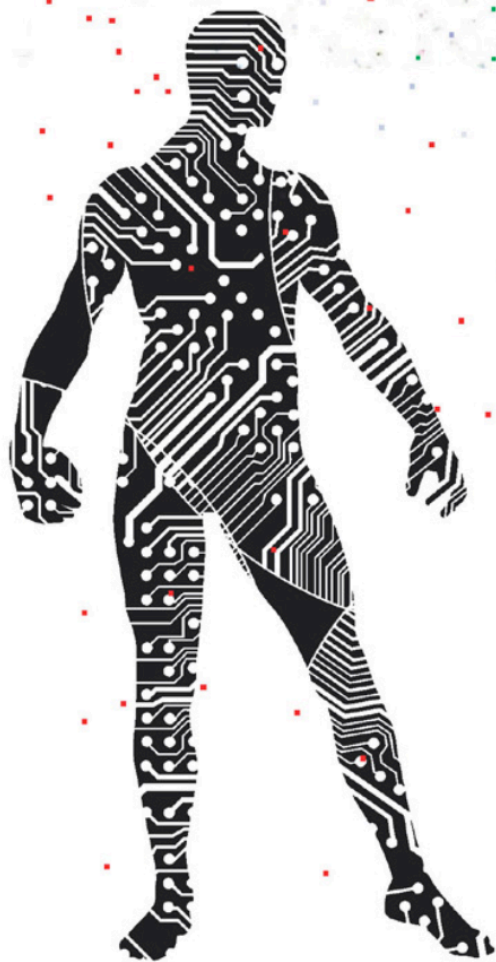


# HUMANIDADES

# DIGITALES

La cultura frente a las nuevas tecnologías

DOMINIQUE VINCK



gedisa  
editorial



# HUMANIDADES

# DIGITALES

La cultura frente a las nuevas tecnologías

DOMINIQUE VINCK

gedisa  
editorial

Título original en francés: *Humanités numériques Le Cavalier Bleu, Paris*  
© Dominique Vinck, 2018

Corrección: Borja Criado

Cubierta: Juan Pablo Venditti

Primera edición: abril, 2018

Reservados todos los derechos de esta versión castellana de la obra.

© Editorial Gedisa, S.A.  
Avda. Tibidabo, 12, 3º  
08022 Barcelona (España)  
Tel. 93 253 09 04  
gedisa@gedisa.com  
www.gedisa.com

Con la colaboración de



Preimpresión:  
Moelmo, SCP

ISBN: 978-84-17341-10-7  
Depósito legal: B.6716-2018

Impreso por Ulzama

Impreso en España  
*Printed in Spain*

Queda prohibida la reproducción total o parcial por cualquier medio de impresión, en forma idéntica, extractada o modificada, en castellano o en cualquier otro idioma.

# Índice

<b>Introducción</b> .....	7
<b>¿Qué son las humanidades digitales?</b> .....	11
«Con lo digital entramos en una nueva civilización.» .....	13
«Las humanidades digitales son la desmaterialización del patrimonio cultural.» .....	21
«Las humanidades digitales son la cuantificación de las humanidades.» .....	33
«Las humanidades digitales son cosa de hombres de letras que juegan a ser <i>geeks</i> .» .....	41
«Las humanidades digitales son un efecto de Internet.» ..	49
«Las humanidades digitales son también una cuestión de las ciencias sociales.» .....	57
<b>¿Para qué sirven las humanidades digitales?</b> .....	65
«Las humanidades digitales salvarán las humanidades.» ..	67
«Las humanidades digitales democratizarán el saber, la cultura y el acceso al patrimonio cultural.» .....	73
«Las humanidades digitales harán posible el diálogo entre los pueblos y las culturas.» .....	79
«Con las humanidades digitales, las herencias culturales se volverán nuevos recursos para la innovación y el desarrollo económico.» .....	85
<b>¿Temor fundado?</b> .....	91
«Con lo digital vamos a perder mucho.» .....	93

«Lo digital supone la muerte del libro y las bibliotecas.» . .	103
«Las tecnologías digitales replazarán a los investigadores de las ciencias humanas y sociales.» . . . . .	109
<b>Las humanidades digitales en el mundo</b> . . . . .	117
«Con las humanidades digitales se despoja al Sur de su patrimonio cultural.» . . . . .	119
«Con lo digital, la hegemonía del inglés es cosa del pasado.» . . . . .	129
«Las humanidades digitales crearán nuevas brechas.» . . . .	135
<b>Conclusión.</b> . . . . .	141
<b>Para saber más</b> . . . . .	145
Libros sobre lo digital, Internet y el acceso abierto. . . . .	145
Libros sobre las humanidades digitales. . . . .	146
Revistas . . . . .	149
<b>Bibliografía y proyectos en la Web.</b> . . . . .	151
Recursos digitales generales . . . . .	151
Proyectos . . . . .	151
Ejemplos de recursos especializados (texto, imágenes) . . . .	153
Visualización . . . . .	154
<i>De datos convencionales</i> . . . . .	154
<i>De sitios web</i> . . . . .	154
<i>De datos de Facebook</i> . . . . .	154
<i>De datos de Filckr</i> . . . . .	154
<i>De datos de Twitter</i> . . . . .	155
Herramientas . . . . .	155
Instituciones y redes . . . . .	156

## Introducción

Los promotores de las humanidades digitales las consideran compuestas por todas las disciplinas científicas que capturan, analizan y presentan las dinámicas culturales y sociales pasadas, presentes y emergentes, mediante herramientas informáticas y el cálculo. Se refieren a todas las ciencias humanas y sociales (CHS) y a las actividades patrimoniales (incluyendo los archivos y las bibliotecas) culturales y sociales que ellas tratan (por ejemplo, redes sociales informáticas). Forman un espacio interdisciplinario que reúne CHS y ciencias y tecnologías de la información (CTI) en la investigación y desarrollo de nuevas herramientas.

El campo tiene algunas de sus raíces en los principios de la computación, hacia los años cuarenta, cuando el jesuita Roberto Busa convenció al fundador de IBM para elaborar concordancias de la obra de santo Tomás de Aquino (*Index Thomisticus*), es decir, la vinculación de cada palabra a su contexto en el texto utilizando procesos computacionales. Emerge bruscamente desde los años sesenta, en lingüística, en historia y en sociología en particular, con el desarrollo de métodos cuantitativos y la programación informática para el procesamiento automático de *corpus* de textos y datos. Esto condujo, por ejemplo, a la creación del *Trésor de la langue française*, diccionario de 100.000 palabras del siglo XIX y XX, con su historia, sus definiciones (270.000) y ejemplos, que permite búsquedas complejas en línea o en la *Base historique du vocabulaire français*, pues registra los trazos de las primeras apari-

ciones de citas de las palabras, su grafía, fonética, morfología y semántica.

Luego siguió una serie de acontecimientos relacionados con la difusión de los ordenadores personales, nuevos idiomas para la codificación de información (SGML, XML), incluyendo de textos (*Text Encoding Initiative* - TEI), el establecimiento de enlaces de hipertexto, la diseminación de Internet, de la Web y de los motores de búsqueda. Todo esto conduce a la creación de archivos de textos, imágenes y materiales audiovisuales que se pueden buscar en línea, por ejemplo, sobre las mujeres escritoras de la época previctoriana (<http://www.wwp.northeastern.edu/>); la obra del poeta romántico William Blake con sus propias anotaciones y los análisis que han publicado los investigadores (<http://www.blakearchive.org/blake/>); las historias paralelas de dos comunidades norteamericanas durante la guerra civil (*The Valley of the Shadow*: <http://valley.vcdh.virginia.edu>) o el análisis de los relatos producidos durante las elecciones presidenciales de Estados Unidos vistos en forma de redes de personas, de acciones (verbos) y de objetos.

Por otra parte, las grandes empresas como Google y Facebook han crecido mediante el aprovechamiento de las teorías y métodos desarrollados por las CHS, especialmente en lingüística computacional y análisis de redes sociales. Generan beneficios considerables por hacer el tratamiento por ordenador del patrimonio cultural y los rastros de acciones e interacciones sociales (incluyendo lo que escriben quienes navegan en Internet, las llamadas telefónicas o los desplazamientos de los coches o teléfonos con sistemas de geolocalización).

Estas humanidades digitales también se debaten. Ya rodeadas de ideas acerca de sus beneficios y sus daños, plantean preguntas



que merecen ser discutidas. Este libro trata de aportar algunos elementos de reflexión sobre ellas. Se busca iluminar lo que se está haciendo y por qué es objeto de debate. ¿En qué sentido son nuevas? ¿Para qué sirven? ¿Qué aportan como solución o problema?

Las humanidades digitales ya nos acompañan todos los días. Incluso se han convertido en un asunto mundial. Los políticos se preocupan por ellas y financian su desarrollo. Los investigadores, empresas e instituciones públicas compiten por ocupar los mejores asientos y obtener beneficios económicos e ideológicos. Algunos grupos sociales se apoderan de ellas y cuestionan la legitimidad de los investigadores. Surgen nuevas profesiones que amenazan las instituciones y otras profesiones muy consolidadas. Las humanidades digitales ahora afectan a todo el mundo porque tienen que ver con nuestro patrimonio y nuestras identidades hasta el punto en que se relacionan con algunas tragedias recientes, como la matanza de personas en nombre de argumentos religiosos, reformados y distribuidos por las tecnologías digitales. Influyen en la formación humana de los miembros de nuestras sociedades y en el desarrollo de la oferta educativa. También alimentan demandas de autonomía y justicia, y contribuyen a la transformación de los equilibrios culturales en el mundo. Desempeñan opciones fundamentales para la sociedad como son: la extensión de la economía del mercado hacia el patrimonio cultural, o, por el contrario, el intercambio abierto y generalizado de un patrimonio considerado básicamente como un bien común. Las humanidades digitales son entonces un asunto de sociedad, de comunidades, de ciudadanos, de generaciones, de consumidores de bienes culturales y de empresas a veces hegemónicas.

No son cosas abstractas que existen sólo entre algunos estudiosos. Este libro se ocupa de lo que ocurre en el mundo de hoy alre-

dedor de las humanidades y de las ciencias sociales que se vuelven digitales y de lo que preocupa a sus actores: aspectos de política científica y cultural, de creación de nuevos mercados, de relaciones hegemónicas, de patrimonio y de nuevas sociedades en proceso de devenir, además de los debates entre investigadores.

Los análisis presentados en el libro se alimentan de la investigación en sociología de la ciencia y la innovación, que llevo a cabo personalmente en el campo de las humanidades digitales. También de las posiciones que he ocupado en la estructuración de la investigación y en comités estratégicos donde se estableció relación con actores de diferentes orígenes: investigadores en CHS e CTI, grupos sociales interesados tanto en los países del Norte como en los del Sur.

En primer lugar, examinaremos algunas ideas básicas sobre el tema, y después de los usos de las humanidades digitales, de los riesgos que conllevan y del estado de su desarrollo en el mundo.

# ¿QUÉ SON LAS HUMANIDADES DIGITALES?



## «Con lo digital entramos en una nueva civilización.»

«Esto no es una revolución digital,  
sino una civilización digital.»

Irina Bokova, directora de la Unesco

Cuando surge una nueva tecnología, los medios de comunicación a veces hablan de *revolución tecnológica*, es decir, de una transformación de toda la sociedad, su economía y su funcionamiento social y político, como fue el caso de la máquina de vapor, el desarrollo de las minas de carbón, la industria y el surgimiento de una nueva clase social (el proletariado). Tales revoluciones rara vez se explican por la mera invención de una nueva tecnología. En general, el interés por ésta tiene que ver con las transformaciones que ya estaban ocurriendo en la sociedad antes de su invención, mientras que la capacidad de la tecnología innovadora para transformar el mundo depende de la invención de nuevas prácticas, habilidades y oficios, formas de organización del trabajo, de democracia, de comercio, de relaciones sociales, etcétera.

En cuanto a lo digital, las profecías van aún más lejos al hablar de la era digital, como se habla de la era cristiana desde el siglo XVI, a raíz de la diseminación del libro impreso, en referencia a la cultura del libro que se había desarrollado en Occidente. El concepto de

«era digital» apareció en el mundo de las telecomunicaciones en la década de 1980 y se asoció con el anuncio de que, veinte años más tarde, todos los intercambios serían digitales. Gran parte de las relaciones humanas estarían vinculadas a las tecnologías digitales, así como todas las prácticas, estilos de vida, valores, creencias y conocimientos constitutivos de una *cultura digital*. Por lo tanto, con lo digital, se habría formado una nueva manera común de ser, de pensar, de actuar y de comunicar.

Hablar de *civilización digital* conjetura que esta cultura digital no es el acto de un grupo social distinto, sino que caracteriza el estado técnico, intelectual, político y moral de toda una sociedad. El término «civilización» se utiliza para hablar de los cambios importantes en sociedades como la transición de la barbarie al estado civilizado, o de las civilizaciones agrarias a las civilizaciones urbanas. En *Choque de civilizaciones*, Samuel Huntington identifica ocho civilizaciones vinculadas a regiones del mundo y con frecuencia a las religiones: la civilización occidental, latinoamericana, soviética, africana, musulmana, hindú, budista y china. Con la globalización de los mercados, del flujo de mercancías y finanzas, y de las tecnologías digitales, se plantea la cuestión de si, en efecto, estamos entrando en una nueva civilización llamada digital.

Este tipo de anuncio de un cambio radical de la sociedad, la cultura, la era o la civilización no es nuevo. El advenimiento de la radio llevó a la profecía de una sociedad de comunicación global y de la circulación de las ideas mediante la construcción de una noosfera (esfera del pensamiento humano) en todo el mundo. En 1934, Paul Otlet profetizó la desaparición del libro en favor de fichas microfotográficas transmitidas a través de un «telescopio eléctrico» (televisión) que permitiría la lectura en casa de libros «telefotados»,

mientras que los centros de documentación, unidos entre sí, formarían una red de documentación universal, cambiando radicalmente la relación de los humanos con el conocimiento y entre sí. En la década de 1990, con el advenimiento de Internet, se popularizó la idea de «comunidad virtual» y de «aldea global», que revalorizaban la proximidad en lugar de los lazos funcionales y contractuales predominantes en la sociedad. Con esta tecnología estaríamos asistiendo a la fundación de una convivencia más satisfactoria, vinculada a la posibilidad de compartir en línea una comunicación sincera, inmediata y emocional; el advenimiento de un ser social fortalecido.

Estos anuncios de cambios radicales de la sociedad relacionados con la introducción de nuevas tecnologías no sólo no son nuevos, sino que, además, rara vez se han materializado. Los cambios ocurrieron, de hecho, pero la nueva sociedad rara vez es lo que se había anunciado. Ni el teléfono ni la televisión ni Internet han llevado hasta ahora la aldea global a donde las ideas y emociones circulan de manera sincera e inmediata, uniendo a las personas en una comunidad. Hay muchos cambios, pero rara vez están a la altura de las expectativas o utopías proclamadas. La tentación mayor es entonces relativizar la magnitud de los cambios que conllevan las tecnologías digitales.

La relativización de los efectos sociales de la tecnología es aún mayor cuando los cambios esperados provienen de la acción de grupos y movimientos sociales antes de ser efectos tecnológicos. El sueño de una comunidad virtual es el resultado, en particular, de los movimientos de protesta y de la experimentación, en la década de 1960, de formas de sociedades alternativas. Estos movimientos alimentaron la acción de los pioneros de Internet, el diseño de los ordenadores y la configuración de la red sociodigital. Después, frente

al hallazgo de un vínculo social debilitado, sobre todo en sus formas colectivas, se cuestionarán las tecnologías digitales de la información y de la comunicación por sus efectos positivos en términos de revitalización del vínculo social ahora en línea, o negativos por la individualización y fragmentación de lo social. Los investigadores de las ciencias sociales estudian las nuevas formas colectivas, mientras que los gobiernos invierten en infraestructura para supuestamente favorecer esas nuevas formas sociales, tanto a nivel local (*smart cities* - ciudades inteligentes) y nacional (nueva forma de democracia), como internacional (*digital divide* - la reducción de la brecha digital). Las tecnologías digitales están bien provistas de expectativas y esperanzas; muchos creen que por fin van a traer la salvación: una sociedad mejor.

Sin embargo, una vez pasada la emoción de la novedad, las desilusiones llegan una tras otra: las comunidades virtuales son efímeras; son comunidades de interés que incluyen de forma transitoria individuos débilmente unidos; Internet se ha convertido en un área de comercio y de relaciones hegemónicas, entiéndase de bandolerismo; sirve principalmente a algunos (ya sean consumidores de información sobre foros o grandes empresas que implementan las herramientas del *big data*) para hacerse con el valor agregado obtenido de las acciones y las huellas producidas por otros; permite la fidelización de las relaciones y recomendaciones comerciales (sugerencias de sitios para visitar, de compras, de lectura, de música, de «amigos» o de almas gemelas) a la imagen del perfil que deja rastros en las herramientas digitales; fortalece al mismo tiempo las identidades de cada uno en lugar de invitar al descubrimiento de otros y de diferentes pensamientos.

Tal desencanto (ver el estallido de la burbuja financiera ligada a Internet en 2000) se refiere a las tecnologías que ya están allí



y cuyos límites ha demostrado la experiencia, pero nuevas tecnologías llegan, y ellas deberían, por fin, traer el cambio; hablamos de la Web 2.0 (conocida como social) y Web 3.0 (Web Semántica o la Internet de las cosas). Web 2.0 supuso un hito para el desarrollo de herramientas (blogs, sitios para compartir, wikis, sitios de redes sociales) que facilitan e impulsan la edición y la manipulación de contenido en tiempo real. Permite a una amplia audiencia aportar contenido (texto, fotos, vídeo, etcétera) y modificarlo. Internet se convierte en un nuevo espacio de intercambio que permite la libre circulación de las ideas y la información producidas por un sinnúmero de usuarios. Estos aportes agregados y procesados hacen posible el diseño y la implementación de nuevas actividades o servicios (producción colectiva de conocimiento o novelas, coordinación horizontal de acciones colectivas, organización de intercambios, incluidos comerciales, ilegales, etcétera). Los observadores demuestran qué sucede actualmente: nuevos modelos económicos o libres (Wikinomía); una organización social reticular; nuevos poderes (*empowerment*) de los ciudadanos (democracia directa, cortocircuito de los medios de comunicación tradicionales con WikiLeaks), de los clientes (cortocircuito de las empresas y profesiones establecidas, incluidos los conductores de taxi o agencias de viaje) y de los pacientes (desafío a la experiencia médica); así como producciones de otro modo inimaginables tales una enciclopedia (Wikipedia) o innovaciones (*crowdsourcing* y la innovación abierta) en manos de la multitud de fans en lugar de expertos en Web o investigadores profesionales. La comunidad se convierte en la multitud de individuos que contribuyen voluntariamente a un mismo proyecto, por ejemplo, la transcripción de una obra colosal (véase el *Jeremy Bentham project*) que los investigadores no serían capaces de realizar por fal-

ta de personal suficiente. La multitud se moviliza para contribuir libremente con su trabajo, conocimiento, información, ideas y finanzas (*crowdfunding*).

Las críticas y desilusiones de la Web 2.0, sin embargo, alcanzan una mayor amplitud: poner a trabajar a las multitudes y sus contribuciones voluntarias enmarcadas por las herramientas y las organizaciones (a menudo empresas) activas en la Web 2.0 es objeto de controversia; redes sociales digitales que se convierten en poderosos vectores de conformación de normas sociales identitarias; los gigantes de Internet animan a los usuarios para que dejen aún más rastros de sus acciones e interacciones (compras, desplazamientos, consultas, etcétera), y por lo tanto proporcionen un nuevo oro negro, haciendo posible el capitalismo cognitivo, contra el cual emergen varias formas de resistencia. La esperanza se desplaza entonces hacia la Web 3.0.

Todo esto sugiere que los cambios pueden no ser tan importantes como se había planeado, o no van necesariamente en la dirección de una civilización mejor.

Sin embargo, parece que los cambios son de tal magnitud que sí podemos hablar de una nueva civilización. Las huellas y los datos conducen a nuevas formas de agregación de las intenciones y los comportamientos, a la instrumentalización de la sociabilidad y la confianza, y a un nuevo orden social basado en los datos y su procesamiento algorítmico, lo que genera una prescripción (recomendación), un valor (económico) y una clasificación (medida de la grandeza de las personas a partir del número de «amigos» o «like» o citas aparecidas en índices de publicaciones científicas). La medida generada por las tecnologías digitales y las empresas se vuelve una herramienta de persuasión. La relevancia, la legitimidad, des-

de el cómo hacer la sociedad hasta la identidad son en gran medida sugeridas por las tecnologías digitales que constituyen cada vez más nuestro medio ambiente. Potencialmente, esto afecta a todos los aspectos de la vida en la sociedad, incluso los objetos, lo que contribuye a recomponer el mundo haciendo de todas las cosas una interfaz digital que se conecta con códigos, con el *big data* (grandes volúmenes de datos), el *long data* (datos históricos) e infraestructuras y un nuevo capitalismo que media nuestra relación con el mundo, los demás y nosotros mismos. El humano se redefine, el humanismo, según Milad Doueihi, se vuelve digital después de haber sido clásico (retorno letrado hacia la Antigüedad y el análisis crítico contra la hegemonía de la Iglesia), exótico (descubrimiento del otro y de las diferencias de la humanidad) y democrático (lo que incluye a todas las personas y todos sus hechos). El humanismo digital estaría preparando una nueva civilización, es decir, una nueva condición de existencia con sus transformaciones de categorías, valores, relaciones, objetos, representaciones, territorios y prácticas. El humanismo digital, virtual o numérico sería un nombre equivocado si nos olvidamos de que la tecnología digital está hecha de técnicas corporales (gestos de lectura con los dedos y de escritura, comunicación y desplazamientos, etcétera), interfaces, códigos, plataformas, servidores de datos e infraestructuras cada vez más pesadas tanto física como energéticamente.



## «Las humanidades digitales son la desmaterialización del patrimonio cultural.»

«La gente no tiene nada que hacer con los medios físicos.»

Désilets, *Journal du gamer*, 28 de junio de 2013

Desde la década de 2000, cuando hablamos de lo digital pensamos en la «desmaterialización». Los gruesos libros de papel y el mobiliario de la biblioteca son reemplazados por archivos de ordenador que caben en dispositivos muy pequeños; ahora podemos caminar con bibliotecas inmensas en nuestro bolsillo. La música guardada en los estantes de los CD se encuentra también en el mismo dispositivo, o bien tenemos acceso inmediato a ella desde cualquier lugar con una conexión a Internet. Los grandes edificios de nuestras herencias culturales, catedrales que contienen millones de libros, pueden ya no ser necesarios. Del mismo modo, nuestras hemerotecas, equipadas con una infraestructura capaz de controlar la temperatura, la humedad y los posibles ataques de fuego, seres humanos, roedores o insectos, conforman kilómetros de cajas de cartón que preservan los rastros de nuestra historia; podrían ser digitalizadas, desmaterializadas e inmediatamente accesibles desde cualquier lugar.

Esta desmaterialización implica la digitalización del patrimonio cultural, es decir, la conversión de ciertos aspectos de los obje-

tos culturales (por ejemplo, el contenido textual de libros, la música grabada en discos de vinilo, las imágenes guardadas en viejas fotos amarillentas y dobladas o la forma de cerámicas) en un conjunto de datos digitales que sean: caracteres alfabéticos e información de formato en el caso de un texto; una paleta de sensaciones visuales para cada punto (*pixel*), o bien una serie de instrucciones para compresión / descompresión o su vectorización, en el caso de las imágenes. A partir de esos datos se podrían reconstruir los objetos digitalizados. La calidad de la digitalización depende en particular de la finura del muestreo (el número de puntos por centímetro cuadrado en el caso de una imagen). Estos datos digitales se traducen, en los circuitos electrónicos, en señales (impulsos eléctricos que circulan o cambios en el estado del material en un lugar específico, por ejemplo, en la memoria de un servidor, en un disco duro o en un CD). Son manipulados por medio de conjuntos de instrucciones (programas informáticos) que los agregan, comparan, clasifican, almacenan, procesan, etcétera. Los programas de ordenador procesan la información a alta velocidad o de manera repetitiva (varios miles de operaciones por segundo). Los dispositivos electrónicos e informáticos pueden así mantener y gestionar grandes cantidades de información de la digitalización y luego procesarlas, por ejemplo, para mostrar colores, imágenes y texto en la pantalla, difundir sonido por nuestros altavoces o activar una impresora 3D.

La desmaterialización consiste en sustituir los soportes de almacenamiento de hoy en día, a menudo de papel, por archivos de ordenador. Antes de los ordenadores, estos archivos estaban hechos de papel. Ya Paul Otlet, en la década de 1930, había creado el *Mundaneum*, que reuniría todo el conocimiento del mundo. Se creó un

sustituto del libro sobre un soporte dinámico, flexible y actualizable: la ficha que resume el conocimiento contenido en los libros. Las fichas, clasificadas por un sistema de índice universal, formarían una enciclopedia mundial, nunca terminada, que aumentaría constantemente en función del progreso del conocimiento. Los libros serían microfotografiados, lo que permitía reproducir en dimensiones muy pequeñas sus páginas y mantenerlas en rollos de película que luego se presentarían frente a un equipo de ampliación en el momento de la lectura, y se proyectarían en una pared a través de la «bibliofoto» (máquinas que leían bibliopelículas). Por otra parte, el telescopio eléctrico permitiría el acceso remoto a los libros y las fichas; los libros ya no saldrían de las bibliotecas, ya que se podrían leer desde casa, en pantallas, después de pedir a la biblioteca, por teléfono, los libros o las páginas que «telefotear».

