

¡Hola, mundo!

Humanidades digitales en América Latina

The background of the cover is a complex, abstract pattern of diagonal lines and circles. The lines are thin and grey, creating a sense of depth and movement. The circles are solid black and vary in size, some appearing as if they are on the end of the lines, others floating independently. The overall effect is a dynamic, digital aesthetic.

HUMANIDADES DIGITALES



¡Hola, mundo!

Humanidades digitales
en América Latina

{ Maria José Afanador Llach,
Gimena del Rio Riande,
Ernesto Priani Saisó };
Edición académica

Nombre: Afanador Llach, María José, autor, compilador. | Rio Riande, Gimena del, autor, compilador. | Priani Saisó, Ernesto, autor, compilador.

Título: ¡Hola, mundo! Humanidades digitales en América Latina / María José Afanador Llach, Gimena del Rio Riande, Ernesto Priani Saisó.

Descripción: Bogotá: Universidad de los Andes, Facultad de Artes y Humanidades, Ediciones Uniandes, 2025. | 443 páginas: ilustraciones; 14,5 x 19,5 cm. | Colección Humanidades Digitales

Identificadores: ISBN 978-958-798-736-2 (rústica) | 978-958-798-737-9 (e-book) | 978-958-798-786-7 (e-pub)

Materias: Humanidades digitales – América Latina | Humanidades digitales – Enseñanza -- América Latina

Clasificación: CDD 001.30285-dc23

SBUA

Primera edición: marzo del 2025

© María José Afanador Llach,
Gimena del Rio Riande
y Ernesto Priani Saisó,
autores compiladores

© Universidad de los Andes,
Facultad de Artes y Humanidades,
Ediciones Uniandes
Carrera 1.ª n.º 18A-12, bloque Tm
Bogotá, D. C., Colombia
Teléfono: 601 339 4949, ext. 2133
<http://ediciones.uniandes.edu.co>
ediciones@uniandes.edu.co

CORRECCIÓN DE ESTILO:
Camilo Sierra Sepúlveda

DISEÑO DE COLECCIÓN Y DIAGRAMACIÓN:
Boga Visual

IMPRESIÓN:
DGP Editores S. A. S.
Calle 63 n.º 70D-34
Teléfonos: 601 721 7641 - 601 721 7756
Bogotá, D. C., Colombia

ISBN: 978-958-798-736-2

ISBN e-book: 978-958-798-737-9

ISBN epub: 978-958-798-786-7

DOI: <http://doi.org/10.51573/Andes.9789587987362.9789587987379.9789587987867>

Impreso en Colombia – *Printed in Colombia*

Universidad de los Andes |
Vigilada Mineducación.

Reconocimiento como universidad:

Decreto 1297 del 30 de mayo de 1964.

Reconocimiento de personería jurídica:

Resolución 28 del 23 de febrero de 1949,

Minjusticia.

Acreditación institucional de alta calidad,

10 años: Resolución 000194 del 16 de enero

del 2025, Mineducación.

Todos los derechos reservados. Esta publicación no puede ser reproducida ni en su todo ni en sus partes, ni registrada en o transmitida por un sistema de recuperación de información, en ninguna forma ni por ningún medio, sea mecánico, fotoquímico, electrónico, magnético, electro-óptico, por fotocopia o cualquier otro, sin el permiso previo por escrito de la editorial.

CONTENIDO

Introducción

MARIA JOSÉ AFANADOR LLACH

< pág. 10 />

{ PARTE I

Datos e infraestructuras

La gestión del patrimonio bibliográfico en América Latina y su relevancia en el desarrollo de las humanidades digitales: un análisis histórico desde la Biblioteca Nacional de México

ISABEL GALINA RUSSELL

GUILLERMO MORALES

< pág. 26 />

La interoperabilidad de los datos digitales: el uso compartido de información en el proyecto ArqueoData, una base de datos radiocarbono para Latinoamérica

ALBERTO ALCÁNTARA

< pág. 64 />

Código y reflexividad en la investigación histórica
de repositorios digitales: exploración de las ruinas
de GeoCities

NICOLÁS QUIROGA

< pág. 98 />

};

PARTE II

Enseñanza de las humanidades digitales

De la enseñanza de las humanidades digitales
en México, ¿qué hemos aprendido?

ERNESTO PRIANI SAISÓ

MIRIAN PEÑA PIMENTEL

ADRIANA ÁLVAREZ SÁNCHEZ

< pág. 150 />

O ensino de humanidades digitais no Brasil:
visões a partir de um *survey* nacional

JIMMY MEDEIROS

JULIANA MARQUES

SUEMI HIGUCHI

< pág. 188 />

Un puente entre la elección y la necesidad:
reflexiones sobre la *minimal computing* para
las humanidades digitales y la ciencia abierta
en América Latina

GIMENA DEL RIO RIANDE

RAFFAELE VIGLIANTI

< pág. 236 />

};

PARTE III

Formas de las humanidades digitales en la región

Biteratura en español: posibilidades literarias
generativas e interactivas para archivos digitales
dinámicos

SERGIO RODRÍGUEZ GÓMEZ

JUAN ÁLVAREZ UMBARILA

< pág. 268 />

A politização do consumo mediada pela tecnologia:
o caso do *Sleeping Giants* no Brasil

DANIEL COELHO DE OLIVEIRA

< pág. 314 />

*La fiebre del banano: una aproximación
a la curaduría de arte contemporáneo a través
de las humanidades digitales*

JUANITA SOLANO ROA

BLANCA SERRANO ORTIZ

< pág. 354 />

**Movimientos sociales y humanidades digitales:
el estallido del 18 de octubre en Chile**

DAVID BORDONABA-PLOU

< pág. 398 />

} ;

Sobre los autores

< pág. 438 />

Introducción



MARIA JOSÉ AFANADOR LLACH

{ En la última década, múltiples voces latinoamericanas se han escuchado en espacios formales e informales discutiendo e indagando problemáticas que convergen en las humanidades digitales (HD). La formación de redes y asociaciones latinoamericanas desde el 2011, la creación de cursos, diplomados y maestrías, y los encuentros y conferencias han dinamizado el campo. En este proceso, los debates latinoamericanos sobre las HD se han nutrido de las discusiones sobre el multilingüismo, la diversidad geográfica, la hegemonía del código, el colonialismo de datos, las infraestructuras digitales y la cooperación norte-sur, entre otros (Fiormonte *et al.*, 2015; Liu, 2012; Galina, 2014; Rojas Castro, 2020; Spence *et al.*, 2021; del Rio Riande 2022). Estas miradas han contribuido a cuestionar las visiones hegemónicas del campo, que tienden a “constrañir e invisibilizar el trabajo que se realiza en lenguas distintas al inglés y en latitudes distintas a las de Estados Unidos, Canadá y el norte de Europa” (Galina, 2018, p. 10).

El inglés es el idioma dominante en la academia y, por ende, en las HD. A esto se suma que en el campo operan desigualdades globales e injusticias epistémicas en la producción del conocimiento y en las infraestructuras que lo hacen posible (Fiormonte *et al.*, 2022; del Rio Riande *et al.*, 2022). El problema del anglocentrismo en las HD se observa en el hecho de que los recursos, las lecturas, el código, las herramientas digitales y la documentación en español y portugués para la investigación digital son escasos (Isasi y del Rio Riande, 2022). Si bien existe una brecha de información y conocimiento que afecta la forma en que nos involucramos en las HD en la región, se está articulando una comunidad de práctica que impulsa una agenda de investigación unida por lenguas y problemas comunes, pero a su vez heterogéneos¹.

Las redes latinoamericanas han trabajado constantemente en la promoción de diálogos, la producción de conocimientos y la gestión de publicaciones. Por ejemplo, con la creación de la Red de Humanidades Digitales de México, en el 2011, se han publicado más de cien entradas en el blog, lecturas de referencia centrales para quienes quieren conocer sobre el campo². Por su parte, la Red Colombiana de Humanidades Digitales, fundada en el 2016,

- 1 La Red de Humanidades Digitales de México, la Asociación Argentina de Humanidades Digitales y la Red Colombiana de Humanidades Digitales, con el apoyo de la Coordinación de Universidad Abierta, Innovación Educativa y Educación a Distancia, de la Universidad Nacional Autónoma de México (UNAM), han organizado tres ediciones de la SemanaHD, un evento gratuito, en línea, abierto y colaborativo, para que personas y grupos de investigación propongan charlas, talleres, paneles, etc., sobre humanidades digitales. Véase <https://semanahd.cuaieed.unam.mx/>
- 2 Véase <http://humanidadesdigitales.net/blog-rhd/>

gestiona el blog *Cultura y Bits*, el cual ha publicado más de treinta entradas con contenidos multimedia, que tienen como propósito servir como plataforma de divulgación del campo³. Así mismo, la Asociación Argentina de Humanidades Digitales (AAHD) se ha dado a la tarea de publicar con frecuencia los resultados de sus congresos y encuentros, por medio de libros de actas y depósitos de presentaciones y trabajos⁴. La AAHD gestiona además un *podcast*, que incluye entrevistas a expertos sobre HD, y publica su propia revista de investigación gratuita, multilingüe y abierta a la comunidad global de las HD⁵. En el contexto brasileño existe la Associação das Humanidades Digitais (AHDig), la primera asociación, y la Associação Brasileira de Humanidades Digitais (ABHD), que es la actual⁶.

A partir de los esfuerzos orientados a abrir espacios de conversación pública en la web sobre las HD en español, se comenzaron a consolidar iniciativas editoriales de índole científico. En el 2016 se estrenó *The Programming Historian* en español y en el 2020 en portugués, una publicación en línea de tutoriales revisados por pares sobre herramientas digitales, técnicas computacionales y flujos de trabajo útiles para la investigación y la pedagogía en humanidades⁷. En el 2017 inició la *Revista de Humanidades Digitales* (RHD), una propuesta de la Universidad Nacional de Educación a Distancia (UNED, Madrid), el Consejo Nacional de Investigaciones

3 Véase <https://redcolhdblog.wordpress.com/>

4 Véase <https://aahd.net.ar/publicaciones/>

5 Véanse <https://aahd.net.ar/podcast/> y <https://revistas.unlp.edu.ar/publicaahd/>

6 Véase <https://ahdig.hypotheses.org/> y <http://abhd.org.br/>

7 Véase <http://programminghistorian.org/>

Científicas y Técnicas (Conicet, Argentina) y la Universidad Nacional Autónoma de México (UNAM, México), que asume “la defensa del español como vehículo de comunicación científica” (del Rio Riande, 2017, p. 3) y que a la fecha ha publicado nueve números.

En el 2018, la Red de Humanidades Digitales de México comenzó la publicación de la Biblioteca de Humanidades Digitales junto con Bonilla Artiga Editores (Galina *et al.*, 2018). La colección de tres volúmenes reconstruye problemas, posibilidades y trayectorias del campo desde una perspectiva no anglosajona. Se trata de la primera publicación de su longitud y seriedad sobre el tema en español y portugués. Contiene veintiún ensayos que ofrecen una mirada sobre qué son las HD, por qué importan y cómo se ha conformado el campo en América Latina, España y Portugal. En los tres volúmenes, investigadores de diferentes países, mayoritariamente de México, exploran temáticas relacionadas con las instituciones, las epistemologías, las pedagogías y las tecnologías de las HD, desde una perspectiva no anglófona. En el 2021, se publicó un nuevo volumen sobre corpus y literatura en México, continuando así con la producción científica. La más reciente publicación, editada por miembros de la RHD de México, es el libro *Pautas para el desarrollo y la evaluación de proyectos digitales en las humanidades* (2022).

La colección Humanidades Digitales de la Facultad de Artes y Humanidades de la Universidad de los Andes (Colombia) busca contribuir a la comprensión, análisis y divulgación sobre la investigación-creación en HD en América Latina. El llamado para el libro se propuso como un ejercicio para mapear y descubrir a una comunidad de práctica latinoamericana que se reconoce en las HD. Para ello, el proyecto editorial se gestó entre miembros de

la comunidad LatamHD, una iniciativa que comenzó en el 2018 en una reunión de humanistas digitales latinoamericanos durante el congreso DH2018, celebrado en Ciudad de México. La iniciativa se fortaleció en el congreso de la AAHD ese mismo año, en la ciudad de Rosario, Argentina. Como punto de encuentro de las iniciativas comunes, miembros de LatamHD convocaron al Encuentro Latinoamericano sobre Humanidades Digitales, en julio del 2021. Al evento se inscribieron más de doscientas cincuenta personas y se conectaron de manera remota alrededor de ciento veinte personas de Colombia, Argentina, México, Brasil, Perú, Chile, Guatemala, Cuba, Uruguay, Bolivia, Estados Unidos, Canadá, entre otros. En dicho evento se socializó el proyecto editorial, con el interés de difundirlo dentro de una comunidad más amplia de la que se reunió en México y Argentina en las primeras reuniones de LatamHD.

El primer volumen de la colección, *¡Hola, mundo! Humanidades digitales en América Latina*, recoge diez contribuciones de investigadores de Argentina, México, Colombia, Chile y Brasil. Aludiendo a la icónica frase de introducción al aprendizaje de lenguajes de programación, *¡Hola, mundo!* propone un recorrido heterogéneo desde Latinoamérica sobre problemas, métodos y proyectos en las HD. Las formaciones de los autores cubren disciplinas tan diversas como historia del arte, sociología, biblioteología, arqueología, lingüística, filosofía, literatura, entre otras. Cabe mencionar que la convocatoria para recibir capítulos tuvo una gran acogida: se recibieron más de sesenta resúmenes, lo cual refleja un creciente interés de la comunidad académica latinoamericana heterogénea por situarse y reconocerse dentro del campo. En el llamado abierto para recibir contribuciones nos planteamos las siguientes preguntas: ¿qué investigaciones, proyectos y

reflexiones teóricas están emergiendo en la región? ¿De qué maneras los investigadores, instituciones y proyectos están articulando los medios digitales y la computación en sus agendas? ¿Qué enfoques metodológicos y diálogos interdisciplinarios emergen de la práctica? ¿Qué espacios institucionalizados e informales existen en la región sobre el campo y qué están haciendo? De acuerdo con la pregunta que nos hace Domenico Fiormonte (2016), ¿existen unas humanidades digitales no angloamericanas?, y, si es así, ¿cuáles son sus características?, el volumen aporta algunas pistas sobre temáticas, propuestas metodológicas y aproximaciones pedagógicas que investigadores de Latinoamérica están gestando.

Estructura y contenido del libro

La primera parte de esta antología, “Datos en infraestructuras”, contiene tres capítulos que ofrecen perspectivas sobre la interoperabilidad de repositorios digitales, la gestión del patrimonio bibliográfico en las HD y el trabajo con archivos web para la investigación histórica. El primer capítulo, “La gestión del patrimonio bibliográfico en América Latina y su impacto en el desarrollo de las humanidades digitales: un análisis desde la Biblioteca Nacional de México”, propone estudiar la multifacética y completa relación entre las bibliotecas y las HD. A partir de un análisis del papel de estas instituciones como encargadas de reunir, organizar y proveer el acceso al registro cultural como fuente de investigación para los humanistas digitales, el capítulo argumenta que el tipo y calidad de los repositorios digitales creados por las bibliotecas emerge del proceso histórico. A través de un recorrido

por las formas en las que se conformaron las colecciones, el cuidado y atención a los inventarios y la creación de sus catálogos a lo largo de los años, esta investigación establece que la incertidumbre y la falta de control bibliográfico de las colecciones heredadas en la Biblioteca Nacional de México son el resultado de la situación sociopolítica, la cual provocó un rezago en las labores del registro del patrimonio documental, que en algunos casos persiste hasta nuestros días.

En el segundo capítulo, “La interoperabilidad de los datos digitales: el uso compartido de información en el proyecto ArqueoData, una base de datos radiocarbono para Latinoamérica”, plantea la importancia de la interoperabilidad de bases de datos para las investigaciones históricas latinoamericanas. Partiendo de los objetivos de la libre consulta de la información para las comunidades académicas, este trabajo analiza el papel de las bases de datos para organizar, compartir y estandarizar datos de investigación. Se concentra en el caso de la base de datos arqueológicos ArqueoData, para mostrar su funcionalidad y las maneras como usa la interoperabilidad, permitiendo la colaboración y la reutilización de la información por parte de la comunidad de investigación. El último capítulo de la sección se titula “Código y reflexividad en la investigación histórica de repositorios digitales: exploración de las ruinas de GeoCities”; este se ocupa de un repositorio masivo del sitio web GeoCities (1994-2009) como posible fuente para hacer historia. Con base en el análisis exploratorio de los datos de los sitios, la investigación examina técnicas estadísticas y visualizaciones siguiendo recomendaciones de trabajo con páginas web como fuentes históricas, recientemente exploradas por abanderados de la historia digital en el mundo angloparlante.

La segunda parte, “Enseñanza de las HD”, está compuesta por tres capítulos que indagan sobre los espacios e iniciativas de pedagogía con medios digitales y computación en disciplinas de las humanidades. El capítulo “Enseñar humanidades digitales en México: ¿qué hemos aprendido?” presenta una reflexión sobre las experiencias pedagógicas en HD desarrolladas en la Facultad de Filosofía y Letras de la UNAM. Si bien la experiencia narrada da cuenta de una apertura multidisciplinar, en la práctica ha tenido resistencia en relación con la incorporación de tecnologías digitales en el currículo de disciplinas tradicionales de las humanidades. Este trabajo muestra tres estrategias de incorporación de contenidos de HD a nivel de pregrado y un curso abierto masivo en línea (*massive online open course*, MOOC por su sigla en inglés) para todo tipo de públicos, lo que revela la fluidez en los modelos de enseñanza de las HD y la importancia de adaptar constantemente sus contenidos. El caso presentado en el capítulo “O ensino de humanidades digitais no Brasil: visões a partir de um *survey* nacional” pone en cuestión tensiones similares a las presentadas en el caso mexicano, pero en el contexto brasileño y a partir del problema de la integración entre investigación y docencia en la formación en competencias digitales en las humanidades. Los autores muestran que, aun cuando las prácticas de investigación en humanidades y ciencias sociales con métodos computacionales han aumentado desde el 2010 en Brasil, estas no se han reflejado en cambios curriculares sustanciales en la docencia. La formación en temas de HD ha ocurrido sobre todo a partir de iniciativas individuales y no necesariamente institucionales. En esta realidad, anotan los autores, los estudiantes buscan contenidos digitales en la web que suplan sus necesidades de alfabetización digital. Así, se muestra que

existen más grupos de estudio y laboratorios de HD que contenidos de corte digital en los cursos de humanidades en Brasil.

El último capítulo de la sección, “Un puente entre la elección y la necesidad: reflexiones sobre la *minimal computing* para las humanidades digitales y la ciencia abierta en América Latina y el Caribe”, contribuye al debate de la enseñanza al analizar las oportunidades que la computación mínima (*minimal computing*) aporta a la investigación y pedagogía en HD en el contexto global. Este caso de estudio es una investigación y el desarrollo de un curso bilingüe entre dos universidades, una en Argentina y otra en Estados Unidos. Al ser las infraestructuras informáticas elementos centrales en la investigación y la pedagogía, el capítulo muestra cómo el trabajo con línea de comandos, sitios web estáticos e instalación de dependencias incentiva el trabajo autónomo sin depender de plataformas ajenas o servidores. El enfoque de la *minimal computing*, señalan los autores, tiene el potencial de fortalecer los vínculos entre la comunidad académica de humanistas digitales y los movimientos del *software* y *hardware* abierto, la ciencia abierta y los movimientos de acceso abierto.

La última sección, “Formas de las humanidades digitales en la región”, se compone de cuatro capítulos sobre proyectos e investigaciones que articulan los medios digitales y la computación con preguntas relacionadas con literatura, archivos digitales, curaduría de arte, activismo en redes y movimientos sociales. Cada trabajo propone mover fronteras disciplinares para ampliar las preguntas, posibilidades metodológicas y problemas de investigación que emergen en las HD latinoamericanas. El capítulo “Biteratura en español: posibilidades literarias generativas e interactivas para archivos digitales dinámicos” hila la posibilidad creativa

de la *biteratura* con el desarrollo de una pieza de *software* para movilizar y dinamizar archivos digitales culturales, con el fin de potenciar mecanismos de descubrimiento para los usuarios. El siguiente capítulo, “La fiebre del banano: una aproximación a la curaduría de arte contemporáneo a través de las humanidades digitales”, presenta un análisis del proyecto digital *La fiebre del banano/Banana Craze*. Se trata de un proyecto colaborativo y pedagógico compuesto por una exposición virtual, un archivo digital y un módulo de investigación, que propone formas posibles de lectura no lineal de obras de arte sobre el banano en América Latina. Las prestaciones del entorno digital son exploradas para mostrar cómo en la exhibición digital se aumentan las oportunidades de interacción, la hipertextualidad y las formas de filtrado, para establecer relaciones difícilmente logrables en un espacio físico.

Los últimos dos capítulos del volumen se ocupan del tema del activismo en redes y de las posibilidades metodológicas digitales para el estudio de los movimientos sociales. El ecosistema digital de circulación de información falsa y las prácticas de consumo en plataformas de redes sociales es la preocupación de “A politização do consumo mediada pela tecnologia: o caso do Sleeping Giants no Brasil”. El texto plantea un cuestionamiento de los límites y posibilidades del activismo de redes para el actuar dentro de la esfera de acción pública en el proceso de polarización política de Brasil. De esta manera, delinea un ejercicio de análisis discursivo en X (antes Twitter), presentando un modelo para comprender las formas del activismo en redes en América Latina. Por último, el capítulo “Movimientos sociales y humanidades digitales: el estallido del 18 de octubre en Chile” expone un caso de estudio de los debates parlamentarios que tuvieron lugar después del estallido

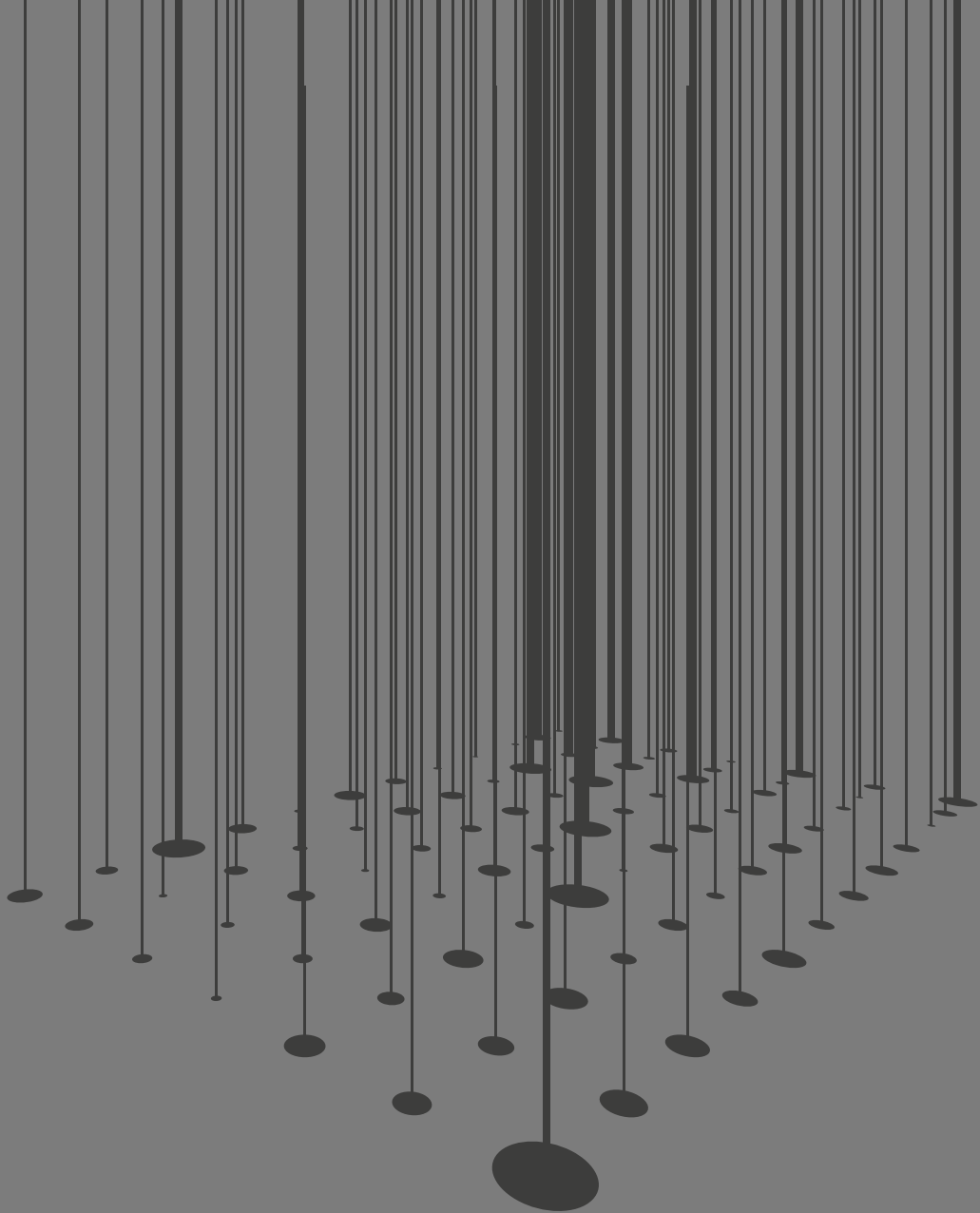
social en Chile. Tras aplicar metodologías de corpus lingüísticos, el proyecto compiló y editó estos utilizando RStudio y los paquetes XML y RCurl.



Referencias

- del Rio Riande, G. (2017). La Revista de Humanidades Digitales: Año 1, volumen 1. *Revista de Humanidades Digitales*, 1, 1-3.
- del Rio Riande, G. (2022). Digital humanities and visible and invisible infrastructures. En D. Fiormonte, S. Chaudhuri y P. Ricaurte (Eds.), *Global debates in the digital humanities* (pp. 247-258). Minnesota University Press.
- del Rio Riande, G. y Fiormonte, D. (2022). Una vez más sobre los sures de las humanidades digitales. *Acervo: Revista do Arquivo Nacional*. <https://www.aacademica.org/gimena.delrio.riande/180.pdf>
- Fiormonte, D., Chaudhuri, S. y Ricaurte, P. (2022). *Global debates in the digital humanities*. University of Minnesota Press. <https://public.ebookcentral.proquest.com/choice/PublicFullRecord.aspx?p=6921430>
- Fiormonte, D., Tomasi, F. y Numérico, T. (2015). *The digital humanist: A critical inquiry*. Punctum Books.
- Galina Russel, I. (2014). Geographical and linguistic diversity in the digital humanities. *Literary and Linguistic Computing*, 29(3), 307-316.

- Galina Russel, I. (2018). Introducción. En *Humanidades digitales: recepción, institucionalización y crítica*. Biblioteca de Humanidades Digitales, Red de Humanidades Digitales, Bonilla Editores.
- Isasi, J. y del Rio Riande, G. (2022). ¿En qué lengua citamos cuando escribimos sobre humanidades digitales? *Revista de Humanidades Digitales*, (7), 127-143.
- Liu, A. (2012). What is cultural criticism in the digital humanities? En M. K. Gold y L. F. Klein (Ed.), *Debates in the digital humanities 2012* (pp. 490-509). University of Minnesota Press.
- Rojas Castro, A. (2020). *FAIR enough? Building dh resources in an unequal world* [conferencia]. Digital Humanities Kolloquium, Berlín. <https://hcommons.org/deposits/item/hc:32187/>.
- Spence, P. y Faria Brandao, R. (2021). *Disrupting digital monolingualism: A report on multilingualism in digital theory and practice*. Language Acts and Worldmaking Project. <https://languageacts.org/digital-mediations/digital-mediations-publications/disrupting-digital-monolingualism-a-report-on-multilingualism-in-digital-theory-and-practice/>



PARTE I

{ Datos e infraestructuras

La gestión del patrimonio bibliográfico en América Latina
y su relevancia en el desarrollo de las humanidades
digitales: un análisis histórico desde la Biblioteca Nacional
de México

ISABEL GALINA RUSSELL

GUILLERMO MORALES

La interoperabilidad de los datos digitales: el uso
compartido de información en el proyecto ArqueoData,
una base de datos radiocarbono para Latinoamérica

ALBERTO ALCÁNTARA

Código y reflexividad en la investigación histórica de
repositorios digitales: exploración de las ruinas de
GeoCities

NICOLÁS QUIROGA

};

**La gestión
del patrimonio
bibliográfico y su
relevancia en el
desarrollo de las
humanidades digitales:
un análisis histórico
desde la Biblioteca
Nacional de México**

Para citar este capítulo:

<http://doi.org/10.51573/Andes.9789587987362.9789587987379.9789587987867.1>

ISABEL GALINA RUSSELL

GUILLERMO MORALES

Introducción

{ En su ponencia “Archivos y humanidades digitales en América Latina”, para el 5.º Encuentro de Humanistas Digitales de la RedHD, María José Afanador (2021) preguntó qué similitudes y diferencias tienen los repositorios digitales gestionados desde América Latina y el norte global. Dicha pregunta se desprende de una búsqueda por identificar las características particulares que las humanidades digitales (HD) presentan en otras áreas del mundo, en especial en el sur global, pues las diferencias existentes entre los repositorios digitales disponibles para los investigadores condicionan, en gran medida, las posibilidades y los alcances del trabajo que puede realizarse.

En el quehacer de las HD en América Latina encontramos que bibliotecas, museos y archivos —y los respectivos proyectos de digitalización de sus colecciones— han sido espacios pioneros en la divulgación y experimentación con materiales digitales, por

ejemplo, en Colombia existe una importante relación entre instituciones de memoria y la gestión del patrimonio cultural (Afanador *et al.* 2020); asimismo, se identifican algunos vínculos de las HD con el movimiento de acceso abierto (del Rio Riande, 2018). En estas primeras aportaciones consideramos que las bibliotecas, vinculadas a las HD por tradición, deben estudiarse con mayor detalle. Así, y retomando la pregunta planteada por Afanador, es factible creer que por lo menos una parte de la respuesta podría encontrarse analizando su situación histórica en América Latina.

A lo largo de las décadas, en diversos países, las bibliotecas han actuado como colaboradores importantes en el desarrollo de las HD, ya sea al ofrecer espacios para la gestión de proyectos, proporcionar los recintos para la capacitación y la enseñanza, suministrar infraestructura o bien como custodias y facilitadoras de fuentes de información digital en la investigación.

En ese escenario, existe abundante literatura disponible sobre los distintos aspectos que las relacionan con las HD, al menos desde el norte global (Kear y Joranson, 2018; Millson-Martula y Gunn, 2020; Posner, 2013), aunque también encontramos estudios sobre su desvinculación, como en el caso de España (Morales del Castillo *et al.*, 2020).

En una investigación realizada por Silvia Gutiérrez (2020), se analizaron los resultados de una encuesta a diversas bibliotecas latinoamericanas efectuada entre el 2017 y el 2018, cuyo objetivo fue identificar y visibilizar la relación entre el personal bibliotecario y su participación activa en proyectos de HD. Gutiérrez compara los resultados de su encuesta con otra realizada a bibliotecas estadounidenses para contrastar similitudes y diferencias. Entre algunos hallazgos destacados, la autora identifica la falta de

infraestructura, la precariedad y la sobrecarga laboral como una percepción común en ambas regiones que trunca el ejercicio profesional, desmitificando la idea de que solo el avance tecnológico es apremiante en la diferencia de los proyectos con el norte global. Asimismo, resalta la perspectiva sociopolítica e histórica en el papel de las bibliotecas y sus colecciones en Latinoamérica, a partir de opiniones comunes de los encuestados sobre la relevancia de comprender al usuario de los proyectos de HD en instituciones de la memoria, sus necesidades como sujeto social e histórico, e incluso entender más sobre los cambios a nivel estatal que han favorecido el desarrollo de las HD en diferentes países (Gutiérrez, 2020), por lo que, para nosotros, son de especial interés los archivos y repositorios digitales de materiales patrimoniales.

La relación entre las bibliotecas y las HD es multifacética y compleja; por esta razón, en este capítulo nuestro interés se centrará en el papel de dichas instituciones como encargadas de reunir, organizar y proveer el acceso al registro cultural como fuente de investigación para los humanistas digitales. Consideramos que el tipo y calidad de los repositorios digitales que elaboran tiene una correspondencia directa con el proceso histórico y depende de cómo fueron conformadas las colecciones, el cuidado y atención a los inventarios y la creación de sus catálogos a través de los años, así como la eficiencia con la que fueron automatizadas en su momento, los sistemas de catalogación y clasificación utilizados, el manejo de los proyectos de digitalización y, por último, la visión en la creación de las plataformas digitales, incluida la oferta de herramientas digitales para la investigación. De igual manera, pensamos que el repositorio digital no es solo un producto de la actividad reciente en una biblioteca, sino que sus características

dependen, en gran medida, de una larga trayectoria y de las decisiones y acciones que se tomaron décadas y hasta siglos atrás; en este aspecto se distinguen de las bibliotecas en países del norte global, donde los fondos documentales descansan en tradiciones y trayectorias disímiles.

Es posible encontrar un ejemplo de esto en un estudio del uso de minería de datos para investigación histórica en el Reino Unido y Colombia, realizado por Crymble y Afanador Llach (2021), en el cual se realiza una importante comparativa, entre ambos países, de sus respectivas “costumbres” al elaborar y preservar registros. Los autores destacan cómo la tradición imperialista británica del siglo XIX engendró una cultura de documentar, coleccionar y organizar archivos. Estos documentos registrados —resultado de dicha tradición— forman la base de los acervos digitales que los investigadores tienen hoy a su alcance. En contraste, el archivo originado por la administración colonial en Colombia no servía a los propósitos ni a los intereses de los colombianos; además, gran parte de este se encuentra ahora en el Archivo General de Indias en Sevilla, España, lo que ocasiona que Colombia tenga nula injerencia sobre la digitalización y la disponibilidad de los materiales para realizar estudios con minería de datos sobre su propio pasado.

Por lo tanto, para identificar de qué forma las características de las bibliotecas latinoamericanas y la conformación de sus colecciones influyen en la generación de repositorios digitales empleados por humanistas digitales, proponemos acudir al estudio histórico de sus acervos, pues representan el nicho en el que surgen la gran mayoría de los proyectos de digitalización de fondos patrimoniales.

En gran parte de América Latina, las estructuras coloniales crearon escenarios con elementos particulares, los cuales impactaron

en la formación de las instituciones y símbolos que durante el siglo XIX dieron forma a las nacientes identidades nacionales. A su vez, las complicadas realidades sociopolíticas de los periodos posteriores a las independencias en cada territorio significaron un verdadero reto no solo para establecer los nuevos órdenes de Estado y sus instituciones —incluidas las bibliotecas y los archivos—, sino también para mantener la continuidad y estabilidad de los proyectos culturales durante los años inmediatos a su puesta en marcha. En la práctica, se puede afirmar que frecuentemente la administración de las bibliotecas y los archivos quedó en manos de personas cuyos intereses políticos no permitieron la continuidad de los proyectos planteados por sus antecesores, o bien su administración se vio atropellada por decisiones arbitrarias o improvisadas sobre la organización, descripción y acomodo de las colecciones, la asignación de inmuebles inadecuados para alojarlas y dar servicio, la precariedad de los presupuestos o su mala administración, el rezago en los inventarios y la catalogación, y el robo frecuente o desaparición de materiales de los acervos que se heredaron; todo en conjunto creó el particular escenario del patrimonio bibliográfico en América Latina.

Nuestra hipótesis es que la incertidumbre y falta de control bibliográfico de las colecciones heredadas, resultado de la situación sociopolítica de la época, provocó un rezago en las subsecuentes labores del registro del patrimonio documental que se ha arrastrado, incluso en algunos casos, hasta nuestros días. Por ejemplo, en un estudio de la Fundación Histórica Tavera (2000) sobre el estado de archivos con patrimonio documental en Latinoamérica se destaca la situación de extrema vulnerabilidad en la que se hallan, pues necesitan especialistas, adolecen de escaso apoyo de

los sectores público y privado, y no pueden adaptarse al nuevo proceso informático. Esas dificultades iniciales se traducen en una posición desventajosa para iniciar proyectos de digitalización: se limitan las opciones, se elevan los costos o se dificulta la generación de herramientas digitales de mayor sofisticación. Crymble y Afanador Llach (2021) también destacan la falta de presupuesto, una débil aplicación de leyes archivísticas y una falta de catálogos en los archivos regionales colombianos. En un caso de estudio sobre proyectos de digitalización mexicanos, encontramos que la ausencia de una infraestructura mayor, tanto tecnológica como administrativa, es una de las grandes debilidades para la visibilidad y el alcance de estos proyectos en dicho país (Galina, 2018). Suponemos también que el papel sociopolítico de las bibliotecas en la creación de estos nuevos Estados formados con los procesos de independencia ha trastocado, a lo largo de los años, la percepción social sobre el patrimonio cultural y bibliográfico, y lo que este representa. Sin duda, es un planteamiento ambicioso que implica toda una nueva área de investigación, por lo que en este capítulo nos hemos ceñido, a modo de ejercicio, a un caso de estudio en México, en particular la Biblioteca Nacional, pues se explica a través de dos procesos trascendentales: la nacionalización de las antiguas bibliotecas conventuales y la formación, con sus acervos, de una biblioteca nacional y pública (Nava, 2017), que presume además dos importantes repositorios digitales nacionales, la Hemeroteca Nacional Digital de México (HNDM) y la Biblioteca Nacional Digital de México (BNDM). No obstante, entendemos las limitantes existentes para ahondar de manera exhaustiva en nuestro análisis, pero, del mismo modo, lo consideramos un primer ejercicio para conectar la historia de las bibliotecas en América

Latina con los proyectos de digitalización y, por ende, con el desarrollo de acervos digitales para las HD.

Breve historia de la Biblioteca Nacional de México

Para explicar de una manera más clara el trasfondo del estado del patrimonio documental que resguarda la Biblioteca Nacional de México (BNM), creemos prudente realizar un breve repaso de su historia, con el fin de ilustrar su complejo proceso de fundación como institución de Estado, el origen de sus colecciones, las dificultades de los controles bibliográficos y, por consiguiente, el punto de partida de los posteriores proyectos de digitalización de su actual acervo.

Como otros territorios americanos bajo el gobierno de la monarquía española, México obtuvo su independencia tras diversos conflictos armados que culminaron hacia 1821. De esa manera, iniciaba el camino para colocar los cimientos de una república, el cual no pudo escapar de complejos fenómenos sociopolíticos empapados de crisis internas por constantes luchas de poder entre las clases gobernantes y sus opositores.

Según palabras de Jaime Rodríguez (1986), el México colonial era un amplio territorio que gozaba de cierta estabilidad, pero que, para mediados del siglo XIX, no solo había perdido más de la mitad de su territorio, sino que se hallaba sumergido en una situación de extrema inestabilidad económica y política. Como ejemplo de este contraste, se puede mencionar que tan solo en sus primeros años como nación independiente se rigió bajo tres

constituciones, veinte gobiernos y más de cien gabinetes. Esta fue la atmósfera en medio de la cual un grupo de élite política e intelectual manifestó su interés de formar bibliotecas públicas que hicieran un servicio a la sociedad, pues una de las ideas primordiales del pensamiento liberal de dicho siglo fue la secularización y divulgación de la cultura. Existía la firme creencia de que la educación representaba un elemento clave en el desarrollo social y abriría la puerta a la oportunidad de formar mejores ciudadanos a través de la lectura.

En México, la idea de una biblioteca pública nacional existió desde mucho antes del anuncio oficial de su creación (Nava, 2017), pero es hasta 1828 cuando se presentó a la Cámara de Diputados una petición para que, por un lado, el Gobierno extinguiera colegios de religiosos en la capital y fusionara sus acervos y, por el otro, designara un espacio para la formación de una biblioteca nacional en el Palacio, con la decidida intención de formar una nueva sociedad liberal, culta e ilustrada a través de la secularización. No obstante, la propuesta fue estudiada y rechazada por la Comisión de Instrucción Pública (Vázquez *et al.*, 2007).

Más tarde, en 1833 apareció en *El Telégrafo*¹ un artículo sin firma con el encabezado “Biblioteca Nacional”, donde se planteaba la idea de un recinto en el que se hallaran “todas las obras, opúsculos,

1 El 10 de enero de 1833 dejó de editarse el *Registro Oficial*, publicación que daba a conocer las disposiciones del Gobierno desde enero de 1830. Al día siguiente, inició la publicación de *El Telégrafo*, órgano oficial de la nueva administración, el cual imprimió solo un año, hasta diciembre de 1834. Con las constantes luchas de poder, la prensa desempeñó un papel esencial en la comunicación oficial y de oposición; a su vez, esto explica el vasto universo de títulos conservados en archivos y hemerotecas nacionales.

impresos y colecciones de periódicos, mapas y planos cuyos autores fueran mexicanos y que en lo sucesivo los impresores del Distrito y Territorios estuviesen obligados a remitir un ejemplar de todo cuanto publicasen” (Spell, 1959, p. 452), así como colecciones de leyes, decretos, memorias y periódicos oficiales donados por el Gobierno. Este texto dejó plasmadas nuevamente las intenciones acumuladas años atrás, con la diferencia de que no se señaló la compra de libros europeos ni la expropiación de acervos conventuales, sino el abastecimiento a partir de obras de autores mexicanos e impresos nacionales, apelando además a la buena voluntad de las donaciones, muy probablemente debido a la falta de recursos por parte del Estado que había hecho imposible ejecutar las propuestas de años anteriores.

Más adelante, se aprobó la organización de un sistema de educación pública, que clausuró la Real Universidad y autorizó el establecimiento de una biblioteca nacional pública bajo la dirección de Manuel Eduardo de Gorostiza, quien redactó el reglamento que señala la organización y administración que debería seguir el recinto, incluidos los detalles sobre las funciones del personal, el horario de atención y la elaboración de numerosos índices de toda la colección (Spell, 1959).

Como podía esperarse, la empresa no resultó sencilla: el edificio tenía que ser acondicionado para ofrecer al público salas de lectura y el reglamento exigía la catalogación detallada y la elaboración de un inventario totalmente nuevo para uso de los empleados, sin mencionar la importante tarea de trasladar los documentos y poner en cada volumen el sello de la biblioteca. Sin embargo, el presidente Antonio López de Santa Anna abolió todos los cambios realizados y ordenó que todo regresara a su estado anterior en

un plazo de treinta días; así, echaba atrás el proyecto, aunque la idea de establecer una gran biblioteca nacional y pública quedó latente. Después, en 1846, se expidió otro decreto en el que se creaba por segunda ocasión la biblioteca nacional (Nava, 2017), cuando todos los libros habían sido trasladados de un edificio a otro, debido a las órdenes expedidas con anterioridad. En dicha ocasión, se propuso como fondo inicial de la biblioteca los libros del extinto Colegio de Todos los Santos, así como los duplicados de la Biblioteca Turriana, de la Real Universidad, y de las bibliotecas de comunidades religiosas —lo que no representaba necesariamente el ideal liberal de biblioteca “moderna”, tal como lo perseguían sus antecesores—, pero la inestable situación del Gobierno y la guerra con Estados Unidos hicieron fracasar el proyecto.

Uno de los eventos más trascendentales para la conformación de colecciones bibliográficas patrimoniales se dio en julio de 1859, cuando por ley se extinguieron los claustros y conventos, se declararon nacionales todos sus bienes (incluidas las bibliotecas) y se ordenó que “los libros impresos, manuscritos, pinturas, antigüedades y demás objetos pertenecientes a las comunidades religiosas suprimidas, se aplicarán a los museos, liceos, bibliotecas y otros establecimientos públicos” (Ley de 12 julio de 1859, art. 12), con lo que se sumarían otros noventa mil libros de aquellos acervos declarados propiedad de la nación (Spell, 1959). A partir del establecimiento del nuevo Gobierno liberal en 1861 y el anuncio explícito de la ley, los religiosos de la Ciudad de México abandonaron sus conventos voluntariamente en medio de un hueco de autoridad y dejaron atrás sus libros, aún sin un espacio designado o plan establecido para su reacomodo y resguardo. A manera de ejemplo,

en ese mismo año se sacaron del Convento de San Francisco alrededor de diecisiete mil libros y se trasladaron al edificio de la universidad. Sobre este suceso, algunas crónicas de la época relatan cómo las carretas de carga, jaladas por mulas, iban tirando libros a su paso, mientras la gente los tomaba para revenderlos a bajo precio, obras cuyo valor alcanzaba mejores cifras si se ofrecían a coleccionistas más instruidos en la bibliofilia de la época (Ramírez, 2020); así, fue un año en el que propios y extraños surtieron sus bibliotecas particulares.

No es sino hasta 1867, después de más confrontaciones por el poder, las cuales incluyeron un breve periodo de un gobierno de corte imperial, traslados y éxodos de colecciones, que por fin se designa a la antigua Iglesia de San Agustín como sede de la BNM y a José María Benítez como su encargado. Al acervo expropiado de conventos y resguardado en el recinto universitario se le sumarían los libros de la catedral y de los jesuitas; de acuerdo con Iguíniz (1940), el acervo recaudado de los fondos primitivos sumaba 116 631 volúmenes, de los cuales 10 652 se extraviaron en el traslado desde las bibliotecas conventuales y solo 1642 se vendieron o destinaron a otras instituciones. De ese modo, la BNM nació como una respuesta a la monopolización de la educación y la cultura por parte del clero mexicano (Vázquez *et al.*, 2007), pero con un fuerte *aliento* de su pasado colonial y religioso en paredes y estantes.

No obstante, debido a la situación política del país y la falta de dinero, los libros trasladados al Convento de San Agustín de la Ciudad de México quedaron hacinados entre el agua y el polvo, las adecuaciones necesarias al recinto para colocarlos, conservarlos

y consultarlos tardaron en llegar y el público tuvo que esperar de nuevo por su apertura, pues la BNM se inauguró oficialmente hasta 1884, medio siglo después de su primera concepción.

Entre tanto, se fueron añadiendo otras grandes colecciones, tanto de algunos particulares ilustrados que donaron sus libros como de otros millares de volúmenes de instituciones religiosas suprimidas en 1917, que perturbaron la relativa calma con la que había estado operando la institución —incluso en medio del periodo de la Revolución, nuevo conflicto armado que estalló en el país en 1910—, pues no se tenía el espacio suficiente para agregarlas ni dinero para catalogarlas. Debido al gran crecimiento de la biblioteca, la estructura del edificio se vio amenazada y hubo que encajonar muchos libros y manuscritos durante años, mientras se desataba una controversia sobre su destino y se avanzaba lentamente con los trabajos de catalogación, tras el escaso personal preparado y la dificultad de conservarlo por los bajos salarios (Spell, 1959).

Por mandato del Gobierno federal, en 1929 la BNM se integró a la Universidad Nacional Autónoma de México (UNAM). Al crearse la Ciudad Universitaria en la década de los cincuenta, se consideró su traslado junto con el de la Hemeroteca Nacional de México (HMN), instaurada en el ex Templo de San Pedro y San Pablo en 1944, pero se desestimó la idea. Posteriormente, y después de varios intentos por mantener el edificio en pie, a finales de los años setenta se construyeron unas nuevas instalaciones en la zona cultural de la Ciudad Universitaria para albergar a la BNM, la HMN y el Instituto de Investigaciones Bibliográficas (IIB). En 1979 se ocuparon los dos edificios principales; en 1988 fue entregado el Almacén de Periódicos y Revistas, para

concentrar el material duplicado de la HMN, y en 1992 se inauguró el edificio anexo del Fondo Reservado, donde se resguarda gran parte de nuestro patrimonio documental (BNM, s. f.). Al estar bajo custodia de la UNAM, se convierte en una de las pocas bibliotecas nacionales integrada dentro de una universidad nacional, así como en el recipiente del patrimonio legado. Fue como dependencia de la UNAM que se llevaron a cabo los procesos de modernización de los servicios de la BNM en los años cruciales hacia la automatización, los cuales darían pie a los proyectos de digitalización. Debemos imaginar —tras todo este escenario de avatares— colecciones heredadas en malas condiciones, cuyo estado de conservación dificultará su identificación y posterior registro, así como su proceso de digitalización décadas más tarde.

La historia de la clasificación y catalogación de fondos patrimoniales en México es una cuestión que requiere de una revisión mucho más profunda y aún no existe un único trabajo que recoja de manera retrospectiva este tema tan específico, pues, como mencionamos, el desarrollo del registro de los bienes patrimoniales de los catálogos impacta de forma importante en los proyectos de HD. Sin embargo, investigaciones como la de Martínez (2002) recogen datos interesantes que vale la pena señalar y consideramos relevantes aquí. Dicho texto revela una fuerte influencia de la escuela norteamericana en la bibliotecología en México, en particular por la adopción de las reglas angloamericanas de catalogación en muchas bibliotecas en nuestro país. Asimismo, se relatan las dificultades que presentaron los trabajos pioneros de catalogación moderna, en especial por las barreras del idioma —los materiales no estaban traducidos al español— y por la falta de personal capacitado para llevar a cabo estas tareas, pues simplemente no

había suficientes bibliotecarios. Por fortuna, estos problemas paulatinamente se solucionaron gracias a la traducción de los textos guías y otros materiales de apoyo, al igual que por la profesionalización de la disciplina en las escuelas mexicanas y la inclusión de sus egresados en las bibliotecas. No obstante, en nuestra opinión, es probable que la difícil tarea de identificación del patrimonio bibliográfico, la múltiple dispersión de las colecciones y la falta de estandarización en los registros, por mencionar algunos aspectos, sean el motivo de la gran dificultad que el país ha enfrentado para la creación de catálogos colectivos, que hasta la fecha continúa siendo un reto por superar.

En la segunda mitad del siglo xx encontramos notables avances en la bibliotecología en México; se destaca el trabajo de Gloria Escamilla, suscrita al IIB, quien realizó una estancia de trabajo en la Biblioteca del Congreso en Estados Unidos e incorporó mucho de lo aprendido a la BNM. Sus manuales *Interpretación catalográfica de los libros* (1979) y *Manual de catalogación descriptiva* (1981) fueron textos clave en español sobre las reglas de catalogación; sin embargo, como es natural entender, dichos textos estaban mayormente enfocados en la catalogación de materiales modernos. Para el caso de fondos patrimoniales, en particular libros impresos antiguos, trabajos como el de García Aguilar (2008) describen los retos que ha presentado la catalogación y la descripción de esta parte del legado bibliográfico mexicano.

La UNAM también figura como una organización relevante en el desarrollo de la bibliotecología en el país. Cuenta con más de cien bibliotecas y su desarrollo de sistemas de automatización, al igual que catálogos y otros servicios de información, que han sido pioneros y de gran envergadura. La BNM no pertenece al sistema

bibliotecario de la UNAM; su catálogo no está integrado con el universitario, pues se ha desarrollado independiente, aunque ambos utilizan el mismo *software* para la administración de bibliotecas. No obstante, al encontrarse dentro de la misma organización, comparten características de contratación de personal, infraestructura tecnológica, apoyo administrativo, entre otras. Consideramos que esto genera una situación muy particular que afecta cuestiones relacionadas con los proyectos de digitalización, las tecnologías disponibles, la contratación de personal capacitado, el presupuesto y, sobre todo, la noción de responsabilidad de la BNM en la socialización y difusión del patrimonio. Revisemos ahora algunos datos relacionados con los proyectos de digitalización en América Latina y, en particular, con la BNM.

La era de la digitalización

Ya son varias décadas de historia de proyectos de digitalización alrededor del mundo. Muchas de las bibliotecas se han enfocado en los materiales patrimoniales, es decir, los que se encuentran en fondos antiguos, reservados, colecciones especiales o, en general, aquellos que por su valor histórico, cultural, artístico o por ser únicos, raros o de limitada disponibilidad son considerados de carácter patrimonial por la institución que los resguarda (Galina, 2018). Los primeros proyectos de digitalización se centraron en la obtención de imágenes digitalizadas por dos razones principales: reducir la manipulación —para la conservación de los originales— y proveer nuevas formas de acceso. Si bien no hay mucha documentación, sabemos que, además de los retos técnicos en los

proyectos de digitalización, se presentaron frecuentemente problemas de orden y control bibliográfico. A menudo, el registro catalográfico de la versión digital estaba desvinculado del catálogo general, y se creaban repositorios con metadatos que no estaban integrados en el *online public access catalogue* (OPAC), el catálogo en línea de la biblioteca. Sabemos por experiencia que en México muchos proyectos que han incursionado en la digitalización de los acervos patrimoniales no cuentan con inventarios depurados de todas sus colecciones, o bien sus registros catalográficos están desactualizados y no representan una herramienta fiable para soportar la digitalización.

Con el paso del tiempo, las bibliotecas han expandido su visión de la digitalización más allá de un simple reemplazo de consulta del original y ofrecen opciones que permiten optimizar el uso de los documentos digitales para fines de investigación a través de metodologías de cómputo, esto es, usar las colecciones como datos (Padilla *et al.*, 2019). Un primer paso es la posibilidad de descargar no solo las imágenes de los documentos, sino también, por ejemplo, sus metadatos o el texto plano, producto del *reconocimiento óptico de caracteres* (OCR, por sus siglas en inglés). Otras bibliotecas han incursionado en ofrecer a sus usuarios herramientas digitales que posibilitan visualizar y manipular las colecciones de otras formas. Para poder realizar esto a nivel escalable, cada una de estas fases en la digitalización descansa sobre la anterior. En este sentido, los procesos de digitalización pueden ir desde una simple imagen escaneada, hasta el acceso a datos y herramientas digitales.

Los discursos en torno a los proyectos de digitalización, en especial de fondos patrimoniales, suelen ser positivos y se presentan como beneficiosos para la sociedad y la difusión de la cultura. En

México, por ejemplo, podemos ver en las últimas décadas un cambio en el discurso de las instituciones encargadas de resguardar y proteger los materiales patrimoniales de forma restrictiva (Ángeles *et al.*, 2005) hacia una mentalidad más abierta, lo que amplía el acceso para brindar mayores posibilidades para su uso. Como comenta Thylstrup (2018), la digitalización masiva ha “adquirido el estatus de un imperativo y obligación cultural y moral” (p. 3).

Existe poco trabajo de investigación sobre los proyectos de digitalización masivos de bibliotecas alrededor del mundo, incluida América Latina. Algunas de las iniciativas de gran impacto mediático han sido analizadas, como Google Books (2002), Europea (2008) y World Digital Library (2009). En Latinoamérica encontramos proyectos pioneros como la Biblioteca Luis Ángel Arango y la Biblioteca Nacional en Colombia², Memoria Chilena de la Biblioteca Nacional de Chile, o la BNM con su HNDM (2002). En años más recientes, se destaca la Biblioteca Digital del Patrimonio Iberoamericano (2012), en la que participan actualmente bibliotecas nacionales de catorce países de América Latina, incluida la de México.

Así mismo, es importante mencionar proyectos de digitalización internacionales en los que participaron instituciones latinoamericanas. Uno de los pioneros de digitalización en español es la Biblioteca Virtual Cervantes (BVC), proyecto creado en España en 1999 por la Universidad de Alicante, con financiación de la Fundación Marcelino Botín y el Banco Santander. Se trata de un proyecto que pretendía digitalizar y poner a disposición de forma abierta textos hispánicos en internet, motivados por la preocupación del

2 Para una historia detallada sobre este proceso, véase Tobón Restrepo (2022).

dominio del inglés en la red; con esto en mente, la BVC buscó la participación de países latinoamericanos. En México, el proyecto comenzó en el 2003, con la participación de la BNM y las bibliotecas del Colegio de México y de la Universidad Iberoamericana, ambas con importantes fondos reservados. La BVC aportaba los recursos para que las bibliotecas pudieran adquirir escáneres y, con el apoyo de la Coordinación de Publicaciones Digitales de la UNAM, responsable técnico, se procedía a la digitalización de los documentos. Una vez obtenidas las imágenes digitales, según las propias especificaciones técnicas designadas por la BVC, se enviaban en CD a España, donde se trabajaban, marcaban en XML, catalogaban e incorporaban a su plataforma. En el 2014, el proyecto se convirtió en la Biblioteca Virtual de las Letras Mexicanas, y hasta la fecha se continúan enviando imágenes periódicamente; del mismo modo, el trabajo de marcado, así como el desarrollo y administración del sistema, se sigue realizando en España.

Pero ¿por qué los primeros proyectos de digitalización de fondos patrimoniales en México enviaban las imágenes para ser procesadas y alojadas en servidores de otro país? Es curioso que los acervos de la BNM, provenientes en su mayoría de las bibliotecas coloniales, regresaran siglos después al mismo país colonizador. ¿Se problematizaron en algún momento las implicaciones legales, jurídicas, socioculturales y políticas de que el acervo patrimonial digitalizado de nuestros países se estuviera gestionando, administrando y representando desde España? Se requeriría, por supuesto, de un trabajo más profundo de reconstrucción histórica, pues, como se mencionó, la bibliografía o documentación disponible escasea. Sin embargo, suponemos que el énfasis estuvo principalmente en aspectos y lineamientos técnicos, ya que, sobre todo en aquella época,

la digitalización se entendía en esencia como un problema técnico para ser resuelto por ingenieros e informáticos, asignado a los departamentos de cómputo y no, como se empieza a entender mejor hoy en día, como una cuestión sociopolítica y de representación de la memoria cultural de una nación (Thylstrup, 2018).

También es probable que la BVC ofreciera la posibilidad a las bibliotecas de digitalizar sus acervos y ponerlos en línea, porque con los presupuestos y recursos tecnológicos que tenían disponibles no lo podían hacer. La primera década de este siglo se caracteriza por una falta de apoyo institucional cohesivo, integrado y general. Encontramos muchos proyectos, esfuerzos aislados e iniciativas solitarias heroicas, pero no un reconocimiento de apostar desde la administración, incluyendo recursos humanos y presupuestales constantes y asegurados a la creación de colecciones digitales, así como asignaciones de presupuesto para proyectos de digitalización, pero sin continuidad (Galina, 2018).

Otro emblemático proyecto contemporáneo es Primeros Libros, iniciado en el 2012; este es liderado por la Universidad de Texas en Austin y cuenta con la participación de alrededor de veinticinco instituciones de cinco países. Su objetivo es digitalizar y poner a disposición los primeros libros impresos en las Américas. Al igual que la BVC, el resultado final se encuentra en una plataforma desarrollada y administrada por una institución extranjera: la Universidad de Texas en Austin. Para este proyecto, se reconoce la importancia de que las instituciones participantes tengan acceso a los archivos terminados, y mucho del desarrollo de la plataforma se realiza con la participación activa de todos, aunque sigue siendo centralizada. En años recientes, se ha trabajado para que la plataforma y sus herramientas estén disponibles

en español y portugués. Otro importante aporte al tema es el trabajo de Reading the First Books, un proyecto para desarrollar *software* de OCR para el reconocimiento óptico de los materiales, lo que produce archivos de texto digital para realizar trabajos de investigación de las HD, como minería de datos; nuevamente, la gran mayoría de la inversión viene de la Universidad de Texas en Austin. ¿Qué conlleva esta dependencia tecnológica? ¿Qué implicaciones tiene sobre cómo nos representamos digitalmente nosotros y nuestro patrimonio bibliográfico? Sin duda, el trabajo colaborativo es de suma importancia y permite que las bibliotecas, de forma cooperativa, compartan recursos y sus colecciones. ¿Dónde está el balance?³

En el 2012 se creó otro proyecto importante, promovido por la Biblioteca Nacional de España, en el que de nuevo participan varios países de Latinoamérica: la Biblioteca Nacional del Patrimonio Iberoamericano. En este proyecto hay un cambio importante en la gestión y administración, ya que cosecha los objetos digitales de las plataformas de cada una de las bibliotecas en vez de centralizarlas (Gutiérrez de la Torre y Cuadros-Sánchez, 2020). Sin duda, es un cambio que responde a cuestiones tecnológicas, pues la misma tecnología posibilita el traslado de grandes proyectos centralizados, donde era necesario hospedar los objetos digitales dentro de la misma plataforma de búsqueda y recuperación, hacia proyectos distribuidos, donde el uso de cosechadores aprovecha la creciente interoperabilidad de estándares de metadatos.

3 Otro ejemplo que habría que analizar es el involucramiento de las bibliotecas latinoamericanas en proyectos de digitalización como Google Books. Esto es material para mucho más trabajo, pero no queremos dejar de mencionarlo.

Asimismo, creemos que refleja, por un lado, la creciente autonomía de las bibliotecas nacionales latinoamericanas, al tener su propia infraestructura tecnológica para crear colecciones digitales en línea; y, por el otro, quizá una conciencia de la importancia de hacerlo. Sin embargo, es importante enfatizar en que nuevamente la iniciativa parte de España.

Regresando a nuestro caso de estudio, durante este tiempo, la BNM fue artífice de dos megaproyectos de digitalización importantes: la HNDM y, posteriormente, la digitalización del Fondo Reservado y la creación de la BNDM. Como hemos comentado, un problema frecuente es la falta de documentación en torno a los proyectos de digitalización (Galina *et al.*, 2020), aunque en el caso particular de la BNM se ha investigado precisamente para subsanar esta carencia (Curiel y Jiménez, 2006, 2015; Galina *et al.*, 2020; Ramírez, 2019), lo que ha permitido vincular algunos de los puntos sobre su historia como institución con los proyectos de digitalización que se desarrollaron en las últimas décadas.

Para iniciar el proyecto digital, la HNM, como muchas hemerotecas del mundo, utilizó microfilmes como punto de partida, y fueron estos los materiales que se digitalizaron, no las publicaciones periódicas en papel. A principios del 2000, la empresa canadiense Cold North Wind se acercó a la HNM para participar en su plataforma digital, que contendría periódicos digitalizados de Estados Unidos, Canadá y México, a través de un sistema de suscripción. Debido a la naturaleza privada y con fines de lucro, la HNM no continuó con su participación, pero este esfuerzo temprano de una compañía extranjera por digitalizar contenidos culturales mexicanos generó suficiente interés para que la UNAM misma diera fondos a través de la Fundación UNAM, así como

aportes de Fomento Social Banamex y de la Fundación Miguel Alemán (Curiel y Jiménez, 2006), para comenzar con la digitalización de sus microfilmes.

Es importante notar el interés internacional en digitalizar contenidos culturales mexicanos desde los inicios. El trabajo de digitalización se llevó a cabo con una empresa privada en México, y aunque los periódicos digitalizados estaban disponibles de forma gratuita dentro de las instalaciones de la BNM, la plataforma digital que proveía acceso a la colección digitalizada, la HNDM, era desarrollada y alojada en servidores de dicha empresa. Esta situación duró varios años hasta que, debido al interés de incluir la HNDM en el proyecto Toda la UNAM en Línea, se consiguió financiamiento nuevamente de la universidad para migrar el sistema y todos los objetos digitalizados a servidores en la BNM. También hubo apoyo adicional de infraestructura de la Dirección General de Cómputo y Tecnologías de la Información (DGTIC), encargada de las telecomunicaciones y la infraestructura tecnológica de la UNAM (Curiel y Jiménez, 2015). La plataforma de la HNDM es un desarrollo propio, por lo que está adaptado a las necesidades y requerimientos de la BNM. Como es de esperarse, la desventaja del desarrollo propio es el mantenimiento y la actualización del sistema, que requiere de recursos humanos y financieros.

Galina *et al.* (2020) documentan su experiencia tras utilizar la HNDM para realizar un proyecto de investigación de minería de datos. En este estudio se constata que, si bien existe la búsqueda por palabra, la cual descansa sobre el OCR, no están disponibles los archivos para descarga a través de la plataforma. Para este trabajo, los investigadores solicitaron —y recibieron— los archivos para uso. Que la interfaz de búsqueda no contemple la

descarga de este tipo de información y se haya tenido que solicitar puede explicarse porque quizá los encargados del proyecto no contemplaron que estos datos también pueden ser de interés para los humanistas digitales o —como comentamos— por una visión del cuidado del patrimonio documental, la cual sigue instalada en limitar el acceso como manera de protegerlo y resguardarlo, lo que se extiende incluso hasta los datos.