«El escritorio ya no tiene un libro. Su lugar lo ocupan una pantalla y un teléfono. Allí, en la distancia, en un enorme edificio, están todos los libros y toda la información [...]. A partir de ahí se manda en las pantallas la página a leer para responder la pregunta planteada por teléfono.» (Otlet, 1934; pág. 428). Además, los centros de documentación, unidos entre sí, formarían una «red universal de documentación».

El desarrollo de la tecnología informática en la década de 1940 contribuye de manera significativa a la «desmaterialización» de los libros y documentos en papel. A finales de 1960, algunas iniciativas tratan de superar las limitaciones físicas de la información impresa pasando los textos a archivos digitales almacenados en un disco flexible (15") que contiene el equivalente de un centenar de páginas, más ligero para llevar y fácilmente duplicable.

A principios de 1970, el Proyecto Gutenberg inicia el escaneo de libros existentes en papel con la intención de formar una biblioteca digital universal y accesible para mucha gente. Miles de voluntarios (52.000 en 2008) escanean los libros, produciendo imágenes de cada página, luego los «ocerizan», es decir, les aplican un reconocimiento de caracteres (OCR) que transforma la imagen del texto en letras, palabras y textos, y por último hacen la corrección de lo que producen estas operaciones (ciertas letras o palabras están distorsionadas o no son reconocidas), dan formato al texto y editan la versión electrónica de estos libros para diseminarlos gratuitamente (en forma de CD, que contiene hasta 600 libros, o de DVD, que puede contener 17.000 libros). El proyecto se basa en obras de dominio público y pretende dar al público el gusto por la lectura y la herencia literaria. Después de la digitalización de los textos de gran valor simbólico (Declaración de los Derechos Humanos, la Biblia) o literario, el proyecto se extiende a diccionarios y enciclopedias y obras de entretenimiento (*Alicia en el país de las maravillas*, *Las aventuras de Robin Hood*).

La desmaterialización no se trata sólo de los libros. Se refiere en primer lugar a los metadatos, es decir, los datos sobre los objetos recogidos (libros, pinturas, películas, la evidencia arqueológica, etcétera) y sus inventarios (catálogos). Esto se conoce como la desmaterialización de los datos documentales. En Francia, el Inventario General del Patrimonio Cultural, por ejemplo, se inició en la década de 1980 con la creación de los bases de datos Mérimée (arquitectura), Palissy (para objetos móviles), Mémoire (para la ilustraciones) y Archidoc (para las referencias de los documentos). Los metadatos de estos objetos y colecciones son registros descriptivos de su contenido, no los objetos en sí. Permiten a los archivistas



y bibliotecarios, y también a los investigadores y el público en general, buscar objetos o documentos y localizar dónde se encuentran (en qué biblioteca, museo o archivo). Las preocupaciones de los profesionales en cuestión se centrarán principalmente en la normalización del vocabulario y la interoperabilidad para que el uso de sus datos sea posible mediante sistemas técnicos heterogéneos. Trabajan también en la sostenibilidad del archivado de datos a pesar de los rápidos cambios tecnológicos.

Más allá de los metadatos están también los mismos objetos que estas instituciones están digitalizando. Numerosas bibliotecas y hemerotecas han empezado la digitalización de libros, periódicos, mapas, grabados, fotografías, películas, etcétera. Surgen también iniciativas académicas y/o comerciales que compiten con el proyecto Gutenberg, y proliferan los libros electrónicos y sus versiones. Hablamos entonces de millones de documentos que se escanean y catalogan para alimentar bases de datos enormes.

Un importante punto de inflexión se produce en la década de 1990 con el desarrollo de Internet y el acceso a bases de datos en línea. Miles de bibliotecas y hemerotecas se embarcan en la aventura de poner a disposición su material en la Web, lo que implica el diseño y la programación de interfaces entre la Web y las bases de datos que permitan la búsqueda y consulta a la distancia. La Biblioteca Nacional de Francia (BNF) crea Gallica en 1997, lo que proporcionará más de dos millones de libros, periódicos y mapas accesibles en formato de imagen y texto. En 2015 está compuesta por 1,3 millones de números de periódicos y revistas, más de 30.000 grabaciones, 60.000 manuscritos, 350.000 fotos de objetos (monedas, muebles, muestras de tejidos, trajes, instrumentos, etcétera), 800.000 imágenes, casi 40.000 partituras y 80.000 mapas.

Mientras tanto, los editores empiezan a publicar libros digitales. En 1993, la Online Books Page enumera los libros electrónicos disponibles libremente en la Web. En 1994, Amazon lanzó una librería comercial en línea (aparecen 28 millones de libros impresos en 2000 y 1000 *e-books*, y después de haber lanzado su propio lector de libros electrónicos (*Kindle*) en 2009, un catálogo de 230.000 *e-books*). Además de estos enormes catálogos y tiendas en línea, la compañía se expandió y anima sus páginas en la Web mediante la presentación de resúmenes de libros, consejos, entrevistas con los autores y comentarios de los lectores. Otras librerías virtuales o híbridas (Barnes & Noble, Alapage, FNAC.com, Chapitre.com, etcétera) también se lanzan y compiten con las librerías tradicionales.

En ese momento, en 1998, se creó una pequeña compañía llamada Google que proporciona en línea un motor de búsqueda y pretende organizar la información y su acceso al mundo. Las librerías virtuales desarrollan entonces una función de búsqueda directa en el texto de los libros que venden.

Desde la década de 2000 se ponen en marcha en todo el mundo numerosas operaciones de digitalización de diferentes herencias culturales, mientras aumenta el número de libros digitalizados: en 2002, las obras puestas en línea por el Proyecto Gutenberg representaban una cuarta parte de las obras de dominio público de libre acceso; en 2009, ofrecía 75.000 libros digitalizados por otras instituciones. Entre los proyectos de grandes magnitudes, podemos mencionar:

- **Google Book Search:** en 2006 digitaliza grandes bibliotecas y las indexa a su contenido a través de un robot (*crawler* o rastreador) que evalúa la proximidad entre las obras a partir de las

citas que reciben. Más tarde ofrece la oportunidad de compartir extractos (*snippets*), construir una biblioteca personal (My library), buscar libros a partir de un lugar en Google Earth, imprimir libros bajo demanda, traducir porciones de texto, resaltar, anotar y editar el texto. En 2012, Google Books supera los veinte millones de libros (más de 6 mil millones de páginas), 35.000 editores y 40 bibliotecas de 100 países. Esta compañía se convirtió en la biblioteca más grande de la historia, pero da acceso público a sólo una pequeña parte (un millón) de su fondo. Con este fondo, al que asocia las palabras que los internautas utilizan para consultar las obras, constituye un capital lingüístico.

- **Open Content Alliance (OCA)**: en respuesta a Google Books, ya en 2006, esta asociación americana incluye bibliotecas, organizaciones gubernamentales y culturales, archivos públicos, asociaciones y empresas (Adobe, Hewlett Packard, Microsoft, Yahoo!, Xerox, Internet Archive) para escanear y cargar una base de datos compuesta de documentos multimedia, accesible para todos, sostenible y multilingüe, con la aprobación previa de los editores de libros con derechos de autor, que se puede consultar y descargar a partir de varios motores de búsqueda.
- **Europeana**: también en respuesta a la iniciativa de Google, las bibliotecas nacionales y los jefes de Estado de varios países europeos crean una biblioteca digital europea que, en 2008, abrió un portal multilingüe. Los participantes incluyen las bibliotecas nacionales de 27 países europeos y varios países de América Latina, Quebec, y la Biblioteca de Alejandría en Egipto. En 2011, el proyecto se presentó con el anuncio de 24 millo-

nes de horas de documentos multimedia, 75 millones de obras de arte, 77 millones de libros, 358 millones de fotografías y 10 mil millones de páginas de archivo. En 2013, son 1.500 las instituciones involucradas en construir un catálogo en línea de más de 26 millones de objetos digitales, accesibles a través de las instituciones que los han digitalizado. El objetivo es alcanzar los 30 millones de objetos en línea en 2015 con un coste total de 100 millones de euros.

Los documentos archivados en formato digital también se procesan de múltiples formas para que estén disponibles para diferentes públicos y como productos específicos: registros descriptivos y de indexación, documentos de origen en múltiples formatos, recorridos virtuales en el patrimonio orientados al turismo (por ejemplo, en un patrimonio industrial regional), ensamblajes temáticos, atlas con indexación de objetos dentro de un sistema de información geográfica (SIG), o presentaciones multimedia escenificadas.

Este movimiento está lejos de terminar. Después de casi medio siglo de digitalización, en 2014 se estima que sólo el 20% del patrimonio cultural europeo ha sido ya digitalizado: el 13% de los libros y el 3% de las películas.

Si continúa, ¿eso significa el fin de los medios físicos? No parece, por dos razones. La primera es que la digitalización nunca captura más que una parte del objeto digitalizado. Escanear un libro, por ejemplo, consiste por lo general en convertir el texto impreso en texto digital, pero deja de lado toda la información con respecto a la consistencia del papel, el tipo de encuadernación, la degradación sufrida por el papel, las anotaciones hechas por lectores, etcétera. Estos elementos son el objeto de ciencias específicas, como

la codicología, que estudia los manuscritos como objetos materiales o la paleografía, que estudia la formación de las letras y los aspectos físicos de la escritura. Estos aspectos son a veces objeto de otra digitalización que conduce a más representaciones digitales del mismo objeto físico en función del tipo de información que ha sido digitalizado. Además, las técnicas de digitalización evolucionan y, a veces, es útil volver a los objetos físicos para volver a digitalizarlos con mejoras técnicas. Por estas razones, como medida de precaución, los conservadores generalmente prefieren mantener estos objetos incluso cuando han sido digitalizados.

La segunda razón es que hay un público que prefiere la versión física a la versión digital, ya sea por razones de conveniencia (la materialidad del objeto ofrece oportunidades que la materialidad digital no ofrece) o necesidad (lo digital depende de conexiones con las infraestructuras sociotécnicas eléctricas e informáticas que no siempre están disponibles ni en todas partes) o por razones sentimentales (para ver y mostrar su biblioteca, su colección de discos o DVD; oler el papel o sentir el peso el libro). La potencia de las conexiones, por ejemplo, limita el acceso a contenidos desmaterializados tales como películas de alta definición, o algunos juegos que pueden tardar horas o incluso días en descargarse; esta potencia depende ampliamente del lugar, incluso dentro del territorio europeo.

La desmaterialización nos priva de una parte de los atributos físicos de los objetos que muchas veces son importantes para los usuarios. En 2014, en Francia, los productos culturales físicos se seguían utilizando más que su versión «desmaterializada»: el 28% de los usuarios dicen consumir varias veces a la semana los productos culturales físicos frente a sólo el 21% que hacen uso de los productos culturales digitales. Respecto al papel, a pesar de que la llegada

de la microinformática en la década de 1970 se asoció con la idea de la oficina sin papel, éste no sólo no ha desaparecido, sino que el número de páginas impresas sigue aumentando un 3% por año, mientras que el consumo de papel ha aumentado un 40% desde la aparición del correo electrónico. Con el desarrollo de Internet en la década de 2000 se pensó que algunos medios físicos caerían en declive y se volverían obsoletos.

En términos de patrimonio cultural, la resistencia a la digitalización también se debe a la alta valoración de la materialidad y sus funciones nemotécnicas, simbólicas e identitarias. El patrimonio material (el libro más que el texto, los objetos, el patrimonio inmobiliario) sigue siendo hegemónico; el patrimonio inmaterial (prácticas perpetuadas por las comunidades humanas que son portadores de éste y lo transmiten de generación en generación, de acuerdo con la definición dada por la UNESCO) sigue siendo el pariente pobre. Los amantes de la historia de la informática prefieren conservar ordenadores a programas, medios de almacenamiento físicos a conocimientos de programación. Los coleccionistas prefieren los viejos discos de vinilo que perpetuar la atmósfera de un concierto. Los museos etnológicos regionales aún prefieren exponer las herramientas antiguas y no los gestos de los artesanos, mostrar los trajes de época antes que revivir las antiguas fiestas o las procesiones. ¿El aumento del patrimonio digital, que preserva el contenido más que los soportes y que «desmaterializa» el patrimonio, socavará la hegemonía del patrimonio material a favor de una mayor atención al espectáculo (musical, profesional, relacional, social), a la emoción y al colectivo? La desmaterialización podría desfeticizar la materia y la conservación de los soportes y permitir a las sociedades dar más importancia al patrimonio inmaterial y vivencial. La

autenticidad y la integridad de una herencia tal vez podría redefinirse en la medida en que la desmaterialización relativiza su dimensión material.

El movimiento de la «desmaterialización» parece avanzar. Sin embargo, se enfrenta al hecho de que la digitalización no significa desmaterializar en el sentido de «privar a una cosa de cualquier materialidad». Si la desmaterialización es privar a una cosa de los atributos físicos de su soporte originario, en realidad consiste en una re-materialización. El libro ya no está hecho de tinta, papel y tapa dura, sino de archivos digitales, con formatos que se vuelven rápidamente obsoletos, y se almacenan en discos duros o servidores que hay que vigilar, ya que son frágiles, o circulan en redes de cables y ondas, pasando a través de fibras ópticas y satélites. Esta materialidad digital se vuelve relevante cuando se mide su consumo de recursos naturales (energía y tierras raras) y su impacto sobre el medio ambiente (desechos electrónicos). La desmaterialización de los intercambios ahora electrónicos, el acceso en línea a los datos, las búsquedas en Internet (6 mil millones de consultas realizadas en Google cada día en 2015) y los múltiples datos digitales que producimos todos los días (500.000 horas de comunicación de vídeo, actualización de 4 millones de horas de vídeo de *Youtube* y 7,5 mil millones de *Like* en Facebook) contribuyen a la necesidad de crear más y más centros de tratamiento y almacenamiento de datos (*data centers*). Estos centros incluyen ordenadores (servidores con almacenamiento rápido y seguro), herramientas de telecomunicaciones (*routers* y conmutadores), energía eléctrica y un generador de energía de reserva, un sistema de aire acondicionado para enfriar servidores, sistemas de aireación y de prevención y lucha contra los incendios que no dañen los ordenadores, así como sistemas de seguridad

contra los intrusos. Desde 2010 se estima que Google tiene más de un millón de servidores repartidos por todo el mundo, Microsoft llega a este número a partir de 2013, Amazon, 450.000 desde 2012, y Facebook, 180.000 desde 2013. Sólo los servidores web en el mundo son más de 230 millones desde 2010. Cada año se abren 10 millones de servidores. Habría alrededor de 6 millones de centros de datos en el mundo, con una superficie total de 150 km<sup>2</sup> (una vez y media la superficie de París).

Hablar de desmaterialización se vuelve irónico. Algunos incluso creen que esta desmaterialización no reduce el consumo de materias primas ni de energía. El sector de las tecnologías de la información nada tiene de virtual: servidores, antenas repetidoras, terminales, cables transoceánicos, fibras ópticas. La demanda de metales en este sector se ha triplicado entre 1980 y la década de 2010.



## «Las humanidades digitales son la cuantificación de las humanidades.»

«Es difícil cuantificar y sacar conclusiones precisas cuando se trata de un fenómeno cualitativo, pero la tentación de hacerlo persiste.»

Maria Konnikova, «Humanities aren't a science. Stop treating them like one». *Scientific American*, 10 de Agosto de 2012

Quien dice «digital» dice cifra. Las humanidades digitales se encargan, de hecho, de convertir los fenómenos sociales y culturales en series de datos digitales. Una vez hecho esto, el contenido de los textos, imágenes, películas e interacciones sociales se puede manipular por diversos procedimientos, cada vez más informatizados. Las palabras de un texto, incluso de miles de páginas, se cuentan a la velocidad del rayo, lo que permite descubrir cuáles son las más utilizadas, cuáles surgen tardíamente a través de las páginas, cuáles pierden su importancia o cuáles están asociadas entre ellas. El historiador Dominique Pestre analiza las tendencias del repertorio lingüístico del Banco Mundial contando las palabras que se usan en sus informes anuales durante 60 años. Encontramos predominantes algunas palabras reveladoras: banco, crédito, desarrollo, fiscal,

económico, financiero y privado. A éstas les siguen otras que especifican la naturaleza de las principales preocupaciones de esta institución: país, inversiones, intereses, programas, proyectos, asistencia, crecimiento, costo, deuda, comercio y precios. Pero muestra también que el Banco cambia. Al principio se utilizó un lenguaje más concreto que hablaba de carreteras, de riego, ferrocarril, agricultura, industria, etcétera, en relación con logros tangibles, pero estos términos han ido desapareciendo y han sido reemplazados por términos más abstractos, desconectados de la realidad local, y asociados a comportamientos y términos indeterminados: administración, gobierno, estrategia, ajuste, esfuerzo, promoción, responsabilidad, cooperación y participación, global y duradera. Este simple cómputo de palabras ya muestra importantes fenómenos sociales y políticos.

La cuantificación de los fenómenos sociales y culturales no es nueva. Los aritméticos políticos del siglo xvii, en ausencia de un censo, trataron de estimar la evolución de la población y se enzarzaron en discusiones para determinar si la ciudad más grande en el mundo era Londres o París. Otro ejemplo: en 1851, el matemático Augustus de Morgan se propuso llevar a cabo un análisis cuantitativo del vocabulario de las epístolas paulinas, mientras que en 1901, el físico y meteorólogo Thomas C. Mendenhall contó las palabras de las obras de Shakespeare y Bacon para analizar su frecuencia en función de su tamaño para determinar si Bacon fue el autor de ciertas obras atribuidas a Shakespeare.

A principios del siglo xx, las ciencias sociales e históricas buscaban «cientificarse», por lo que rechazaron las interpretaciones especulativas en favor de un enfoque neutral y objetivo, basado principalmente en los hechos y la medición. La cuantificación se convirtió, por tanto, en una misión política para racionalizar la di-

rección de los asuntos humanos y desideologizar las discusiones sustituyendo con la medición y el cálculo el arbitrio de las pasiones y la relación de fuerzas. Se creó una nueva forma de pensar, representar y actuar en el mundo.

En sociología recurrir a la cuantificación, las probabilidades y las matemáticas para descubrir, argumentar y explicar los fenómenos sociales es común desde la década de 1940. Se promovió la cuantificación en esta disciplina con la intención de probar hipótesis concernientes a relaciones causales (por ejemplo, entre el estatus social de los padres y la trayectoria educativa de los hijos) y poner de manifiesto la regularidad de ciertos procesos. Estos análisis tratan de descubrir las leyes que rigen las relaciones sociales reduciéndolas a diversidad, multiplicidad y singularidad para hacer posible el recuento y la medición basados en el procesamiento estadístico. En respuesta a la reducción de los fenómenos sociales a las variables, indicadores y relaciones causales, muchos autores señalan también la necesidad de un análisis cualitativo con el fin de obtener un significado en lugar de sólo medir los fenómenos. El análisis se hace entonces por impregnación, a base de hacer circular el material de investigación y de manipularlo, mientras que los enfoques cuantitativos delegan el tratamiento a procedimientos específicos, formales y explícitos, aplicados de forma sistemática (cómputo, inclusión de tablas, procesamiento estadístico o programas informáticos). Los métodos cuantitativos de hecho son métodos para ordenar los datos cualitativos.

En lingüística, el uso de estadísticas se origina al menos a principios del siglo xx, cuando Jean-Baptiste Estoup, taquígrafo francés, inventa los métodos de cálculo de las frecuencias de palabras y categorías gramaticales en un texto. Sin embargo, el uso de ordenadores no ocurrirá hasta la década de 1950 (*Computational linguis-*

*tics*), con la idea de traducir automáticamente textos de un idioma a otro, empezando con publicaciones científicas del ruso al inglés. Ya que el ordenador puede calcular mucho más rápido que los humanos y con menos errores, en aquella época se creyó que en pocos años se generalizaría la traducción automática de una lengua a otra. Cuando los expertos se dieron cuenta de la mala calidad de estas traducciones mecánicas, el tema pasó a ser **temade** investigación interdisciplinaria. La llegada de la inteligencia artificial en la década de 1960 llevó a formalizar la gramática, la morfología de las palabras, la sintaxis de las frases, la semántica (significado de las palabras), el léxico (vocabulario), la fonética y la pragmática (el estudio del contexto y de la manera de expresar las cosas). Desde hace 50 años se ha abierto un enorme campo de investigación interdisciplinario que se encarga del procesamiento de textos, la creación automática de resúmenes, el análisis del discurso, el estudio comparativo de las lenguas y su evolución en el tiempo, el reconocimiento y la síntesis de voz, la formulación automática de un léxico de base, la creación de avatares informáticos capaces de conversar con personas, el reconocimiento óptico de caracteres escritos por seres humanos, incluso en manuscritos antiguos, el análisis automático de millones de consultas realizadas por los internautas en los motores de búsqueda, de los tuits o de conversaciones telefónicas, etcétera. Los lingüistas y los investigadores de la rama de las Letras también trabajan en la constitución de importantes bases de datos lingüísticas, como el *Tesoro de la lengua francesa* o la *Penn Treebank*, que contiene más de 4,5 millones de palabras anotadas del inglés americano.

El uso de estos recursos digitales permite realizar estudios gigantescos. Los análisis digitales se utilizan para identificar fenómenos como la tendencia a acortar las palabras más usadas o para

umentar el contenido de información de un comunicado empleando palabras poco frecuentes.

En historia la informática también se traduce en la digitalización de archivos y la cuantificación de fenómenos. Una corriente de la historia cuantitativa, la Escuela de los Annales francesa, suscitó un gran interés a finales de 1960. Otra rama, la cliometría, incluso afirma que la historia tiene que transformarse en una verdadera ciencia mediante la producción de datos cuantitativos, procesándolos matemáticamente (mediante modelado y simulación informáticos) y construyendo teorías formales de los fenómenos históricos, como por ejemplo, la alternancia secular entre el crecimiento de la población y la inestabilidad política. Así, sería posible poner a prueba las predicciones hechas por dos teorías concurrentes. Para esto, los historiadores y arqueólogos deberían excavar en los archivos para extraer los datos. El análisis matemático de un gran número de eventos debería permitir que se estudie cómo las diferentes variables interactúan mediante la construcción de aproximaciones cuantitativas (*proxies*) de fenómenos (la población, la esperanza de vida, las estructuras sociales y las desigualdades, la estabilidad del Estado, los procesos judiciales, las conversiones religiosas, el clima, la productividad agrícola, etcétera). Jack Goldstone, miembro de la Political Instability Task Force, creada por la CIA para anticipar acontecimientos que pueden ocurrir fuera de los Estados Unidos, se ha centrado en las revoluciones en el pasado para predecir nuevos episodios revolucionarios en Egipto y el tiempo que haría falta para estabilizar sus instituciones. Convencidas de la relevancia de tales enfoques, las instituciones invierten ahora mucho en la construcción de grandes bases de datos históricos, antropológicos y sociológicos que deberían permitir nuevas ciencias sociales e históri-

cas digitales. «Es hora de que la historia se convierta en una ciencia analítica, incluso predictiva», escribe Peter Turchin (experto en dinámica de la población) en la revista *Nature* en 2008. O, escribe Armand Marie Leroifeb, biólogo, en *The Opinion Pages* en 2015: «Usted puede estar seguro de una cosa: se expresará no en palabras, sino en ecuaciones».

Es un tema muy controvertido entre los historiadores, que recuerdan la complejidad y la singularidad de los acontecimientos históricos y su desconfianza en las grandes teorías (marxistas, darwinianas, estructuralista, posmoderna) y la necesidad de construir relatos detallados de lo que realmente ocurrió. Estos historiadores no siempre rechazan la tecnología digital, pero eso no les lleva necesariamente a pasar a la cuantificación de los fenómenos. La creación de bases de datos facilita el acceso a un gran número de documentos que de otra forma quedarían dispersos (por ejemplo, las 5.700 versiones griegas del Antiguo Testamento que estudian los exégetas), y también su comparación. Las herramientas de análisis y visualización de redes de personas permiten explorar otros fenómenos históricos. Las tecnologías de la información y de la Web permiten la edición multimedia para que los usuarios vean y sientan los fenómenos y obtengan una comprensión más íntima que lo que lograría sólo por leer un artículo o un libro. Lo digital no se reduce a la idea de cuantificación; es también un desvío útil para enfoques cualitativos.

Por otra parte, los propios algoritmos matemáticos son construcciones que reflejan supuestos, intuiciones, preferencias, pensamientos, tradiciones, influencias culturales, limitaciones económicas y compromisos entre los investigadores. No pueden reducirse a una objetividad que escaparía a cualquier dinámica social. Deben ser analizados, descodificados e interpretados. La digitalización,

cuantificación y matematización no son sólo las operaciones mecánicas para obtener «datos en bruto» (datos que siempre se extraen, fabrican, ponen en forma) que introducir en una fórmula o una máquina que dice la verdad.

En cuanto a la cuantificación, también tiene bases cualitativas. Contar asume que los objetos contados son equivalentes, de tal forma que su adición ( $1 + 1 = 2$ ) tiene sentido; no se suman peras y manzanas salvo si ambas se consideran frutas equivalentes. ¿Pero qué hacer con el grano o las uvas? De hecho, cuantificar siempre implica, según Alain Desrosières **historiador**, de estadísticas, construir primero acuerdos de equivalencia entre objetos, personas, relaciones o eventos. Pero tratar las cosas como equivalentes implica redefinir su estado. Para poder sumar los trabajadores agrícolas y los propietarios de tierras se debe, de antemano, afirmar, por ejemplo, que todos los seres humanos son iguales. Así, podemos descuidar sus diferencias y considerarlos como equivalentes. Este trabajo preliminar al recuento es eminentemente cualitativo y, en este caso, la cuantificación presupone un acto político.

Así que no es de extrañar que la cuantificación haya sido a menudo objeto de controversia. Tiene que ver con la construcción de convenciones de equivalencia, que generalmente entablan negociaciones y compromisos. Viviana Zelizer, profesora de sociología económica, muestra cómo en los procesos de divorcio en los Estados Unidos, el amor entre personas es objeto de negociaciones que determinan su valor, en dólares, y fijan beneficios compensatorios. La traducción de algo en números nunca es evidente; en cualquier caso, se trata de pasar por alto algo del objeto, persona o fenómeno así cuantificado. Se debe recordar lo que se pierde durante el proceso de cuantificación.





## «Las humanidades digitales son cosa de hombres de letras que juegan a ser *geeks*.»

«El historiador del mañana será programador o no será.»

Emmanuel Le Roy Ladurie, *Le Nouvel Observateur*, 1968

Si las humanidades digitales son una cuestión de digitalización y procesamiento de datos, entonces, ¿qué son los *DHer* (humanistas digitales)? ¿Un humanista, un hombre de letras, un informático, un apasionado *geek*, un híbrido?

Los desarrolladores de las humanidades digitales están acostumbrados a trazar el origen de las *digital humanities* en la *literary and linguistic computing*, cuyo pionero sería el jesuita italiano Roberto Busa. Exégeta, es decir, hombre de letras, se vio involucrado en 1946 en un gigantesco proyecto que consistía en la identificación de 11 millones de ocurrencias de palabras en latín en la obra medieval de Santo Tomás de Aquino, pensaba que la máquina **nade** la que hablaba el jefe de IBM podría ayudarle. Explotando oportunamente el lema de IBM «Lo difícil lo hacemos de inmediato; lo imposible requiere un poco más de tiempo», convenció al fundador de IBM para automatizar la identificación de las concordancias de la obra monumental de Santo Thomas Aquino (*Indice Thomisticus*), donde cada palabra del texto se vincularía a su

contexto. Para ello, la obra de Santo Tomás se tradujo en forma de tarjetas perforadas y su equipo concibió un programa para buscar e identificar cada palabra y sus variantes, así como todas las concordancias en el texto que contenía dicha palabra. El nivel de exigencia que se impuso y la incorporación de nuevos equipos informáticos puso a punto esta disciplina emergente ponía la tecnología informática al servicio de las ciencias humanas. En 30 años, el proyecto logró primero una versión impresa de las concordancias (en 1972), luego un CD-ROM y un DVD que incluían hipervínculos (en 1992) y, por último, su puesta en línea. En 1998 se sugirió que la Web también integrara herramientas de análisis sofisticados. Busa habría sido el primero en comprender que las nuevas herramientas, que se utilizaban para hacer estadísticas con fines científicos o en la guerra comercial, también podrían servir a las humanidades.

Desde 1960, otros humanistas llevaron a cabo por su cuenta búsquedas de concordancia en un autor (los poemas de Matthew Arnold, en 1962), una obra, una época (textos alemanes de la Edad Media en 1963) o un idioma (descripción de la lengua común inglesa en todas las variantes de discursos escritos: *Brown Corpus for Use on Digital Computers*). Estos proyectos de estudios sistemáticos normalizaron la construcción de un corpus según unos principios estadísticos explícitos. En el caso del *British National Corpus*, supone 100 millones de palabras tratadas así. Al mismo tiempo, se crearon centros informáticos en las universidades (el centro de procesamiento de documentos electrónicos de Louvain-la-Neuve en 1968 para los textos medievales y clásicos) y en las grandes academias de la lengua para compilar diccionarios como el *Trésor de la Langue Française* y el Instituto Holandés de Lexicon.

La disponibilidad de estos recursos de TI hizo que algunos investigadores las utilizaran para estudiar cuantitativamente el estilo de un autor (estilometría) a mayor velocidad y precisión que los seres humanos, y de paso, confeccionar corpus más grandes. En algunos casos, se trataba de determinar quién era el autor de textos cuyo origen resulta controvertido; el uso frecuente de ciertos elementos lingüísticos específicos de un autor, una época o género, permite atribuir obras sin firma a ciertos autores, detectar «falsos» y reconstruir las filiaciones.

Los humanistas desarrollaron así nuevos métodos para el tratamiento de textos y corpus. Sin embargo, sólo la introducción de un texto en mayúsculas en las tarjetas perforadas resultaba muy trabajosa, y algunas veces generaba toneladas de tarjetas perforadas que debían transportarse hasta los centros informáticos, donde el investigador observaba el resultado sólo después de que el programa informático hubiera completado su tarea. Poco a poco, se han desarrollado nuevos métodos que tienen en cuenta los acentos (importante para los idiomas distintos del inglés) y que convierten (transliteración) los alfabetos no latinos en alfabetos reconocidos por los ordenadores. Las limitaciones de las máquinas (entre otras, las cintas) exigen que se organice la información en serie, lo que sirve para los datos de texto, pero resulta un problema para los investigadores de otras disciplinas (historiadores, en particular), que describen sus materiales en varias dimensiones. En conferencias y revistas de literatura en lingüística computacional se trata el tema de la codificación de los manuscritos, la clasificación automatizada de las concordancias, la consideración de variantes ortográficas, de programación, la lexicografía, la estilística, la edición de textos, la enseñanza de idiomas y el archivo de documentos electrónicos. La re-

vista *Computers and the Humanities*, lanzada en 1966, da a conocer los proyectos de unos y otros.

Al mismo tiempo, entre los historiadores, la formación de series de datos cuantitativos llevó a algunos a acercarse a la informática. Impresionado por el trabajo presentado en una conferencia en los EE. UU. en 1967, Leroy Lardurie profetiza que el historiador del futuro será programador. En 1974, dos historiadores económicos inauguran un estudio sistemático del sistema americano de esclavitud racial mediante el recuento de los hechos, lo que incluye el número de veces que se azotaba a un esclavo, para evaluar la viabilidad económica del sistema. Las críticas a este tipo de estudio, que deja de lado el sentido de las cosas, contribuyeron a que el entusiasmo por la historia cuantitativa decayera rápidamente y la profecía de Leroy Lardurie **no** se cumpliera.

todavía

En cuanto a la ciencia de la literatura, en la década de 1970 se digitalizaron cada vez más textos, y la comunidad de investigadores se consolidó en particular en Gran Bretaña y Estados Unidos. Se empieza a enseñar esta disciplina y surge la pregunta de si se debe enseñar programación en humanidades, lo que conlleva el riesgo de sustituir el rigor del latín por el rigor del Fortran (lenguaje de programación). En 1976 nace una nueva revista: *Literary and Linguistic Computing*. La lingüística domina el campo de la investigación y se constituye la «lingüística del corpus». **p**oco a poco se extiende a nuevas materias: la música, el arte y la arqueología. Dentro de las universidades también empezó a considerarse en los centros informáticos. Debido a que los usuarios rechazaban los programas de análisis de concordancia, los británicos financiaron el desarrollo de un nuevo programa (*Oxford Concordance Program*), que se diseminó por el mundo en los años 1980 y atrajo a nuevos usuarios.

*liberó* La creciente participación de los bibliotecarios y archivistas en la digitalización **libró** a los investigadores de una gran parte del trabajo de la constitución de recursos digitales (bases de datos), así como del archivado, mantenimiento, documentación (metadatos) y acceso a los textos electrónicos. Esto no impidió que algunos grupos de investigadores crearan bases de datos impresionantes, como el *Thesaurus Linguae Graecae* (texto en griego de Homero) en la Universidad de California, Irvine, que invirtió millones de dólares en tratar 70 millones palabras, o corpus de textos del latín clásico o del antiguo inglés. Las nuevas herramientas de diseño de bases de datos evolucionan hacia bases de datos relacionales que se adaptan mejor a los investigadores (incluyendo los historiadores) que trabajan con varias fuentes o más dimensiones. Por otra parte, la proliferación de recursos y análisis abre el camino a herramientas matemáticas más complejas. Este campo de investigación, todavía en actividad, trata el texto como un fenómeno estadístico (textometría), a veces asociado con otras variables tales como las corrientes de pensamiento o la afiliación social de los autores. Los resultados de este trabajo, sin embargo, no circulan mucho más allá de las pocas revistas e investigadores involucrados en la *literary and linguistic computing*. Sus colegas de disciplina (literatura, historia, arqueología, etcétera) no parecen estar tan interesados.

La década de 1980 supone un punto de inflexión para todas las disciplinas de CHS debido a la propagación del ordenador personal, las listas de difusión y el correo electrónico. Se desarrollan en esta época herramientas informáticas genéricas, evitando así que los propios investigadores tengan que programar. Si el investigador finalmente no llega a ser un programador, se debe a la

simplificación de las herramientas que mejoran las interfaces y a la reducción del coste del *hardware* y del *software*. La difusión desde 1980 de los ordenadores personales y del *software* fácil de usar impide que los propios investigadores hayan metido «las manos en el código».

La formación en nuevas tecnologías de la información y la comunicación (NTIC) está cada vez más extendida. El número de usuarios crece firmemente, así como el número de proyectos de TI aplicada a las CHS. Los investigadores son ahora menos dependientes de los servicios de TI e incluso pueden trabajar desde su casa, se organizan para discutir sus experiencias y se estructuran para evitar la duplicación de esfuerzos; los intercambios entre sí aumentan debido a los nuevos medios de comunicación. Los *software* de análisis (estadístico o textual) y la gestión de bases de datos van en aumento. El Apple Macintosh cautivó a causa de su interfaz gráfica de usuario que permite mostrar los caracteres no estándares (en griego, cirílico, antiguo inglés o alemán, etcétera); la capacidad de crear enlaces de hipertexto (HyperCard) entre las fichas (por ejemplo, entre un texto antiguo, su versión moderna y sus anotaciones lingüísticas); y la posibilidad de programarse fácilmente a sí mismo. Las primeras listas de correo permitieron a los encargados de dar soporte informático a los buscadores mantenerse al tanto de las nuevas herramientas y avances de unos y otros, como si se tratara de seminario electrónico permanente. Constituyen una nueva comunidad (*humanities computing*) consolidada por la publicación de un libro de 1000 páginas que revisa el estado del arte de las publicaciones, proyectos y programas: el *Humanities Computing Yearbook*.

El humanista finalmente no se ha vuelto un programador, aunque se trate de literatos apasionados por el código y el tratamiento

cuantitativo, ansiosos por proporcionar el rigor matemático a la interpretación crítica de textos, la historia, las culturas y los fenómenos sociales. **N**Quizá no programen, pero han aprendido a colaborar con informáticos.





## «Las humanidades digitales son un efecto de Internet.»

«[...] podría parecer que todos los humanistas digitales  
están interesados en proyectos alternativos,  
no tradicionales, pero no es el caso»

Patrik Svensson, «The Landscape of Digital Humanities»,  
*Digital Humanities Quarterly*, 4(1), 2010

De hecho, el *humanities computing* aparece, antes que la Web, gracias al establecimiento de un estándar para la codificación de textos: el *Text Encoding Initiative* (TEI). Los numerosos investigadores que codifican los elementos que se deben distinguir en un texto (fecha, nombre del lugar y persona, etcétera) tenían cada uno su propia manera de hacer las cosas. Construyeron una torre de Babel y se obsesionaron con recodificar sus textos de acuerdo al *software* utilizado. En 1986, se repensó el lenguaje de descripción del documento de IBM para dar paso al SGML (*Standard Generalized Markup Language*), que distingue la estructura lógica de un documento de su edición en una página. Al ser válido para todo tipo de textos y sus metadatos, un grupo de investigadores consiguió un millón de dólares para definir los patrones (por ejemplo, la secuencia en un diálogo, la identificación de hablantes) de los textos estudiados por las ciencias humanas (incluyendo rastros de letras o pa-

labras suprimidas que permiten reconstruir la génesis de un texto) y establecer un estándar para su codificación. Publicado en 1994, se convirtió en el estándar para la constitución de nuevos corpus y bases de datos para la edición crítica de textos por parte de bibliotecarios y archivistas. También será el punto de partida de una reflexión que condujo al desarrollo de una teoría de marcas (*markup theory*). En realidad, el *humanities computing* se constituyó como una nueva disciplina, conectada sólo parcialmente con la lingüística computacional (*computational linguistics*) en la medida en que los investigadores necesitan herramientas desarrolladas por ésta (bases de datos y el análisis morfológico, sintáctico y léxico).

El siguiente giro resulta a su vez del desarrollo de Internet (1983), los protocolos de intercambio de archivos de datos y, sobre todo, la Web (puesta en relación de documentos con enlaces de hipertexto HTML - *Hypertext Markup Language*), y de los navegadores para moverse entre las páginas-pantallas vía los enlaces hipermedia (texto, gráficos, imágenes, sonido, vídeo), así como formularios que hacen la Web interactiva (el usuario puede introducir los datos). Se hace más fácil de usar y es ahora una de las herramientas básicas de cualquier investigador, que tiene aún menos necesidad de poner sus manos en el código. A finales de 1990, los jóvenes investigadores encuentran que el término *humanities computing* es poco atractivo y hace pensar en los departamentos de informática que dan soporte a las humanidades. Cuando se prepara un libro de referencia, los editores proponen hablar de humanidades digitalizadas (*digitized humanities*), pero insistiendo en que no es una simple digitalización, sino una codificación de textos que implica una reflexión teórica y plantea interesantes cuestiones científicas, forjan la expresión *digital humanities* (DH) (humanidades digitales) y

popularizan el término con esta publicación en 2004: *A Companion to Digital Humanities*.

Sigue después una primera oleada de las humanidades digitales, que corresponde a la digitalización de los archivos, la codificación de documentos y la comparación de bases de datos con la Web. Los grandes proyectos de digitalización, el establecimiento de la infraestructura tecnológica y la publicación en línea se multiplican; numerosas instituciones se lanzan y abren sus servicios a un público más amplio. Surgen nuevos grupos de usuarios, mientras que los investigadores producen y ponen a disposición nuevos materiales y análisis, sin limitarse por las restricciones de la publicación impresa, o de la escritura. Con los motores de búsqueda, como Google, y el movimiento internacional para la codificación TEI de documentos textuales que se pueden buscar en los textos y cosechar en muchas bases de datos, se anuncia un cambio radical en el modo de estudiar los textos. El nuevo nombre DH permite marcar la diferencia. A esto le sigue una ola de codificación de textos, lo que abre muchas posibilidades para su edición y tratamiento; un verdadero sitio DH debe entonces proporcionar un mayor acceso a las bases de datos digitales; debe ofrecer potentes herramientas de búsqueda personalizables para los investigadores que abordan cuestiones de fondo sobre la forma de representar el conocimiento.

Nacen nuevos tipos de publicaciones, abiertas a la anotación, al comentario, el enriquecimiento y a la actualización por los visitantes en Internet, investigadores profesionales o el público en general. El mundo editorial está cambiando.

Brotan nuevas ideas, algunos proyectos consiguen los recursos necesarios para salir a la luz, pero muchos no van más allá de la «demo» o prototipo, ya que estos recursos se gastan en el duro

trabajo de digitalización y el marcado del texto. No quedan los suficientes recursos para el diseño de una interfaz amigable; las habilidades de los diseñadores se utilizan poco en los proyectos y rara vez se realizan estudios de usabilidad, lo que se traduce en herramientas que fracasan a la hora de atraer a un público más amplio. La audiencia es limitada, los proyectos se detienen debido a su financiación, nadie mantiene los dominios web, los enlaces no funcionan, la investigación queda encapsulada en herramientas que no se difunden. Además, los proyectos se reducen a menudo a una publicación en línea de los materiales y a la navegación por palabras clave; se proponen pocas herramientas de análisis. A finales de 1990, la cosa se vuelve aburrida. A pesar de esto, el número de actores involucrados (especialmente los archivos y bibliotecas) aumentó.

Surgen nuevas prácticas a partir de las herramientas diseñadas con base en esta perspectiva: anotación colaborativa de textos, extracción de partes de documentos para establecer cronologías o historias (historia de las escritoras británicas), la relación de textos con imágenes o películas.

La proliferación de los recursos digitales, su subida a las redes, sus usos y rastros (incluyendo textuales) dejados por los usuarios en Internet generan, de hecho, un nuevo objeto de investigación: la Web y sus comunidades de usuarios. Así, se desarrolla una nueva comunidad de investigación que tiene el ciberespacio como objeto de estudio. Amplía de hecho el campo de la investigación del *humanities computing*, pero conduce a una polarización interna entre los que hacen y los que hablan. Incluso cuando se proponen definiciones inclusivas, muchos observan a sus colegas y su producción científica considerando que los que tienen razón son ellos, es decir, los que meten las manos en el código. Stephen Ramsay, en su char-

la *Who's In and Who's Out* en 2011 dijo: «Creo que las humanidades digitales se ocupan de hacer las cosas. [...] También creo que la disciplina incluye y debe incluir a los que teorizan la realización, los que diseñan para que otros puedan lograr, y los que supervisan la ejecución. Incluso me atrevería a incluir a las personas que trabajan para reconstruir nuestros sistemas, irreparablemente destruidos, de publicación académica. Pero si usted no fabrica nada, no es un humanista digital.» El denominador común de los *DHers* sería el paso de la lectura crítica al hecho de crear y realizar.

Las humanidades digitales tienen menos de 20 años, pero los investigadores ya están hablando de la segunda ola, de las humanidades digitales 2.0, que se centran en la creación de entornos y herramientas de trabajo para la producción, gestión y tratamiento de los datos digitalmente nativos, como micromensajes, los tuits, las consultas en los motores de búsqueda o los metadatos asociados con las fotografías registradas en los sitios y compartidas.

Los historiadores que hacen de la distancia histórica una virtud se han impuesto una «ley de 30 años» antes de abordar los archivos y las huellas de la historia. Pero para los historiadores del siglo xx, 30 años es ya 1980, es decir, ordenadores personales, Internet y la Web. Los archivos sobre los que tendrá que hacer frente ya son bases de datos, *software*, millones de correos electrónicos, mensajes de Twitter, minivídeos y protocolos informáticos, a pesar de que sus objetos de estudio sean la integración europea, los movimientos sociales, el comunismo, etcétera. Se necesitarán nuevos métodos y habilidades para hacer frente a estos nuevos tipos de materiales.

Por otra parte, la distribución multimedia, en donde se mezclan textos, imágenes, sonidos y películas, nos invita a considerar también la experiencia emocional que puede vivir el usuario. Mien-

tras que los enfoques anteriores reproducen las formas de hacer académicas centradas en el texto, la investigación en bases de datos, el procesamiento automático de corpus y el análisis cuantitativo, el sonido y la imagen permiten al usuario experimentar y sentir la complejidad de las situaciones y secuencias, la profundidad analítica, los contextos y los tiempos, la interpretación del sentido y la crítica. Lo cualitativo prima sobre lo cuantitativo, la interacción con las masas de datos sobre su tabulación, la inmersión sobre lo analítico. Las humanidades digitales se desvían del texto sólo para dar más espacio a las visualizaciones, al diseño, a la puesta en escena y la experiencia espacial, la escenificación y la experiencia temporal, al sonido y la imagen. Se acercan a los artistas que manipulan y distorsionan los materiales (en este caso, las masas de datos), lo que les permite conocer mejor y hacer las cosas con sentido.

Bien sea que se trate de textos, registros históricos o rastros en Internet de nuestros contemporáneos, la nueva ola se preocupa por el diseño de tecnologías e interfaces que permiten explorar la riqueza de las interpretaciones posibles, vivir la experiencia iluminadora en lugar de aplicar un razonamiento sistemático y controlado. La informática se vuelve especulativa, lo que permite poner a prueba las interpretaciones, por ejemplo, por estar inmersos en reconstrucciones tridimensionales de las viejas situaciones a descubrir, circulando en ellas como si estuviéramos allí: caminar por las calles de un pueblo del siglo XIV gracias a un *street view* histórico, protegerse de una epidemia al sumergirse en la situación y el estado de los conocimientos disponibles en el siglo XVII, revivir una obra como si estuviéramos en Pompeya, simular la evolución de una controversia actual en un sujeto cargado de cuestiones y poblado por muchos actores con intereses divergentes.

Estas nuevas formas de hacer las ciencias humanas y sociales vuelven obsoletos los viejos métodos de la investigación solitaria, que parte del texto y publica sus resultados en forma de texto, centrada en un caso, un fragmento de archivo, una región o un período muy limitado. Ahora, las tecnologías digitales permiten estudios comparativos a gran escala y la movilización de muchos archivos, así como diferentes disciplinas. También pueden procesar y combinar datos heterogéneos, textos (desde manuscritos antiguos hasta tuits), números (de los registros de tiempo hasta los *logs* en los sitios web), imágenes (desde mapas hasta *selfies*). Se abren a una amplia variedad de prácticas que están lejos de ser estabilizadas, validadas y consensuadas. Las humanidades digitales a veces se asemejan al arte digital o la investigación en museología. Se definen entonces en términos generales como el estudio de las culturas a través de las posibilidades que ofrecen las tecnologías digitales.

Por todo esto, sociólogos, historiadores, antropólogos y humanistas deben aprender a usar el archivo digital y el *big data*, aprender cómo codificar para desarrollar herramientas originales. En THAT-Camp (*The Humanities and Technology Camp*), encuentros de humanistas digitales que ellos llaman «no conferencias», que se definen informales y no jerárquicas, los participantes se autoorganizan para definir los temas propuestos; a menudo hay un intercambio de experiencias en torno a ciertas herramientas específicas. Algunos autores consideran que el estudiante de CHS ahora debe ser capaz de recuperar los datos desde la Web o base de datos, seleccionar y ejecutar algoritmos, elaborar tablas y visualizaciones e interpretarlas. La cuantificación ha triunfado en casi todas las disciplinas de las ciencias naturales; las CHS no deberían escapar. La profecía de Le Roy Ladurie en la programación quizá se haga finalmente realidad.





## «Las humanidades digitales son también una cuestión de las ciencias sociales.»

«En la sociología de Internet nos surge una pregunta desde hace mucho tiempo para la que no se encuentra una respuesta coherente: ¿nos encerró Internet en nuestras creencias o nos abre al otro [...]?»

Antonio Casilli, *Obs* con *Rue89*, 2015, <http://rue89.nouvelobs.com/2015/01/03/antonio-casilli-peut-encore-aimer-internet-256885>

Las humanidades digitales abogan por la cooperación interdisciplinaria entre las humanidades, las ciencias sociales y la informática, así como por la invención de metodologías híbridas. Esto se entiende aún más cuando se observa lo minúscula que es la frontera entre el trabajo de los investigadores en literatura que exploran nuevas formas de escritura (micromensajes tipo SMS y tuits, conversaciones en línea tipo chat, cuadernos y discusiones en línea tipo *blog*, incluso escritos por *bots*, es decir, robots que corrigen, traducen, sintetizan, reúnen textos, etcétera) y los sociólogos que estudian las comunidades en línea y los nuevos tipos de movimientos sociales, como la «primavera árabe».

Estos nuevos fenómenos sociales y lingüísticos implican la invención de nuevos enfoques y métodos para recopilar y analizar la

información, por ejemplo: el *scraping* o raspado, es decir, la extracción del contenido de sitios web o de metadatos a través de una secuencia de comandos del ordenador; el *crawling*, la exploración automática de la Web para recoger recursos como páginas web, imágenes, vídeos, documentos, etcétera, archivarlos y ponerlos en un índice; la recuperación de *logs*, el historial de los acontecimientos tales como la consulta de páginas web, apertura de archivos o comandos por medio del teclado o la pantalla táctil, con el fin de estudiar las prácticas de los usuarios, las interacciones humanos-máquinas o el rendimiento de un sistema informático.

Los investigadores de literatura necesitan aventurarse en terrenos que no sean libros, revistas, correspondencia y archivos; los de ciencias sociales deben imaginar otros enfoques distintos a las entrevistas, cuestionarios y la observación directa del campo.