En el mismo artículo se revisan tanto los procesos de digitalización como la forma en que los rezagos en la administración documental, provenientes de crónicas deficiencias en el presupuesto y designación de un espacio óptimo para su cuidado, afectaron a largo plazo la catalogación y consulta de los periódicos. A su vez, el sistema de gestión documental automatizada no se vinculó en un inicio al proyecto de digitalización, lo que resultó en la creación de sistemas paralelos sin conexión. Hoy en día se cuenta con diferentes sistemas de inventario, el catálogo de microfilmes, el catálogo de la HNDM y el catálogo general llamado Nautilo. Del mismo modo, la HNDM se manejó desde el inicio como un proyecto de digitalización, siempre con presupuestos extraordinarios, lo que revela la necesidad de la asignación de un presupuesto adicional fijo para su funcionamiento.

Por su parte, la BNDM utiliza una plataforma diferente a la de su contraparte hemerográfica. La HNDM emplea un sistema de índices (ciudad, estado, periodicidad, fecha) y proporciona búsquedas por título o por palabra a través del OCR recogido; las imágenes están disponibles. La BNDM proporciona el acceso a los materiales en formato PDF y no tiene búsqueda en texto completo, tal como la HNDM; sin embargo, sí está vinculada al catálogo general Nautilo y facilita búsquedas por diversos campos. Una revisión de los

antecedentes da cuenta de esta situación. En 1994, la BNM utilizaba Dynix para automatizar los procesos de catalogación (Ramírez, 2019), lo que sirvió de base para la migración a Aleph, sistema para la administración de bibliotecas utilizado para el catálogo automatizado Nautilo. Los primeros esfuerzos de digitalización directa con documentos originales se dieron durante los años noventa, ya que, a diferencia de los periódicos, estos documentos no contaban con alguna reproducción previa pensada para facilitar su consulta sin comprometer su integridad física, como lo fueron los microfilmes en su momento. Estos primeros trabajos de digitalización se llevaron a cabo en respuesta a peticiones de usuarios que requerían fuentes de información para la investigación; la entrega de las digitalizaciones se hacía generalmente en un CD grabado por la misma institución. Con el tiempo, los investigadores del IIB fueron quienes solicitaron bastantes digitalizaciones para apoyar sus proyectos de investigación; de esta manera, se crearon los primeros proyectos digitales, por ejemplo, la Bibliografía Mexicana del Siglo XIX, el Sistema Bibliográfico Sor Juana Inés y Españoles en México Siglo XIX⁴. Este hecho dio como resultado una proliferación de proyectos innovadores, pero dispersos entre sí y, sobre todo, sin conexión al catálogo. En los últimos años se ha realizado un importante esfuerzo por tratar de localizar e incorporar todos estos trabajos, con el fin de administrar de mejor forma la riqueza digital generada. No obstante, esta tarea ha resultado complicada, ya que la falta de personal y la demanda por crear

4 El reporte *Creación, uso y diseminación de los recursos electrónicos para las humanidades* ofrece un recuento de los primeros proyectos digitales en humanidades del IIB y la UNAM (Galina, 2012).

nuevos proyectos agudizan la problemática (Ramírez, 2019). Para la creación de la BNDM, se tomó la decisión de utilizar el mismo *software* Aleph. Las ventajas: no era necesario desarrollar y administrar un sistema nuevo, como sucedió con la creación de la HNDM, y fue sencilla su vinculación directa al catálogo Nautilo. Las dificultades: sus limitantes para adecuar el sistema a necesidades particulares, ya que es *software* propietario.

De estas dos experiencias cabe resaltar la importancia de la existencia previa de un catálogo automatizado y su relación con los proyectos de digitalización, lo que nuevamente apunta a la necesidad de revisar la trayectoria histórica de las bibliotecas, debido a que estas herramientas documentales también se van construyendo sobre trabajo previo, como los inventarios, las memorias, los pocos registros conservados de las colecciones dispersas, los primeros esfuerzos de catalogación y clasificación, los estándares de catalogación extranjeros adoptados, la transferencia de catálogos en fichas a los automatizados, las posteriores iteraciones de sistemas de *software* y migraciones de un sistema a otro, y los proyectos de digitalización y cómo se insertan y relacionan con todo el trabajo anterior.

Cabe anotar que actualmente no están disponibles los archivos OCR para descarga en la HNDM; existe la experiencia de utilizarlos para fines de investigación y se espera que en el futuro próximo esto se modifique. Sin embargo, como hemos demostrado, las diferencias entre las plataformas de la HNDM y la BNDM condicionarán necesariamente las posibilidades del uso de estas colecciones como datos y como base para el desarrollo de herramientas digitales de investigación más sofisticadas.

Conclusiones

Para entender el desarrollo de las HD en América Latina no basta con analizar los entornos tecnológicos en los que se vive, también es preciso prestar atención a los contextos históricos de los recintos y sus colecciones. En ese sentido, conocer la problemática específica que instituciones como la BNM atravesó a lo largo de los años nos ayuda a comprender y poder responder con mejores herramientas a los retos de organización y puesta en marcha de las múltiples iniciativas de proyectos de digitalización que se desprenden de sus colecciones.

Las ideas de la Ilustración que alimentaron el posterior pensamiento positivista y liberal en México tuvieron un fuerte impacto en lo que se refiere a bibliotecas contemporáneas, sobre todo en la configuración del entendido patrimonial de la memoria escrita y la conformación de sus fondos. Esto significa que en los espacios pensados para que la población pudiera educarse con los ideales liberales germinó el ideal de lo que más adelante serían aquellas colecciones de *interés nacional*, rechazando al inicio el legado colonial, por considerarlo obsoleto y opuesto a uno moderno y progresista. No obstante, ese mismo pasado colonial forma parte de la construcción histórica y subyace asimismo en la creación de nuestras identidades latinoamericanas.

Por consiguiente, es entendible el porqué, una vez conseguida la victoria liberal sobre la conservadora y formulada la intención de quitarle poder a la Iglesia sobre la educación de los mexicanos, el Gobierno de Juárez dio prioridad a acciones contundentes para sus objetivos en lugar de buscar una sutil transformación; por esta misma razón no hubo el orden y escrutinio para el legado

bibliográfico que persiguieron pensadores como Gorostiza. Este hecho provocó, al menos en el caso de la BNM, que gran parte de su fondo de origen fuera mermado por los éxodos del coleccionismo bibliográfico al extranjero, tal y como sucede con el ejemplo de proyectos como los de España o el de la Universidad de Texas en Austin, en Estados Unidos, donde los acervos mexicanos funcionan como una mina que alimenta las colecciones en otras latitudes. Esta podría ser una primera diferencia de los proyectos existentes en países del sur global frente a los del norte global, donde los procesos de transformación del siglo XIX mantuvieron cierto orden y continuidad respecto a sus instituciones de memoria, como detectaron Crymble y Afanador Llach (2021).

Los testimonios conservados poco detallan los mecanismos de traslado de documentos, la metodología para el levantamiento de inventarios, los sistemas de clasificación y ordenamiento empleados, y las políticas de selección bibliográfica o de conservación de los libros, lo que genera un vacío que después fue más difícil subsanar con la integración de nuevas colecciones al recinto. Consideramos que estas trayectorias históricas también impactan en el buen desarrollo de los proyectos de digitalización, los cuales, por lo general, parten de catálogos automatizados pobremente alimentados con registros bibliográficos que no reflejan la riqueza documental que yace en los acervos; en consecuencia, tampoco permite tener un panorama completo de lo que conservamos para la mejor toma de decisiones al momento de emprender un proyecto de digitalización. Esta falta de documentación sobre los procedimientos técnicos también puede ser hallada al momento de intentar rastrear proyectos de digitalización en México. Claro está, la disciplina ha aprendido a lo largo del tiempo

y mucha de su metodología de trabajo ha sido fruto de años de desarrollo; sin embargo, parece persistir un fuerte hábito heredado de no atender con rigor la documentación explícita de los proyectos que involucran el manejo —quizá a nivel operativo— de grandes volúmenes de información, y esto se entiende cuando existe poco personal encargado de un sinfín de tareas a la vez.

Para la BNM, el antecedente de un depósito legal fue dado por decreto en 1857, con el cual se obligaba a los impresores de la capital a delegar en la biblioteca dos ejemplares de los impresos que publicaran; lamentablemente, esta disposición no tuvo la acogida esperada en la construcción de las colecciones nacionales, pues, incluso para la década de los cuarenta, Iguíniz (1940) refiere el incumplimiento de esta ley. Un ejemplo opuesto: en España, el antecedente del depósito legal se remonta a 1616 en la Biblioteca de El Escorial y a 1716 en la Biblioteca Real de Madrid, un ejercicio con trayectoria mejor arraigada en la construcción de sus fondos patrimoniales. Claramente, España también enfrentó conflictos armados y desamortización de bienes clericales durante el siglo XIX; no obstante, su tradición bibliográfica y la formación de su biblioteca nacional tuvo un antecedente mucho más directo, pues la Real Biblioteca solo cambió de nombre, hecho que en Nueva España no tuvo un símil. Hasta la fecha de hoy, es notoriamente bajo el cumplimiento de esta disposición en México.

Los fondos patrimoniales conservados son fruto de eventos históricos producto de una inestabilidad económica y social, la cual no permitió mantener proyectos continuos y progresivos, sino que se fueron adaptando conforme a los intereses y necesidades de la época. Por tal motivo, procurar una continuidad política en la dirección de las instituciones que resguardan la memoria

documental del país ha representado un reto tan grande como la organización y ejecución de los proyectos que buscan la consulta del patrimonio bibliográfico a través de las herramientas tecnológicas. Existen trabajos como el de García Aguilar (2008) que han descrito el problema de la representación bibliográfica de los objetos patrimoniales y los conflictos del valor cultural que esto acarrea. Históricamente, el inventario de los fondos patrimoniales ha presentado tropiezos para la disciplina bibliotecológica que aún no han sido sorteados con pleno éxito; tal es el caso de la propuesta todavía inconclusa de un catálogo colectivo del patrimonio bibliográfico del país, que sí existe en otras naciones como España o Francia. ¿Es un problema exclusivamente de financiamiento o continúa representando una carencia en la formación de personal especializado? La respuesta aún está en el aire; no obstante, los sentires demostrados por las entrevistas de Gutiérrez (2020) a los bibliotecarios latinoamericanos o los retos de organización e infraestructura descritos en Galina *et al.* (2020) tienen paralelismos históricos en la conformación del proyecto de la BNM durante el siglo XIX. Las mismas claves que han imposibilitado la creación de catálogos colectivos mexicanos pueden darnos un norte del porqué de las diferencias en proyectos de digitalización. La tecnología existe, pero ¿está al alcance de todas las bibliotecas que custodian la diáspora del patrimonio? Hay una fuerte tendencia a que el diverso origen de los proyectos, la naturaleza de los acervos, la administración de los fondos, el grado de capacitación del personal y un largo etcétera trunquen su despunte.

En general, los proyectos de digitalización son percibidos como inherentemente “buenos”, ya que incrementan la difusión y el acceso a las colecciones patrimoniales, y sirven como recurso

para la investigación y la docencia. Sin embargo, esta lógica generalizadora limita la capacidad de dimensionar las probables consecuencias negativas o inesperadas que pueden acarrear dichos proyectos, como los costos ocultos relacionados con aspectos geopolíticos —de acuerdo con los participantes involucrados—, las tecnologías utilizadas, los metadatos asignados o los responsables de determinar las características de las plataformas, por mencionar algunos ejemplos. En la revisión de los proyectos de digitalización internacionales, es interesante ver la colaboración que ha tenido México con España, Estados Unidos y Canadá; sería interesante examinar los vínculos que otras bibliotecas latinoamericanas tienen con el norte global. Las posibilidades de desarrollo de proyectos de digitalización de gran nivel están condicionadas también a vinculaciones geopolíticas; a su vez, esto puede determinar la infraestructura disponible para las HD. Como hemos visto, las posibilidades de uso de colecciones digitales para las HD dependen de cómo se conciben y ejecutan los proyectos de digitalización, los cuales a su vez dependen del estado en el que se encuentra la colección y la automatización de los catálogos, además de cuestiones de infraestructura tecnológica. Es importante, en particular para los proyectos de HD y las bibliotecas mismas, que los proyectos de digitalización se entiendan en toda su complejidad, incluida su trayectoria histórica.

Sin duda, gracias al trabajo colectivo de muchos —personal de la BNM, investigadores del IIB y otros profesionales— se han logrado importantes avances. Los acervos de la BNM constituyen un invaluable patrimonio para el país y, pese a las dificultades, resguardan y proveen el acceso a nuestro legado bibliográfico. En el ámbito digital, proyectos como la HNDM fueron pioneros

y continúan siendo un punto de referencia para el desarrollo de colecciones digitales, además de que han sido utilizados por miles de usuarios.

Como mencionamos al inicio de estas líneas, los planteamientos expuestos en este esbozo precisan no solo un espacio de reflexión mucho mayor, sino además una ardua labor de investigación colaborativa entre especialistas de distintas áreas del conocimiento, pues en la confección de herramientas digitales también yace la responsabilidad compartida por la preservación, el acceso y la socialización del patrimonio bibliográfico para las generaciones futuras. Consideramos que este ejercicio podría realizarse con otras bibliotecas en América Latina y, a partir de este trabajo, comenzar a identificar características que nos permitan vislumbrar un panorama de la situación de los repositorios digitales en América Latina y su uso para las HD.



Referencias

- Afanador Llach, M. J. (2021, 11 de noviembre). *Archivos y humanidades digitales en América Latina* [ponencia magistral]. Quinto Encuentro de Humanistas Digitales RedHD.
- Afanador Llach, M. J., Murcia Galindo, J. C., Rincón Rodríguez, O. G., Gallini, S., Méndez Mahecha, T. y Jaramillo Liévano, J. N. (2020). Humanidades digitales “a lo colombiche”: Cadáver exquisito de la Red Colombiana de Humanidades Digitales. *Revista de Humanidades Digitales*, 5, 217-235. <https://doi.org/10.5944/rhd.vol.5.2020.27837>
- Ángeles, P., Aureliano, R., Bañuelos, J., García Aguilar, I., Michel, P. y Morfin, M. (2005, 18 de noviembre). *Panorama de la digitalización de bienes culturales en México* [documento de trabajo]. Foro Digitalización y Patrimonio, Auditorio del Instituto de Investigaciones Estéticas, Universidad Nacional Autónoma de México, Ciudad de México.
- Biblioteca Nacional de México (BNV). (s. f.). *Antecedentes*. <https://bnm.iib.unam.mx/index.php/quienes-somos/antecedentes>
- Crymble, A. y Afanador Llach, M. J. (2021). Digital history: The globally unequal promise of digital tools for history: UK and Colombia's case study. En J. Clark y A. Nye (Eds.), *Teaching history for the contemporary world tensions, challenges and classroom experiences in higher education* (pp. 85-98). Singapore.
- Curiel Defossé, G. y Jiménez Rivera, R.J. (2006). Hemeroteca Nacional Digital de México: un esfuerzo en favor de la memoria. En *La Biblioteca Nacional: Triunfo de la República* (pp. 259-267). Universidad Nacional Autónoma de México.

- Curiel Defossé, G. y Jiménez Rivera, R. J. (2015). Diez años de la Hemeroteca Nacional Digital de México (HNDM). Breve reseña de una larga gestión de preservación y acceso, 2002-2012. En *Textos, píxeles y bits: Reflexiones sobre la publicación digital* (pp. 39-86). Universidad Nacional Autónoma de México, Instituto de Investigaciones Bibliográficas, Biblioteca Nacional/Hemeroteca Nacional.
- del Rio Riande, G. (2018). Humanidades digitales bajo la lupa: Investigación abierta y evaluación científica. *Exlibris*, (7), 136-149.
- Fundación Histórica Tavera. (2000). *Los archivos de América Latina. Informe exporte de la Fundación Histórica Tavera sobre la situación actual*. Fundación Tavera y Banco Mundial.
- Galina Russell, I. (2012). Creación, uso y diseminación de los recursos electrónicos para las Humanidades. Reporte final. *Humanities Commons*. <https://doi.org/10.17613/qf6b-ef29>
- Galina Russell, I. (2018). La digitalización de fondos patrimoniales. El caso de México. *Relaciones Estudios de Historia y Sociedad*, 39(155), 135-166. <https://doi.org/10.24901/rehs.v39i155.300>
- Galina Russell, I., Priani Saisó, E., Peña Pimentel, M., Martínez Domínguez, L. y Castellanos Rueda, R. (2020). Hemeroteca Digital Nacional de México, su historia y alcance como fuente para proyectos digitales. *Relaciones Estudios de Historia y Sociedad*, 41(163), 200-219. <https://doi.org/10.24901/rehs.v41i163.707>
- García Aguilar, M. I. (2008). Entre páginas de libros antiguos: La descripción bibliográfica material en México. *Investigación Bibliotecológica*, 22(45), 13-40.
- García Aguilar, M. I. y Aureliano Alarcón, R. (2007). *La catalogación del libro antiguo en México: Revisión de un problema de*

- control patrimonial*. https://ru.iibi.unam.mx/jspui/handle/IIBI_UNAM/CL907
- Gutiérrez de la Torre, S. E. (2020). Bibliotecas y humanidades digitales en América Latina. *Revista de Humanidades Digitales*, 5, 113-131. <https://doi.org/10.5944/rhd.vol.5.2020.27826>
- Gutiérrez de la Torre, S. E. y Cuadros-Sánchez, M. D. (2020). Digital resources: The Digital Library of Ibero-American Heritage. En *Oxford Research Encyclopedia of Latin American History*. Oxford University Press. <http://dx.doi.org/10.1093/acrefore/9780199366439.013.798>
- Iguíniz, J. B. (1940). La Biblioteca Nacional de México. *Revista de Historia de América*, 8, 57-86.
- Kear, R. y Joranson, K. (Eds.). (2018). *Digital humanities, libraries, and partnerships: A critical examination of labor, networks, and community*. Chandos Publishing.
- Ley de 12 de julio de 1859. Ley de nacionalización de bienes eclesiásticos. <https://archivos.juridicas.unam.mx/www/bjv/libros/12/5625/23.pdf>
- Martínez Arellano, F. F. (2002). Cataloging and classification history in Mexico. *Cataloging & Classification Quarterly*, 35(1-2), 227-254. https://doi.org/10.1300/J104V35N01_14
- Millson-Martula, C. y Gunn, K. B. (Eds.). (2020). *The digital humanities: Implications for librarians, libraries, and librarianship*. Routledge, Taylor & Francis Group.
- Morales del Castillo, J. M., Tamayo Ramírez, R. M. y Peis Redondo, E. (2020). Bibliotecas de investigación y humanidades digitales en España: Una relación en construcción. *Revista de Humanidades Digitales*, 5, 86-112. <https://doi.org/10.5944/rhd.vol.5.2020>

- Nava Martínez, O. (2017). Destruir y luego organizar: La nacionalización de las bibliotecas conventuales y la formación de una biblioteca nacional y pública en la Ciudad de México. En L. Suárez de la Torre (Coord.), *Estantes para los impresos. Espacios para los lectores. Siglos XVIII-XIX* (pp. 309-342). Instituto Mora.
- Padilla, T., Allen, L., Frost, H., Potvin, S., Russey Roke, E. y Varner, S. (2019). *Always already computational: Collections as data: Final report*. <https://zenodo.org/record/3152935#.X6WOf-LPzIU>
- Posner, M. (2013). No half measures: Overcoming common challenges to doing digital humanities in the library. *Journal of Library Administration*, 53(1), 43-52.
- Quirarte, V. (Ed.). (2006). *La Biblioteca Nacional: Triunfo de la República*. Universidad Nacional Autónoma de México.
- Ramírez, A. Y. (2019). Biblioteca Nacional Digital de México: Orígenes, encrucijadas, tropiezos y logros. *Boletín de la Biblioteca Nacional de México*, 3, 69-80.
- Ramírez López, J. E. (2020). La Biblioteca John Carter Brown: Del éxodo bibliográfico a la conservación del patrimonio mexicano. *Bibliographica*, 3(2), 15-50. <https://doi.org/10.22201/iib.2594178xe.2020.2.78>
- Rodríguez, J. E. (1986). La crisis de México en el siglo XIX. *Estudios de Historia Moderna y Contemporánea de México*, 10(10). <https://doi.org/10.22201/iih.24485004e.1986.010.68957>
- Ruiz Castañeda, M. del C., Schneider, L. M. y Castro, M. Á. (2004). *La Biblioteca Nacional de México: Testimonios y documentos para su historia*. Biblioteca Nacional de México.
- Spell, L. M. (1959). La fundación de la Biblioteca Nacional. *Historia Mexicana*, 8(4), 449-473.

- Thylstrup, N. B. (2018). *The politics of mass digitization*. MIT Press.
- Tobón Restrepo, I. (2022). *Del fichero al ciberespacio: Historias orales de la transformación digital del patrimonio cultural colombiano y su llegada a la web* [tesis de maestría]. Universidad de los Andes. <http://hdl.handle.net/1992/64315>
- Vázquez Mantecón, C., Herrero Bervera, C. y Flamenco Ramírez, A. (Eds.). (2007). *La Biblioteca Nacional de México 1810-1910: Las bibliotecas mexicanas en el siglo XIX*. Nueva Biblioteca Mexicana Herreriana, Universidad Autónoma Metropolitana Iztapalapa.

**La interoperabilidad
de los datos digitales:
el uso compartido
de información en el
proyecto ArqueoData,
una base de datos
radiocarbono para
Latinoamérica**

Para citar este capítulo:

<http://doi.org/10.51573/Andes.9789587987362.9789587987379.9789587987867.2>

Introducción

{ El uso de las nuevas tecnologías a través de las redes telemáticas en la investigación histórica ha sido estudiado desde diversas aristas. El giro digital de la historia demostró las consecuencias de aquello denominado *humanidades digitales* (HD)¹ desde un punto de vista historiográfico, al llevar al plano del análisis histórico dicho campo de investigación (Bresciano, 2010; Pons, 2013). Uno de los objetivos principales de las actuales investigaciones y desarrollos por parte

- 1 El concepto de HD es muy amplio y está abiertamente discutido por parte de la comunidad académica, que abarca desde historiadores, sociólogos, antropólogos, artistas, hasta comunicólogos, filósofos, entre otros. Esto ha sido un reto para su definición, pues no se quiere limitar su práctica. En este sentido, se ha optado por definir a estas como un conjunto de prácticas que comparte una serie de comunidades académicas y que establecen un punto de intersección en este concepto. Por lo tanto, para propósitos de este ensayo, utilizaré el concepto de HD en este último sentido. Para conocer más, véase Gayol y Melo Flórez (2017), Priani Saisó (2019), del Rio Riande (2014) y Russell (2018).

de la comunidad académica es la libre consulta de la información, principalmente bajo los principios FAIR: rastreadable, accesible, interoperable y reusable (*findable, accesible, interporable, reusable*) (Wilkinson *et al.*, 2016). Estos objetivos guían actualmente la mayor parte de herramientas digitales, con base en otros supuestos conceptuales como el *linked open data*, los datos abiertos (Ríos-Hilario *et al.*, 2012) o la cultura de la documentación (Morales *et al.*, 2021).

En este ensayo me propongo hacer una reflexión sobre el objetivo de la interoperabilidad de las herramientas digitales, tomando como estudio de caso la base de datos radiocarbono ArqueoData². Al entender el concepto de *interoperabilidad* como la capacidad de una plataforma digital para intercambiar y consumir información de otros sistemas (Lamprecht *et al.*, 2020), debido a la estandarización y, en consecuencia, a la semejanza entre metadatos, resulta posible consultar, cosechar e interactuar con estos entre ambas plataformas. De igual forma, discutiré sobre las posibilidades que se abren para la interoperabilidad con otras bases de datos radiocarbono, y con otros sistemas web o locales que quieran utilizar la información que se resguarda en ArqueoData, así como invitar a otros sistemas web, que se desarrollan por individuales o instituciones, para que la interoperabilidad sea una realidad del giro digital en las investigaciones sobre HD latinoamericanas.

Así, sostengo que el ejercicio de interoperabilidad puede ser satisfactorio para los proyectos en HD a partir del uso de una *interfaz de programación de aplicaciones* (API)³. La utilización de

2 Véase <https://www.arqueodata.com>

3 API son las siglas de *application programming interface*; esta se define como una serie de protocolos que permiten la comunicación e integración de aplicaciones

este tipo de tecnologías, propias de la ciencia de la computación, permite visualizar un entorno de comunicación masivo de datos entre sistemas con diferentes arquitecturas de *software*⁴.

En suma, a partir de un proyecto de realización de una base de datos sobre radiocarbono, abrí la discusión hacia la utilización de las HD en la comunidad arqueológica, sus usos y retos ante un espacio en constante cambio, al igual que los aportes que este tipo de herramientas puede ofrecer a la investigación, en busca de nuevos cuestionamientos en la ciencia arqueológica. Esto es posible al preguntarme para qué sirve contar con este tipo de herramientas y su empleo en la web, qué diferencias tiene este proyecto con otras propuestas y cómo este tipo de herramientas digitales puede llevar a la generación de una investigación más fructífera en la arqueología.

El uso de bases de datos

Las bases de datos nos ayudan a hacer más eficiente y rápido el control de información en un proyecto⁵. Desde el siglo XVIII, las

.....
 y usuarios externos. Una API se compone, básicamente, por una petición —*request*— y una respuesta del servidor —*response*— (Haider, 2021, p. 14).

- 4 La arquitectura de *software* es el conjunto de estructuras necesarias para que un sistema pueda funcionar. El desarrollo de esta requiere no solo pensar en un grupo de individuos, sino también en la satisfacción de requerimientos, integración e implantación (Cervantes Maceda *et al.*, 2016, p. 4).
- 5 Un sistema de manejo de bases de datos (DBMS, *database management system*) sirve, en las tecnologías de la información, para guardar registros y relacionarlos entre sí.

colecciones, los museos, las bibliotecas y los archivos llevaron a cabo clasificaciones de los objetos; estas se anotaban en compendios, libros o tarjetas, con el fin de crear ficheros e informes de registros (López Zahar, 2020). A partir de 1970, con el auge de las computadoras, a pesar de que aún no eran una herramienta de oficina, varias instituciones comenzaron a emprender proyectos para digitalizar su información (Kahler, 1978). Una de las disciplinas que se acercó a estas prácticas con mayor antelación fue la bibliotecología. Sus bases de datos realizadas como ficheros, principalmente con el estándar MARC⁶, se fueron adaptando a nuevos estándares de clasificación e incluso se mudaron las fichas a un formato XML⁷ (Baca, 2016).

A nivel internacional, grandes museos y colecciones de arte también crearon bases de datos digitales, aunque hay que destacar que la mayoría de estos trabajos se dieron en instituciones con grandes recursos estatales y ayuda privada; ejemplo de este tipo de proyectos se pueden apreciar en la Biblioteca Nacional de España (D'Amaro, 2019). Posteriormente, se llevó a cabo el gran proyecto de digitalización por parte de la Unión Europea, denominado Europea (Isaac y Haslhofer, 2013). En Latinoamérica,

- 6 MARC significa *machine-readable cataloguing format* y hace referencia a un conjunto de datos estandarizados para describir materiales bibliográficos que faciliten la catalogación y el intercambio de información entre sistemas de control bibliográfico (Baca, 2016).
- 7 El *lenguaje de marcado extensible* (XML) es un sistema de etiquetado que se utiliza en la web para facilitar la estructura de datos, como consecuencia de las carencias del *lenguaje de etiquetas de hipertexto* (HTML), aunque este tiene sus inicios en la década de los sesenta, con el *general markup language*, creado por Charles Goldfarb, Edward Mosher y Raymond Lorie de IBM (Bates, 2003).

la práctica de la digitalización no se presentó hasta ya entrado el siglo XXI, y es una tarea que aún hoy en día se sigue desarrollando (Gayol y Melo Flórez, 2017). Algunos proyectos pioneros en este ámbito fueron las bases de datos que se realizaron en países como Chile (Rodríguez Matta, 2009) y Colombia (Tobón Restrepo, 2022).

Las bases de datos nos permiten organizar, compartir y estandarizar una serie de información de manera global. Desde la década de 1960, estas funcionaban de forma local, en muchas ocasiones solo en el mismo *software* que se había diseñado para tal función; sin embargo, a partir de 1990, con el comienzo del uso masivo de la *world wide web*⁸, las bases de datos comenzaron su incursión en este mundo interconectado (Hendler y Berners-Lee, 2010). En el caso de una comunidad académica que se dedica a la investigación, tener una base de datos pública en la web, la cual pueda ser consultada y alimentada desde el mismo internet sin necesidad de descargar archivos y tener que actualizar cada cierto tiempo, representa un gran avance en las herramientas digitales para la investigación en comunidad.

Bajo estas premisas surgió en el 2019 el proyecto ArqueoData, en el que me enfoqué en recopilar dataciones radiocarbono de la década de 1960, con la finalidad de contabilizar la producción de estos en la comunidad arqueológica mexicana. Sin embargo, al

8 La *web* se refiere a un tipo específico de comunicaciones mediante el internet. La web tuvo como base una estructura y lenguaje estandarizado llamado HTML, que permitió a los usuarios receptores interactuar con los elementos en pantalla (Banks, 2008). La utilización del internet nos da una visibilidad pública global, por lo que no tenemos la limitante espacial para la consulta de los datos; no obstante, nuestro límite es el acceso al internet, denominado por algunos autores *brecha digital* (Santoyo y Martínez, 2003).

recopilar los datos me percaté de que la información podía servir a otros fines, así como ser comentada y complementada a partir del diálogo con otros colegas arqueólogos. Ante esto, pensé que sería útil un canal de comunicación y propagación de esta base de datos a través de la web.

Las bases de datos radiocarbono

En historia, las bases de datos tienen sus principales aplicaciones para gestionar archivos (Melo Flórez, 2011); en arqueología, para guardar la información proveniente de los contextos arqueológicos (Jiménez Badillo y Gándara, 2016), además de un abanico amplio de herramientas que han sido creadas para la historia digital y la arqueología computacional (Bresciano, 2010; Jiménez Badillo, 2017; Schroeder y Brügger, 2017).

En arqueología, específicamente en la recopilación de resultados que ofreció la técnica de datación radiocarbono⁹, fue una tarea que comenzó en paralelo a la creación de la técnica misma. Desde mediados de 1940 se publicaron números especiales de la revista *Science* dedicados solo a presentar tablas de los resultados de diversos laboratorios; a partir de 1958 esta tarea estuvo a cargo de la

9 El radiocarbono es una técnica de datación arqueométrica que permite conocer el intervalo de temporalidad en el que un ser orgánico murió. Esta técnica inició sus aplicaciones en la arqueología durante la década de 1940, gracias a los trabajos del Dr. Willard Frank Libby y su equipo en la Universidad de Chicago (Libby, 1955). Desde entonces, se ha perfeccionado la técnica e incluso se han producido varios métodos, de los cuales destaca el conteo por centelleo líquido (LSC) y la espectrometría de masas con aceleradores (AMS); esta última es la más reciente y utilizada en el mundo, incluso en México se cuenta con un equipo y un laboratorio que proporciona este servicio a los proyectos arqueológicos (Solís *et al.*, 2014).

revista *Radiocarbon*. Desde 1990 se vio reducido el número de laboratorios que participaban en estas tablas; sin embargo, la llegada del internet abrió un horizonte de posibilidades para compartir información. Así, desde la última década del siglo xx se comenzó la publicación de artículos sobre bases de datos de radiocarbono que pueden ser consultadas o descargadas mediante sitios web (Ashmore *et al.*, 2000; Gulliksen, 1983; Moffett y Webb, 1983).

Las bases de datos sobre resultados radiocarbono han sido realizadas por diversas instituciones y proyectos, y concentran, principalmente, información de sitios arqueológicos estadounidenses y europeos, por ejemplo: Canadian Archaeological Radiocarbon Database (CARD, 2015), Infraestructura de Datos Espaciales de Investigación Arqueológica (IDEArq, 2019), IntChron (Ramsey, 2010), AndesC14 (Ziólkowski *et al.*, 2015) y p3k14c (Bird *et al.*, 2022), aunque esta última no tiene un ambiente web¹⁰ y solamente está disponible para su descarga en csv¹¹, o a través de un paquete en R¹². Sin embargo, los trabajos anteriores solo presentan la información en tablas o mapas con buscadores de información; a excepción de CARD, ninguna tiene una API para extraer información de sus bases de datos (tabla 2.1).

- 10 Por ambiente web me refiero a que no se puede consultar en la web mediante internet, lo que limita su uso a la descarga de los datos.
- 11 Un archivo csv, *comma separated value*, es un tipo de archivo de texto que permite representar datos en forma de tabla a partir de comas y saltos de línea.
- 12 R es un lenguaje de programación utilizado para el computo estadístico y la visualización grafica de resultados; desde la década de los noventa es muy utilizado por la comunidad académica de ciencias a nivel internacional. Véase <https://www.r-project.org/>.

TABLA 2.1. Proyectos de bases de datos radiocarbono en la web y sus características principales

CAMPOS	IDEARQ	INTCHRON	ANDESC14	CARD	ARQUEODATA
Año	2019	2010	2015	2015	2021
Aplicación web	X	X	X	X	X
Paquete R		X			
Comunidad				X	X
Reportes de error				X	X
Exportar	X		X	X	X
Usuarios		X		X	X
API		X			X

FUENTE: elaboración propia.

Bases de datos y metadatos

Los datos de los datos, así es como comúnmente se denomina a los metadatos (Gilliland, 2016), los cuales comprenden una descripción de un objeto digital y su misma composición nos permite lograr una interoperabilidad entre estos, por lo cual se han generado estándares de metadatos. Para una mejor organización de los objetos, digitales o físicos, debemos considerar estándares de catalogación que nos ayuden a ordenar de una mejor forma y, por ende, a entender los objetos con los que trabajamos.

Los estándares digitales para el patrimonio cultural iniciaron su uso en la década de 1990, principalmente en la bibliotecología

y la organización de archivos (Baca, 2016). Los estándares físicos ya existían desde antes, pero en la última década del siglo xx se adaptaron al mundo digital, por ejemplo, el MARC y su ajuste al XML. Los estándares de catalogación pueden clasificarse en vocabularios controlados, modelos de metadatos y reglas de catalogación. Los primeros hacen referencia a identificar y unificar lo que se denomina una *ontología*; esta se compone de conceptos y sus relaciones. *Grosso modo*, la ontología puede representarse bajo el *modelo de entidad-relación* (MER) en el diseño de una base de datos¹³; sin embargo, su creación necesita de expertos en la materia y no solo puede ser creada por especialistas en informática, lo que conlleva un constante trabajo de transdisciplinariedad. Los modelos de metadatos son estructuras, casi siempre en XML o JSON¹⁴, que permiten estandarizar etiquetados; uno de los más utilizados es Dublin Core. Por último, existen reglas de catalogación, las cuales pueden ser modelos como el *cataloging cultural objects* (cco), el *conceptual reference model* (CIDOC CRM) y el *object id*.

13 Una base de datos se compone de tablas y estas, a su vez, se pueden relacionar unas a otras; por lo tanto, a este tipo de bases se les ha denominado *diseños lógicos* bajo el MER. Este modelo permite crear una serie de rutinas y funciones para acceder a los datos, crearlos, actualizarlos y eliminarlos. El MER representa una percepción del mundo real contenida en conjuntos de datos: “una entidad es un objeto que existe y puede distinguirse de otros, dado que cada uno posee ciertos atributos que le identifican. Una relación es una asociación entre varias entidades” (Rivera, 2008, p. 17).

14 Estándar de etiquetado que permite una lectura más limpia para el usuario y está diseñado, principalmente, para su uso en *single page applications*. Este formato fue denominado *notación de objetos de JavaScript* (JSON) (Mora-Castillo, 2016).

Los metadatos en el campo de la arqueología no están estandarizados, lo que puede generar un problema a la hora de elaborar bases de datos. No obstante, a través de una revisión previa de los campos usados por otras bases de datos radiocarbono —como las presentadas—, todas contienen la información de la datación bajo los estándares de los laboratorios de radiocarbono. Aun así, la información que proviene del contexto arqueológico queda como un mero anexo en un campo de descripción, y esto atañe a la historia de la arqueología y cómo ha construido sus estándares de clasificación desde el siglo XIX, al menos para el caso de la arqueología mexicana (Tschauner, 1985).

La naturaleza del dato arqueológico

La estandarización de los datos arqueológicos aún se dificulta en el mundo analógico; así, las tareas principales, excavación y clasificación de materiales, se ven condenadas solo a seguir la forma de descripción de los proyectos en turno. En el caso mexicano, es abundante llevar el registro de excavación por niveles métricos y clasificar a través de los principios del sistema tipo-variedad (López Aguilar, 2012). Esto ha tenido como consecuencia la creación de un sinnúmero de culturas arqueológicas que se han formado desde finales de las últimas décadas del siglo XIX (López Hernández, 2018).

En la práctica de excavación a nivel internacional, uno de los modelos de registro más utilizados es el *single-context-recording*, que utiliza las unidades estratigráficas (Harris, 1979) y su organización durante el proceso de excavación (Sandoval García, 2019).

Incluso este modelo de registro es el que mejor se adapta a las variables de los *softwares* de calibración bayesiana, como OxCal (Ramsey, 1995) y Calib¹⁵. Sin embargo, los contextos prehistóricos pueden no abundar en información estratigráfica y se emplean otros métodos de registro, lo que hace que ya no sigan un modelo único, y las muestras radiocarbono de sitios prehistóricos son muy abundantes.

Lo anterior dificulta el registro de toda la información proveniente del contexto arqueológico de cada muestra. De igual forma, la mayor parte de los repositorios han sido realizados por colegas arqueómetras, lo que ha hecho que solo se tomen en consideración los estándares de la técnica como variables indispensables, sin unificar el campo del contexto arqueológico y dejándolo como un campo de comentario anexo a cada muestra. En el caso particular de las bases de datos radiocarbono, se pueden identificar patrones en la información que presenta cada sitio web y la forma en que se separan los campos de registro en la base de datos (tabla 2.2).

TABLA 2.2. Datos públicos de los proyectos y los campos que emplean para guardar la información

CAMPO	IDEARQ	INTCHRON	ANDESCI4	CARD	ARQUEODATA
Clave laboratorio	x	x	x	x	x
Material	x	x	x	x	x
Taxonomía	x				

Continúa...

15 Véase <http://calib.org/calib/calib.html>

Continúa...

CAMPO	IDEARQ	INTCHRON	ANDESC14	CARD	ARQUEODATA
Sitio	X	X	X	X	X
Técnica	X				X
Contexto	X			X	X
Periodo	X	X			X
Edad R14	X	X	X	X	X
Error	X	X			X
D13C		X	X	X	X
Investigador			X		X
Referencias	X	X	X	X	X
Calibración		X			X

FUENTE: elaboración propia.

En miras hacia una interoperabilidad entre sistemas, la estandarización de los metadatos de cada muestra debe presentar similitudes entre sistemas, incluso si se tiene en cuenta que ninguna base de este tipo utiliza los estándares internacionales de catalogación, como Dublin Core, porque sería difícil equiparar los campos de una datación con las variables de dicho sistema de estandarización de metadatos. Aun así, valdría la pena realizar una revisión para tener una mejor comunicación entre sistemas.

Ninguna base de datos radiocarbono considera el valor de la interoperabilidad ni el trabajo en comunidad, solo sirven como catálogos en línea, los cuales son consultables y, en ocasiones, es posible descargar su información. Sin embargo, mi propuesta a partir de las HD es que este tipo de proyectos puede escalar para hacer un

trabajo tanto de registro como de uso compartido de la información a través de las plataformas web, y la interoperabilidad por medio de una API, la cual es una forma de acceso rápido a la información si se quiere enlazar a otras bases de datos, con la finalidad de no repetir información, así como método de cosecha de información para otros usos académicos.

El duplicado de registros en las bases de datos es un problema que concierne a la falta de interoperabilidad entre las bases internacionales y a la falta de una estandarización común a todos los proyectos. En el caso de ArqueoData, las dataciones presentes en otras bases sobre México eran muy escasas, lo cual minimizó el problema de duplicación.

Sin embargo, la presencia de registros duplicados ha sido un problema recurrente en los *datasets*, que son colecciones estructuradas de datos, disponibles para descarga y organizadas por versiones. Un ejemplo destacado es el *dataset* p3k14c (Bird *et al.*, 2022), desarrollado por el proyecto Paleoclimate and the Peopling of the Earth (PEOPLE) 3000. Este conjunto incluyó, en su versión 2021, 170 000 dataciones de diversas partes del mundo y se presenta en formato CSV.

A pesar de su utilidad, enfrenta desafíos significativos, como la ausencia de metadatos y enlaces por medio de identificadores uniformes de recursos (URI), los cuales pueden proporcionar un enlace claro y directo a la ubicación exacta de donde se obtuvo la información en la web. Esta limitación constituye un problema grave de interoperabilidad, que afecta tanto la referencia cruzada de los datos en otras bases de datos como su verificación en las fuentes originales, sean artículos o libros, donde se publicó la información.

Las aplicaciones web en la investigación

Una aplicación web es una serie de código y bases de datos con los que se puede interactuar a través de internet, por medio de un navegador web y etiquetas y lenguajes que pueden ser interpretados por los navegadores, por ejemplo, HTML y JavaScript. Para la interacción con la base de datos se utilizan los servicios web que comunican el código del navegador (*front-end*) con el código e instrucciones del servidor (*back-end*); este último puede estar programado en PHP, C#, Java, entre otros lenguajes. Las acciones que realice el usuario en el *front-end* —como dar clic a un elemento del menú, una noticia, una imagen, un video— enviarán una petición (*request*) al *back-end*, que lo recibirá, procesará y enviará una respuesta (*response*); en ocasiones, puede actualizar todo lo que el usuario visualice en el navegador (figura 2.1).

A diferencia de un sitio web, donde el objetivo solo es presentar información, una aplicación web presenta funcionalidades y herramientas para los usuarios. Por ejemplo: en el campo de la creación de sitios web se destacan aplicaciones como WordPress, Joomla o Drupal, los cuales se denominan *content management system* (CMS); en el ámbito de la investigación y gestión de colecciones sobresalen Collective Access y el complemento Tinacan para WordPress. Estos últimos son muy utilizados en el mundo para crear sitios web dinámicos que hospeden repositorios digitales institucionales.

En la investigación, el concepto de *aplicación web* aún no es tan utilizado, destaca la expresión *sitio web*; sin embargo, en términos informáticos, esta aún se queda corta y puede variar entre

sitios estáticos y dinámicos, a pesar de que un sitio web dinámico cuenta con un panel de administración para, principalmente, incluir artículos en formato de blog. Las aplicaciones web contienen un sitio y un panel de administración, pero están pensadas para que no solo los administradores puedan publicar información, sino que sean utilizadas por los usuarios dentro del entorno web. Con lo anterior quiero enfatizar en la importancia del trabajo en comunidad dentro de una aplicación web.

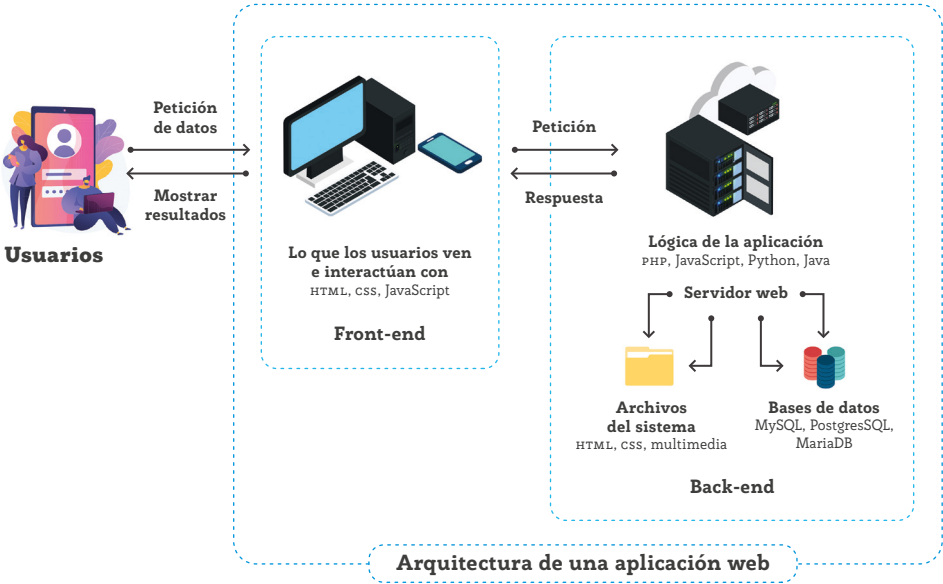


FIGURA 2.1. Representación de la comunicación entre el *back-end* y el *front-end*

FUENTE: elaboración propia.

Los retos de la interoperabilidad: las API

La interoperabilidad es definida como la capacidad de una plataforma digital para intercambiar y consumir información (Lamprecht *et al.*, 2020). Debido a la estandarización de los resultados y la semejanza entre metadatos, resulta posible aplicar la interoperabilidad de dicha información. En el campo de la informática, esta interoperabilidad se consigue cuando un sistema puede conectarse con otro, pero ambos siguen siendo independientes. Para conseguir este tipo de trabajos, se crearon las consultas AJAX durante la primera década del siglo XXI. La información se consulta por medio de “servicios web”, los cuales pueden responder con un formato estándar XML; no obstante, a partir del 2009 surgió un nuevo estándar de visualización de datos que hoy en día es el principal para aplicaciones web: el formato JSON.

El reto más grande de la interoperabilidad es que los estándares de metadatos sean iguales o similares, con el fin de intercambiar de forma más rápida la información. En el caso de la bibliotecología y los museos, como se describió, los vocabularios controlados y los estándares de metadatos ya son una realidad; sin embargo, en arqueología y específicamente en radiocarbono los únicos estándares que seguimos son los de la técnica en sí, pero ninguna en sentido del contexto arqueológico.

En las bases de datos radiocarbono los estándares de datación implican tener en todas las bases los siguientes atributos: clave de laboratorio, material, sitio, edad radiocarbono, incertidumbre y referencia. Por otro lado, los atributos que pueden o no estar presentes son: taxonomía del material, técnica utilizada, contexto, periodo, delta 13, investigador que colecto la muestra, año de recolección y opción de calibración (véase la tabla 2.2).

A pesar de tener atributos similares, aún falta una estandarización completa en todas las bases de datos. Recientemente, el proyecto p3k14c (Bird et al., 2022) ha intentado abordar esta cuestión al compilar resultados de radiocarbono de múltiples bases existentes en un *dataset*. Sin embargo, aunque el proyecto se enfocó en la estandarización, la referencia bibliográfica es proporcionada en formato parenético, lo que dificulta la localización de la fuente original. El objetivo entonces no debería ser compilar dataciones en un archivo de texto plano como un csv, sino facilitar la interpolación de datos entre bases mediante peticiones que utilicen un canal de comunicación HTTP a través de una API. Este enfoque representa un ejercicio práctico de interoperabilidad y permite compartir información sin necesidad de repetirla.

En la década de los noventa, la posibilidad de descargar tablas de sitios web para trabajar con ellas de manera local fue una innovación en la historia computacional. No obstante, esta práctica ha perdido preponderancia hoy en día, debido a la disponibilidad de aplicaciones web que permiten consultar, editar, compartir y visualizar información directamente en internet. Estas aplicaciones ofrecen una solución más dinámica y accesible para el manejo de datos en tiempo real.

El caso de ArqueoData

ArqueoData es un proyecto mexicano para la recopilación de dataciones radiocarbono de sitios arqueológicos, principalmente pensada para Latinoamérica. Este surgió como una herramienta de investigación creada para una tesis de licenciatura en

arqueología, del autor de este trabajo, en la Escuela Nacional de Antropología e Historia de la Ciudad de México (Alcántara, 2021).

El proyecto ArqueoData inició como una hoja de cálculo que tenía como fin acumular una cantidad de dataciones radiocarbón y llevar una cuantificación de los lugares datados, su localización, los arqueólogos responsables, los laboratorios donde se realizaron los estudios, entre otras variables. Conforme avanzó el proyecto, la base de datos se convirtió en algo más que una cuantificación de información, pues se comenzaron a trazar constantes en los trabajos de datación, las áreas más estudiadas, las redes intelectuales, el tratamiento de la información por los autores, entre otras. Dicha información resultó fructífera para el trabajo de tesis, lo que planteó nuevas interrogantes que se trasladaron al plano de estudio antropológico, histórico y arqueológico.

Al percatarme de que este tipo de herramientas eran proyectos comunes en otros países, e identificar que en México y el resto de Latinoamérica son escasos, me di la tarea de crear una base de datos pública en la web, gestionando el servidor, la programación y la alimentación del proyecto de forma individual, con el asesoramiento de mi director y mis asesores de tesis, así como el apoyo en materia de radiocarbón del Laboratorio de Espectrometría de Masas con Aceleradores de la Universidad Nacional Autónoma de México (Solís *et al.*, 2014).

Además de la recopilación de información, este proyecto busca crecer y ofrecer una herramienta a los investigadores en materia arqueológica, para concentrar la información de los trabajos de investigación, compartir resultados e incluso ofrecer la calibración desde la misma aplicación web.

La aplicación se construyó mediante una arquitectura de *single page application* (SPA)¹⁶, utilizando el lenguaje JavaScript (VueJS2) y comunicándose con un servicio web programado en PHP, a través del Framework Laravel 8. La presencia en internet, en esta forma de sitio SPA, repercute en el primer principio FAIR, *findability*, pues la aplicación web tiene un único nombre y hospedaje, lo que hace su localización en internet más fácil y su consulta más accesible.

ArqueoData usa la interoperabilidad entre sistemas al conectarse a servicios API como las bases de datos de CrossRef¹⁷, Google Books¹⁸ y la calibración de resultados a través del *software* Calib¹⁹. Así mismo, cuenta con su propio servicio de API pública para cualquier otro sistema web o local que quiera consultar o extraer información de la base de datos ArqueoData. Mediante la API se puede consultar la información y las dataciones arqueológicas.

La estructura de ArqueoData

Hasta ahora, ArqueoData se compone de un total de diecinueve tablas para guardar la información proveniente de la datación de cada muestra. Estas se pueden dividir en cuatro grupos de trabajo en

16 Este término designa un sitio web que reescribe dinámicamente una página actual con nuevos datos del servidor, en lugar del método predeterminado de un navegador que carga páginas nuevas completas.

17 Véase <https://www.crossref.org/documentation/retrieve-metadata/rest-api/>

18 Véase <https://developers.google.com/books/docs/v1/using?hl=es-419>

19 Véase <http://calib.org/calib/calib.html>

la base de datos: *samples*, *methods*, *places* y *bibliographic*; a su vez, cada uno se compone de tablas, catálogos (cat) y relaciones (rel).

Samples (muestras): este grupo de tablas guarda la información de la muestra datada y sus características principales; está conformada por las siguientes tablas:

- *samples*
- *samples_labs*
- *samples_provenance*
- *cat_labs*
- *cat_institutions*

Methods (métodos): este grupo de tablas guarda las relaciones entre métodos y técnicas, ya que ArqueoData se planeó para que a futuro no solo se usen dataciones radiocarbono, sino también de otros métodos de datación, como arqueomagnetismo y dendrocronología. Por el momento, todas las muestras forman parte del método radiocarbono y sus técnicas derivadas, como el método de Libby, el conteo de gas proporcional, el LSC y el AMS. El grupo está conformado por las siguientes tablas:

- *cat_methods*
- *cat_techniques*
- *rel_method_technique*

Places (lugares): este grupo de tablas guarda la información de los sitios arqueológicos relacionados con cada muestra y, a su vez, la relación del sitio con su ubicación en estado o provincia, país y continente. Las tablas que conforman este grupo son las siguientes:

- cat_sites
- cat_states
- cat_countries
- cat_continents

Bibliographic (bibliografía): este grupo guarda la información de cada referencia bibliográfica y su relación con cada muestra. Se considera que cada registro en la base de datos debe estar asociado con las referencias bibliográficas donde aparece la datación. Para lograr este objetivo, ArqueoData trabaja con la API de Cross-Ref y Google Books, con el fin de obtener, por medio del identificador DOI y el registro ISBN, para artículos científicos o libros, respectivamente, la ficha bibliográfica de cada referencia; si no se encuentran en estas bases de datos, se da la opción de agregar manualmente la ficha bibliográfica. Las tablas que conforman este grupo son las siguientes:

- cat_references
- cat_doctype
- cat_authors
- cat_journals
- rel_reference_author
- rel_reference_journal
- sample_rel_reference

Este ordenamiento de las tablas permite realizar una búsqueda detallada en cada elemento de la muestra concerniente a datos de fecha, espacio y bibliografía. Sin embargo, sobre la información del contexto arqueológico, aún no se incluye mayor

estandarización de datos que lo identificado como comentario. En ArqueoData se pueden incluir tres comentarios por muestra: uno sobre la proveniencia del material (`samples_provenance.cCommentProv`), otro sobre la muestra en sí (`samples.cComment`) y un tercero que puede ser efectuado por el laboratorio que fechó la muestra (`samples_lab.cCommentLab`). El contexto arqueológico puede estar presente en cualquiera de los dos primeros (figura 2.2).

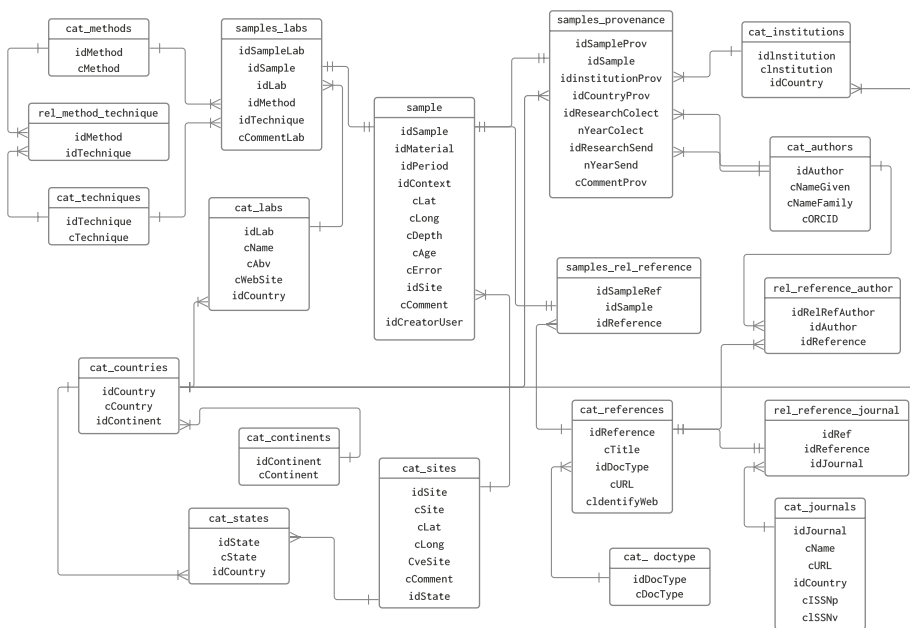


FIGURA 2.2. Modelo entidad-relación de las tablas correspondientes a guardar la información de cada muestra en ArqueoData

FUENTE: elaboración propia.

Todas las muestras radiocarbono de ArqueoData son de sitios mexicanos, en vista de que ninguna otra base proporciona, hasta ahora, un servicio de API pública; sin embargo, ahora cuenta con la base de datos p3k14c de forma interna para su consulta pública. Lo anterior requirió una tarea de estandarizar los datos de esta base global con los de ArqueoData. Aun así, hay muchos datos en el aire que se pueden complementar en un futuro, pensando en el trabajo en comunidad y en la interoperabilidad al relacionar esas dataciones con sus iguales en otras bases de datos.

El trabajo colaborativo en ArqueoData

Uno de los objetivos comunes de todos los proyectos de bases de datos en radiocarbono es servir como herramientas para guardar, consultar, compartir y comparar información. En ArqueoData no solo es posible consultar la información de diversos sitios arqueológicos, también se ha diseñado como una pequeña red social, en la cual los usuarios pueden crear colecciones y, en cada colección, guardar dataciones de un trabajo que estén desarrollando. Dicha colección puede arreglarse por secuencias y fases, anexas archivos y enlaces, y compartirse con los usuarios.

Cada colección recibe un URL único en ArqueoData para poder encontrarlo fácilmente. De esta forma, no solo se presentan resultados en una base de datos, sino que la interacción entre la base de datos y los usuarios permite generar una comunidad abierta al intercambio de ideas e información en el espacio virtual, además de poner en contacto a investigadores o gente interesada en el tema. De esta forma, se busca que la reutilización de la información sea más accesible y fácil de localizar, así como expuesta a

una mayor parte de la comunidad académica, cumpliendo el objetivo de reusabilidad de la información de los principios FAIR.

Una propuesta de API pública y su utilidad

Mediante el servicio de consulta pública de ArqueoData es posible acceder a todos los registros y sus relaciones en la base de datos. Las peticiones se realizan por el protocolo HTTP GET, siguiendo algunas reglas en las URL de cada consulta; finalmente, se responde de la petición en un formato JSON²⁰. Las respuestas del servidor están ordenadas bajo la siguiente estructura:

```
{
  element: "Tipo de elemento",
  message: "Mensaje",
  status: "Estatus de la petición",
  type: "Tipo de búsqueda",
  version: "Versión de la API que se utiliza"
}
```

En este apartado me interesa reflexionar sobre la importancia de estas herramientas de interoperabilidad en el campo de las bases de datos web, en especial de las bases radiocarbono, y su relación con este proyecto.

Como argumenté, la interoperabilidad en este proyecto la visualizo como una forma de entablar un diálogo entre todas las bases de datos existentes. ArqueoData se centró en los sitios

²⁰ Para una revisión más técnica sobre las respuestas que ofrece la api pública del proyecto, véase el sitio web de ArqueoData (<https://www.arqueodata.com>).

arqueológicos mexicanos, y es una base de datos pionera en el país; sin embargo, hay bases tan grandes, como las de Estados Unidos, Inglaterra o la península ibérica, en ocasiones con información repetida, que la comunicación entre estas, en términos de interoperabilidad, permitiría conectar la información de varios proyectos, con la finalidad incluso de refinar los metadatos presentados.

Por otro lado, se encuentra el trabajo en comunidad, el cual concibo más en el ámbito del registro de dataciones y la difusión de estas, ya que uno de los problemas con el que inició el proyecto fue la gran cantidad de información depositada en tablas de artículos y libros científicos; sin embargo, esta es una información que no se difunde entre la comunidad académica y lleva a vacíos a la hora de iniciar un proyecto de investigación. Ubicar de una forma más rápida la información de dataciones, mediante métodos cuantificables, puede dar indicios para nuevas propuestas de estudios y formas de acción en la práctica arqueológica al definir un proyecto; obviamente la datación por radiocarbono solo es una pieza de todo el reensamblaje de una sociedad pretérita.

Perspectivas de un futuro cercano

ArqueoData es una herramienta en proceso dentro del abanico de prácticas de las HD para la investigación. Este es un campo que en Latinoamérica debemos seguir adaptando a nuestras necesidades; de esta forma, podemos generar nuestros propios recursos, pero siempre tomando en cuenta la perspectiva global y considerando que siempre será mejor trabajar en multidioma, para que las herramientas también se conozcan fuera del ambiente latinoamericano.

A lo largo de este capítulo me enfoqué en esclarecer las bondades que la web, una aplicación, la interoperabilidad a través de una API y el trabajo en comunidad tienen para la creación de estas herramientas de investigación. No obstante, también es necesario puntualizar los retos que conllevan este tipo de proyectos; en el caso de ArqueoData, es una iniciativa individual a partir de un proyecto personal que no está inscrito en ningún ambiente institucional. Esto lleva a cuestionar la duración del soporte de esta base de datos y las acciones que a futuro deben tomarse para que el proyecto pueda seguir funcionando.

Además, debemos pensar que, para las humanidades y, en específico para la historia y la arqueología, más allá de los métodos cuantitativos de una base de datos, nos permite realizar preguntas de investigación cualitativas al indagar en las presencias y las ausencias de información, pues la finalidad de los científicos sociales y humanistas no es la aplicación web en sí —aunque es una parte importante—, pero el tratamiento de la información con la finalidad de crear narrativas sobre la sociedad y el pasado es lo que mueve la información incluso en una base de datos.

Aún queda un largo camino por recorrer para que el trabajo en comunidad sea una realidad, no solo por parte de las herramientas digitales, sino también del ámbito individual de sobrepasar nuestras fronteras y hacer realidad un trabajo conjunto entre la academia con mayor altura de miras.



Referencias

- Alcántara, L. A. (2021). *Entre Tepalcates: Hacia una comprensión historiográfica de la datación arqueométrica en la arqueología mexicana (1950-2020)* [tesis de licenciatura]. Escuela Nacional de Antropología e Historia.
- Ashmore, P. J., Cook, G. T. y Harkness, D. D. (2000). A radiocarbon database for Scottish archaeological samples. *Radiocarbon*, 42(1), 41-48. <https://doi.org/10.1017/S0033822200053030>
- Baca, M. (Ed.). (2016). *Introduction to metadata*. Getty Publications. <http://www.getty.edu/publications/intrometadata>
- Banks, M. (2008). *On the way to the web: The secret history of the internet and its founders*. Apress. <https://doi.org/10.1007/978-1-4302-0870-9>
- Bates, C. (2003). *XML in theory and practice*. John Wiley & Sons, Ltd.
- Bird, D., Miranda, L., Vander Linden, M., Robinson, E., Bocinsky, R. K., Nicholson, C., Capriles, J. M., Finley, J. B., Gayo, E. M., Gil, A., d'Alpoim Guedes, J., Hoggarth, J. A., Kay, A., Loftus, E., Lombardo, U., Mackie, M., Palmisano, A., Solheim, S., Kelly, R. L. y Freeman, J. (2022). P3k14c, a synthetic global database of archaeological radiocarbon dates. *Scientific Data*, 9(1), 27. <https://doi.org/10.1038/s41597-022-01118-7>
- Bresciano, J. A. (2010). *La historiografía en el amanecer de la cultura digital: Innovaciones metodológicas, discursivas e institucionales*. Ediciones Cruz del Sur. https://www.academia.edu/453401/La_Historiograf%C3%ADa_en_el_amanecer_de_la_cultura_digital_Innovaciones_metodol%C3%B3gicas_discursivas_e_institucionales.
- Canadian Archaeological Radiocarbon Database (CARD). (2015). *Welcome to card 2.0*. <https://www.canadianarchaeology.ca/>

- Cervantes Maceda, H., Velasco-Elizondo, P. y Castro Careaga, L. (Eds.). (2016). *Arquitectura de software. Conceptos y ciclo de desarrollo*. Cengage Learning Editores.
- D'Amaro, F. (2019). *La memoria digital de España: El archivo web como nueva fuente para la historia del presente*. Biblioteca Virtual Miguel de Cervantes. <http://rua.ua.es/dspace/handle/10045/95804>
- del Rio Riande, G. (2014). *¿De qué hablamos cuando hablamos de humanidades digitales?* I Jornadas Nacionales de Humanidades Digitales. <https://www.aacademica.org/jornadasaahd/3>
- Gayol, V. y Melo Flórez, J. A. (2017). Presente y perspectivas de las humanidades digitales en América Latina. *Mélanges de la Casa de Velázquez*, 47(2), 281-284. <https://doi.org/10.4000/mcv.7907>
- Gilliland, A. J. (2016). Setting the stage. En M. Baca (Ed.), *Introduction to Metadata* (3rd. Edition). Getty Publications. <http://www.getty.edu/publications/intrometadata/setting-the-stage/>.
- Gulliksen, S. (1983). Radiocarbon database: A pilot project. *Radiocarbon*, 25(2), 661-666. <https://doi.org/10.1017/S0033822200005993>
- Haider, R. (2021). *Web API development with Python*. CloudBytes.
- Harris, E. (1979). *Principios de estratigrafía arqueológica*. Crítica.
- Hendler, J. y Berners-Lee, T. (2010). From the semantic web to social machines: A research challenge for AI on the world wide web. *Artificial Intelligence*, 174(2), 156-161. <https://doi.org/10.1016/j.artint.2009.11.010>
- Infraestructura de Datos Espaciales de Investigación Arqueológica (IDEArq). (2019). *IDEArq: Infraestructura de Datos Espaciales de Investigación Arqueológica*. <http://www.ideaarqueologia.org/>

- Isaac, A. y Haslhofer, B. (2013). Europeana linked open data-data.europeana.eu. *Semantic Web*, 4(3), 291-297. <https://doi.org/10.3233/SW-120092>
- Jiménez Badillo, D. (Ed.). (2017). *Arqueología computacional: Nuevos enfoques para la documentación, análisis y difusión del patrimonio cultural*. Instituto Nacional de Antropología e Historia.
- Jiménez Badillo, D. y Gándara, M. (Eds.). (2016). *El patrimonio cultural y las tecnologías digitales. Experiencias recientes desde México*. Instituto Nacional de Antropología e Historia.
- Kahler, M. E. (1978). Nettie Lee Benson and the Latin American collection. *Latin American Research Review*, 13(1), 173-177.
- Lamprecht, A. L., García, L., Kuzak, M., Martínez, C., Arcila, R., Martín del Pico, E., Domínguez del Ángel, et al. (2020). Towards FAIR principles for research software. *Data Science*, 3(1), 37-59. <https://doi.org/10.3233/DS-190026>
- Libby, W. (1955). *Radiocarbon dating*. University of Chicago Press.
- López Aguilar, F. (2012). Arqueología y complejidad: Sobre la naturaleza del dato arqueológico. En M. Cervantes (Ed.), *17 temas de la antropología mexicana* (pp. 237-264). Academia Mexicana de Ciencias Antropológicas.
- López Hernández, H. (2018). *En busca del alma nacional: La arqueología y la construcción del origen de la historia nacional en México (1867-1942)*. Instituto Nacional de Antropología e Historia.
- López Zahar, J. G. (2020). *La biblioteca del Museo Nacional de México y la práctica del coleccionismo de libros (1825-1867)* [tesis de licenciatura]. Universidad Nacional Autónoma de México.
- Melo Flórez, J. A. (2011). Historia digital: La memoria en el archivo infinito. *Historia Crítica*, 43, 82-103.

- Moffett, J. C. y Webb, R. E. (1983). Database management systems, radiocarbon, and archaeology. *Radiocarbon*, 25(2), 667-668. <https://doi.org/10.1017/S0033822200006007>
- Mora-Castillo, J. A. (2016). Serialización/deserialización de objetos y transmisión de datos con JSON: Una revisión de la literatura. *Revista Tecnología en Marcha*, 29(1), 118-125. <https://doi.org/10.18845/tm.v29i1.2544>
- Morales, J., Ángeles Jiménez, P., Molina Salinas, C., Miranda, E. y Ramírez, V. (2021). Apretando el paso: Un modelo de datos para el Repositorio Digital del Patrimonio Cultural de México. En D. Jiménez Badillo (Ed.), *Patrimonio digital. Métodos computacionales y medios interactivos para estudiar y divulgar el patrimonio cultural*. (pp. 203-214). Instituto Nacional de Antropología e Historia.
- Priani Saisó, E. (2019). Codificación y buenas prácticas: Crítica a la delimitación de las humanidades digitales en América Latina. *Relaciones: Estudios de Historia y Sociedad*, 40(158), 129-144. <https://doi.org/10.24901/rehs.v40i158.343>
- Pons, A. (2013). *El desorden digital. Guía para historiadores y humanistas*. Siglo XXI Editores.
- Ramsey, C. B. (1995). Radiocarbon calibration and analysis of stratigraphy: The OxCal Program. *Radiocarbon*, 37(2), 425-430. <https://doi.org/10.1017/S0033822200030903>
- Ramsey, C. (2010). *IntChron*. <https://intchron.org>
- Ríos-Hilario, A., Martín-Campo, D. y Ferreras-Fernández, T. (2012). Linked data y linked open data: Su implantación en una biblioteca digital. El caso de Europeana. *Profesional de la Información*, 21(3). <https://doi.org/10.3145/epi.2012.may.10>
- Rivera, F. L. O. (2008). *Base de datos relacionales*. Instituto Tecnológico Metropolitano.

- Rodríguez Matta, P. (2009). *Asimetrías en un proyecto estatal de reproducción digital: El caso de memoria chilena* [tesis de maestría]. <http://repositorio.flacsoandes.edu.ec/handle/10469/894>
- Russell, I. G. (2018). La institucionalización de las humanidades digitales. En I. Galina (Ed.), *Humanidades digitales 1: Recepción, institucionalización y crítica* (pp. 16-37). Bonilla Artigas Editores.
- Sandoval García, G. (2019). *An epistemology of archaeological excavation. A comparative study of British field practice* [tesis doctoral]. University of York.
- Santoyo, A. S. y Martínez, E. M. (2003). *La brecha digital: Mitos y realidades*. Universidad Autónoma de Baja California.
- Schroeder, R. y Brügger, N. (Eds.). (2017). *The Web as History: Using Web Archives to Understand the Past and the Present*. University College London Press. <https://doi.org/10.14324/111.9781911307563>
- Sivarajah, U., Irani, Z. y Weerakkody, V. (2015). Evaluating the use and impact of Web 2.0 technologies in local government. *Government Information Quarterly*, 32(4), 473-487. <https://doi.org/10.1016/j.giq.2015.06.004>
- Smith, K. (2006). Simplifying Ajax-style web development. *Computer*, 39(5), 98-101. <https://doi.org/10.1109/MC.2006.177>
- Solís, C., Chávez-Lomelí, E., Ortiz, M. E., Huerta, A., Andrade, E. y Barrios, E. (2014). A new AMS facility in Mexico. *Nuclear Instruments and Methods in Physics Research Section B: Beam Interactions with Materials and Atoms*, 331, 233-237. <https://doi.org/10.1016/j.nimb.2014.02.015>
- Tobón Restrepo, I. (2022). *Del fichero al ciberespacio: Historias orales de la transformación digital del patrimonio cultural colombiano y su llegada a la web* [tesis de maestría]. Universidad de los Andes. <https://repositorio.uniandes.edu.co/handle/1992/64315>

- Tschauner, H. (1985). La tipología: ¿Herramienta u obstáculo? La clasificación de artefactos en arqueología. *Boletín de Antropología Americana*, 12, 39-74.
- Wilkinson, M. D., Dumontier, M., Aalbersberg, I. J., Appleton, G., Axton, M., Baak, A., Blomberg, N., Boiten, J.-W., da Silva Santos, L. B., Bourne, P. E., Bouwman, J., Brookes, A. J., Clark, T., Crosas, M., Dillo, I., Dumon, O., Edmunds, S., Evelo, C. T., Finkers, R. y Mons, B. (2016). The FAIR guiding principles for scientific data management and stewardship. *Scientific Data*, 3(1), 160018. <https://doi.org/10.1038/sdata.2016.18>
- Ziólkowski, M. S., Pazdur, M. F., Krzanowski, A. y Michczyński, A. (2015) *Andes: Radiocarbon Database for Central Andes*. <https://andesc14.pl/en/>

Código y reflexividad en la investigación histórica de repositorios digitales: exploración de las ruinas de GeoCities*

* Agradezco la lectura de una primera versión de este texto a Emiliano Calomarde y a los integrantes del proyecto de investigación en el que participo: “Comunidades virtuales: historia, prácticas e imaginarios” (Universidad Nacional de Mar del Plata, Argentina).

Para citar este capítulo:

<http://doi.org/10.51573/Andes.9789587987362.9789587987379.97895879878673>

NICOLÁS QUIROGA

Introducción

{ Este capítulo utiliza una técnica denominada *análisis exploratorio de datos*, para inspeccionar una copia enorme de las páginas web puestas en línea por una empresa muy conocida en los primeros años de la web: GeoCities. Mi propósito es avanzar en la construcción de un corpus de fuentes documentales para el análisis histórico; para ello, trataré de comprender contenidos y contextos de un repositorio incógnito al procesar computacionalmente los datos. Este texto analiza esa exploración, pero no tiene a GeoCities como objeto de investigación, ni refiere a un campo de investigación en particular: trata sobre los comienzos de cualquier investigación en la que haya que trabajar con lo que Brügger (2012) denomina *web guardada*¹.

1 Niels Brügger distingue tres tipos de materiales: *digitalizado* (“previamente analógico”), *digital* (“nunca ha existido en otra forma que no sea la digital”) y

El objetivo es exponer los primeros pasos del análisis a partir de una pregunta básica y funcional: ¿cómo podrían los historiadores reconocer fuentes documentales en este repositorio? La exploración que desarrollo consiste en “probar” algunos procedimientos conocidos en las humanidades digitales (nubes de palabras, modelado de tópicos) para mapear un repositorio sin curaduría, heterogéneo y enorme. Sigo la estela de un texto seminal de Carlo Ginzburg, “Conversar con Orión” (2004), en el que, a partir de las consultas realizadas al *software* que gestionaba el catálogo de la biblioteca de la Universidad de California en Los Ángeles (UCLA), el reconocido historiador italiano refiguraba una pregunta raigal de la interpretación histórica: ¿cómo abordar la desproporción entre “una ingente masa documental y el limitado tiempo humano”? El auge de las computadoras le permitió a Ginzburg interactuar con el catálogo de la biblioteca y utilizar el “recurso al azar, orientado por conjeturas razonables”, para hallar temas de futuras investigaciones.

Mi propuesta puede pensarse como una etapa de reconocimiento en la investigación (Webster, 2021), por medio del análisis exploratorio de datos (Tukey, 1977; Unwin, 2010) y algunas técnicas que abarcan tres perspectivas diferentes, centradas en el usuario, en los datos y en las visualizaciones (Holzmann y Nejd, 2021), como modos de variar la escala de “la vista” de la colección. En “Conversar con Orión”, Ginzburg (2004) se declaró “en cuanto al plano

.....
renacido-digital (“se ha recopilado y conservado y que, en cierta medida, se ha modificado en este proceso”). El repositorio con el trabajamos es un archivo web en ese último sentido: materiales de la web que han sido guardados. Según Brügger (2012), el archivado web no produce copias, sino versiones de “un original que se pierde para siempre”, y tiene al menos tres características: incompleto, escasamente documentado e inconsistente en términos de tiempo y espacio.

de la informática, un analfabeto” y apenas reproducía unas pocas instrucciones del *software* que utilizó, pero dos décadas después las posibilidades que habilitan técnicas y métodos computacionales están más a la mano para quienes se interesen por utilizarlas; por eso, este trabajo introduce complejos fragmentos de código, de modo que la reflexión historiográfica surja desde el trabajo con el código y no desde la experiencia del usuario de la interfaz de consulta de una aplicación, como sucede en el texto de Ginzburg².

Entiendo por *código* una práctica que contiene la escritura de un lenguaje de programación (en este caso Python), pero no se limita a esta³. “Codificar” (junto con “construir” y “hackear”) es una actividad fundamental para las humanidades digitales

- 2 Winters y Prescott (2019) han mostrado que la búsqueda por palabras clave es incierta frente a la abundancia, por lo que proponen utilizar nuevas técnicas y enfoques, como los de las humanidades digitales y la historia digital, frente a repositorios enormes (su ejemplo son los documentos provenientes del caso *WikiLeaks*).
- 3 “Código es un término general que designa una amplia variedad de lenguajes de programación concretos y prácticas asociadas. Cuando queremos observar el código, vemos una serie de perspectivas y escalas diferentes dependiendo del tipo de código que estemos viendo (ensamblador, C++, Pascal), de su estado (fuente, compilado, desensamblado), ubicación (incrustado, sistema, aplicación) o su forma (textual, visual, mapeado como un gráfico) [...] Llama la atención la capacidad del código para actuar a la vez como actor que ejecuta acciones sobre los datos y como recipiente que los contiene dentro de sus límites” (Berry, 2011, p. 33). Para Hayles (2005), el código es “compañero” del lenguaje e “implica una asociación entre los seres humanos y las máquinas inteligentes en la que las prácticas lingüísticas de cada uno influyen en el otro y se interpenetran” (p. 61). Para otros autores, como Kittler (2008), la relación con el lenguaje es “insoluble”, pero programar implica un distanciamiento progresivo con respecto al lenguaje (hace que la cabeza del programador se “vacíe de palabras”). Sobre el concepto de código, además de los autores citados,

(Ramsay y Rockwell, 2012), y como técnica implica práctica y especulación, como bien ha señalado Chun (2011), siguiendo a P. Ricouer⁴. Mark Marino (2020), uno de los autores más reconocidos de los llamados *critical code studies*, sostiene que tenemos que aprender a leer críticamente el código, sabiendo que

su forma final no es el resultado de la certeza matemática, sino del acopio de conocimientos y convenciones culturales (culturas del código y lenguajes de codificación), prisas y perspicacia, inspiraciones y observaciones, evoluciones y adaptaciones, retórica y razón, paradigmas del lenguaje, avances en el enfoque y fracasos en la conceptualización. (p.8)

A la luz de estas referencias, codificar implica tanto el análisis de los resultados producidos por la ejecución de programas o aplicaciones como la vigilancia epistemológica sobre el trabajo con el código. Utilizo *reflexividad* en ese sentido (Piovani, 2018), con la intención de recordar que las técnicas y métodos computacionales no son solo agentes que devuelven resultados, sino un trabajo de *desplazamientos* (Kittler, 2008) que realizan humanos y máquinas, y que como producto “técnico y social, y material y simbólico

.....

véanse Ilcid (2022), Cramer (2008), Krysa y Sedek (2008). [Salvo indicación específica, las traducciones de las citas textuales son mías.]

- 4 “El código no siempre o automáticamente hace lo que dice, sino que lo hace de una manera astuta y especulativa en la que se crean tanto el significado como la acción. Conlleva la posibilidad de desviación: nuestra creencia de que los compiladores simplemente expanden comandos de nivel superior —en lugar de alterar o insertar otros comportamientos— es simplemente eso, una creencia, una de las muchas que sustentan la informática como tal” (Chun, 2011, p. 25).

simultáneamente” (Berry, 2011, p. 36) requiere de “pensar en lo que se hace” (Hidalgo, 2006, p. 48)⁵. Algunos autores han sugerido que, a diferencia del uso de aplicaciones o plataformas para analizar datos (que acepta algo de opacidad en los procesamientos en sus operaciones), codificar es una instancia analítica fundamental para las humanidades (Morgan, 2018; Venturini *et al.*, 2021; Rawson y Muñoz, 2019; Van Es *et al.*, 2018).

GeoCities y la investigación sobre sus archivos guardados

GeoCities (1994-2009) fue una empresa que permitía publicar páginas web. Su plataforma estuvo pensada como una ciudad: tenía barrios y suburbios temáticos, los cuales agrupaban con dificultad creciente tópicos y comunidades. En 1999, en el cénit de su evolución, fue adquirida por Yahoo! y permaneció, en caída, hasta el 2009. Durante este año, Archive Team pudo hacer una copia “pirata” bastante amplia de la ciudad y la publicó. Millones de páginas web se conservan en ese repositorio (cerca de dieciocho millones de archivos), el cual solo puede recorrerse como un conjunto de carpetas y archivos. La incertidumbre del alcance del salvataje colectivo, el caótico orden de las ruinas y la casi absoluta imposibilidad de replicar la navegación por sus barrios nos deja frente a un conjunto de datos inmenso, del que no sabemos casi nada. Sin embargo, en

5 El concepto de reflexividad tiene un andarivel muy relevante en la conexión arte y tecnologías, aunque escapa a los objetivos de este texto. Véase Soon (2016) y su *reflexive coding practice*.

la casi indescifrable organización de la copia de urgencia, algunos subsitios se destacan: los que se agrupan por dominios y subdominios de países o idiomas (*ar, br, de, asia, espanol*, etc.). Este capítulo se ocupa de indagar esos “restos” salvados, que estuvieron por “fuera” de la organización barrial. De esos directorios, presentaré datos generales de las secciones de Argentina (ar.geocities.com), Brasil (br.geocities.com) y México (mx.geocities.com), y seguiré en detalle la primera⁶.

La historia de GeoCities, desde sus orígenes hasta su caída, pasando por la compra de la empresa por Yahoo!, en el momento en el que era una las páginas web más visitadas, ha sido narrada muchas veces (Wikipedia, s. f.; Major y Gomes, 2021; Howard, 2019; Milligan, 2019). Apenas sucedió el cambio de manos, se decidió desactivar la formación de barrios y suburbios temáticos, administrados por líderes comunitarios⁷. Tal vez allí radica el origen de los subdominios locales e idiomáticos de GeoCities, aunque no resulta fácil conocer ese aspecto relevante, que en parte determina el tipo de “comunidad” conformada (las páginas que analizo no están “ubicadas” por afinidad y solo las une su pertenencia a la empresa)⁸.

6 Utilizo *repositorio, copia, backup* y *sitios web* como sinónimos, aunque no lo sean, para referirme al material colectado de dos proyectos de reconstrucción de GeoCities.

7 Para un listado de barrios, véase <http://www.bladesplace.id.au/geocities-neighborhoods-suburbs.html>; para *snapshot* de líderes comunitarios en Archive.org, véase <https://web.archive.org/web/19990128104452/http://www7.geocities.com/Athens/leader.html>.

8 Katie Mackinnon (2022) ofrece detalles sobre el impacto de los cambios en la estructura y gobernanza de GeoCities, al analizar mensajes que los propietarios de páginas web escribieron cuando abandonaron la plataforma (“elegías”).

La decisión de borrar GeoCities fue, como bien dice Jason Scott (2010), un episodio trágico para el patrimonio cultural, pero sobre todo para la muchedumbre que había construido esa comunidad:

[Los sitios web y los servicios de alojamiento] representan innumerables horas de escritura, de edición, de pensamiento, de creación. Representan tiempo y representan los pensamientos y los sueños de personas ahora mucho más viejas, o fallecidas. Aquí hay historia. Historia real, honesta, verdadera. Así que el *Archive Team* hizo lo que pudo, al igual que otros equipos independientes de todo el mundo, y se salvó una parte de *Geocities*.

El Archive Team se define como un colectivo informal de archiveros, programadores y escritores dedicados a salvar el patrimonio cultural (Archive Team, s. f.; Odgen, 2021), quienes durante el 2009 lograron copiar una considerable cantidad de páginas web de GeoCities. Debido a cómo fue realizada y luego distribuida (a través de un archivo *torrent*), y porque no hay datos sobre la plataforma, no puede medirse la representatividad de la copia, ni puede evaluarse la fidelidad del proceso de copiado.

No fue la única empresa que se propuso reconstruir esa colmena. También tienen ese objetivo The Geocities Gallery⁹, One Terabyte of Kilobyte Age¹⁰, Cameron's World¹¹, Deleted City¹² y

9 Véase <https://geocities.restorativland.org/>.

10 Véase <https://anthology.rhizome.org/one-terabyte-of-kilobyte-age>.

11 Véase <https://www.cameronsworld.net/>.

12 Véase <http://deletedcity.net/>.

OoCities¹³. Son muchas las diferencias entre los repositorios. Para esta investigación, unifiqué la copia de Archive Team con la de OoCities, dado que ambas proponen reconstrucciones totales y permiten el acceso a las páginas (aunque poseen diferencias en el tratamiento y la visualización de lo que guardan).

En la copia de Archive Team, uno de los subdirectorios (“sub-sites1”) guarda grupos de páginas web de subdominios o dominios asociados a GeoCities (“categorías de locación o lenguaje” es el modo que OoCities los denomina), Argentina (“ar.geocities”), Asia (“Asia.geocities.com”), español (“espanol.geocities.com”), entre muchos otros. Como indiqué, no resulta fácil confirmar que esas secciones hayan sido consecuencia de cambios decididos por Yahoo! luego de la adquisición (eliminación de la forma barrial), pero es la hipótesis más económica¹⁴. El propósito de mapear un conjunto de páginas muy heterogéneo debe mantenerse lejos de conclusiones o conclusiones parciales, lejos de un cierre de sentido sobre esos restos del pasado: si están juntos en un directorio es debido a políticas de una empresa comercial, dos jornadas de “raspado” (OoCities, Archive Team) y mi decisión de unificar lo obtenido.

Este acervo es fundamental para escribir la historia de la web temprana en América Latina; eso significa que, más allá de los contenidos escritos en esas páginas web personales, también las

13 Véase <http://www.oocities.org/>.

14 Algunos sitios web que quedaron guardados en la copia poseen páginas asociadas a fechas previas a ese cambio; algunos sitios relacionados con Argentina, guardados en subdirectorios barriales, tienen páginas de fechas posteriores. Por otro lado, en las secciones que voy a revisar, además de páginas en español (*ar*, *mx*) y portugués (*br*), hay también algunos sitios web en inglés y (pocos) en otros idiomas.

tecnologías de la web, la expansión de internet, el diseño y la historia de los medios del periodo son aspectos relevantes para comprender el posible alcance de esta copia como fuente histórica¹⁵. Por otro lado, esto implica la necesidad de comprender problemas específicos del tipo de documento con el trabajamos: las páginas web guardadas, entendidas como unidades de análisis, contenedores de objetos igual de relevantes y elementos de una ecología o de una “esfera” (Brügger, 2009, 2018; Foot y Schneider, 2006, 2010; Major, 2021; Helmond, 2017).

La literatura que investiga sobre el repositorio de Archive Team abarca áreas como historia de la web y sus tecnologías (Reynolds y Hallinan, 2021), literatura (Baker, 2020) y archivística (Major y Gomes, 2021). Uno de los proyectos más destacados es Archives Unleashed Project¹⁶, en especial las contribuciones de Ian Milligan (2016, 2019, 2020) y su equipo (Ruest *et al.*, 2020; Fritz *et al.*, 2021). El proyecto Archives Unleashed viene desarrollando herramientas para el análisis de la copia de GeoCities. Sus enfoques han ido cambiando con el tiempo: conviven en una línea de trabajo muy original —que incluye *datathons* y un programa de mentorías de proyectos de investigación— los usos de complejas arquitecturas como Scala y Apache Spark y herramientas simples como *WARC*, la biblioteca *Pandas* para Python y *notebooks*. En los últimos tiempos, su enfoque se basa en curar parte de la copia de

15 Tomo como referencia el libro *The past web: Exploring web archives* (Gomes *et al.*, 2021), un excelente punto de partida para comprender aspectos fundamentales en el uso de fuentes para la investigación sobre el pasado de la web.

16 Véase <https://archivesunleashed.org/>.

GeoCities —“derivados” (Ruest, 2021)— y trabajar con *notebooks* para la exploración de esos datos (Deschamps *et al.*, 2019).

Análisis exploratorio de datos¹⁷

Mi interés en el análisis exploratorio de datos proviene de algunas lecturas que cuestionan las relaciones dadas por sentado entre técnicas estadísticas y visualizaciones en la construcción de los datos (Taylor y Tilton, 2019, 2015), pero también del seguimiento de las actividades de Archives Unleashed, en particular sus *notebooks* con análisis del *back-up* de GeoCities¹⁸. Filtrar, analizar, agregar y visualizar son verbos en esta instancia, pero no “pasos”; porque el trabajo no es “programable”. El reconocimiento de formatos e interfaces y la “familiaridad” con el repositorio suceden durante el proceso de indagación y obligan a rehacer los procedimientos.

17 El análisis exploratorio de datos no consiste en una lista cerrada de procedimientos. Algunas otras exploraciones pueden servir para dimensionar la variedad de estrategias utilizadas. Para el GitHub de Archives Unleashed, véanse https://github.com/archivesunleashed/notebooks/blob/main/Parquet%20Examples/parquet_pandas_example.ipynb y https://github.com/archivesunleashed/notebooks/blob/main/geocities/geocities_binary_information.ipynb. Una búsqueda en Kaggle devuelve miles de *cuadernos*; véase <https://www.kaggle.com/search?q=exploratory+data+analysis>. Así mismo, en Nbviewer hay un visor de cuadernos Jupyter; véanse https://nbviewer.org/github/TanuN-Prabhu/Python/blob/master/Exploratory_data_Analysis.ipynb y <https://nbviewer.org/github/donnemartin/data-science-ipython-notebooks/blob/master/kaggle/titanic.ipynb>.

18 Véase <https://github.com/archivesunleashed/notebooks>

El análisis de páginas web es un campo que paulatinamente adquiere densidad, en especial debido a su importancia para quienes estudian la historia de la web temprana. La página web como fuente histórica es un tema fundamental en esta investigación. Mi trabajo de archivo consiste en un prolongado paseo por la heurística y la crítica de las fuentes, con herramientas poco utilizadas en la investigación histórica y con “documentos” que desafían el sentido común forjado por mi trato con papeles y libros¹⁹. A medida que el trabajo va dando frutos, resulta mucho más fácil advertir problemas comunes por “debajo” de las diferencias entre una página y una página web (Milligan, 2019). Pero en el momento del análisis exploratorio, objetos digitales, formatos, protocolos, lenguajes y, en particular, esos desplazamientos de sentido a los que llamamos *conversiones* ocupan la mayor parte del tiempo²⁰.

- 19 Uso comillas porque el término “documentos” oculta particularidades del tipo de materialidad de las páginas web guardadas y el peso de las transformaciones en el trabajo con código. El primer desplazamiento, decisivo, consiste en que esas páginas están desactivadas al haber perdido su referencia con el servidor. Pueden abrirse en un navegador, pero algunos objetos que apuntan al servidor ya no funcionan y otros nunca fueron guardados. Tengo que estar atento a páginas iniciales (“index.htm”, “principal.htm”, “home.htm”), por lo que recorrer directorios tampoco es simple. OoCities recreó la ciudad inyectando código en cada una de las páginas web y, de ese modo, es posible “abrirlas”. Para este trabajo también tuve que eliminar el código fuente de OoCities, no de las fuentes, pero sí en el análisis.
- 20 Conversiones y transformaciones de formatos “originales” a formas de representación resueltas en el marco de la exploración de datos, a partir de decisiones a nivel de código, lenguaje y visualidad. En Python debo moverme del HTML a diccionarios, listas y tuplas, al igual que tratar con formatos de Pandas (*dataframes*), además de formatos como el csv y el JSON. En la pugna por “ver”, hay que anotar lo que se pierde y deriva en cada paso.

Cantidades

Traté de identificar cantidades y aspectos relacionados con la estructura de la web guardada (aproximadamente, setecientos ochenta mil archivos). Decidí visualizar esa información agrupando los archivos `HTM` y `HTML`, así como los archivos `JPG` y `JPEG` (a partir de aquí, `HTMX` y `JPGX`, respectivamente), y filtrando los tipos de archivos por cantidad, de modo que ese valor permita mantener la escala del gráfico (figuras 3.1, 3.2 y 3.3)

Aunque no conviene utilizar este tipo de aproximación, debido a que sacrifica valores pequeños en favor de la visibilidad del gráfico, los cálculos mostraron tanto la fuerza de la dupla `HTMX` + imágenes como el indicio de los archivos `GIF` y `SWF`: el estudio de objetos y tecnologías puede indicar usos de época y lo que habilitan en la comunicación, y son sólidos los abordajes que proponen articular su diseminación con otros medios y consumos (Anker-son, 2012; Miltner y Highfield, 2017; Owens y Thomas, 2019).

Por otro lado, procesé los archivos con la biblioteca de Python —Beautiful Soup—, la cual utilicé para extraer datos de la estructura de los archivos `HTMX`²¹. Recorrí los directorios, esta vez buscando archivos `HTMX` con nombre “index”, “home” y “principal”²². De ese subconjunto, observé la presencia de algunas etiquetas: tablas e *iframes* (para comprender si la página principal funcionaba como marco o no) y metadatos en los encabezados, para poder

21 Beautiful Soup es un paquete de Python para analizar documentos `HTML` y `XML`. Para más referencias, véase <https://www.crummy.com/software/BeautifulSoup/>

22 En particular “home” e “index” son archivos que se cargan por defecto cuando accedemos a un sitio web en el navegador.

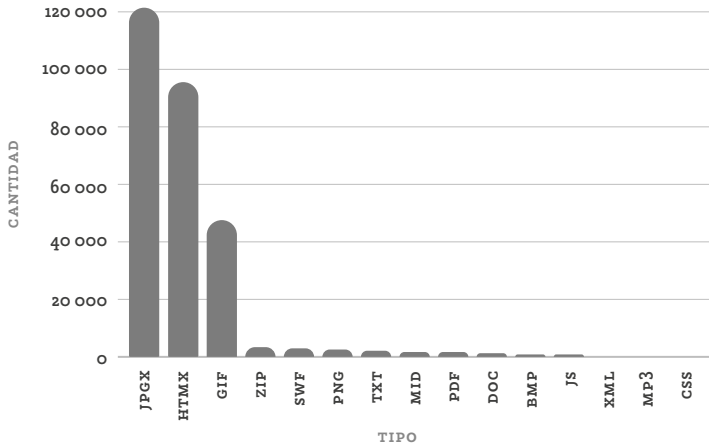


FIGURA 3.1. Cantidad de tipos de archivos, sección Argentina (n = 291819, > 500)

FUENTE: elaboración propia.

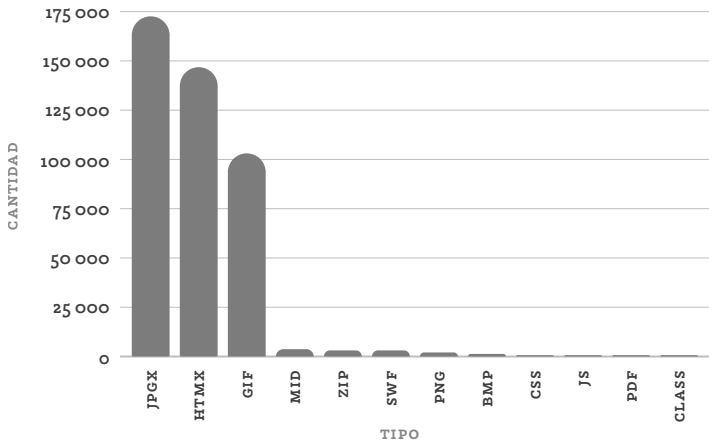


FIGURA 3.2. Cantidad de tipos de archivos, sección Brasil (n = 418836, > 700)

FUENTE: elaboración propia.

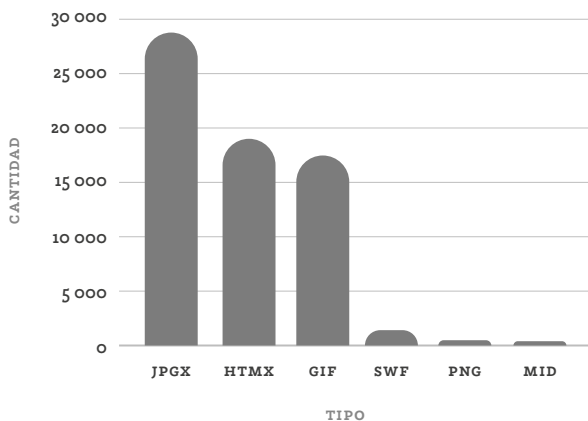


FIGURA 3.3. Cantidad de tipos de archivos, sección México (n = 64820, > 500)

FUENTE: elaboración propia.

utilizarlos en aproximaciones más específicas²³. Finalmente, recurrí a una función de preprocesamiento de texto (más adelante me referiré a ella) para obtener la cantidad de palabras en el contenido textual de esas páginas, luego promediadas (tabla 3.1).

Esto me permitió establecer algunas balizas sobre el conjunto de las fuentes: no todos los directorios poseen archivos como “home” o “index”. Al menos en las secciones Brasil y México, la diferencia es significativa y me obligó a pensar en otras maneras de quitar del análisis aquellos directorios sin páginas de inicio (luego

23 Como puede leerse en la tabla 3.1, las tablas fueron un recurso muy utilizado para organizar la información en las páginas principales, mientras que el uso del elemento *iframe* —que permite anidar páginas— no fue significativo. Este último dato es relevante, porque robustece otras características de la tabla (como la cantidad de palabras en “home”).

TABLA 3.1. Primeras exploraciones sobre los sitios web de las secciones de Argentina, Brasil y México en las secciones locales de GeoCities

	DIRECTORIOS (CARPETAS)	ARCHIVOS	TABLAS	IFRAMES	META = DESCRIPCIONES	META = KEYWORDS	META = AUTOR	META = FECHAS	PROMEDIO PALABRAS
Argentina	6404	7081	64 %	2 %	13 %	17 %	5 %	0	188
Brasil	19997	9544	69 %	2 %	8 %	8 %	5 %	0	110
México	3334	1469	59 %	1 %	7 %	6 %	3 %	0	115

NOTA: las dos primeras columnas indican totales de carpetas y archivos en cada sección. Las columnas en gris contienen datos relativos a las *home pages* de cada sección.

FUENTE: elaboración propia.

buscaré otras páginas que podrían haber funcionado como *home* y detectaré directorios parcialmente guardados). Se trata de un problema del guardado web, por lo que la decisión de no considerar en algunos procedimientos los directorios sin página de inicio es discutible, provisoria. Los datos obtenidos no son esperanzadores en cuanto a la posibilidad de ampliar el conocimiento del acervo con los metadatos. No hay muchos: la sección Argentina posee cerca de un 15 % de descripciones y palabras clave. Por otro lado, está claro que tendré que probar otras maneras de pensar fechas

de edición y publicación de las páginas web, así como tratar de adelantarme al código mal escrito (errores en la confección de las páginas web que el navegador puede ignorar): el listado de descripciones de la sección Argentina mostró claramente esa práctica.

Incorpore otras características al conjunto de datos de las secciones y las páginas web grabadas que no enumeraré. Con todo, el más relevante de los procedimientos fue la obtención de contenido textual significativo de los HTML.

Preprocesamiento

Elegí mantener sin cambios los directorios y archivos guardados, por lo que alojé en otra carpeta los archivos que obtuve con el preprocesamiento. Recorrí los directorios de Argentina, Brasil y México, recuperé los archivos de cada “sitio web” en su primer nivel y, con Beautiful Soup, retuve el texto de los HTML; luego, “limpié” las cadenas obtenidas y, para cada sitio web, extraje un archivo TXT con las palabras relevantes para mi análisis.

En algunos casos, es posible ver la página web servida (a través de la Wayback Machine de Archive.org o en OoCities). Pero la página web guardada en Archive.org es muy diferente (en muchos casos, la copia que poseo solo tiene un HTML, y ese archivo es distinto al primer *snapshot* en la Wayback Machine).

“Limpiar” es una de las actividades más importantes del análisis; en términos demasiados sencillos, significa decidir algunas transformaciones y eliminaciones de los textos del corpus o repositorio: el reemplazo de la letra “ñ” y los acentos, la conversión a minúsculas, la eliminación de palabras no significativas para el análisis —“de uso frecuente” es un argumento sólido—, remoción de números, símbolos, puntuación, entre muchas opciones. La

tokenización (dividir un texto en unidades como palabras, oraciones o párrafos), el *stemming* y la *lematización* —reducir las formas flexivas y derivadas de una palabra a su base común (Manning *et al.*, 2008) y la lista de palabras (*stopwords*) desestimadas por irrelevantes (Luhn, 1958)— son operaciones del preprocesamiento de textos conocidas. Resulta de mucha utilidad identificar y cotejar distintas funciones relacionadas con el preprocesamiento, algunas de ellas incluidas en bibliotecas de *natural language processing* (NLP) en R y Python (Karsdorp *et al.*, 2021). Sin embargo, es indispensable comprender las implicaciones, en especial para las humanidades y las ciencias sociales, del trabajo interpretativo mal sintetizado en el sintagma “limpieza de datos”. Como han señalado oportunamente Rawson y Muñoz (2019):

Muchas veces los detalles de la limpieza de datos no se describen en ningún sitio, sino que residen en las prácticas profesionales generales, los materiales, las historias personales y las herramientas de los investigadores. El hecho de que empleemos un lenguaje tan oscuro como el de la “limpieza de datos” debería ser una fuerte invitación a escudriñar, quizás a reimaginar y, casi con toda seguridad, a renombrar esta parte de nuestra práctica. (p. 279)

Tuve que tomar un conjunto de decisiones en este punto. Opté por remover números²⁴, no quitar acentos y no lematizar. Para

24 Me pareció apropiado recuperar años entre 1994 y 2009 por medios de *expresiones regulares*, para luego cruzar esa información con el *dataframe* de objetos del repositorio.

este tipo de aproximación, considero que los acentos pueden ser importantes para algunas búsquedas y no conozco en profundidad las bibliotecas para lematizar en idioma español. No es la mejor decisión, sino una buena para permitirme escapar de horas de control sobre los lemas.

A partir de esos archivos, tal como ha sido propuesto en los *notebooks* del Archives Unleashed, puedo trabajar con *dataframes* y *series* de la biblioteca Pandas²⁵. Por ejemplo, puedo leer y cargar en una columna de un *dataframe* los archivos de texto plano, con una simple función:

```
directorio = "C:/geocities/parseados/"
df['texto'] = [open(os.path.join(directorio, f), 'r',
encoding='utf-8').read() for f in lista]
```

Con al menos dos columnas (nombre del directorio y “texto”) puedo obtener nuevos datos, como puede leerse en la tabla 3.2.

El análisis de esos valores no es formalizable, se parece más a un vagabundeo con resultados de agrupamientos inciertos y variados dispositivos para visualizarlos. Decidí calcular el número de carpetas y HTMX para saber un poco más sobre la organización del sitio web: algunos usuarios disponían en el directorio raíz todos los HTMX, mientras que otros organizaban los archivos HTMX en directorios e incluso hacían copias de versiones anteriores. Los datos sobre cantidades de algunos tipos de archivos

25 Pandas es una biblioteca para Python que se utiliza para manipulación y análisis de datos. El *dataframe* es una estructura de datos bidimensional (filas y columnas) para el manejo de datos heterogéneos. Para más información, véase <https://pandas.pydata.org/>.

TABLA 3.2. Fragmento del *dataframe* con algunas características de los archivos HTMX de cada directorio

NOMBRE	TEXTO	TOKENS	PALABRAS	CARPETAS	HTMX	JPG	PDF	GIF	SWF	NÚMERO DE LINKS
a2qvegeta.txt	guía capítulos sagas guía capítulos sagas guí...	[guía, capítulos, sagas, guía, capítulos, saga...	187	3	3	13	0	0	0	28
a77acados_online.txt	home guest-book letras fechas fotos censo biog...	[home, guestbook, letras, fechas, fotos, censo...	506	0	10	6	0	0	0	199
abaguik.txt	alquiler departamento villa gesell ubicación ...	[alquiler, departamento, villa, gesell, ubicac...	148	1	7	6	0	0	0	84

NOTA: las características de la tabla corresponden al total de archivos HTMX de cada directorio.

FUENTE: elaboración propia.

me permitieron retener algunos sitios por escasez o exceso. El hápax es una de las pocas cosas que resaltan para mí en una tabla como esa, incluso si puedo leerla con Excel, R o Visual Code, aplicaciones que me permiten interactuar con estas, a diferencia de los cuadernos (Jupyter)²⁶, que requiere agregar más código para ese tipo de experiencia, pero sin que pueda trabajar con ventanas. Los archivos PDF, así como los SWF, siguen apareciendo como posibilidad; ambos pueden indicarme tipos de sitios de web específicos, en tanto no todos los sitios con fuerte contenido textual o artístico adoptaron necesariamente esos tipos de archivos²⁷.

Los valores de los enlaces están puestos en esta tabla como un ejemplo, en cuanto son un objeto muy preciado para la historia de la web (Rogers, 2013; Liu, 2020). Sin embargo, en mi caso, el dato más relevante para no utilizar técnicas de análisis de redes sociales y *software* como Gephi o la biblioteca NetworkX es que la copia de las secciones locales ya no tiene estructura de barrios. El indicio más significativo para esa decisión fue que no había muchos enlaces a otros sitios de GeoCities entre los enlaces de cada sitio de la sección Argentina, más allá de los internos. Entiendo que los sitios web que analizo no se miraban entre sí, es decir, están lejos

26 “Jupyter Notebook es una interfaz web de código abierto que permite la inclusión de texto, video, audio, imágenes, así como la ejecución de código a través del navegador en múltiples lenguajes. Esta ejecución se realiza mediante la comunicación con un núcleo (Kernel) de cálculo” (Cabrera Granado y Díaz García, 2018).

27 En el caso de los PDF, me interesé por rastrear usos durante los primeros años de la web. Pude leer sus metadatos con bibliotecas para Python (Pikepdf, por ejemplo) y recuperar fechas, pero no muchos archivos PDF tienen ese metadato, y solo diecinueve archivos de los que sí lo tienen están datados entre 1990 y 1999.

de la comunidad de GeoCities que Milligan (2019) identificó entre los participantes del barrio “Enchanted Forests”, pero eso requiere nuevas exploraciones.

A partir del análisis de la información parcialmente expuesta en la tabla 3.2, tomé otras notas. Existían sitios web sobresalientes por sus números: tanto por la cantidad de carpetas o archivos HTML, enlaces y *tokens*, esa veintena de proyectos conformaban un borrador de corpus, pequeño en cantidad, extenso en contenidos (texto, metadatos, etc.), algunos de ellos con más de doscientos mil *tokens*. Identificarlos me resultó de mucha importancia para la siguiente etapa en la exploración, en la que trabajé con texto.

Algunos directorios enormes son copias de noticias (un sitio de Uruguay con estilo *weblog*, pero sin bases de datos; el sitio sobre información del mundo ferroviario; un portal con columnas de opinión y muchos enlaces a otros sitios; una página sobre astronomía), reservorios de libros y reseñas de cine y literatura, páginas personales que acumularon muchos escritos (en algún caso ocupando varios directorios), una página organizada por una persona que se define como militante político y una página sobre bibliografía católica²⁸.

Nubes de palabras

Lo sorprendente de esta técnica es la distancia entre su empleo extendido y la escasa repercusión, en esos usos, de su creciente

28 En las secciones que revisé había sitios web que no son de los países a los que las secciones hacen referencia (por ejemplo, el sitio web uruguayo que menciono, muy voluminoso), y eso implica nuevas balizas para futuras conclusiones.

sofisticación técnica. La bibliografía reconoce varias generaciones en el desarrollo de esta forma de resumen visual de texto (Khusro, 2018) y revisa cuantiosos experimentos actuales, enfocados a utilizar algoritmos relacionados con procesamiento, lenguaje natural e imágenes para superar las maneras menos sutiles y más comunes de esta herramienta (Cui *et al.*, 2010). En la pila de producción, circulación y consumo de esta técnica se entrelazan (muchos) procedimientos y factores de eficacia: desde la importancia del tipo y composición de las fuentes documentales, pasando por distintos aspectos del procesamiento de lenguaje natural (preprocesamiento, *stop words*, características evaluadas y formas de calcularlas), hasta los diferentes desafíos de la visualización y sus procedimientos para espacializar gráficamente relaciones semánticas. A esto hay que sumar un elemento decisivo: la usabilidad. De acuerdo con algunas investigaciones (Hearst *et al.*, 2020), no queda claro si las versiones más sofisticadas de las nubes de palabras son más eficaces que el paradigma *wordle* —“la droga de entrada al análisis textual” (Feinberg, 2010, p. 37)— o que las versiones de la primera generación, que ordenan los términos alfabéticamente²⁹. La usabilidad nos recuerda que la lectura de textos es, principal pero no solamente, un ejercicio visual y que el “golpe de vista” —el acto intelectual que hace deseable el resumen visual— es un elemento

29 Wordle.net fue uno de los sitios más conocidos para producir nubes de palabras. A través de un *applet* de Java, parte privado (IBM) y parte propio, Jonathan Feinberg diseñó una visualización que ubicaba en un lienzo palabras y representaba su frecuencia por medio del tamaño, con notables cambios estéticos con respecto a las primeras versiones de nubes de palabras (especialmente las de Flickr). Para más información sobre el paradigma *wordle*, véase Davies (s. f.), Khosravi (2018), Feinberg (2010) y Viégas *et al.* (2009).

clave en el funcionamiento de la técnica. Con ese reconocimiento, la usabilidad le devuelve sentido y pertinencia (utilidad) a los debates sobre la materialidad del libro y la lectura, al menos para este ejercicio.

La nube de palabras puede ser comprendida como parte de la *toma de notas*, antes que como una técnica para exponer resultados. Sin embargo, incluso concebida como una forma de anotar, similar a las marginalias, el subrayado, las fichas, el cuaderno de notas, entre otras, sigue distante de la literatura que discute la técnica, en tanto es notable la presencia del estilo *wordle* en textos académicos.

La decisión de utilizar nubes de palabras fue el resultado de lo que tengo *incorporado* en mi trabajo cotidiano: necesito recorrer el repositorio. Como afirmé, requiero deambular por esas carpetas, tratando de entresacar información de la lectura de documentos HTML como si fueran texto plano, del visionado de imágenes y de la lectura del orden que adopta el conjunto de archivos y carpetas en el sistema operativo Windows. Estoy convencido de que, en cierto pero importante sentido, ese procedimiento es perjudicial, además de ineficaz: produce evaluaciones apresuradas y sesgadas allí donde la lectura no funciona, como cuando leo informes, proyectos, artículos académicos, sobre los cuales puedo imprimir algo de “lectura oblicua” o “saltos de lectura”, en virtud de las convenciones fuertemente establecidas de esos géneros y de mi experiencia previa con ellos.

Pero, además, y esto fue muy importante a la hora de comenzar mi exploración, la lectura de documentos HTML con un visor de texto plano no arrojó buenos resultados: nunca verifiqué un proceso de aprendizaje (aunque sigo pensando que sería posible

para algunas personas leerlos tan bien como el personaje Cypher lee la “lluvia verde” en *Matrix*).

Elegí generar una nube de palabras para el conjunto de archivos HTML de cada directorio (sin recursividad, es decir, no mapeo los directorios por debajo del primer nivel) y examinar los gráficos resultantes en una galería, como si leyera resúmenes visuales de sitios web. El primer intento lo hice con la biblioteca Wordcloud, un recurso conocido. Mi segunda opción fue utilizar nubes semánticas; intenté con Wordmesh, una biblioteca para Python con muchos años sin actualizaciones. Primero, esto resultó un desafío y, luego, un problema. La base de Wordmesh para procesamiento de lenguaje natural es Spacy + Textacy, poderosas bibliotecas de NLP para Python. La actualización regular de esta última biblioteca “rompió” el código en diferentes funciones y tuve que trabajar con versiones anteriores de Textacy o directamente “tocar” el código de Wordmesh. Aunque utilicé este programa para visualizar bigramas, decidí probar con otras bibliotecas.

Uno de los recursos más conocidos en ese campo es el sitio Semantic Word Cloud Visualization³⁰, el cual permite probar distintas funciones en el sitio web, desde la línea de comando o bien instalando su versión abierta en una *web app*. No hubo manera de poner en funcionamiento ninguna de esas estrategias. Finalmente, elegí WordCloud.jl, una biblioteca escrita en Julia, un lenguaje de programación más nuevo y menos conocido que Python. Basada en la selección de términos por coocurrencia y similaridad (conecta con varias librerías de representación y clasificación de textos: CoreNLP, Word2Vec y fastText, pero la única que tiene

30 Véase <http://wordcloud.cs.arizona.edu/>

un modelo para el español es la última), y con algoritmos para el diseño que buscan mejorar los del tipo *force-directed* (Schubert *et al.*, 2017), WordCloud.jl ofreció algunas mejoras en mi trabajo: el tamaño de los documentos y la distancia entre términos disímiles favoreció la mirada “oblicua” que buscaba. En concreto, fui “afinando” (*tuning*, como cuando buscamos la señal de radio) una clasificación incipiente de términos no significativos que Beautiful Soup no quitó de escena (“endif”, por ejemplo); por otro lado, me hice una idea general sobre tipos de expresiones y géneros culturales que circularon por GeoCities a principios de los años 2000.

Contar objetos y palabras, poner en funcionamiento algoritmos para disponer de otro modo información que atravesó varias conversiones hasta mostrarse, operar con tablas en lugar de páginas, mirar antes que leer son procedimientos más precisos que la búsqueda de palabras tradicional. La construcción de vocabularios relativos a temas de importancia para los contemporáneos es una mejor aproximación, pero no puede surgir solo del conocimiento de otro tipo de fuentes (medios, archivos, etc.). Géneros como el *weblog* o el álbum de fotos (galería), formas emergentes de presentación de algunas instituciones, mecanismos de comunicación como el libro de visitas, imágenes que circulan mucho, entre otros, son objetos que pueden servirnos para configurar vocabularios parciales o inestables para seguir probando en el laboratorio. Esto significa que un ejercicio puede no servir para generar conclusiones, pero incluso los más simples nos permiten comenzar una exploración. James Baker optó por buscar en el enorme listado de *seeds* de la copia de GeoCities, los nombres de archivos que tuvieran la palabra “diary” como subcadena. Su intuición fue que muchos diarios de vida incluían ese término en

la URL inicial de su sitio web. Parece un buen comienzo, porque a partir de allí no solo trabajó de modo tradicional (conectando esos textos con el género), sino que también amplió su conocimiento sobre las inflexiones del género en relación con las tecnologías. En esa investigación, por ejemplo, los enlaces redefinen la escena nativa, al igual que los límites éticos de la investigación (Baker, 2020). Ensayé algunas tácticas que terminaron mal. A continuación, un ejemplo.

Localidades y provincias

Una vez preprocesados los archivos HTML, las búsquedas son más rápidas; con un mínimo conocimiento de expresiones regulares es posible recuperar cadenas y subcadenas, y contrastar listas y diccionarios con los textos asociados a cada directorio. Decidí recuperar un listado de localidades de Argentina³¹ y buscarlas en ese corpus (variable “listar”, en el código fuente de la figura 3.4). Sabía que se trataba de un procedimiento confuso: algunos nombres de localidades son términos que tienen otro uso o no existían a fines del siglo xx, otros se repiten en distintas provincias, los sitios web no solo mencionan esas palabras como el lugar de residencia de su autor o autores, entre otros. Una opción es buscar esos términos como texto de algunas etiquetas o partes de la estructura del documento (h1, h2, h3, *footer*, etc.), pero preferí probar primero con el texto sin marcar. Además de normalizar la lista (acentos,

31 Véase https://datos.gob.ar/dataset/jgm-servicio-normalizacion-datos-geograficos/archivo/jgm_8.11

minúsculas, quitar fórmulas como “comisión de fomento” o “municipio de”, entre otras), tenía que obtener la cantidad de veces que las provincias (no las localidades) se repetían, para cada sitio (figura 3.4).

```
def find_all(a_str, sub):
    start = 0
    while True:
        start = a_str.find(sub, start)
        if start == -1: return
        yield start
        start += 1 # use start += 1 to find
                   overlapping matches
# https://stackoverflow.com/questions/4664850/
# how-to-find-all-occurrences-of-a-substring
# Patrick Doghare

def buscar(listal, df):
    dic_provs = {}
    pags = []
    for x, row in df.iterrows():
        for ele in listal:
            cantidad = 0
            dato = find_all(row['texto'], ele[0])
            cantidad = sum(1 for x in dato)
            if cantidad > 0:
                if ele[2] in dic_provs.keys():
                    dic_provs[ele[2]] += 1
```

```
        else:
            dic_provs[ele[2]] = 1
    for d in dic_provs.keys():
        pags.append([row.nombre.split('.')[0], d,
                    dic_provs[d]])
    dic_provs.clear()
    return pags
paginas = buscar(listal, df)
```

FIGURA 3.4. Código fuente de la búsqueda de localidades

FUENTE: elaboración propia.

Luego de obtener esos resultados (una lista con tres valores: directorio, provincia y cantidad de veces que aparecen localidades de esa provincia), los transformé en una grilla que tenía a las provincias como columnas y agregué dos columnas más: la cantidad de ceros de cada fila (es decir, las provincias que no se mencionan en ese sitio web) y la suma de menciones de localidades (sin importar las provincias).

Hice entonces un gráfico para saber qué sitios web mencionaban más de una vez a solo una provincia, esto es: las filas con solo una columna de las columnas de provincias con valor mayor a uno. Luego repetí la búsqueda, pero solo con nombres de provincias en lugar de localidades (tabla 3.3).

La parte del trabajo diario que se pierde en intentos truncos es mucha. Son borradores, en el sentido tradicional del término, en cuanto no solo buscamos producir *capta* (en lugar de *data*),

TABLA 3.3. Lista de únicas provincias mencionadas más de una vez en las páginas HTMX (búsqueda de provincias) (n = 19)

```

[[‘biolembrio’, ‘tucumán’, 2],
 [‘qrvlanus’, ‘tucumán’,2],
 [‘patarcas’, ‘tucumán’, 2],
 [‘alfabetaboletin7’, ‘tucumán’, 2],
 [‘universidadtuc’, ‘tucumán’, 2],
 [‘plusherbel’, ‘tucumán’, 2],
 [‘starwarsrosario’, ‘tucumán’, 2],
 [‘biolgene’, ‘tucumán’, 2],
 [‘billetesargentinos’, ‘tucumán’, 2],
 [‘roxanacarabajal’, ‘tucumán’, 2],
 [‘ligadeajedrezescolardetucumán’, ‘tucumán’, 2],
 [‘fedelanus’, ‘tucumán’, 2],
 [‘partidasajedrezhistorico’, ‘tucumán’, 2],
 [‘biolme’, ‘tucumán’, 2],
 [‘gerihalliwel100’, ‘tucumán’, 2],
 [‘biolcito’, ‘tucumán’, 2],
 [‘albertojesusbonnet’, ‘tucumán’, 2],
 [‘carlosasuarez1’, ‘tucumán’, 2],
 [‘enseanzacad’, ‘tucumán’, 2]]
  
```

FUENTE: elaboración propia.

feliz término que acuñó Johanna Drucker (2011)³², sino también pensar con fuentes nacidas digitales. El borrador es precario solo en algún sentido: la “vista” del repositorio es más precisa luego de este desafío con pobre resultado, pero el procedimiento no es formalizable a nivel de código. En este caso, aunque el ajedrez tucumano está presente en varias páginas de la época (como muestra la tabla 3.3), lo más importante fue advertir que también por este procedimiento algunos directorios resultaron clave por la densidad técnica y textual de sus páginas HTML. Este experimento confirma mi interés por revisarlos por medio de una lectura intensiva.

Modelado de tópicos

El modelado de tópicos es una herramienta extensamente utilizada en el análisis de texto, imágenes y otros materiales. Uno de los algoritmos más conocidos en ese campo es el *latent dirilichet allocation* (LDA), pero existen muchos otros. Cuando un investigador conoce rudimentos de lenguajes informáticos y dispone de

32 “Las diferencias en las raíces etimológicas de los términos dato y capta dejan clara la distinción entre los enfoques constructivista y realista. Capta se ‘toma’ activamente, mientras que los datos se asumen como algo ‘dado’ que se puede registrar y observar. De esta distinción surge un mundo de diferencias. La investigación humanística reconoce el carácter situado, parcial y constitutivo de la producción de conocimiento, el reconocimiento de que el conocimiento se construye, se toma, no simplemente se da como representación natural de un hecho preexistente” (Drucker, 2011, p. 3). Sobre los datos en humanidades, véase Allés Torrent (2019); sobre Drucker y las humanidades digitales, véase Mournier (2021).

muchas aplicaciones creadas para aplicar esas fórmulas sobre un corpus, puede quedar prendado de lo que arrojan esos misterios. La *Encyclopedia Dramatica* (s. f.) reconoce esa figura extática en su definición políticamente incorrecta de comportamientos sociales en internet:

La tragedia de un típico *script kiddie* es que se aproxima al software de la misma manera en que el cura se relaciona con la eucaristía: está tan fascinado por la magia de las computadoras que las percibe como un sacramento de Dios. No tiene por eso la paciencia ni la resistencia mental para “hackear”.³³

El LDA busca afinidades probabilísticas entre palabras y temas (latentes) en el corpus analizado. Es *generativo* porque nos propone la ficción de la producción de un texto (futuro), a partir de elegir tópicos y palabras, basada en probabilidades. Es una historia contada muchas veces, de muchos modos, doctrinales y evangélicos (Blei, 2011; Underwood, 2012; Templeton, 2011; Jockers, 2013). Son muy conocidas algunas herramientas para ese método (*machine learning for language toolkit* [MALLETT], por ejemplo) y cada ensayo introduce pequeños o grandes variaciones, inteligentes acotaciones sobre el uso de la técnica, más allá de las particularidades del corpus utilizado: Jockers (2013), para analizar novelas; Rosatti (2022), con letras de tango; Nelson (s. f.), con un diario; Mercado *et al.* (2020), para analizar la alineación política

33 Solo retengamos de esa definición la ironía al cubo: experticia significa extrañar una práctica y eso implica conocerla. Véase *Encyclopedia Dramatica* (s. f.).

de periodistas argentinos; Karsdorp *et al.* (2021), para clasificar dictámenes de la Corte Suprema estadounidense.

Milligan (2017) utilizó el modelado de tópicos para testear la hipótesis sobre las relaciones efectivamente densas (“comunitarias”) entre las páginas de un barrio de GeoCities (tópicos latentes temáticamente relacionados con el tópico nominal del barrio). Como mi repositorio no es un barrio, ni es un corpus documental, este procedimiento puede funcionar como un modo de continuar la pesquisa. En este caso, el supuesto es avanzar en la construcción de un vocabulario que me permita rastrear un conjunto de páginas web relacionadas con un tema historiográfico. Al probar técnicas no supervisadas para la producción de tópicos, el riesgo no es solo el embelesamiento, sino también no ponderar adecuadamente el peso de regularidades y las simplificaciones de la visualización impuestos sobre la heterogeneidad manifiesta de la copia de las secciones de GeoCities.

La biblioteca que más utilicé en este ejercicio fue Tomotopy. Pude aplicar diferentes algoritmos (LDA, HLDA, PAM, HPAM, etc.)³⁴; sin embargo, fue muy difícil comprender parte de su funcionamiento. Opté entonces por comparar resultados bajo un mismo algoritmo (LDA) con diferentes bibliotecas (figura 3.5). Realicé pequeños ajustes para acercarme a páginas y temas, y evaluar los hallazgos, además de utilizar una herramienta de visualización para R y Python, LDAvis³⁵. Retuve resultados de procesamiento de

34 Las siglas remiten a algunos algoritmos, supervisados y no supervisados, de la docena que pone en funcionamiento la biblioteca Tomotopy. Para más información, véase <https://bab2min.github.io/tomotopy/vo.12.3/en/>

35 LDAvis y su versión para Python, pyLDAvis, son bibliotecas que generan una visualización interactiva del modelado de tópicos. Para profundizar en ellas,

subtópicos (HPAM). Desarrollé otras tácticas caseras para conectar este ejercicio con otros: copié en una carpeta temporal las imágenes de las nubes de palabras realizadas con Wordcloud.jl y de aquellas páginas fuertemente asociadas con un tópico (lo que me permitía tener un “pantallazo” del texto preprocesado); así mismo, produje enlaces que apuntaban a la copia en Archive.org de esas páginas, buscando esa información en la lista de *seeds* de la copia de GeoCities de Archive Team.

```

import tomotopy as tp
n_topicos = 20
n_palabras = 20
topicos = []
topico_palabras = []

#defino el modelo
modelo = tp.LDAModel(tw=tp.TermWeight.ONE, min_cf=3,
                    rm_top=5, k=n_topicos)

# Cargo los textos del dataframe de páginas
for _, row in df.iterrows():
    modelo.add_doc(row['texto'].split())

print("Entrenamiento de Tomotopy LDA...\n")
iteraciones = 10
for i in range(0, 100, iteraciones):

```

.....
 véase <https://github.com/cpsievert/LDAvis> y <https://github.com/bmabey/pyLDAvis>.

```

modelo.train(iteraciones)
print(f'Iteración: {i}\tLog-likelihood: {modelo.
ll_per_word}')

print("Resultado:\n")
for tn in range(n_topicos):
    p_t = ' '.join(p for p, prob in modelo.get_topic_
words(topic_id=tn, top_n=n_palabras))
    topicos.append(p_t)
    topico_palabras.append(p_t.split())
    print(f"Tópico {tn}\n{p_t}\n")

```

FIGURA 3.5. Código fuente de la función LDAModel de la biblioteca Tomotopy
FUENTE: elaboración propia.

Las variaciones sobre los parámetros del modelo (número de tópicos [10-35] y número de palabras [10-25]) mostraron resultados que interpreté similares:

Algunos datos del sumario:

```

<Basic Info>
| LDAModel (current version: 0.12.2)
| 1862 docs, 11121996 words
| Total Vocabs: 553851, Used Vocabs: 180532
| Entropy of words: 10.13520
| Entropy of term-weighted words: 10.13520
| Removed Vocabs: vida argentina años mundo forma

```

<Training Info>

| Iterations: 150, Burn-in steps: 0
| Optimization Interval: 10
| Log-likelihood per word: -9.92906

<Topics>

| #0 (248424) : color perro cabeza dinosaurios pelo
| #1 (592846) : dios iglesia jesús cristo señor
| #2 (1467759) : social sistema desarrollo proceso sociedad
| #3 (641051) : internet resultados html curso programa
| #4 (197080) : motor fiand repuestos autopartes autores
| #5 (656391) : that with your this from
| #6 (156780) : write window width html diet
| #7 (624196) : aires gobierno pueblo nacional país
| #8 (765533) : luis ciudad mail aires nacional
| #9 (251576) : krishna dios señor nombre mente
| #10 (530215) : banda disco música metal rock
| #11 (1188553) : personas hombre persona realidad libro
| #12 (344133) : club equipo river jugador torneo
| #13 (433923) : alas artículo mensajera nacional
derecho
| #14 (157198) : brasilien tempo kriegler flug nach
| #15 (285847) : juan ciudad maría josé carlos
| #16 (485438) : agua rapidshare files part aire
| #17 (1029696) : amor quiero noche mujer corazón
| #18 (759470) : casa cabeza nombre padre cuerpo
| #19 (305887) : castellano ingles original español
director

El tópico 1 (que comparte “dios” con el tópico 9, pero no “Iglesia”) fue recurrente en las distintas pruebas. Si selecciono los documentos de ese tópico y me quedo con valores altos de correlación, puedo comenzar a revisar una lista de páginas (tabla 3.4).

TABLA 3.4. Filtrado de documentos asociados al tópico “Iglesia” con valores altos [> 0.7]

PÁGINAS	CORRELACIÓN
acercadelcristianismo.txt	0,80524373
catolicos_denuncian.txt	0,7716869
catolicos_denunciano3.txt	0,82800573
Integrismo.txt	0,8327203
Integrismoo02.txt	0,82698774
integrismoo03.txt	0,8053105
iobeargentina.txt	0,7160498
ministeriomisionnuevavida.txt	0,8338355
misaldiario.txt	0,794178
misa_tridentina.txt	0,7452652
misa_tridentinao6.txt	0,71824
padres_y_doctores.txt	0,7276514
porlavirgen.txt	0,7322862

Continúa...

Continúa...

PÁGINAS	CORRELACIÓN
predestinacionyelecciondivina.txt	0,725962
sacrificioyexpiacion.txt	0,80524373
sedevacancia.txt	0,73008
sonlosjudioselpueblolejidadedio.txt	0,779047
verdadunica.txt	0,7408422
verdadunica01.txt	0,7506057

FUENTE: elaboración propia.

Me serví de LDavis (figura 3.6) para obtener una visión del modelado reducido a dos dimensiones (saliencia y relevancia, PC₁ y PC₂, respectivamente) (Sievert y Shirley, 2014). Esto me permitió identificar algunos problemas y especificidades del repositorio y el preprocesamiento (por ejemplo, el tópico 5). A medida que cambié esas operaciones, también pude reajustar estas aproximaciones. Está claro que este método no supervisado revela información como la presentada en la tabla 3.4, que podría ser relevante sin la exploración que estoy exponiendo; sin embargo, la lectura intensiva que haremos de los sitios web clasificados no poseerá información valiosa sobre objetos digitales y codificación HTML. El mismo problema aparece cuando usamos el modelo de tópicos “automáticamente” y observamos la interfaz de la figura: la atención que presté al tópico 5 solo es posible al conocer detalles del guardado web y el preprocesamiento de los archivos.