El uso de la tecnología de la información y de lo cuantitativo en las ciencias sociales, sin embargo, no es nuevo. En sociología, desde la fundación de la disciplina en el siglo XIX, las encuestas se realizan para producir datos cuantitativos de las variables que caracterizan a la población (edad, sexo, estado civil, ocupación, lugar de residencia, etcétera) y el fenómeno estudiado. En su estudio sobre el suicidio, Durkheim también anota la residencia de los suicidios, la época del año, la hora del día del momento del suicidio y la religión. Los datos presentados en forma de tabla en la que cada línea corresponde a una persona y cada columna a una característica de esa persona se prestan fácilmente al **el** procesamiento estadístico. Con la aparición de los primeros procesadores, los sociólogos cuantitativos los emplearon de manera casi natural. En la sociología computacional de la década de 1960, los datos se introducían en tarjetas perforadas; se reformularon los programas en Fortran (len-

guaje de programación) para tratar de transformar las matrices de datos en matrices de variables (agregación de varios indicadores), para probar las hipótesis que asocian las relaciones entre estas variables y las estructuras sociales (estratificación, afiliación, etcétera) o la explicación causal de los fenómenos sociales (la desviación, la movilidad social, el racismo, la elección de una profesión, el apego a las tradiciones, etcétera). Los años 1960 y 1970 corresponden a un aumento significativo de la sociología cuantitativa y de las grandes encuestas por cuestionarios; la informática ya está ahí para muchos.

Simultáneamente, el modelado (determinista y/o estocástico) y la simulación de los fenómenos sociales surgieron como sustitutos para el enfoque experimental, imposible de practicar al menos en macrosociología. Estos métodos vivieron una moda pasajera a principios de 1960. A continuación, se aplicaron en el estudio de la difusión de innovaciones tecnológicas en áreas agrícolas y permitirían demostrar la importancia de las redes de comunicación interpersonal. Se formó así una pequeña comunidad investigadora en torno a la simulación de fenómenos sociales en el ordenador. La sociología computacional se relacionó entonces con el estudio de los sistemas sociales complejos, y las interconexiones entre los micro y macro procesos. Otra corriente, llamada socioinformática, se esforzó por modelar y simular sociedades artificiales.

Las diferentes comunidades de ciencias sociales se estructuran y se especializan en torno a enfoques específicos (sociología cuantitativa, análisis de redes sociales con la revista *Social Networks* fundada en 1979, sociología computacional, socioinformática, ciencia-metría, informetría y webometría, estudio social de las TIC y sus usos, etcétera). Un ejemplo: en los años 1950-1960, en el campo de la sociología de la ciencia, los investigadores establecieron la obser-

vacación de las citas de una publicación a otra (la lista de los artículos citados en la bibliografía en un artículo determinado). Este método permitía luego contar el número de veces que alguien citaba un artículo (indicador de influencia) y analizar las relaciones entre las redes de artículos, dando cabida a una nueva disciplina, fruto de este estudio cuantitativo de las publicaciones científicas: la cientometría. El primer volumen del *Science Citation Index* vio la luz en 1963, donde se propusieron los primeros indicadores de la influencia que más tarde inspirarían la informetría, a los fundadores de Google y la webometría (las relaciones entre sitios web). El estudio cuantitativo de la dinámica científica se consolidó con la creación de la revista *Scientometrics* en 1978, mientras que, en la década de 1980, los sociólogos de la ciencia de la École des Mines de París desarrollaron un método de análisis de palabras clave para la elaboración de mapas de actores y de contenido científico con el fin de estudiar la dinámica de la ciencia y la innovación.

Aparte de los análisis cuantitativos en las ciencias sociales, los investigadores también tratarían de desarrollar herramientas informáticas para el análisis de corpus cualitativos: entrevistas, relatos etnográficos, documentos de todo tipo, incluyendo los producidos o disponibles en la Web (*blogs*, revistas en línea, hipervínculos entre documentos). A continuación, reinventarían sus herramientas tradicionales (transcripciones de entrevistas y conversaciones con varias personas, codificación temática de materiales, análisis automático o textometría, clasificación de contenido, interpretación automática a partir de una referencia teórica, construcción inductiva de las variables, el modelado), aprovechando la informática, a veces con pocas sinergias por parte de lo que estaba sucediendo en las humanidades. Así pues, otra corriente llamada socioinformática creó y uti-

lizó estas herramientas para investigar las disputas y controversias mediante la elaboración de la cartografía de los asuntos públicos, los debates que generan, los repertorios argumentativos movilizados y la reconfiguración de los actores.

En la década de 1970, la sociología tomó la informatización (del trabajo, del hogar) y las nuevas tecnologías de la comunicación (Minitel, Internet, correo electrónico) como un estudio de caso. La revista *Réseaux*, fundada en 1982, se centraba en los usos sociales de la tecnología de medios de comunicación y las TICs. En la década de 2000, los científicos sociales se enfrentaron a la aparición de las redes digitales y la Web, lo que los llevó a fijarse en una nueva categoría de actores: quienes navegan en Internet. Por otra parte, en el estudio social de la ciencia y la tecnología (*science and technology studies*), los investigadores estudiaron el diseño e implementación de la tecnología digital y se organizaron en comunidades subespecializadas: estudio social de los programas informáticos (*software studies*), protocolos (*protocole studies*), infraestructura (*infrastructure studies*) e Internet (*Internet studies*). También estudiaban las formas de expresión en la red digital, la transformación de las relaciones sociales (la confianza, el conflicto y la regulación, la coordinación, la pertenencia, la transmisión entre generaciones, relaciones hombres/mujeres, privado/público, autonomía/interdependencia, libertad/supervisión, etcétera) vinculada a la aparición de diversos dispositivos, como las aplicaciones de todo tipo en teléfonos inteligentes, las relaciones más o menos efímeras entre internautas, las transformaciones del yo (*cyberself*) y de la intimidad, los cambios en las instituciones culturales, incluyendo las guerras mediáticas, o las estructuras sociales (conectividad, desigualdad, individualismo en red, relaciones sociales jerárquicas y de poder, capitalismo digital, etcétera).

La adopción de herramientas informáticas, sin embargo, no despertó a las ciencias sociales ni las llevó a la creación de una comunidad científica genérica ni a la profecía de una revolución científica, como es el caso en las humanidades desde la década de 2000. El concepto de las humanidades digitales, a veces demasiado asociado a las «Letras», no se impone en las ciencias sociales, puesto que éstas estaban ya acostumbradas a usar herramientas informáticas, incluyendo el análisis de redes sociales y semánticas, que han empezado a descubrir ahora las humanidades.

En otras disciplinas, como la antropología, el uso de ordenadores parece marginal. Existen muchas fotografías, notas de campo, dibujos, grabaciones y películas etnográficas, así como millones de utensilios conservados en museos que son objeto de la catalogación y la indexación informáticas, del cruce de bases de metadatos con la Web y herramientas de búsqueda. Las aplicaciones de tratamiento informático de los datos en sí siguen siendo raras y controvertidas. Los investigadores cuestionan el interés por el uso de unos métodos que consideran farragosos y simplistas. Los arqueólogos se enfrentan al mismo tipo de problema, pero hacen uso de la catalogación y del procesamiento informático de sus colecciones de objetos y su descripción (características de los objetos, de su origen, del modo de adquisición, las referencias al objeto en diversos documentos, el contexto que ayuda para comprender el significado del objeto, etcétera).

Los etnólogos, sin embargo, se enfrentan a un problema diferente: la cultura inmaterial (rituales, interacciones sociales, creencias, forma de hablar y técnicas corporales, etcétera) son menos fáciles de catalogar que los objetos asociados a estas culturas. Más importante aún, el discurso sobre el objeto y sus usos a menudo cuenta

más en esta disciplina que el objeto mismo. Codificar, comparar y clasificar los sistemas indígenas de pensamiento *a priori* se presta menos al procesamiento informático que los corpus de literatura o las colecciones de objetos arqueológicos. Cuando se implementa la informática, ésta sirve sobre todo para constituir bases de datos de descripciones de objetos, a gestionar su colección y buscar objetos en dichas colecciones (que implica tener una descripción y un vocabulario estandarizados). Los análisis tratan de comparar las culturas, las prácticas (por ejemplo, la poligamia o la afiliación de la pareja), su evolución, su distribución geográfica y dibujar atlas etnográficos. Este tipo de trabajo ya se hizo con Claude Levi-Strauss mediante fichas de cartón antes de que se pudiera disponer de ordenadores. A veces se analizan corpus de cuentos, canciones, mitos y proverbios. Así fue como vieron la luz los primeros programas de simulación informática de comportamiento colectivo (composición de la familia) para, por ejemplo, calcular la probabilidad de matrimonios endogámicos.

Con la digitalización de nuestras sociedades han aparecido la *sociología digital* y la *antropología digital* con la intención de abordar los nuevos contenidos digitales. Por lo tanto, los antropólogos que estudian las migraciones contemporáneas analizan los miles de vídeos de peregrinaciones que suben a Youtube los migrantes, o participan en juegos en línea para analizar la vida social digital. La antropología digital, como la sociología digital, se basa en dos componentes principales: el uso de herramientas digitales para estudiar los fenómenos sociales y culturales (visualización, análisis de red), y el estudio social y antropológico de la cultura digital (de los diseñadores a los usuarios y viceversa). Más fundamentalmente, las transformaciones del mundo con lo digital son una oportunidad

para comprender con más profundidad lo que somos como especie humana, sacando provecho de la diversidad y la proliferación de las singularidades sociodigitales.

Otras ciencias sociales (la geografía humana, la ciencia política, la ergonomía, las ciencias de la educación, los estudios de la organización, etcétera) también utilizan y estudian estas tecnologías digitales. En la ciencia política, por ejemplo, la sociología electoral sigue de cerca el flujo de los tuits entre votantes, los movimientos sociales, políticos, los candidatos electorales, los encuestadores, etcétera, mientras que otras ramas estudian cómo se gobierna Internet. En geografía, el uso de sistemas de información geográfica (GIS) ha dado lugar a cambios profundos en la disciplina desde la década de 1960. Se utilizan para introducir la información geográfica en formato digital, para archivar datos y gestionar y consultar bases de datos, para procesar los datos y generar nuevas formas de visualización cartográfica o representaciones más abstractas y simular fenómenos (la migración u ocupación de la tierra, por ejemplo). Corren buenos tiempos para las humanidades digitales y cada día forman más parte de las ciencias sociales.



# ¿PARA QUÉ SIRVEN LAS HUMANIDADES DIGITALES?



## «Las humanidades digitales salvarán las humanidades.»

«Una orientación hacia lo digital podría salvar las “humanidades moribundas”» dixit.

Alexander Huang, profesor de inglés en la Universidad George Washington

En los Estados Unidos, a diferencia de Europa, las humanidades están en crisis. En estos países los alumnos de enseñanza secundaria deben elegir una orientación. Desde la década de 1950, en los Estados Unidos el número de alumnos que ha elegido las humanidades se ha reducido en un 50%. En el año 2010, ya no son más de un 8%. A esto se suma una gran reducción de puestos de trabajo en estas disciplinas desde la crisis económica de 2007 y los recortes presupuestarios (entre 2010 y 2013, el presupuesto del *National Endowment for the Humanities* americano se redujo en un 20%, mientras el de otras ciencias se mantuvo) que llevaron a las instituciones a reducir el presupuesto destinado a las humanidades. En 2010, la Universidad de Albany incluso cerró varios programas de formación en francés, italiano, ruso, letras clásicas y teatro, mientras que otras universidades «venden» su formación con el argumento de que preparan al alumno para el mercado laboral en lugar de entrenarle en la reflexión y la ciudadanía. Por otra parte, los sectores crí-

ticos argumentan a favor de la extinción de las humanidades diciendo que son un lujo innecesario, sin interés económico. En tiempos de crisis lo primero sería formar a los jóvenes para el empleo en lugar de molestar a la gente con filosofía analítica, historia medieval o teorías feministas posmodernas. La gente de estas formaciones no inventa nuevas tecnologías; difícilmente podrán generar actividad económica y podría causar dificultades económicas en los Estados Unidos. Se les acusa de formar un club subvencionado de solitarios obsesionados (*nerds*) por sus propios temas de interés intelectual y cultural. El aporte de las humanidades para la sociedad es objeto de controversia; se valoran y descalifican en virtud de argumentos como la importancia de la creatividad frente al conservadurismo, la formación de la mente y del ciudadano, la calidad de la educación, la sed intelectual y la búsqueda de significado, el agotamiento de la vida política y social, etcétera. Las humanidades están en peligro no sólo a causa de la caída del número de estudiantes y los recortes presupuestarios, sino también porque se ataca a nivel político e ideológico su valor para la sociedad y los valores que defienden.

En este contexto, en el que las humanidades están preocupadas por su destino, los académicos presentan lo digital como la manera de salvarlas de una muerte segura. Sus facultades se salvarían de la extinción si dieran un giro hacia la tecnología digital que les concediera el apoyo de la sociedad y les permitiera atraer nuevas subvenciones para la investigación. Podrían atraer a un nuevo público estudiantil hacia las humanidades y sus objetos de investigación (lenguas en peligro de extinción, manuscritos antiguos, objetos mantenidos en archivo, historias del pasado, discusión de las ideas), «espontáneamente atraídos» por las nuevas tecnologías, así como otras

audiencias más allá del medio académico. Lo digital es la solución: una manera de revivir el interés en la investigación en estas áreas.

Lo digital permitiría, en particular, que los objetos, datos y conocimientos tratados en las humanidades sean más atractivos y accesibles y, abre el mundo de la investigación a un público más amplio. Tanto para los estudiantes y como para el público en general, ayudaría a concebir maneras más estimulantes de explorar el patrimonio cultural. Por lo tanto, con la invención de máquinas informáticas capaces de retroceder en el tiempo (por ejemplo, un *software* interactivo de visualización del pasado a partir de datos de los archivos), las humanidades permitirían a sus usuarios descubrir ciudades y civilizaciones antiguas, reconstruidas por medio de la realidad virtual. La colaboración internacional entre arqueólogos, arquitectos e informáticos ha dado lugar a la simulación de 7.000 edificios de la época en que Roma vivía en pleno apogeo. En el proyecto *Venice Time Machine*, el tratamiento de 80km de archivos de manuscritos, libros de cuentas y mapas antiguos permiten al usuario caminar por Venecia, como lo haría con Google Street View si se le hubiera añadido un cursor cronológico que permitiría pasar fácilmente tanto de una calle a otra como de un siglo a otro, desde el siglo XVIII hasta el siglo XIII, escrutar los detalles de una fachada o abrir el plano para ver la topografía de la ciudad, o acceder con un clic a la información sobre un barco que se dispone a entrar (en esa época) en el puerto, conocer su carga, su origen y la ruta marítima que ha seguido.

Otros investigadores están desarrollando herramientas para la enseñanza de la historia basada en los juegos de rol. Con *Desperate Fishwives*, dan a conocer a los jugadores-estudiantes las prácticas sociales y culturales del siglo XVII sumergiéndoles virtualmente en

un pueblo británico de la época. Frente a los problemas de esta sociedad, como el alcoholismo de un vecino, las acusaciones de brujería o el embarazo antes del matrimonio, el jugador adquiere una sutil comprensión del orden y el desorden en la Inglaterra moderna, donde no existe la policía. El aprendizaje de cómo mantener el tejido social se da sin la sensación de estudiar la historia. En el aula o a través de un blog, los estudiantes deben reportar a continuación la complejidad de la historiografía que descubrieron a través de esta experiencia.

En otras disciplinas, las máquinas virtuales de antropología o sociología podrían simular la experiencia de las comunidades, culturas, grupos sociales o de la dinámica social. En Letras, los algoritmos de búsqueda en los textos (*text mining*) permitirían a los usuarios dar sentido a un texto y luego representarlo de una manera atractiva, de forma visual, en lugar de una narración argumentada. En cuanto a la historia, con *Orbis*, los usuarios pueden explorar el espacio y los movimientos dentro del Imperio romano a través de herramientas digitales que procesan e incorporan muchos datos dispersos y heterogéneos. El portal, desarrollado por un equipo de investigadores de la Universidad de Stanford, ofrece un servicio de preparación al viaje (como la función «itinerario» de Google Maps), pero para el Imperio romano. La aplicación dibuja varias rutas (por ejemplo, entre Roma y Lyon) sobre la base de la configuración elegida por el usuario a medida que viaja en invierno o en verano: si es un caballero portador de un mensaje urgente o alguien que debe moverse con cien legionarios a pie; si puede tomar prestado un barco haciendo cabotaje o embarcarse en alta mar. La herramienta también calcula el tiempo de viaje y el consumo de alimentos para el viajero y su montura.

A pesar del interés que estas tecnologías generan, no es seguro que vayan a salvar a las humanidades. Algunas voces que enfatizan el alto coste de este tipo de proyectos, y pocas universidades encontrarán los recursos necesarios para desarrollar este tipo de aplicaciones espectaculares. El proyecto Orbis, por ejemplo, supuso el trabajo de un historiador experimentado, 12 estudiantes, cuatro técnicos para el desarrollo de un Sistema de Información Geográfica (GIS), interfaces y sitios web, el apoyo de la biblioteca de Stanford y la contribución voluntaria de seis investigadores externos a la Universidad de Stanford. El coste total del proyecto, que no ha recibido subsidios específicos, es de alrededor de 150.000 dólares.

La mayoría de las instituciones probablemente utilizarán las tecnologías digitales a pequeña escala, involucrando a estudiantes de licenciatura, sin recibir fondos públicos importantes. El objetivo principal seguirá siendo el procesamiento de documentos sin dejar de cuestionar los métodos utilizados. Por otra parte, insistir en las tecnologías digitales (*data mining*, *text mining*) y decir que las humanidades son una disciplina tan relevante como las otras, lejos de salvarlas, sólo ayuda a devaluarlas aún más, pues atribuye todo el crédito del rigor conceptual y metodológico a la informática y las tecnologías digitales.





## «Las humanidades digitales democratizarán el saber, la cultura y el acceso al patrimonio cultural.»

«Las humanidades digitales poseen una base utópica debido a una genealogía asociada a las contraculturas y ciberculturas de los años 1960 y 1970. Por esta razón promueven valores como el acceso abierto, la expansión sin fin, la universidad/el museo/el archivo/la biblioteca sin paredes, la democratización de la cultura y el conocimiento.»

*The Digital Humanities Manifesto, 2009*

La desmaterialización ligada a la digitalización sugiere que los contenidos de los libros, el conocimiento y las representaciones de los objetos culturales ahora pueden fluir sin restricciones materiales y ser fácilmente accesibles en todo momento y en todo lugar con una simple conexión a Internet. Esta posibilidad ahora posible, retoma viejas utopías: la de Paul Otlet en 1930; las contraculturas de los hippies de los años 1960 que proponían abolir las fronteras entre el arte, la ciencia, la cultura y la vida social; aquellas de las comunidades informáticas académicas que tuvieron lugar en los años 1970-1980 (arpanautas), quienes conectaron en red los ordenadores y promovieron los intercambios igualitarios, lo que daría pie a Internet

y la Web; las de los hackers que reivindican el acceso total y sin restricciones a los ordenadores, una información libre y el rechazo de la informática centralizada, con la intención de crear una nueva sociedad armoniosa y una informática comunitaria; aquellas de los cyberpunks para quienes la informática es sinónimo de liberación y vida social, etcétera. Por lo tanto, no es sorprendente que el movimiento de las humanidades digitales esté relacionado con la del acceso libre y abierto (*open access*) y se exprese en manifiestos que declaran:

«Nosotros, actores de las humanidades digitales, nos constituimos en una comunidad de práctica solidaria, abierta, acogedora y de libre acceso.

Somos una comunidad sin fronteras. Somos una comunidad multilingüe y multidisciplinaria.

Nuestros objetivos consisten en profundizar en el conocimiento, mejorar aún más la calidad de la investigación en nuestras disciplinas, y enriquecer el saber y el patrimonio colectivo más allá de la esfera académica.

Abogamos por la integración de la cultura digital en la definición de la cultura general del siglo XXI.

Reivindicamos el libre acceso a los datos y metadatos. Éstos tienen que estar documentados y ser interoperables, tanto desde el punto de vista técnico como del conceptual.

Estamos a favor de la difusión, de la circulación y del libre enriquecimiento de los métodos, del código, de los formatos y de los resultados de la investigación.» (*Manifiesto por unas Humanidades Digitales*, 2010)

Con las humanidades digitales, la cultura y el conocimiento son finalmente accesibles a todos. De hecho, con los ordenadores,

tabletas, lectores electrónicos o teléfonos inteligentes, desde los estudiantes hasta los jubilados, todo el mundo puede tener acceso a enormes bibliotecas, discotecas, filmotecas, archivos y museos, y viajar virtualmente por el mundo. La cultura ya no está encerrada entre los libros y las paredes de las a veces remotas bibliotecas. Sin salir de la clase, los estudiantes pueden pasear por los estantes de la biblioteca pública o en las páginas de Wikipedia, mientras que el profesor no está obligado a hacer fotocopias. Sin cruzar el país o hacer cola en el mostrador, el retirado aficionado a la historia y genealogía puede tener acceso a archivos remotos. El paciente puede circular por las páginas oficiales de las Academias de Medicina, de las enciclopedias médicas, de los blogs de otros pacientes o asociaciones, o bien ver las publicaciones científicas.

Esto es resultado de la voluntad política (por ejemplo, de introducir lo digital en el aula o de imponer el acceso abierto a los resultados de la investigación científica pública), así como de enormes inversiones de instituciones y empresas culturales para digitalizar sus colecciones. En Francia, en 2015, un proyecto de Ley sobre lo digital condujo al debate sobre el tema: neutralidad de la red, acceso abierto a las publicaciones científicas y a los datos de la investigación, derecho sobre los datos científicos para crear servicios con valor añadido. El tema es objeto de tensión especialmente entre los editores y las comunidades científicas con respecto al derecho exclusivo de los datos, ilustraciones y escritos de la investigación pública; al libre acceso digital a los datos y a los resultados sin que haya periodo de embargo; y a la libre comercialización del desarrollo de servicios de registro de datos, lo que permitiría su revalorización. El debate se refiere, en particular, a la viabilidad económica de dicho acceso libre, sin la cual revistas, colecciones de libros y editores

podrían desaparecer, lo que puede afectar especialmente a las ciencias humanas y sociales y a la publicación en otros idiomas distintos al inglés, salvo si se replantea la financiación pública del trabajo editorial. A pesar de que la técnica posibilita el libre acceso a la información, ésta se ve limitada por consideraciones económicas (su sostenibilidad y el coste de su publicación y su difusión) y de propiedad intelectual.

Así pues, lo digital no es sinónimo de libre acceso para todos. Este acceso libre se debe construir, negociar y financiar. E incluso cuando se paga, el acceso universal no está garantizado; los editores limitan la distribución de las versiones digitales de sus obras en algunas partes del mundo (especialmente en América del Norte y Europa) y la lectura del soporte informático a ciertas zonas geográficas (como ocurre con los DVD) o a un determinado *hardware* (tableta, lector electrónico, teléfono inteligente, etcétera) o uso (copia, anotación, extracción de citas y apropiación personal prohibida; limitación a la lectura, *read only* o *streaming*, o el acceso gratuito a la versión HTML, mientras que la versión en PDF se paga; el bloqueo de la lectura rápida mediante anuncios). Así, los medios digitales están equipados con medidas técnicas de protección y gestión de derechos digitales (DRM), que imponen límites en términos de territorio, tiempo y uso, basados en estrategias de negocio y en el derecho de los autores. Los editores digitales tatúan sus obras para facilitar el seguimiento de las copias no autorizadas, y en última instancia, el acceso también depende de la disponibilidad de Internet, los formatos y equipos.

Se espera que las humanidades digitales trabajen en la democratización del conocimiento y en un mayor control democrático de la ciencia por parte de la sociedad. Lo digital permite a los inves-

tigadores comunicar y difundir su producción científica de forma distinta, por ejemplo, haciendo caso omiso de las revistas académicas y conservadoras poco abiertas a las producciones innovadoras, o mediante la liberación de los formatos convencionales para comunicarse de manera interactiva (cuaderno de investigación en línea, blog, nanopublicaciones, etcétera). Al abrir la posibilidad de publicar fuera de las estructuras tradicionales de edición (con comités de lectura, editores privados, etcétera) y de dirigirse directamente a sus audiencias, estas tecnologías democratizan el acceso a las ciencias humanas y sociales: no sólo el acceso a los resultados de la investigación, sino también el acceso a los datos primarios utilizados o producidos por los investigadores. El público podría entonces controlar por sí mismo la validez de los resultados y discutir la elección de los datos y su procesamiento e interpretación. El público podrá debatir los análisis y proponer otros, conduciendo su propia investigación a partir de los datos que son ahora accesibles. La evaluación del trabajo de los investigadores ya no dependería sólo del escepticismo organizado por pares, sino también de la revisión y discusión por parte del público. Por otra parte, puesto que los datos y las fuentes son técnicamente más fáciles de circular, aproximar, comparar, agregar y comentar —también gracias a servicios como Google Custom Search, Rollyo o Yahoo Pipes, que procesan los datos sin tener que construir una base de datos o escribir códigos informáticos—, el proceso de producción científica se abrirá a diversos grupos sociales que podrían desafiar la prepotencia de los investigadores y cuestionar la distinción entre profesionales y aficionados. Las ciencias humanas y sociales serían más democráticas, participativas, e incluso divertidas. Por el contrario, las tecnologías digitales permiten a los investigadores movilizarse (*crowdsourcing*: la movi-

lización de las masas) y apoyarse en una multitud de aficionados a los que asignar en línea tareas específicas, como la transliteración de manuscritos antiguos; descifrar escritos antiguos (en griego o español medieval, jeroglíficos mayas, etcétera); la identificación de los lugares y las personas de más de miles de fotos antiguas; la indexación de colecciones de películas o archivos de la Iglesia y del registro civil; la investigación y recopilación de documentos en archivos familiares; la descripción de las ilustraciones de partituras musicales; la clasificación basada en las categorías que vienen de los aficionados (folcsonomía).

Lo digital abre muchas oportunidades para evolucionar hacia la democratización del conocimiento, pero las dinámicas sociales, económicas y tecnológicas pueden conducir a la creación de nuevas barreras. Hace falta una vigilancia social sobre los desarrollos en curso.

## «Las humanidades digitales harán posible el diálogo entre los pueblos y las culturas.»

«El libre flujo de información, posible gracias a la llegada de las nuevas tecnologías como Internet, conlleva la promesa de un diálogo intercultural.»

UNESCO, [http://portal.unesco.org/fr/ev.php-URL\\_ID=34309&URL\\_DO=DO\\_TOPIC&URL\\_SECTION=201.html](http://portal.unesco.org/fr/ev.php-URL_ID=34309&URL_DO=DO_TOPIC&URL_SECTION=201.html)

Las humanidades contribuyen a la vida y al progreso intelectual de nuestras sociedades, incluido el diálogo entre civilizaciones, ayudando a comprender mejor el patrimonio cultural de cada uno. Las tecnologías digitales estimulan su imaginación en este sentido con la creación de herramientas de colaboración e intercambio. De los antiguos mayas a las religiones contemporáneas, de la música de Bach a las músicas del mundo, de los primeros embriones de sociedades humanas a las nuevas culturas digitales, las tecnologías ofrecen la oportunidad de reconsiderar el avance hacia una mejor comprensión de nuestros respectivos pasados y un mayor respeto mutuo.

Algunos filólogos de latín y griego, por ejemplo, pusieron en marcha un proyecto de «filología abierta» para revivir las fuentes lingüísticas de las culturas grecorromanas que han penetrado en Oc-

cidente y el mundo árabe. Jugaron un papel importante en la vida intelectual durante más de 2000 años, pero sufrieron un duro golpe desde que el inglés, el negocio y la alta tecnología se convirtieron en el espejismo de nuestras sociedades. Este espejismo se disipa cuando uno se descubre como actor involuntario en una guerra de civilizaciones, ignorantes de las culturas con las que compartimos una amplia base común. Los filólogos ven en lo digital una oportunidad para apoyar el diálogo entre civilizaciones, convencidos de que las lenguas muertas todavía pueden desempeñar un papel importante para la humanidad. Con lo digital y el acceso a múltiples y variadas fuentes podríamos desarrollar una mejor comprensión de nuestras raíces culturales, por mucho tiempo reducidas a un corpus limitado de textos canónicos que han acaparado la atención de los expertos y educadores, y encontrar la riqueza de civilizaciones pasadas y las hibridaciones que han experimentado.

De hecho, los investigadores en esta área, en peligro de extinción por falta de un público estudiantil suficiente, se ven obligados a innovar. En lugar de publicar sólo monografías impresionantes, adquiridas por cientos de bibliotecas y leídas por un puñado de eruditos, los filólogos digitales tratan de inventar formas de llegar a los mil millones de internautas de todas las civilizaciones. Se esfuerzan por proporcionar acceso a las fuentes textuales repartidas por todo Occidente con su digitalización y su subida a las redes, además de inventar herramientas informáticas para buscar de forma remota en todos estos textos, incluso manuscritos, en pergamino o piedras talladas. Se podría así descubrir más de un millón de palabras en griego o latín, sus usos, su circulación y sus evoluciones. Un ejemplo: la publicación bilingüe en inglés de clásicos griegos y latinos (la colección Loeb) incluye 500 volúmenes impresos, 25 millones de



palabras, un logro monumental de los filólogos del siglo xx. Sin embargo, en una encuesta realizada en el primer millón de libros digitalizados por Gallica, Google Books e Internet Archives se muestra que los usuarios de Internet ya tienen acceso a los mil millones de palabras latinas y griegas y a su contexto.

Más allá del acceso a los textos y sus traducciones en nuestras lenguas modernas, incluido el árabe clásico, entender estas civilizaciones también significa inventar formas de interactuar con la herencia que nos han dejado y renovar nuestras preguntas contemporáneas: herramientas de alineación de textos para compararlos, analizadores morfológicos de escritura antigua, herramientas de clasificación de términos de acuerdo a su forma o función, herramientas de visualización de red de palabras a través de masas de texto (incluyendo los textos raros, anteriormente censurados o considerados como secundarios), los países y épocas. Este trabajo es titánico: no hay suficientes académicos y bibliotecarios. Entonces se organizan redes de investigación extensas, como la propuesta de «filología abierta» de la Universidad de Leipzig, con la participación de estudiantes y ciudadanos aficionados, incluidos los jubilados y toda persona a la que le entusiasme desenterrar estos tesoros de la civilización. Se trata de un proyecto de formación digital para las lenguas antiguas que trata de superar las barreras lingüísticas, culturales e históricas.

Otros proyectos contribuyen a esta ambición de construir un diálogo global entre civilizaciones pasadas y presentes con la creación de recursos digitales y la invención de herramientas de traducción y de búsqueda en los textos de diferentes tradiciones, así como con interfaces más fáciles de usar para interactuar con estas herencias culturales.

Las políticas internacionales invitan a hacerlo, sobre todo en un contexto de grandes migraciones y mezcla de poblaciones. La UNESCO ha hecho del diálogo intercultural una prioridad con la adopción de la Declaración Universal sobre la Diversidad Cultural en el año 2001, que llama a un diálogo justo entre los pueblos, culturas y civilizaciones, fundado en la comprensión mutua y el respeto por la igual dignidad de todas las culturas, requisito previo para la construcción de la cohesión social y la paz entre las naciones. La Unión Europea, en 2007, estableció una agenda para la «Cultura en un mundo globalizado». El diálogo intercultural evitaría las tensiones culturales, lingüísticas y religiosas para lograr lo que la modernidad científica y legal no ha podido construir homogeneizando (universalismo) a los humanos. Las humanidades, con su conocimiento de estas culturas, y con las tecnologías de la Web 2.0, contribuyen a la creación de este nuevo mundo, abierto, híbrido, cosmopolita, multicultural, donde vivir juntos, conectados (por red) y respetando nuestras diferencias.

Las convocatorias de proyectos invitan a los investigadores a ser innovadores, ofrecen herramientas como el espacio de Internet colaborativo y globalizado *diverCities* donde establecer un diálogo entre residentes de las grandes ciudades de todo el mundo, empezando por Sidney, Bombay y Singapur. A estas tres ciudades las une haber formado parte del Imperio británico, a pesar de que su población no haya tenido la oportunidad de conocerse. Suponiendo que sus habitantes viven prácticas urbanas similares o diferentes, el sitio intentaba facilitar el conocimiento mutuo y construir la solidaridad global a través de las distancias. Para comenzar a compartir las historias y el diálogo, los investigadores actuaron como los usuarios que generan su propio contenido en las comunidades que


pueblan sus respectivas ciudades y eventos clave, historias surgidas a partir de esta diversidad, etcétera. Para evitar que los usuarios no se limitaran a su propia ciudad y favorecieran sus enlaces y el diálogo entre las historias de los demás, el sitio se estructuró en torno a temas transversales (trabajar, comer, vivir, moverse, pertenecer, etcétera) y a las variaciones culturales. También proporcionó la expresión multimedia que permitía la presentación de fotos, vídeos, canciones, copias de periódicos, etcétera.

Sin embargo, el proyecto nunca llegó a buen término; las tecnologías digitales no son una solución mágica. La reflexión sobre el fracaso ayuda a explicar las dificultades del diálogo intercultural digital. La primera es una paradoja: aunque la intención es promover una expresión abierta y democrática, y saltarse el control de los expertos, por razones de seguridad y para evitar la proliferación de expresiones que exacerbarán las tensiones entre las comunidades se llevó a cabo un control de la integridad intelectual. Por otra parte, aunque el sitio promueve la comunicación multimedia, los investigadores han recurrido principalmente a su modo habitual de expresión, es decir, la escritura densa, precisa y rigurosa. Por último, aunque la idea era que todo el mundo, desde su casa, contribuyera al sitio, los propios investigadores lo alimentaron realmente sólo cuando pasaron varios días juntos físicamente; después de este punto álgido, de regreso a casa su compromiso decayó, lo que destaca la debilidad del vínculo social distante, incluso cuando lo media la tecnología. Este fracaso, como muchos otros, plantea la cuestión de si las humanidades digitales logran realmente contribuir al diálogo entre civilizaciones.

Una solución puede estar en contratar a diseñadores y utilizar la ergonomía en el diseño de las herramientas digitales. Esto per-

mite tener en cuenta las motivaciones y las prácticas de los usuarios, ya que muchas herramientas y recursos digitales sólo reflejan formas de ver y actuar de los investigadores, despertando poco interés entre los usuarios de Internet. Una segunda solución sería la de acercarse a los artistas para crear experiencias que tengan sentido (*performances* artísticas), que puedan tocar y poner en marcha los internautas, ofreciendo instalaciones en línea, acciones colectivas en Twitter, algoritmos poéticos, juegos masivos de colaboración como *PlaceStorming v 3.0*, movilizandando así a una comunidad de usuarios involucrados en la definición y resolución de problemas en apoyo de la comunidad académica.

Aún así, el diálogo entre civilizaciones no es automático. Las tecnologías digitales deberían mejorar la capacidad pública para examinar y criticar las producciones científicas, comprobar la selección de datos y análisis, sugerir otras interpretaciones y difundirlas en la Web. Como las fuentes digitales fácilmente se copian, anotan, reimprimen y distribuyen, cualquier individuo o grupo puede tomar posesión de ellas, como se ha observado de los exégetas. Acostumbrados a trabajar sobre la totalidad o parte de las 5.700 versiones griegas del Nuevo Testamento, identificadas por una institución académica de referencia, encuentran hoy en día en los sitios web de los movimientos religiosos extremistas versiones de las que no eran conscientes, con comentarios en árabe, y cuya interpretación refuerza los argumentos en contra de otras comunidades religiosas amenazadas por su existencia. Las humanidades digitales pueden contribuir involuntariamente a dividir a las personas mientras intentan unirlas. Una vez abierto este campo de posibilidades, será necesaria una vigilancia colectiva para que se produzcan los beneficios esperados.



«Con las humanidades digitales,  
las herencias culturales se volverán  
nuevos recursos para la innovación  
y el desarrollo económico.»

«Cuando las palabras valen oro.»

F. Kaplan, *Le Monde diplomatique*, nov. 2011

Las humanidades digitales no se limitan al hecho de utilizar las tecnologías digitales para hacer ciencias sociales y humanas; pretenden también inventar nuevos métodos o herramientas.

En el proyecto *Venice Time Machine*, cuando se trata de la digitalización de 80km de archivos (manuscritos y libros impresos, libros de contabilidad, documentos notariales, mapas, etcétera), los investigadores se han fijado el reto de inventar robots que podrían digitalizar 1.000 páginas por hora, y máquinas para escanear los libros sin necesidad de abrirlos; para ello utilizan técnicas de imagen médica. Cuando se trata de avanzar en la transcripción automática de los manuscritos, diseñan *softwares* que sirven para el reconocimiento de las formas de las palabras y para la detección automática de las líneas de texto (para diferenciarlas, por ejemplo, de lo que aparece en la transparencia de la escritura en la parte posterior de la página). En el proyecto *Visual Page* se desarrollaron herramien-

tas de reconocimiento automático de los aspectos gráficos de los textos publicados (diseño de página, márgenes, sangrado, espaciado, etcétera) y no sólo del contenido lingüístico de los textos, importantes sobre todo para la poesía y el análisis de sus cambios en el tiempo en grandes conjuntos de libros. En el proyecto *sx:archipelagos* se desarrollan nuevos *workflows* (flujos operativos de trabajo y *software* asociados) para la publicación y la conservación a largo plazo de las producciones digitales de la investigación en humanidades en el Caribe. En el caso del proyecto AMPACT (*Automatic Music Performance Analysis and Comparison Toolkit*), el objetivo es desarrollar un conjunto de herramientas para el análisis de los cantos y la música polifónica. Con el diPH (*digital public humanities toolkit*) se proporciona un conjunto de herramientas personalizable, de libre acceso, para todos los que quieren editar y administrar sitios web, proporcionándoles el acceso y la visualización a un gran número de materiales de patrimonio cultural (mapas, fotos, copias de periódicos, historias orales, manuscritos), lo que incluye *tours* y exposiciones virtuales, enseñanza interactiva, y la posibilidad de añadir otros documentos y testimonios por parte de los internautas. El proyecto *Scriptorium* desarrolla el libre acceso a tecnologías digitales para la investigación y la anotación de un corpus de los jeroglíficos de la época faraónica, publicados en varios formatos digitales codificados con etiquetas de búsqueda, que se pueden explorar a través de herramientas de búsqueda y de visualización.

Las herencias culturales plantean un reto a los investigadores: inventar nuevas herramientas. *A priori*, son recursos al servicio de la innovación y, potencialmente, del desarrollo económico. Los investigadores informáticos y los ingenieros no se equivocan; en la actualidad muchos se acercan a las ciencias humanas y sociales porque

entienden el potencial de estos datos y este tipo de información que desafía las herramientas existentes.

La Escuela Politécnica Federal de Lausane, por ejemplo, creó un laboratorio y un Instituto de Humanidades Digitales; recluta, junto con sus ingenieros, filólogos, musicólogos, profesionales de museos y sociólogos que trabajan en el *Big data* cultural. La *Venice Time Machine* es uno de sus proyectos estrella. La otra es la digitalización de los archivos del Festival de Jazz de Montreux: 50 años de grabación de audio y vídeos de conciertos, sin contar las decenas de miles de fotos. Alrededor de este corpus que ya ocupa casi tres petabytes de datos en servidores seguros, se han puesto en marcha más de 25 proyectos de investigación y desarrollo tecnológico con el objetivo de producir conocimiento tecnológico, inventar nuevas tecnologías, producir patentes y crear nuevas empresas. Más de 30 profesores del Politécnico y varios estudiantes de doctorado de diez laboratorios (en física del sonido, informática, procesamiento de señales, etcétera) gravitan alrededor del archivo digital. La digitalización de 10.000 horas de grabación realizada en 18 formatos y medios de comunicación diferentes, el control de calidad (invención de herramientas de detección y corrección automática de varios tipos de defectos) y el establecimiento de una base de datos del contenido de este archivo musical (canciones, músicos, instrumentos, formatos de grabación de originales, derechos de difusión y reproducción, etcétera) ya han generado una gran cantidad de creatividad técnica. La búsqueda automática de patrones (por ejemplo, la identificación de los solos de batería, saxofón, guitarra o canto o los aplausos) o la identificación automática de los principios y finales de cada obra musical para su indexación son otros campos de investigación tecnológica. Para los ingenieros, este corpus gigante

les permite desarrollar y comprobar los nuevos algoritmos que inventan, por ejemplo, para recomendar canciones basadas en características del sonido que prefiere el oyente. También permite a los ingenieros de sonido, por ejemplo, inventar dispositivos de restitución controlada del sonido (por ejemplo, para proponer una música diferente en cada mesa de un restaurante sin la interferencia de una mesa a la otra, o la restitución de un concierto con la acústica reconstruida de un auditorio que ha desaparecido). La constitución y el acceso privilegiado a tales corpus culturales son recursos estratégicos para la escuela de ingeniería.

Así pues no es sorprendente observar que surjan competencias entre escuelas, empresas y países para constituir o hacerse con dichos corpus. Los corpus sociales y culturales ya proporcionan la base para un desarrollo económico significativo. Empresas como Google, Amazon, Twitter, Facebook, Flickr, etcétera, viven de eso. No es caridad que Google Books digitalizara gratuitamente millones de libros de las bibliotecas públicas u ofrezca servicios gratuitos como Google Ancient Places. La creación de archivos digitales y la recuperación de datos digitales (las palabras que usamos al buscar en Google, los textos que traducimos, nuestros mensajes de correo electrónico a través de Gmail, etcétera) son los recursos culturales que nutren el desarrollo de nuevos algoritmos y el diseño de servicios generadores de ingresos. Cuando se nos invita a leer una palabra distorsionada (CAPTCHA) para demostrar que somos un ser humano y no un robot, sin saberlo, participamos de forma involuntaria a la mejora de los algoritmos de reconocimiento de caracteres. Con estos tesoros idiomáticos (libros escaneados, escritos en Internet), Google ha demostrado que podemos emprender e incluso crear nuevos mercados lucrativos. Con la organización de mi-



llones de subastas de palabras que venden a los anunciantes con la esperanza de atraer a los usuarios de Internet a sus páginas, Google gana más de 75.000 US\$ por minuto. Las subastas de palabras se recalculan cada vez que se hace clic (millones de veces por segundo) para evaluar el estado del mercado lingüístico global y multilingüe, y tener una idea precisa de los movimientos semánticos globales (estacionalidad, tendencias, etcétera). También existe la competencia entre algoritmos para sacar partido de estos valiosos recursos culturales, mientras que la proliferación de los robots informáticos (encargados de la traducción, síntesis, corrección automática de las páginas wiki, la escritura automática de correo no deseado o tuits, etcétera) hace que los algoritmos, cada uno con su propio dialecto, se mezclen en nuestras conversaciones y creen nuevos lenguajes como el Googlish (inglés estilo Google Translation). Los investigadores de literatura no sólo deben desarrollar herramientas de reconocimiento de autores humanos (gracias a su propio estilo), sino también no humanos, y estudiar las transformaciones de la lengua viva.

Seguramente las humanidades digitales son recursos para la innovación y el desarrollo económico. Sin embargo no se limitan a ponerse al servicio de estas dinámicas económicas. Por el contrario, los diversos manifiestos producidos por los investigadores en estos campos abogan por el acceso libre y abierto a los datos y algoritmos. Lo hacen en nombre de la humanidad, para que la supervivencia de los recursos digitales no dependa de la supervivencia de una economía basada en el *software* privado, lo que realmente es complejo, ya que la dinámica económica también contribuye a la democratización de la herencia y la innovación. La solución para mantener estas dos dinámicas consiste en diseñar estándares abiertos y ampliamente compartidos.

Por último hay que recordar que el movimiento de las humanidades digitales también lo impulsa la competencia entre las instituciones y entre los países en cuanto a su prestigio y poder de influencia. Los proyectos, a veces muy caros, son medios para posicionarse en un ámbito académico y cultural en el que se da la rivalidad entre disciplinas, estándares (TEI, código Dublín, etcétera), instituciones culturales (museos, medios de comunicación, etcétera), y tradiciones nacionales, que se comparan a partir del número de visitantes, citas, traducciones o factores de impacto.

¿TEMOR FUNDADO?



## «Con lo digital vamos a perder mucho.»

«El mejor sistema de conservación que se ha inventado  
es el libro premoderno obsoleto.»

Robert Darnton, director de las bibliotecas  
de la Universidad de Harvard, *The New York  
Review of Books*, 2008

Se considerará que la digitalización tiene *a priori* numerosas ventajas: ahorro de espacio (100 libros: 30 kg; un teléfono celular: 200 g; 10.000 en un ordenador o en la biblioteca de un servidor) gracias al hecho de archivar versiones digitales en vez documentos originales; la reunión en el mismo servidor de diferentes herencias culturales (texto, fotos, música, películas); el rápido acceso a una amplia gama de objetos (en 2015, Deezer ofrece 35 millones de canciones en su catálogo, algo imposible de guardar en una discoteca) y la valorización de los fondos; el coste reducido (de las copias, los medios de comunicación, el almacén y la distribución); la fácil generación de información agregada (catálogos, estadísticas); un efecto de halo (escanear una colección aumenta el valor de cada uno de los objetos que contiene); se puede duplicar de forma ilimitada (el *stock* nunca se agota) y no se degrada. A esto se añade la idea de preservar y proteger los documentos contra los riesgos de alteración, en

particular contra los daños relacionados con el uso y suministro al público de documentos raros o antiguos. Se impone la idea de que, con lo digital, finalmente seremos capaces de conservar todo, un archivo que sea, por ejemplo, de 350 millones de nuevas fotos subidas por los usuarios a diario en Facebook; ya en 2013, tienen mil millones de fotos que mantener (400 petabytes).

Sin embargo, muchas personas ponen en duda la fiabilidad de los sistemas digitales para la conservación de todo este material a largo plazo, debido a los costos de su conservación y al riesgo de pérdida relacionada con esta digitalización. ¿Qué parte de la información que producimos se preservará para las generaciones futuras?

Los medios de almacenamiento han demostrado ser más frágiles de lo esperado. Si un grabado en piedra se conserva durante 10.000 años, el papiro 3.000 años, los libros impresos 200 años, las películas 100 años, los discos duros pueden ser destruidos (*crash*) cada vez que los activa un sistema informático (desgaste mecánico, choque con la cabeza de lectura); los discos ópticos (CD, DVD), aunque se mantengan en buenas condiciones de temperatura y humedad, son fiables sólo entre 5 y 10 años (oxidación, arañazos en la superficie); la lectura de estos materiales depende de tecnologías que se vuelven obsoletas rápidamente (las cintas perforadas, cintas de audio y discos flexibles ya han desaparecido; el acceso a las cintas magnéticas, los discos ópticos, los discos duros externos y las memorias *flash* no son eternos). El problema tiene que ver con la estabilidad física de ceros y unos de según qué soporte, y con la conexión y acceso a la parte lógica de su agenciamiento de acuerdo con los dispositivos de lectura (pérdida repentina o total de acceso al contenido, a diferencia de los medios tradicionales, que se degradan gradualmente), en particular en el caso de las bases de da-

tos desarrolladas por los investigadores de humanidades digitales (colecciones originales de los documentos anotados y conectados entre sí).

Las tecnologías están pensadas para el futuro (memoria holográfica, nanotecnología, etcétera), pero tal vez no se puedan controlar o no sean económicamente accesibles. Los investigadores de Hitachi han demostrado que podemos almacenar datos a muy largo plazo grabándolos, en forma binaria, en el cuarzo, que luego se puede leer con un microscopio óptico sencillo; este grabado es eterno (resiste altas fluctuaciones de temperatura, microondas, agua y muchos productos químicos); pero su capacidad de almacenamiento es muy baja. Otra solución es la transcripción de la información en forma de hebras de ADN, una larga molécula que traspa el código genético de generación en generación. Se ha codificado ya un libro, imágenes y un programa en JavaScript de 5,37 megabits, lo que demuestra la posibilidad de conservar 5,5 petabytes por centímetro cúbico de ADN (el equivalente a 1.000 años de música en formato MP3). El ADN es muy estable (se han encontrado en hielo genes de 700.000 años y han sido leídos por técnicas de secuenciación) y podría ser una solución para archivar a muy largo plazo. Sin embargo, este almacenamiento depende de tecnologías avanzadas de síntesis y secuenciamiento del ADN, así como de algoritmos de procesamiento de datos; por consecuencia, no es de fácil acceso.

La fiabilidad de los soportes es muy relativa; las herencias conservadas no son necesariamente seguras. Por este motivo, los conservadores aplican estrategias diseñadas para copiar los datos a través de múltiples dispositivos, renuevan con regularidad las copias de seguridad en los nuevos medios y los migran (reescriben) a nuevos

sistemas a medida que aparecen cambios en los formatos y soportes. La conservación sostenible depende entonces de una actividad permanente de seguridad.

Entonces surge otra pregunta: la de los costes de conservación. Se trata de un debate muy debatido que conduce a la siguiente conclusión: a largo plazo (después de 50 años), el método económicamente más atractivo es digitalizar los libros y destruir los originales, ahorrando espacio y costes de almacenamiento. También reduce el coste de la digitalización, ya que permite no tomar ninguna precaución a la hora de escanear el libro (no importa su encuadernación, ya que será destruida). Al menos es así para trabajos cuya materialidad no posee un valor intrínseco. Sin embargo, debido al riesgo que supone perder las versiones digitales tras la destrucción de las versiones impresas, las bibliotecas deberán organizarse entre sí para retener una versión física.

La digitalización, sin embargo, plantea muchas preguntas acerca de los riesgos de pérdidas. Empresas, gobiernos e individuos se ven tentados a deshacerse de sus archivos físicos voluminosos (papel, negativos fotográficos, etcétera) para conservar sólo los ligeros archivos digitales. O bien la digitalización no mantiene todo de los originales. Captura y convierte sólo ciertos aspectos del documento original y descuida, por ejemplo, el peso del libro, el olor del papel, el tipo de encuadernación, incluso el grano y el color del papel. Algunos usuarios también se preocupan por su capacidad de probar la integridad y la autenticidad de un documento (un diploma, un auto de fe, una obra, facturas y recibos de pago, correspondencia electrónica). A otros les preocupa lo difícil que resulta reparar documentos digitales después de su degradación (mientras que un fragmento de pergamino es utilizable, tal no es el caso de un fragmento de



disco duro) y leerlos después de la desaparición de los dispositivos (*hardware* y *software*) relacionados con su escritura y lectura.