<136,

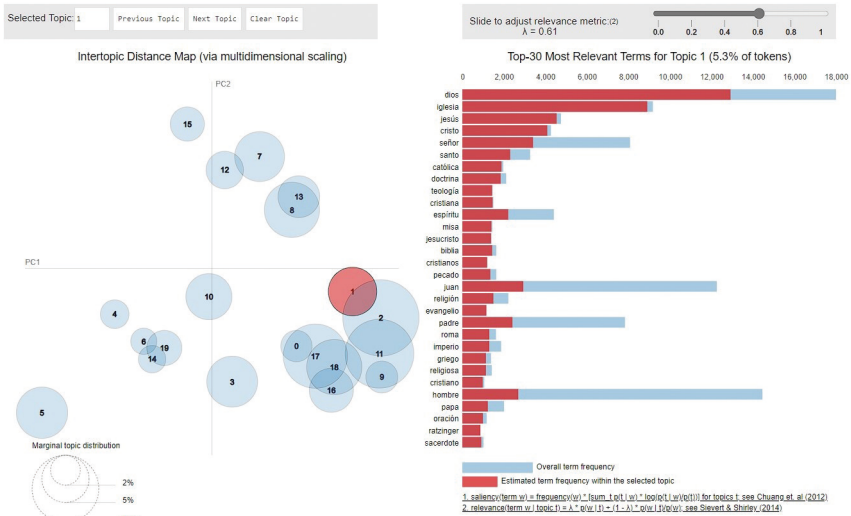


FIGURA 3.6. Interfaz de LDAVIS para el repositorio .ar de GeoCities

FUENTE: elaboración propia.

Conclusiones

El agrupamiento provisorio de “atados” de sitios web relacionados con temas de historia argentina y de historia de la web (“católicos”, “páginas personales”, “política”, “intercambios de software”, “rock nacional”, etc.) se debe en parte a la recuperación de algunos resultados del ejercicio expuesto. La búsqueda deberá avanzar hacia el aprendizaje de vocabularios, fórmulas, temas y soportes de expresión en esos “campos”, pero manteniendo el principio de no clausura y desplazando paulatinamente formas probabilísticas

en la producción de hallazgos. El objetivo, recuperar fuentes para la investigación histórica, implica además conversaciones con expertos en esos temas y problemas. Como Carlo Ginzburg señaló años después de “Conversar con Orión”, él escribió ese texto contra ideas cuestionadoras y simplistas sobre internet, con el propósito de conjugar velocidades en la interpretación textual; así, sostenía la necesidad de utilizar “la velocidad increíble de internet [...] en simultáneo con el ‘arte de la lectura lenta’, aquella famosa definición de Nietzsche sobre la filología” (citado en Garzón Rogé, 2012). Aunque Ginzburg apunta a la velocidad, allí donde otras reflexiones vuelven sobre la distancia (“lectura distante” vs. “lectura cercana”), sus expresiones enmarcan estas “pruebas” para construir corpus documentales, en la medida en que no hemos pretendido, con el uso de técnicas reconocidas en el ámbito de las humanidades digitales, alcanzar conclusiones sobre el repositorio, ni obtener hipótesis, ni generar nuevos interrogantes.

Como he tratado de mostrar con esta exploración, la relación entre interpretación de datos cuantitativos y lectura intensiva de archivos destacados no es directa ni se encuentra formalizada, pero el trato con problemas del almacenamiento, las características de las tecnologías digitales y sus usos, y la codificación (en especial con los procesos de transformación de datos) resultan aspectos fundamentales en ese vínculo. Sin atender esos elementos, la experiencia de la interpretación histórica con repositorios digitales puede ingenuamente concebirse compuesta de dos tramos: utilizar como usuarios herramientas que devuelven información (alejando de ese modo de la vista del historiador o historiadora la faena que recorrimos en este artículo) y hacer lo que siempre hicimos con los textos. Mi propuesta ha sido mostrar que el código es una práctica

fundamental para comprender en qué radica la importancia de los repositorios digitales, la web guardada y el conocimiento sobre las tecnologías y objetos digitales. El código es un habla, pero también una práctica (como el ensayo en teatro); a medida que se incorpore al escritorio de trabajo, como sucede con otras actividades propias de la investigación social, ya no hará falta recordar y subrayar todo el tiempo que gastamos “probando” funciones y aplicaciones.



Referencias

- Allés Torrent, S. (2019). Sobre la complejidad de los datos en humanidades, o cómo traducir las ideas a datos. *Revista de Humanidades Digitales*, 4, 1-28.
- Ankerson, M. S. (2012). Writing web histories with an eye on the analog past. *New Media & Society*, 14(3), 384-400.
- Archive Team. (s. f.). *History is our future*. <https://wiki.archive-team.org/>
- Baker, J. (2020). GeoCities and diaries on the early web. En B. Ben-Amos y D. Ben-Amos (Eds.), *The diary* (pp. 441-456). Indiana University Press.
- Berry, D. (2011). *The philosophy of software: Code and mediation in the digital age*. Palgrave Macmillan.
- Blei, D. (2011). Introduction to probabilistic topic models. *Communication of the ACM*, 55(4), 77-84. <https://doi.org/10.1145/2133806.2133826>

- Brügger, N. (2009). Website history and the website as an object of study. *New Media & Society*, 11, 1-2.
- Brügger, N. (2012). When the present web is later the past: Web historiography, digital history, and internet studies. *Historical Social Research*, 37(4), 102-117. <https://doi.org/10.12759/HSR.37.2012.4.102-117>
- Brügger, N. (2018). *The archived web: Doing history in the digital age*. Massachusetts Institute of Technology Press.
- Chun, W. (2011). *Programmed visions: Software and memory*. Massachusetts Institute of Technology Press.
- Marino, M. (2020). *Critical code studies*. Massachusetts Institute of Technology Press.
- Cramer, F. (2008). Language. En M. Fuller (Ed.), *Software studies: A lexicon* (pp. 168-173). Massachusetts Institute of Technology Press.
- Cui, W., Wu, Y., Liu, S., Wei, F., Zhou, M. y Qu, H. (2010). Context-preserving, dynamic word cloud visualization. *IEEE Computer Graphics and Applications*, 30(6), 42-53.
- Davies, J. (s. f.). *How the word cloud generator works*. <https://www.jasondavies.com/wordcloud/about/>
- Deschamps, R., Ruest, N., Lin, J., Fritz, S. y Milligan, I. (2019). The archives unleashed notebook: Madlibs for jumpstarting scholarly exploration. En S. Downie, Q. Zhu y J. Hahn (Coords.), *Proceedings of the ACM/IEEE Joint Conference on Digital Libraries* (pp. 337-338). Institute of Electrical and Electronics Engineers. <https://doi.org/10.1109/JCDL.2019.00059>
- Díaz García, E. y Cabrera Granado, E. (2018). *Manual de uso de Jupyter Notebook para aplicaciones docentes*. <https://eprints.ucm.es/id/eprint/48304/>

- Drucker, J. (2011). Humanities approaches to graphical display. *Digital Humanities Quarterly*, 5(1). <http://www.digitalhumanities.org/dhq/vol/5/1/000091/000091.html>
- Encyclopedia Dramatica*. (s. f.). Skript kiddie. <https://archive.org/details/dramaticaonline-20210409-history.7z>
- Feinberg, J. (2010). Wordle. En J. Steele y N. Illinsky (Eds.), *Beautiful visualization* (pp. 37-58). O'Reilly.
- Foot, K. y Schneider, S. (2006). *Web campaigning*, Cambridge. Massachusetts Institute of Technology Press.
- Foot, K. y Schneider, S. (2010). Object-oriented web historiography. En N. Brügger (Ed.), *Web history* (pp. 61-79). Peter Lang.
- Fritz, S., Milligan, I., Ruest, N. y Lin, J. (2021). Fostering community engagement through datathon events: The archives unleashed experience. *Digital Humanities Quarterly*, 15(1). <https://www.digitalhumanities.org/dhq/vol/15/1/000536/000536.htm>
- Ginzburg, C. (2004). Conversar con Orión. En *Tentativas* (pp. 321-336). Prohistoria.
- Garzón Rogé, M. (2012). Entrevista a Carlo Ginzburg. *Red-Historia*, (2). <https://web.archive.org/web/20160822214520/http://historia-politica.com/redhistoria/2012/08/entrevista-a-carlo-ginzburg/>
- Gomes, D., Demidova, E., Winters, J. y Risse, T. (Eds.). (2021). *The past web: Exploring web archives*. Springer-Nature.
- Hearst, M., Pedersen, E., Patil, L., Lee, E., Laskowski, P. y Franceneri, S. (2020). An evaluation of semantically grouped word cloud designs. *IEEE Transactions on Visualization and Computer Graphics*, 26(9), 2748-2761. <https://doi.org/10.1109/TVCG.2019.2904683>
- Helmond, A. (2017). Historical website ecology. Analyzing past states of the web using archived source code. En N. Brügger

- (Ed.), *Web 25: Histories from the first 25 years of the World Wide Web* (pp. 139-155). Peter Lang.
- Hidalgo, C. (2006). Reflexividades. *Cuadernos de Antropología Social*, 23, 45-56.
- Holzmann, H. y Nejdil, W. (2021). A holistic view on web archives. En D. Gomes, E. Demidova, J. Winters y T. Risse (Eds.), *The past web. Exploring web archives* (pp- 85-99). Springer-Nature.
- Howard, T. (2019, 22 de enero). How Geocities suburbanized the internet. *Bloomberg*. <https://www.bloomberg.com/news/articles/2019-01-22/remembering-geocities-the-suburbia-of-the-early-web>.
- Icld, A. (2022). Código. En D. Parente, A. Berti y C. Celis (Coords.), *Glosario de filosofía de la técnica* (pp. 91-94). La Cebra.
- Jockers, M. (2013). *Macroanalysis: Digital methods and literary history*. University of Illinois Press.
- Karsdorp, F., Kestemont, M. y Riddell, A. (Eds.). (2021). *Humanities data analysis. Case studies with Python*. Princeton University Press.
- Khosravi, S. (2018). Lab 7. Word clouds. *Observable*. <https://observablehq.com/@058a3c74a4f31700/lab-7-word-clouds>
- Khusro, S., Jabeen, F. y Khan, A. (2018). Tag clouds: Past, present and future. *Proceedings of the National Academy of Sciences, India Section A: Physical Sciences*, 91, 369-381.
- Kittler, F. (2008). Code (or, how you can write something differently). En M. Fuller (Ed.), *Software studies. A lexicon* (pp. 40-47). Massachusetts Institute of Technology Press.
- Krysa, J. y Sedek, G. (2008). Source code. En M. Fuller (Ed.), *Software studies. A lexicon* (pp. 236-242). Massachusetts Institute of Technology Press.

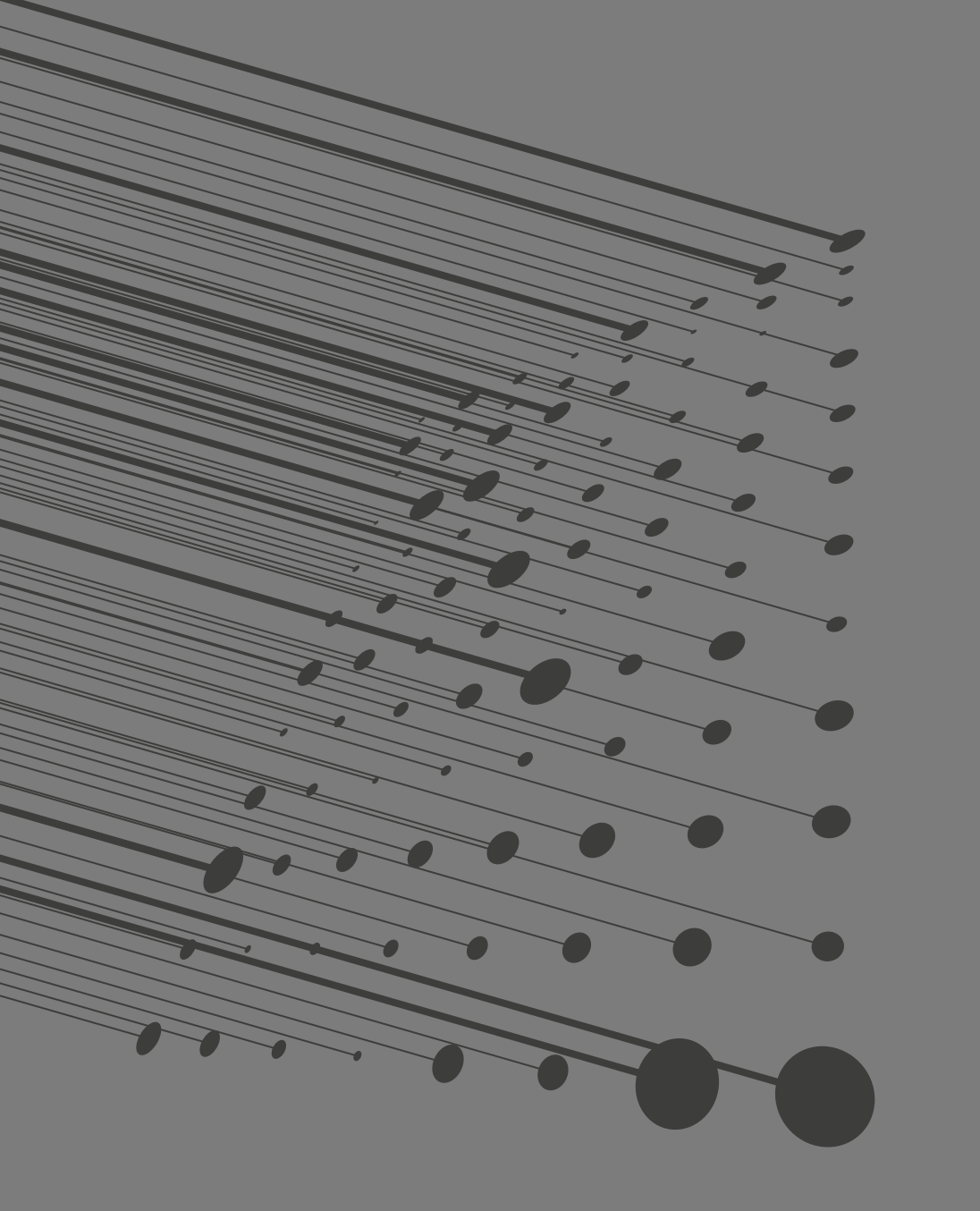
- Liu, A. (2020). ¿Qué significación tienen las humanidades digitales para las humanidades? *452ªF*, (23). <https://revistes.ub.edu/index.php/452f/article/download/32121/32240/76016#:~:text=Respecto%20a%20las%20humanidades%2C%20las,nuevos%20campos%20que%20a%20veces>
- Luhn, H. P. (1958). The automatic creation of literature abstracts. *IBM Journal of Research and Development*, 2(2), 159-165. <https://doi.org/10.1147/rd.22.0159>
- Mackinnon, K. (2022). The death of GeoCities: Seeking destruction and platform eulogies in Web archives. *Internet Histories*, 6(1-2), 237-252. <https://doi.org/10.1080/24701475.2022.2051331>
- Major, D. (2021). The problem of web ephemera. En D. Gomes, E. Demidova, J. Winters y T. Risse (Eds.), *The past web. Exploring web archives* (pp. 5-10). Springer-Nature.
- Major, D. y Gomes, D. (2021). Web archives preserve our digital collective memory. En D. Gomes, E. Demidova, J. Winters y T. Risse (Eds.), *The past web. Exploring Web archives* (pp. 11-21). Springer-Nature.
- Manning, C. D., Raghavan, P. y Schütze, H. (Eds.). (2008). *Introduction to information retrieval*. Cambridge University Press. <https://nlp.stanford.edu/IR-book/html/htmledition/stemming-and-lemmatization-1.html>
- Mercado, V., Villagra, A. y Errecalde, M. (2020). Análisis exploratorio de los datos para la identificación de la alineación política de periodistas argentinos. *Informes Científicos Técnicos-UNPA*, 12(3), 203-218.
- Milligan, I. (2016). Lost in the infinite archive: The promise and pitfalls of web archives. *International Journal of Humanities*

- and Arts Computing*, 10(1), 78-94. <https://doi.org/10.3366/ijhac.2016.0161>
- Milligan, I. (2017). Welcome to the web: The online community of GeoCities during the early years of the world wide web. En N. Brügger y R. Schroeder (Eds.), *The web as history: Using web archives to understand the past and the present* (pp. 137-158). University College London Press.
- Milligan, I. (2019). *History in the age of abundance?: How the web is transforming historical research*. McGill-Queen's University Press.
- Milligan, I. (2020). We are all digital now: Digital photography and the reshaping of historical practice. *Canadian Historical Review*, 101(4), 251-262. <https://doi.org/10.3138/chr-2020-0023>
- Miltner, K. M. y Highfield, T. (2017). Never gonna GIF you up: Analyzing the cultural significance of the animated GIF. *Social Media + Society*, 3(3). <https://doi.org/10.1177/2056305117725223>
- Morgan, P. (2018). The consequences of framing digital humanities tools as easy to use. *College & Undergraduate Libraries*, 25(2), 1-21. <https://doi.org/10.1080/10691316.2018.1480440>
- Mournier, P. (2021). *Humanidades digitales. Una historia crítica*. Mármol Izquierdo Editores.
- Nelson, R. K. (s. f.). Introduction. En *Mining the Dispatch*. <https://dsl.richmond.edu/dispatch/introduction>
- Ogden, J. (2021). "Everything on the internet can be saved": Archive Team, Tumblr and the cultural significance of web archiving. *Internet Histories*, 6(1-2). <https://www.tandfonline.com/doi/full/10.1080/24701475.2021.1985835>
- Owens, T. y Thomas, G. H. (2019). The invention and dissemination of the spacer gif: Implications for the future of access and

- use of web archives. *International Journal of Digital Humanities*, 1(1), 113-132. <https://doi.org/10.1080/24701475.2021.1985835>
- Piovani, J. I. (2018). Reflexividad en el proceso de investigación social: entre el diseño y la práctica. En J. I. Piovani y L. Muñiz Terra (Eds.), *¿Condenados a la reflexividad? Apuntes para repensar el proceso de investigación social* (pp. 74-92). Consejo Latinoamericano de Ciencias Sociales.
- Ramsay, S. y Rockwell, G. (2012). Developing things: Notes toward an epistemology of building in the digital humanities. En M. K. Gold (Ed.), *Debates in the digital humanities* (pp. 75-84). University of Minnesota Press.
- Rawson, K. y Muñoz, T. (2019). Against cleaning. En M. K. Gold y L. F. Klein (Eds.), *Debates in the digital humanities 2019* (pp. 279-292). University of Minnesota Press.
- Reynolds, C. y Hallinan, B. (2021). The haunting of GeoCities and the politics of access control on the early web. *New Media & Society*, 23(11). <https://doi.org/10.1177/1461444820951609>
- Rogers, R. (2013). *Digital methods*. Massachusetts Institute of Technology Press.
- Rosati, G. (2022). Procesamiento de lenguaje natural aplicado a las ciencias sociales. Detección de tópicos en letras de tango. *Revista Latinoamericana de Metodología de la Investigación Social*, 23, 38-60. http://relmis.com.ar/ojs/index.php/relmis/article/view/procesamiento_lenguaje_natural/217
- Ruest, N. (2021). *GeoCities web archive collection derivatives*. <https://archive.org/details/geocities-webarchive-collection-derivatives>
- Ruest, N., Lin, J., Milligan, I. y Fritz, S. (2020). The Archives Unleashed Project: Technology, process, and community to improve scholarly access to web archives. En R. Huang, D. Wu

- y G. Marchionini (Coords.), *Proceedings of the ACM/IEEE Joint Conference on Digital Libraries* (pp. 157-166). ACM. <https://doi.org/10.1145/3383583.3398513>
- Schubert, E., Spitz, A., Weiler, M., Geiß, J. y Gertz, M. (2017). Semantic word clouds with background corpus normalization and t-distributed stochastic neighbor embedding. *arXiv*. <https://arxiv.org/abs/1708.03569>.
- Scott, J. (2010). Archiveteam! The Geocities torrent. ASCII. <http://ascii.textfiles.com/archives/2720>
- Sievert, C. y Shirley, K. (2014). LDAvis: A method for visualizing and interpreting topics. En J. Chuang, S. Green, M. Hearst, J. Heer y P. Koehn (Ed.), *Proceedings of the workshop on interactive language learning, visualization, and interfaces* (pp. 63-70). Association for Computational Linguistics. <https://doi.org/10.3115/v1/W14-3110>
- Soon, W. (2016). *Executing liveness. An examination of the live dimension of code interactions in software (art) practice* [tesis de doctorado]. School of Communication and Culture, Aarhus University.
- Taylor, A. y Tilton, L. (2015). *Exploring humanities data in R: Exploring networks, geospatial data, images and texts*. Springer.
- Taylor, A. y Tilton, L. (2019). New data, new histories: The role of statistics in DH. En M. K. Gold y L. F. Klein (Eds.), *Debates in the Digital Humanities 2019*. University of Minnesota Press. <https://dhdebates.gc.cuny.edu/read/untitled-f2acf72c-a469-49d8-be35-67f9ac1e3a60/section/a2a6a192-fo4a-4082-afaa-97c76a75b21c>
- Templeton, C. (2011). Topic modeling in the humanities: An overview. *MITH Blog*. <http://mith.umd.edu/topic-modeling-in-the-humanities-an-overview/>.
- Tukey, J. (1977). *Exploratory data analysis*. Addison-Wensley.

- Underwood, T. (2012). What kinds of “topics” does topic modeling actually produce. *The Stone and the Shell*. <http://tedunderwood.com/2012/04/01/what-kinds-of-topics-does-topic-modeling-actually-produce/>.
- Unwin, A. (2010). Exploratory data analysis. En P. Peterson, E. Baker y B. McGaw (Eds.), *International encyclopedia of education*. Elsevier.
- Van Es, K., Wieringa, M. y Schäfer, M. (2018). *Tool criticism. From digital methods to digital methodology* [conferencia]. *International conference on Web Studies (WS.2 2018)*, París, Francia.
- Venturini, T., Jensen, P. y Jacomy, M. (2021). What do we see when we look at networks: Visual network analysis, relational ambiguity, and force-directed layouts. *Big Data & Society*, 1(13). <https://doi.org/10.1177/20539517211018488>
- Viégas, F., Wattenberg, M. y Feinberg, J. (2009). Participatory visualization with Wordle. *IEEE Transactions on Visualization and Computer Graphics*, 15(6), 1137-1144.
- Webster, P. (2021). Digital archaeology in the web of links: Reconstructing a late-1990s web sphere. En D. Gomes, E. Demidova, J. Winters y T. Risse (Eds.), *The past web. Exploring web archives* (pp. 155-164). Springer-Nature.
- Wikipedia. (s. f.). *Geocities*. <https://es.wikipedia.org/wiki/GeoCities>.
- Winters, J. y Prescott, A. (2019). Negotiating the born-digital: A problem of search. *Archives and Manuscripts*, 47(3), 391-403.
- York, C. (2017). Exploratory data analysis for the digital humanities: The Comédie-Française Registers Project analytics tool. *English Studies*, 98(5), 1-10.



PARTE II

{ Enseñanza de las humanidades digitales

De la enseñanza de las humanidades digitales en México,
¿qué hemos aprendido?

ERNESTO PRIANI SAISÓ

MIRIAN PEÑA PIMENTEL

ADRIANA ÁLVAREZ SÁNCHEZ

O ensino de humanidades digitais no Brasil: visões a
partir de um *survey* nacional

JIMMY MEDEIROS

JULIANA MARQUES

SUEMI HIGUCHI

Un puente entre la elección y la necesidad: reflexiones
sobre la *minimal computing* para las humanidades digitales
y la ciencia abierta en América Latina

GIMENA DEL RIO RIANDE

RAFFAELE VIGLIANTI

};

De la enseñanza de las humanidades digitales en México, ¿qué hemos aprendido?

Para citar este capítulo:

<http://doi.org/10.51573/Andes.9789587987362.9789587987379.9789587987867.4>

ERNESTO PRIANI SAISÓ
MIRIAM PEÑA PIMENTEL
ADRIANA ÁLVAREZ SÁNCHEZ

Introducción

{ Las humanidades digitales (HD) son un campo de investigación que se ha extendido a lo largo y ancho del ámbito académico en México y en América Latina durante los últimos años, lo que ha generado nuevas aproximaciones a problemas de investigación desde múltiples disciplinas a través del uso de metodologías digitales (Svensson, 2014). Desde nuestra experiencia en México, el humanista digital, casi siempre formado en diferentes espacios de autoaprendizaje relacionados con el trabajo colaborativo, y ante la ausencia de cursos formales de enseñanza, ha buscado ser también enseñante por medio de los espacios institucionales a su alcance, para articular la enseñanza y así ampliar el conocimiento de las HD. Es importante señalar que la formación de cada uno de quienes presentamos este capítulo es muy diversa en cuanto a la perspectiva disciplinar y como humanistas digitales: una es historiadora, otra estudió letras hispánicas y el último, filosofía;

además, dos somos autodidactas y solo una de las autoras tiene estudios formales en HD. Sin embargo, cada uno de los autores vimos la oportunidad de incorporar en nuestra investigación el uso de recursos y herramientas digitales, lo que dio paso a la interdisciplina teórica y práctica en la que actualmente nos desempeñamos.

En este capítulo reflexionaremos sobre las experiencias de enseñanza de las HD desarrolladas en la Facultad de Filosofía y Letras de la Universidad Nacional Autónoma de México (UNAM), un espacio multidisciplinario, porque en ella confluye la enseñanza de las humanidades y de las ciencias sociales¹. No obstante, la facultad es un espacio donde la tradición de los estudios humanísticos tiene un peso importante en el desarrollo de las disciplinas, y si bien las prácticas metodológicas y didácticas se han modificado en los últimos años, este cambio ha sido distante de la incorporación de tecnologías digitales, hasta la llegada de la pandemia, que llevó a toda la comunidad a valorarlas. De hecho, desde hace un par de

1 La Facultad de Filosofía y Letras de la UNAM ofrece doce grados de licenciatura en modalidad escolarizada, ocho en modalidad a distancia/abierta y quince posgrados. La modalidad escolarizada se estructura bajo colegios, los cuales coordinan al menos un grado (licenciatura). Dos de los casos de estudio revisados en este capítulo se rigen bajo la estructura de colegios; el primero es la Licenciatura en Historia, perteneciente y coordinado desde el Colegio de Historia; el segundo es la Licenciatura en Filosofía, perteneciente y coordinado por el Colegio de Filosofía. Cada colegio designa el tipo de cursos que imparte: los cursos de “tronco común”, los seminarios, las cátedras, las asignaturas teóricas, las asignaturas prácticas, las optativas libres, los seminarios/taller, etc.; es posible que los estudiantes de un colegio obtengan autorización para cursar asignaturas en otros colegios e, incluso, en otras facultades si el plan de estudios correspondiente lo permite (siempre y cuando se justifique la pertinencia a su desarrollo profesional). Véase <http://www.filos.unam.mx/>.

décadas existen proyectos digitales desarrollados por académicos de la facultad, aunque la introducción de cursos para la formación de humanistas digitales ha sido un poco más lenta y compleja.

Es necesario considerar que la mayor parte de los planes de estudio actuales no contemplan contenidos cercanos a las HD y que la introducción de nuevas asignaturas de cualquier área requiere de cambios cuyo proceso puede llevar años. Ante esta situación, las experiencias docentes que presentaremos a continuación han encontrado su lugar a partir de asignaturas ya existentes o en otros ámbitos de enseñanza. Así, las experiencias que aquí se abordan son una muestra de las formas en las que se han podido introducir las HD en la formación de estudiantes de las licenciaturas de Historia y Filosofía, pero también de los nuevos caminos que este campo de investigación y docencia ha encontrado para llegar a un alumnado más diverso, a través del diseño de un *massive online open courses* (MOOC) dentro de una plataforma global como lo es Coursera UNAM.

El hilo conductor de este trabajo es la forma como sorteamos la dificultad de converger los estudios tradicionales de humanidades y ciencias sociales con el conocimiento crítico y el uso de herramientas digitales. Como se verá, hemos desarrollado dos acercamientos al tema. En un caso hemos entendido las HD como una herramienta metodológica más dentro de un campo disciplinar; en otros dos casos, hemos asumido que las HD son un campo por sí mismo, el cual mantiene relaciones interdisciplinarias con el resto de las disciplinas humanísticas.

Si bien una definición de lo que son las HD y cuál es su relación con las disciplinas humanísticas es todavía un tema de debate, después de nuestra experiencia como investigadores y

docentes consideramos que, parafraseando a Scott Rettber y Alex Saum-Pascal (2020), las HD no son una metodología o un cuerpo de textos y proyectos; no son historia, filosofía o estudios literarios, ni siquiera una mezcla de todos estos elementos. A pesar de algunas críticas hechas a las HD en años pasados, sobre todo en el campo de los estudios literarios sobre el valor de sus aportes, hoy en día son una práctica cada vez más consolidada, la cual está creando conceptos, estructuras, métodos e instituciones nuevas que amplían la diversidad temática y metodológica del conocimiento humanístico, así como las formas de adaptación a los entornos institucionales en los que emerge.

El primer curso curricular a nivel licenciatura de humanidades digitales en América Latina

En la primavera del 2014 llevamos a cabo una investigación relacionada con la enseñanza de las HD en la docencia universitaria; esta nos llevó al trabajo de Burdick (2012), con respecto a la pragmática y los aspectos institucionales de la docencia en HD. Nuestra propuesta se desarrolló dentro de la Facultad de Filosofía y Letras, institución en la que una de nosotras era profesora de tiempo completo de Historia y otra realizaba una estancia posdoctoral de investigación dentro de las HD y, como parte de ella, organizó el seminario de HD en esta facultad. El seminario representó un punto de encuentro entre quienes, desde distintas formaciones y experiencias, se interesaron en él; así, se inició una colaboración que hizo posible proponer un primer curso curricular. Tras

la solicitud de la titular de la coordinación del Colegio de Historia², se propuso incluir un curso regular en la programación de la Licenciatura en Historia, cuyo plan de estudios incluye un área de investigación, en la que se encuentra el Seminario-Taller Especializado (STE). Las temáticas del STE son propuestas por quienes lo imparten y se cursa en el tercer año, el penúltimo de la carrera. En nuestro caso, el seminario se tituló Humanidades Digitales e Historia y combinó reflexiones epistemológicas con práctica y análisis de recursos y herramientas digitales para la investigación histórica. Los objetivos de este curso fueron:

- [El estudiante] Conocerá las principales discusiones acerca de la utilización de herramientas y modelos de las humanidades digitales aplicados a la investigación histórica.
 - Reflexionará sobre los problemas metodológicos relacionados con la construcción del conocimiento en entornos digitales.
 - Conocerá las distintas herramientas digitales para la investigación humanística.
- 2 A lo largo de la licenciatura se deben aprobar tres distintos tipos de seminarios, los cuales, de acuerdo con el plan de estudios, introducen a los estudiantes a la práctica profesional de la historia, por lo que es necesario realizar una investigación que tenga como resultado final un producto académico concreto. En el caso del STE, que debe tener máximo quince estudiantes, este “se enfocará al uso y aplicación de técnicas específicas de investigación, modalidades particulares de docencia y de medios de comunicación para la difusión del conocimiento histórico, tales como video, cine, CD-ROM, teatro, guion de radio, etcétera” (Facultad de Filosofía y Letras, 1999, p. 52). Es importante mencionar que el plan de estudios data de 1999, por lo que no incluía productos digitales publicados en internet.

- Diseñará un producto digital de investigación, mismo que desarrollará en el segundo semestre del curso. (Álvarez Sánchez y Peña Pimentel, 2016)

A pesar de que las HD no son un campo nuevo, la inclusión de este “apelativo” en el título de un curso de pregrado generó desconcierto, debido al poco o nulo conocimiento del término entre el alumnado. Usamos este desconocimiento en nuestro favor, con la intención de generar curiosidad entre estudiantes y también lograr la incorporación de las HD en un curso curricular. Consideramos que la propuesta de seminario para el alumnado tenía que complementarse con cursos semejantes para docentes. Así, también aprovechamos el programa de actualización docente, que incluye cursos de veinte y cuarenta horas, dirigidos a todo el personal académico de la institución. Estos cursos pueden ser impartidos por docentes de la universidad y las temáticas son variadas. Impartimos distintos cursos en este programa y constatamos el interés de docentes de la Facultad de Filosofía y Letras y de otras dependencias de las UNAM; por lo tanto, parte de los cursos incluyeron la discusión en torno a las HD y el uso de tecnologías digitales (así como de metodologías y recursos *ad hoc*), cuestionando su función en los estudios humanísticos y estableciendo la diferencia entre ser un usuario avanzado de tecnologías de la información y las comunicaciones (TIC) y ser un humanista digital³.

- 3 Aunado al curso regular, en agosto del 2015 (Álvarez, 2015; Peña, 2015), impartimos un curso-taller para los alumnos de nuevo ingreso a la Licenciatura en Historia, con el objetivo de facilitar su transición del bachillerato a la universidad; en este contexto, se aplicó una encuesta que nos permitiera: (1) estar al tanto de sus prácticas académicas; (2) conocer las ideas del alumnado sobre

El STE fue pensado desde las HD para incorporar en los contenidos dos disciplinas humanísticas —la filología y la historia— y fue impartido por dos profesoras en un laboratorio de cómputo, para lograr el desarrollo de las prácticas; es importante señalar que la Facultad de Filosofía y Letras cuenta con dos tipos de laboratorios de cómputo: (1) salas con equipos de cómputo y acceso a internet de uso libre para los alumnos; y (2) salas con equipo de cómputo, acceso a internet y proyector para impartir clases. En estos últimos es posible la instalación en todos los equipos del *software* necesario para el desarrollo de los cursos. No es común que en estos laboratorios se impartan cursos de todas las licenciaturas de humanidades; antes del STE, los principales usuarios de estos espacios eran grupos de las asignaturas de las licenciaturas de Geografía, Bibliotecología y Pedagogía. Para lograr los objetivos del STE —que el alumnado fuera capaz de conocer los métodos y debates de las licenciaturas que estudian, al igual que aprovechar las herramientas y los recursos a su alcance, entre los que se encuentran los digitales, y participar de los debates relacionados con el uso de estas (Fyfe, 2011)—, se requería de un laboratorio para desarrollar las clases.

Este seminario se impartió durante tres ciclos académicos (2015, 2016 y 2017), con un total de seis semestres y aproximadamente

.....
herramientas, metodologías y contenidos digitales; y (3) consolidar la inclusión de una materia que cubriera esas temáticas. Los resultados de estas encuestas se tomaron en cuenta durante la elaboración de los materiales que se iban a incorporar en el STE y nos ayudó a conocer el perfil general sobre los estudiantes, sus prácticas académicas y su nivel de conocimiento (o desconocimiento) de estas en entornos digitales, además de conocer sus hábitos digitales y de búsqueda/recuperación de información.

cuarenta y cinco estudiantes en total. Sin embargo, el contenido del programa y el modelo docente se ajustó en cada ciclo, con base en la experiencia. Durante el primer ciclo escolar nos quedó claro que los estudiantes no contaban con conocimientos sólidos ni reflexivos sobre las modificaciones en la creación de conocimiento en las disciplinas humanísticas producto de la “incorporación tecnológica”. Percibimos que en su mayoría eran consumidores ávidos de contenidos en internet y usuarios de más de un dispositivo electrónico (*laptop*, *tablet* y teléfono inteligente); no obstante, era necesario revisar continuamente términos básicos. Por ejemplo, aclarar la diferencia entre sistema operativo y programas, o explicar la existencia de licencias de publicación. Notamos que el uso de herramientas digitales, desde buscadores, estaba principalmente orientado a la búsqueda de temas de entretenimiento y ocio. Los alumnos no solían utilizar bases de datos especializadas y la mayoría de sus búsquedas bibliográficas las llevaban a cabo, casi exclusivamente, desde los catálogos electrónicos de las bibliotecas. No declararon utilizar publicaciones como blogs, páginas de proyectos, entre otras, ni habían considerado incluir herramientas para sistematizar la información (p. ej. gestores de bibliografías) o producir visualizaciones académicas (p. ej. creación de líneas del tiempo, inclusión de visualización tipo mapa, visualización en red de tópicos).

Los alumnos que cursaban este seminario ya tenían conocimientos más que básicos de la disciplina humanística principal (historia) y estaban familiarizados con las diferentes propuestas docentes de la Facultad de Filosofía y Letras (seminario, cátedra, asignaturas teóricas, asignaturas prácticas, optativas libres, etc.), además de contar con acceso a recursos digitales para la

investigación y la consulta de materiales, aunque no siempre los utilizaran. El alumnado conocía también la existencia de las redes sociales de la facultad y de cada uno de sus colegios como herramienta de divulgación de eventos, actividades académicas e información administrativa útil para la comunidad estudiantil, a pesar de no emplearlas como medio de comunicación entre administración, docentes y estudiantado. Sin embargo, los alumnos de este STE reconocieron, en una retroalimentación al final de la clase, que poco utilizaban estos recursos, más allá del ámbito recreativo —internet como un espacio para el ocio, las redes sociales para el contacto entre amistades—, y afirmaron no tener conocimiento sobre los recursos digitales/electrónicos con los que cuenta la universidad (bases de datos especializadas en humanidades, publicaciones periódicas en acceso abierto, blogs académicos, entre otros).

La primera generación de estudiantes que realizó el STE —habituada a dinámicas de discusión dentro del aula, pero también a recibir cátedra—, en principio, se desconcertó, pues el seminario era en un laboratorio de cómputo, espacio al que asisten únicamente para imprimir, utilizar internet o incluso para redactar sus trabajos, pero no para una clase formal (no en la Licenciatura en Historia). La dinámica de clase —lecturas⁴, debates y, sobre todo, las prácticas realizadas— hicieron que los estudiantes se convirtieran en usuarios y participantes activos. Por tanto, este primer curso curricular funcionó no solo como un aula, sino como un laboratorio experimental para que estudiantes de historia, con el

4 Un ejemplo de estas lecturas son Brescio (2010), Dixon (2012), Galina Rusell (2012, 2013), Lamarca Lapuente (s. f.) y Pons (2013).

acompañamiento de las docentes, exploraran nuevas formas de construir conocimiento y difundirlo; una experiencia novedosa en ese momento y que hoy podría considerarse pionera en el ámbito universitario latinoamericano. Por ejemplo, desconocían que es posible construir un producto académico en línea, que pueden difundirse productos intermedios o abrirse a la discusión de otros usuarios.

La mayoría de los alumnos presentaron como trabajo final productos de investigación en los que incorporaron diferentes elementos aprendidos en clase: publicación de las investigaciones en páginas web, creación de mapas, creación de líneas del tiempo, incluso la creación de una red social orientada a estudiantes de historia. Lamentablemente, sus creadores no mantuvieron los productos actualizados, por lo que hoy no es posible consultarlos.

El STE no se orientó únicamente al uso y experimentación con herramientas digitales; parte fundamental de este fue la presentación de los debates internacionales en torno a las HD. Se discutieron textos sobre qué son y cuáles son sus orígenes, vertientes y principales debates, y se propuso retomar a D'Iorio y Barbera (2011) y lo que denominan *conditions of possibility*, resumidas en tres condiciones: la primera es la citación (*quoting*), ya que la labor académica es una conversación basada en argumentos, demostraciones y pruebas —y los hechos, en las humanidades, a menudo están contenidos en documentos, en textos—. La segunda condición es el consenso: los autores señalan que, para recibir apoyo y ser incluidos en las empresas comunes de investigación, los académicos deben producir trabajos considerados interesantes por sus colegas, pues en la academia este reconocimiento se entiende basado en pruebas, y debe ser lo más justo y transparente

posible; por lo tanto, la confianza, el prestigio, la revisión por pares y otros sistemas de evaluación justa se cuestionan continuamente. Finalmente, explican la tercera condición de posibilidad: preservación/diseminación; de las tres condiciones de posibilidad, esta es la única que implica una característica dual. Reconocemos la importancia de la preservación del conocimiento; sin embargo, el resguardo de este no se concibe útil si no se disemina: sin la difusión del conocimiento no se cumple la condición de la citación, y sin la preservación no se puede garantizar la permanencia del conocimiento (D'Iorio y Barbera, 2011). No obstante, dadas las tendencias internacionales con relación al uso de tecnologías y herramientas digitales, y su inclusión en la investigación y la docencia, la forma de hacer humanidades no abandona sus orígenes, sino que propone la inclusión crítica o reflexiva de estas tecnologías y las tendencias competentes en cada paso del proceso de creación de contenido original. A lo largo de seis semestres (tres grupos distintos, dos semestres cada uno), guiados por las lecturas, la revisión de proyectos de HD, el aprendizaje en el uso de herramientas y la creación de contenidos propios en los que las utilizaron, los estudiantes concluyeron que el ejercicio de las HD es cambiante, adaptativo y, sobre todo, inclusivo en sus creaciones.

Dentro de las actividades prácticas del seminario se llevaron a cabo evaluaciones de proyectos digitales, con base en la primera versión del “Formulario de evaluación para proyectos de humanidades digitales” (Galina *et al.*, 2015). El formulario (que funciona como una rúbrica) tenía la intención de servir como una guía para quienes evaluaran proyectos financiados de HD; consideramos que era y sigue siendo una herramienta valiosa que puede ser utilizada para conocer los elementos básicos de un proyecto

de esta naturaleza: documentación, mantenimiento del recurso, licencias de uso, información de los responsables, accesibilidad de los materiales, etc. El uso de esta herramienta reveló una confusión sobre los distintos elementos por valorar (por ejemplo, se confundía documentación con descripción del recurso). Sin embargo, esto abrió la posibilidad de que el alumnado reconociera la importancia de evaluar los sitios que consultan y, de esta manera, demostrar que las publicaciones y los proyectos académicos en formato digital pueden cumplir con los criterios de calidad de los productos editados e impresos en papel. Así, la aplicación de criterios académicos que el alumnado ya utilizaba para seleccionar bibliografía impresa se amplió a productos académicos publicados en un nuevo soporte, por lo que confirmamos que el uso del formulario como material didáctico es acertado.

Así, para el STE Humanidades Digitales e Historia el trabajo final esperado era la creación de un recurso digital de tema histórico, en el que se utilizaran las herramientas y metodologías de las HD. Los alumnos se convirtieron en investigadores y creadores de sus propios sitios web, bases de datos y contenidos multimedia; al llegar a este punto fue necesario cuestionar la forma en la que iban a desarrollar y publicar sus recursos. Es necesario aclarar que el trabajo colaborativo es poco frecuente en este grado y en esta disciplina; por ello, también cuestionamos la manera en la que el alumnado estaba dispuesto a divulgar su producción. Estos cuestionamientos buscaban hacer conciencia entre los alumnos, pero como usuarios de recursos y herramientas digitales, ya que es importante generar una cultura digital —buenas prácticas—, que les permita aprovechar los medios con los que cuentan para

desarrollarse como profesionales de la historia, usuarios críticos y consumidores éticamente responsables de este tipo de recursos.

El resultado del primer año del seminario curricular dejó clara la necesidad de revisar los procesos de formación en humanidades desde una perspectiva que considerase las producciones “alternativas”, esto es, los proyectos y publicaciones digitales. En general, el reto para el alumnado fue aproximarse a una forma distinta de realizar un trabajo académico; para nosotras también implicó evaluar de manera distinta: hicimos una evaluación continua y, más allá del tema histórico elegido por cada estudiante, valoramos el proceso que llevaron a cabo, corroboramos que cada trabajo contara con los elementos básicos del formulario y pusimos especial atención a los objetos digitales utilizados y sus licencias —por supuesto, revisamos los textos incluidos y las referencias—. Una vez evaluados los productos, recurrimos a elementos más formales, como la asistencia y la entrega de avances conforme al calendario previo; así evaluamos a cada estudiante.

Como se indicó, el STE Humanidades Digitales e Historia se impartió en tres ciclos escolares, por lo que para su segunda emisión fue necesario replantear el programa y, sobre todo, cambiar la forma de creación de conocimiento hacia el trabajo colaborativo. A diferencia del ciclo 2015, en el 2016 el STE no tuvo como objetivo la creación de recursos de manera individual, sino que se estableció un proyecto único en el que el grupo colaboraría, siguiendo la propuesta de Rockwell y Sinclair (2012), orientada a la creación de competencias y habilidades necesarias para su incorporación al mercado laboral; de esta manera, los alumnos del grupo participarían en todas las etapas de investigación y recolección de fuentes y, más adelante, en la creación de contenidos

(bibliografías, visualizaciones y publicación). La idea detrás de esta iniciativa era potenciar el conocimiento de múltiples herramientas (bases de datos, visualizaciones en diferentes formatos, líneas de tiempo, mapas, etc.) y metodologías (extracción y estructuración de la información, gestión de contenidos digitales, publicación digital, entre otras) orientadas a un objetivo común: guiar la publicación de un recurso único y decidir colectivamente la forma de registro y divulgación de este, logrando así un debate sobre el consumo de recursos existentes, la necesidad de recursos libres y el uso de licencias alternativas para la producción y registro de conocimiento nuevo y auténtico. Es decir, el objetivo fue realizar trabajo colaborativo, por lo que parte de las decisiones se tomaron de esa manera.

La tercera vez que se impartió este STE, en el 2017, su nombre cambió a Historia Digital, debido al desconocimiento generalizado sobre las HD entre los estudiantes y a sugerencia de la coordinadora del programa de la licenciatura⁵. A pesar del cambio en el nombre, esta tercera versión contó con un número similar de alumnos que las dos anteriores. Los estudiantes inscritos mostraron interés, pues algunos sí habían oído hablar de Historia Digital y querían conocer el tema del STE, aunque tampoco estaban familiarizados con usos de herramientas, licencias de publicación o

5 A lo largo de las tres emisiones del curso, no se debatió sobre lo que es o no *historia digital*, ya que el cambio en el título del curso (de Humanidades Digitales e Historia a Historia Digital) obedeció a la necesidad de atraer alumnos con términos más familiares para ellos; sin embargo, mantuvimos la misma línea de enseñanza de las HD y su aplicación en un modelo interdisciplinario orientado a la historia.

rastreo de recursos digitales, característica común en el alumna-
 do de las generaciones anteriores.

Esta tercera y última emisión permitió revisar los contenidos
 del curso —tanto la bibliografía complementaria como las herra-
 mientas digitales—, al igual que las estrategias de enseñanza. Se
 mantuvo la misma línea de trabajo: incorporar herramientas digi-
 tales, metodologías digitales y estudios básicos de la formación
 en historia, pero se puso mayor énfasis en la participación en las
 discusiones acerca de las HD, centrándonos en Latinoamérica y
 México, pues para este momento había crecido la bibliografía y la
 divulgación de proyectos y estudios sobre y en esta región. Se buscó
 que los proyectos finales fueran una mezcla de productos originales
 individuales y de trabajos colaborativos. En general, es posible afir-
 mar que, a lo largo de los tres años y seis semestres que se impartió
 el STE, los alumnos no solo consolidaron su conocimiento en mate-
 ria de la disciplina histórica, sino que demostraron que, gracias a
 la libertad en la producción de contenidos, son capaces de generar
 recursos propios de diferentes tipos, en varias modalidades y por
 medio de distintas plataformas, sin descuidar la calidad académica.

Para el ciclo 2017 cambió el titular de la coordinación del Co-
 legio de Historia, quien señaló que no era necesario contar con
 un seminario con esta temática, pues consideró que la relativa
 familiaridad de la comunidad con la búsqueda y el uso de recur-
 sos digitales era suficiente, centrando la formación de los histo-
 riadores en temáticas y prácticas metodológicas tradicionales; por
 esta razón, el STE dejó de ofertarse en la Licenciatura en Historia.
 No obstante, la existencia de ese seminario abrió la posibilidad
 de que otras licenciaturas de la facultad incorporaran asignatu-
 ras semejantes, ya sea en modalidad optativa o como seminarios

especializados, con temáticas cercanas a las HD, como parte de algunos programas de estudio: tales son los ejemplos del Colegio de Lengua y Literatura Hispánicas (Humanidades Digitales y Estudios Literarios), el Colegio de Bibliotecología (Bibliotecas Digitales y Archivos) y el Colegio de Filosofía.

El contexto de la pandemia cambió por completo las condiciones sobre las cuales impartimos estos cursos. La crisis sanitaria reveló la importancia de contar con unas mínimas habilidades digitales entre docentes y estudiantes, así como las carencias a las que nos enfrentábamos. En este contexto se formularon las siguientes dos experiencias docentes.

Seminario Optativo Humanidades Digitales

La decisión de impartir un curso de HD dentro de un seminario optativo de la carrera de Filosofía⁶, en la Facultad de Filosofía y Letras de la UNAM, tuvo como antecedente no solo el STE, sino también la propuesta de iniciar la planeación de una especialización en HD como parte de un programa que se desarrollaría en la

6 El currículo de la carrera de Filosofía en la UNAM contempla dos tipos de materias: las obligatorias, aquellas que todos los alumnos deben cursar necesariamente para poder obtener el grado, y las optativas, las cuales, si bien deben tomarse en un número determinado por todos los alumnos, son aquellas que pueden ser elegidas libremente entre diversas opciones. Estas no tienen un contenido preestablecido y este puede variar de acuerdo con la iniciativa del profesor de semestre a semestre, previa autorización del Comité Académico del colegio.

facultad. De hecho, el curso comenzó a impartirse en el semestre 2020-1, cuando los autores de este texto habíamos iniciado el desarrollo de ese plan de estudios de nivel posgrado⁷.

Constituimos un seminario optativo porque, dentro del plan de estudios de la carrera de Filosofía, ese tipo de seminarios están pensados como espacios de especialización, por lo que se ajustaba a nuestros intereses y nos libraba de dificultades administrativas, pues la elección temática es libre por parte del profesor. Además, los seminarios funcionan de manera más horizontal. A diferencia de la cátedra tradicional, en la que el profesor es el principal exponente, en el seminario se busca una interacción más rica entre alumnos y profesores, a partir de exposiciones hechas por los primeros. El modelo resultaba conveniente para el curso porque permitía a los alumnos experimentar con métodos y herramientas, trabajar de manera colaborativa y mostrar iniciativa, mientras el profesor funge como guía y asesor. De esta manera, en clase, los estudiantes dejan de ser unos meros consumidores de información y pasan a tener un papel activo en la producción de conocimiento.

El seminario optativo podía servir tanto para ir delineando en la práctica algunos de los contenidos y las estrategias pedagógicas que podrían aplicar más adelante en la especialización como para comenzar a preparar estudiantes que pudieran estar interesados en especializarse en el área, ya que al impartirse formalmente en un curso regular se podían alcanzar ambos objetivos.

En cuanto a qué y cómo enseñar HD en este contexto, el mayor reto lo representaba, de nuevo, la disciplina de origen del

7 La especialidad continúa su proceso institucional para crearse; a la fecha de escritura de este capítulo, continúa siendo un proyecto.

estudiantado. Este proviene en su totalidad de la carrera de Filosofía, una disciplina que ha mostrado poco interés por las HD. De acuerdo con *The index of digital humanities conferences*, que reúne las ponencias presentadas en quinientos congresos de HD, en sesenta y un años solo el 0,8% del total, es decir, sesenta ponencias, han sido clasificadas “filosofía” como tema principal o rector (Weingart *et al.*, 2020).

La escasa participación de la filosofía en las HD tiene dos implicaciones principales a la hora de pensar en un curso dirigido a estudiantes de esta disciplina. La primera es de enfoque: qué relación ha de establecerse entre las HD y la filosofía. La segunda es didáctica: qué tipo de recursos hay para la enseñanza de las HD para filósofos en español y al alcance de una población con recursos limitados.

Respecto a la primera, hay algunas reflexiones. Por ejemplo, hace una década Peter Bradley (2011) se preguntaba en un *post* dónde están los filósofos tras asistir a un The Humanities and Technology Camp (THATcamp). En esa misma entrada, después de dar cuenta del interés filosófico por la tecnología y de proyectos digitales como PhilPapers⁸, afirmaba: “Sin embargo, no hay ejemplos de filósofos que utilicen técnicas de humanidades digitales para hacer *filosofía* o que utilicen herramientas digitales para enseñar filosofía”⁹. La precisión es importante por dos cosas. Primero, porque distingue adecuadamente entre la aproximación

8 Véase <http://philpapers.org/>

9 “There are not, however, numerous examples of philosophers using techniques of the digital humanities to do *philosophy* or using digital tools to teach philosophy”. Traducción propia.

teórica a la tecnología de la incorporación de los métodos y las técnicas de las HD al campo filosófico. Segundo, porque propone las HD como metodologías útiles para hacer filosofía.

Una aproximación similar, en este segundo aspecto, la encontramos en Cristina Marras (2019), una de las pocas filósofas en las HD. En su curso del 2019 sobre HD para filósofos para la Universidad de la Sapienza, en Roma, este tenía como objetivo: “desarrollar la alfabetización básica y las habilidades digitales necesarias para considerar críticamente las formas de conducir, presentar y organizar la información y los contenidos (datos) en y para la investigación filosófica (alfabetización informacional)”¹⁰.

En estos dos ejemplos (no hay muchos más) vemos cómo las HD aparecen como una metodología y unas herramientas para la instrucción de capacidades digitales, con el fin de llevar a cabo investigación filosófica, haciendo eco de cómo se estudian en otras disciplinas como la historia o la literatura. Estas aproximaciones, sin embargo, no contemplan la posibilidad de formar filósofos dentro de las HD para que se desarrollen en este último campo, es decir, enseñar métodos digitales no implica necesariamente incursionar en los problemas epistémicos, lógicos o éticos que subyacen a la aplicación de esas herramientas y que constituyen los fundamentos teóricos de las HD. En este sentido, se puede enseñar a los filósofos a utilizar las herramientas para reflexionar sobre sus problemas subyacentes y no solo como herramienta de trabajo. A pensar, pues, como humanistas digitales.

10 “to develop the basic literacy and digital skills needed to critically consider the ways to conduct, present and organize information and contents (data) in and for philosophical research (information literacy)”. Traducción propia.

En nuestro caso, ya que consideramos las HD como una práctica colaborativa y transdisciplinar, que se enriquece cuando no queda circunscrita a los candados impuestos por una disciplina, no intentamos enseñar metodología a filósofos, sino impulsarlos a conocer cómo los métodos se aplican y problematizan en las HD, para que, si así lo desean, ellos puedan ser teóricos de las HD y no solo practicantes.

Con esta idea en mente, Ernesto Priani Saisó diseñó el programa del curso; para ello, primero, estableció que el curso se impartiera en dos semestres no seriados, es decir, los alumnos pueden asistir al segundo semestre sin haber tomado el primero. Esto obligó a pensar cada semestre como una unidad autónoma, en la cual los alumnos se introdujeran a lo que son las HD en general, descubrieran alguno de los objetos epistémicos centrales de estas humanidades (el texto digital, los datos, por ejemplo), conocieran y aplicaran al menos una metodología de las HD y desarrollaran un proyecto personal en torno a uno de esos objetos con la metodología propuesta.

El resultado fueron dos programas de estudio independientes. El primero, correspondiente al primer semestre, se enfoca en el texto digital, su definición y problemas epistémicos, una introducción al mercado *Text Encoding Initiative* (TEI) y ejercicios específicos, con el fin de que puedan marcar un texto o grupo de textos de su elección. El segundo, en cambio, se enfoca en los datos, sus problemas epistémicos y éticos, y en una introducción al desarrollo de bases de datos, el uso de metadatos y la utilización de Voyant Tools como herramienta de análisis de textos, ambas para ser aplicadas en un proyecto a elección de cada estudiante. Dada las diferencias en el contenido de cada semestre, cada uno tiene sus

propios retos desde el punto de vista didáctico, que se han tenido que sortear durante los tres años que se lleva impartiendo.

El marcado TEI es un conjunto de etiquetas XML estandarizadas, creado por una comunidad amplia para codificar en los textos elementos de interés para los humanistas. Es una herramienta flexible y fácil de adaptar a distintos requerimientos tanto profesionales como pedagógicos. Además, constituye una de las herramientas de mayor uso en las HD; sin embargo, como ya han hecho notar Susanna Allés-Torrent y Gimena del Rio Riande (2019), las etiquetas TEI han sido desarrolladas sobre todo en el marco de la tradición de crítica textual angloamericana y británica, y aplicadas en particular en la edición digital académica y la creación de corpus digitales, con una orientación fundamentalmente filológica. Así mismo, señalan estas autoras que la adopción de TEI para proyectos digitales en español, en especial en Latinoamérica, ha sido escasa, por lo que no hay muchos recursos de enseñanza de TEI en español orientadas a un trabajo con textos filosóficos. En el curso utilizamos las lecciones y los ejemplos sobre el uso de TEI de Text Technologies Hub¹¹, pero estas tienen una orientación literaria y filológica en su exposición y en sus ejemplos que no es fácil trasladar al campo filosófico.

Por lo general, los filósofos no están interesados en la descripción del texto y su estructura tanto como en el análisis particular de dos componentes: conceptos y argumentos. Aunque las etiquetas son muy flexibles, y es posible adaptarlas, a menudo las usamos con un sentido distinto al que tienen en las guías, lo que no es un problema en términos de uso de las etiquetas, pero sí tiene

11 Véase <https://tthub.io/>

repercusiones. Un texto etiquetado puede tener dos características: estar bien formado, es decir, que sus etiquetas sigan correctamente la sintaxis de un documento XML, y ser validado contra un esquema¹², para la aplicación consistente de las etiquetas. En este sentido, la variación en el uso de las etiquetas y sus aplicaciones específicas con textos filosóficos nos permite evaluar que estén bien formadas, pero no validarlas, pues los esquemas existentes y habilitados en el *software* gratuito no lo permiten. Esto es un reto por resolver mediante un trabajo más sistemático en el uso y aplicación de TEI a textos filosóficos.

Para el caso del segundo semestre, que tiene como tema los datos y la exploración de textos asistido por computadoras, los retos didácticos tienen otras características. En primer lugar, y a diferencia de lo que ocurre con TEI, contamos con Voyant Tools¹³, desarrollado por Stefan Sinclair y Geoffrey Rockwell, un magnífico instrumento didáctico para enseñar el uso de herramientas lingüísticas, no solo por el número de visualizaciones/análisis que concentra, sino también por uso intuitivo, su acceso web gratuito y su interfaz en español.

Esto desplaza el reto didáctico hacia la explicación de los fundamentos de la herramienta y su uso, los problemas de la formación del corpus para su análisis y, por supuesto, cómo deben formularse las preguntas para indagar a través de este tipo de

12 Un esquema XML es básicamente un conjunto de reglas predefinidas con las que se describe qué características debe tener un documento XML. En el esquema se definen los elementos y los atributos asociados a estos que pueden utilizarse en el documento. También establece la forma como pueden aparecer los elementos. Es decir, formular las reglas a partir un documento XML es válido.

13 Véase <https://voyant-tools.org/>

herramientas. Como ocurre también con TEI, hay mucho material didáctico en inglés, pero poco se encuentra en español, lo que en muchos casos representa una barrera para los estudiantes¹⁴.

Sin embargo, el curso no gira únicamente alrededor del uso de Voyant Tools: este se enmarca, por un lado, en la discusión sobre qué son los datos en humanidades, cómo se trabaja en ello y, por otro, en torno a la formación de bases de datos textuales.

Este marco nos permite darle al curso la orientación que buscamos: ofrecer una idea integral de lo que se trabaja en HD, cuáles son sus nuevos objetos y qué metodologías se utilizan.

Durante casi la totalidad de los tres años, el curso se ha impartido a distancia sin ser propiamente un curso a distancia, debido a la pandemia. Este ambiente ha permitido trabajar de manera mucho más intensiva con las herramientas y los recursos en línea que como suele ocurrir de forma presencial, por limitaciones de infraestructura de la universidad. En esos años han pasado por el aula grupos que no alcanzan la decena de alumnos, compuestos en su mayoría por mujeres, con un porcentaje de conclusión del curso muy alto, quizás por el interés que desde el principio tienen en estos temas, así como por su conformación, pues las alumnas tienden a tener un compromiso mayor que los alumnos. Parte de la experiencia ha sido observar las enormes desigualdades en la formación digital de los estudiantes, pues mientras unos están más familiarizados con el uso de herramientas y conceptos, otros las desconocen.

14 Para un tutorial de Voyant Tools en español, véase <https://programminghistorian.org/es/lecciones/analisis-voyant-tools/>.

En síntesis, este curso ofrece una integración distinta entre las HD y la disciplina, aprovechando que ciertos temas de las HD, concretamente su epistemología, pueden ser abordados desde la filosofía, al mismo tiempo que los estudiantes adquieren las herramientas y los métodos de trabajo de las HD. En ese sentido, la aproximación depende de las características específicas de la disciplina a la que se integran las HD, menos generalizable que la observada en el caso previo. Un ejemplo de la estrategia de integración adoptada aquí, pero para funcionar de manera más transversal a las disciplinas, es la que presentamos a continuación.

Desarrollo de un mooc de humanidades digitales

En el 2020, la Coordinación de Universidad Abierta, Innovación Educativa y Educación a Distancia (CUAIEED) de la UNAM nos propuso el desarrollo de un MOOC sobre HD. La propuesta nos pareció atractiva por las características que definen a los MOOC: cursos de formación autónoma, masivos, abiertos y en línea (Mercado del Collado, 2018), lo que nos permitiría alcanzar a un grupo más amplio de interesados en las HD que los que pueden acceder a un curso universitario como los que hemos descrito antes. Es importante señalar que el hecho de que nuestra propuesta forme parte de los cursos en esta modalidad, dentro de una plataforma de la universidad, contribuye a la institucionalización de la enseñanza de las HD.

A diferencia de los cursos que hemos examinado antes, el MOOC está dirigido a todo tipo de públicos, entre ellos los profesionales

de distintas disciplinas que estén interesados en conocer qué son las HD, pero que no son estudiantes en un sentido formal (no están inscritos en un grado o pregrado) y recurren a este tipo de objeto de aprendizaje de manera autodidacta y autogestiva. Es posible que los usuarios busquen en este tipo de recursos elementos que abordan sus investigaciones, que tengan curiosidad con respecto al campo de estudio o que encuentren el curso complementario a sus labores/estudios, pues mantiene la característica de interdisciplinariedad de las HD.

En este sentido, el proyecto ofrecía dos retos interesantes desde el punto de vista pedagógico. El primero era poder desarrollar un curso de formación autónoma, que es muy distinto al modelo escolarizado, pero que es particularmente útil para el tipo de conocimientos que se desarrollan en las HD. El segundo era poder reflexionar sobre cómo trasladar el conocimiento del ejercicio de las HD, de una manera práctica, a un curso en línea.

Para lo primero tuvimos la suerte de contar con la orientación de un grupo de profesionales dedicados a la educación a distancia dentro de la CUAIEED, quienes nos ofrecieron su experiencia para que pudiéramos formular un programa que considerara algunos lineamientos que la literatura especializada recomienda en el diseño de un MOOC. Para diseñar el curso, tuvimos especial cuidado en los siguientes aspectos: contenidos significativos, involucramiento del estudiante, mensurabilidad de lo aprendido, accesibilidad y escalabilidad, pero ante todo que, como constantemente nos lo recordaban, tuviera una ventaja competitiva al preguntarse: ¿qué hace diferente mi curso a otros? (Vadillo y Bucio, 2018).

El proceso para diseñar el MOOC partió, por un lado, de verter los contenidos que pensábamos que deberían formar parte del curso,

mientras explorábamos diversas plataformas en busca de cursos introductorios a las HD. La demanda de objetos digitales de aprendizaje¹⁵ ha crecido de forma exponencial desde el 2007 (Kaplan y Haenlei, 2016) y es prácticamente imposible encontrar un tema sobre el que no haya un material digital donde este haya sido abordado. Sin embargo, el número de cursos encontrados no fue alto: se identificaron dos concretamente en español y tenían sobre todo un enfoque teórico, semejantes al que tenía nuestra primera aproximación. Por este motivo, tuvimos que reconsiderar lo que habíamos planteado inicialmente y preguntamos cómo podríamos formular el curso de una manera distintiva.

En términos generales, lo que nos interesaba era llevar al alumno a la reflexión sobre la importancia de los medios digitales e internet en las disciplinas humanísticas y de las ciencias sociales, además de reflexionar sobre algunas de las prácticas de las HD en el ámbito de la docencia, como las implicaciones de la mediación digital, el conocimiento abierto, la protección de datos personales y el uso de herramientas digitales para la creación de contenidos y la evaluación de objetos digitales. Recurrimos a nuestras experiencias de cursos curriculares para extraer los elementos que podrían configurar un curso de la naturaleza de un MOOC. La manera en que nos propusimos alcanzar esos objetivos fue pensar que el MOOC debería reflejar la experiencia de concebir los aspectos de un proyecto de HD propio, de manera tal que fuera una

15 Los objetos digitales de aprendizaje son conjuntos de elementos digitales articulados pedagógica, didáctica y curricularmente, con el objetivo de ser utilizados por docentes y estudiantes en la enseñanza y el aprendizaje, respectivamente.

suerte de “manos a la obra”, lo que llevará a la consecución de los objetivos propuestos.

Ahora bien, un curso de esta naturaleza requiere de una planeación distinta, ya que será el usuario quien, de forma autónoma, adquirirá una serie de habilidades y conocimientos. Por ello, en el diseño deben ser claros los objetivos y los indicadores correspondientes a través de los cuales se concretarán. En nuestro caso, se plantearon cuatro objetivos, con su correspondiente indicador: (1) esbozar un proyecto de HD: crear una propuesta; (2) aplicar criterios de evaluación para proyectos de HD y herramientas digitales: evaluar un recurso digital y autoevaluar la propuesta de proyecto; (3) utilizar herramientas para el análisis de texto, visualización de información y publicación digital: crear un corpus de texto, utilizar Voyant Tools para visualización, interpretar los resultados y socializarlos a través de la publicación en línea; (4) identificar y aplicar licencias de uso a contenidos propios: registrar un objeto digital propio con una licencia Creative Commons (cc).

Como se puede observar, cada objetivo cuenta con un indicador y este es medible a través de rúbricas u otro tipo de instrumentos, que permiten al usuario conocer el grado de cumplimiento de los objetivos al término del curso. Esta es otra de las características del MOOC que lo diferencian de los cursos escolarizados, donde la autoevaluación o evaluación por pares se realiza en el contexto de una clase. De este modo, nuestro MOOC quedó dividido en cuatro módulos y se contempla que cada uno pueda cursarse en el lapso de una semana, por lo que en un mes el participante puede terminar en su totalidad el curso. Todos los módulos se componen de conocimiento teórico y prácticas, las cuales se reflejan en actividades dentro de la plataforma y, en

ocasiones, se socializan por diferentes medios digitales. Al ser un modelo de curso autodidacta, los contenidos se componen de videos explicativos, materiales de consulta, materiales de referencia e instrucciones para las prácticas.

De acuerdo con el modelo MOOC, cada módulo debe plantear los valores, conocimientos y habilidades que cada usuario puede desarrollar. Así, decidimos promover el conocimiento abierto, tanto por la naturaleza de la plataforma que hospeda el curso como por la apertura que los contenidos tienen al estar orientados a todo tipo de público y el trabajo colaborativo e interdisciplinario, con un uso adecuado de recursos y contenidos digitales. Consideramos que los conocimientos básicos que corresponden a los valores mencionados implican aprender: los principales debates y líneas de investigación de las HD, un material que permite acercar al participante a las buenas prácticas para elaborar y evaluar proyectos de HD, las herramientas de análisis de texto, visualización y publicación, y los fundamentos de licencias para contenidos en línea.

Al término de los cuatro módulos, el participante será capaz de evaluar distintos elementos de los recursos digitales disponibles para las humanidades, al determinar la calidad, confiabilidad, documentación y usabilidad de estos. También será capaz de elegir críticamente entre distintas herramientas digitales —análisis de corpus textuales, evaluación de recursos digitales, visualizaciones en mapa, etc.— para el trabajo en humanidades y utilizarlas para crear contenidos, los cuales podrá socializar bajo alguna de las licencias de publicación.

Los MOOC deben contar con una colección de materiales (videos, lecturas y cuestionarios) organizados en secuencia para guiar el aprendizaje autónomo de los participantes. Esto implicó

pensar en qué materiales —textuales, audiovisuales, herramientas e instrumentos de evaluación— y en qué orden deberían estar estructurados para que los participantes alcanzaran los objetivos planteados: el desarrollo de un proyecto propio, así fuera de manera esquemática, y los distintos aspectos que en este se involucran, como la evaluación, la experimentación con herramientas de análisis textual, la visualización de datos y la publicación de libre acceso. Esta fase del curso conllevó la selección de materiales textuales en español, principal pero no exclusivamente, y de herramientas —las cuales solo se encuentran traducidas de forma parcial o en caso de que interactúen con el reconocimiento del idioma de la IP—, al igual que la realización de un guion y una serie de videos cortos que buscan la interacción con el participante de manera individual y la adquisición de un panorama general de las HD, sus prácticas, sus preocupaciones, etc.

Hoy en día el MOOC ya ha sido publicado y ha cumplido un año en línea. La experiencia de su creación nos ha dejado enseñanzas tanto pedagógicas como en preguntas de investigación, pues no solo es la práctica frente a cómo desarrollar y articular un curso de estas características, sino también la reflexión en cuanto a la naturaleza de los cursos de HD cuando no se tiene en mente una disciplina en específico. Si bien la articulación de los ejemplos y los materiales provienen necesariamente de proyectos o trabajos relacionados con alguna disciplina (con base en los proyectos presentados en los materiales del curso, encontramos las siguientes disciplinas rectoras: literatura, historia, filosofía y educación, principalmente), nuestro enfoque fue centrarnos en la elaboración de proyectos sin importar la disciplina de origen, es decir, hicimos énfasis en las formas como se desarrolla el campo

de las HD en cualquier espacio: cómo se plantea un proyecto, qué se toma en cuenta para desarrollarlo, cómo se valoran las alternativas tanto en métodos como en herramientas. De esta manera, es posible encontrar los puntos de intersección para que cada participante, sin importar su disciplina, alcance los objetivos del curso.

Conclusiones

Cada una de las experiencias docentes que hemos presentado aquí responden de distinta forma a la dificultad de hacer converger las HD y los campos disciplinares humanísticos en dos momentos significativamente diferentes: antes de y durante la pandemia. Además, lo hacen en el contexto de una institución donde es compleja la creación de nuevas asignaturas y la cultura académica ha sido distante de la incorporación de los métodos digitales a la enseñanza.

En el primero de los casos, que corresponde a los años previos al cierre de la universidad por la emergencia sanitaria, las HD fueron integradas como metodologías para el campo de la historia, alrededor del espacio físico del laboratorio como un lugar de experimentación colaborativa, donde estudiantes y docentes participaron del mismo proceso de construcción de conocimientos. La experiencia se extendió a la formación transversal de docentes de la facultad en los métodos de las HD y en algunas de sus discusiones, aprovechando que el modelo adoptado para la integración de las HD con la historia podía extenderse a otros campos.

El desplazamiento de los cursos presenciales al mundo virtual acercó el tema de la incorporación de la tecnología a las

disciplinas humanísticas en nuestra universidad, lo que facilitó una convergencia distinta de las HD con aquellas. Tanto en el curso de filosofía como en el MOOC, se buscó integrar las HD como campo y no solo como un conjunto de metodologías: en el primero, a partir de los problemas epistémicos de las HD que pueden explorarse desde la propia filosofía; en el segundo, a partir de la idea de proyecto como eje articulador de las HD con cualquiera de las áreas de humanidades y ciencias sociales.

En los cursos impartidos en la facultad, en las carreras de Historia y Filosofía, consideramos que logramos despertar la curiosidad del estudiantado y hacer visible la necesidad de que la enseñanza de las HD, en cualquiera de los dos modelos que hemos visto, se incorpore desde la licenciatura, porque esta debe ofrecer las bases para que las HD se integren de manera plena a las humanidades y no solo encontrarlas como una rama especializada en los posgrados.

No tenemos preferencia por ninguno de los dos modelos que hemos mostrado aquí. Pensamos que cada uno responde a condiciones y oportunidades concretas, y que conforme se avance podrán encontrarse otros equilibrios para la enseñanza de las HD en disciplinas humanísticas. Un resultado importante que se debe considerar es que en los tres casos convergimos en la búsqueda de productos finales para la evaluación y aprobación de los cursos que fueran diferentes a lo hasta ahora reconocido como producto académico. Por supuesto, hacerlo fue el resultado natural de incorporar las HD a los cursos, pero también una estrategia para mostrar las ventajas que estos productos ofrecen como recurso pedagógico, porque en muchos casos rebasan el espacio del aula y se conforman como productos académicos en un sentido más amplio.

A pesar de que, como docentes, logramos encontrar espacios para proponer y desarrollar asignaturas relacionadas con las HD, esta posibilidad sigue siendo excepcional, por lo que su creación y continuidad —o discontinuidad— depende aún de muchas coyunturas en el pregrado. Mientras esa apertura sea ocasional, es necesario seguir impulsando programas que permitan una mayor institucionalización de las HD, como diplomados o especialidades, que sean transversales a las disciplinas humanísticas. Se debe considerar también que, contando o no con una asignatura de HD, en los contenidos de los cursos del área de investigación se debe pensar en introducir estas discusiones y prácticas para acercar las HD a quienes se están formando como profesionales en las distintas disciplinas.

La labor docente de las HD en la Facultad de Filosofía y Letras, y en la universidad en general, no ha concluido, seguimos desarrollando propuestas nuevas de cursos curriculares: en el ciclo 2024-1, la Dra. Adriana Álvarez impartirá una materia optativa sobre las HD en la Licenciatura en Historia; así mismo, se dará continuidad a algunos de los cursos ya existentes, ya que el curso del Dr. Ernesto Priani se seguirá ofertando en la Licenciatura en Filosofía, además de mantener la oferta del MOOC, de la cual hemos comenzado a analizar los productos resultantes de los participantes, con la finalidad de actualizar sus materiales y contenidos. Podemos concluir que esta labor es continua y que, a pesar de las dificultades iniciales en el planteamiento de los cursos, con estos ejemplos hemos logrado pasar de “llamar la atención” a construir un currículo sólido para la enseñanza de las HD en la UNAM.

Referencias

- Allés-Torrent, S. y del Rio Riande, G. (2019). The switchover: Teaching and learning the Text Encoding Initiative in spanish. *Acta Académica*, 12. <https://www.aacademica.org/gimena.delrio.riande/174>
- Álvarez Sánchez, A. y Peña Pimentel, M. (2015, 10 de diciembre). *eLaboraHD, grupo de capacitación para humanistas. Un proyecto de la FFyL-UNAM* [conferencia]. Segundo Congreso Internacional: El Patrimonio Cultural y las Nuevas Tecnologías, Museo Nacional de Antropología e Historia.
- Álvarez Sánchez, A. y Peña Pimentel, M. (2016). *Programa de curso: Historia digital*. <http://ru.ffyl.unam.mx/bitstream/handle/10391/4890/PROGRAMA%20HISTORIA%20DIGITAL%201.%202017-1.pdf?sequence=1&isAllowed=y>
- Blackwell, C. y Martin, T. R. (2009). Technology, collaboration, and undergraduate research. *DHQ*, 3(1). <http://www.digitalhumanities.org/dhq/vol/3/1/000024/000024.html>
- Bradley, P. (2011). Where are the philosophers? Thoughts from THATCamp pedagogy. *Journal of Digital Humanities*, 1(11). <http://journalofdigitalhumanities.org/1-1/where-are-the-philosophers-by-peter-bradley/>
- Brescia, J. A. (2010). *La historiografía en el amanecer de la cultura digital. Innovaciones metodológicas, discursivas e institucionales*. Ediciones Cruz del Sur.
- Burdick, A. (Ed.). (2012). *Digital humanities*. MIT Press.
- D'Iorio, P. y Barbera, M. (2011). Scholar source: A digital infrastructure for the humanities. En T. Bartscherer y R. Coover (Eds.),

- Switching codes: Thinking through digital technology in the humanities and the arts* (pp. 61-88). University of Chicago Press.
- Dixon, D. (2012). Analisis tool or research methodology? Is there an epistemology for patterns? En D. M. Berry (Ed.), *Understanding digital humanities* (pp. 191-209). Palgrave MacMillan.
- Facultad de Filosofía y Letras. (1999). *Plan de Estudios. Licenciatura en Historia*. Universidad Nacional Autónoma de México. <http://historia.filos.unam.mx/inicio/plan-de-estudios/>
- Fyfe, P. (2011). Digital pedagogy unplugged. *DHQ*, 5(3). <http://digitalhumanities.org/dhq/vol/5/3/000106/000106.html>
- Galina Rusell, I. (2012, 18 de diciembre). Lectura en medios digitales: Apuntes sobre nuevas perspectivas. <http://humanidades-digitales.net/blog/2012/12/18/lectura-en-medios-digitales-apuntes-sobre-nuevas-perspectivas/>
- Galina Rusell, I. (2013, 8 de noviembre). Las humanidades digitales globales. <http://humanidadesdigitales.net/blog/2013/11/08/las-humanidades-digitales-globales/>
- Galina Rusell, I. et al. (2015). Formulario de evaluación para proyectos de humanidades digitales. <http://web.archive.org/web/20150321041957/http://humanidadesdigitales.net/evaluacion/>
- Kaplan, A. y Haenlein, M. (2016). Higher education and the digital revolution: About MOOC, SPOC, social media, and the Cookie Monster. *Business Horizons*, 59(4), 441-450. <https://doi.org/10.1016/j.bushor.2016.03.008>
- Koh, A. (2014). Introducing digital humanities work to undergraduates: An overview. *Hybrid Pedagogy*. <https://hybridpedagogy.org/introducing-digital-humanities-work-undergraduates-overview/>

- Lamarca Lapuente, M. de J. (s.f.). *Hipertexto: El nuevo concepto de documento en la cultura de la imagen*. <http://www.hipertexto.info>
- Marras, C. (2019). *Digital humanities for philosophy* [tesis de maestría]. Univesitá degli Studi di Roma. <https://www.lettere.uniroma1.it/node/5601/14996>
- Mercado del Collado, R. (2018). Los cursos masivos abiertos en línea y el bien común: Una reflexión frente a los objetivos de desarrollo sostenible. En R. Mercado del Collado (Coord.), *Háblame de TIC. Cursos masivos abiertos en línea (MOOC): El caso de México* (pp. 25-42). Editorial Brujas.
- Peña Pimentel, M. (2015). *Uso de tecnologías: panorama* [seminario]. Universidad Iberoamericana, Ciudad de México.
- Pons, A. (2013). *El desorden digital. Guía para historiadores y humanistas*. Siglo XXI Editores.
- Rettberg, S. y Saum, A. (2020, 2 de agosto). Introduction: Electronic literature as a framework for the digital humanities. *Electronic Book Review*. <https://doi.org/10.7273/vyqe-dq93>
- Rockwell, G. y Sinclair, S. (2012). Acculturation and the digital humanities community. En B. D. Hirsch (Ed.), *Digital humanities pedagogy: Practices, principles and politics* (pp. 177-221). Open Book Publishers. <https://doi.org/10.11647/OBP.0024>
- Svensson, A. (2014). El término humanidades digitales y los estudios latinoamericanos una revisión bibliográfica. *Anuario Americanista Europeo*, 12, 1-28. <https://dialnet.unirioja.es/servlet/articulo?codigo=5072581>
- Trinkle, D. (2015). Guidelines for evaluating digital media activities in tenure, promotion, and review. En D. L. Andersen (Ed.), *Digital scholarship in the tenure, promotion and review process* (pp. 211-213). ME Sharp Inc. <http://site.ebrary.com/id/11043072>

- Vadillo, G. y Bucio, J. (2018). Un MOOC, muchos MOOC: Diseño multinivel en cursos masivos del área de la salud. *Investigación en Educación Médica*, 7(26), 92-98. <https://doi.org/10.22201/facmed.2007865x.2018.26.03>
- Weingart, S. B., Eichmann-Kalwara, N., Lincoln, M., *et al.* (2020). *The index of digital humanities conferences*. Carnegie Mellon University. <https://dh-abstracts.library.cmu.edu>. <https://doi.org/10.34666/k1de-j489>

**O ensino de
Humanidades Digitais
no Brasil: visões a
partir de um survey
nacional**

Para citar este capítulo:

<http://doi.org/10.51573/Andes.9789587987362.9789587987379.97895879878675>

JIMMY MEDEIROS
JULIANA MARQUES
SUEMI HIGUCHI

Introdução

f Neste capítulo, exploramos o processo de integração entre pesquisa e ensino em resposta à demanda crescente por profissionais com habilidades tecnológicas e formação humanística. Nosso objetivo é contribuir para a compreensão desse fenômeno por meio da consolidação de discussões realizadas em eventos internacionais no ano de 2018, a saber: o 1 Congresso Internacional em Humanidades Digitais no Rio de Janeiro, o HDRIO2018 (Ribeiro e Higuchi, 2018), a *Digital Humanities Conference* na Cidade do México (Marques e Castro, 2018) e o *Congreso de la Asociación Argentina de Humanidades Digitales* na cidade de Rosário (del Rio Riande, 2018). A contribuição principal procura apresentar e discutir resultados preliminares de nossa pesquisa atual, intitulada “Literacia digital: modelando competências digitais para humanistas e cientistas sociais”, desenvolvida na Escola de Ciências Sociais da Fundação Getúlio Vargas (FGV CPDOC) desde 2021.

Em termos gerais, a pesquisa busca experimentar, avaliar e monitorar a relação aluno-conteúdo-professor nas disciplinas de Humanidades Digitais (HD) ofertadas pela FGV CPDOC, a fim de identificar as principais dificuldades e barreiras enfrentadas e, por conseguinte, propor ajustes e melhorias no programa do curso e em práticas pedagógicas. No bojo deste projeto, realizamos um *survey on-line* nacional com representantes dos cursos de graduação e pós-graduação do Brasil da grande área de conhecimento “Humanidades”, cujas respostas serão, em parte, apreciadas neste capítulo, sobretudo no que tange à oferta de conteúdo e à percepção sobre as Humanidades Digitais, na medida em que estamos preocupados não só com o quê e de que forma os estudantes aprendem, mas também com as posições e avaliações dos educadores.

No Brasil, o termo Humanidades Digitais (HD)¹ ganhou adesão a partir da década de 2010 e passou a figurar cada vez mais frequentemente em publicações, eventos², além de nomear

- 1 No mundo anglófono, a expressão ganha notoriedade com a publicação, em 2004, do livro *A Companion to Digital Humanities*, editado por Susan Schreibman, Ray Siemens e John Unsworth, pela editora Blackwell. Um dos primeiros guias a compilar reflexões sobre História Digital foi organizado por Roy Rosenzweig e Daniel Cohen (2005), logo no ano seguinte.
- 2 Os autores compilaram um banco de dados com grupos de pesquisa, nomes de pesquisadores, eventos e publicações no Brasil encontrados a partir da utilização de uma variedade de palavras-chave, em pesquisas no site de busca Google, na Plataforma Lattes de currículos acadêmicos e no Diretório de Grupos de Pesquisa, ambos do Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), e por meio do serviço automatizado Google Alerts, para buscas com o termo “Humanidades Digitais”. Os dados estão sendo sistematizados e já permitem ver o aumento brusco de iniciativas a partir do ano de 2015. Daniel Alves (2021) também chega a uma avaliação similar.

duas associações, a Associação das Humanidades Digitais em língua portuguesa (AHDig), criada em 2013, e a Associação Brasileira de Humanidades Digitais, criada em 2021. O 1º Seminário Internacional em Humanidades Digitais, ocorrido em 2013 na Universidade de São Paulo, para além do supramencionado HDRio2018, amplificaram a organização de encontros de pesquisadores que se identificam nesta grande área. Assim como Ribeiro *et al.* (2020) no panorama que fazem das HD no país, também identificamos que são nestes espaços de disseminação científica que se buscou o fortalecimento das iniciativas existentes, o impulsionamento da produção acadêmica e o favorecimento de parcerias e trabalhos colaborativos.

Trabalhamos com a hipótese inicial de que a oferta de conteúdos de Humanidades Digitais nos cursos de graduação e de pós-graduação no Brasil é pouco expressiva se comparada à produção intelectual e à proliferação de grupos de estudos e laboratórios sobre o tema. Com o questionário, investigamos essa suposição e as principais características do ensino *com* e *sobre* tecnologias digitais, na expectativa de que também possamos extrapolar achados e reflexões para outros países e regiões. Por HD, entendemos o uso de métodos e recursos computacionais mais sofisticados na pesquisa e no trabalho em Humanidades, além do estudo ou da reflexão sobre como as tecnologias digitais impactam a experiência humana e a vida em sociedade.

O capítulo está organizado da seguinte forma: nesta introdução, registramos o contexto de desenvolvimento de nossas reflexões sobre o ensino de Humanidades Digitais no Brasil e o escopo do capítulo. Na próxima seção, oferecemos um breve histórico do processo de dupla afetação entre a configuração do campo

das Humanidades e o desenvolvimento de tecnologias digitais. Na terceira seção, descrevemos a metodologia do *survey* sobre ensino de HD no Brasil, além do processo de seleção dos participantes da amostra. Na quarta seção, compartilhamos alguns dos achados da pesquisa, selecionando os resultados que consideramos mais relevantes para a presente discussão. Na última seção, propomos uma reflexão sobre os modelos atuais de ensino de Humanidades Digitais no Brasil e traçamos as considerações finais.

50 Anos de Revolução Digital e as respostas das Humanidades

A popularização dos computadores pessoais a partir da década de 1970, dos *smartphones* nos anos 2000 e a massificação do acesso à Internet encurtaram as distâncias entre as pessoas, permitiram maior interatividade, sobretudo com a web 2.0³, intensificando,

- 3 As primeiras redes a conectar computadores só conseguiam fazê-lo em comunidades específicas, por exemplo, em uma empresa ou no campus de uma universidade. Na década de 1990, um protocolo de comunicação digital padronizado permitiu conectar “redes de redes” de computadores, na famosa World Wide Web (www). Apesar das muitas e profundas desigualdades sócio-digitais, podemos afirmar que o acesso à Internet se massificou, possibilitando a comunicação instantânea numa escala sem precedentes. No Brasil, são mais de 150 milhões de usuários de Internet e, no mundo, são cerca de sete bilhões (Comitê Gestor da Internet do Brasil, 2020). A web 2.0, que emerge no início dos anos 2000, se caracteriza pela interatividade e maior participação dos usuários na produção e compartilhamento de conteúdo e recursos digitais. Agora, está em marcha a evolução dessa web para uma chamada Web Semântica ou Web 3.0, que dá maior autonomia para os usuários ao passo que é também extremamente apoiada em inteligência artificial.

por um lado, a comunicação e o intercâmbio, e por outro, a digitalização da vida social, tanto na sua dimensão institucional quanto na individual. Há desafios e oportunidades nesta era de grandes volumes de dados e de transformações nas relações culturais, econômicas e sociais, mediadas por tecnologias digitais, que impactam desde atividades comuns do cotidiano até tarefas ultrassofisticadas, inclusive na geração de conhecimento (Orton-Johnson e Prior, 2013; Castells, 2013; Rainie e Wellman, 2012).

Nas atividades de pesquisa científica, a ampliação do universo de fontes de informação e o uso de ferramentas computacionais proporcionaram uma renovação tanto em termos de métodos quanto de reflexão teórico-crítica e epistemológica. Hoje, quase todas as problemáticas tradicionais do ofício de um pesquisador – da delimitação de uma hipótese à descoberta, acesso e gestão dos dados, da construção dos fundamentos narrativos à comunicação dos resultados – passam, em parte ou no todo, pela tela do computador. Inevitavelmente, as formas de estudar a sociedade e refletir sobre ela sofreram grande impacto também, alcançando as práticas de ensino-aprendizagem.

No Brasil, antes da popularização da expressão “Humanidades Digitais”, várias áreas das Ciências Humanas⁴, da Linguística e das Letras e Artes experimentaram a emergência de “novos subcampos” moldados sob uma natureza mais interdisciplinar e que

4 No Brasil, a área “Ciências Humanas” abarca todos os campos de conhecimento que tem como foco o estudo do ser humano enquanto indivíduo e como parte da sociedade, como a Arqueologia, as Ciências Sociais, a Educação, a Filosofia, a Geografia, a História, a Psicologia e a Teologia. Neste capítulo, quando nos referimos às Humanidades, incluímos nelas as Ciências Humanas, a Linguística e as Letras e Artes.

visavam abarcar as transformações decorrentes do desenvolvimento e da massificação do uso de eletrônica, da computação e da automação avançada, que vão aumentando a conexão entre o mundo material ou físico e o mundo digital ou virtual. Este é o caso da Arte Digital, da História Digital, da Netnografia ou da Antropologia da cibercultura, ou ainda da Linguística Computacional, entre tantas outras denominações.

Na virada do milênio, nas Ciências Sociais, que é o campo principal de atuação dos autores, os impactos das tecnologias emergentes logo foram debatidos em trabalhos como os de Abbot (2000), que defende a necessidade de se adotar novos métodos e abordagens, como análise de rede e simulações de computador, ou de Witte (2012) que, além de sumarizar novas abordagens de pesquisa, aponta a resistência de cientistas sociais em admitir que as mudanças produzidas pela revolução digital são tão profundas quanto aquelas provocadas pela revolução industrial.

Em publicações brasileiras do mesmo período, o debate é colocado de forma ampla por Baltar e Baltar (2013) e por Conte *et al.* (2013), que chamam a atenção para o volume e a variedade de dados de relevância para a pesquisa social, incluindo dados textuais, numéricos, sonoros e de imagem provenientes de fontes nato-digitais e digitalizadas, ou ainda por Noiret (2015), que fala de uma “perturbadora virada digital” a evidenciar o caráter público da produção do conhecimento, bem como a pressionar pela adoção e a criação de novos métodos de investigação.

Na área da História, no mundo lusófono, Carvalho e Lucchesi (2016) passam a discutir a História Digital e consolidam-se grupos que advogam pela importância de abordagens quantitativas e da construção de recursos eletrônicos – como bases de dados e

coleções de documentos históricos – para análises computacionais variadas, dentre as quais destacamos o mapeamento digital de fontes históricas (Villa, 2015; Villa e Gil, 2017; Ferla *et al.*, 2020b) e a identificação de eventos e padrões a partir do uso de *corpora* (Gao *et al.*, 2020; Marques, 2021; Higuchi, 2022).

Nos estudos literários, Moretti (2000) e Jockers (2013) defendem há algum tempo o emprego de métodos estatísticos e uso de automação para a observação de vastas coleções textuais, numa abordagem que passou a ser chamada de Leitura Distante, aliçada principalmente na Informática e na Linguística Computacional. Por fim, movimentos análogos acontecem também nas Artes, que, modernizadas, propõem linhas de estudos abrangendo novas mídias, design, realidade virtual, expressões digitais, dentre outras, conforme discutido em Gómez (2019). Mais recentemente, as artes digitais têm contribuído para as discussões sobre pedagogias das Humanidades Digitais, em iniciativas inovadoras como as de Graham (2021) e de Musiol (2021).

Inegavelmente, a expansão da capacidade dos recursos informáticos e a proliferação dos *softwares* de pesquisa – principalmente os concebidos sob licenças gratuitas – permitiram esse movimento de renovação nas Humanidades, transformando a forma como os dados são tratados, examinados e, por vezes, ressignificados. No Brasil, proliferaram-se análises críticas acerca do impacto das tecnologias digitais na cultura, na política e na sociedade, em suas diversas manifestações, como as pesquisas de Tarcízio Silva (2022) sobre racismo algorítmico e de Luiz Valério Trindade (2022) sobre discursos de ódio nas redes sociais. A articulação entre estudos humanísticos e tecnologias digitais constitui, assim, o cerne das chamadas Humanidades Digitais.

Ora as HD representam um novo campo transdisciplinar de estudos (Schreibman *et al.*, 2004), ora são enquadradas como um simples conjunto de novas práticas de pesquisa e de valores adotados por uma comunidade, conforme postula o *The Digital Humanities Manifesto 2.0*, publicado originalmente em 2008 (Multitudes, 2015). Elas são, sem dúvida, recentes, perfazendo cerca de duas décadas, mas impõem de imediato o investimento na aquisição de novas habilidades por parte daqueles interessados em atuar neste campo. Tanto os nativos digitais quanto os chamados imigrantes digitais, pessoas que se socializaram antes da popularização da Internet, necessitam de “letramento” ou “literacia” digital, ou seja, conhecer a lógica computacional e um conjunto de ferramentas relevantes para o seu trabalho, como pode ser o caso da estatística ou do uso de programas de computador específicos, além de habilidades sócio-emocionais que facilitem o trabalho em grupo, com profissionais de outras áreas. Esse investimento em novas capacidades representa, para alguns, um muro elevado ou mesmo intransponível e, diferentemente, para outros, uma pequena barreira de pista de atletismo.

Embora existam pontos de contato entre as áreas de pesquisa e de ensino, suspeitamos que o conteúdo formativo dos cursos no Brasil não tem acompanhado as mudanças percebidas nas práticas de pesquisa. Ou seja, discussões e experimentações voltadas para as humanidades digitais não são levadas para as salas de aula em disciplinas obrigatórias. Alves (2021 e 2016) percebe o mesmo padrão em Portugal e, em análise sobre o ensino na área de História desde os anos 1990, ele argumenta que “a prática se tem sobreposto à teoria, a investigação ao ensino e o trabalho em comunidade à institucionalização acadêmica” (Alves, 2021, p. 7).

Hirsch (2012) e Cordell (2016) identificaram que, no mundo anglo-saxão, passou-se quase uma década desde a “fundação” das HD para que a comunidade passasse a tratar de temas relacionados à pedagogia digital, mas ainda com muita resistência. A primeira autora atribui esse descompasso a uma disputa que hierarquiza pesquisa e ensino, relegando a segunda dimensão à condição de parênteses metafóricos das reflexões desenvolvidas na academia. Sem tomar partido nesta contenda específica, gostaríamos de, por ora, sublinhar apenas o potencial crítico e criativo da prática das HD em sala de aula, em que professores, instrutores, técnicos e alunos pensam e criam juntos, fazendo emergir novas práticas pedagógicas que estão ainda por ser mais descritas e sistematizadas, como em trabalhos inspiradores como os de Ferla *et al.* (2020a), Alves (2021) e Christian-Lamb e Shrout (2017), todos eles destacando o poder de agência de alunas e alunos.

Nossa avaliação sobre a mesma discrepância no Brasil tomou como base a ausência do tema - ensino de HD - nos principais congressos e eventos acadêmicos, bem como nas revistas mais bem avaliadas no país⁵. Além disso, até o lançamento do *Programming Historian* em Português⁶, em 2021, não havia nenhuma iniciativa reunindo recursos pedagógicos voltados para as HD no Brasil, como o *Digital Pedagogy Toolkit* (Christie, 2017) ou o grupo

- 5 Revistas do estrato “Qualis A”, de acordo com a avaliação da Capes - Coordenação de Aperfeiçoamento de Pessoal de Nível Superior, agência ligada ao Ministério da Educação.
- 6 Portal que publica tutoriais acessíveis, avaliados por pares, que ajudam os humanistas a aprender uma ampla gama de ferramentas digitais, técnicas computacionais e metodologias de trabalho que facilitam a pesquisa e o ensino. Mais informações em <https://programminghistorian.org/pt/sobre>.

no *Zotero* sobre Educação em HD⁷. Cursos brasileiros tampouco são registrados no *DH Course Registry*⁸ (Wissik *et al.*, 2020) ou em outros registros internacionais, tornando-os um *proxy* ineficaz para a compreensão do cenário de ensino de HD do país. Por fim, programas inteiramente dedicados às Humanidades Digitais só agora estão sendo discutidos, e o único Mestrado em Humanidades Digitais do país foi instituído em 2019 na Universidade Federal Rural do Rio de Janeiro em Nova Iguaçu (Ribeiro *et al.*, 2020). Como o acesso aos programas (objetivos e bibliografias) de disciplinas ofertadas no ensino superior no país é difícil, optamos por investigar esse e outros pontos por meio de um questionário de amplitude nacional.

Resultados preliminares mostram que, quando há alguma oferta de conteúdo nesse sentido, ela ocorre mais por iniciativa individual do docente do que em função de uma ação planejada do colegiado do curso. Por outro lado, os discentes buscam este conteúdo fora das suas instituições de ensino, em cursos livres, *sites* e portais da Internet. Esta é apenas uma porta de entrada em uma discussão sobre como as Humanidades Digitais enfrentam desafios organizacionais, geopolíticos e teórico-práticos, principalmente em sua democratização por meio do ensino. Acreditamos que compartilhar experiências e soluções para problemas comuns é um dos caminhos a seguir. Nas próximas seções, apresentamos a metodologia do *survey* e seus principais resultados.

7 https://www.zotero.org/groups/25016/digital_humanities_education.

8 <https://dhcr.clarin-dariah.eu/>.

Metodologia da pesquisa de *survey*

Conforme já foi mencionado, para realizar o mapeamento do ensino das Humanidades Digitais no Brasil, conduzimos um *survey on-line* (Babbie, 1999) com representantes – preferencialmente, coordenadores e coordenadoras⁹ – dos programas de pós-graduação e dos cursos de bacharelado da grande área das Humanidades. Cabe registrar, uma vez mais, que o acesso a informações sobre ensino no nível superior não é fácil, tendo em vista que muitos cursos não disponibilizam os programas de suas disciplinas na página das suas instituições na Internet. A minoria que o faz, geralmente disponibiliza apenas os programas do semestre corrente, quando não apenas na Intranet. Além disso, os professores também não costumam incluir informações completas e atualizadas das disciplinas ministradas em seus currículos na Plataforma Lattes, um sistema de currículos mantido pelo governo. Dessa forma, julgamos que a melhor forma de levantar essas informações seria por meio de uma consulta direta.

A princípio, não estendemos o convite a outros atores importantes do processo, como corpo docente e discente, porque buscamos informações institucionais que acreditamos ser pertinentes aos representantes dos cursos. Ademais, entendemos que a autodeclaração como humanistas digitais não é suficiente para

9 No Brasil, coordenadores/as de programas de graduação e de pós-graduação são professores/as permanentes da instituição, ou seja, pesquisadores/as experientes que, usualmente, acumulam funções de pesquisa, ensino e gestão. É comum que a posição seja ocupada por mandato, sendo rotativa entre os professores permanentes, mas há variação neste formato institucional em instituições de ensino privadas.

responder à rica dinâmica de transformação das disciplinas, e o próprio enquadramento da pesquisa poderia incentivar certos grupos de respondentes enquanto excluía outros, na medida em que existem disputas políticas, institucionais e epistêmicas que podem levar à adoção de diferentes rótulos e categorias para a definição de grupos e de abordagens, conforme vimos na seção anterior – Ciências Sociais Computacionais, Linguística Computacional, História Digital, dentre outras.

O recorte da área de conhecimento “Humanidades” que aqui empreendemos, parte das definições do Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) para as áreas de “Ciências Humanas” e de “Linguística, Letras e Artes” e inclui os seguintes cursos: Antropologia, Arqueologia, Artes, Ciências Sociais, Ciência Política, História, Letras, Linguística, Relações Internacionais e Sociologia. Para efeitos desta primeira incursão ao mundo das HD brasileiras, ficam de fora disciplinas que têm uma história mais específica de articulação com ferramentas digitais, como a Geografia e seus Sistemas de Informação Geográfica e outros suportes tecnológicos, sobretudo na Geografia Física, ou que nos parecem mais distantes da comunidade de práticas que procuramos abarcar, como a Filosofia, a Teologia e a Educação.

Tendo em vista o universo delimitado para este estudo no âmbito das Humanidades, contabilizamos 803 departamentos de graduação e 226 programas de pós-graduação no país. A quantidade elevada de cursos justifica, dessa forma, a seleção de uma amostra aleatória para a pesquisa. Para compor a amostra da graduação, levamos em consideração a categoria administrativa – a saber: privada sem fins lucrativos, privada com fins lucrativos, pública

estadual, pública federal e pública municipal – e a Classificação Internacional Normalizada da Educação Adaptada para Cursos da Graduação (CINE), que, em linhas gerais, define o enquadramento dos cursos. Para a pós-graduação, a seleção se deu por área de avaliação da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (Capes).

Após o sorteio de 285 cursos de graduação e 218 programas de pós-graduação, estratificados por região do país, identificamos uma primeira dificuldade operacional para a pesquisa: grande parte dos *sites* dos cursos selecionados não informa o nome do(a) coordenador(a), nem o *e-mail* de contato, inviabilizando nestes casos o envio do convite. Deixamos, assim, uma primeira recomendação aos departamentos e programas: verifiquem se esta informação está pública, visível e atualizada, pois é do interesse de alunos e pesquisadores, e confere transparência à gestão dos seus cursos.

Após empreender esforços extras para encontrar os contatos dos representantes sorteados para a amostra, obtivemos 212 candidatos aptos a preencher o questionário *on-line*. Deste total, 14 acabaram optando por não receber novos contatos, recusando a pesquisa de forma explícita (opção *opt-out* disponível no convite), ao passo que 54 preencheram o questionário em sua completude e outros 7 de forma parcial. Assim, a análise dos dados para este trabalho contou com um conjunto de 61 respostas válidas (taxa de resposta de 29%).

A pesquisa transcorreu por 33 dias, entre os dias 1 de abril e 3 de maio de 2022. Foram, portanto, 33 dias para captar participantes, usando apenas contatos via Internet, sem uso de abordagens

telefônicas. Através da plataforma Lime Survey¹⁰, todos os 212 integrantes da amostra foram contatados via *e-mail*. O disparo dos convites e lembretes teve periodicidade semanal, variando o dia e o horário para buscarmos diferentes possibilidades de receptividade do respondente.

Cada *e-mail* de convite foi personalizado com um *link* único e individual, e controlado por uma chave distinta para evitar ser preenchido por pessoas fora da amostra. Paralelamente, criamos uma página de divulgação da pesquisa no portal de nossa instituição¹¹ e divulgamos a realização do *survey* por meio de nossas redes institucionais e pessoais, abrindo a possibilidade de outros coordenadores se voluntariarem para participar (tivemos dois interessados). Também solicitamos às associações nacionais de algumas áreas que divulgassem a pesquisa para seus membros mas, com a exceção da Associação Nacional de Pós-Graduação e Pesquisa em Ciências Sociais (ANPOCS), não temos informações sobre o atendimento da solicitação.

Com relação ao questionário em si, buscamos contemplar questões que nos permitissem mensurar variáveis relevantes sobre o ensino das Humanidades Digitais no Brasil, na atualidade. São 32 perguntas agrupadas em cinco seções dispostas da seguinte forma: i) informações gerais do curso e da instituição de ensino a que se vincula; ii) percepção geral sobre as Humanidades Digitais e o espaço que elas ocupam nos cursos; iii) atividades e conteúdos relevantes e/ou em oferta; iv) governança, parceria e sustentabilidade

10 Software livre para aplicação de questionários online. Mais informações em <https://www.limesurvey.org/>.

11 <https://cpdoc.fgv.br/ensino-hds>.

para o seu estabelecimento; e v) espaço livre para comentários, sugestões e críticas à pesquisa ou ao tema. Nem todas as perguntas são discutidas neste trabalho, dado o espaço e o recorte proposto.

Percepções sobre o ensino das Humanidades Digitais no Brasil

A amostra obtida para este estudo é composta por 59% de respostas concedidas por coordenadores(as) de programas de pós-graduação – equivalente a 36 questionários – e 41% de coordenadores(as) de cursos de graduação – equivalente a 25 questionários. A pesquisa pode ser considerada nacional dado a distribuição das respostas por 16 unidades da federação das cinco regiões do país, a saber: 49% das respostas são do Sudeste, 28% do Nordeste, 12% do Sul, 8% do Centro-Oeste e 3% do Norte do país.

Uma primeira variável da pesquisa selecionada para este trabalho tratou da oferta de conteúdos que versam sobre Humanidades Digitais no âmbito dos cursos pesquisados. A pesquisa identificou que 44% dos participantes declararam que os cursos de graduação ou pós-graduação que coordenavam já ofereceram algum curso relacionado ao tema. Neste universo, as opções de modalidade poderiam ser: uma disciplina, uma oficina, um minicurso ou outra modalidade qualquer. O gráfico 5.1 mostra que a distribuição ficou bastante equilibrada, com proporções de ocorrência similares. A modalidade que ganhou um pouco mais de destaque foi a oficina – cujo formato é menor e, em geral, adotado para o ensino de um recurso computacional específico, como R, SPSS, Python ou Omeka – seguida pela disciplina, que demanda um

planejamento maior mas permite um aprofundamento no conteúdo, além de proporcionar créditos para a formação dos alunos.

Decompondo os 44% de cursos que já ofereceram algo relacionado às Humanidades Digitais nas diferentes opções mostradas no gráfico 1, encontramos a menção de uma modalidade em 18% deles, duas modalidades em 12%, três modalidades em 10% e quatro modalidades em apenas 4% do total dessa amostra. Assim, se por um lado a pesquisa identificou uma elevada oferta de conteúdo das Humanidades Digitais na amostra pesquisada, por outro lado, a ocorrência das diferentes modalidades listadas acima não é uniforme nos diferentes cursos. Ou seja, nem todas são empregadas nos cursos presentes na amostra. O que chama a atenção, porém, é que a modalidade “disciplina regular” (obrigatória ou eletiva), que configura a célula básica de estruturação de conteúdo em um curso

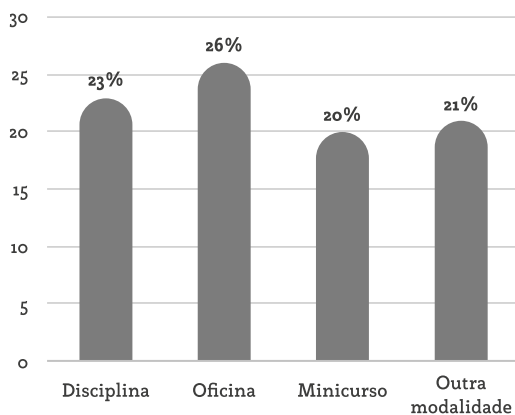


GRÁFICO 5.1. Modalidade de curso sobre Humanidades Digitais oferecido (% de sim)

FONTE: elaboração própria.

de formação, figurou em apenas 23% dos cursos participantes na pesquisa. A baixa incorporação do tema no ementário dos cursos nos leva a crer que a inclusão do tema na grade curricular é recente. Em seguida, abordamos esta temporalidade, face à hipótese da circulação transnacional de ideias e de formatos institucionais.

Como vimos discutindo, a literatura especializada no Norte Global marca a emergência da HD a partir do ano de 2004 e uma busca simples por livros indexados pela base Google Books nos ensina que só houve um crescimento significativo e sustentado de publicações a partir de 2008-2010 (gráfico 5.2). Não temos dados a respeito da utilização dessa literatura no ensino no Brasil, mas nossa pesquisa mostra que houve um aumento significativo na oferta de cursos a partir de 2015 (gráfico 5.3), acompanhando a já mencionada explosão de publicações, eventos e criação de grupos de pesquisa. Considerando que professores brasileiros também fazem e pensam as HD desde o início dos anos 2000, e assumindo que parte deles já estivesse em contato com a literatura em inglês a respeito desse campo emergente, isso não implicou, também no caso brasileiro, que o debate ou que as novas técnicas de pesquisa tenham sido logo levados para a sala de aula.

A ampliação do uso da programação e diversos *softwares* livres nas humanidades passou a ter novos contornos e relevância, e o resultado presente no gráfico 3 ilustra a hipótese de aproximação temporal entre a popularização de ferramentas computacionais, como a difusão do uso da linguagem de programação em R nas Ciências Sociais, sobretudo na Ciência Política, e o período em que os cursos presentes na amostra da pesquisa passaram a oferecer conteúdos das Humanidades Digitais. Uma hipótese que levantamos é que o surgimento de novos projetos de pesquisa

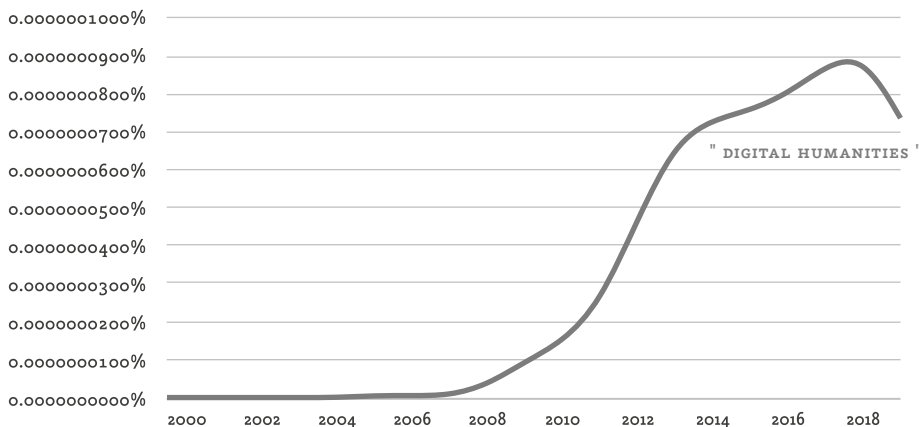


GRÁFICO 5.2. Ocorrências do termo “Digital Humanities” no corpus de livros do Google Books de 2000 a 2019

FORTE: Gerado a partir de busca dos autores na ferramenta Ngram Viewer do Google Books.

usando ferramentas computacionais acessíveis, com interfaces amigáveis que não exigem programação ou conhecimento avançado de bancos de dados, contribuiu significativamente para a disseminação das Humanidades Digitais no Brasil. Esses projetos inovadores foram divulgados por meio de publicações acadêmicas, como os trabalhos de Lage e Godoy (2008) e Campos (2014), incentivando outros pesquisadores a explorar esses recursos em suas próprias pesquisas e em experiências de ensino de curta duração, como oficinas e cursos de verão.

Em particular, os representantes dos cursos que já ofereceram conteúdos próprios das Humanidades Digitais na modalidade “disciplina” – equivalente a 14 respondentes ou 23% da amostra

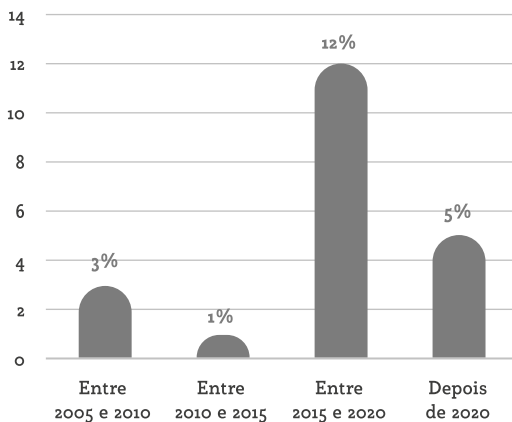


GRÁFICO 5.3. Período em que os cursos passaram a oferecer conteúdos das Humanidades Digitais (quantidade de menções)

FONTE: elaboração própria.

(gráfico 5.1) – foram questionados a respeito das características de como elas são instituídas. Em apenas um curso a disciplina é definida como “obrigatória”, sendo, portanto, oferecida de forma sistemática ao corpo discente e presente no Projeto Pedagógico do Curso, ou seja, certamente demandou uma discussão anterior do colegiado. Outros 4 entrevistados informaram que a disciplina é eletiva e disponibilizada com certa regularidade, no entanto, sem obrigação de oferta, ao passo que outros 9 entrevistados apontaram que as suas disciplinas também são eletivas, mas oferecidas eventualmente, e sendo assim, não há a obrigatoriedade, nem recorrência sistemática.

Internacionalmente, a participação em projetos de pesquisa em HD parece ter sido a porta de entrada para muitos estudantes e profissionais adquirirem experiência e treinamento com tecnologias digitais. Spiro (2012) argumenta que esse caminho não é escalável e não distribui equitativamente as oportunidades de aprendizagem, o que só seria possível por meio de reformas de currículo que, para garantir continuidade e consistência, precisam incluir disciplinas obrigatórias de base, que promovam o letramento digital específico para a área.

A oferta de disciplinas e minicursos passa pela *expertise* do corpo docente. É preciso que ao menos um professor ou professora da instituição tenha condições – e interesse – de lecionar sobre a temática. Neste sentido, uma das perguntas colocadas no estudo tenta dimensionar este aspecto, questionando “qual o percentual do seu corpo docente é dotado de conhecimento teórico e/ou prático para ministrar um curso do campo das Humanidades Digitais?”. Trazendo as respostas para o gráfico 5.4, observamos que 67% das ocorrências figuram na parcela dos 20 a 50% com aptidão para lecionar sobre este conteúdo e somente 10% dos cursos têm menos de 20% de docentes aptos. O alto percentual surpreende e talvez seja explicado pelo fato de a pergunta não abordar em detalhes que tipo de conteúdo, *softwares* e temáticas estes docentes têm condições de ministrar. De todo modo, ao nosso ver, há claramente um terreno fértil para a alavancagem das HD no ensino superior no país. O resultado nos parece extremamente positivo, e um aumento da demanda discente por disciplinas, minicursos e oficinas nesta seara tende a pressionar por mudanças.

Ter um corpo docente preparado para esta nova realidade é, portanto, fundamental. Perguntamos se o departamento ou

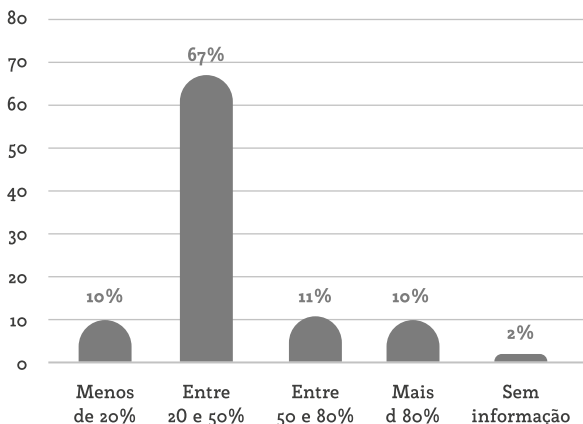


GRÁFICO 5.4. Parcela do corpo docente do Departamento de graduação ou do Programa de Pós-graduação dotado de conhecimento teórico e/ou prático para ministrar um curso do campo das Humanidades Digitais (%)

FONTE: elaboração própria.

programa de pós-graduação havia admitido especialistas em HD ou métodos digitais nos últimos 10 anos, e 15% dos questionados responderam positivamente. Em pergunta aberta, consultamos os coordenadores a respeito de ações voltadas a treinar ou familiarizar o corpo docente em ferramentas e questões das Humanidades Digitais. Identificamos que, em decorrência da pandemia do coronavírus, foram oferecidos treinamentos no manejo de ambientes digitais para aprendizado à distância, como Moodle, Google Classroom, Google Meet e Zoom. Algumas respostas ressaltaram o caráter não obrigatório dessas atividades básicas de letramento em sistemas digitais de ensino à distância, além da baixa adesão

do corpo docente¹². Como era de se esperar, no entanto, o isolamento social provocado pela pandemia também aparece como catalisador, entre docentes e discentes, do interesse em pesquisas em meios digitais, o que suscitou iniciativas de debate e de capacitação em grupos de pesquisas. Esse movimento pode ter efeito sobre os currículos nos próximos anos.

Dois pontos se fazem notar a partir das respostas à essa mesma questão. O fato de apenas 23% dos respondentes ter relatado iniciativas das coordenações ou colegiados no sentido de promover a capacitação e o debate em torno das HD entre docentes revela a baixa institucionalização da resposta dada por diferentes áreas aos desafios impostos pela crescente digitalização da vida social e das formas de trabalhar, de forma que o engajamento dos professores depende quase que exclusivamente da disposição individual de cada um. Um dos respondentes afirma que uma desvantagem da forma como o ensino de HD tem operado em seu programa está relacionada à “falta de articulação com demais disciplinas no curso”. A esse respeito, gostaríamos de sublinhar a importância da ideia de projeto para uma incorporação proveitosa dos temas e oportunidades das HD e na lida com os seus desafios. Não falamos apenas em projetos pedagógicos, mas de projetos mais amplos, enquanto áreas de conhecimento e campo

12 Referência a sistemas de gestão de ensino à distância apareceu espontaneamente em respostas abertas e, embora tais recursos digitais não estejam no centro da nossa definição básica do que são as HD, há autores favoráveis a uma definição mais alargada das Humanidades Digitais, que inclua a utilização de fontes digitais de pesquisa e de quaisquer ferramentas digitais no trabalho cotidiano de ensino, pesquisa e disseminação.

profissional. Esse debate será retomado ao longo do capítulo e nos leva, de imediato, ao segundo ponto que gostaríamos de sublinhar.

Do ponto de vista qualitativo, surgem menções a disputas políticas relacionadas à “alteração de forças” dentro de departamentos de área, com a valorização de saberes e práticas que não são canônicas nas Humanidades¹³. Esse é um tema delicado, que também aparece em conversas informais em congressos e outras esferas de socialização acadêmica. Não temos dados para discutir a prevalência desse tipo de percepção; porém, desde logo, gostaríamos de defender que o estabelecimento de incentivos institucionais a pesquisas interdisciplinares e realizadas coletivamente tem grande potencial de atenuar, quando não dirimir, preocupações e disputas dessa natureza, com resultados positivos para a pesquisa, que, à medida que agrega pessoas e abordagens, ganha em diversidade e em escala.

Internacionalmente, a oferta de cursos compartilhados e a realização de pesquisas por grupos que reúnem diferentes saberes tem sido uma tendência na prática das HD (Zorich, 2008; Sula *et al.*, 2017). No entanto, no Brasil as iniciativas de ensino de ferramentas e de debates das Humanidades Digitais com parceiros internacionais ainda são bastante incipientes, uma vez que

- 13 Algumas das opiniões que corroboram esta percepção incluem “percebo que talvez um dos grandes entraves seja realmente geracional, no sentido em que claramente há uma disparidade entre pesquisadores mais jovens e aqueles com 30 anos de trabalho sobre o tema”, “[...]quando eles perceberam que, não apenas não iriam aprender, como aquele tipo de conhecimento (de ferramentas de digitais de pesquisa) iria mudar a “correlação de forças”, começaram a boicotar”, “O cerne da coisa também pode estar relacionado a uma mudança na balança de poder entre docentes ‘do passado’ (analógicos) e os pesquisadores mais jovens (digitais). Seria precisa matizar/ponderar isso”.

apenas 5 participantes indicaram a existência deste tipo de ação. No que tange à colaboração com outros departamentos dentro e fora da grande área das Humanidades com vistas ao ensino de HD, 15% dos respondentes registraram haver algum tipo de parceria e a maioria não sabia informar. A maior parte das parcerias existentes (5 casos) se dá com as áreas de Ciência da Informação ou Ciência da Computação (que constavam como opção única).

Embora nem todos os cursos participantes da amostra tenham oferecido conteúdo das Humanidades Digitais em qualquer modalidade e construído pontes colaborativas com instituições estrangeiras, a percepção dos seus representantes quanto ao grau de importância de ensinar as ferramentas computacionais ao corpo discente demonstra uma tendência nova para as Ciências Humanas. A respeito deste assunto, uma variável da pesquisa identificou que 54% dos entrevistados declararam considerar “muito importante” e 43% informaram estimar “importante” oferecer este tipo de conteúdo aos alunos, a fim de que possam desenvolver as habilidades e competências necessárias para lidar com os desafios presentes e se diferenciarem dos demais. Outros 3% optaram por deixar a pergunta sem registro (gráfico 5.5).

Desse modo, os números apontam que a oferta deste conteúdo nas Humanidades é tida como extremamente relevante para os participantes da pesquisa, e denota que será ampliada nos próximos anos. Parece importante registrar que nenhum participante marcou a opção “não é relevante”, presente no questionário. Mas isso pode ter acontecido por um fator anterior ao preenchimento do questionário, pois aqueles que consideram o tema marcadamente como irrelevante podem ter optado por não participar do estudo.

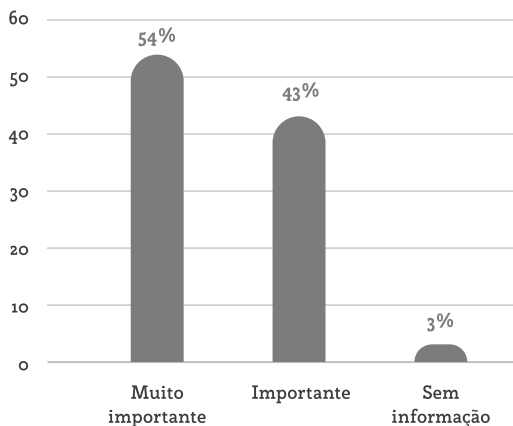


GRÁFICO 5.5. Importância do ensino das ferramentas computacionais das Humanidades Digitais para o corpo discente (%)

FONTE: elaboração própria.

Apesar do elevado grau de importância atribuído ao ensino das ferramentas computacionais, os entrevistados apontam que isso não configura um novo paradigma (gráfico 5.6). Afinal, mais de 2/3 dos participantes da pesquisa discordam da afirmativa apresentada aos entrevistados. Essa dualidade abordada na pesquisa está presente em discussões da literatura trazidas por diversos autores. Em artigo sobre a “virada computacional”, Berry (2011), que é editor e autor de duas obras de referência, *Understanding Digital Humanities* (Palgrave, 2012) e *Digital Humanities: knowledge and critique in a digital age* (Polity, 2017), discute as chamadas “ondas” de desenvolvimento das Humanidades Digitais e defende que a pesquisa mediada pelo digital precisa lidar com mudanças não apenas

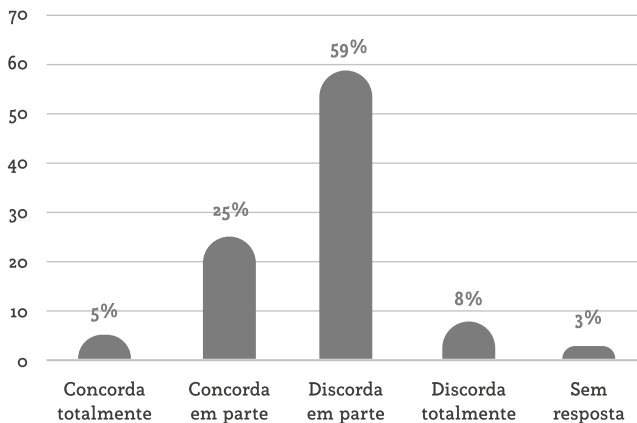


GRÁFICO 5.6. A popularização das ferramentas computacionais das Humanidades Digitais representa um novo paradigma para as Ciências Humanas (%)

FONTE: elaboração própria.

metodológicas mas também epistemológicas e ontológicas, relativas à construção do conhecimento e à natureza dos próprios objetos que buscamos conhecer. Neste ponto de vista, os especialistas em Humanidades precisam incorporar o pensamento computacional ao seu já bem consolidado ferramental teórico-metodológico e possuir um entendimento técnico de base para criticar e/ou operar metodologias e dados digitais em seus contextos sociotécnicos.

Vejam os como a discussão se dá em algumas áreas. Na Sociologia, Marres (2017) aponta a existência de duas perspectivas sobre o tema. De um lado, uma linha de pensamento que considera a sociologia digital como uma novidade e inovação para a compreensão da sociedade – uma abordagem, portanto, mais radical

e transformadora – e, de outro lado, uma segunda linha de pensamento que aborda o assunto com maior cautela e vê a “sociologia digital” como sendo a sociologia já conhecida acrescida do uso de ferramentas computacionais. Não há destaque para uma transformação do campo, apenas uma nova abordagem instrumental. Esse acréscimo não representaria uma mudança da sociologia em si, mas apenas uma forma nova de fazer análises. Já Nascimento chama a atenção para a “necessidade de os sociólogos e a sociologia fazerem um *‘upgrade’* de sua clássica fundamentação teórica e metodológica para lidar com novos fenômenos que simplesmente não existiam nas sociedades de cerca de 150 anos atrás, aquando do seu surgimento” (Nascimento 2016, p. 230). Portanto, o adensamento do ensino das Humanidades Digitais é um imperativo para análises adequadas e novas perspectivas da realidade, além de permanecer vibrante, criativa e inovadora.

Já a Antropologia Digital propõe refletir sobre as conexões existentes entre o digital, a cultura e as redes de sociabilidade (Machado, 2018), sem colocar em xeque os alicerces teórico-metodológicos da disciplina. O cerne da análise dos autores deste campo, como Miller (2011), Cesarino (2021) e Machado (2018), é compreender e refletir sobre as sociabilidades digitais, assim como sobre os usos e apropriações das tecnologias nas rotinas pelos grupos sociais. Cabe ainda apontar que os autores estão interessados em compreender as experiências a respeito da apropriação e (re)produção dos sentidos que os dispositivos tecnológicos e as redes sociais passam a ter na sociedade.

A mesma discussão sobre o impacto do digital na identidade própria da disciplina acontece em outras áreas, com diferentes desdobramentos. Na História, alguns autores mostram-se céticos

ou não veem muito sentido em criar novos rótulos para uma disciplina onde em breve – como todas as outras –, o digital estará tão incorporado nas suas práticas que será parte inerente dela (Cohen *et al.*, 2008). Na mesma linha, Alves argumenta que hoje o digital já envolve muito do que os historiadores fazem, mesmo que nem todos se apercebam, e, nessa perspectiva, daqui a uns anos falar de Humanidades será essencialmente falar de Humanidades Digitais. Atualmente, falar de História é, em grande medida, falar de História Digital (Alves, 2017). Não há, portanto, consenso e as diferentes áreas estão lidando com o avanço das tecnologias digitais de várias maneiras, dependendo de suas tradições disciplinares em âmbito nacional, questões de pesquisa e recursos disponíveis.

Outra questão aberta pedia aos coordenadores para discorrer sobre a principal desvantagem das HD. 54% dos respondentes identificaram algum malefício no ensino das HD, ainda que apenas em potencial. A sumarização das desvantagens elencadas inclui (sem ordem de frequência): 1. Considerações sobre desigualdades de acesso, por parte do corpo discente, a tecnologias digitais, incluindo conexão à Internet; 2. Falta de literacia digital prévia, tanto do corpo docente quanto do discente, o que gera alto custo na preparação de recursos humanos; 3. Custos de infraestrutura computacional para as instituições de ensino; 4. Perigo de fomentar uma percepção de que o digital (seus dados, métodos e/ou resultados) substitui ou é melhor do que o analógico/tradicional; 5. Falta de integração com as disciplinas tradicionais; 6. Riscos de um tratamento superficial de teorias e metodologias clássicas; 7. Diminuição da vivência coletiva presencial; 8. Limite ao potencial criativo/subjetivo da produção artística ou

intelectual; e 9. Necessidade de atualização constante a respeito das ferramentas e técnicas disponíveis.

Embora não seja possível distinguir até que ponto as respostas refletem uma opinião meramente pessoal ou uma percepção mais coletiva a partir de debates internos aos departamentos e programas de pós-graduação, eis aqui uma lista de temas a serem debatidos internamente, em associações de área, congressos, etc. Embora o fenômeno de digitalização da vida seja, em grande medida e sobretudo para as novas gerações, inescapável, argumentamos que a excelência no ensino de HD não será resultado de um processo espontâneo, com baixo grau de institucionalização e debates ainda muito atravessados por receios de perda de qualidade, prestígio, espaço e/ou recursos. Não se trata de dizer que as preocupações levantadas não são legítimas ou fundamentadas, mas a concentração das HD em nichos de pesquisa aponta para um debate ainda restrito e muito aquém das necessidades do tempo presente. Esperamos que a discussão em torno de novos currículos e de pedagogias digitais extravase essas questões para todo o campo acadêmico, mesmo que não estejamos todos fazendo pesquisa digital.

A respeito da formação discente em Humanidades Digitais, o questionário também indaga se ela é útil para a atuação acadêmica, para a atuação profissional fora da academia ou se para ambas. De acordo com os dados do gráfico 5.7, a percepção hegemônica é para uso generalizado, ou seja, tanto na academia como em outras posições no mercado de trabalho. Esse tipo de percepção é interessante, sobretudo, por levar a crer um entendimento de que os discentes atuam ou devem atuar também fora da academia.