En cuanto a las técnicas como la OCR para extraer la información textual de un libro o un periódico y luego aplicar el reconocimiento automático del nombre de entidades (persona, lugar, fecha, organización, evento, concepto) para facilitar la búsqueda en el texto completo, la fiabilidad y la calidad son relativas: quedan letras y palabras no reconocidas o distorsionadas, y se omiten ciertas partes del texto. Se dan errores de reconocimiento que necesitan de la masiva movilización de los usuarios (*crowdsourcing*) para ayudar a mejorar la calidad de OCR y de la indexación (detección manual de errores, análisis, corrección de defectos para que esa versión sea más acorde al original).

El color de los documentos y su fiel reproducción también están en juego. A pesar del establecimiento de un patrón de prueba, la exactitud de la digitalización es siempre relativa y no siempre puede prevenir la pérdida, que puede no siempre ser visible para el ojo, dependiendo del tipo de compresión o sensor utilizado y demás riesgos técnicos, incluyendo el ruido electrónico.

Para poder acercarse a la original, los conservadores producen varias versiones digitales en formatos distintos: las imágenes de la página en formato TIFF sin comprimir en color o en escala de grises de alta resolución para hacer *zoom* en los detalles; un archivo de metadatos para el documento (para asociar la versión digital de los originales con otras versiones del documento y con otros archivos digitales), la posición de la imagen (incluyendo su paginación), las opciones técnicas y la historia de la digitalización, la información técnica de la lectura de los archivos, los derechos de acceso (privacidad, derechos de autor) y de uso (impresión, reproducción, modi-

ficación, anotación, extracción, etcétera); índices para la navegación en los documentos; el texto ocerizado, incluyendo la información sobre la ubicación de cada palabra en la página y la identificación de la estructura del texto (palabra, línea de texto, párrafo, diseño de página, imagen, título, leyenda, etcétera). Todo eso, además, se reproduce en diferentes medios de comunicación: un sitio de almacenamiento principal con las tecnologías de lectura y escritura rápidas y fiables; uno o más sitios de respaldo con el método de recuperación en caso de fallo o desastre en el sitio principal; sitios de almacenamiento que facilitan el acceso y la consulta rápida; además de sitios de almacenamiento temporal de acuerdo a las actividades específicas (investigación, preparación antes de verterlo a un archivo). La cuestión es que la información digital siga siendo legible y utilizable a largo plazo incluso si cambia el entorno técnico y los usos (incluyendo las prácticas de lectura).

Cuando se trata de materiales distintos a libros o periódicos, los riesgos de pérdida asociada a la digitalización se multiplican: grandes planos, fotografías, calcos, pinturas y acuarelas, muestras textiles, pergaminos, monedas y sellos, manuscritos iluminados, álbumes fotográficos, paquetes y archivadores, muebles, esculturas, juguetes, monumentos, etcétera. A menudo, la digitalización vuelve a transformar los objetos tridimensionales (un libro con su volumen y encuadernación) en imágenes bidimensionales (imágenes de la página). También transforma el ensamblaje que conforma el objeto, por ejemplo, en el caso de las fichas de cartón que conserva los estados del servicio militar de un soldado de la primera guerra mundial, sobre la que se pegan láminas curvadas en todas direcciones. La digitalización de una obra de un gran maestro, independientemente de la resolución seleccionada, pierde la información

oculta debajo del pigmento (otras capas de pintura, retoque, dibujado en lápiz) que puede ser descubierto por otras técnicas (de rayos X o por la tecnología del futuro). La digitalización volumétrica de máscaras aztecas pierde la información de la estructura y los materiales internos. La digitalización pierde inevitablemente algo del documento original. La cuestión que se plantea es cómo preservar los objetos cuando sus características físicas son también importantes.

La digitalización es también una oportunidad para corregir el original, por ejemplo, para encontrar los colores originales de fotos antes de perder el color o volverse amarillentas a causa del tiempo; en este caso, la restauración introduce cambios (saturación de color, contraste) dependiendo de las creencias y del conocimiento de los restauradores con respecto a lo que debía ser el original, pero no necesariamente se corresponde a lo que era el original. ¿Qué puede hacerse en el caso de una imagen borrosa y cómo saber si la falta de definición fue intencional o no, y si hay algo que corregir? ¿Qué hacer con los rastros de pliegues, poros e imperfecciones, lo que demuestra el uso y lo que les sucedió a estas viejas fotos? La digitalización difiere según el uso que se pretende hacer (estudio del objeto fotografiado, historia de las condiciones de disparo o la historia de la circulación de fotografías como objeto manipulado, ofrecido o expuesto).

La valoración del archivo también llega a alterarlo, por ejemplo, mediante la transformación de conciertos en *playlist* para su reproducción en el teléfono móvil omitiendo los aplausos y la atmósfera del concierto. La digitalización (al igual que la conservación o restauración, por ejemplo, de un edificio) implica un compromiso entre la preocupación por la conservación a largo plazo de

la autenticidad del objeto y la valorización de la herencia de acuerdo a las aspiraciones del momento.

Por su parte, la migración de un formato o medio a otro supone la transformación del contenido nunca libre de pérdidas de información similares al paso de una lengua a otra, especialmente para contenidos complejos, como una voz, una música o un vídeo; con el fin de emular en 20 o 50 años lo que era el original (un videojuego, un correo electrónico, un muro de Facebook o el arte digital) conviene también que el formato original haya sido bien documentado (que no es siempre el caso), que no sea un formato «privado» e inaccesible (un código fuente mantenido en secreto por la empresa) y que el formato de marcado del contenido (TEI, por ejemplo) no se haya utilizado constantemente a pesar de sus complicadas características. En el arte digital, que también juega con formatos singulares y no estándar, los artistas documentan lo que en su trabajo es esencial (resolución, velocidad de procesamiento, etcétera) y lo que es secundario para su recreación en el futuro a través de otros medios de comunicación y *software*. La más mínima negligencia en la preservación y documentación de los formatos puede hacer que los contenidos sean inaccesibles para siempre. Por el contrario, un libro en una estantería se puede encontrar casi sin cambios un siglo más tarde. Las nuevas profesiones de archivistas y curadores de datos se han dado a esta misión, pero la sostenibilidad de la misma también depende de las instituciones (bibliotecas, archivos públicos, derechos de autor, etcétera), de modelos de negocio y de los interesados cuyos intereses pueden a su vez ser inconstantes.

Sean cuales sean las nuevas técnicas y las capacidades de almacenamiento, lo más probable es que no todo pueda ser conservado. Los archivistas digitales, como siempre han hecho con los objetos

físicos, elijen lo que merece ser preservado por su ejemplaridad o singularidad: listas de compras, intercambios de correo electrónico, televisión, puntuaciones de videojuegos, consultas de sitios web extremistas, fotos tomadas por teléfonos móviles, visualizaciones interactivas diseñadas por los investigadores de humanidades digitales, sitios web de jóvenes entre los que algunos seguramente se han vuelto celebridades políticas y artísticas, *street view* del mundo en la década de 2010, etcétera. La selección, casi siempre colectiva, se debe hacer por lo tanto para no perder irremediable e involuntariamente algunos patrimonios.



## «La digitalización supone la muerte del libro y las bibliotecas.»

«El libro está muerto. [...] En cinco años habrá desaparecido.»

Nicholas Negroponte, profesor del MIT, *CNN*,  
6 Agosto 2010

A partir de 2015, en Francia, las ventas de libros aumentaron un 5% los primeros cinco meses del año. En 2014 se publicaron cerca de 100.000 libros, un 3% más que en 2013, con más de 43.000 novedades. Se vendieron un total de 422 millones de copias (un 1,2% menos que en 2013). En cuanto a volumen de negocio, el libro es más resistente que la música (que disminuyó un 7%) y el vídeo (cuya disminución fue del 10%).

El libro digital está creciendo pero todavía no tiene mucho peso; representó sólo el 6,4% de las ventas de libros en 2014. A nivel internacional tuvo mayor impacto (un 12% en 2013 con vistas a un 25% en 2018) y en Francia, en las ciencias sociales y humanidades (casi el 25%). El libro impreso aún no está muerto. En general, el libro sufre porque la lectura tiende a disminuir en relación con otras prácticas culturales: el tiempo de lectura semanal pasa de 5h48 en 2011 a 5h27 en el año 2012 y 5h20 en 2013; las personas que no compran libros se elevan del 18% al 22%, mientras que los grandes

lectores (mayoritariamente mujeres), que compran más de quince libros al año, cae del 10% en 2011 al 7% en 2013. La lectura, segundo pasatiempo favorito de los franceses, está en declive y no se ve respaldada por la lectura en formato digital, aun cuando los hogares están adquiriendo tabletas. No es seguro que estas tabletas vayan a animar a leer libros, mientras que los teléfonos móviles hacen leer y escribir más correos que lo que hicieron nuestros antepasados, pero en la forma de micromensajes. Las novelas clásicas y contemporáneas retroceden mientras avanzan las novelas románticas y policiales, los libros prácticos y los libros sobre historia.

En cuanto a las bibliotecas, la digitalización de las colecciones y su subida a Internet continúa hasta el punto de que algunos consideran que «se encuentra todo en Internet, así que ya no se necesitan bibliotecas». Las bibliotecas ponen el acento en la digitalización, se crean bibliotecas nuevas con colecciones exclusivamente para el formato digital, vistas a distancia, mientras que una selección de las colecciones raras y antiguas, los tesoros más preciados de bibliotecas, se destacan y toman vida. Un efecto paradójico de este acceso libre es el hecho de hacer más invisible lo que no se digitaliza: «Si no está en línea, es que no existe». Los estudiantes tienen dificultades para imaginar la magnitud de los recursos no digitales aún disponibles. Peor aún, con el pretexto de resolver el dilema entre la puesta a disposición del público y la preservación, la digitalización llevó a las bibliotecas a negar el acceso físico al manuscrito original.

Además, el acceso en línea a través de algunos motores de búsqueda sugiere que el acceso es inmediato y transparente. No se necesita un bibliotecario para orientar una búsqueda documental ni para clasificar los libros. Estos profesionales de la gestión del conocimiento se vuelven más invisibles de lo que eran cuando el acceso



a los documentos dependía de un importante trabajo de descripción de los objetos catalogados y de la organización de las colecciones que ayudaba a contextualizar y dar sentido a estos objetos. La mediación entre el contenido y los usuarios, que llevan a cabo bibliotecarios y documentalistas, ya no está tan clara, debido a la idea de la comunicación directa, inmediata y transparente a través de Internet, que median a su vez las herramientas de investigación e indexación, los algoritmos a menudo diseñados por empresas que sacan provecho económico de esa mediación, así como por la gestión de ramas de la información (revistas o libros, en particular) que venden a bibliotecas ciertas compañías multinacionales de la edición, las cuales, de paso, recogen información sobre los lectores (sus búsquedas, sus lecturas) y la explotan para refinar sus estrategias de negocio. La mediación también requiere bibliotecarios que organicen sus fondos en subcategorías de documentos en función de las prioridades de conservación (periódicos viejos que se degradan), de las peticiones de consulta, de la coherencia temática a partir de su valoración y de nuevos factores (por ejemplo, un cambio de ley que permite la digitalización de obras huérfanas, agotadas o rara vez consultadas). A veces, estos documentos digitalizados se ponen en línea total o parcialmente (*snippets*), por lo menos para cierto público estudiantil o discapacitados, o bajo ciertas condiciones (por ejemplo, documentos cuya consulta sólo se permite en cubículos de la biblioteca o a través de una red segura). Los usuarios a veces tienen derecho a descargar las obras para su trabajo de investigación. En algunos casos, especialmente en los EE. UU., la ley autoriza el uso de tecnologías de registro de datos (*text y data mining*), mientras que las grandes editoriales incluyen este derecho en los contratos de autor y cobran un canon a bibliotecas e investigadores. Estos usos

y difusiones restringidas o condicionales, ya sea en relación a la ley o a las estrategias comerciales de los editores de contenido, limitan el interés por lo digital y confirman que el acceso inmediato para todos es un mito.

Ni el libro ni las bibliotecas están muertos, pero la lectura disminuye por falta de tiempo y por otros tipos de ocio que atraen a la gente. El hecho de navegar por Internet, sin embargo, va de la mano de ser un gran lector, mientras que el 43% de aquellos que no utilizan Internet no son lectores. Por el contrario, las maneras de leer cambian; la navegación por Internet, pasando por el hecho de hacer clic para activar unhipervínculo o las sugerencias generadas por algoritmos, rompen la linealidad del libro (del autor), y generan una lectura fragmentada en la que interfieren las empresas de contenido editorial. Se le impone al lector la selección de la información y se le define un camino con puntos de referencia para calificarla en función de las páginas donde la encuentra, sin ser capaz de tener una comprensión global. Por su parte y debido a su materialidad, el libro proporciona algunas indicaciones en cuanto a la importancia de la obra, la ubicación de la información en el libro y las secuencias entre la información. Promueve una lectura lenta, una atención minuciosa y seguida del desarrollo de un pensamiento, mientras que los medios digitales exigen más bien a una hiperatención furtiva y habilidades de circulación y de asociación veloz. El riesgo, dice Nicholas Carr, es que «la lectura clásica se vuelva insoportable, incluso físicamente». La lectura de libros podría convertirse en una forma de resistencia a la hipermovilidad en las redes de información y los flujos de atención.

Con la digitalización, el libro se transforma. No existe hasta que se regenera mediante una serie de operaciones y *software*, con

un diseño que depende del dispositivo de reproducción utilizado por el lector y los ajustes seleccionados (tamaño de la fuente en particular), y traducido a la lengua de su elección, lo que hace que el libro sólo exista en esa lengua mientras se lee. Por otra parte, el libro digital se puede actualizar más rápido que la impresión de una nueva edición; el conocimiento que transmite parece ser menos estable y más fluido que el conocimiento relacionado con el libro impreso. Nuestra concepción de la verdad también sufrirá quizá las consecuencias de ser más efímera. Con lo digital, la lectura en sí cambia el texto, ya que deja rastros (el tiempo que el lector se quedó en una sola página, sus anotaciones, etcétera) que pueden acumularse en bases de datos y ser procesados para producir nueva información que se agregue al libro. Si los libros electrónicos pueden reproducir el sonido de las páginas que se giran con la punta del dedo, el doblar de la hoja que comienza a girar y la foliación del volumen controlado por el pulgar separa a la obra de su encuadernación y autonomiza sus páginas. Por los enlaces que se pueden agregar para conectarlo a un comentario o a otros textos, el libro se convierte en un terminal desde el que el lector puede navegar por un paisaje de datos y generar otras visualizaciones de la obra, por ejemplo, en forma de tablas de datos o gráficas (*analytics*), el número de visitas o de *like*, libros comprados por otros lectores de este libro, la red de comentarios sobre el libro, etcétera.

Si el lector se reconfigura, el autor y su autoridad también. ¿Qué más hace el autor cuando el libro leído se genera con base en algoritmos del editor, ajustes del lector y las relaciones con otros textos y lectores? ¿No sería entonces la muerte del autor?



## «Las tecnologías digitales reemplazarán a los investigadores de las ciencias humanas y sociales.»

«Existe un plan secreto para sustituir  
a los científicos humanos por robots.»

Brett Bobley, director de la ODH  
(Kolowitch, 2011)

Las tecnologías digitales deberían aumentar la capacidad de los investigadores para inventar nuevas formas de analizar e interpretar de forma automática grandes cantidades de datos, de tal forma que sus estudiantes, los públicos y ellos mismos puedan recuperar el contexto de un documento antiguo, dar acceso al significado de los acontecimientos históricos y la experiencia ordinaria de una profesión o una comunidad remota, y así mejorar nuestra comprensión de la humanidad. Con estas tecnologías, las CHS nos ayudarían a sumergirnos en otros lugares y otros tiempos, por ejemplo, sentir la dinámica espacial y acústica de un discurso o un sermón en la plaza de St. Paul para entender la eficacia de la predicación pública como el método preferido de relaciones públicas de las autoridades eclesiásticas y políticas del Londres de principios del siglo xvii.

Estas tecnologías ayudarían a cerrar la brecha entre las ciencias «de verdad» y las CHS, permitiendo a sus investigadores, a través

de un análisis automatizado, distanciarse del trabajo de interpretación subjetiva y de sus enfoques para defender sus tesis. Gracias a las herramientas informáticas, ahora se pueden tratar muchos casos (eventos, lugares, personas, objetos y textos), compararlos y evaluar su relevancia, su representatividad o especificidad. Esto debería llevarnos a confirmar firmemente cosas ya conocidas, a desafiar resultados y descubrir otros nuevos. De este modo, el estudio de los procesos de la corte criminal central de Londres (*Old Bailey*), basado en la digitalización y la transcripción de 198.000 casos entre 1674 y 1913 y el uso de herramientas de registro de datos, permitió descubrir un aumento inusual del número de declaraciones de culpabilidad, lo que representa un punto de inflexión en el desarrollo del sistema acusatorio moderno de justicia en Inglaterra y en la América colonial. Las herramientas de análisis digital sirven para progresar en la comprensión del pasado y para que las CHS sean más rigurosas.

En los Estados Unidos, la *Office of the Digital Humanities* (ODH) de la *National Endowment of the Humanities* (NEH) apoya financieramente estos avances científicos y tecnológicos de las CHS. Para hacer frente a estos retos, se están dando impresionantes colaboraciones científicas internacionales e interdisciplinarias, dejando al investigador solitario en su aislamiento para sumergirse en la *big science* y las redes de la megaciencia como la física, las ciencias de la vida y la nanotecnología. Con el programa *Digging into Data* financiado por las agencias de apoyo a la investigación británica, canadiense, holandesa y estadounidense, se busca responder a preguntas como: «¿Qué hacer con un millón de libros? ¿O un millón de páginas de los periódicos? ¿O un millón de fotografías de obras de arte?»

Las CHS se volverían así verdaderas ciencias hasta el punto de seducir a los investigadores de física, matemáticas, informática e ingeniería para enfrentarlos a nuevos retos científicos. De hecho, las CHS les interesan, ven en ellas un nuevo campo de investigación (*computational social science, social physics*, sistemas complejos, teoría de redes, *big data*, etcétera) y proyectan la creación de grandes infraestructuras *CERN-like* para medir nuestras prácticas y aumentar la conciencia que tenemos de nosotros mismos y de nuestro pasado como sociedad (*ICT-enabled Self-Awareness*), así como nuestra capacidad de resiliencia a los cambios venideros (crisis financieras, crisis ambiental, innovación).

Con el uso de las tecnologías digitales aún más sofisticadas (la modelización matemática, *machine learning*, etcétera) sería incluso posible gracias a una serie de algoritmos rastrear 1.000 años de historia en unos pocos clics para buscar en 80km de archivos (*Venice Time Machine*), reconstruir el aprendizaje de una lengua y sus transformaciones históricas y anticipar sus evoluciones futuras, hasta diseñar máquinas que piensan, o al menos, máquinas a las que los seres humanos no serían capaces de distinguir de un ser humano (prueba de Turing). El investigador en CHS corre el riesgo de ser reemplazado gradualmente por otras disciplinas consideradas como más científicas y que tienen la ventaja de dominar la tecnología.

El fenómeno es evidente cuando los expertos en *data science* se enorgullecen de estudiar los fenómenos sociales a través de los mil millones de huellas digitales que dejamos mediante nuestro uso de teléfonos móviles, GPS, micromensajes, consultas y visitas a sitios de Internet, metadatos relacionados con las imágenes que colocamos en sitios de intercambio, del Bluetooth y sus sensores, de la información confiada a nuestras redes sociales digitales, de nuestra

tarjeta de crédito o de cliente, etcétera. Con esta cantidad de datos que proporcionamos muy a la ligera se estudian fenómenos de carácter social global, las redes de relaciones sociales, la ubicación de las acciones individuales (telefonar, tomar fotos, comprar) y los desplazamientos. Así pues, asistimos a una proliferación de visualizaciones como el *Vizualizing Facebook Friends*, que cartografía diez millones de relaciones de «amistad», la ubicación de los turistas a través de las fotos que publican en Flickr o las tuitografías que revelan la densidad de los mensajes transmitidos en cada esquina de ciudades como París, Londres o Nueva York.

Si los científicos sociales tardaron mucho en explotar estos datos, no fue el caso de otras disciplinas que publican sus resultados en revistas de prestigio como *Nature* o *Science*, a pesar de existir una evidente falta de reflexividad metodológica sobre la calidad, el significado y origen de los datos. Cometan grandes errores debido a su ignorancia sobre el conocimiento acumulado de las CHS, o bien redescubren evidencias demostradas hace tiempo por estas disciplinas (como el papel fundamental del trabajo, la vivienda y la escuela). Los sesgos de representatividad son fenomenales: Twitter, por ejemplo, ampliamente utilizado como fuente porque los datos son abundantes, proporciona la ilusión de visión global, pero se ignora que esta aplicación se utiliza sólo marginalmente por la población. Otras aplicaciones para teléfonos móviles también producen una gran cantidad de datos, por ejemplo, el *Street Bump* que permite señalar degradaciones en las calles en Boston, a pesar de que provienen de una parte específica de la población, jóvenes de las zonas ricas, dando la impresión de que el suelo está más degradado en esos barrios que en otros lugares. El fácil acceso a las huellas digitales alienta su uso, pero tiene el efecto de orientar la investigación hacia



temas en los que hay rastros digitales (efecto de reverberación) en lugar de abordar las cuestiones para las que hay que producir datos originales. Con estos conjuntos de datos, algoritmos, procesamientos y muchas visualizaciones estetizadas, estas disciplinas miran con condescendencia a los competentes pero repentinamente obsoletos investigadores de las CHS.

Estos datos sugieren que los investigadores de CHS les evalúan seriamente. En la práctica, son principalmente investigadores de cualquiera de otras disciplinas interesados en esta rama, o empresas como Google, Facebook y otros, o agencias gubernamentales como la NSA (*National Security Agency*) de los Estados Unidos que vigilan y analizan el mundo con cantidades colosales de información a su disposición.

No obstante, estos datos tienen su utilidad. Pueden ayudar a estimar la población (en las ciudades del norte o en países en vías de desarrollo), los viajes (para el trabajo, el ocio o la acción colectiva), las actividades, y repensar las divisiones territoriales (barrios, regiones). También pueden alentar un análisis crítico de los discursos o de las decisiones políticas, como hicieron Antonio y Paola Casilli Tubaro, investigadores de ciencias sociales, sobre los disturbios del 6 al 9 abril de 2011 en Londres. Tras el estallido, el primer ministro David Cameron reaccionó inmediatamente acusando a la mensajería instantánea y a las redes sociales de haber permitido coordinarse a los alborotadores; hay dos personas condenadas a una pena de prisión por haber hecho un llamamiento en Facebook, a pesar que dicho llamamiento no obtuvo repercusión alguna. Para evitar este tipo de disturbios, la policía ahora puede cortar las comunicaciones electrónicas. Los científicos sociales invalidaron este análisis político fue rechazado de inmediato por haciendo uso de la

simulación por ordenador para poner a prueba la teoría de D. Cameron que insistía en que las redes sociales habían amplificado los disturbios. Sus simulaciones, sin embargo, mostraron que si la comunicación es libre, la violencia insurgente es breve, mientras que si se impide la comunicación, los disturbios tienden a ser prolongados y tardan más en volver a la normalidad. Esta «*just in time sociology*», que publicó rápidamente sus resultados de forma libremente accesible y que fue ampliamente traducida e seguida por los medios de comunicación, muestra que las ciencias sociales también pueden iluminar casi en directo la acción política y participar en el debate público, sin necesitar un tiempo extenso para comprender los hechos pasados. Algunos se aventuran a continuación a exigir que las CHS se pasen a lo «justo a tiempo» digital, o de lo contrario desaparecerían y serían reemplazadas por una física social o por unas humanidades digitales mejor gestionada por matemáticos que por humanistas.

Las CHS deberían entonces adaptarse aun a riesgo de abandonar la reflexión crítica sobre la obtención de los datos (su digitalización, el desarrollo de modelos de datos, procesamiento de OCR, la creación de bases de datos, procesamiento y visualización), pues se supone que así se salvarían de la subjetividad. El hecho de no recurrir a las herramientas digitales supone la prueba de esta «tardanza». Los jóvenes diplomados en CHS, *nativos digitales* (nacidos con lo digital y que lo tendrían en la sangre) seguramente se adapten mejor. La «*just in time sociology*», que ha logrado introducir una voz disonante en el debate sobre el papel de las redes sociales, ve sin embargo su alcance muy limitado. Para hacerse visible, se sitúa en el mismo plano epistemológico que el debate que supone que los alborotadores son agentes racionales e individuales. Esta «*just in*

*time sociology*» evita ser descalificada y puede ser escuchada por los medios de comunicación y los políticos, pero no ayuda a ganar perspectiva sobre su propio modelo de análisis de la realidad social, que podría haber sido una investigación exhaustiva e inédita. Además, el riesgo para las CHS, incluso si se apoderan de lo digital y no se dejan superar por los *data scientists*, es que las absorban las ecologías digitales (la masa de datos y la potencia de las herramientas) y los marcos de pensamiento, elementales e ingenuos, asociados con éstas. La fascinación por las herramientas (donde basta con preguntar, como en Google) y las visualizaciones puede debilitar las construcciones intelectuales necesarias para la comprensión adecuada de lo que nos ayudan a conocer.