Todavia, chamam a atenção os 13% de entrevistados que apontaram que este tipo de conteúdo é aplicado mais para a atuação fora da academia, reduzindo a importância do seu uso acadêmico.

Em resposta qualitativa a respeito de possíveis prejuízos das HD, está presente não só a percepção de que o investimento em habilidades com ferramentas digitais favorece a formação voltada para o mercado de trabalho fora da academia, como a de que este mesmo tipo de formação representa um perigo para as Humanidades, por responder a pressões do mercado, enviesadas pela busca e, em última instância, pelo lucro, em detrimento de valores e desafios sociais de outras ordens. Não podemos ignorar a possibilidade de instrumentalização de currículos com objetivos políticos que visam esvaziar a capacidade crítica típica das Humanidades e coibir a sua liberdade de cátedra em função de demandas mercadológicas específicas. Contudo, as discussões sobre o papel distintivo das Humanidades no contexto da estrutura universitária atual, sobre a relação entre pesquisa pura e pesquisa aplicada, e sobre a interface entre formação universitária e preparação para o mercado de trabalho, fogem ao escopo deste capítulo.

Cabe apenas sublinhar que, apesar de ainda encontrarmos resistência à popularização das HD, há inúmeros trabalhos acadêmicos que aplicam ferramentas computacionais em seu desenvolvimento e que a comunidade das HD no Brasil cresce em número de participantes e publicações, assim como no impacto no debate público, com participação em embates sobre desigualdades territoriais, de gênero e de raça, sobre a relação entre capitalismo e opressão algorítmica, entre outras pautas caras às Humanidades. Nesta seara, para além das obras citadas na Introdução, destacamos dois livros recentemente traduzidos para o português, das

autoras O’Neil (2020) e Noble (2021), sobre aspectos danosos de ferramentas digitais, quando permitem que análises e decisões mal concebidas ou mal intencionadas, no sentido de irem de encontro à ética, à democracia e à justiça social, se massifiquem sob uma pretensa aura de legitimidade ou de neutralidade, que advém de concepções equivocadas sobre o poder e os limites dos métodos e dos recursos computacionais.

Dominar a lógica computacional e as ferramentas úteis para o ofício do humanista em meio digital é fundamental para que se possa criticar e contribuir tanto com os avanços na fronteira do conhecimento quanto no campo da pesquisa aplicada, fora do ambiente estritamente universitário. Portanto, é necessário adequar



GRÁFICO 5.7. Percepção da aplicação da formação discente em Humanidades Digitais entre a atuação acadêmica e profissional fora da academia (%)

FONTE: elaboração própria.

os cursos a esta realidade, ainda que utilizar dados e métodos digitais seja uma das opções existentes, a depender dos temas, perguntas e interesses em vista. Compartilhamos esta percepção com a maioria dos coordenadores participantes da pesquisa, como vemos no gráfico 7.

Mapeamos as formações acadêmicas em nível de pós-graduação, de cursos regulares na graduação, cursos e oficinas avulsos que se auto-identifiquem no domínio das Humanidades Digitais ou que tenham como característica distintiva a aplicação de métodos digitais e abordagens computacionais a questões, desafios, dados e fontes de pesquisa. Em pergunta aberta¹⁴ sobre os conteúdos abordados em disciplinas ou cursos avulsos, obtivemos as seguintes respostas, por nós sumarizadas: pesquisa em bases de dados digitais, tecnologias digitais e comunicação, tecnologias digitais e ensino, tecnologias digitais e história pública, estatísticas populacionais, cartografia, história da informática, netnografia/ antropologia digital, sociologia digital, história digital, ciência de dados para a ciência política, mineração de texto, estudos em redes sociais, estudos críticos de algoritmos e da inteligência artificial, construção de bases de dados, produção audiovisual, pesquisa com arquivos digitais, protocolos de edição digital, processamento de linguagem natural, linguagens de programação (Python e R), análise de redes, análise de conteúdo assistida por computadores.

Neste levantamento inicial, como o número de respondentes é ainda pequeno, preferimos não quantificar a ocorrência desses

14 Também incluímos perguntas fechadas, nas quais listamos conteúdos e ferramentas específicas, mas discutiremos esses resultados mais específicos em outra oportunidade.

conteúdos, mas marcar a presença de distintas comunidades de HD em atuação no ensino de nível superior brasileiro. No geral, estão presentes os grandes tópicos da comunidade internacional, com exceção talvez de análises computacionais de fontes visuais e de patrimônio histórico-cultural. Algumas referências típicas das HD estão ausentes, como “visualização de dados”, “simulações”, “gamificação”, “realidade virtual” e “realidade aumentada”, mas isto não significa que esses tópicos não estejam subsumidos em outros listados de forma espontânea pelos coordenadores. Esse resultado reafirma o crescente envolvimento da comunidade científica brasileira com as Humanidades Digitais.

Por fim, na nossa amostra, apenas 6,5% dos respondentes consideram que o interesse do colegiado de professores não é necessário para a integração de conteúdos das HD nos currículos; 69% afirmam que é baixo o apoio de uma Superintendência de Tecnologia de Informação e Comunicação ou órgão equivalente; e 72% afirmam que é baixo o apoio da administração superior. Além da carência de infraestruturas digitais para a pesquisa e o ensino, não existem sistemas de incentivo aos docentes, com treinamento, mentoria vinculada às suas agendas e interesses de pesquisas e ações que ensejem oportunidades de colaboração interdisciplinar. Por fim, há dificuldades específicas relacionadas à defasagem no letramento digital no país. Os entraves a disseminação do ensino de Humanidades Digitais são, portanto, desafiadores, mas conhecê-los e reconhecê-los é o primeiro passo para superá-los.

Considerações finais

Este capítulo discutiu a percepção de professores-coordenadores do ensino superior brasileiro na área das Humanidades a respeito da inovação no ensino de nível superior suscitada pelo enquadramento teórico-metodológico das Humanidades Digitais. Em reflexão à luz da literatura especializada, apresentamos dados de um *survey* nacional, realizado em 2022, que buscou mapear atividades e conteúdos ofertados pelos departamentos a nível de graduação e de pós-graduação, além de informações sobre formas de cooperação e incentivos institucionais.

Internacionalmente, nos principais centros de prestígio e de acúmulo de recursos no ensino superior, a tendência consolidada é de mudança na natureza do treinamento dos alunos de graduação e de pós-graduação de Ciências Humanas e Sociais. O *Digital Humanities Course Registry*¹⁵, um inventário aberto e colaborativo de cursos de formação em Humanidades Digitais, que se concentra principalmente no contexto europeu, registra 9 programas de doutorado e 86 programas de mestrado ativos, além de centenas de disciplinas, escolas de verão, entre outras modalidades formativas. Além disso, e não menos importante, programas europeus e norte-americanos mudaram os tipos de parceria e de pedagogia e passou-se a incentivar fortemente práticas colaborativas e orientadas ao diálogo com diferentes comunidades, para além dos muros departamentais e das universidades. Ou seja, as HD do Norte têm se desenvolvido, sobretudo, em centros de pesquisa e não em departamentos acadêmicos tradicionais. Muitas das vezes, entretanto, esses centros têm

15 <https://dhcr.clarin-dariah.eu/>.

como função principal atender os diferentes departamentos, ou com eles fazer parcerias, constituindo-se em verdadeiros *loci* – ou *hubs* – de colaboração interdisciplinar (Fraistat, 2012).

No caso brasileiro, foi identificado que 44% dos cursos de graduação e programas de pós-graduação presentes na amostra já ofereceram alguma disciplina, oficina ou minicurso relacionado ao campo das Humanidades Digitais. Grande parte destes cursos iniciou a oferta destes conteúdos a partir de 2015. Dentre todos os respondentes, 85% afirmam que não têm nenhum tipo de parceria para a oferta de conteúdos das HD, nem com programas da mesma área, nem com outras áreas nas Ciências Humanas ou fora delas, o que revela que muitas iniciativas se dão de forma mais isolada dentro de cada departamento e por iniciativas individuais, com pouco ou nenhum apoio institucional.

Em uma perspectiva comparada ao contexto europeu e estadunidense, há grande espaço para a expansão da oferta de conteúdos das HD nas universidades brasileiras, mas há muitos desafios relacionados, principalmente, à falta de infraestrutura computacional e de apoio à pesquisa, além de problemas crônicos na educação básica que se refletem no baixo letramento digital, no geral. De um lado, é imprescindível aumentar o percentual de instituições a ofertar este tipo de conteúdo para configurar uma literacia digital mais ampla aos egressos de cursos das Humanidades. E, de outro lado, é fundamental estabelecer novas parcerias – nacionais e internacionais – para fortalecer os laços, ampliar as trocas acadêmicas e a atualização dos debates, que mudam e avançam em ritmo acelerado.

Os participantes do estudo destacaram elevada importância para a temática, dado que todos os entrevistados marcaram

“muito importante” ou “importante” oferecer este tipo de conteúdo aos alunos. Apesar desta percepção positiva sobre as HD, os entrevistados apontam que elas não configuram um novo paradigma para as Humanidades. Sob a ótica predominante, o uso de ferramentas computacionais propicia novas perspectivas e olhares sobre os objetos de estudo, no sentido metodológico, mas sem impacto sobre modelos teóricos.

Há um entendimento hegemônico de que a aplicação de conteúdos das HD é útil tanto para a atuação acadêmica, quanto para a atuação fora da academia. Ou seja, oferecem um novo destaque para a participação de humanistas em frentes de atuação profissional que não se limitam à universidade. Não obstante, há também desconfiança de que o alargamento da atuação profissional de formados em Humanidades prejudique a qualidade e a liberdade da formação e atuação estritamente acadêmicas.

Ao entendermos as transformações que têm ocorrido no contexto brasileiro em resposta à crescente digitalização do mundo cotidiano e da nossa prática profissional, estaremos melhor preparados para atuar na construção de caminhos que fortaleçam boas práticas científicas, a cidadania e a valorização do conhecimento humanístico e crítico. Neste sentido, os coordenadores indicaram que, na maioria dos cursos, de 20 a 50% dos docentes estão preparados para lecionar sobre a temática. Esse percentual alto surpreende e aponta para a necessidade de pesquisa empírica sobre os enquadramentos das HD no ensino brasileiro. Além disso, há sempre a necessidade de atualização dos conteúdos e ferramentas por conta da velocidade em que eles são transformados.

Um desafio geral está relacionado à resistência, seja na sala de aula, seja dentro das próprias instituições de ensino, à

incorporação de metodologias orientadas “para o digital” e “para uma aplicação prática e pública”, especialmente quando consideramos o custo político e organizacional de mudanças institucionais e o custo financeiro do investimento necessário, sobretudo em um contexto nacional de cortes de orçamento e de ataque político à legitimidade da grande área de conhecimento das Humanidades.

Outro grande desafio, desta vez de ordem propriamente pedagógica e teórico-prática, reside em desenvolver habilidades técnicas, de um lado, e perspectivas críticas em relação à tecnologia digital, de outro, posto que são competências interligadas. Unir o pensamento algorítmico (*computational thinking*) e a “perspectiva humanista” em um mesmo domínio epistêmico nos parece fundamental. Profissionais de Humanas precisam saber como um computador funciona para poder aplicar princípios da Ciência da Computação na resolução de problemas ou tarefas típicas das Ciências Sociais. Sem esta base, a curva de aprendizado de ferramentas e métodos computacionais se torna muito acentuada e as oportunidades de aprendizagem continuam sendo distribuídas de forma desigual.

Compreender que algoritmos funcionam a partir do uso da abstração para resolver problemas específicos, da decomposição de tais problemas e do reconhecimento de padrões é um primeiro passo, mas não há consenso (e talvez nem mesmo tanta discussão) sobre o quão profundo precisa ser esse conhecimento. Um segundo passo consiste em saber traduzir questões da nossa área em termos que podem ser entendidos e processados pelo computador, ao mesmo tempo que conseguimos fazer o caminho inverso e analisar criticamente quais os impactos do funcionamento dessas tecnologias digitais sobre a organização social, a comunicação e as

artes. No campo das HD, a tônica das discussões sobre pedagogia digital é a de que a prática (com programas e metodologias) e o desenvolvimento de habilidades analítico-teóricas se retroalimentam, mesmo em experiências com uso pouco intensivo de recursos propriamente computacionais (Davis *et al.*, 2020; Fyfe, 2011).

No Brasil e no mundo, as discussões sobre que habilidades devem compor os currículos e sobre as experiências já desenvolvidas seguem dispersas, provavelmente concentradas em grupos de pesquisa, logo, são difíceis de acessar. Talvez uma iniciativa em rede voltada para refletir sobre esse ensino possa acelerar ou desdobrar o movimento já em curso, dado que é imperativo que o Humanista sendo formado hoje tenha uma literacia digital sólida que o habilite a criticar ou manejar um amplo leque de ferramentas computacionais, capacitando-o a propor reflexões prementes e soluções eficazes aos problemas da nossa realidade.

A *Red de Humanidades Digitales* (RedHD), estabelecida de forma muito precursora em 2011, além das outras redes falantes de espanhol e de português mencionadas ao longo do capítulo, talvez possam contribuir para a promoção do ensino das HD na América Latina, em um esforço coordenado para atingir os currículos básicos de graduação e de pós-graduação. Em histórico sobre a formação de Centros de HD pelo mundo, Russell (2018) destaca a importância desse arranjo institucional para o desenvolvimento de muitas comunidades de práticas e para o reconhecimento desse perfil profissional. Nossa avaliação do caso brasileiro é que esse modelo não se repetiu no país, estando as iniciativas mais circunscritas aos grupos de pesquisa ligados a departamentos tradicionais. De toda forma, não há receita universal, o importante, nos parece, é adensar o intercâmbio.

O envolvimento entre parceiros internacionais ou dentro de uma mesma universidade pode ocorrer de forma gradual, e com o apoio de especialistas de outras áreas que já dominem o conhecimento das ferramentas computacionais, mas que, todavia, têm baixo domínio teórico ou da realidade social analisada. No futuro, uma maior autonomia dos Humanistas no campo das HD permitirá uma combinação mais fina entre os dois mundos. A interdisciplinaridade, na prática, é uma grande propulsora das Humanidades Digitais. Com este panorama, esperamos que este capítulo possa contribuir para reflexões que incentivem a generalização e a consolidação do ensino de Humanidades Digitais no Brasil e na região, e para um sempre renovado debate acerca do imprescindível papel das tecnologias digitais no mundo acadêmico do século XXI.



Referências bibliográficas

- Abbott, A. (2000). Reflections on the Future of Sociology. *Contemporary Sociology*, 29(2), 296-300. <https://doi.org/10.2307/2654383>
- Alves, D. (2021). Ensinar Humanidades Digitais sem as Humanidades Digitais: um olhar a partir das licenciaturas em História. *Revista EducaOnline*, 15(2), 5-26.
- Alves, D. (2017, 17 de julho). História e Humanidades Digitais: conexões para um novo tempo. Entrevista concedida a Bruno Leal Pastor de Carvalho. *Café História – história feita com cliques*. <https://www.cafehistoria.com.br/historia-e-humanidades-digitais>

- Alves, D. (2016). Humanidades Digitais e Investigação Histórica em Portugal: perspectiva e discurso (1979-2015). *Práticas Da História. Journal on Theory, Historiography and Uses of the Past*, 1(2), 89-116.
- Babbie, E. (1999). *Métodos e técnicas da pesquisa de Survey*. UFMG.
- Baltar, R. e Baltar, C. (2013). As Ciências Sociais na Era do Zetabyte. *Mediações*, 18(1), 11-19.
- Berry, D. M. (2011). The computational turn: thinking about the digital humanities. *Culture Machine*, 12, 1-22. <https://culturemachine.net/wp-content/uploads/2019/01/10-Computational-Turn-440-893-1-PB.pdf>
- Berry, D. M. (Ed.). (2012). *Understanding Digital Humanities*. Palgrave Macmillan UK. <https://doi.org/10.1057/9780230371934>
- Berry, D. M. e Fagerjord, A. (2017). *Digital Humanities: Knowledge and Critique in a Digital Age*. Polity Press.
- Campos, L. A. (2014). A identificação de enquadramentos através da análise de correspondências: um modelo analítico aplicado à controvérsia das ações afirmativas raciais na imprensa. *Opinião Pública*, 20(3), 377-406. <https://doi.org/10.1590/1807-01912014203377>
- Carvalho de, B. L. e Lucchesi, A. (2016). História Digital: Reflexões, experiências e perspectivas. In A. M. Mauad; J. R. de Almeida; R. Santhiago. (Orgs.). *História Pública no Brasil: sentidos e itinerários*, v.1 (pp. 123-137). Letra e Voz.
- Castells, M. (2013). *A sociedade em rede: A era da informação: Economia, Sociedade e Cultura*, v. 1. Paz & Terra.
- Cesarino, L. (2021). Antropologia digital não é etnografia: explicação cibernética e transdisciplinaridade. *Civitas, Revista de Ciências Sociais*, 21(2), 304-315.

- Christian-Lamb, C. e Shrout, A. (2017). “Starting From Scratch”? Workshopping New Directions in Undergraduate Digital Humanities. *Digital Humanities Quarterly*, 11(3). <https://www.digital-humanities.org/dhq/vol/11/3/000311/000311.html>
- Christie, A. (2017). Building a toolkit for digital pedagogy. *Digital Humanities Quarterly*, 11(3). <http://www.digitalhumanities.org/dhq/vol/11/3/000310/000310.html>
- Cohen, D.; Frisch, M.; Gallagher, P.; Mintz, S.; Sword, K.; Taylor, A. M.; Thomas, W.; Turkel, W. (2008). Interchange: The promise of digital history. *The Journal of American History*, 95(2), 442-491. <http://dx.doi.org/10.2307/25095630>
- Conte, R.; Gilbert, N.; Bonelli, G.; Cioffi-Revilla, C.; Deffuant, G.; Dertész, J.; Loreto, V.; Moat, S.; Nadal, J.-P.; Sánchez, A.; Nowak, A.; Flache, A.; San Miguel, M.; Helbing, D. (2013). Manifesto da ciência social computacional. *Mediações-Revista De Ciências Sociais*, 18(1), 20-54. <https://doi.org/10.5433/2176-6665.2013v18n1p20>
- Cordell, R. (2016). How Not to Teach Digital Humanities. In M. K. Gold e L. Klein, (Eds.), *Debates in the Digital Humanities. Part 5. Digital Humanities and Its Critics*. University of Minnesota Press. <https://doi.org/10.5749/9781452963761>
- del Rio Riande, G. (Coord.). (2019). *La cultura de los datos: Actas del II Congreso Internacional de la Asociación Argentina de Humanidades Digitales*. La Plata.
- Davis Frost, R.; Gold, M. K.; Harris, K. D.; Sayers, J. (2020). *Digital Pedagogy in the Humanities: Concepts, Models, and Experiments*. Modern Language Association.
- Digital Humanities Manifesto 2.0, 2009. <http://manifesto.humanities.ucla.edu/>

- Ferla, L. A. C.; Lima, L. F. S.; Feitler, B. (2020a). Novidades no front: experiências com humanidades digitais em um curso de história na periferia da grande São Paulo. *Estudos Historicos*, 33(69), 111–132. <https://doi.org/10.1590/s2178-14942020000100007>
- Ferla, L.; Ferreira, K. R.; Atique, F.; Britt, A. G.; Fook, K. D.; Lesser, J.; Miyasaka, C.; Musa, D.; Rogers, T. D. e Vijaykumar, N. (2020b). Pauliceia 2.0: mapeamento colaborativo da história de São Paulo, 1870-1940. *Historia, Ciencias, Saude--Manguinhos*, 27(4), 1207–1223. <https://doi.org/10.1590/s0104-59702020000500010>
- Fraistat, N. (2012). *The Function of Digital Humanities Centers at the Present Time*. In M. K. GOLD (Ed.), *Debates in the Digital Humanities* (pp. 281-291). University of Minnesota Press.
- Fyfe, P. (2011). Digital Pedagogy Unplugged. *Digital Humanities Quarterly*, 5(3). <https://dhq-static.digitalhumanities.org/pdf/000106.pdf>
- Gao, Y.; Goetz, J.; Connelly, M.; Mazumder, R. (2020). Mining events with declassified diplomatic documents. *The Annals of Applied Statistics*, 14(4), 1699–1723. <https://doi.org/10.1214/20-aos1344>
- Gómez, M. (Ed.). (2019). *Digital Arts: Challenges and Opportunities at the Intersections between Arts, Society and Technology*. Brill. <https://doi.org/10.1163/9781848884274>
- Graham, S. (2021). Sonificação de dados (uma introdução à sonificação para historiadores) (G. Kucuruza, Trans.). *Programming Historian em português*, 1. <https://doi.org/10.46430/phptoo20>
- Higuchi, S.; Freitas, C.; Santos, D. (2022). Automatic Information Extraction: A Distant Reading of the Brazilian Historical-Biographical Dictionary. In V. Pinheiro et al (Orgs.). *Computational Processing of the Portuguese Language. PROPOR 2022. Lecture*

- Notes in Computer Science* (pp. 148-155). vol 13208. Springer, Cham. https://doi.org/10.1007/978-3-030-98305-5_14
- Hirsch, B. D. (2012). Digital humanities and the place of pedagogy. In B. D. Hirsch (Ed.), *Digital Humanities Pedagogy: Practices, Principles and Politics* (pp. 3-30). Open Book Publishers. <https://doi.org/10.11647/OBP.0024.01>
- Jockers, M. (2013). *Macroanalysis. Digital Methods and Literary History*. University of Illinois Press.
- Kosteles, C. e Haythornthwaite, C. (2002). Undergraduate Programs in Information Science: A Survey of Requirements and Goals. *Journal of Education for Library and Information Science*, 43(2). <https://doi.org/10.2307/40323975>
- Lage, M. C. e Godoy, A. S. (2008). O uso do computador na análise de dados qualitativos: questões emergentes. *RAM. Revista de Administração Mackenzie*, 9(4), 75-98. <https://doi.org/10.1590/S1678-69712008000400006>
- Machado, M. (2018). A teoria da antropologia digital para as humanidades digitais. *Revista Z Cultural*. <http://revistazcultural.pacc.ufrj.br/a-teoria-da-antropologia-digital-para-as-humanidades-digitais/>
- Marques, J. (2021). *Disputas intra-elites e agendas distributivas no Brasil rural de 1920 a 1945: mineração de texto no Arquivo Juarez Távora*. [Tese de doutorado inédita]. Universidade Estadual do Rio de Janeiro.
- Marres, N. (2017). *Digital Sociology: The Reinvention of Social Research*. Polity Press.
- Miller, D. (2011). *Tales from facebook*. Cambridge Press.

- Moretti, F. (2000). Conjectures on world literature. *New Left Review*, 1. <https://newleftreview.org/issues/iii/articles/franco-moretti-conjectures-on-world-literature>
- Multitudes, C. (2015). Digital Humanities Manifesto 2.0. *Multitudes*, 59, 181-195. <https://doi.org/10.3917/mult.059.0181>
- Musiol, H. (2021). 'Beyond the Word': Immersion, Art, and Theory in Environmental and Digital Humanities Prototyping". *Digital Humanities Quarterly*, 15(2). <https://www.digitalhumanities.org/dhq/vol/15/2/000557/000557.html>
- Noble, S. U. (2021). *Algoritmos da opressão: como o Google fomenta e lucra com o racismo*. Tradução de Felipe Damorim. Rua do Sabão.
- Nascimento, L. F. (2016). A Sociologia Digital: um desafio para o século XXI. *Sociologias*, 18(41), 216–241. <https://doi.org/10.1590/15174522-018004111>
- Noiret, S. (2015). História Pública Digital | Digital Public History. *Liinc Em Revista*, 11(1). <https://doi.org/10.18617/liinc.v11i1.797>
- Presner, T. (2010). Digital Humanities 2.0: A Report on Knowledge. In M. Bailar (Org.), *Emerging Disciplines: Shaping New Fields of Scholarly Inquiry in and beyond the Humanities*. Rice University. <https://archive.org/details/cnx-org-col11201>
- O'Neil, C. (2020). *Algoritmos de destruição em massa: como o big data aumenta a desigualdade e ameaça a democracia*. Tradução de Rafael Abraham. Rua do Sabão.
- Orton-Johnson, K. e Prior, N. (2013). *Digital Sociology: Critical Perspectives*. Palgrave Macmillan.
- Raine, L. e Wellman, B. (2012). *Networked: The New Social Operating System*. MIT Press. <https://doi.org/10.7551/mitpress/8358.001.0001>

- Ribeiro, C. J. S.; Higuchi, S.; Ferla, L. (2020). Aproximações ao cenário das humanidades digitais no Brasil. *Digital Humanities Quarterly*, 14, 1–10.
- Ribeiro, C. J. S. e Higuchi, S. (2018). *Anais do Congresso Internacional em Humanidades Digitais*. CPDOC/FGV.
- Ribeiro dos Santos, F. A. *et al.* (2020). Explorando os potenciais da História Digital: a experiência do Centro de Documentação e Imagem da Universidade Federal Rural do Rio De Janeiro - Campus de Nova Iguaçu. *Estudos Históricos*, 33(69), 152-172. <https://doi.org/10.1590/S2178-14942020000100009>
- Rosenzweig, R. e Cohen, D. (2005). *Digital History: a guide to gathering, preserving and presenting the past on the web*. University of Pennsylvania Press.
- Russell, I. G. (2018). La Institucionalización de las Humanidades Digitales. In I. G. Russell *et al.* *Humanidades digitales 1: Recepción, institucionalización y crítica*. Bonilla Artigas Editores, 2018.
- Schreibman, S.; Siemens, R.; Unsworth, J. (Eds.). (2004). *A companion to digital humanities*. Blackwell Publishing. <http://www.digitalhumanities.org/companion/>
- Silva, T. (2022). *Racismo algorítmico: inteligência artificial e discriminação nas redes digitais*. Edições Sesc SP.
- Spiro, L. (2012). *Opening Up Digital Humanities Education*. In B. D. Hirsch (ed.). *Digital Humanities Pedagogy: Practices, Principles and Politics*. Open Book Publishers. DOI: 10.11647/OBP.0024
- Sula, C.; Hackney, S. E.; Cunningham, P. (2017). A Survey of Digital Humanities Programs. *Journal of Interactive Technology & Pedagogy*, 11.

- Trindade, L. V. (2022). *Discurso de ódio nas redes sociais*. Editora Jandaíra.
- Villa, C. E. Valência (2019). Os Sistemas de Informação Geográfica como Curso de Extensão para Historiadores. *Educação Em Foco*, 24(2). <https://doi.org/10.34019/2447-5246.2019.v24.27865>
- Villa, C. E. Valência e Gil. T. (Orgs.). (2017). *O retorno dos Mapas. Sistemas de informação Geográfica em História*. Ladeira Livros.
- Villa, C. E. Valência (2015). Georreferência De Mapas históricos: Os Casos De Cartagena De Indias Colonial e do Rio De Janeiro Imperial. *Locus: Revista De História*, 21(2).
- Wissik, T.; Edmond, J.; Fischer, F.; De Jong, F.; Scagliola, S.; Scharnhorst, A.; Schmeer, H.; Scholger, W.; Wessels, L. (2020). Teaching Digital Humanities Around the World: An Infrastructural Approach to a Community-Driven DH Course Registry. *Library Tribune*, 40, 1-27. <https://hal.archives-ouvertes.fr/hal-02500871/document>
- Witte, J. C. (2012). A Ciência Social digitalizada: avanços, oportunidades e desafios. *Sociologias*, 14(31). <https://seer.ufrgs.br/index.php/sociologias/article/view/34927>
- Zorich, D. (2008). *A survey of digital humanities centers in the United States*. Council on Library and Information Resources.

**Un puente entre la
elección y la necesidad:
reflexiones sobre la
minimal computing
para las humanidades
digitales y la ciencia
abierta en América
Latina**

Para citar este capítulo:

<http://doi.org/10.51573/Andes.9789587987362.9789587987379.9789587987867.6>

GIMENA DEL RIO RIANDE

RAFFAELE VIGLIANTI

f ¿Podría la *minimal computing* extenderse más allá de la “computación realizada bajo algunas limitaciones tecnológicas”, para servir como un terreno común entre diferentes investigaciones y dinámicas pedagógicas en humanidades digitales (HD) entre los comúnmente denominados *norte y sur global*? ¿Podría la *minimal computing* proporcionar un conjunto de principios y tecnologías compartidos para capacitar a los estudiantes e investigadores de cualquier latitud para trabajar de manera autónoma y tener más control sobre el futuro de sus propios proyectos? ¿Cuáles serían sus alcances y limitaciones? ¿Cuánto aportaría el contexto académico latinoamericano a esta propuesta?

Como es sabido, a pesar de las ventajas que podríamos esperar de unas HD globales, es crucial recordar que el concepto *global* es complejo e incluso contradictorio, en especial cuando se relaciona con la tecnología. En este trabajo exploramos las preguntas iniciales y comentamos nuestra experiencia de uso de la *minimal computing* en proyectos de investigación, así como en

el desarrollo y la enseñanza de un curso universitario bilingüe para estudiantes matriculados en la Universidad de Maryland, College Park (Estados Unidos), y la Universidad del Salvador, en Buenos Aires (Argentina). Desde nuestro punto de vista, tanto la enseñanza como la investigación a través de una infraestructura informática basada en la instalación de paquetes y dependencias, el manejo de línea de comandos y el trabajo con repositorios y construcción de sitios web estáticos exhortan a los estudiantes e investigadores en cualquier lugar del mundo a pensar su quehacer de forma autónoma, sin depender de plataformas ajenas o servidores, global y localmente, y desde la elección o la necesidad, o, mejor dicho, como un puente entre ambas. Aún más, entendemos que este enfoque puede contribuir a fortalecer los vínculos entre la comunidad de humanistas digitales y el *software* y *hardware* de código abierto, u otros movimientos de innovación sobre el acceso abierto y la ciencia abierta, los cuales, en muchas regiones como América Latina o África, han estado liderando o cocreando proyectos *open source* dentro de la academia.

¿Con qué tecnologías podemos hacer humanidades digitales en América Latina?

Pasan los años y resulta cada vez más difícil, y tal vez inútil, pensar en una definición desagregada para las HD a escala global¹.

1 Por una cuestión de espacio, no entramos aquí en el debate de los manuales que buscaron definir a las *digital humanities*, como Kirschenbaum (2010, 2014)

Como se ha discutido en otros lugares, uno de los problemas para definir las HD radica en que estas no son una disciplina, sino una línea de investigación transversal (Klein, 2015). Las HD se han venido definiendo como un campo científico extendido, constituido por diferentes disciplinas o áreas de interés, como la filología digital o la edición digital de textos (Nyhan *et al.*, 2013; del Rio Riande, 2016b; entre otros), o por temas, como los de las Computational Humanities, las FeministDH o las QueerDH, entre otros (Gallon, 2016; Losh y Wernimont, 2018). Y si bien todas las disciplinas, líneas de investigación o áreas de interés forman parte de un sistema científico global, estas poseen asimismo un cierto nivel de autonomía, favorecido por las tradiciones, las teorías y las prácticas locales, al igual que por el uso de diferentes lenguas para la comunicación de la ciencia (Babini, 1955; Bourdieu, 1999; Kreimer *et al.*, 2014). Por ello, muchas veces resulta complejo llegar a zonas de acuerdo entre lo que es o no es HD en actividades académicas, como revisiones de trabajos presentados a congresos, o evaluaciones de proyectos o publicaciones científicas internacionales (Spiro, 2012). Se suma al debate el hecho de que la práctica de las HD se encuentra, siguiendo a Svensson (2016),

y Nyhan *et al.* (2013), o la primera parte de los *Debates in digital humanities* (Gold, 2012), enteramente dedicada a esta empresa. En nuestra región, el tono asertivo no parece ser el modo en el que las HD se están construyendo, sino que se observa la necesidad de contar desde los datos (del Rio Riande, 2016a) o desde la experiencia (Afanador Llach *et al.*, 2020). Ejemplo de esto último también es la serie de podcasts “Cinco preguntas a humanistas digitales”, impulsada por la Asociación Argentina de Humanidades Digitales (AAHD). Véase <https://aahd.net.ar/podcast/>.

integrada o embebida en la tecnología². En esta construcción del conocimiento mediada por la tecnología quedan aún más a la luz todas las diferencias mencionadas, ya que es imposible obliterar que el acceso a la tecnología y la alfabetización digital a nivel global es ampliamente desigual (Krull, 2016). En síntesis, aunque nos esforcemos por pensar en unas HD globales, el componente *localizado* y *situacional* no parece fácil de superar (del Rio Riande, 2016a; Fiorimonte y del Rio Riande, 2017).

En el caso de América Latina y el Caribe, si bien se observa un crecimiento global sostenido en el acceso a internet en los últimos diez años, sus niveles de penetración en las capas sociales más bajas siguen siendo bajos (37%), si los comparamos con los de Europa (82,5%) (Baladron, 2020). Asimismo, la denominada *brecha digital* se verifica no solo entre países, sino también en el interior de estos (Baladron, 2020; Gray *et al.*, 2016). Acerca de la desigualdad en el acceso a la tecnología y de su importancia en la educación informacional de niños y jóvenes, desde la Cumbre Iberoamericana de San Salvador del 2008 se viene sosteniendo, por ejemplo, la necesidad de:

impulsar políticas, que incluyan el marco de la colaboración público-privada, encaminadas a facilitar la integración plena de las y los jóvenes en la Sociedad de la Información y del Conocimiento a través del acceso universal a las Tecnologías de la Información y de la Comunicación (TIC) y el desarrollo de contenidos digitales, mediante programas de alfabetización digital que reduzcan la brecha existente y con la mira puesta

2 El autor las define como “a technologically embedded field” (Svensson, 2016).

en facilitar el acceso al empleo, el emprendimiento y la realización personal. (INTEF, 2013)

Entonces, ¿cómo puede consolidarse, a escala global, un campo científico que basa gran parte de sus abordajes en la tecnología, si esta es en muchos casos una barrera para el acceso, la construcción y la compartición del conocimiento? ¿Cómo puede hacerse investigación en HD en una región donde la desigualdad sigue siendo parte de su matriz? (Comisión Económica para América Latina y el Caribe, [Cepal], 2016).

Aunque contamos con algunas propuestas previas, principalmente pensadas desde el movimiento del acceso abierto en América Latina (del Rio Riande, 2016b; Schallier *et al.*, 2018; Ruiz Vaca, 2022), hasta el momento no se ha producido un debate global acerca del tipo de propuestas metodológicas y tecnológicas que podrían favorecer la equidad y la sostenibilidad en el campo de las HD, tanto desde la investigación como desde la educación. Desde nuestro punto de vista, América Latina cuenta con interesantes experiencias previas relacionadas con el acceso abierto y el uso de *software* y *hardware* libre y abierto, que podrían producir intervenciones significativas para pensar unas HD globales³. Por un lado, el acceso abierto a la publicación científica es —desde hace mucho tiempo en nuestra región— el estándar de difusión en el ámbito de las universidades e instituciones de ciencia

3 Además de los ejemplos mencionados, añadimos el mayoritario uso de Open Journals System (OJS), un *software* libre y colaborativo desarrollado por Public Knowledge Project, para la publicación de revistas en acceso abierto, que además ofrece datos abiertos de uso. Véase <https://github.com/saurabh-khanna/pkp/>.

públicas, ampliamente entendido además como motor clave de democratización del conocimiento (Packer *et al.*, 2006); de hecho, nuestras iniciativas son anteriores a la primera Declaración de Budapest y al término *acceso abierto* en sí mismo. Por poner un ejemplo, SciELO, una de las más importantes bibliotecas digitales y *metapublisher* en abierto en la región, fue fundada en 1997, cinco años antes de que empezaran a producirse las reuniones de las declaraciones de acceso abierto⁴.

Una razón importante para el desarrollo de las infraestructuras abiertas en América Latina fue la premisa de que la investigación financiada con fondos públicos debe ser gratuita para la sociedad, en términos de bien público y en pos de la mencionada democratización del conocimiento (Babini y Rovelli, 2020), y que la investigación debe publicarse en las lenguas de comunicación científica de la región, el español y el portugués. De igual manera, es destacable que este abordaje de lo abierto, desde universidades y otras instituciones de investigación, se construye para y por la comunidad académica, siendo esta la que cumple roles de edición, publicación y mantenimiento de las infraestructuras de publicación científica (Alperin y Fischman, 2015). Así, hoy en día América Latina es la región del mundo con el mayor porcentaje de revistas disponibles en acceso abierto y financiadas como parte del costo de la investigación desarrollada con fondos públicos, sin cobrar por la lectura de los artículos ni por publicarlos en acceso abierto —sin cobrar por *article processing charges* (APC)—. La región es dueña de un sistema que no solo favorece la publicación

4 Existe en Wikipedia una categoría para las declaraciones de acceso abierto; véase https://en.wikipedia.org/wiki/Category:Open_access_statements.

científica, sino que amplía significativamente su público lector. Un estudio acerca de las redes SciELO y Redalyc demostró que el 75 % de sus lectores son estudiantes, investigadores y profesores universitarios, quienes las utilizan para el aprendizaje, y el restante 25 % está constituido por profesionales, pero también por ciudadanos no relacionados con el ámbito académico que buscan información en la web (Alperin, 2016).

En los últimos años, el lugar central de las tecnologías web a la hora de investigar amplió las propuestas del acceso abierto en lo que viene denominándose *ciencia abierta*, una serie de principios que incluyen el acceso abierto, los datos abiertos de investigación y la revisión por pares abierta, que, a su vez, se complementan con otros componentes, como prácticas abiertas de investigación, uso de licencias abiertas, y *software* y *hardware* abierto (Fostering the Practical Implementation of Open Science in Horizon 2020 and beyond [FOSTER], 2018). Acerca de estos dos últimos, América Latina cuenta con importantes propuestas; por ejemplo, las del Plan Ceibal, en Uruguay, o Vía Libre y Conectar Igualdad, en Argentina, que apostaron no solo por el acceso a *netbooks* y formación⁵, sino que también procuraron generar redes e infraestructuras sobre uso de *software* y espacios de educación y capacitación de docentes desde “lo abierto”⁶. Asi-

5 Para leer una interesante nota sobre el programa Conectar Igualdad y el plan Ceibal, véase <https://ladiaria.com.uy/mundo/articulo/2022/1/argentina-retoma-el-programa-conectar-igualdad-similar-al-plan-ceibal-y-abre-debate-sobre-software/>. Para el sitio de Vía Libre, véase <https://www.vialibre.org.ar/>.

6 Algunas iniciativas en la región son la Red Argentina de Educación Abierta (AREA) o el Instituto de Educación Abierta (INTEA) de Uruguay. Véanse <http://a-rea.org/> y <https://www.intea.edu.uy/>.

mismo, redes como regosh⁷ (Arancio, 2021) o grupos derivados del DIY Book Scanner⁸ trazaron puentes entre el activismo sobre el *hardware* abierto y la academia. Creemos que las lecciones aprendidas sobre lo abierto en América Latina aportan a la construcción conjunta de unas HD, si no globales, al menos más abiertas y basadas en un uso y distribución del conocimiento y la información más equitativos e inclusivos. A partir de este lugar de enunciación, este capítulo busca reflexionar sobre el modo en el que un abordaje desde la ciencia abierta y para las HD puede colaborar con la disminución de la brecha digital o siquiera con una distribución más equitativa de la alfabetización digital, así como puede ayudar a la sostenibilidad de proyectos de HD. Si bien somos conscientes de las limitaciones de nuestra contribución, que incluye el aprendizaje y la puesta en práctica de una serie de lenguajes de codificación web y de escritura de código, en lo que denominamos aquí *tecnologías de minimal computing*, creemos que, al menos, la propuesta apoya la construcción de un paradigma que, desde las humanidades, usa crítica y responsablemente la tecnología.

7 Véase <https://regosh.libres.cc/en/home-en/>.

8 Sobre este proyecto, véase: <https://diybookscanner.org/es/index.html>. Para algunas reflexiones sobre la creación de un catálogo colaborativo y uso de *software* abierto, véase Afanador Llach y Lombana Bermúdez (2022).

¿Qué entendemos por *minimal computing*?

La *minimal computing* puede entenderse como un conjunto de principios y tecnologías web de código abierto, que permiten capacitar a los estudiantes e investigadores para trabajar de manera autónoma y tener más control sobre el futuro de sus proyectos digitales. En este término compuesto, el adjetivo *minimal* no hace referencia a lo poco o lo simple en el aprendizaje informático o la destreza técnica, sino a las características de arquitectura de *software*, la infraestructura de *hardware* y el mantenimiento a largo plazo de los proyectos de investigación que la utilizan. Desde este abordaje, la gestión y publicación de datos se realiza desde repositorios de tamaño pequeño o mediano y de carácter personal o compartido, como los que ofrece GitHub¹⁰. Así, para crear un proyecto con *minimal computing*, no se requiere de una gran cantidad de recursos; de ahí su definición. Esta escasa

- 9 Las reflexiones sobre la *minimal computing* vertidas en esta sección pertenecen a un trabajo en prensa de Allés Torrent y del Río Riande (2020).
- 10 GitHub es uno de los recursos más usados en los proyectos de *minimal computing* para oficial de repositorio de datos y documentación de un proyecto. Se trata de una plataforma de código abierto que aloja repositorios de código, donde se pueden crear proyectos abiertos. GitHub nació como un espacio para que los desarrolladores compartieran el código de sus aplicaciones y herramientas, pero rápidamente se extendió a espacios académicos o a su uso como repositorio personal. El código de los proyectos que se pone en abierto puede ser descargado y revisado por cualquier usuario de la plataforma, lo que ayuda a mejorarlos y reproducirlos, al poder ejecutar operaciones de ramificación o clonación de los repositorios. GitHub se caracteriza también por permitir el trabajo colaborativo.

necesidad de recursos materiales y de procesamiento y almacenamiento redundante en una mejor accesibilidad y una mayor estabilidad para conexiones con acceso limitado, además de favorecer la autogestión. En segundo lugar, el concepto *minimal* atiende a la posibilidad de generar una publicación en formato web de tipo estático¹¹, con un diseño simple, donde se dispone de forma concisa de la información que allí se aloja. En tercer lugar, el concepto puede relacionarse con las tecnologías de código abierto u *open source*, por ejemplo, el uso de editores de código como Visual Studio Code. Finalmente, la noción de *minimal* puede aplicarse al aprendizaje de estándares de código abierto que sirven para interactuar con la mayor parte de los objetos web, como los lenguajes de marcado TEI-XML¹², markdown¹³, HTML¹⁴,

- 11 En un sitio web estático, el código fuente y las páginas se encuentran fijos, es decir, tal como fue una página diseñada y almacenada en un repositorio, así las recibe el navegador y las ve el usuario. Una página web estática está compuesta por archivos HTML individuales que son pregenerados y presentados al usuario a través del navegador web. En un sitio estático, las páginas existen como archivos individuales, mientras que en un sitio dinámico se generan en función de la demanda del usuario.
- 12 TEI, o Text Encoding Initiative, es un sistema de marcado para textos de humanidades y ciencias sociales. Este marcado es de naturaleza modular y se expresa a través del lenguaje estándar web *extensible markup language* (XML). Para conocer el sitio web de la TEI, con información sobre el consorcio, herramientas, proyectos y novedades, véase <https://tei-c.org/>.
- 13 Markdown es un lenguaje de marcado ligero, creado por el desaparecido activista de internet y acceso abierto Aaron Swartz y John Gruber, que busca la máxima legibilidad y facilidad de publicación.
- 14 HTML, o *hypertext markup language* (lenguaje de marcado hipertextual), se utiliza para estructurar y desplegar una página web y sus contenidos en el navegador.

css¹⁵, y a la comunicación directa con el ordenador a partir de la línea de comandos¹⁶. Evidentemente, donde la noción de *minimal* no entra es en los lenguajes y tecnologías web por aprender. Aun en un proyecto simple de publicación digital, nos vemos obligados a realizar diferentes instalaciones y, al menos, a comprender el comportamiento de estos distintos lenguajes de marcado. En consecuencia, la curva de aprendizaje inicial en proyectos de *minimal computing* es elevada, y su aprendizaje lleva mucho más tiempo que los pasos que deben aprenderse a la hora de trabajar con aplicaciones o plataformas web ya desarrolladas para la publicación en línea; sin embargo, resulta un insumo básico que puede replicarse más tarde en diferentes contextos de investigación. Por ejemplo, el conocimiento de lenguajes web —como los mencionados markdown, XML, HTML, CSS y sus derivados— es aplicable no a una plataforma o *software* específico, sino a la mayoría de los objetos web.

¿Qué necesitamos?

A pesar de la variedad de objetos de estudio y metodologías que implican, los proyectos de investigación en HD incluyen muy a menudo la creación de un recurso digital, el cual puede ser

- 15 CSS, o *cascading style sheets* (hojas de estilo en cascada), es el lenguaje que estructura el diseño y la presentación de una página web.
- 16 La línea de comandos (en inglés, *command-line interface* [CLI]) es un tipo de interfaz de usuario que, en cualquier sistema operativo, permite a los usuarios dar instrucciones a algún programa informático o al mismo sistema operativo por medio de una línea de texto simple.

consultado en línea por otros académicos e investigadores. Para proyectos relacionados con ediciones, bibliotecas o archivos digitales, por ejemplo, la creación de un recurso web es indispensable y forma parte integral de la investigación. Incluso aquellos proyectos que no se centran en la creación de un sitio web y que implementan técnicas de análisis textual producen una cantidad de datos y códigos que deben ser preservados y puestos a disposición en línea. La creación, pero sobre todo el mantenimiento de los recursos digitales, supone un obstáculo nada desdeñable que suele ser gestionado a nivel estructural por entidades como universidades, bibliotecas o consorcios de investigación nacionales y supranacionales, que no siempre van de la mano con lo que la comunidad académica necesita (Van Zundert, 2012). Es decir, muchas veces, al terminarse el dinero que financió un proyecto, las instituciones no continúan su compromiso de mantenimiento de este. Se suma a ello que el acceso a estas infraestructuras de sostenibilidad o preservación no está garantizado para todos los académicos, en especial para aquellos que aún son estudiantes o comienzan a acercarse a las HD después de haber adquirido, a través de capacitaciones formales e informales, suficiente competencia en las herramientas necesarias para sus estudios (Allés Torrent y del Río Riande, 2018, 2020). Incluso grupos de investigación más organizados pueden encontrarse con un acceso limitado a la infraestructura de su institución o con problemas de seguridad al utilizar servicios externos para servidores (del Río Riande, 2020, 2022a, 2022b).

La metodología de la *minimal computing* surgió en Estados Unidos como una reacción a la falta de acceso a las infraestructuras institucionales, así como a su inadecuación para responder

a las necesidades de los proyectos de HD, en particular a los que tienen cierta urgencia en dar respuesta a acontecimientos socioculturales de actualidad (Gil y Ortega, 2016)¹⁷. En el 2015 se formó un grupo de trabajo en *minimal computing*, auspiciado por varios miembros de Implementing New Knowledge Environments (INKE) y de Global Outlook Digital Humanities (GO::DH), grupo de interés de la Alliance of the Digital Humanities Organizations (ADHO). La conformación del grupo sobre *minimal computing* fue un salto cualitativo para los objetivos de GO::DH, centrados en el debate sobre la posibilidad de construir unas HD globales, desde su fundación en el 2013¹⁸. De algún modo, el grupo era también una respuesta a las HD de ese entonces, caracterizadas por el surgimiento de grandes proyectos que desarrollaban contenido e infraestructuras digitales y luego las ponían en código abierto para uso de la comunidad. Tapor¹⁹, Omeka²⁰ o los extintos Juxta

17 Uno de los primeros proyectos notables del movimiento *minimal computing* fue Torn Apart/Separados, una iniciativa *rapidly deployed* (“desarrollada rápidamente”) para documentar, a través de mapas y datos públicos del Gobierno, el impacto de los centros de detención durante la cruda separación de familias solicitantes de asilo en la frontera sur de Estados Unidos durante la administración del entonces presidente Donald Trump (2016-2020). El proyecto fue creado de forma independiente y rápida por un grupo de académicos que utilizaron tecnologías web abiertas e infraestructura gratuita (como Jekyll y GitHub), para producir un sitio inicial al que se le fueron añadiendo contenidos. Véase <https://xpmethod.columbia.edu/torn-apart/>

18 Véase <https://go-dh.github.io/mincomp/about/>.

19 Véase <https://tapor.ca/>.

20 Véase <https://omeka.org/>.

Commons²¹, Bamboo/DiRT²² o DH Press²³ son buenos ejemplos de este tipo de infraestructuras, que necesitan de grandes equipos de investigación y desarrollo, abultada financiación, espacio físico en servidores, constante mantenimiento, y adopción y uso por parte de la comunidad de humanistas digitales. Como sabemos, la mayoría de estas iniciativas fracasó u hoy en día encuentran problemas de mantenimiento y financiación.

El grupo sobre *minimal computing* exploró, en primera instancia, los elementos que se utilizan por necesidad o por elección en los proyectos de HD. En una entrevista con el arquitecto cubano Ernesto Oroza, Alex Gil (2016) introdujo el concepto de *arquitectura de la necesidad* y lo aplicó a los proyectos de HD y a la infraestructura que los sustenta. Oroza había acuñado este concepto para describir la expansión de La Habana, que se produjo de forma espontánea y en respuesta a las necesidades inmediatas de sus habitantes, a veces en contraste con las normativas gubernamentales y los intentos de reorganizar y regular su desarrollo. Según Gil, esto es en gran medida comparable con el desarrollo de proyectos de investigación en HD que han surgido y siguen surgiendo a pesar de las dificultades para obtener financiación y acceso a la infraestructura. Una consecuencia importante de la falta de acceso a la financiación es el acercamiento de los investigadores de

21 Juxta Commons es un *software* para la colación de variantes textuales, desarrollado en el marco del grupo Nineteenth-century Scholarship Online (NINES). Para más información, véase <http://pedagogy-toolkit.org/tools/Juxta.html/>.

22 Véase <https://digitalhumanities.berkeley.edu/projects/dirt-digital-research-tools/>.

23 Véase <https://digitalinnovation.web.unc.edu/projects/dhpress/>.

humanidades a las herramientas técnicas, como la programación web básica (Gil, 2015; Gil y Ortega, 2016). Para esto, Gil (2016) también se basa en un concepto del arquitecto Oroza que describe al *modulor moral* (*moral modulor*) como un individuo que construye y aprende a construir por necesidad, centrándose en lo que es útil y necesario; una perspectiva de escala moral que reelabora las propuestas de Le Corbusier²⁴.

Esta centralidad en el concepto de necesidad —“an injunction to constantly repeat the question ‘what do we need?’” [“una petición a repetirnos constantemente la pregunta ‘¿qué necesitamos?’”] (Gil y Ortega, 2016)— hace que el enfoque de la *minimal computing* sea aplicable en múltiples contextos y sirva de denominador común para unas HD más abiertas y equitativas: un enfoque que sea accesible globalmente y adaptable localmente, que pueda ser un puente entre la necesidad y la elección. Como es sabido, a pesar de las ventajas que podríamos esperar de unas HD globales, Boaventura de Sousa Santos (2006) nos enseñó en su día que “el término *global* se refiere tanto a los procesos como a los resultados de la globalización, y lo que llamamos globalización es siempre la globalización exitosa de un localismo particular” (p. 396)²⁵.

Es crucial recordar, como adelantábamos en el apartado anterior, que el concepto *global* es complejo e incluso contradictorio, en particular cuando se relaciona con la tecnología. En las HD, la

24 El modulor es un sistema de medidas matemáticas entre el hombre y la naturaleza desarrollado por el arquitecto suizo Le Corbusier, en colaboración con André Wogenscky.

25 “The term ‘global’ refers both to the processes and to the results of globalization, and what we call globalization is always the successful globalization of one particular localism”. Traducción propia.

tecnología en uso incluye la arquitectura de *software*, es decir, la disponibilidad y el coste de las soluciones de *software* y las habilidades necesarias para hacer uso de ellas; la infraestructura de *hardware*, o la disponibilidad y el coste de los ordenadores, los periféricos y otras herramientas necesarias para utilizar el *software* y las metodologías requeridas para un proyecto de HD; y la infraestructura de mantenimiento a largo plazo, es decir, la gestión de datos y el apoyo institucional para garantizar que los resultados de los proyectos de investigación se salvaguarden en el futuro. Según la disponibilidad y el estado de la tecnología así definida, que varía considerablemente en función de la posición de los investigadores y del contexto local o hiperlocal en el que se encuentran, varían también las motivaciones para adoptar los principios de la *minimal computing*, los cuales oscilan entre la necesidad y la elección.

Minimal computing: entre la elección y la necesidad

Si bien el abordaje norteamericano de la *minimal computing* se presenta, como hemos visto, desde la pregunta por lo que se necesita para comenzar un proyecto de HD, es entendible que, en muchos casos, esta cuestión también se enmarque simplemente en la elección entre varias opciones. En el caso del Maryland Institute for Technology in the Humanities (MITH) de la Universidad de Maryland, institución a la que pertenece uno de los autores de este trabajo, el uso de la *minimal computing* se destina principalmente a cuestiones de preservación digital. El MITH migró varios sitios y aplicaciones web a sitios estáticos para reducir las demandas de

infraestructura y aumentar su longevidad (Summers, 2016)²⁶. No obstante, el MITH lleva adelante otros proyectos que no solo implican el trabajo con tecnologías de *minimal computing*.

En el caso de los proyectos del Laboratorio de Humanidades Digitales (HD Lab) del Consejo Nacional de Investigaciones Científicas y Técnicas (Conicet) de Argentina, espacio que dirige una de las autoras de este capítulo, la *minimal computing* se convirtió en la única solución para “producir nosotros mismos nuestra propia investigación” (Gil, 2015). Esta elección autoimpuesta es parte de los estándares del laboratorio y se integran a las HD desde los principios de apertura de la ciencia abierta en el país. Desde este abordaje, lo abierto recorre un flujo de trabajo que hace a corpus, documentación, colaboración, *software*, infraestructura, publicación y reúso de datos (del Rio Riande *et al.*, 2018). Dado que el laboratorio no puede comprar *software* o pagar servidores y alojamiento, la adopción de estrategias mínimas creó las condiciones necesarias para que nuestro trabajo forme parte del panorama global de las HD. Como tal, practicamos la *minimal computing* más allá de su definición de simplemente “computación realizada bajo algunas limitaciones tecnológicas”, ya que se convierte en nuestro principal y casi único instrumento para cualquier forma de HD.

En diciembre del 2019 propusimos conjuntamente un curso titulado Humanidades Digitales: Ediciones Digitales con Minimal Computing²⁷ al programa Global Classrooms Initiative (GCI), de

26 Un ejemplo es el Lakeland Community Archive Project; véase <https://mith.umd.edu/research/lakeland/>.

27 Para el sitio web público bilingüe del curso, véase <https://mith.umd.edu/minimaldigipub/>.

la Universidad de Maryland. Este programa ofrece apoyo para el desarrollo de cursos en colaboración con una institución de educación superior fuera de los Estados Unidos, con el objetivo de exponer a los estudiantes a un trabajo intercultural, basado en proyectos y virtual. Los cursos financiados por el programa preparan a los estudiantes para participar en un mundo globalizado, fomentando las competencias interculturales. En nuestro caso, la pandemia de covid-19 hizo que nos tuviésemos que adaptar aún más a depender de los vínculos tecnológicos para colaborar y elaborar nuestro proyecto.

En esta formación, que desarrollamos desde el 2020 y finalizó en el 2022, han participado estudiantes de la Universidad del Salvador, en Argentina, y de la Universidad de Maryland, en Estados Unidos²⁸, en un abordaje de aprendizaje en línea y presencial²⁹. El curso basa su metodología en la elaboración de un proyecto grupal, en el que los estudiantes colaboran virtualmente para crear una edición digital bilingüe (español e inglés), fundamentada en tecnologías de *minimal computing*, de un texto multilingüe de la época colonial³⁰, al mismo tiempo que aprenden sobre los enfoques de las HD en los estudios literarios, la edición digital y la historia.

28 Los instructores de Argentina enseñaron en la Universidad del Salvador, pero forman parte del Conicet en Argentina.

29 Nuestro plan era tener clases presenciales con un enlace de audio/video compartido que uniera las dos aulas. Sin embargo, dada la pandemia mundial, la primera sesión del curso fue totalmente en línea para todos los estudiantes en el 2020, y en línea para la Universidad del Salvador e híbrido para los estudiantes de la Universidad de Maryland en el 2021 y el 2022.

30 Este texto es una relación de viaje escrita por un comerciante vasco llamado Acarete Du Biscay, y su historia editorial es realmente multilingüe. Los viajes

El proyecto grupal, creemos, facilita la colaboración intercultural y es también un paradigma de aprendizaje eficaz, en el que los “students actively construct their knowledge by participating in real-world activities similar to those that experts engage in, to solve problems and develop artifacts”³¹ (Krajcik y Blumenfeld, 2005, p. 319). Este enfoque no es ajeno a la pedagogía que, principalmente desde las academias norteamericanas, se propuso para las HD (Clement, 2012), dado que estas tienden a ser andamiadas a través de un trabajo colaborativo centrado en la idea de proyecto, el cual resulta en el desarrollo de artefactos como herramientas o publicaciones digitales (Burdick *et al.* 2012). También está relacionado con un enfoque de multiliteracidad, que reúne la diversidad lingüística y las formas multimodales de expresión y representación en respuesta a los cambios en los entornos tecnológicos

.....

de Acarete se publicaron en su lengua materna, el francés, en la *Relation des voyages du Sr... dans la rivière de la Plata, et de-là par terre au Pérou*, en 1672, como parte del cuarto volumen de *Collection of relations de divers voyages curieux*, de Melchisedec Thevenot; y en 1696, de forma independiente, en la *Relation des voyages dans la rivière de la Plate*. Dos años más tarde, la edición londinense de 1698 salió a la luz en una colección titulada *Voyages and discoveries in South America*, y posteriormente como libro individual por la imprenta de Samuel Buckley como *An Account of a voyage up the River de la Plata, and thence over land to Peru: with observations on the inhabitants, as well as Indians and Spaniards, the cities, commerce, fertility, and riches of that part of America*. Luego fue traducido del inglés al español por Daniel Maxwell y publicado en la *Revista de Buenos Aires*, en mayo y junio de 1867, con el título *Relación de los viajes de Monsieur Ascarate du Biscay al Río de la Plata, y desde aquí por tierra hasta el Perú, con observaciones sobre estos países*.

31 “Los estudiantes construyen activamente su conocimiento al participar en actividades del mundo real similares a las que realizan los expertos, para resolver problemas y desarrollar artefactos” (la traducción es nuestra).

globalizados, como internet, y a la creciente diversidad lingüística y cultural, generada por el aumento de la migración transnacional (New London Group, 1996; Cope y Kalantzis, 2009).

El trabajo en este curso nos ha demostrado que la enseñanza desde la óptica de la *minimal computing* se beneficia enormemente de los proyectos que exhortan a los estudiantes a pensar tanto global como localmente, al reconocer las posibilidades tecnológicas a las que tienen acceso (así como el porqué y el cómo) y al enfrentarse a las limitaciones y restricciones que juegan en su contra, ya sea en el *hardware*, el *software*, la educación, la capacidad de la red, la energía o, de hecho, las limitaciones autoimpuestas con fines pedagógicos (Calarco *et al.*, 2021)³². En otras palabras, entrenamos a los estudiantes para que aprendan a reconocer los privilegios de tener acceso a recursos informáticos de última generación, al igual que para que ideen estrategias para sortear las limitaciones que puedan encontrar, adoptando técnicas de la *minimal computing*. También hacemos hincapié en los beneficios a la hora de construir una comunidad desde el uso de herramientas libres o de código abierto (Sala *et al.*, 2014). Es probable que los estudiantes (y los expertos) de cualquier contexto sociocultural se encuentren con esas limitaciones en distintos grados en un futuro. Incluso aquellos que tengan acceso a la infraestructura a través de sus instituciones o futuros empleadores pueden enfrentarse

32 En Calarco *et al.* (2021) pueden encontrarse interesantes reflexiones de los alumnos que se inscribieron desde la Universidad del Salvador (Argentina). El trabajo surgió de una encuesta llevada a cabo por los estudiantes al finalizar la primera sesión del curso en el 2020; además, añade información sobre los procesos y métodos utilizados desde su experiencia en el curso.

a problemas de conservación y portabilidad de los proyectos de publicación digital que hayan creado.

Conclusiones

En efecto, tal como sostienen Boido y Baldatti (2012), las soluciones técnicas o tecnológicas no necesariamente resuelven los problemas de desigualdad en el seno de las sociedades. Aprender a desplegar enfoques mínimos puede contribuir a la viabilidad y la sostenibilidad a largo plazo de un proyecto de HD y dar a investigadores y estudiantes una mayor participación y un mayor control sobre el futuro de su propio trabajo. Asimismo, la *minimal computing* puede desempeñar un rol central en el establecimiento de enfoques abiertos y globales para la investigación y la pedagogía de las HD, contribuyendo a determinar unos principios soberanos para la ciencia abierta en América Latina y reducir la brecha entre las contribuciones del sur y el norte global.

Entendemos que la *minimal computing* tiene limitaciones a la hora de enseñarse y ponerse en práctica en proyectos de investigación y que no es una solución única, pero creemos que puede convivir con otros abordajes de las HD más escalables. Desde la región más desigual, pero tal vez una de las más ricas en propuestas de investigación abierta, buscamos que la *minimal computing* promueva efectivamente unas HD que puedan participar de forma equitativa en una conversación de voluntad global, que no se defina por lengua, financiamiento o región, sino que lo haga a partir de un conjunto de valores comunes, que incluya el uso de

tecnologías abiertas, el control sobre los datos y el código, y la reducción de la infraestructura informática.



Referencias

- Afanador Llach, M. J., Murcia Galindo, J. C., Rincón Rodríguez, O. G., Gallini, S., Méndez Mahecha, T. y Jaramillo Liévano, J. N. (2020). Humanidades digitales “a lo colombiche”: Cadáver exquisito de la Red Colombiana de Humanidades Digitales. *Revista de Humanidades Digitales*, 5, 217-235. <https://doi.org/10.5944/rhd.vol.5.2020.27837>
- Afanador Llach, M. J. y Lombana Bermúdez, A. (2022). Developing new literacy skills and digital scholarship infrastructures in the Global South. A case study. En D. Fiormonte, S. Chaudhuri y P. Ricaurte (Eds.), *Global debates in the digital humanities*. Minnesota University Press. <https://dhdebates.gc.cuny.edu/read/6b455755-8b93-4033-9b78-781cacc32e8d/section/14f2991b-6c29-4d33-ac23-b286b19boc18#ch17>
- Allés Torrent, S. y del Rio Riande, G. (2018). Enseñar edición digital con TEI en español. Aprendizaje situado y transculturación. En G. del Rio Riande, G. Calarco, G. Striker y R. de León (Eds.), *Humanidades digitales: Construcciones locales en contextos globales*. Editorial de la Facultad de Filosofía y Letras Universidad de Buenos Aires. <https://www.academica.org/aahd2016/36.pdf>

- Allés Torrent, S. y del Rio Riande, G. (2020). The switchover: Teaching and learning the text encoding initiative in spanish. *Journal of the Text Encoding Initiative*, 12. <https://doi.org/10.4000/jtei.2994>
- Alperin, J. P. (2016, 20 de abril). *Research is also for non-scholars: Lessons from Latin America* [presentación]. <https://doi.org/10.6084/m9.figshare.3187551.v1>
- Alperin, J. P. y Fischman, G. (2015). *Hecho en Latinoamérica: Acceso abierto, revistas académicas e innovaciones regionales*. Consejo Latinoamericano de Ciencias Sociales (CLACSO). <http://biblioteca.clacso.edu.ar/clacso/se/20150722110704/HechoEnLatinoamerica.pdf>
- Arancio, J. C. (2021). Opening up the tools for doing science: The case of the global open science hardware movement. *International Journal of Engineering, Social Justice and Peace*, 8(2), 1-27. <https://ojs.library.queensu.ca/index.php/IJESJP/article/view/13997/9834>.
- Babini, J. (1955). *¿Qué es la ciencia?* Editorial Columba.
- Babini, D. y Rovelli, L. (2020). *Tendencias recientes en las políticas científicas de ciencia abierta y acceso abierto en Iberoamérica*. Consejo Latinoamericano de Ciencias Sociales (CLACSO) y Fundación Carolina. <https://www.memoria.fahce.unlp.edu.ar/libros/pm.5293/pm.5293.pdf>
- Baladron, M. I. (2020). Apropiación de tecnologías en las redes comunitarias de internet latinoamericanas. *Trípodos*, 46, 59-76.
- Boido, G. y Baldatti, C. T. (2012). Nuevas tecnologías: ¿para quiénes? El caso de la nanotecnología. *Revista Iberoamericana de Ciencia, Tecnología y Sociedad*, 7(21), 11-21. <https://www.redalyc.org/articulo.oa?id=92424175002>

- Bourdieu, P. (1999). El campo científico. En *Intelectuales, política y poder* (pp. 75-110). Eudeba.
- Burdick, A., Drucker, J., Lunenfeld, P. Presner, T., y Schnapp, J. (2012). *Digital_humanities*. MIT Press.
- Calarco, G. A., Gionco, P., Méndez, R., Merino Recalde, D., Striker, G. y Suárez-Giraldo, C. (2021). Digital publishing with minimal computing (UMD-USAL, 2020): Nuestra experiencia como estudiantes. *Publicaciones de la Asociación Argentina de Humanidades Digitales*, 2. <https://revistas.unlp.edu.ar/publicaahd/article/view/13742>
- Clement, T. (2012). Multiliteracies in the undergraduate digital humanities curriculum: Skills, Principles, and Habits of Mind. En Hirsch, B. (Ed.), *Digital humanities pedagogy: Practices, principles and politics* (pp. 365-88). Open Book Publishers. <http://www.openbookpublishers.com/htmlreader/DHP/chap15.html>
- Comisión Económica para América Latina y el Caribe (Cepal). (2016). *La matriz de la desigualdad social en América Latina*. Cepal. <https://www.cepal.org/es/publicaciones/40668-la-matriz-la-desigualdad-social-america-latina>
- Cope, B. y Kalantzis, M. (2009). “Multiliteracies”: New literacies, new learning. *Pedagogies: An International Journal*, 4(3), 164-195.
- del Rio Riande, G. (2016a). ¿De qué hablamos cuando hablamos de humanidades digitales? En *Actas I Jornadas de humanidades digitales de la AAHD. Humanidades digitales: Tecnologías, culturas, saberes* (pp. 50-62). Editorial de la Facultad de Filosofía y Letras. <https://www.aacademica.org/jornadasaaahd/3.pdf>
- del Rio Riande, G. (2016b). *Humanidades digitales: Estándares para su consolidación en el campo científico argentino*

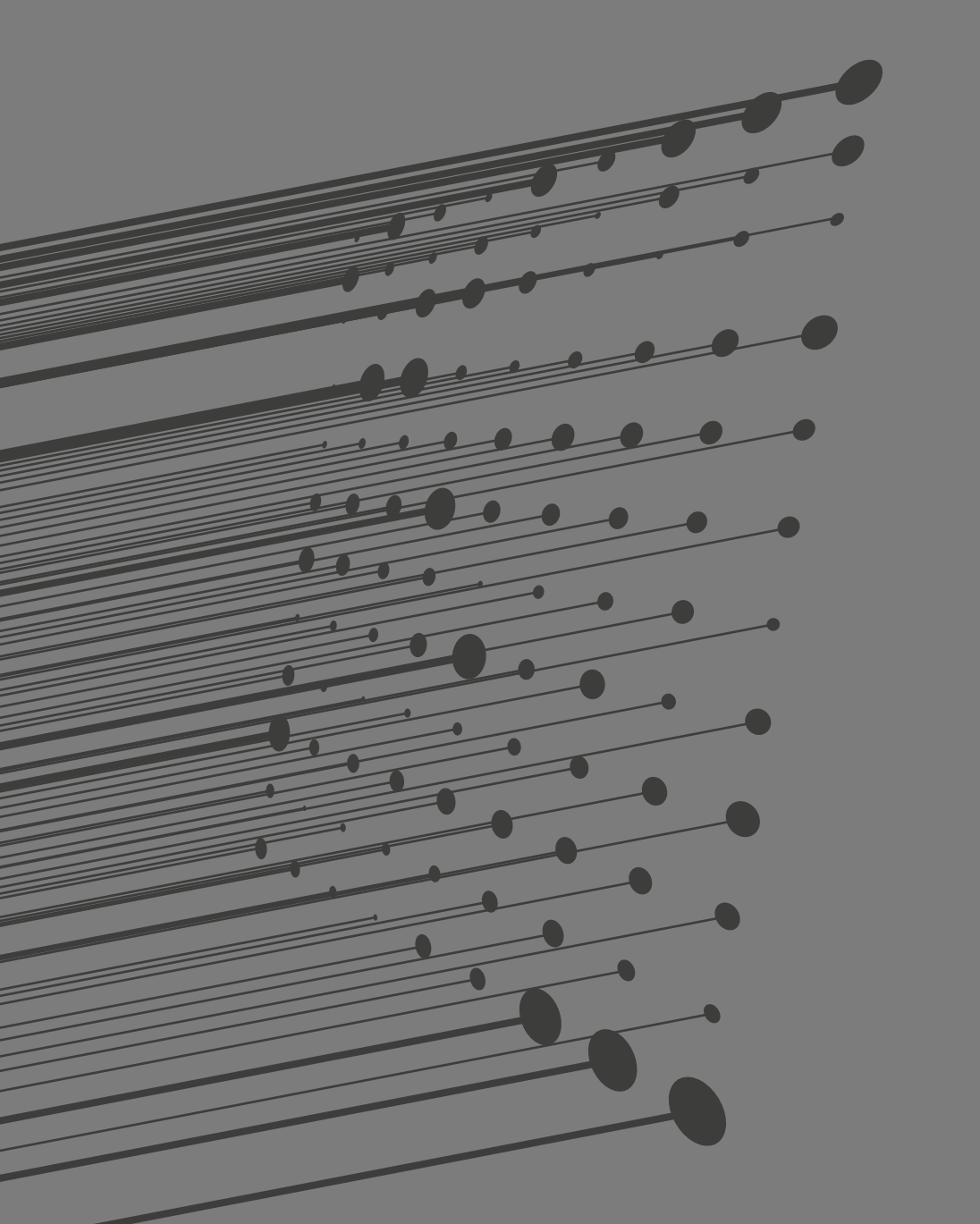
- [ponencia]. XLVIII Reunión de Bibliotecarios de la Asociación de Bibliotecarios Graduados de la República Argentina, Buenos Aires. <http://sedici.unlp.edu.ar/handle/10915/62008>
- del Rio Riande, G. (2018). Humanidades Digitales bajo la lupa: Investigación abierta y evaluación científica. *Ex Libris*, 7, 136-149. <http://revistas.filo.uba.ar/index.php/exlibris/article/view/3177>
- del Rio Riande, G. (2020). Editing a medieval hispanic poetry corpus with TextGrid. *Text+*. <https://www.text-plus.org/en/research-data/user-story-407/>
- del Rio Riande, G. (2022a). Digital humanities and visible and invisible infrastructures. En D. Fiormonte, S. Chaudhuri y P. Ricaurte (Eds.), *Global debates in the digital humanities* (pp. 247-258). Minnesota University Press. <https://dhdebates.gc.cuny.edu/read/6b455755-8b93-4033-9b78-781cacc32e8d/section/7383e66b-52ce-48ff-b9e8-f63abdafc9a3#ch19>
- del Rio Riande, G. (2022b). Humanidades digitales o las humanidades en la intersección de lo digital, lo público, lo mínimo y lo abierto. *Publicaciones de la Asociación Argentina de Humanidades Digitales*, 3. <https://doi.org/10.24215/27187470e038>
- del Rio Riande, G. y Tóth-Czifra, E. (2019). El metablog OpenMethods o cómo abrir la investigación en humanidades digitales. *Hipertext.net*, 19, 63-73. <https://raco.cat/index.php/Hipertext/article/view/360128>
- de Sousa Santos, B. (2006). Globalizations. *Theory, Culture & Society*, 23(2-3), 393-399. <https://doi.org/10.1177/026327640602300268>
- Dombrowski, Q. (2014). What ever happened to Project Bamboo? *Literary and Linguistic Computing*, 29(3), 326-339. <https://doi.org/10.1093/llc/fqu026>

- Fiormonte, D. y del Rio Riande, G. (2017). Por unas humanidades digitales globales. *Infolet*, 10. <https://infolet.it/2017/10/09/humanidades-digitales-globales>.
- Fostering the Practical Implementation of Open Science in Horizon 2020 and Beyond (FOSTER). (2021). *Open Science Definition*. <https://www.fosteropenscience.eu/foster-taxonomy/open-science-definition>.
- Gallon, K. (2016). Making a case for the black digital humanities. En M. K. Gold y L. F. Klein (Eds.), *Debates in the digital humanities 2016* (pp. 42-49). University of Minnesota Press.
- Gil, A. (2015). The user, the learner and the machines we make. Minimal Computing. <https://go-dh.github.io/mincomp/thoughts/2015/05/21/user-vs-learner/>
- Gil, A. (2016). Interview with Ernesto Oroza. En M. K. Gold y L. F. Klein (Eds.), *Debates in the digital humanities 2016*. University of Minnesota Press. <https://dhdebates.gc.cuny.edu/read/untitled/section/f2df3413-259e-46fe-9982-f1dba0c386fb>
- Gil, A. y Ortega, É. (2016). Global outlooks in digital humanities: Multilingual practices and minimal computing. En C. Crompton, R. J. Lane y R. Siemens (Eds.), *Doing digital humanities. Practice, training, research* (pp. 22-34). Routledge.
- Gray, T. J., Gainous, J. y Wagner, K. M. (2016). Gender and the digital divide in Latin America. *Social Science Quarterly*, 98(1), 326-340.
- Gold, M. (2012) *Debates in the digital humanities*. University of Minnesota Press.
- Instituto Nacional y del Profesorado (INTEF). (2013). Declaración sobre innovación y TIC del Foro de Ministros de Educación de las Américas. *Educalab* (blog). <http://blog.educalab.es/>

- intef/2013/06/26/declaracion-sobre-innovacion-y-tic-del-fo-ro-de-ministros-de-educacion-de-las-americas/
- Isasi, J. y Rojas Castro, A. (2021). ¿Sin equivalencia? Una reflexión sobre la traducción al español de recursos educativos abiertos. *Hispania*, 104(4), 613-624. <https://doi.org/10.1353/hpn.2021.0130>
- Kirschenbaum, M. (2010). What is digital humanities and what is doing in english departments? *ADE Bulletin*, (150), 1-7. https://mkirschenbaum.files.wordpress.com/2011/01/kirschenbaum_ade150.pdf
- Kirschenbaum, M. (2014). What is digital humanities and why are they saying such terrible things about it? *Differences*, 25(1). <https://mkirschenbaum.files.wordpress.com/2014/04/dh-terrible-things-kirschenbaum.pdf>
- Klein, J. T. (2015). *Interdisciplining digital humanities: Boundary work in an emerging field*. University of Michigan Press.
- Krajcik, J. S. y Blumenfeld, P. C. (2005). Project-based learning. En R. K. Sawyer (Ed.), *The Cambridge handbook of the learning sciences* (pp. 317-334). Cambridge University Press. <https://doi.org/10.1017/CBO9780511816833.020>
- Kreimer, P., Vessuri, H., Velho, L. y Arellano, A. (2014). *Perspectivas latinoamericanas en el estudio social de la ciencia y la tecnología*. Siglo XXI Editores.
- Krull, S. (2016). *El cambio tecnológico y el nuevo contexto del empleo: Tendencias generales y en América Latina*. Comisión Económica para América Latina y el Caribe (Cepal).
- Leeds-Hurwitz, W. y Stenou, K. (2013). *Intercultural competences: Conceptual and operational framework*. Organización de las Naciones Unidas para la Educación, la Ciencia y la Cultura (Unesco). <https://unesdoc.unesco.org/ark:/48223/pf0000219768>

- Losh, E. y Wernimont, J. (Eds.). (2018). *Bodies of information: Intersectional feminism and digital humanities*. University of Minnesota Press.
- Nyhan, J., Terras, M. y Vanhoutte, E. (Eds.). (2013). *Defining digital humanities. A reader*. Ashgate Publishing.
- Organización de las Naciones Unidas para la Educación, la Ciencia y la Cultura (Unesco). (2013). *Intercultural competences: Conceptual and operational framework*. <https://unesdoc.unesco.org/ark:/48223/pf0000219768>.
- Packer, A. L., Prat, A. M., Luccisano, A., Montanari, F., Santos, S. y Menghini, R. (2006). El modelo SciELO de publicación científica de calidad en acceso abierto. En D. Babini y J. Fraga (Comps.), *Edición electrónica, bibliotecas virtuales y portales para las ciencias sociales* (pp. 191-208). Consejo Latinoamericano de Ciencias Sociales (CLACSO). <http://bibliotecavirtual.clacso.org.ar/ar/libros/secret/babini/Parcker%20Part%20Lucisano.pdf>.
- Risam, R. y Gil, A. (2022). Introduction: The questions of minimal computing. *Digital Humanities Quarterly*, 16(2). <http://www.digitalhumanities.org/dhq/vol/16/2/000646/000646.html>
- Ruiz Vaca, J. O. (2022). *El acceso abierto y las humanidades digitales* [tesis de doctorado]. Universidad Nacional Autónoma de México. http://ru.atheneadigital.filos.unam.mx/jspui/handle/FFYL_UNAM/6277
- Sala, H. E. y Núñez Pölcher, P. N. (2014). Software libre y acceso abierto: Dos formas de transferencia de tecnología. *Revista Iberoamericana de Ciencia, Tecnología y Sociedad*, 9(26), 115-128.
- Schallier, W., del Rio Riande, G., O'Donnell, D., Hathcock, A. y Champieux, R. (2018). Where is the open in DH? *DH2018*. <https://dh2018.adho.org/en/where-is-the-open-in-dh/>.

- Spiro, L. (2012). "This is why we fight": Defining the values of the digital humanities. En M. K. Gold (Ed.), *Debates in the Digital Humanities 2012*. University of Minnesota Press. <https://dhdebates.gc.cuny.edu/read/untitled-88c11800-9446-469b-a3be-3fdb36bfbdie/section/9e014167-c688-43ab-8b12-of6746095335>
- Summers, E. (2016). *The web's past is not evenly distributed*. <https://mith.umd.edu/webs-past-not-evenly-distributed>.
- Svensson, P. (2016). Sorting out the digital humanities. En S. Schreibman, R. Siemens y J. Unsworth (Eds.), *A new companion to digital humanities* (pp. 476-492). John Wiley and Sons.
- The New London Group. (1996). A pedagogy of multiliteracies: Designing social futures. *Harvard Educational Review*, 66(1), 60-93.
- Van Zundert, J. (2012). If you build it, will we come? Large scale digital infrastructures as a dead end for digital humanities. *Historical Social Research*, 37(3), 165-186. <https://www.jstor.org/stable/41636603>
- Viglianti, R., del Rio Riande, G., Hernández, N. y De León, R. (2022). Open, equitable, and minimal: Teaching digital scholarly editing North and South. *Digital Humanities Quarterly*, 16(2). <http://www.digitalhumanities.org/dhq/vol/16/2/000591/000591.html>



PARTE III

{ **Formas de las humanidades digitales en la región**

**Biteratura en español: posibilidades literarias generativas
e interactivas para archivos digitales dinámicos**

SERGIO RODRÍGUEZ GÓMEZ

JUAN ÁLVAREZ UMBARILA

**A politização do consumo mediada pela tecnologia: o caso
do *Sleeping Giants* no Brasil**

DANIEL COELHO DE OLIVEIRA

***La fiebre del banano*: una aproximación a la curaduría de
arte contemporáneo a través de las humanidades digitales**

JUANITA SOLANO ROA

BLANCA SERRANO ORTIZ

**Movimientos sociales y humanidades digitales: el
estallido del 18 de octubre en Chile**

DAVID BORDONABA-PLOU

} ;

**Biteratura en español:
posibilidades
literarias generativas
e interactivas para
archivos digitales
dinámicos**

Para citar este capítulo:

<http://doi.org/10.51573/Andes.9789587987362.9789587987379.9789587987867.7>

SERGIO RODRÍGUEZ GÓMEZ

JUAN ÁLVAREZ UMBARILA

Introducción

- { En este capítulo proponemos integrar el campo de la creatividad computacional con la dinamización de archivos digitales. Definimos la *creatividad computacional* como una práctica en la que, por medio de procedimientos mecánicos, algorítmicos, hipertextuales o de interfaz, se potencian las formas en las que alguien puede expresarse artísticamente. Crear sistemas de reglas, jugar con el azar, o manipular datos y píxeles, más que una limitante para la creación, representa una oportunidad para el descubrimiento y la sorpresa. Como afirma la precursora del arte digital Vera Molnar en una entrevista, con estos métodos “siempre terminas con algo diferente y siempre es la misma cosa y nunca la misma

cosa” (Szántó, 2018, p. 85)¹. La *biteratura*², una forma particular de creatividad computacional, utiliza estos procedimientos para crear obras textuales y multimedia esencialmente digitales: textos nuevos que se generan a partir de remezclas, permutaciones, azar y condiciones probabilísticas; historias que se navegan de forma no lineal, en las que se pueden tomar decisiones acerca del desarrollo de los acontecimientos o en las que es posible perderse por múltiples ramificaciones narrativas.

El archivo digital es en principio diferente al físico porque su propósito no es únicamente proteger, conservar y preservar un conjunto de objetos, sino abrirse y reproducirse en múltiples espacios e integrarse a las interfaces con las que interactúan todo tipo de personas, no solo las conocedoras expertas. Sin embargo, el archivo digital a menudo imita las lógicas estáticas de los archivos físicos, pero con carpetas, cajas y documentos virtuales, que se hacen igualmente difíciles de acceder a los que se encuentran en los edificios físicos de la memoria, en especial para públicos no especializados o que carecen de credenciales académicas (Álvarez Umbarila, 2017). Por esto, para dar lugar a todo el potencial del archivo digital, hay que movilizarlo y abrirlo a la lógica dinámica de la internet.

Los ejercicios de creatividad computacional como los experimentos biterarios pueden servir para dinamizar los archivos

- 1 Todas las fuentes citadas que están originalmente en inglés fueron traducidas al español por los autores de este texto.
- 2 El término *biteratura* fue acuñado por Swirski (2014), quien lo propone como alternativa a *literatura electrónica*, porque, según este autor, *biteratura* hace una referencia más clara a los aspectos computacionales y digitales de este tipo de literatura.

digitales, de manera que potencien el descubrimiento y la sorpresa a través de estrategias algorítmicas e hipertextuales, y así invitar a los usuarios a descubrir documentos culturales de formas interesantes y complejas. En este sentido *dinamizar* quiere decir incorporar en el archivo digital estrategias de interactividad, sorpresa, personalización y posibilidades múltiples, para que, en lugar de tener un solo archivo general disponible para todos, se pueda contar con un archivo plural y cambiante para cada usuario.

Como caso de estudio, describiremos el desarrollo de una pieza de *software* que permite producir biteratura en español de múltiples maneras: la librería Aventura.js, para el lenguaje de programación JavaScript, creada por uno de los autores de este capítulo. También expondremos una serie de experimentos que realizamos aprovechando las prestaciones de Aventura.js como medio de dinamización de diferentes archivos públicos digitales: la Colección de Arte del Banco de la República de Colombia, la colección de poesía Un Libro por Centavos y el diario de campaña del artista Peregrino Rivera Arce durante la guerra de los Mil Días. Adicionalmente, reflexionamos acerca del desarrollo de la librería como respuesta a las brechas idiomáticas e infraestructurales existentes con respecto al *software* usado en humanidades digitales.

¿Qué entender por biteratura?

El libro *Magnitud imaginaria*, del autor de ciencia ficción Stanisław Lem (2017), es una recopilación de prólogos de libros académicos que no existen. Uno de estos pertenece al libro *Historia de la literatura bítica en cinco volúmenes*. Allí, tomándose la voz

de eruditos críticos literarios del futuro, Lem examina la extensa obra textual que producen las máquinas una vez cuentan con el tiempo libre, la inteligencia y la potencia computacional para hacerlo. En ese prólogo se cuenta que las máquinas empiezan imitando las estructuras lingüísticas de los humanos, luego son capaces de hacer crítica literaria y, por último, producen sus propios textos clásicos, imposibles de entender para nosotros los humanos. Con las complejas categorías de temas, formatos y estilos que surgen, la literatura bítica se convierte en un tema digno del estudio humanístico.

Basado en este juego futuroológico, en su libro *From literature to biterature*, Peter Swirski (2014) propone el término *biteratura* para referirse a todas las formas de creación literaria en las que interviene sustancialmente una máquina digital: ya sea en el proceso de autoría, el sustrato material de la obra o el proceso de su lectura. Con intervenir sustancialmente queremos decir que la obra no es solo un texto digitalizado, es decir, que surgió en otro medio y simplemente se reproduce lo más fielmente posible, sino que es nativamente digital, en el sentido en que fue producida por medio de estrategias informacionales, algorítmicas o es navegable de forma hipertextual. Por supuesto, *biteratura* no es el único término que se ha usado para este tipo de obras y prácticas literarias, también se usan los términos *literatura electrónica*, *literatura digital* o *ciberliteratura*, casi siempre de forma intercambiable. Sin embargo, creemos que el término *biteratura* es el que mejor transmite la idea de la intervención esencial del medio digital en la estructura y el significado de la obra; no solo como un adjetivo añadido a la palabra *literatura*, sino como una forma constitutiva de ciertas formas de creación literaria. En especial es ambiguo el término

más usado, *literatura electrónica*, porque puede llevar a pensar en circuitos, voltajes y corrientes, más que en las profundas implicaciones que tiene intentar crear y navegar obras literarias en el medio digital en su sentido más abstracto, es decir, en relación con la manipulación de la información a través de algoritmos computacionales y las interfaces digitales multimodales en las que se manifiesta tal información. Por el contrario, biteratura entremezcla la palabra “literatura” con la palabra “bit”, la unidad mínima de la información digital, a partir de la cual las máquinas ejecutan todas sus operaciones.

Los orígenes de la biteratura como forma de cooperación creativa entre máquina y humano pueden remontarse incluso a ejercicios previos al uso de computadores digitales. Podemos pensar, por ejemplo, en los juegos de la escritura automática, la escritura limitada y el cadáver exquisito de la vanguardia surrealista, los *cut-ups* —recortes remezclados de fragmentos de texto hechos por el autor William Burroughs—, las historias interactivas en los libros de la serie *Elige tu Propia Aventura* o librojuegos, o en la novela *Rayuela* de Julio Cortázar, que puede leerse en órdenes diferentes. Aunque no se usan computadores, en estos ejemplos existen estrategias que podríamos llamar *algorítmicas*, porque su creación involucra procesos mecánicos basados en reglas, o *interactivas*, porque piden la participación del lector en el desenvolvimiento del sentido de la obra; así, de cierta forma, son protodigitales y protobiterarias.

Desde el surgimiento y uso extendido de los computadores digitales como herramientas creativas surge la biteratura propiamente dicha. De acuerdo con el académico puertorriqueño Leonardo Flores (2017, 2021), lo que nosotros aquí llamamos

biteratura puede pensarse desde las categorías de tres generaciones históricas, además de la mencionada protobiteratura. La primera estuvo marcada por la producción de obras literarias que surgían de procesos computacionales, pero que eran distribuidas en medios físicos: impresiones, cintas magnéticas, CD-ROM, televisores, pantallas led, etc. Luego, a mediados de los años noventa empezó la segunda generación, que se vio potenciada por la distribución y las prestaciones de internet y la *world wide web* en sus primeras versiones, tímidamente multimedia. Finalmente, la tercera generación emplea dispositivos móviles, dispositivos de realidad virtual y aumentada, e interfaces que propician interacciones hápticas, y aprovecha las nuevas formas de lectura y relacionamiento común que origina la internet contemporánea a través de las redes sociales, como es el caso saliente de los bots de X (antes llamada Twitter) o los videojuegos en línea.

Las formas concretas en las que una obra literaria podría ser biteratura son muchas, y, como sucede con cualquier otra forma de creación, su horizonte está en expansión. Aquí nos referiremos principalmente a tres estrategias y daremos algunos ejemplos latinoamericanos que las han usado:

Una primera estrategia biteraria consiste en la manipulación de textos ya existentes por medio de procedimientos como el reemplazo de palabras, la remezcla de fragmentos de un mismo texto, la unión fragmentada de textos disímiles o la elaboración de juegos estéticos que transforman la presentación, la diagramación o la modalidad del texto original. En Latinoamérica encontramos ejemplos como la obra *Promesas*, de la artista colombiana Ana María Montenegro (2017), que junta discursos de diferentes figuras políticas mediante palabras comunes, o los *Tipoemas* y

anipoemas de la argentina Ana María Uribe (1995), una serie de exploraciones literarias que surgen de juegos tipográficos y animados en la interfaz de la web³.

Una segunda estrategia consiste en la producción de texto generativo. De acuerdo con Jean-Pierre Balpe (2005), la literatura generativa es:

una literatura en donde los textos se producen usando un computador por medio de un conjunto de reglas formales, el uso de cualquier clase de algoritmo, diccionarios específicos y representaciones enciclopédicas. Eso quiere decir que es una literatura en la cual el autor no escribe el texto final, sino que solamente trabaja en el nivel de componentes abstractos de alto nivel.

Estos sistemas pueden partir de métodos diversos: por ejemplo, la generación de texto basado en probabilidades derivadas del procesamiento estadístico de un corpus; el entrenamiento de redes neuronales y otras formas de inteligencia artificial, a partir de grandes conjuntos de datos; o la construcción de sistemas de expertos, gramáticas u otras formas de estructuración de procedimientos. Una particularidad del texto generativo es que suele combinar estrictos sistemas de reglas con elementos aleatorios e impredecibles; esta combinación paradójica produce diversidad dentro de un mundo de posibilidades limitado. Como afirma Evens (2020),

3 Para un panorama más amplio de la biteratura electrónica en Latinoamérica, <https://www.cartografiadigital.cl/> y <https://www.observatorioldigital.ufscar.br/atlas-da-literatura-digital-brasileira/>.

“los métodos combinatorios promueven una potencial reautoría al dar lugar a yuxtaposiciones azarosas, moviéndose deliberadamente entre la intención y el accidente y así también entre el significado y su ruptura” (p. 235). Un ejemplo latinoamericano es el libro *A noise such a man might make* del escritor argentino Milton Läufer (2018), que fue creado usando el método de cadenas de Markov y está basado en el corpus de las obras *The man and the sea*, de Ernest Hemingway, y *The road*, de Cormac McCarthy.

La tercera estrategia biteraria que queremos destacar consiste en la producción de obras hipertextuales y narraciones interactivas: la producción de textos que responden, se modifican y cambian su curso narrativo a partir de interacciones por parte de los usuarios. Por ejemplo, la personalización estética de la historia a través de interfaces de usuario, la navegación no lineal entre diferentes textos o, como sucede con los videojuegos, la inmersión de las acciones del usuario en el mundo de la historia. Ensslin (2007) afirma que

Hipertexto se refiere a una forma de literatura interactiva computacional, más que todo basada en lenguajes de programación, que traduce formas previas de escritura lineal en un formato no lineal que promueve procesos de lectura múltiples. El componente estructural y estético más crucial es el hipervínculo, pues abre al lector muchas posibilidades de navegación. (p. 20)

Un ejemplo en nuestro contexto de tales tipos de obras es el cómic interactivo colombiano *Muerte en el bosque*, escrito por Catalina Holguín e ilustrado por Joni B. (2017), que ofrece a los lectores una historia de misterio en la que se deben tomar decisiones que llevan por diferentes caminos hacia resoluciones distintas.

El desarrollo de Aventura.js

Parte del proceso de adopción de prácticas biterarias, en el marco amplio de la creación artística y la investigación humanística digital, consiste en el desarrollo de herramientas que permitan y faciliten su producción. Sin ellas, la puerta de acceso a la biteratura queda restringida solo a personas con conocimientos especializados tanto en campos computacionales como en prácticas artísticas. Esto es un problema de infraestructura, como lo entendería Star (1999); para ella, las infraestructuras son entramados tecnosociales que sirven para organizar sistemas conceptuales, crear convenciones de práctica, servir como interfaz técnica y, en general, permitir y agenciar prácticas humanas. Así como para producir textos convencionales es necesario tener elementos materiales, como papel, lápiz o un computador con un editor de texto, pero también teorías, comunidades, experiencias y educación, para producir biteratura es necesario contar con algoritmos, *software* y métodos mecánicos que den resultados complejos y sirvan como medio de expresión, al igual que comunidades que tengan interés en discutir, mejorar y organizar sus prácticas. Todos estos son elementos infraestructurales. En esta sección del capítulo queremos dar cuenta del desarrollo de Aventura.js, una librería biteraria en español para el lenguaje de programación JavaScript, que fue usada para dar lugar a todos los experimentos de dinamización del archivo digital que mencionamos en secciones posteriores. El desarrollo de Aventura.js busca aportar a la construcción infraestructural de la biteratura en español, así como establecer métodos que faciliten su aplicación en campos como las humanidades digitales y la creatividad computacional.

Hay que decir que, en el presente, existe una moderada oferta de *software* que permite la creación de piezas de biteratura a partir de las diferentes estrategias creativas que resumimos en secciones anteriores. Por ejemplo, desde la estrategia de manipulación de textos existentes encontramos *software* como la suite experimental Applied Poetics, desarrollada por Douglas Luman (2016), o The N+7 Machine, de Peter Christian (2007). Para la estrategia de producción de texto generativo, por su parte, pueden destacarse las librerías Tracery.js, hechas por Kate Compton Galaxykate, (Compton *et al.*, 2015) y Rita.js, por Daniel C. Howe (2009), ambas para el lenguaje de programación JavaScript, que permiten la creación de gramáticas libres de contexto, la generación de texto mediante modelos estadísticos y el procesamiento de lenguaje natural básico. También existen modelos de inteligencia artificial propietarios basados en métodos de *deep learning* con los que es posible generar texto automáticamente —por ejemplo, el masivo modelo GPT-3 de la empresa OpenAI (Brown *et al.*, 2020)—. Por último, para la estrategia de la literatura hipertextual y las historias interactivas, se destacan el *software* Twine y la librería Ren'Py para el lenguaje de programación Python. Con estas herramientas es posible diseñar historias no lineales, formas de navegación interactiva y diferentes formas de presentación en la interfaz computacional: imágenes, sonido, video, texto, entre otras.

No obstante, aquí salta a la vista que la gran mayoría de estas opciones fueron creadas para funcionar dentro de un contexto de producción anglófona y son nulas o mínimas las alternativas de *software* para la creación biteraria en español. En ocasiones es irrelevante, pues hay *software* que puede usarse en cualquier lenguaje, pero en otras ocasiones implica la imposibilidad de aprovechar el

potencial creativo del *software* en su totalidad. Por ejemplo, con la librería *Rita.js* se pueden hacer algunas operaciones de análisis lingüístico, como el etiquetado de partes del discurso, pero solo para textos escritos en inglés. Así, en nuestro contexto es común el uso limitado de las herramientas en inglés o la creación de herramientas *ad hoc* para proyectos específicos. Esto tiene como consecuencia que sea común el uso de infraestructuras prestadas de otros contextos, que no necesariamente se adaptan a las idiosincrasias de la creación en nuestro idioma o a nuestras formas de creación, o, por otra parte, que se requiera de un conocimiento especializado en programación para poder obtener resultados interesantes. Es decir, la puerta de acceso a la biteratura es más estrecha si se busca hacer en español, pues operar dentro del mundo de las humanidades digitales y de la creatividad computacional implica, por defecto, cierta familiaridad con un idioma no nativo. Para Takhteyev (2012), se trata de una “situación ‘diaspórica’ de los practicantes de la periferia, quienes se involucran simultáneamente en dos culturas: la cultura central local y el mundo de práctica global” (p. 6). Adicionalmente, debido al contexto de producción externo de estas herramientas, se deriva un efecto de caja negra, es decir, quienes usan las herramientas solo lo hacen como usuarios, pero no conocen los mecanismos que operan en el fondo y, por lo tanto, no poseen la agencia para modificar o cuestionar los presupuestos del *software*; en otras palabras, cuando la infraestructura falla, es muy difícil auditarla o mejorarla. Desde el punto de vista de Fiorimonte (2018), las humanidades digitales desde las márgenes presentan una desventaja creativa e investigativa, porque existe una brecha idiomática en los mundos digitales y en la producción de herramientas, y se debe trabajar activamente para reducirlas.

Esto va desde el código de programación a los textos canónicos de las humanidades digitales.

De todo lo anterior se puede concluir que es necesaria la producción de *software* y dependencias ancladas a lenguajes de programación que no operen como cajas negras y que permitan tanto la expresión creativa situada, y en nuestro idioma, como el desarrollo de una literacidad procedimental para construir sobre las infraestructuras disponibles. En los términos de Mateas (2008), la literacidad procedimental se refiere a “la habilidad de leer y escribir procesos, de involucrarse en representaciones y estéticas procedimentales, en entender la interacción entre las prácticas incrustadas culturalmente de la producción de sentido humana y los procesos mediados técnicamente” (p. 67).

Dentro de este panorama, y al pensar en estos vacíos idiomáticos y de accesibilidad, uno de los autores de este capítulo, Sergio Rodríguez Gómez, desarrolló *Aventura.js*, una librería de programación, esto es, una pieza de infraestructura digital que facilita y agencia, para la experimentación biteraria en español. *Aventura.js* funciona en el marco del lenguaje de programación JavaScript. La librería está en código abierto, no requiere de dependencias adicionales y está disponible en un repositorio de GitHub⁴.

El proceso de creación de la librería ha sido iterativo y experimental, debido a que su creador no tiene formación en desarrollo de *software*, sino en artes y humanidades. La versión actual, al momento de publicación de este texto, es la 2.5.0, y desde el 2020 se han lanzado cinco versiones. Así, ha sido un proceso de aprendizaje sobre la marcha y de retroalimentación constante, en buena

4 Véase <https://github.com/srsergiorodriguez/aventura/>.

medida alimentado por las discusiones y la construcción de proyectos como los que presentamos en este capítulo. En ese proceso hemos aprendido que una librería de programación es una colección de funciones que debe ser aplicable para propósitos y entornos generales. El desarrollo de la librería también ha estado marcado por la escritura de su documentación, es decir, las instrucciones y las referencias de su uso. En términos de sostenibilidad, desde el punto de vista financiero y humano, la librería no tiene ninguna forma de financiamiento, ni está anclada a ninguna institución por el momento, sino que ha sido creada y actualizada por su autor en su tiempo libre; desde el punto de vista técnico, debido a que la librería no tiene dependencias adicionales al código básico de JavaScript y este, a su vez, se considera un lenguaje de programación sólido y estable, hay menores posibilidades de que Aventura.js caduque rápidamente. Un asunto que merece mención adicional es el de las plataformas de terceros donde se despliega la librería, por ejemplo, el caso de los bots de X basados en Aventura.js que se describen en este capítulo. El funcionamiento de estos bots depende en buena medida de las políticas que establece X para el uso de sus API⁵; estas políticas son mutables y dependen de asuntos técnicos, y coyunturas políticas y económicas. En el caso en el que la posibilidad de usar ciertas plataformas se imposibilite, Aventura.js está construida de manera que se puedan migrar los proyectos a otras plataformas con relativa facilidad⁶. Finalmente, para garantizar

- 5 Las API son interfaces de programación de aplicaciones que permiten a dos piezas de *software* diferentes comunicarse e interactuar entre sí.
- 6 Desafortunadamente, en el tiempo de la escritura de este capítulo y su publicación, las políticas de uso de x han cambiado y no es posible alojar los

mayor sostenibilidad a futuro, es clave encontrar otras personas o apoyo institucional que deseen mantener la librería actualizada. Como todo *software* sujeto a cambios tecnológicos y convenciones de uso, la pregunta sobre su futuro es hasta cierto punto abierta e incierta, y está sujeta a la recepción y utilización por parte de los usuarios, su adopción institucional, formas de financiación, etc.

Con respecto a los usos concretos de Aventura.js, permite tres formas de creación biteraria que se describirán a continuación, y que corresponden a las estrategias de texto generativo y texto hipertextual. Para la parte de creación generativa es posible producir textos a partir de dos métodos algorítmicos: gramáticas libres de contexto y cadenas de Markov.

La teoría de gramáticas libres de contexto fue propuesta originalmente por Chomsky (1956) como un sistema de descripción formal de las transformaciones sintácticas posibles en un lenguaje natural. Posteriormente, la teoría se extendió al análisis de objetos digitales, como los lenguajes de programación, y versiones alternativas derivaron en métodos para la producción de literatura generativa. En términos generales, una gramática libre de contexto describe la estructura de un lenguaje a partir de un vocabulario y un sistema de reglas de combinación de ese vocabulario. De acuerdo con Eco (2011), un sistema de signos puede entenderse a través de dos dimensiones: una sintagmática y otra paradigmática. La dimensión sintagmática define la estructura o el orden de las unidades que componen una expresión, mientras que la paradigmática define las posibles opciones que pueden

.....
proyectos gratuitamente o acceder a las funcionalidades de la API de la plataforma sin hacer pagos adicionales.

ocupar esa estructura. Por ejemplo, si pensamos en la organización sintáctica convencional de una frase simple en español, podemos definir una estructura sintagmática en la que primero se pone un artículo, luego un sustantivo y después un verbo; en términos paradigmáticos, podríamos decir que los artículos disponibles en el vocabulario para llenar parte de la estructura son: el, la, los, las, un, una, unos y unas. Las gramáticas pueden construirse como complejos árboles de reglas que permiten muchas permutaciones; incluso el concepto de gramática puede extenderse a otros sistemas de signos. Podríamos imaginar que existe una gramática de los rostros, una organización sintagmática que define que, verticalmente, arriba están los ojos, en la mitad la nariz y abajo la boca, y unas posibles opciones paradigmáticas, por ejemplo, un inventario de ojos dentro del que podemos escoger.

Aventura.js implementa en código de programación estas ideas y permite describir gramáticas en las dimensiones sintagmáticas y paradigmáticas. Así, usando la sintaxis del lenguaje de programación, un árbol gramatical que contiene una multitud de variables de animales y acciones podría dar como resultado frases como “la nutria come” (figura 7.1)

```

const gramatica = {
  base: [“<parteNominal> <parteVerbal>”],
  parteNominal: [“<articulo> <sustantivo>”],
  parteVerbal: [“corre”, “come”, “salta”],
  articulo: [“la”, “una”],
  sustantivo: [“ardilla”, “gata”, “nutria”]
}
    
```

```
const a = new Aventura("es").fijarGramatica(gramatica);
const textoGenerado = a.expandirGramatica("base");
//Un resultado de texto generado podría ser: "la nutria
come      "
```

FIGURA 7.1. Un ejemplo básico que demuestra el uso de la librería *Aventura.js* para la creación de textos generativos a partir de gramáticas libres de contexto

FUENTE: elaboración propia.

Por otra parte, el sistema de cadenas de Markov, inventado por Andrey Markov y descrito por Chomsky (1956) en su versión lingüística, es un proceso estadístico a través del cual se mide la probabilidad de que una observación siga después de otra. En el caso de un generador de texto, la unidad lingüística se denomina un *n-grama* y puede estar compuesta por un número arbitrario de caracteres o palabras. Por ejemplo, un *bi-grama*, o un *n-grama* $n=2$, puede estar compuesto por dos letras o dos palabras. El modelo estadístico se deriva de un corpus en el que se procesan los posibles *n-gramas* que podrían seguir a otro y sus probabilidades derivadas de la frecuencia en la muestra. Así, se puede construir un algoritmo o sistema de reglas en el que se parte de una palabra inicial, llamada *semilla*, y a continuación se agregan palabras de acuerdo con una escogencia probabilística aleatoria.

Aventura.js implementa un método que se alimenta de un corpus en texto crudo (TXT), del cual se deriva un modelo estadístico basado en un tamaño arbitrario para los *n-gramas*, y con tal

modelo se generan nuevas versiones de texto. La figura 7.2 muestra un ejemplo de la generación de un texto basado en las probabilidades estadísticas de una selección de cuentos de Julio Cortázar.

```

const a = new Aventura("es");
const ngrama = 2;
const nombreArchivo = "cortazar.txt";
const numeroPalabrasGeneradas = 320;
const semilla = "el auto";

a.modeloMarkov(nombreArchivo, ngrama).then(modelo => {
  const textoGenerado = a.fijarMarkov(modelo)
    .cadenaMarkov(numeroPalabrasGeneradas, semilla);
});

// Un posible texto generado: "El auto de juguete sobre
// los caracoles y su madre. Se golpea la cabeza. Ustedes
// no querrían que yo fuese un microbio. Rema, pero no era
// el verdadero paseo: una pausa, luego la conversación
// como 'el del olor que más temía, y antes de acostar-
// nos le dije sin mirarlo de frente, me duele tanto desde
// ayer, un tópico tras otro"
    
```

FIGURA 7.2. Un ejemplo básico que demuestra el uso de la librería `Aventura.js` para la creación de textos generativos a partir de cadenas de Markov

FUENTE: elaboración propia.

Finalmente, las historias interactivas o hipertextuales se producen a partir de la navegación con vínculos entre un texto y otro. Aquí debemos entender *texto* en un sentido ampliado, pues en las interfaces digitales existen muchos modos de presentación: sonido, imágenes, textos convencionales, video, entre otras. En *Aventura.js* es posible crear historias interactivas en las que se navega de forma no lineal entre diferentes escenas. La interfaz de navegación permite crear botones o áreas en porciones de imágenes que remiten a lugares diferentes de la narración. Así, es posible crear estructuras narrativas basadas en árboles de decisión, en las que se pueden definir caminos diferentes del desarrollo narrativo que pueden ser tomadas por la audiencia. Un ejemplo de una estructura simple de narración interactiva se muestra en la figura 7.3.

```
const escenas = {
  inicio: {
    texto: "Érase una vez un círculo...",
    opciones: [
      {
        btn: "dejar tranquilo",
        texto: "dejas al círculo en paz",
        escena: "final1"
      },
      {
        btn: "aplastar",
        texto: "...aplastamiento activado...",
        escena: "final2"
      }
    ]
  }
}
```

```

    },
    final1: {
      texto: "Perfecto, un final redondo",
      sinSalida: true
    },
    final2: {
      texto: "Parece que la historia formó una elipsis",
      sinSalida: true
    }
  }
}

const a = new Aventura("es")
a.fijarEscenas(escenas).iniciarAventura("inicio")

```

FIGURA 7.3. Un ejemplo del diseño y ejecución de un sistema de escenas de acuerdo con la sintaxis de Aventura.js

FUENTE: elaboración propia.

Cabe añadir que todos estos sistemas son combinables. De esta manera, es posible generar historias interactivas en las que además se producen sobre la marcha fragmentos generativos de texto, de acuerdo con las decisiones de quien lee.

Adicionalmente, la librería Aventura.js cuenta con documentación y ejemplos detallados en el repositorio⁷ que guarda el código fuente. Esta documentación está escrita en inglés y en español.

7 Véase https://github.com/srsergiorodriguez/aventura/blob/master/README_es.md

Como se ve en los ejemplos anteriores, los nombres de las funciones de la librería pueden escribirse en español (aunque también hay versiones de las mismas funciones en inglés). Aunque esta pueda parecer una decisión menor, es una forma de recalcar que esta es una infraestructura pensada y creada intencionalmente para asumir la brecha idiomática mencionada, y para permitir la creación biteraria en español desde un lugar más cómodo, más familiar y menos intimidante.

El archivo digital dinámico

Los archivos digitales plantean posibilidades adicionales a la lógica tradicional del archivo, que son sobre todo evidentes cuando se examinan en paralelo a sus contrapartes físicas. El archivo físico, por un lado, obedece generalmente a los principios fundamentales de proteger, conservar y preservar un conjunto de objetos o documentos considerados dignos del futuro. Sin embargo, como señala Jeffrey Schnapp (2008), esos objetos permanecen de cierto modo en el pasado, y solo resurgen en el presente de formas estrictas y controladas. Los objetos del archivo físico están contenidos y ordenados en cajas y paredes, separados del mundo exterior y accesibles, en su mayor parte, para quien tenga la intención específica (y muchas veces las credenciales apropiadas, de especialista) para abrirlo.

El archivo digital, por otro lado, es potencialmente un espacio ilimitado de exhibición, almacenamiento y exploración. Al no ser ya su propósito más saliente proteger y conservar para el futuro, sino proveer acceso amplio en el presente, podría decirse que su

enfoque está más direccionado a los procesos que a los objetos mismos y que la lógica para su acceso es más la de multiplicar que la de restringir (Burdick *et al.*, 2012). En principio, el archivo digital tiene el potencial de movilizar y exteriorizar sus contenidos en entornos múltiples y simultáneos, desacralizados y cotidianos, por personas que no necesariamente son especialistas ni buscan en el archivo con una especificidad obligatoria.

En términos metafóricos, podría hablarse con cierta generalidad del archivo físico como un cofre cuyos tesoros enterrados hay que proteger y desvelar, y del archivo digital como una ventana que trae los tesoros a la superficie, a los ojos de los transeúntes y a los caminos que transitan día a día (Álvarez Umbarila, 2017). Aunque no son contrarios ni antagonistas, sino que forman parte de un mismo sistema archivístico complementario, sí es posible observar en cada uno características propias que los hacen más proclives a cumplir unos roles que otros: sean de orden y preservación en el archivo físico, o de posibilidades múltiples de exhibición y experimentación en el archivo digital.

Sin embargo, y aunque el archivo digital tenga hasta cierto punto aquellas características propias que lo hacen más como una ventana, muchas veces estas características no están ejercitadas en su totalidad o simplemente están durmientes, existiendo apenas de forma potencial. Cabe aquí recordar la distinción que plantea Johanna Drucker (2013) entre lo *digitalizado* y lo *digital*: mientras que lo primero es mayormente un conjunto de archivos convertidos y depositados en un repositorio digital en el que pueden ser consultados, lo segundo es una forma de pensar y relacionar objetos mediante procesos digitales. Lo digital exige situarlos en una red dinámica de procesamiento que los resignifique de

modo constante, y exige necesariamente “repensar la identidad, el propósito, el uso y la sustancia de los objetos” (p.8). En otras palabras, mientras que el archivo digitalizado es más como una caja estática con objetos depositados que es posible recuperar, el archivo plenamente digital se comporta como una red abierta y móvil, en la que los objetos son en relación con otros objetos, espacios, procesos y usos.

En efecto, los archivos digitales también pueden *ser*, y a menudo son estructurados y ejecutados mediante las lógicas analógicas de los archivos físicos, que son estáticas. Lo *estático* puede entenderse aquí como las formas unívocas y generales de experimentar contenidos archivísticos, que, aunque son recuperables mediante búsqueda, no son experimentables de formas plurales, interactivas o creativas. Pero no existe una sola línea definitoria entre un tipo de archivo y otro. Un ejemplo sencillo de dinamización en entornos estáticos puede ser introducir aleatoriedad en el orden de los resultados de un buscador de un archivo: una función de resultados aleatorios agrega dinamismo a la interacción en el azar de la información recuperada (la hace nueva cada vez) y, en principio, fomenta la exploración. Este puede ser el caso en el archivo digital de la colección de arte de un museo.

Los elementos de dinamismo son especialmente importantes en los archivos culturales digitales, porque su rango de audiencias y usos se extiende más allá de un público especializado, y forma parte del objetivo democratizante de este tipo de archivos poder llegar y conectar de manera significativa con esa multiplicidad de personas. Como señala Angeliki Tzouganatou (2021), “indiscutiblemente, los archivos culturales deberían poder ser usados de formas creativas, y no almacenados en ‘bodegas digitales’ para ser

consultados solo por los archivistas y otros expertos” (p. 1). De hecho, los elementos de exploración y descubrimiento deben existir activamente en y para los archivos digitales culturales, y no deben atribuirse como una obligación de los usuarios, por cuanto estas personas no necesariamente llegan allí con un propósito claro, y la navegación incierta entre “cantidades excesivas de material puede convertirse en una experiencia insatisfactoria” (Ojamaa e Ibrus, 2021, p. 393).

Lo que aquí proponemos es una conceptualización del archivo digital dinámico a partir de dos ideas fundamentales: movimiento y apertura; basados en ella, exponemos tres experimentos de dinamización del archivo mediante algunas herramientas de la biteratura, en particular la generación computacional y la narrativización interactiva a partir de la librería *Aventura.js*.

El espacio del archivo digital dinámico existe en tanto existen las relaciones entre sus objetos, y mientras existan interacciones con otros espacios externos y con agentes humanos con objetivos plurales y diversos para aquellas interacciones. Las tres proposiciones para definir el espacio que hace la geógrafa Doreen Massey (2005) son útiles para conceptualizar el archivo dinámico: “el espacio es el producto de interrelaciones dadas por interacciones, basadas en la existencia de pluralidad, en construcción permanente” (p. 9). El archivo así concebido está en constante transformación, se debe a una pluralidad en interacción y está abierto al cambio y a un mundo con el que conjuntamente adquiere significado, resignificándose en cada interacción y en cada uso. Contrariamente al archivo estático, cuyo objetivo es ser más definitivo, más general, que espera a que saquen algo de él; el archivo dinámico respira, es plural, cambia y, como siempre está

construyéndose, nunca se cierra, pues no tiene la pretensión de ser o estar terminado.

Si cambiamos el énfasis y lo ponemos, ya no en los objetos del archivo, discretos entre sí y almacenados juntos, sino en las relaciones entre ellos y sus usos posibles, y hacia los usuarios que con diferentes propósitos los utilizan, podemos imaginar un archivo digital en movimiento. No se trata únicamente de la recuperación de un documento en un entorno digital, sino de cómo un documento existe en relación con otros. Tal y como explica Wolfgang Ernst (2013), el enfoque se desplaza hacia una “dinámica algorítmica en lugar de una quietud documentaria” (p. 97). Relacionar los objetos se logra en una conjunción entre el proceso computacional de la máquina, que genera relaciones, y una intencionalidad humana del usuario, que es circunstancial y cambia según la identidad, el momento, el espacio, etc.

Para imaginar el archivo digital dinámico como abierto, es necesario un proceso similar: hay que alejarse de la idea del archivo como un contenedor con paredes que lo separan del mundo. En lugar de eso, hay que percibirlo y estructurarlo como una red de conexiones que forma parte de otras redes de conexiones mayores. El archivo dinámico no es solo una entidad “aislada y autosuficiente, sino un aparato social enmarcado en interacciones que tienen que ver con la memoria, información, conocimiento, agentes, usos y usuarios” (Álvarez Umbarila, 2017, p. 21). La apertura no se detiene en los “datos abiertos” (Tzouganatou, 2021); más bien, requiere de un proceso que permita que los datos lleguen a las manos de los usuarios y sean usados y, asimismo, que a partir de los datos existentes nuevos datos sean creados: “lo que se necesita no es necesariamente más datos abiertos, sino más

datos ‘útiles’” (Tzouganatou, 2021, p. 337). Así, el proceso de apertura, para ser realmente abierto, tiene que recorrer activamente la distancia adicional para pasar de información a conocimiento, ya que solo con la participación real del archivo se puede concretar su democratización (Derrida, 1995). Por ello, la lógica del archivo dinámico tiene que parecerse menos a la del archivo físico, que almacena y contiene objetos consultables, y más a la de internet, que se define no por sus contenidos, sino por sus protocolos de hipervinculación (Ernst, 2013).

De esta manera, podemos ver que movimiento y apertura son interdependientes en el archivo digital dinámico, y por fuerza se siguen y suceden mutuamente. Los principios relacionales (entre objetos del archivo, con las intenciones y entornos de los usuarios) implican necesariamente movimiento, y si del movimiento se sigue el cambio, significa que el archivo nunca está terminado, nunca está cerrado. Igualmente, derrumbar las paredes, comunicar y abrir el archivo, hacia el exterior y hacia lo nuevo, obliga a que el movimiento no cese, porque el archivo en comunicación con los usos y con el mundo es siempre un archivo nuevo, en flujo constante.

Experimentos para un archivo digital dinámico

A partir de los principios de la biteratura en general y de las prestaciones de Aventura.js en particular, proponemos a continuación tres experimentos de movimiento y apertura para dinamizar el archivo cultural digital de formas generativas e interactivas.

El archivo abierto desplazado hacia otros espacios digitales: experimento con gramáticas libres de contexto⁸

Derribar las paredes del archivo y hacerlo móvil exige revertir la direccionalidad tradicional del usuario que va al archivo; en cambio, el archivo va hacia el usuario, hacia sus espacios digitales cotidianos. Eso hace, primero, que ejercitar el archivo no sea el resultado de un propósito específico, sino que lo convierte en una sorpresa en medio del flujo de la vida común y corriente; y, segundo, que los objetos del archivo sean resignificados según los espacios y actividades cotidianas de quien los usa. Así, el archivo abandona su lugar institucional sacralizado y se convierte en parte de la vida común de las personas; explorarlo se vuelve una actividad orgánica y aleatoria, que conduce a la sorpresa del encuentro significativo y no planeado entre el usuario y el archivo. ¿Dónde significan los objetos?, ¿para quién? Si abrimos el archivo y lo llevamos a las manos de los usuarios, sus elementos van a significar en relación con la experiencia particular de esa persona, que no tiene que buscar una caja y abrirla para encontrarlos, sino que los encuentra en su propio camino y los relaciona con emociones y circunstancias de su propia vida.

Este experimento de gramáticas libres de contexto con Aventura.js propone una serie de interacciones basadas en la unión entre textos autogenerados y elecciones del usuario por medio de un bot en X (@OyeArchiBot). De acuerdo con el texto generado por la máquina y completado por el usuario, se provee una imagen y sus metadatos de la colección de arte del archivo digital del

8 Véase <https://twitter.com/OyeArchiBot>

Banco de la República de Colombia. El objetivo del experimento es vincular una pieza artística al contenido (y a las emociones que de este se desprenden) de un texto producido por el computador y un humano.

La situación —es decir, el texto autogenerado que va a ser autocompletado— es presentada por el bot a partir de un *prompt* del usuario: un trino que contenga los términos “@OyeArchiBot” y “situación” en cualquier orden y con las palabras adicionales que sean. Por ejemplo, el trino: “@OyeArchiBot: dame una situación para hoy”. A partir de ese *prompt*, el bot desenvuelve la gramática y genera una situación cotidiana que va a ser completada en un trino respuesta al trino original. La gramática de la versión beta puede producir oraciones en cuatro macrocategorías de lo cotidiano (desplazamientos, familia, trabajo/estudio y ocio) que se descomponen en diferentes subcategorías, que, a su vez, tienen diferentes subopciones, lo que crea un árbol que se ramifica en una amplia gama de combinaciones posibles. Un ejemplo de una interacción con un usuario sería el siguiente:

Usuario: dame una situación para hoy.

Archibot: aceleración en la entrevista de trabajo, soy una bola de (a) tranquilidad, (b) angustia, (c) melancolía.

Usuario: angustia.

En seguida, el usuario completa la oración, en otro trino respuesta, con la opción que más acorde le parezca, a lo que el bot responde con la parte final de la interacción: un resumen de la oración completa y una imagen que corresponde con la información del texto autocompletado. La imagen es traída hasta la interfaz de X

desde el archivo digital del Banco de la República, se presenta con su título correspondiente, y se incluye en el trino un enlace para que el usuario pueda visitarla en el archivo original, descubrir más de ella y posiblemente explorar más de sus contenidos⁹.

Lo que permite vincular la imagen con el texto generado y completado es un sistema de metadatos de etiquetado (figura 7.4), con el que se asigna a cada imagen de un conjunto de datos una serie de atributos en varias categorías (para la versión beta, cuatro categorías: emoción, color, acción y emoji), que a su vez se descomponen cada una en cinco posibilidades. Estos metadatos están expresados en el texto generado-completado; en cada caso, uno es autogenerado por la máquina y el otro es la elección del usuario (de tres posibilidades también autogeneradas). En nuestro ejemplo actual: “[etiqueta de acción autogenerada: *Aceleración*] en la entrevista de trabajo, soy una bola de... [etiquetas de emoción a elegir: a) tranquilidad b) angustia c) melancolía]”. A partir de ahí, el sistema escoge una de las imágenes del conjunto de datos que cumpla con las condiciones de las dos etiquetas y la asigna como respuesta a la oración generada-completada.

Esta interacción en cuatro movimientos es en esencia la concatenación de dos gramáticas libres de contexto: una textual, generada por la máquina y completada mediante la interacción del usuario; y otra de metadatos y visual, que junta las etiquetas del texto y la imagen, y completa el sentido del texto con una imagen escogida del archivo, en lo que podría considerarse una sintaxis

9 Para consultar la imagen y el resultado de la interacción se puede consultar visitando el siguiente enlace: <https://twitter.com/OyeArchiBot/status/1525284004672552961>

33	Transmisión	angustia, rabia	verde	estruendo, despl	🤖	https://www
34	El cafetal	tranquilidad	verde	quietud	❤️	https://www
35	Nocturno medieval	angustia	azul	ascenso, acelera	🙏	https://www
36	Bodegón	tranquilidad, melanc	azul	quietud	❤️, 🙏	https://www
37	Color amazonía	regocijo	ámbar	estruendo	😎	https://www
38	Sin título	melancolía	verde	aceleración	🤖	https://www

```

const categorias = {
  emocion:
    ['melancolía', 'angustia', 'rabia', 'tranquilidad'],
  color: ['gris', 'azul', 'púrpura', 'verde', 'ámbar'],
  accion: ['quietud', 'aceleración', 'ascenso', 'desplo-
me', 'estruendo'],
  emoji: ['❤️', '🙏', '🤖', '😎', '🤖']
}

```

FIGURA 7.4. Una muestra del sistema de metadatos subyacente al bot @OyeArchibot en su modo de gramáticas libres de contexto

FUENTE: elaboración propia.

de significado textual y visual, que está mediada por los metadatos y tiene una interpretación abierta.

A partir de la creación de gramáticas libres de contexto autogeneradas, junto con la participación interactiva de humanos, este experimento también busca profundizar en el uso de bots culturales, que en los últimos tiempos se ha hecho más común en redes como X (en especial bots de imágenes y textos de autores

específicos)¹⁰. Con el ejercicio interactivo propuesto se pretende superar algunas limitaciones de los bots culturales que son meramente expositivos de un objeto o un fragmento de información aleatorio; en palabras de Angeliki Tzouganatou (2018), son “puramente informativos o centrados en simples objetos o exhibiciones, lo que permite pocas oportunidades significativas de interacción, expresión creativa o involucramiento crítico” (p. 337).

Contenido nuevo generado a partir de elementos del archivo: el experimento de cadenas de Markov¹¹

Una de las características del archivo dinámico es la apertura, no solo en sentido metafórico espacial, sino en un sentido del archivo mismo, de sus propios contenidos abriéndose hacia contenidos nuevos, generados a partir de lo ya existente.

El segundo experimento, que busca probar esa potencia generativa, usa como archivo la colección de poesía *Un Libro por Centavos*, un proyecto de la Decanatura Cultural de la Universidad Externado de Colombia. La colección pretende servir como medio de divulgación en Colombia de poesía nacional e internacional, y circula libremente en internet y en publicaciones físicas que se obsequian con los números de la revista *El Malpensante*. Para nuestro experimento, creamos un corpus con el texto crudo de treinta de los libros publicados en la colección y utilizamos *Aventura.js* para el procesamiento estadístico de cada uno de ellos. Como resultado,

10 Dos ejemplos de bots culturales en América Latina son el Bot Ulises Carrión, por Élika Ortega (@UC_Poesias_Bot), que genera poemas basados en la obra de Ulises Carrión, y el Hostos Bot, por Leonardo Flores (@HostosBot), que genera textos basados en la obra de Eugenio María de Hostos.

11 Véase <https://twitter.com/OyeArchiBot>

obtuvimos modelos de probabilidades de los n-gramas (en este caso, $n=1$) de los libros que sirvieron como insumo para crear generadores de texto personalizados. Con estos generadores y por medio del sistema definido en *Aventura.js*, es posible usar una semilla, es decir, una palabra o secuencia de palabras iniciales, para desencadenar un texto que se deriva de las probabilidades del texto original. El texto generado puede verse como una remezcla de fragmentos del texto original que puede detonar lecturas novedosas o revelar cualidades estilísticas de la obra.

Con los modelos se crea una interfaz en la que la audiencia puede interactuar con la cuenta de X @OyeArchiBot, proponer una semilla y solicitar un texto generado mediante el sistema de cadenas de Markov. En concreto, los usuarios pueden mandar un trino dirigido a @OyeArchiBot que contenga al menos la palabra “poema” y la semilla entrecomillada con la que se generará el texto¹². La cuenta de X, usando la interfaz creada y la librería *Aventura.js*, responde al usuario con el texto nuevo y con los metadatos de la obra original. Así, a través de la interacción del usuario y la producción de una obra biteraria, se incentiva el acceso al archivo desde la plataforma X, de una forma casual y entretenida. El ejercicio puede repetirse indefinidamente y, dependiendo de la diversidad de las probabilidades, pueden producirse muchos textos diferentes con la misma semilla y referenciarse varios de los libros de la colección. El siguiente es un ejemplo de un nuevo poema generado, basado en el libro *Luz en lo alto* de Juan Felipe Robledo (2006):

12 Para consultar la imagen y el resultado de la interacción, véase <https://twitter.com/OyeArchiBot/status/1525176090276286464/photo/1>

Insecto en esta cantina sosa
que no podamos ver
el brillo del
sol, no invoques a los dedos.

Así, el poema nuevo, que imita los existentes y al tiempo es distinto, es una cocreación entre la máquina y la decisión-intención del usuario. Por lo tanto, la obra del archivo se abre hacia sí misma y hacia una continuación de sí, hacia algo nuevo y diferente.

Líneas narrativas que conectan/movilizan objetos del archivo: experimento de historia interactiva¹³

Los objetos que habitan los archivos pueden conectarse narrativamente unos con otros de formas virtualmente infinitas. Los objetos archivísticos pueden organizarse y ligarse de maneras diversas, según criterios narrativos e interacciones potenciales variables. Esta forma de dinamización del archivo propone historias interactivas basadas en la función de árboles de decisión de Aventura.js, en la que el usuario escoge entre varios caminos (opciones) para navegar el archivo, y dichos caminos hilan una línea narrativa coherente entre objetos distintos, que es variable según decisiones de uso. En este caso, un objeto archivístico no es una entidad aislada ni un fin en sí mismo, sino necesariamente un camino hacia otros objetos; y cuando se recorre el camino, se cuenta asimismo una historia.

13 Véase <https://aventurainteractiva.github.io/diarioperegrino/>

Para este experimento utilizamos el álbum de dibujos *Recuerdos de campaña*, hecho por el artista colombiano Peregrino Rivera Arce, quien además de dibujante fue comandante militar en la guerra civil colombiana conocida como de los Mil Días (1899-1902). Durante su paso por la guerra, Rivera Arce llevó este diario, donde dibujó impresiones de personajes, eventos, paisajes y anécdotas. El álbum no tiene textos más allá de algunas descripciones para los dibujos; no tiene un orden cronológico establecido con fechas, ni una ilación narrativa distinta de la secuencia de imágenes (cuarenta y cinco bocetos hechos a lápiz) que presenta en sus páginas. Aun así, el álbum de Rivera, al tiempo que es un importante relato documental de un periodo fundacional en la historia de Colombia, transmite una visión particular de la guerra que es a veces contemplativa, a veces melancólica, a veces humorística y a veces aterradora.

Recuerdos de campaña está disponible de forma digitalizada en la Biblioteca Virtual del Banco de la República de Colombia, un archivo digital cultural de acceso abierto que publica documentos diversos, en su mayoría, de autores colombianos o sobre Colombia. Allí, el álbum de Rivera es recuperable para consultar en web o para descargar en formato PDF, y se presenta como un único documento digitalizado que muestra los dibujos en la misma manera secuencial de las páginas de la libreta original¹⁴.