Entonces, en lugar de una colonización de las CHS por los *data scientists* o de su sustitución por robots, ¿no deberían las CHS colonizar las *data sciences* y la informática? Stanley Fish, profesor de la Universidad Internacional de Florida en 2011, escribe sobre lo que está sucediendo en última instancia en las humanidades digitales: «Las humanidades han penetrado de forma activa, han colonizado, los campos que se suponía iban a desalojarlas».



# LAS HUMANIDADES DIGITALES EN EL MUNDO



## «Con las humanidades digitales se despoja al Sur de su patrimonio cultural.»

«El atractivo de la ayuda financiera ha creado una nueva forma de imperialismo que fortalece la brecha digital: los países del Norte se dedican a saquear la propiedad intelectual del patrimonio africano en nombre de la preservación.»

Michele Pickover y Dale Peters, «DISA: An African Perspective on Digital Technology». *Innovation*, 2002

La digitalización del patrimonio cultural se presenta a menudo como una forma de preservarlo de la destrucción del paso del tiempo (una degradación debida a los cambios físicos y químicos de la materia y a los insectos, incendios, terremotos y daños de las inundaciones) y de los humanos (conflictos armados, saqueo, vandalismo, destrucción deliberada, «limpieza cultural», etcétera). Entre los ejemplos más conocidos de destrucción del patrimonio cultural se incluyen: los archivos y edificios durante la Revolución francesa; el patrimonio cultural de los nativos de América del Norte por los jóvenes Estados Unidos en el siglo XIX; el patrimonio de Palestina, en 1948, por milicias sionistas para presentar a los palestinos como «un pueblo sin cultura»; el patrimonio de miles de años destruido por la

Revolución Cultural en China y el Tíbet en 1966; la destrucción de los budas de Bamiyan en Afganistán (dinamitados en 2001 por los talibanes), los mausoleos de Tombuctú (en 2015), las esculturas preislámicas en Mosul (Irak, en 2015), y monumentos como el arco de triunfo de Palmyra en Siria en 2015, etcétera. En cuanto al patrimonio cultural inmaterial, incluida la diversidad lingüística, la erosión proviene tanto de las políticas nacionales de armonización como de las hegemonías culturales relacionadas en particular con el surgimiento de la Web, donde dominan idiomas incluidos en los algoritmos de recomendación.

Para luchar contra estas amenazas para el patrimonio cultural de la humanidad, las instituciones, como la UNESCO, desarrollan políticas y programas de conservación, entre otras cosas mediante la elaboración de inventarios y la digitalización del patrimonio, asegurando su acceso en línea. En el caso del patrimonio cultural inmaterial, estos inventarios ayudan a concienciar a sus poblaciones de la importancia de este patrimonio en el refuerzo de su identidad colectiva, su autoestima, su creatividad y, sobre todo, del desarrollo económico del turismo. En Brasil, en la década de 1930, el Instituto del Patrimonio Histórico y Artístico Nacional (IPHAN) trabajaba para proteger el patrimonio cultural material y preservarlo, en primer lugar, en su forma original. En la década de 1970, el concepto se fusiona con el de patrimonio intangible y pasa por su identificación, documentación e interpretación, actualizándose regularmente, como un patrimonio vivo y en evolución. El Inventario nacional de referencias culturales tiene así la intención de preservar la diversidad étnica y cultural del país (pueblos indígenas, afrobrasileños y el medio urbano multicultural) en un contexto marcado por la homogeneización. Un bien (sea éste una celebración, una



forma de expresión, la artesanía y los conocimientos tradicionales, o los lugares o espacios materiales), una vez reconocido como parte del patrimonio cultural nacional, cuenta con el apoyo para su promoción y protección, incluida la integración social de los tutores de estas tradiciones y la mejora de sus condiciones de vida.

Los patrimonios en peligro de extinción son objeto de constantes y urgentes copias de seguridad. Así, en las Filipinas este trabajo se llevó a cabo para documentar el conocimiento tradicional relacionado con las plantas, cuya biodiversidad es cada vez más reducida, y cuyos usos tienden a no transmitirse de generación en generación. Este conocimiento se presenta en medios audiovisuales utilizados particularmente en las escuelas. En Mali, tras el golpe de 2012, las prácticas culturales (rituales, música, fiestas y artesanías) están prohibidas y las identidades colectivas se debilitan. La salvaguardia de este patrimonio está relacionado con la reconstrucción de una armonía social después de los conflictos. En otros casos, como la epopeya tradicional mongola, se trata de revitalizar y transmitir el patrimonio a través de la creación de bases de datos de las costumbres y de su enseñanza. En el caso de ciertas tradiciones orales en Vietnam, donde canciones populares, cuentos y el lenguaje tradicional desaparecen debido al entorno escolar y profesional dominado por el idioma nacional, se desarrolla un material audiovisual para transmitir esta herencia a los niños y mujeres mediante la movilización de los jefes del pueblo, los ancianos y artesanos. En Uruguay es la llamada de los tambores de candombe, específico de cada barrio, lo que se documenta y perpetúa mediante la promoción de intercambios entre los barrios y las generaciones y la documentación en Internet. En el caso del carnaval de Barranquilla, en Colombia, se crea un carnaval infantil para promover el aprendizaje

de la mezcla de tradiciones indígenas, africanas y europeas, así como una base de datos de las expresiones carnavalescas (coreografía, música, teatro). En el caso de los manuscritos medievales preciosos de Tombuctú, se inauguró en 2008 un proyecto para escanear los documentos a cargo del Instituto Nacional de Ciencias Aplicadas (INSA) de Lyon.

La digitalización es una forma de asegurar la transmisión de la herencia cultural. También es una manera de compartir (por ejemplo, a los turistas potenciales) y hacer circular dicha herencia, por ejemplo, para repatriar a su país o comunidad de origen copias digitales de objetos, fotos, películas y textos que habían sido robados por las potencias coloniales o por saqueadores y traficantes de antigüedades. Este tipo de repatriación virtual es controvertido, ya que no responde necesariamente a las expectativas de las comunidades de origen, que exigen los documentos originales o señalan los riesgos de (fuera de su contexto ritual en particular, o comercialización).

La digitalización del patrimonio cultural en los países en vías de desarrollo, con el pretexto de preservación, supone también una amenaza de despojo, porque las versiones digitales se llevan a cabo por instituciones o empresas del Norte, que controlan el acceso y uso, incluida su comercialización. Los principales proyectos de digitalización del patrimonio cultural africano, por ejemplo, son financiados por los donantes en función de sus propios intereses y prioridades (particularmente en términos de la selección de materiales para escanear), lo que puede conducir a una nueva forma de imperialismo. La digitalización del juicio de Nelson Mandela, originalmente grabada en cintas cuyos lectores ya no existen en Sudáfrica, se llevó a cabo por el Instituto Mellon en los EE. UU., que ahora

es dueño de los derechos (*copyright*) de la versión digital. Otros archivos importantes, como las luchas de liberación o la literatura en Sudáfrica, están en servidores de Estados Unidos; la digitalización de archivos del arzobispo Desmond Tutu la asume una universidad británica. Tal «cooperación» digital se traduce en una «fuga de archivos» de los países de origen y en un retraso en el dominio del proceso de su digitalización, y por lo tanto, en la profundización de la brecha digital entre los países hegemónicos y no hegemónicos.

Algunos países proporcionan dispositivos que incitan a sus comunidades indígenas a grabar sus saberes ancestrales con el fin de protegerlos y valorizarlos. Es el caso de la ley peruana, que permite al Estado desarrollar económicamente este patrimonio, entre otros, a través de contratos con empresas o naciones extranjeras. Estas comunidades podrían recibir entre el cinco y el diez por ciento de los contratos sobre sus conocimientos ancestrales.

Los proyectos de cooperación digitales Norte-Sur proporcionan asistencia a los países que no tienen necesariamente los medios para salvar su propia herencia. En el caso de Nigeria, la Universidad de Iowa mantiene los datos digitales sobre las elecciones y proporciona acceso a una intranet a las instituciones africanas que no tienen acceso a Internet de banda ancha. Otros proyectos promueven archivos africanos de fotos en línea y audiovisuales con transcripciones en inglés y francés, diarios de campo de antropólogos que han trabajado en estas áreas o documentos relacionados con la esclavitud. El proyecto Gallica de la BNF en Francia ha digitalizado obras de la época precolonial, mientras que el Museo del África Central en Bélgica ha digitalizado fotografías del Congo.

La restitución pública casi siempre se limita a menudo a las versiones degradadas (de baja resolución), para facilitar la consulta

a distancia y luchar contra la copia comercializable. Tiene el mérito de dar a conocer las obras y atraer al público a los museos o al acceso previo pago a una reproducción de mejor calidad (texto, fotografía, sonido, vídeo, modelo 3D). Muchos sitios web de archivo generan varias versiones digitales de calidades y formatos diferentes de acuerdo al uso. En el caso de modelado 3D de diez estatuas de Miguel Ángel, incluido el *David*, la versión digital de alta definición sólo está disponible para los investigadores bajo estrictas condiciones (por ejemplo, sólo en cubículos especializados) y con fines no comerciales (el mercado de copias de estatuas se apoderaría rápidamente de los modelos de alta definición). La producción de copias también ayuda a difundir el acceso al público, en las escuelas o en los museos, y permitir que se interactúe con el objeto mientras el original está cuidadosamente preservado a distancia del público y su acceso, restringido a una estricta minoría de personas. El acceso extendido a través de copias digitales funciona como un acceso más restringido a los originales o archivos digitales en alta definición, creando así nuevos desequilibrios entre los públicos, las comunidades interesadas y las instituciones que poseen los derechos sobre las versiones digitales.

También supone un riesgo que los países u organizaciones hegemónicas ejerzan un control del patrimonio cultural digital de los países en desarrollo. El riesgo es aún mayor si las que están a cargo son empresas privadas. Las instituciones y empresas del hemisferio Norte son propensas a explotar estos patrimonios culturales del Sur y «proteger» éstos con el Derecho, con la perspectiva de comercializarlos, como hicieron con la biodiversidad y los conocimientos médicos autóctonos de los países que no tienen los recursos para asegurar los costes de protección de la propiedad intelectual. Éste

es el caso de los patrones culturales indígenas que repiten los diseñadores y la industria de la moda. Los aborígenes australianos, por ejemplo, se quejan de ver así comercializado su patrimonio cultural, música, leyendas y estilos, que se copian con mayor facilidad cuando están escaneados. Las instituciones públicas también protegen su digitalización para vigilar su uso (el acceso privilegiado a su propio investigador, censura) y beneficiarse en cuanto a audiencia, prestigio o la venta de productos derivados (que van desde postales, camisetas y catálogos hasta los mismos metadatos).

La reproducción digital de los bienes culturales y su accesibilidad (por ejemplo, los sitios arqueológicos), facilitado por las impresoras 3D, también genera el riesgo de hacerlo menos atractivo como destino turístico, aspecto de vital importancia para la economía de algunos países. Del mismo modo, el enriquecimiento de los videojuegos o juegos serios con contenido cultural de alta calidad (la reconstrucción de la antigua Roma, los Juegos Olímpicos de la antigüedad, un episodio de la guerra 14-18, un ritual de iniciación o la experiencia de la migración y colonización), que permite a los jugadores vivir una experiencia enriquecida, puede tener efectos negativos para las economías turísticas, mientras se enriquecen las empresas que diseñan estos juegos.

En el año 2010, hubo una controversia sobre el uso del patrimonio cultural indígena para transmitir un mensaje político sobre la armonía entre las comunidades indígenas y la sociedad canadiense con motivo de los Juegos Olímpicos de Invierno en Vancouver, promovidos como «indígenas» con un logotipo consistente en una escultura inuit. Las comunidades indígenas entendieron esta publicidad como una agresión por parte de los organizadores, que ignoran u ocultan la dura realidad de su condición.

La apropiación del patrimonio cultural por otros grupos sociales surge con su digitalización. Así, los exégetas cristianos que tienen acceso y trabajan en las versiones de 5.700 manuscritos griegos del Nuevo Testamento descubren que no son los únicos interesados en esta antigua herencia; algunos grupos extremistas musulmanes, en sus propios sitios web, comentan en árabe estos textos bíblicos y ponen en duda la legitimidad de las minorías religiosas amenazadas a nivel local. Lo que ocurre con este patrimonio religioso ocurre con otros conocimientos y prácticas culturales a raíz de intereses comunitarios, científicos, políticos o económicos. Ciertos grupos sociales, instituciones públicas y empresas dependen de la digitalización del patrimonio no sólo para conservarlo, sino para transformarlo en recursos para su trabajo, especialmente en la conformación de su identidad y legitimidad, de una contracultura o de una posición competitiva.

Incluso el uso de versiones digitales por parte de los investigadores es a veces controvertido. Algunas comunidades, como los amerindios Stó:ló en Canadá, están interesadas en su pasado y quieren reapropiárselo. Pero a menudo los objetos de estudio y las preguntas de investigación se definen más por el calendario de la investigación académica que por las comunidades afectadas. Por otra parte, la propiedad intelectual y los datos obtenidos en la búsqueda son propiedad de los investigadores o sus instituciones, y no de esas comunidades. Estas cuestiones de distribución equitativa de la autoridad y de los derechos sobre la investigación del patrimonio son en sí mismas el objeto de la investigación.

De manera más general, se hace un uso fraudulento de las imágenes de los objetos del patrimonio cultural, y existe una tendencia generalizada a la copia digital ilegal de las obras, a menudo aso-

ciada con las redes del crimen organizado, pues se trata de una de las operaciones ilícitas más lucrativas para los copiadores. Además de la explotación económica, este fenómeno priva a las comunidades de su propia herencia. Hablamos de extorsión cultural, fenómeno antiguo desde el saqueo de las pirámides de los faraones y los incas, cuyos objetos se venden a precios elevados en los países desarrollados, donde hay que incluir los museos públicos. Esta extorsión cultural ahora gana dimensión digital.





## «Con lo digital, la hegemonía del inglés es cosa del pasado.»

«El inglés sin duda jugará un papel importante en la conformación del nuevo orden lingüístico mundial, pero su fuerte incidencia dará a luz a nuevas generaciones de hablantes bilingües y multilingües en todo el mundo.»

David Grabbol, «The Future of language», *Science*, 2004

Si el lenguaje es el vehículo preferido de la cultura, ¿qué pasa con él en el mundo digital y en la sociedad?

La influencia del inglés ha crecido en el siglo xx debido a la dominación económica y política de los Estados Unidos. Se convirtió en la principal lengua de comunicación internacional y contribuye a la creación de una forma de hegemonía cultural estrechamente asociada con la economía liberal y la globalización de los intercambios comerciales. Su uso es a veces visto como ideológico; marca la modernidad y la apertura internacional. También se discute: hay una resistencia a la dominación cultural con el objetivo de defender las lenguas nacionales y el multilingüismo; la defensa de una lengua internacional, como el esperanto, más accesible; la equidad política (trabajar en inglés es dar una ventaja competitiva a quienes la tienen como lengua materna).

En este contexto, ¿es Internet el vector de la dominación universal del inglés o promueve la pluralidad lingüística, facilitando la

transición de un idioma a otro, usando la traducción automática y preservando las lenguas en peligro de extinción? En sus inicios, en 1969, Internet fue exclusivamente en inglés, creado por el Pentágono, las agencias gubernamentales y universidades americanas. Después de 1989, con la creación de la Web y los motores de búsqueda, se desarrolló primero en los Estados Unidos y luego en el resto del mundo. El inglés domina aún cuando sólo el 9% de la población mundial es hablante nativo. Lo digital no refleja la diversidad lingüística, como también se observa con el proyecto Gutenberg, que en 2011 había digitalizado y dado acceso libre a cerca de 31.000 libros en inglés, sólo 1.800 en francés, 800 en alemán, 600 en finlandés, 500 holandés y portugués, 400 en chino, 300 en español, etcétera.

La diversidad lingüística, sin embargo, es cada vez mayor. En el año 2015, si bien el inglés sigue imponiéndose con 850 millones de usuarios en Internet, su peso relativo pasa del 75% en 1998 al 45% en 2008, y al 26% en 2015. El chino le sigue de cerca (21,5%), con un crecimiento fuerte. Luego vienen el español (7,5%), árabe (4,8%), portugués (4%), japonés (3,5%), ruso (3,2%), malayo (2,9%), francés (2,8%) y alemán (2,6%). Esta diversidad lingüística es relativa; diez de esos idiomas representan alrededor del 80% de los usuarios de Internet y contenidos disponibles en la Web, mientras que en el mundo se hablan 6.000 lenguas.

Sin embargo, la decadencia del inglés y el surgimiento de otras lenguas sugiere que lo digital promueve el multilingüismo. Los principales actores de la Web contribuyen otorgando más peso a los idiomas locales; el motor de búsqueda de Google soporta 345 idiomas desde el año 2009; Wikipedia, 285; Facebook, 70 en 2012; Twitter, 21, y LinkedIn, 17. Por otra parte, las humanidades digitales y la

Unesco promueven con el proyecto *B@bel* la conservación de las lenguas en peligro de extinción. Para la UNESCO se trata de la defensa de la diversidad lingüística y cultural, lo que incluye el desarrollo de herramientas multilingües en línea y herramientas de traducción automática, la adaptación de sistemas de indexación multilingües de los sitios web y herramientas terminológicas, el fortalecimiento de la interoperabilidad para promover la digitalización y la puesta en línea de los archivos, e influir en las políticas nacionales e internacionales.

Dentro de las humanidades digitales, a pesar de su vocación de trabajo por el patrimonio cultural de la humanidad, la situación también parece estar dominada por el mundo anglosajón, favoreciendo de hecho la hegemonía del inglés. Los primeros centros de investigación surgieron en los Estados Unidos, sobre todo en la Costa Este. El primero, establecido en 1994 en la Universidad de Virginia entre los departamentos de historia e informática y la biblioteca para dar cabida a varios proyectos que marcan la historia de las humanidades digitales: *The Valley of the Shadow*, el *Archive Rossetti*, el *Archive William Blake*, el desarrollo del lenguaje de indexación TEI y ciertas herramientas de visualización y anotación. Luego se crearon otros centros de este tipo en la Universidad de Maryland, en la UCLA (donde se dio el primer manifiesto de las DH), la Universidad de Michigan, en Georges Mason (responsable del diseño de herramientas de acceso libre como Zotero, para la gestión de referencias bibliográficas, y Omeka, para la gestión de bases de datos, la creación de las primeras no conferencias THAT-Camps), etcétera. Crecen en otras partes del mundo de habla inglesa, como en Canadá (Victoria, McGill en Montreal), Australia (Melbourne, Newcastle), Irlanda (University College Dublin) y

Gran Bretaña (Oxford, Glasgow, UCL y el Kings College Londres, Birmingham). Los países de habla no inglesa, a pesar de que están metidos en grandes proyectos de digitalización, en principio estos proyectos son bastante marginales en Alemania (Colonia, Göttingen), los países escandinavos, Suiza (Berna, Zurich, Lausana, y el Portal de las ciencias históricas de la Academia Suiza de Ciencias Humanas y Sociales Infoclio) y en los países latinos (España, Italia, Francia, Suiza).

Las sociedades científicas también reflejan la hegemonía anglosajona. La revista *Literary and Linguistic Computing* se creó en Gran Bretaña en 1986. La asociación correspondiente (ALLC) se convierte, en 2012, en *The European Association for Digital Humanities* (EADH), pero sigue en manos de los británicos. En 1978 se crea la Association for Computers and the Humanities (ACH), con sede en Estados Unidos, para animar el debate, especialmente sobre preguntas frecuentes, *DH Answers*, y la revista *Digital Humanities Quarterly*. Las diferentes asociaciones se unen progresivamente desde 2002 para formar la Alliance of Digital Humanities Organizations (ADHO), que organiza la conferencia anual *Digital Humanities DH*. En el continente, los investigadores italianos, españoles y alemanes se estructuran progresivamente y forman sus respectivas asociaciones nacionales en 2011 y 2012. La comunidad latinoamericana también se estructura en el mismo periodo. En 2014 se crea la asociación francófona Humanística. El EADH también pone en marcha el grupo Global Outlook (GO: :DH), para ampliar la cooperación más allá de las fronteras lingüísticas particulares.

La comunidad se expande y se diversifica, pero la cuestión de la hegemonía lingüística preocupa a los países que no hablan inglés. Durante la primera THATCamp francesa en 2010, el mani-

fiesto de las HD fue escrito y traducido a 11 idiomas. El DH2012 en Hamburgo se centra en el tema de la «Diversidad Digital». Unos meses antes se lanzó sobre el tema una encuesta en línea, disponible en cuatro idiomas, llamada «¿Quién eres tú, humanista digital?», para examinar su espectro lingüístico y geográfico. La encuesta recibió respuestas de 55 países y en ella aparecen Francia, España y América Latina. También demuestra que domina el inglés como lengua de intercambio científico, aun cuando es la primera lengua de menos del 23% de los encuestados.

La cuestión del respeto plurilingüístico aparece regularmente en los debates, las protestas (por ejemplo, llevar una insignia durante el DH2014 en Lausana donde pone las diferentes lenguas habladas por el participante en la conferencia, pasar al español en algunas conferencias para demostrar el multilingüismo, etcétera) y los blogs. Los participantes analizaron el proceso de selección de las contribuciones. En la DH2012 realizada en Hamburgo, 342 expertos participaron en la evaluación: su distribución geográfica muestra que es un evento dominado por colegas norteamericanos y británicos, mientras que los países latinos apenas ocupan un lugar marginal en el proceso de selección de los contribuyentes. Para la DH2014, la llamada a la contribución se hizo en 23 idiomas, prueba de la apertura a la diversidad lingüística. Sin embargo, prácticamente sólo tuvo repercusión en inglés (92%); los no anglófonos temieron sufrir una discriminación, hecho confirmado tras comprobar que el 60% de las propuestas presentadas eran en inglés, frente al 30% presentadas y aprobadas en otros idiomas.

Estas cuestiones de hegemonías culturales también pasan por el sesgo de las herramientas empleadas (UNIX, ASCII, TEI, etcétera) y por el contenido de los trabajos realizados, incluyendo la

digitalización que se hace en gran parte al servicio de las fuentes anglófonas. Para evitar caer en la trampa de estas hegemonías, los autores recomiendan digitalizar igual las huellas digitales de la gente de color y de las mujeres que las de los de los hombres blancos. Responden así a la crítica que se hace a las humanidades digitales, acusadas de no prestar atención a las cuestiones de raza, clase y género, y de interesarse más por temas de investigación que por debates educativos o sociales.

## «Las humanidades digitales crearán nuevas brechas.»

«[El bachillerado en Humanidades digitales]  
¿Por qué reducirlo a un curso en lugar de ampliarlo  
a todos? ¿Cómo vamos a llamar a los bachilleratos  
que no son digitales? ¿Bachillerados analógicos?»

Joël Ronez, ex director de Nouveaux médias de Radio  
France y ex jefe del polo web de ARTE France, 2014,  
[http://www.huffingtonpost.fr/joel-ronez/bac-humanites-  
numerique-1\\_b\\_5943938.html](http://www.huffingtonpost.fr/joel-ronez/bac-humanites-numerique-1_b_5943938.html)

Mientras las herramientas de TI son cada vez más comunes y utilizadas, ya sea sólo para su uso en la Web, enviar mensajes, escribir y editar texto, hacer y organizar fotos, etcétera, los usuarios, como los investigadores, se enfrentan a los desafíos de la apropiación de estas herramientas. Estos desafíos varían de acuerdo con las circunstancias: si han tenido o no experiencia con varias herramientas digitales; si están sujetos a presiones de producción, lo que deja poco tiempo para desarrollar y aprender nuevas rutinas eficaces, frente a tener la oportunidad de trastear y aprender a utilizar la herramienta en sí (por ejemplo, cuando un joven investigador tiene tres años para completar su tesis); o el tener un incentivo (como aquellos jubilados que se apropian de las herramientas de búsqueda y los formatos de intercambio de archivos de datos para hacer su genea-

logía o las herramientas de comunicación para mantenerse en contacto con sus nietos).

Estas dificultades de apropiación explican las fuertes desigualdades frente a lo digital. Pueden aumentar las divisiones sociales, económicas y territoriales, porque la familiaridad con estas herramientas implica recursos relacionales para que alguien te enseñe o para pedir ayuda cuando te aferras a algo. Ahora bien, estos recursos se distribuyen de manera desigual. Para los investigadores de humanidades digitales, estar aislados de grandes proyectos con varias áreas de especialización o de un centro de investigación donde es posible intercambiar fácilmente con informáticos capaz de entender las CHS y proporcionar soluciones (recursos, diseño de nuevas herramientas) cambia radicalmente la situación. Además, si la gran mayoría de los investigadores en CHS ahora usan el ordenador (de escritorio, gestión de base de datos, consulta de la Web, redes sociales digitales), pocos son los que han aprendido a diseñar bases de datos, a crear ontologías, a utilizar herramientas para el procesamiento o la visualización, para editar un sitio web o para hacer la interfaz entre su base de datos y la Web. Aún más raros son aquellos que pueden programar o crear herramientas digitales innovadoras. Los *critical code studies* muestran que, bajo la apariencia de una aparente informatización de todo el mundo, las desigualdades son muy grandes y constituyen divisiones entre los públicos, entre investigadores y entre instituciones de investigación.

Estas divisiones se ven en las reuniones entre los humanistas digitales, especialmente en THATCamp, donde se intercambian consejos sobre herramientas. Algunos se sienten excluidos de los debates demasiado técnicos o coloquiales entre los iniciados. Del mismo modo, en las principales conferencias, los «verdaderos hu-



manistas digitales» son aquellos que programan. Las humanidades digitales crean nuevas brechas que conducen a algunos de sus acólitos a promocionar la formación de las herramientas y de la informática, a esforzarse en sus explicaciones y a ejercer el rol de promotores de esta disciplina. Por otro lado, con el desarrollo de esta disciplina proliferan tanto las herramientas como las especialidades (en texto, imagen, sonido, etcétera). La comunidad que se unificó en general alrededor del texto, la literatura, la lingüística y la historia se está expandiendo y puede ser subdividida en muchas especialidades.

Por otra parte, el uso que hacen los investigadores y las empresas (Google, Facebook, Twitter, Amazon, etcétera) de las huellas digitales de nuestras acciones (correos electrónicos, tuits, localización basada en nuestros teléfonos, el tráfico en la Web, etcétera) sin el consentimiento de las poblaciones afectadas crea un tipo diferente de hendidura que plantea problemas éticos. Para poder analizar grandes cantidades de datos (*big data*, *social physics*) conviene que la gente sea transparente y no se oponga al hecho de que los expertos en *data science* accedan a estos múltiples rastros, incluso si su tratamiento es opaco y dependa de métodos controvertidos, sensibles a la fallas informáticas y al robo masivo de datos. El discurso que acompaña a este movimiento dice que sólo los sospechosos deben temer la transparencia; lo digital contribuye entonces al desarrollo de un seguimiento generalizado de la población. Los humanistas digitales, sin saberlo, contribuyen en el diseño de herramientas de recogida masiva de datos, de análisis automatizado de corpus de texto (paleografía, filología, lingüística), en la intromisión en la vida, las relaciones y las trayectorias de personas y comunidades, en la construcción de perfiles (incluyendo de los autores) y en la visuali-

zación en las redes. Estas herramientas se desarrollan para el progreso de la humanidad pero también sirven a los servicios de seguimiento personalizado del marketing y los servicios de inteligencia, como se hizo evidente con el caso Snowden.

Estas divisiones (las desigualdades digitales entre vigilantes y vigilados, etcétera) conducen a algunos Estados a establecer políticas digitales para promover su acceso, incluso mediante la capacitación, y regular sus usos. Los humanistas digitales proponen entrenar a los estudiantes de una forma distinta a partir de estas nuevas herramientas, a veces ya conocidas por los jóvenes, que tienen que aprender a desarrollarse en el ámbito intelectual, social y tecnológico (elegir las herramientas adecuadas, saber sacar beneficio de ellas, entender lo que está en juego, evaluarlas, reconfigurarlas, crear otras nuevas).

En Francia, en 2014, el informe «Jules Ferry 3.0» abogaba por la creación de un bachillerato general titulado «Humanidades Digitales», para «el individuo creativo de la civilización digital». Se destinaba principalmente a los literatos para cerrar la brecha entre ellos y los matemáticos, iniciándolos en el diseño Web, de videojuegos, de visualización en 3D y experiencias inmersivas, de Internet de las cosas, y en el *big data* y la *data visualisation*. En 2015, el presidente de la República firmó un plan digital con Microsoft orientado a la formación y el apoyo del Plan en las escuelas, en particular los *learning analytics* y el *adaptive learning* (la adaptación de ejercicios basados en los rastros digitales de las acciones realizadas por los estudiantes, rastros que pueden servir a la compañía para diseñar nuevas ofertas); la provisión de soluciones (el Cloud de Microsoft) para los equipos móviles; la enseñanza de programación en las escuelas (la formación del profesorado, plataformas de

videojuegos, una red social interna segura); un apoyo financiero, técnico, operativo y comercial a los actores de la educación, incluyendo fabricantes de teléfonos móviles, los editores de aplicaciones educativas y recursos digitales; una «Carta de confianza» que asegura la protección de los datos personales de los estudiantes y profesores (aunque Microsoft ha facilitado el espionaje con su política de recopilación de información personal sobre los usuarios). Sin embargo, este plan fue criticado debido a la elección de un *software* privado y de formatos cerrados que crean una tutela de la empresa sobre la informática en la escuela y la dependencia de los usuarios vis-à-vis hacia la compañía americana, en perjuicio de las herramientas digitales de código abierto e interoperables solicitadas por las asociaciones de maestros. Se hace caso omiso de la experiencia en las escuelas y en las regiones y de la reflexión política sobre las transformaciones deseadas para la educación nacional. No dice nada de la circulación (incluyendo el derecho a ser olvidado), el intercambio, la propiedad y el análisis de las huellas digitales (eventos de aprendizaje, asiduidad de los alumnos, rendimiento escolar, relaciones entre estudiantes a través de una red social privada, relaciones con los padres). Esas huellas son el nuevo oro negro de la economía digital en beneficio de Microsoft, en lugar de serlo para la industria europea. Lo digital es también un recurso que marca una diferencia y, por lo tanto, conduce a nuevas divisiones en la sociedad.



## Conclusión

La aventura de las humanidades digitales acaba de comenzar. La digitalización del patrimonio y la producción masiva de datos digitales relacionados con nuestras actividades sociales crean depósitos que serán explotados por nuevas herramientas, muchas de las cuales aún no se han inventado. Hay mucho que hacer por parte de los investigadores de las ciencias humanas y sociales e informáticas, así como por las nuevas generaciones de empresarios y usuarios.

El despliegue de la tecnología digital en el ámbito social es tal que pronto ninguna actividad se realizará sin la integración de estas tecnologías, creando una nueva ecología sociotécnica y un humanismo digital. Hoy en día, las humanidades digitales se caracterizan además por ser pioneras en describir las tecnologías de la Web y el procesamiento masivo de datos, su visualización y su inmersión. A veces aventureras, a menudo terminan llenando los cementerios de prototipos y sitios web moribundos, desarrollando gadgets metodológicos que pasan rápidamente de moda o recursos digitales cuyo mantenimiento no está asegurado; las humanidades digitales pueden no ser de largo aliento. No están de moda, ya que se van tan pronto como llegan; lo digital será tan integrado, banalizado y trivializado que ya no tendrá más interés que como otra área específica; el diseño y uso de herramientas podrían llegar a ser tan comunes como lo son ya el *software* de procesamiento de textos y los *workflow* de los editores, los motores de búsqueda y las bases de datos. El trabajo de los investigadores no será el mismo. La in-

vestigación bibliográfica ya ha cambiado con el acceso en línea a distancia, que permite consultar varios registros sin tener que viajar o esperar, y recuperar copias directamente en su ordenador. Lo mismo ocurre con la edición y publicación de textos, y con la comunicación entre los investigadores. Con lo digital, las CHS adquieren otras dimensiones: los sociólogos y los filólogos producen gráficos de las relaciones sociales y semánticas; los historiadores tratan masas de datos dispersos; los geógrafos tienen todo tipo de datos georeferenciados; los lingüistas manejan diversas herramientas de procesamiento del lenguaje. En cuanto a los usuarios, las visitas a museos toman nuevos aires; tenemos acceso a enormes bibliotecas a partir de los equipos móviles; los instrumentos de la enseñanza y el aprendizaje se han rediseñado; los usuarios pueden generar información en forma de indicadores o representaciones visuales de lo que hacen y su relación con estos hechos, etcétera. El mundo está cambiando y la humanidad se vuelve ahora digital.

Para entender lo que son y lo que ocurre con las humanidades digitales habrá que seguir el rastro de lo que hacen los investigadores, los Estados (por ejemplo, sus políticas digitales), las empresas gigantes que viven de los rastros de nuestras actividades, los grupos sociales a los que afectan estas nuevas prácticas. Toca tener en cuenta lo que producen, como objetos, negocios, mercados, costumbres, instituciones, derivas, redes y controversias. No es suficiente una buena introducción a la informática, sino que también se debe ser capaz de decodificar los enlaces, numerosos y complejos, entre el conocimiento producido, los instrumentos y los colectivos de actores que ayudan a formar lo que será nuestra sociedad digital.

Sin embargo, surgen preguntas porque lo digital puede evolucionar de mil maneras, y no todas son iguales. Tienen diferentes

## CONCLUSIÓN

efectos sobre nuestra relación (dependencia, independencia supervisada, libre consentimiento), nuestros proyectos, nuestras culturas y trayectorias, las formas de inclusión y exclusión, las jerarquías y las desigualdades, en la fuerza de los lazos sociales y la evolución de la humanidad. Tenemos ante nosotros varias opciones, algunas de ellas, éticas: ¿qué límites ponemos a la recopilación y uso de datos sobre otros (individuos, comunidades y grupos sociales)? ¿Debemos someternos a la transparencia completa y a la vigilancia mutua generalizada? ¿Qué rol le permitimos a las grandes compañías multinacionales a la hora de guiar nuestro comportamiento, en la gestión de nuestro patrimonio cultural y en la definición de las políticas de nuestros Estados (investigación, educación, salud, seguridad, etcétera)? ¿Qué relación queremos desarrollar con nuestro pasado y con otras comunidades humanas? ¿Cómo regulamos el mal uso (por ejemplo, con fines de espionaje) de las herramientas que desarrollamos pensando en el progreso de la humanidad?

Estas preguntas se refieren a decisiones sobre cómo organizar la producción de conocimiento, la estructuración de los entornos industriales, la formación de las nuevas generaciones, la institución de formas de regulación, la defensa de los intereses de los diversos componentes de la sociedad, incluidos aquellos que puedan no ser útiles u olvidados, o quiénes van a pagar las consecuencias de los acontecimientos cometidos. Las humanidades digitales no son sólo una aventura para los literatos apasionados que se sumergen en una serie de herramientas informáticas; son sobre todo un reto para la sociedad en cuanto a la cultura digital y la nueva humanidad que queremos construir.





## Para saber más

### Libros sobre lo digital, Internet y el acceso abierto

- Beau, F. (2007). *Culture d'univers. Jeux en réseau, mondes virtuels, le nouvel âge de la société numérique*, FYP. Reflexión sobre los avances y desafíos del mundo virtual.
- Beaude, B. (2012). *Internet, changer l'espace, changer la société*, FYP éditions. Análisis de Internet como una cosa real, hecha de actores, cuestiones y relaciones.
- Cardon, D. (2010). *La démocratie internet. Promesses et limites*, Seuil. Síntesis de la problemática política de Internet.
- Casilli, A. (2010). *Les liaisons numériques. Vers une nouvelle sociabilité?*, Seuil. Movilización de herramientas de ciencias sociales para ilustrar las relaciones sociales en la Web.
- Doueihy, M. (2011). *La grande conversion numérique, suivie de Réveries d'un promeneur numérique*, Seuil. Reflexión sobre la manera en que lo digital modifica la vida, la lectura y la escritura, la memoria y la identidad.
- Doueihy, M. (2011). *Pour un humanisme numérique*. Reflexión que propone la creación de un nuevo humanismo para pensar el futuro de las sociedades digitales apoyándose en nuestras tradiciones humanistas.

- \_\_\_\_\_(2013). *Qu'est-ce que le numérique?*, PUF. Intento de definición epistemológica, tecnológica, institucional, económica y política de lo digital.
- Dulong de Rosnay, M. y De Martin, J. C. (2012). *The Digital Public Domain: Foundations for an Open Culture*. Cambridge, Open Book Publishers. Creación de una definición del dominio público digital.
- Mounier, P. (2002). *Les maîtres du Réseau. Les enjeux politiques d'Internet*, La Découverte. Análisis histórico y político de Internet.
- Turkle, S. (2011). *Alone Together*, Basic Books. Análisis de nuevas formas de amistad, de intimidad y de soledad con respecto a las redes sociales digitales.
- Turner, F. (2006). *From Counterculture to Cyberculture*, University of Chicago Press. Análisis de la manera que vino la cultura de la red.

### Libros sobre las humanidades digitales

- Bartscherer, T. y Coover, R. (eds.) (2011). *Switching Codes: Thinking Through Digital Technology in the Humanities and the Arts*, The University of Chicago Press. Análisis de los efectos de la tecnologías digitales en la forma de pensar.
- Berry, D. (2012). *Understanding Digital Humanities*, Palgrave Macmillan. Discusión teórica y practica de diferentes tecnologías de la humanidades digitales.
- Burdick, A., Dricker, J., Lunefeld, P., Presner, T. y Schnapp, J. (2012). *Digital\_humanities*, MIT Press. Exploración de las técnicas y

métodos de las humanidades digitales y su relevancia para la cultura y las ciencias humanas.

Clavert, F., Noiret, S. (2013). *L'Histoire contemporaine à l'ère numérique*, Bruxelles, Peter Lang. Introducción y discusión del uso de la informática en la historia.

Dacos, M. y Mounier, P. (2010). *L'édition électronique*, La Découverte. Análisis de los cambios en el mundo de la edición.

Fiormonte, D., Numerico, T. y Tomasi, F. (2015). *The Digital Humanist: A Critical Inquiry*, Punctum Books, Nueva York. Introducción crítica a las tecnologías digitales de la escritura, del archivo y de la investigación, sensible a los riesgos sociales y a la dominación anglosajona.

Fitzpatrick, K. (2011). *Planned Obsolescence: Publishing, Technology, and the Future of the Academy*, NYU Press. Reflexión sobre el futuro de la investigación académica; obra asociada a un sistema de anotación *Comment Press*.

Gold, M. (dir.) (2012). *Debates in the Digital Humanities*, Univ. of Minnesota Press. <http://dhdebates.gc.cuny.edu/>. Debates teóricos y metodológicos sobre las humanidades digitales, acompañados de una selección de artículos de blogs y una interfaz que permite al lector introducir comentarios e iniciar una discusión en línea gracias a *Comment Press*.

Horst, H. y Miller, D. (2012). *Digital anthropology*, Bloomsbury, Londres. Estudios de caso (Facebook, Second Life, Google Earth) que muestran cómo estudiar antropológicamente lo digital.

Kitchin, R. y Dodge, M. (2011). *Code/Space: Software and Everyday Life*, MIT Press, Cambridge. Análisis que muestra herramien-

tas conceptuales que pueden ayudar a entender las relaciones entre los *software* y la vida cotidiana.

McCarty, W. (2005). *Humanities Computing*, Palgrave Macmillan. Argumento filosófico, histórico y crítico que muestra cómo lo digital ayuda a las ciencias humanas a plantear buenas preguntas.

Moretti, F. (2003). «Gráficos, mapas, árboles. Modelos abstractos para la historia literaria», *New left review*, n.º 20, págs. 83-92; n.º 26, 2004, págs. 47-70; n.º 28, 2004, págs. 43-63 (<http://newleftreview.es/authors/franco-moretti>). Argumento a favor de otra manera de hacer estudios de literatura que contiene y mapea un gran número de obras.

Moretti, F. (2013). *Distant Reading*, Verso, Londres. Colección de artículos que marcaron los estudios de literatura tratándola a nivel mundial y a largo plazo.

Orton-Johnson, K. y Prior, N. (2013). *Digital sociology. Critical perspectives*, Palgrave Macmillan. Serie de contribuciones que dan cuenta de cómo la sociología piensa de una nueva manera sus objetos de estudio en una sociedad digital.

Schreibman, S., Siemens, R. y Unsworth, J. (2008). *A Companion to Digital Humanities*, Wiley-Blackwell. <http://www.digitalhumanities.org/companion/>. Colección de 37 artículos que presentan las humanidades digitales, acompañado de un sitio web donde se encuentra más material.

Terras, M., Nyhan, J., Vanhoutte, E. (eds) (2013). *Defining Digital Humanities. A Reader*, Ashgate. Compilación de textos de la emergencia de las humanidades digitales.

Warwick, C. y Terras, M. (2012). *Digital Humanities in Practice*. Facet, Londres. Presentación de prácticas comunes en la investigación en humanidades digitales y sus desafíos.

## Revistas

*Caracteres: Estudios culturales y digitales de la esfera digital*, <http://revistacaracteres.net/>

*eHumanista / Conversos*, <http://www.ehumanista.ucsb.edu/contact/index.shtml>

*Digital Humanities Quarterly*, <http://digitalhumanities.org/dhq/>

*Digital Studies / Le champ numérique*, [http://www.digitalstudies.org/ojs/index.php/digital\\_studies](http://www.digitalstudies.org/ojs/index.php/digital_studies)

*Digital Studies in the Humanities*, <http://dsh.oxfordjournals.org/>

*Journal on Computing and Cultural Heritage*, <https://jocch.acm.org/>

*Journal of Digital Humanities* (<http://journalofdigitalhumanities.org/>): revista de libre acceso que publica sólo artículos ya publicados y comentados en la Web antes de ser elegidos por un comité de lectores; el proceso se basa en una cadena de selección y agregación automática de contenido *Press Forward* (<http://pressforward.org/>) que saca provecho de más de 500 fuentes de libre acceso y visibles en *Digital Humanities Now* (<http://digitalhumanitiesnow.org/>).

*Journal of the Text encoding initiative*, <http://jtei.revues.org/>

*Literary and linguistic computing*, <http://llc.oxfordjournals.org/>



## Bibliografía y proyectos en la Web

### Recursos digitales generales

*Daria* (infraestructura de investigación para las humanidades digitales) <http://www.dariah.eu/>

*Europeana* (biblioteca digital), <http://www.europeana.eu/>

*Gallica* (recursos digitales de la BNF, Biblioteca Nacional de Francia), <http://gallica.bnf.fr/>

Grupo de Humanidades Digitales en ZOTERO, [https://www.zotero.org/groups/197065/humanidades\\_digitales](https://www.zotero.org/groups/197065/humanidades_digitales) (bibliografía sobre humanidades digitales)

*Huma-num* (infraestructura de investigación con la provisión de herramientas y recursos para las humanidades digitales), <http://www.huma-num.fr/>

*Isidore* (plataforma de acceso a las bases de datos de ciencias humanas y sociales), <http://www.rechercheisidore.fr/>

### Proyectos

*Criminocorpus* (plataforma de edición y archivo hypermedia sobre la historia penal en Francia), <http://criminocorpus.cnrs.fr/>

*Orbis* (representación de las redes geoespaciales de la Roma antigua y cálculo del itinerario al estilo de las herramientas de asistencia al viaje), <http://orbis.stanford.edu/>

*PhotosNormandie* (plataforma en Flickr para compartir fotos de la batalla de Normandía, enriquecida con información del público), <http://www.flickr.com/photos/photosnormandie/>

*Quilt Index* (corpus de 75.000 imágenes de colchas desde la era colonial en los EE. UU., con metadatos que pueden ser enriquecidos por los cibernautas y *software* de reconocimiento de patrones para detectar las afiliaciones). <http://www.quiltindex.org/index.php>

*République des Lettres* (tres proyectos —La Haya, Oxford, Stanford— de visualización de redes intelectuales de toda Europa, a partir de una base de 60.000 cartas desde el siglo xvii hasta el siglo xix, proyectos específicos sobre la correspondencia de Voltaire, D'Alembert y Locke), <http://ckcc.huylgens.knaw.nl/>, <http://www.culturesofknowledge.org/> et <http://republicofletters.stanford.edu/>

*The September 11 Digital Archive* (colección, archivo y la publicación de 40.000 testimonios escritos, 40.000 correos electrónicos, informes de incendios, 15.000 fotos digitales, vídeos y testimonios audiovisuales del 11 de septiembre), <http://911digitalarchive.org/>

*Usines 3D* (reconstrucción 3D de la arquitectura de las fábricas y cadenas de producción de automóviles en Francia a principios del siglo xx y vídeos del montaje de los primeros Citroën), <http://www.usines3d.fr>



## Ejemplos de recursos especializados (texto, imágenes)

*Ancient Lives* (recursos para el estudio de los pergaminos), <http://www.ancientlives.org/>

*Biblissima* (observatorio del patrimonio escrito de la Edad Media y el Renacimiento con la provisión de herramientas), <http://www.biblissima-condorcet.fr/>

*Blake Archive* (archivo del artista William Blake), <http://www.blakearchive.org/blake>

*British National Corpus* (colección de 100 millones de palabras del inglés hablado y escrito), <http://www.natcorp.ox.ac.uk/>

*Corpus Thomisticum* (instrumentos para el estudio de Tomás de Aquino, que incluye el índice *Thomisticus* del padre R. Busa), <http://www.corpusthomisticum.org>

*Cranach Digital Archives* (archivos de la pintura de Lucas Cranach) <http://www.lucascranach.org/digitalarchive.php>

*Frantext* (bases de datos textuales del francés), <http://www.frantext.fr/>

*Hyperprince* (digitalización, codificación TEI y alineación de cuatro traducciones al francés y del texto original en italiano de Maquiavelo con glosario e índice), <http://hyperprince.ens-lyon.fr/>

*Les dossiers de «Bouvard et Pécuchet»* (2.400 páginas de la novela póstuma de Gustave Flaubert), <http://www.dossiers-flaubert.fr/>

*Los diarios de Alexandre Dumas*, <http://alexandredumas.org/>

*Los manuscritos de Stendhal*, <http://www.manuscrits-de-stendhal.org/>

*MONLOE* (obra completa de Montaigne), <http://www.bvh.univ-tours.fr/Montaigne.asp>

*Van Gogh Letters* (digitalización de la correspondencia de Vincent Van Gogh: facsímil, transcripción, traducción al inglés, herramientas de navegación), <http://www.vangoghletters.org/vg/>

## Visualización

### *De datos convencionales*

*Fragile Success* (Nueva York a partir de los datos de empleo, del transporte, del desempeño de las escuelas públicas o del delito). <http://fragile-success.rpa.org/>

### *De sitios web*

*e-Diaspora Atlas* (archivo y análisis de más de 8.000 sitios web o blogs de migrantes y diásporas), <http://www.e-diasporas.fr/>

### *De datos de Facebook*

*Visualizing Facebook Friends*. <http://fbmap.bitasthetics.com>

### *De datos de Flickr*

*A day in Muni, according to NextBus*, <https://www.flickr.com/photos/walkingsf/4521616274>

*Locals and Tourists*, <https://www.flickr.com/photos/walkingsf/sets/72157624209158632/>

*Personnal geography*, <http://www.flickr.com/photos/walkingsf/sets/72157628738161697/>

*Tracing the Visitor's Eye*, <http://www.girardin.org/fabien/tracing/>

### *De datos de Twitter*

*Languages on Twitter*, <https://www.mapbox.com/labs/twitter-gnip/languages/>

*Mobile Devices*, <https://www.mapbox.com/labs/twitter-gnip/brands/>

*Real Time Amsterdam*, <http://realtime.waag.org>

*Real Time Rome*, <http://senseable.mit.edu/realtimerome/>

## Herramientas

*Comment Press* (<http://futureofthebook.org/commentpress/>): herramienta de edición (*plug-in*) desarrollada por el *Institute for the future of the book* a partir de la herramienta de edición de texto en línea WordPress, que permite a los lectores comentar el texto párrafo por párrafo o línea por línea, e iniciar una discusión (a través de BuddyPress y *BuddyPress Groupblog* para crear comunidades de lectores comentaristas) que constituya un sistema de evaluación abierta (*open peer review*).

*DIME-SHS* (infraestructura de datos y métodos de investigación de ciencias humanas y sociales: datos cuantitativos para el es-

tudio longitudinal de un panel por Internet, banco de encuestas cualitativas **BeQuali** para su reanálisis, recopilaciones), <http://www.sciencespo.fr/dime-shs/la-page-d-accueil>

*Mutec* (dispositivo para compartir y diseminar herramientas para la edición crítica y el corpus de humanidades digitales), <http://mutec.huma-num.fr/>

### Instituciones y redes

Asociación HDH (Asociación de Humanidades Digitales Hispánicas) y su lista de comunicación, HDH@LISTSERV.REDIRIS.ES

Rehdi.co (Red colombiana de humanidades digitales), <http://rehdi.co/la-red.html>

RedDH (Red de Humanidades Digitales de la UNAM), <http://www.humanidadesdigitales.net/acerca-de/>

AAHD (Asociación Argentina de Humanidades Digitales), <http://aahd.com.ar/quienes-somos>

ReArte.Dix (Red internacional de estudios digitales sobre la cultura artística), una iniciativa que comenzó en la Universidad de Málaga.

ACH (Association for Computer and Humanities) animando un sitio de preguntas / respuestas), <http://digitalhumanities.org/answers/>

ADHO (Alliance of Digital Humanities Organizations), <http://adho.org/>

- Digital Humanities Now*, <http://digitalhumanitiesnow.org/>
- EADH (European Association for Digital Humanities), <http://eadh.org/>
- HDH (Humanidades Digitales Hispánicas), <http://www.humanidadesdigitales.org/inicio.htm;jsessionid=3EF8CB119C-C56ED0E27D410C92A96754>
- Humanists* (grupo de discusión), <http://dhhumanist.org/>
- Humanistica* (asociación francófona de humanidades digitales), <http://www.humanisti.ca/>
- Infoclio* (portal de las ciencias históricas con el apoyo de la Academia Suiza de Ciencias Humanas y Sociales), <http://infoclio.ch>