La historia interactiva se hizo a partir de las imágenes del álbum, que fueron tratadas como objetos de un subarchivo (los dibujos pueden entenderse como documentos independientes,

14 Para consultar el álbum, véase <https://babel.banrepcultural.org/digital/collection/p17054coll10/id/3636/>.

secuenciados en una libreta que las reúne como obra de un autor en un momento particular). El objetivo fue trazar líneas que unieran los dibujos de tres maneras: cartográfica, cronológica y textualmente, para que el usuario las pudiera seguir en una narrativa con variaciones sujetas a sus propias decisiones. El árbol de decisiones (el sistema que organiza la narrativa de la historia) se estructuró en forma de ramificaciones controladas, es decir, puntos de variabilidad que cambian los senderos de la narrativa hasta un punto, para luego retornar al tronco. Así, existe una línea narrativa principal basada en los desplazamientos históricos y geográficos que llevaron a cabo Peregrino Rivera y su compañía militar en eventos cronológicos de la guerra (figura 7.5); y, luego, en cada uno de los puntos del mapa (geográficos, donde algunos hechos históricos relevantes de la guerra tuvieron lugar) hay estaciones de decisión, en las que los usuarios pueden hacer bifurcaciones según diferentes opciones (botones), para seguir una de las líneas narrativas posibles antes de proseguir a la siguiente estación del tronco narrativo principal.

De este modo, las dos interfaces constitutivas de la historia interactiva son el mapa, donde el usuario sigue la línea temporal y geográfica del paso de Rivera por la guerra; y las estaciones, en las que se presentan las imágenes de la libreta y los textos narrativos que hacen posible la toma de decisiones por parte del usuario. Las bifurcaciones posibles en cada estación están basadas en diferentes aspectos de la visión del artista presentes en los dibujos del álbum (figura 7.6). Estas diferentes facetas se vuelven variables según las decisiones del usuario en cada una de las estaciones; así, cada línea recorrida de la historia interactiva puede mostrar una combinación distinta de los múltiples temas y tonos presentes en el álbum de Rivera. La narrativa desarrollada al final de la

Historia Interactiva – Álbum de Peregrino Rivera

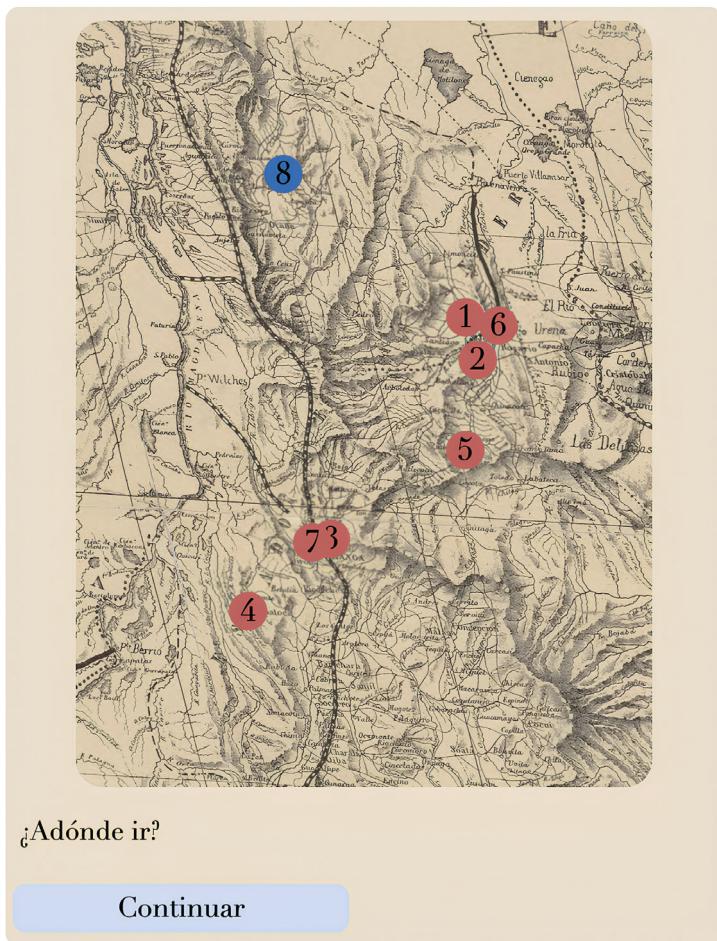
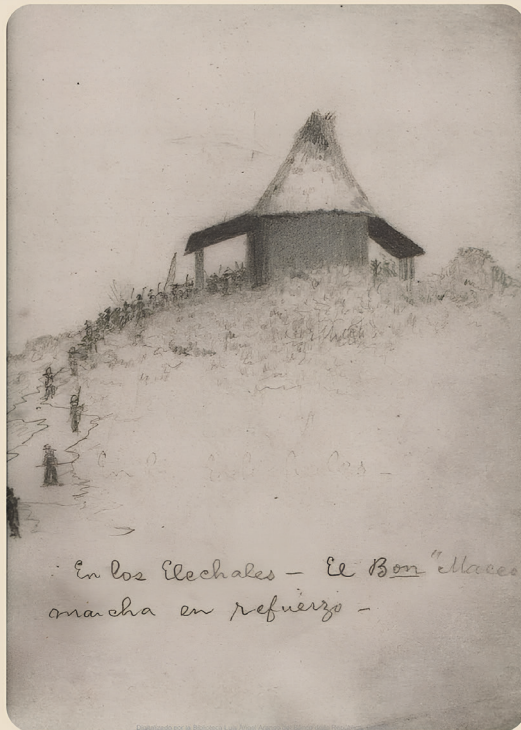


FIGURA 7.5. La interfaz inicial del interactivo creado con Aventura.js basado en el álbum de Peregrino Rivera Arce

FUENTE: elaboración propia.

Historia Interactiva – Álbum de Peregrino Rivera



¿Adónde lleva este camino?

(En los Elechales — el Bon "Maceo" marcha en refuerzo)

¿Un campo abierto?

¿Algún otro laberinto?

FIGURA 7.6. Una escena posible dentro de la navegación hipertextual de la adaptación del álbum de Peregrino Rivera Arce

FUENTE: elaboración propia.

historia termina siendo tanto lineal como variable, o variable dentro de una linealidad.

Los textos que acompañan las imágenes en las estaciones, los cuales inducen la toma de decisiones en el usuario, son una combinación entre algunos fragmentos de un poema transcrito al principio del álbum original *El lápiz de Ribera* y textos de los autores de la historia interactiva, que hacen las veces de textos curatoriales para guiar el camino de los usuarios a través de las imágenes. Los textos curatoriales pretenden, hasta cierto punto, reflejar y señalar el tono poético de las imágenes del álbum, así como provocar interpretaciones propias de los usuarios, que ocurren en el proceso de toma de decisiones. En últimas, los sentidos y significados de la historia interactiva son una combinación entre las imágenes, la interpretación curatorial que guía la historia y las decisiones que toma el usuario en cada interacción particular. La historia se puede recorrer varias veces para encontrar múltiples sentidos.

Este experimento particular se hizo en un ambiente controlado, con un número relativamente pequeño de objetos que forman parte de un mismo corpus original, pero es virtualmente escalable a un número ilimitado de objetos dentro de un mismo archivo con relaciones más o menos estrechas entre sí. El experimento es modificable a historias interactivas más controladas (con una línea narrativa más lineal, en una colección específica, por ejemplo) o más aleatorias (con múltiples líneas narrativas, con mayor número de caminos transitables, con mayor número de entrecruzamientos, en un archivo más amplio y diverso). Las posibilidades de aplicación dependen de la imaginación de los archivistas y de las posibles narrativas que quieran ofrecer a los usuarios del archivo.

Conclusiones

A lo largo de este capítulo hemos propuesto el desarrollo y la implementación de estrategias biterarias aplicadas a tres archivos digitales culturales con el propósito de dinamizarlos, es decir, introducir en ellos movimiento y apertura para que la experiencia de sus usuarios sea más cercana, más interactiva, y suscite más activamente la exploración y el descubrimiento, mediante las prestaciones de la librería Aventura.js: la generación de contenido con gramáticas libres de contexto y cadenas de Markov, y las historias interactivas a partir de árboles de decisión hipertextual.

Creemos que el desarrollo de herramientas digitales culturales fuera del centro de producción anglófono tiene impactos que reverberan más allá de sus campos de acción inmediatos, y que se extienden a la producción de conocimiento y lo incentivan en áreas aledañas, en efectos de cadena. Si la producción de narrativas digitales en español, portugués y otras lenguas distintas al inglés está limitada por la falta de herramientas para su creación que se adecuen a sus particularidades lingüísticas, culturales y de audiencias, enmarcadas en el problema de la *diglosia cultural* (Takhteyev, 2012), entonces el desarrollo de herramientas digitales en esos idiomas generará una proliferación de productos narrativos digitales. A mayor número de narrativas digitales, habrá mayor interés por desarrollar herramientas nuevas y más complejas.

En el caso de la dinamización de archivos culturales latinoamericanos, las herramientas digitales tienen el potencial de acercar a los públicos, poner en circulación los contenidos y tender puentes entre los archivos mismos: volverlos menos como islas de cofres que encierran tesoros y más como redes entretejidas en las

que las audiencias puedan transitar libremente. Para los archivos culturales digitales, la democratización pasa necesariamente por procesos de inclusión y pluralización; esto aplica también para las herramientas y conocimientos digitales que allí se usan. Abrir las cajas negras de lo digital requiere abrir los códigos, compartir la información y colectivizar los procesos, también en nuestros idiomas y culturas.

En nuestro contexto, en el que los recursos, la formación y el desarrollo de *software* tienen más limitaciones que en los lugares hegemónicos, esta simplicidad y recursividad es una estrategia fundamental. Aquí hay que destacar formas de dinamización del archivo no directamente algorítmicas que ejecutan algunas instituciones culturales, de formas creativas y con excelentes resultados, en espacios no intrarchivísticos, como las redes sociales¹⁵. Esta inventiva debe estar en el centro de la incorporación de nuevos métodos digitales y se tiene que sumar a los procesos físicos y analógicos, para que dinamizar el archivo sea un proceso híbrido e integral entre lo digital y lo no digital, entre lo algorítmico y lo humano.

La ejecución de los tres experimentos, para poder cumplir de la mejor manera posible con los objetivos establecidos de dinamización del archivo, exigió algunos procesos técnicos periféricos a la creación del contenido dentro de la librería Aventura.js. El desarrollo de las interfaces visuales, la creación del mapa en la

15 Un ejemplo claro de estas formas de dinamización es el proyecto de Stories del Bicentenario del Banco de la República de Colombia. Allí se usan diferentes estrategias de comunicación a través de redes sociales para enganchar a un público joven a documentos y acontecimientos históricos. Véase <https://www.banrepcultural.org/proyectos/bicentenario-de-una-nacion-en-el-mundo/coleccion-bicentenario/storiesdelbicentenario-las>

historia interactiva o el establecimiento del bot en X fueron retos adicionales que precisaron de conocimientos especializados en programación, y la resolución de problemas derivados fue de naturaleza tanto humanística como técnica. Aunque el uso exitoso de Aventura.js para propósitos esenciales de generación de contenido e historias interactivas no requiere conocimientos técnicos especializados —basta con seguir las instrucciones proporcionadas por la librería—, a menudo los procesos adicionales de publicación y circulación de contenidos sí lo hacen, con niveles variables de complejidad.

Estos retos, más allá de una barrera para los humanistas que quieran embarcarse en proyectos digitales, deben ser un incentivo para empaparse cada vez más, y en la medida de las posibilidades de cada quien, en los aspectos técnicos y tecnológicos de las humanidades digitales. Algo que nos han enseñado las experiencias de desarrollo de proyectos es que separar la ejecución entre “roles técnicos” y “roles humanísticos” nunca produce los mejores resultados, pues los proyectos digitales requieren por fuerza que las dos cosas estén relacionadas estrechamente. Así, los humanistas digitales debemos hacernos cada vez más proficientes en herramientas y procesos tecnológicos; así mismo, debemos involucrar a los especialistas tecnológicos para que sus herramientas y procesos tengan una naturaleza cada vez más humanista en su desarrollo.

Los tres experimentos presentados en este capítulo para dinamizar archivos digitales culturales cumplen sus respectivos propósitos iniciales de movilizar y abrir el archivo y sus objetos. Sin embargo, funcionar según el diseño original y funcionar en las manos de los usuarios son dos procesos distintos y complementarios. Además de los propósitos de desarrollar y analizar los

experimentos, expuestos en este capítulo, cada proyecto digital necesita de pruebas con usuarios reales y de adaptaciones según sus necesidades y observaciones. Para cada proyecto de humanidades digitales, es necesario medir su impacto en escenarios de uso real, fuera de los escenarios iniciales de control, e involucrar en los procesos a otras personas (los verdaderos públicos objetivo) que no sean únicamente quienes desarrollamos, editamos y archivamos.



Referencias

- Álvarez Umbarila, J. (2017). *From chest to window: The literary digital archive and its mediations*. University of Groningen.
- Balpe, J. P. (2005). *Principles and processes of generative literature*. <https://doi.org/10.1515/9783839404935-013>
- Brown, T. B., Mann, B., Ryder, N., Subbiah, M., Kaplan, J., Dhariwal, P., Neelakantan, A., Shyam, P., Sastry, G., Askell, A., Agarwal, S., Herbert-Voss, A., Krueger, G., Henighan, T., Child, R., Ramesh, A., Ziegler, D. M., Wu, J., Winter, C. y Amodei, D. (2020). *Language models are few-shot learners*. <https://doi.org/10.48550/ARXIV.2005.14165>
- Burdick *et al.* (2012). *Digital_humanities*. Massachusetts Institute of Technology Press.
- Chomsky, N. (1956). Three models for the description of language. *IEEE Transactions on Information Theory*, 2(3), 113-124. <https://doi.org/10.1109/TIT.1956.1056813>

- Christian, P. (2007). *The N+7 Machine*. <http://www.spoonbill.org/n+7/>
- Compton, K., Kybartas, B. y Mateas, M. (2015). Tracery: An author-focused generative text tool. En H. Schoenau-Fog, L. E. Bruni, S. Louchart y S. Baceviciute (Eds.), *Interactive storytelling* (pp. 154-161). Springer. https://doi.org/10.1007/978-3-319-27036-4_14
- Derrida, J. (1995). Archive fever: A Freudian impression. *Diacritics*, 25(2), 9. <https://doi.org/10.2307/465144>
- Drucker, J. (2013). Is there a “digital” art history? *Visual Resources*, 29(1-2), 5-13. <https://doi.org/10.1080/01973762.2013.761106>
- Eco, U. (2011). *La estructura ausente*. Debolsillo.
- Ensslin, A. (2007). *Canonizing hypertext: Explorations and constructions*. Continuum.
- Ernst, W. (2013). *Digital memory and the archive*. University of Minnesota Press. <http://site.ebrary.com/id/10716979>
- Evens, A. (2020). Combination and copulation: Making lots of little poems. En J. Tabbi (Ed.), *The Bloomsbury handbook of electronic literature* (pp. 217-236). Bloomsbury Academic.
- Fiormonte, D. (2018). Lenguas, códigos, representación: Márgenes de las humanidades digitales. En I. Galina Russell, M. Peña Pimentel, E. Priani Saisó, J. F. Barrón Tovar, D. Domínguez Herbón y A. Álvarez Sánchez (Eds.), *Humanidades digitales* (pp. 39-81). Bonilla Artigas Editores y Red de Humanidades Digitales.
- Flores, L. (2017). La literatura electrónica latinoamericana, caribeña y global: Generaciones, fases y tradiciones. *Artelogie*, 11. <https://doi.org/10.4000/artelogie.1590>
- Flores, L. (2021). Chapter 2. Third-generation electronic literature. En J. O’Sullivan (Ed.), *Electronic literature as digital humanities:*

- Contexts, forms, & practices*. Bloomsbury Academic. <https://doi.org/10.5040/9781501363474>
- Gainza, C. y Zúñiga, C. (2018). *Cartografía de la literatura digital latinoamericana*. <https://archivocartografia.cl/>
- Holguín, C. y Joni B. (2017). *Muerte en el bosque*. Mákina Editorial. <https://muerteenelbosque.com/>
- Howe, D. C. (2009). RiTa: Creativity support for computational literature. En *C&C '09: Proceedings of the seventh ACM conference on Creativity and cognition*. Association for Computing Machinery. <https://dl.acm.org/doi/10.1145/1640233.1640265>
- Läufer, M. (2018). *A noise such as a man might make*. Counterpath.
- Lem, S. (2017). *Magnitud imaginaria*. Impedimenta.
- Luman, D. (2016). *Applied poetics*. <http://www.appliedpoetics.org>
- Massey, D. B. (2005). *For space*. Sage.
- Mateas, M. (2008). Procedural literacy: Educating the new media practitioner. En D. Davidson (Ed.), *Beyond fun: Serious games and media* (pp. 67-83). ETC Press.
- Montenegro, A. M. (2017). *Promesas*. <https://promesaspromesas.com/>
- Observatório da Literatura Digital Brasileira. (2018). *Atlas da literatura digital brasileira*. <https://www.observatoriodigital.ufscar.br/>
- Ojamaa, M. e Ibrus, I. (2021). Hotspots for textual dynamics: Cultural semiotic approach to digital archives. *Semiotica*, 2021(243), 387-407. <https://doi.org/10.1515/sem-2019-0001>
- Robledo, J. F. (2006). *Luz en lo alto*. Universidad Externado de Colombia.
- Schnapp, J. (2008). Animating the archive. *First Monday*, 13(8). <https://doi.org/10.5210/fm.v13i8.2218>

- Star, S. L. (1999). The ethnography of infrastructure. *American Behavioral Scientist*, 43(3), 377-391. <https://doi.org/10.1177/00027649921955326>
- Swirski, P. (2014). *From literature to biterature: Lem, turing, darwin, and explorations in computer literature, philosophy of mind, and cultural evolution*. McGill-Queen's University Press.
- Szántó, A. (2018). *Bookmarks: Revisiting the hungarian art of the 60s and 70s*. Koenig Books.
- Takhteyev, Y. (2012). *Coding places: Software practice in a South American city*. Massachusetts Institute of Technology Press.
- Tzouganatou, A. (2018). Can heritage bots thrive? Toward future engagement in cultural heritage. *Advances in Archaeological Practice*, 6(4), 377-383. <https://doi.org/10.1017/aap.2018.32>
- Tzouganatou, A. (2021). Openness and privacy in born-digital archives: Reflecting the role of AI development. *AI & SOCIETY*. <https://doi.org/10.1007/s00146-021-01361-3>
- Uribe, A. M. (1995). *Tipoemas y anipoemas*. <https://www.vispo.com/uribe/anipoemas.html>

A politização do consumo mediada pela tecnologia: o caso do *Sleeping Giants* no Brasil*

* Agradeço o apoio do Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), e da Fundação de Amparo à Pesquisa do Estado de Minas Gerais (FAPEMIG). Agradeço os pertinentes e atenciosos comentários de Erika Fernanda Pereira de Souza.

Para citar este capítulo:

<http://doi.org/10.51573/Andes.9789587987362.9789587987379.9789587987867.8>

DANIEL COELHO DE OLIVEIRA

Introdução

{ Em meio às denúncias de propagação de *fake news*¹ relacionadas à pandemia de covid-19, chega ao Brasil o projeto *Sleeping Giants*², que se autodenomina “uma luta coletiva de cidadãos contra o financiamento do discurso de ódio e *fake news*”. O presente capítulo

- 1 Em artigo recente, Recuero e Gruzd (2019) destacam que o conceito de *fake news* é sinônimo de desinformação, utilizado livremente pelos veículos noticiosos para indicar rumores e notícias falsas que circulam, principalmente, na mídia social. Shu et al. (2017) apontam duas características-chave que definem as *fake news*. A primeira, a falta de autenticidade; a segunda, são construídas com o propósito de enganar. Para os autores, sátiras não seriam *fake news*, pois revelam sua intenção no próprio discurso e formato. Não se pode confundir esse tipo de notícia com a mensagem “satírica” que apresenta a sua falsidade por meio do humor, ou seja, não possui o propósito de enganar. Os “boatos e rumores” são diferenciados deste conceito, já que faltam neles o desejo de autenticidade.

- 2 Disponível em: https://twitter.com/slpng_giants_pt

procura entender como a grande repercussão das primeiras ações do *Sleeping Giants* Brasil, denominado aqui de SG-BR, se relaciona com um conjunto de conflitos e de enfrentamentos no ambiente virtual no Brasil. De forma especial, nos últimos 4 anos, o Brasil vivenciou um período de intensa polarização política nas redes sociais: de um lado, os apoiadores do ex-presidente Jair Bolsonaro, alinhados ao campo da direita conservadora; de outro lado, atores do campo político da esquerda progressista.

A proposta do perfil no Twitter³ perpassa pela constante vigiância e denúncia pública de empresas ou organizações que não seguem determinados “princípios éticos” para a divulgação de suas marcas em *sites*. As ações que envolvem o SG-BR são uma clara tentativa de politizar as práticas de consumo a partir da ação nas redes sociais. As práticas de consumo tornaram-se mais complexas com a intensificação das relações sociais mediadas pela tecnologia e pelo processo de “plataformização” da vida cotidiana. Em outros termos, no período atual, o consumo ganhou contornos singulares no ambiente *on-line*⁴. Diante desse panorama

- 3 O Twitter (que mudou de nome para “X” em julho de 2023) foi criado em 2006, por Jack Dorsey, Evan Williams e Biz Stone. Inicialmente, como foi idealizado para ser um programa de *microblogging*, permitia postagens de até 140 caracteres; atualmente, as mensagens podem conter até 280 caracteres. De acordo com Parreiras (2015), ele foi pensado para ser uma plataforma de troca de informações, arquitetada por meio das postagens entre seus usuários. A dinamicidade é uma das suas características básicas, na qual há conexão global de pessoas que produzem e consomem informações, dos mais diferentes tipos, em tempo real.
- 4 O termo “*on-line*” (ou “digital”) contrasta com estratégias investigativas realizadas por meio do contato face a face, no ambiente “*off-line*” ou “analógico”. (Miller e Slater, 2004).

ma, o presente texto apresenta duas perguntas (ou problemas):
 1) Quais são os limites e possibilidades do ativismo dos consumidores nas redes sociais *on-line*? 2) As redes sociais no ambiente *on-line*, de forma especial o Twitter, podem aproximar as esferas de ação pública (cidadão) e particular (consumidor)?

Como objetivo principal, pretende-se contextualizar, caracterizar e descrever o ativismo do *Sleeping Giants* Brasil (SG-BR), assim como suas interlocuções com perfis do mundo político e econômico na rede social Twitter. De forma específica, busca-se entender como o engajamento de consumidores no ambiente *on-line* foi afetado pelo processo de polarização política no Brasil.

As pesquisas sobre a politização do consumo nas redes sociais são uma importante área de estudos das Humanidades Digitais. Do ponto de vista metodológico, a presente pesquisa sobre politização do consumo nas redes sociais envolve uma série de desafios. Por um lado, foi necessário coletar e analisar grandes volumes de dados publicados no Twitter, o que exigiu a utilização de técnicas de mineração e análise de dados. Por outro lado, foi necessário considerar aspectos qualitativos e analisar os discursos dos atores envolvidos na polarização das redes do *Sleeping Giants*, ou seja, busca-se compreender de forma aprofundada os rastros digitais dos principais atores envolvidos.

Este capítulo será dividido da seguinte forma: no primeiro momento, será apresentada a gênese do perfil do Twitter SG-BR; logo após, a discussão teórica acerca do debate sobre “politização do consumo”, “consumo político” e “plataformização do ativismo consumerista”. Na seção seguinte, serão descritos os procedimentos metodológicos e, posteriormente serão apresentados os “resultados e as discussões”. Por fim, serão realizadas algumas

considerações finais, com ênfase nos limites e nas possibilidades do ativismo liderado pelo perfil do SG-BR.

A gênese do *Sleeping Giants* Brasil

Em seu perfil oficial, o SG-BR destaca que seu papel é o de informar empresas e organizações sobre a vinculação em *sites* que disseminam desinformação. A estratégia de comunicação empregada não defende o boicote, pois não incentiva os consumidores a deixarem de comprar bens ou serviços vinculados à determinadas marcas. A estratégia seria “convencer” a não expor sua publicidade em determinados *sites* para, conseqüentemente, não financiar espaços que estariam divulgando *fake news*.

Em análise sobre o tema, Santini (2022) destaca que a estratégia de atuação do movimento *Sleeping Giants* passa por quatro etapas, sendo elas:

- 1) monitoramento e acompanhamento de *sites* que disseminam conteúdo falso e que foram checados; 2) seleção dos *sites* com maior audiência; 3) convocação da colaboração dos usuários; 4) convidar publicamente os anunciantes associados pelos usuários ao conteúdo falso dos *sites* a se manifestarem publicamente. (Santini, 2022, p. 4).

O SG-BR foi inspirado no perfil americano *Sleeping Giants*. A reportagem de João Filho (2020), do *site The Intercept Brasil*, aponta que o projeto *Sleeping Giants* provocou grande prejuízo à estratégia da direita americana. O maior exemplo foi o *site* de

extrema direita, *Breitbart News*, famoso por inventar histórias contra adversários de Trump⁵. Após a ação do *Sleeping Giants*, o site perdeu mais de 4500 anunciantes – um golpe que significou uma perda de mais de 8 milhões de euros. Steve Bannon, o guru da extrema direita internacional, era o proprietário do site e, à época, chamou o *Sleeping Giants* de “a pior coisa que há”. Após investigar o caso, Quimaliñnos, no *El País* de 17 de maio de 2020, destaca que o *Sleeping Giants* deixou quase sem publicidade a mídia ultraconservadora americana. “O site de extrema direita *Breitbart News* perdeu mais de oito milhões de euros (50,7 milhões de reais) em publicidade por culpa de uma conta no Twitter.” (Quimaliñnos, 2020). O perfil do *Sleeping Giants* foi uma criação do publicitário Matt Rivitz, e ele acreditava que o modelo de propaganda AdSense/Google⁶ provocaria danos à reputação das marcas, já que elas não possuem total controle sobre o local onde os anúncios são vinculados. Para Quimaliñnos, o “prejuízo financeiro significativo” à “mídia ultraconservadora nos EUA” foi somente uma das conquistas do projeto *Sleeping Giants*, já que a conquista mais importante teria acontecido na França “onde conseguiu que fosse aprovada uma regulamentação governamental – a Emenda *Sleeping Giants* – para evitar que anunciantes financiem o ódio e o extremismo *on-line*” Quimaliñnos (2020).

- 5 Durante a eleição presidencial, o site publicou a história de que a então candidata Hillary Clinton comandava uma rede de pedofilia e promovia orgias sexuais com crianças no porão de uma pizzaria.
- 6 O AdSense/Google é o serviço de publicidade promovido pelo Google voltado para os donos de *websites* exibirem anúncios publicitários. A exibição das propagandas remunera o *website* de acordo com a quantidade de cliques ou de visualizações.

No Brasil, o perfil do SG-BR disparou seu primeiro tuíte em 18 de maio de 2020. O objetivo defendido pela organização seria a crítica “ética” direcionada a empresas que financiam o “discurso de ódio” e as *fake news* nas redes sociais. É simbólico o fato de o primeiro texto publicado ter sido retirado do perfil @B4STA_, conta que se autodenomina “a segunda conta pedindo o bloqueio dos maiores anunciantes do Twitter pelo banimento de Olavo de Carvalho, Silas Malafaia e Allan dos Santos. #VaiDoerNoBolsoTwitter”.

Sleeping Giants Brasil. [@slpng_giants_pt] (18 de maio de 2020) / Nosso primeiro tweet é uma inspiração e por isso vai integralmente em rt! B.A.S.T.A. – 5 de abr “A gente vai começar esse perfil contando uma história. Vocês devem se lembrar de um cara chamado Alex Jones, dono do infame infowars, especializado em fakanews e que teve a pachorra de dizer que o massacre de Sandy Hook não aconteceu, só para defender a indústria das armas. [Tweet] Twitter. https://twitter.com/slpng_giants_pt/status/1262348397765427201

Um segundo tuíte, publicado no mesmo dia, delimitou os “alvos” da conta SG-BR: “Galera só um adendo: são *sites* oficiais com propaganda falsa. São *sites* extremistas ou de *fake news* que contenham propaganda de alguma marca real” [Tweet] Twitter https://twitter.com/slpng_giants_pt/status/1262381707313008642. O perfil enfatiza a importância do engajamento na proposta defendida. Trechos como “Essa resposta depende de vocês”, “Lembrando que buscamos a conscientização, e não o boicote!” e “Contamos com ajuda de vocês para que *twittem prints* marcando o @slpng_giants_pt” demonstram que a premissa de sucesso do projeto tem

relação direta com a repercussão e a visibilidade das denúncias no espaço virtual. Vale destacar que, durante vários dias, a *hashtag* “@slpng_giants_pt” foi um dos assuntos mais comentados do Twitter no Brasil.

Sleeping Giants Brasil. [@slpng_giants_pt] (19 de maio de 2020) / Essa resposta depende de vocês, é a hora de vocês retuitarem e favoritarem o tweet, além de cobrar constantemente a empresa através da marcação do seu @, mandando mensagem por dm, e etc. Os envios de vocês são verificados, e após essa etapa então publicaremos o tweet buscando uma resposta da empresa (que geralmente não sabe da vinculação dos seus anúncios nesses sites). Lembrando que buscamos a conscientização e não o boicote! [Tweet] Twitter. https://twitter.com/slpng_giants_pt/status/1262899805732438016

Após um período inicial de anonimato, os fundadores do SG-BR revelaram suas identidades em uma entrevista para o jornal *Folha de S. Paulo*⁷. O perfil foi criado por dois jovens estudantes de Direito: Mayara Stelle e Leonardo de Carvalho Leal, residentes em Ponta Grossa, Paraná, Brasil. Em dezembro de 2020, ambos tinham 22 anos e estavam desempregados devido à pandemia de covid-19. Na entrevista, Mayara, destacou os objetivos do SG-BR:

7 Reportagem disponível em: <https://www1.folha.uol.com.br/colunas/monica-bergamo/2020/12/revelacao-de-fundadores-do-sleeping-giants-causa-negacionismo-e-ameacas.shtml>.

O SG não promove a cultura do cancelamento, o SG não promove a censura. Estamos aqui para alertar a empresa. A decisão de manter um anúncio ou retirar, e de banir um usuário por violação dos termos de uso ou não banir, pertence à empresa. Nós só avisamos (Bergamo, 2024).

Em um dos posicionamentos, Leonardo, argumenta que o perfil não segue nenhuma orientação política:

A gente não olha se o *site* é de esquerda ou de direita. Infelizmente, é sabido que o discurso de ódio está presente em todos os espectros políticos. O que acontece é que, nesse momento, não dá para negar que a extrema direita tem um alcance de desinformação muito maior que a esquerda (Bergamo, 2024).

Os fundadores reforçaram na entrevista que, mesmo no anonimato, sofreram ameaças de morte de grupos de extrema direita. A perspectiva de construção de diálogo com o setor empresarial também foi enfatizada em análise sobre a atuação do perfil *Sleeping Giants* Brasil realizada por Freitas (2021). Ela apresenta alguns aspectos sobre as especificidades do perfil brasileiro, entre elas uma certa interlocução amigável e bem humorada que o perfil busca construir com a empresa anunciante na primeira publicação dirigida a ela. De acordo com a autora, destaca-se

o chamamento de seu nome de usuário na plataforma, como é habitual em qualquer tipo de conversa. Esse nome aparece marcado na postagem, ou seja, precedido do símbolo @, que o converte em *link* para seu perfil. Além desse tom simpático,

as publicações que se dirigem às empresas pela primeira vez trazem também em seu texto elementos discursivos que fazem parte da publicidade veiculada pela própria empresa, como uma referência ao seu universo particular, de modo a estabelecer, de saída, uma aproximação. (Freitas, 2021, pp. 184-185).

Deve-se destacar que, em 19 de maio de 2020, ocorreu o reconhecimento público da conta SG-BR pelo perfil SG dos Estados Unidos. Ribeiro, Ribeiro, Almeida e Meira Jr. (2022) destacam que o *Sleeping Giants* é um movimento ativista internacional que concentra sua atuação na rede social Twitter, e busca combater a “desinformação *on-line*” utilizando estratégias de denúncia pública. Ao analisar o caso estadunidense, Braun, Coakley e West (2019) apontam que o movimento busca convencer empresas a remover seus anúncios em páginas da *web* que propagariam *fake news* e, através do bloqueio de publicidades, as páginas seriam desmonetizadas.

Fundamentação teórica: a plataformização do ativismo consumerista

Ao discutir as especificidades da “sociedade informacional”, Castells (2018) defende que a “democracia liberal” entrou em crise, assim como boa parte das instituições políticas que a integram. De um lado, a atual crise da sociedade informacional teria aberto caminho para novas categorias de agenciamentos políticos e de movimentos sociais orientados por identidades de projetos. Por

outro, a estrutura de “rede” construída pela revolução nas tecnologias, principalmente dos meios de comunicação, possibilitou o surgimento de arenas não institucionais de debates políticos.

Além da possibilidade de estabelecer e manter vínculos identitários, parte das práticas de consumo podem ser ressignificadas como formas de participação e de ação política. O conceito de “politização do consumo” parte do princípio de que o ato de consumir é uma prática social dotada de significados privados e públicos. Na visão de Portilho e Barbosa (2016), seria uma maneira de participar de um campo de disputas sobre os modos como a sociedade a produz, assim como sobre os significados e as formas de uso de tal produção. Tal contexto faz com que o consumo expresse várias disputas políticas.

Em abordagem clássica sobre o tema, Canclini (1999) argumenta que transformações na maneira de consumir mudaram as formas e as possibilidades de se exercer a cidadania, pois “junto com a desagregação da política e a descrença em suas instituições, outros modos de participação se fortalecem.” (Canclini, 1999, p. 37). Autores como Hirschman (1983) reforçam a necessidade de se entender o comportamento do indivíduo no eixo público-privado. Há também um conjunto de outros autores, tais como Trentmann (2006), Stolle, Hooghe e Micheletti (2005) e Nestle (2002), que alertam para a tendência crescente de inserção de componentes morais e éticos na ação dos consumidores. Ou seja, analisar as ações do SG-BR é uma possibilidade para entender novos cenários de ativismo e engajamentos *on-line*. Todo o debate sobre o processo de “politização” do espaço de consumo reforça, de algum modo, a tese do deslocamento da esfera de “ação” política. O desgaste

e a falta de confiança nas instituições políticas são razões para a construção de outros caminhos de ação e cidadania.

O “consumo político” pode se manifestar de quatro formas distintas, de acordo com Portilho (2020). Primeiro, o boicote, considerado “consumo político negativo”, é a recusa da compra fundamentada em preocupações éticas, políticas e/ou ambientais. Em segundo lugar, o *buycott* é uma forma de “consumo político positivo”, por referir-se a compras efetivas como maneira de beneficiar ou recompensar produtos, marcas e empresas por seus compromissos políticos, éticos e/ou ambientais. A terceira concerne às ações comunicativas e discursivas, que vão desde críticas pontuais a manifestações públicas (incluindo “falar bem” ou “falar mal”), fundamentadas nas mesmas razões éticas. Já a quarta forma de consumo político envolve mudanças nas práticas cotidianas e são as chamadas políticas de estilo de vida. De alguma forma, essa última estratégia envolve as três categorias de ação anteriores, pois seriam “subpolíticas” propostas por consumidores.

Em sua leitura sobre o fenômeno, Beck (1997) destaca que vivenciamos o crescimento das “subpolíticas”, política direta, à margem e além dos Estados-Nação; esfera não-institucional do político. Por meio da “modernização reflexiva”, surgem novas formas de fazer política, com destaque para atores externos ao sistema político.

Em resumo, um mundo está adquirindo vida, e uma parte dele não pode ser representada na outra: um mundo caótico de conflitos, jogos de poder, instrumentos e arenas que pertencem a duas épocas diferentes, aquela do “não ambíguo” e aquela da

modernidade ‘ambivalente’. Por um lado, está se desenvolvendo um vazio político das instituições; por outro, um renascimento não institucional do político. O sujeito individual retorna às instituições da sociedade. (Beck, 1997, p. 28).

Observam-se em curso batalhas no imaginário de grupos sociais; em outras palavras, pelo controle e pelo direcionamento dos processos de informação e que moldam a comunicação nas redes sociais. Na visão de Alexander (1998), a sociedade pós-materialista, baseada na informação, produz arranjos estruturais que criam formas de estratificação, novos grupos de conflito e padrões de dominação. Segundo Portilho e Barbosa, “nas sociedades contemporâneas, o consumo vem-se tornando, cada vez mais, uma esfera de demandas, reivindicações e lutas políticas.” (Portilho e Barbosa, 2016, p. 256). De maneira geral, as campanhas caracterizadas como “politização do consumo” convidam os adeptos a “votar com o garfo”, conforme Nestle (2002). De forma análoga, Halkier (1999) fala em “voto com a carteira”. Talvez estejam em curso ações como o SG-BR, que fazem um convite para o “voto com uma curtida”, ou “voto com um retuíte”, cenário em que o ativismo *on-line* depende do grau de engajamento e de exposição da campanha.

Se a “politização do consumo” é um “processo” mais amplo, o “consumerismo político”, de forma mais específica, caracteriza-se por um conjunto de “ações” ou “práticas” conduzidas por sujeitos ou grupos engajados em determinadas pautas. De acordo com Hilton (2003), o consumerismo político pode ser compreendido como um fenômeno histórico no qual os indivíduos ou grupos acionam suas relações de compras ou trocas nos mercados com

estratégias para influenciar arenas econômicas ou políticas. Nessa mesma linha, Sassatelli (2006) argumenta que o consumerismo se define como um fenômeno que materializa visões morais a partir de posicionamentos críticos em relação às práticas de consumo. Ou seja, o consumerismo político traduz as ações segundo as quais indivíduos fazem escolhas de serviços e produtos. Ao argumentar sobre o consumerismo político, Micheletti destaca:

Isso representa ações de pessoas que fazem escolhas entre produtores e produtos com o objetivo de mudar instituições ou práticas de mercado. Suas escolhas são baseadas em atitudes e valores em relação a questões de justiça, equidade ou questões não econômicas que dizem respeito ao bem-estar pessoal e familiar e à avaliação ética ou política de práticas comerciais e governamentais favoráveis e desfavoráveis. Os consumidores políticos são as pessoas que se envolvem em tais situações de escolha. (Micheletti, 2003, p. 2; tradução própria)

As ações de consumidores no ambiente *on-line* foram evidenciadas nos últimos anos. Essas atuações ganharam várias denominações, tais como: “netativismo de consumidor”, conforme Di Felice (2011), Schwartz (1996) e Borges (2017); ou “consumerismo de plataforma”, de acordo com Freitas (2021). O fenômeno também recebeu críticas, sendo que Borges (2017), em sua análise sobre “netativismo contra as marcas nas redes sociais”, destaca que o netativismo é criticado sob a argumentação de que ele se trata de manifestação da esfera privada, conduzida por poucos atores; ter-se-ia, assim, uma análise limitada das ações na esfera pública. Para que se deixe clara a especificidade das ações consumeristas

no ambiente *on-line*, faz-se necessário delimitar teoricamente os espaços compreendidos como plataformas.

Em sua análise sobre a plataformização da *web*, Helmond (2019, p. 49) procura situar a “plataformização”, ou o “advento da plataforma”, como o modelo dominante de infraestrutura e economia da *web*. Nas palavras de Helmond, “[o] termo ‘plataformização’ é utilizado para se referir à emergência da plataforma como modelo econômico e infraestrutural dominante da *web* social, bem como às consequências da expansão das plataformas de mídias sociais em outros espaços *on-line*.” (Helmond, 2019, p. 61). Para a autora, o conceito de plataformização se tornou dominante não só para o posicionamento de empresas de mídias sociais no mercado, mas também para a forma como elas se dirigem aos usuários.

Por sua vez, Van Dijk, Poell e De Waal (2018) defendem a visão abrangente de um mundo conectivo onde as plataformas penetraram no coração das sociedades; elas afetam instituições, transações econômicas, práticas sociais e culturais. Nesse cenário, governos e Estados são obrigados a ajustar suas estruturas legais e democráticas. Os autores defendem a utilização do termo “*platform society*”, por enfatizar as conexões entre plataformas *on-line* e estruturas sociais. Ou seja, as plataformas não refletem o social, visto que seriam responsáveis por produzir novos arranjos sociais na sociedade contemporânea. Assim, a implementação de plataformas na sociedade desencadeia um debate sobre benefício privado e ganho corporativo em relação aos interesses públicos. Em outras palavras, segundo os autores:

plataforma” *on-line* é uma arquitetura digital programável projetada para organizar interações entre usuários – não

apenas usuários finais, mas também entidades corporativas e órgãos públicos. Ela é voltada para a coleta sistemática, processamento algorítmico, circulação e monetização de usuários dados. (Van Dijk, Poell e De Waal, p. 3, 2018; tradução própria)

Nessa mesma linha, D’Andréa (2020) destaca que as plataformas não são usadas apenas como meios de intermediação na sociedade, visto que elas se fazem visíveis nas interações sociais, criando um ambiente que condiciona a emergência de um social específico, marcado por assimetrias e articulações multidirecionais. Desse modo, “plataformas se apropriam das lógicas de conexão e as potencializam como parte de uma estratégia – comercial sobretudo – que visa incentivar usuários a deixar rastros de suas relações, preferências, etc.” (D’Andréa, 2020. P. 18). O termo “rede social”, assim, precisa de ser usado de uma forma que evidencie sua dimensão relacional e seus aspectos econômicos, materiais e políticos. Para Gillespie (2010), o termo “plataforma” foi implantado tanto em seus apelos populistas quanto em seus discursos de *marketing*, às vezes como “plataformas” técnicas, às vezes como “plataformas” de oportunidade. O conceito também se encaixa em seus esforços para moldar a política de informação, na qual se busca a proteção para facilitar a expressão do usuário, mas também a responsabilidade limitada pelo que esses usuários dizem.

Após apresentar alguns marcos conceituais do debate sobre politização do consumo, consumerismo e plataformização da sociedade, a proposta apresenta, neste capítulo, a existência de um processo de “plataformização do ativismo consumerista” através da atuação do SG-BR no Twitter. Algumas características da atuação justificam a atualização do conceito. Além do ativismo de

consumidores em relação às empresas, há a presença de outros personagens do campo político partidário. A estratégia de exposição se concentra na tentativa de sensibilização das empresas anunciantes via AdSense/Google e na consequente desmonetização de *sites* vistos como propagadores de *fake news*.

Procedimentos metodológicos

O debate teórico e metodológico sobre as redes não é novo, sendo um dos marcos iniciais as observações de pequenos grupos desenvolvidas por Moreno (1978) e suas descrições do comportamento humano através de representações gráficas, sociogramas compostos por vértices e linhas. Barnes (1954), por sua vez, utiliza o conceito de “rede social” em seu estudo sobre uma organização social na Noruega. Para o autor, as relações sociais podem ser observadas no formato de rede; assim, ao compreender as partes, podem-se entender aspectos significativos de uma determinada comunidade. Ainda, Omena e Amaral destacam que os autores clássicos da análise de redes entendem que “as redes sociais [se] constituem como metáforas da sociabilidade ou como base para a compreensão da formação de interações individuais, do comportamento humano e de relações.” (Omena e Amaral, 2019, p. 123). Estudos recentes têm aproveitado todo o debate acumulado para analisar as “redes” no ambiente *on-line*. Nesse campo, deve-se destacar a Análise de Redes Sociais (ARS). De acordo com Garton, Haythornthwaite e Wellman (1997), a ARS é uma metodologia interdisciplinar que pretende estudar as relações de conjunto de atores, com a finalidade de entender possíveis padrões de interação.

Além da abordagem da ARS, deve-se destacar a análise de Fiscarelli (2022) sobre as redes sociais aplicada às Humanidades Digitais. O autor discute os desafios e oportunidades da análise de redes sociais para entender as dinâmicas sociais e culturais que moldam as interações *on-line*. Ele enfatiza a importância de uma abordagem interdisciplinar para a análise de redes sociais, que envolve a colaboração entre especialistas em humanidades, ciência da computação e ciências sociais.

Em seu estudo, Fiscarelli (2022) apresenta vários casos de uso da análise de redes sociais, como a rede de colaboração acadêmica, a relação de coautoria em publicações científicas e de redes de citações em publicações acadêmicas. Ele demonstra como essas abordagens podem fornecer apontamentos valiosos sobre o comportamento humano em contextos digitais e ajudar a identificar tendências e padrões relevantes. Além disso, Fiscarelli destaca a necessidade de uma reflexão crítica sobre as implicações éticas e políticas da análise de redes sociais nas Humanidades Digitais.

A pesquisa no ambiente *on-line* do Twitter desenvolvida aqui, aproxima-se da abordagem da ARS. De acordo com Recuero, Zago e Bastos (2014), a ARS pode ser entendida como uma ferramenta teórico-metodológica com foco na interpretação da estrutura social, baseada no estudo dos atores sociais (nós ou nodos) e suas interconexões. As variáveis utilizadas na pesquisa foram os dados sobre a frequência de conexão das palavras utilizadas nos tuítes e suas conexões. As variáveis de composição, por outro lado, referem-se às palavras (conceitos) que aparecem com maior frequência nos tuítes analisados. Nesse cenário, a proposta metodológica busca compreender o ativismo do SG-BR por meio das características dos principais autores e seus respectivos *clusters*, bem como

as pautas e o perfil dos léxicos utilizados. Para uma melhor exemplificação, a metodologia foi dividida em três etapas.

A primeira etapa envolve a extração das postagens do Twitter, realizada entre 20/05/2019 e 28/05/2021 (El País Brasil, 2020; Hypeness, 2020). A coleta de dados foi realizada pelo projeto Twitter Monitor (Gomes; Pimenta; Braga, 2021), sistema de extração e armazenamento de tuítes, criado pelo Laboratório em Rede de Humanidades Digitais (LARHUD/UFRRJ). O Twitter Monitor coletou mensagens contendo todas as menções a #SleepingGiants-Brasil e @slpng_giants_pt no período indicado. O total de tuítes coletados foi de 103.756 (cento e três mil e setecentos e cinquenta e seis) e de retuítes foi de 649.587 (seiscentos e quarenta e nove mil e quinhentos e oitenta e sete). Os tuítes foram extraídos e formatados em um arquivo separado por vírgulas (csv).

Na segunda etapa, foi realizada a análise e a visualização dos dados extraídos a partir do aplicativo de código aberto Gephi⁸. Por meio desse aplicativo, foi possível realizar a limpeza, a manipulação do banco de dados e construir grafos com métricas para os nós (indivíduos) ou comunidades (*clusters*) da rede. Após a rotulação dos atores da rede, o banco de dados foi analisado com dois algoritmos: ForceAtlas2⁹, para todo banco de dados de tuítes e retuítes coletados, gráfico 1; OpenOrd, para os gráficos 2 e 3, nos quais foi analisada somente a rede de atores do campo

8 Disponível em: <https://gephi.org>.

9 Após a aplicação do algoritmo ForceAtlas2, utilizou-se, como parâmetro, peso 1.0 para “Influência do peso das arestas”; 2.0 para Afiinação/Dimensionamento; 1.0 para “Gravidade”, aplicação de “Gravidade mais forte”; 1.0 para “Desempenho/tolerância (velocidade), 1.2 para “Desempenho/aproximação. Ao final, foi realizado um processo de ajuste de rótulos dos nós.

político. O objetivo foi aproximar por vizinhança, para criar comunidades ou *cluster*, ou seja, foi analisada a força de repulsão em cada nó com o objetivo de gerar uma melhor visualização. No processo de limpeza, foram retirados tuítes postados por perfis de empresas e do SG-BR. Ficaram preservados somente os perfis de políticos (senadores, deputados, vereadores, entre outros que ocupam cargos na administração pública), de ativistas e de consumidores. Após a limpeza, a nova rede foi construída com 35.244 (trinta e cinco mil duzentos e quarenta e quatro) nós. Na construção do grafo, foram filtradas as comunidades¹⁰ de atores proeminentes, conforme Piaia (2021), e selecionadas aquelas que concentraram 5% ou mais dos perfis que participaram do debate, a seleção resultou em sete *clusters*.

Por último, na terceira etapa, fez-se uma análise qualitativa dos principais termos utilizados nos discursos dos atores envolvidos e selecionaram-se alguns tuítes para exemplificar o cenário de polarização na rede. Os termos utilizados em cada tuíte são como “rastros digitais”. Ao abordar a rastreabilidade digital, Venturini e Latour argumentam que “agora é possível seguir uma multiplicidade de interações e, simultaneamente, distinguir a contribuição específica que cada uma delas tem para a construção de fenômenos sociais” (Venturini e Latour, 2019, p. 43). Ou seja, o objetivo foi analisar de forma qualitativa um conjunto de mensagens dos atores envolvidos no debate.

10 Compreende-se que existem diferenças teóricas entre o debate de “comunidades” e “*clusters*”; mas, neste texto, utilizaram-se os conceitos como sinônimos.

Resultados e discussões

O primeiro grafo, elaborado com bancos completos, foi construído com base “grau de saída” de cada nó. De acordo com Silva e Stabile, o grau de saída (*out degree*) representa a “quantidade de conexões que um nó envia na rede.” (Silva e Stabile, 2016, p. 244). No contexto das mídias sociais, de forma especial o Twitter, o índice é importante, pois ajuda a mensurar a influência, a autoridade, as preferências, o engajamento e a mobilização. Como se trata de um banco de dados direcional, foi observada a quantidade de conexão realizada por cada um dos perfis (nós). Para uma melhor análise, foram retirados do banco todos os perfis excluídos ou bloqueados pelo próprio Twitter¹¹. Para uma melhor visualização do gráfico¹², fez-se um filtro dos maiores *clusters* da rede. O tamanho dos rótulos é proporcional ao “grau de saída” de cada um dos perfis.

No conjunto da rede, observam-se perfis que pertencem a distintos campos de atuação. Nesse conjunto, estão presentes atores do setor empresarial, tais como Café 3 Corações, Vigor Brasil e Smart Fit; do meio jornalístico, como UOL, Estadão e Folha; do campo político, como Ciro Gomes, Manuela D’Ávila e Marina Silva; influenciadores digitais¹³, tais como Paula Marisa, Augusto Botelho e Felipe Neto. Deve-se destacar também o próprio perfil

11 A verificação dos perfis do Twitter e o processo de limpeza foi realizado no dia 1º de maio de 2022.

12 Para ajustar o tamanho da fonte dos rótulos, utilizaram-se como parâmetro o “grau de saída” (tamanho mínimo 1 e máximo 5).

13 O debate acadêmico em relação ao papel de influenciadores digitais é extenso, mas aqui utilizaremos a definição de Joathan e Alves (2020, p. 92). Para os autores, seriam perfis que “apresentam grande influência na rede, pois, por

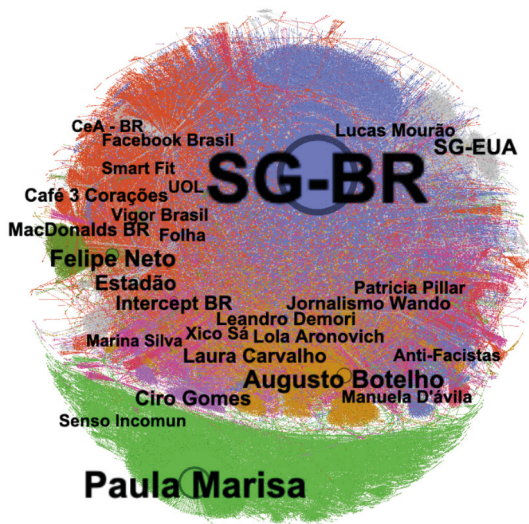


FIGURA 8.1. Rede de Nós SG-BR, parâmetro grau de saída

FONTE: Elaborado pelo autor com base em dados coletados entre 24 e 27 de julho de 2020.

do SG-BR possui o maior grau de saída, ou seja, representa o número de vezes que os usuários citaram o perfil. A rede completa demonstra um certo grau de polarização.

Apesar do claro cenário de polarização, deve-se destacar que, em pesquisa coordenada por Santini sobre a mobilização do SG-BR, verificou-se que boa parte dos ataques às marcas que

.....
terem centenas de milhares de seguidores, suas mensagens são bastante retuitadas.”

tomaram posição a favor do SG-BR “não foi uma reação orgânica da maior parte do público” (Santini, 2022, p. 10). Da mesma forma, grande parte dos perfis que cobrava posicionamentos das empresas junto do SG-BR foram impulsionados por “bots”. Em outras palavras, há fortes evidências de que no fenômeno estudado ocorre a “polarização dos bots”, criando-se, dessa forma, intimidação real das empresas anunciantes.

O debate sobre a polarização política no ambiente *on-line* não é um tema novo, visto que diversos pesquisadores já abordaram a questão, como Gruzd e Roy (2014), Adamic e Glance (2005), Smith, Rainie, Himelboim e Shneiderman (2014), Barberá (2015) e Recuero e Gruzd (2019). A singularidade da análise apresentada aqui se sustenta na constatação de que o fenômeno da polarização extrapola o campo “político”. Assim, a ação do SG-BR e de seus apoiadores e críticos concentra-se no campo do “mercado”, por mais que o objetivo inicial fosse sensibilizar os “consumidores” e não os “cidadãos”.

Para deixar nítido o cenário, após de apresentar o gráfico com a base de dados completa, optou-se por focar os atores que estão diferentemente envolvidos no “processo de politização do consumo”, através da cartografia dos “atores políticos” e dos “influenciadores digitais”. Os gráficos 2 e 3 foram construídos após um processo de filtragem com base em dois parâmetros, sendo eles: 1) retirada de todos os perfis publicados por pessoas jurídicas, empresas e meios de comunicação; e 2) retirada de todos os tuítes publicados pelos perfis do SG estadunidense e do SG-BR. Objetivou-se, assim, priorizar na análise os atores da política institucional e os influenciadores digitais, como estes se posicionaram entre o campo de ação política e o ativismo consumerista.

Para evidenciar a polarização do ativismo nos *clusters* da rede estudada, utilizou-se o algoritmo “OpenOrd¹⁴”. Através da nova rede, ao que tudo indica, o SG-BR colocou no centro do processo de polarização política no Brasil o debate sobre o ativismo político. Há fortes evidências de que o Twitter se transformou em uma arena cujos alinhamentos políticos partidários passaram a interferir nas práticas de consumo. Os dois próximos gráficos demonstram a polarização na matriz de dados estudada.

O processo que identificou os *clusters* foi realizado através de uma métrica de ARS que organiza a amostra em subgrupos, a partir da densidade de suas ligações, por meio do algoritmo de Louvain do Gephi. Conforme destacado por Joathan e Alves (2020), a organização através de *clusters* se fundamenta no princípio da hemofilia ideológica, na qual se agrupam perfis com valores e ideias semelhantes.

Os resultados observados permitem demonstrar a repercussão da rede polarizada. Podem-se visualizar, de forma clara, dois conjuntos de comunidades que se opõem: de um lado, os *clusters* lilás, laranja e vermelho, que mobilizaram, respectivamente, 15,45%, 5,57% e 5,47% dos perfis. Do lado oposto, observam-se, em destaque, os *clusters* verde escuro (9,2%), azul (7,94%) e preto (6,86%) dos perfis. O primeiro conjunto de *clusters* é composto por perfis com espectro ideológico da “direita”, que se definem como conservadores, com destaque para os perfis “Paula Marisa” (@profpaulamarisa), “Senso Incomum” (@sensoinc), “Douglas Garcia” (@DouglasGarcia), “Bela e Moral” (@Bela_e_Moral), “Ian Maldonado” (@ianmndo) e “Ccore” (@ccore_org). Do outro lado, há um conjunto de atores

14 A semente aleatória utilizada foi “3908859194339765496”.



FIGURA 8.2. Matriz completa dos *clusters*

FORNTE: Elaborado pelo autor com base em dados coletados entre 24 e 27 de julho de 2020.

da “esquerda” ou “centro-esquerda”, que defendem pautas progressistas, tais como “Augusto Botelho” (@augustodeAB), “Manuela D’Ávila” (@ManuelaDavila), “Laura Carvalho” (@lauraabcarvalho), “Marina Silva” (@MarinaSilva) e “Ciro Gomes” (@cirogomes).

Na pesquisa aqui realizada, os cinco perfis com maior grau de saída na rede foram: Paula Marisa (3.999), Augusto Botelho (1.672), Felipe Neto (1.438), Ciro Gomes (979) e Laura Carvalho (807). Um campo marcado pela disputa de atores da política institucional da esquerda e da direita. A participação destes atores é relevante, já que as mídias sociais, por meio de plataformas como

o Twitter, possuem uma função proeminente na ampliação dos espaços de debate público.

Ao analisar fragmentos das mensagens publicadas no Twitter, a figura 8.1 e a tabela 8.1 apresentam as 26 palavras mais utilizadas nos tuítes coletados. Na lista das palavras mais frequentes, destacam-se “fake” e “news”, ambas citadas mais de 13000 vezes. Dentre as expressões mais citadas, destacam-se algumas relacionadas à polarização política, como “direita” e “esquerda”, e outras vinculadas ao universo dos consumidores, como “anúncios”, “marca”, “propaganda” e “comprar”. As palavras no conjunto de tuítes evidenciam posicionamentos políticos, expressões e desejos dos consumidores. O quadro de palavras reforça o argumento a que o sujeito “consumidor” e o “cidadão” estão implicados, ou seja, um campo polifônico, no qual em um mesmo tuíte é possível demonstrar alinhamentos políticos e preferências no mercado de consumo.

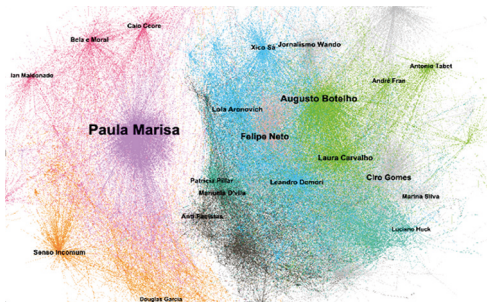


FIGURA 8.3. Clusters de atores políticos

FONTE: Elaborado pelo autor com base em dados coletados entre 24 e 27 de julho de 2020.



FIGURA 8.4. Nuvem de palavras mais utilizadas

FORNTE: Elaborada pelo autor por meio do site Voyant Tools e com base em dados coletados entre 24 e 27 de julho de 2020.

TABELA 8.1. Palavras mais utilizadas nos tuítes coletados

	PALAVRAS	FREQÜÊNCIA
1ª	<i>Fake</i>	154.600
2ª	<i>News</i>	13.934
3ª	<i>Site</i>	7.002
4ª	sleepinggiantbrasil	6.697
5ª	Parabéns	5.760
6ª	Brasil	3.114
7ª	Vocês	4.423
8ª	Eles	3.466

Continúa...

Continúa...

	PALAVRAS	FREQUÊNCIA
9ª	Vamos	2.259
10ª	Empresas	3.046
11ª	Anúncios	2.950
12ª	Ódio	2.668
13ª	<i>Fakenews</i>	2.582
14ª	Perfil	2.134
15ª	Direita	1.880
16ª	Conta	1.805
17ª	Marca	1.783
18ª	Notícias	1.778
19ª	Esquerda	1.741
20ª	Falsas	1.618
21ª	Verdade	1.616
22ª	Dinheiro	1.555
23ª	Mentiras	1.422
24ª	Produtos	1.317
25ª	Propaganda	1.286
26ª	Comprar	1.285

FONTE: Elaborada pelo autor com base em dados coletados entre 24 e 27 de julho de 2020.

A cobrança pública realizada pelo SG-BR em relação à publicidade que várias empresas realizavam em *sites* acusados de propagar *fake news* provocou ampla reação de consumidores e de atores do mundo político. Os tweets @ManuelaDavila reforçam a ideia do surgimento da “plataformização do ativismo consumérista” em torno do posicionamento das empresas. Diferentes de outras estratégias de politização do mundo do consumo, em que atores ou associações de defesa dos consumidores se envolvem em uma causa para mudar uma prática empresarial, o projeto SG-BR parte de um ator das redes sociais, que pauta uma campanha de denúncia contra um *site* específico. O sucesso da proposta depende diretamente da adesão de outros atores das redes sociais e do grau de exposição negativa desta marca. Nos casos aqui analisados, fica evidente que o objetivo principal do SG-BR não foi de promover diálogos sobre padrões éticos adotados pelas empresas nas plataformas de mídia social, mas de fomentar alterações nos mecanismos de financiamento publicitário nesse domínio.

Manuela D’Ávila. [@ManuelaDavila] (21 de maio de 2020)
/ Oi, gente! Vocês tem acompanhado o trabalho do @slp-ng_giants_pt ? entrem lá e entendam: eles identificam sites que distribuem notícias falsas e comunicam as empresas que os banners automáticos de propaganda estão sendo veiculados nesses sites. #SleepingGiantsBrasil. Com isso, eles cortam uma das raízes do mal. Claro q a gente sabe que o dinheiro de quem distribui fakenews no Brasil não vem só de anúncio do AdSense, afinal, gabinete central do ódio e gabinetes locais de ódio estão por aí. Mas causa um grande impacto.

#SleepingGiantsBrasil. [Tweet] Twitter. <https://twitter.com/ManuelaDavila/status/1263461300212023296>

A difusão generalizada da Internet e das redes de comunicação sem fio aumentou a visibilidade das pessoas sobre a conduta das empresas. Em sociedades com elevados graus de relações mediadas pela tecnologia, os “cidadãos” ou “consumidores” conectados às redes sociais estão mais suscetíveis a expressarem suas posições publicamente.

Senso Incomum. [@sensoinc] (20 de maio de 2020) / THREAD ANTI-CENSURA: Notamos que as empresas @DellNoBrasil @philipsbrasil @SamsungBrasil @submarino @americanascom estão retirando anúncios de sites. Tudo porque o perfil @slp-ng_giants_pt está fazendo lobby para censurar desafetos com a desculpa de “fake news”. #NaoCompreDell. [Tweet] Twitter. <https://twitter.com/sensoinc/status/1263211474962481154>

O ativismo do SG-BR e dos diversos apoiadores e críticos estaria entre as “ações comunicativas”, ou seja, “falam mal” das empresas que indiretamente financiam as *fake news* e o “boicote”, e recusam-se a comprar produtos motivados por fundamentos éticos. A repercussão entre os seguidores do perfil SG-BR apresenta ameaças para aquelas empresas que continuam expondo sua marca em *sites* que propagariam notícias falsas. Em outros termos, as disputas não se encaixam totalmente em nenhuma das tipologias descritas por Portilho (2020), visto que as ações de consumo político variam das formas mais “confrontacionais” às mais “cooperativas”.

Utilizando essa tipologia, é possível descrever as ações que envolvem a campanha do perfil SG-BR em três momentos distintos. No primeiro, o perfil do SG-BR se coloca como parceiro, busca “cooperar” com empresas e com organizações cuja publicidade está vinculada em *sites* que propagariam as *fake news* e o discurso de ódio. Já no segundo momento, entram em cena os “seguidores e apoiadores” do perfil SG-BR, que assumem a estratégia confrontacionista ao propor a não compra de produtos e serviços que financiam ações em *sites* com princípios éticos duvidosos. No terceiro e último momento, ocorre o “confrontacionismo dos confrontacionistas”, ou seja, surgem perfis de apoiadores que questionam a retirada dos anúncios de *sites* que propagam *fake news*. No mesmo tom ameaçador de boicote dos apoiadores do SG-BR, os críticos demonstram sua insatisfação pública contra as empresas que teriam retirado a exposição de suas marcas dos *sites*.

Considerações finais

O objetivo deste capítulo foi contextualizar, caracterizar e descrever o ativismo do *Sleeping Giants* Brasil e suas interlocuções com perfis do mundo político e econômico na rede social Twitter. De forma especial, observou-se que a “polarização” atual do campo político brasileiro se reproduz na rede estudada. Ao que tudo indica, a ação política do SG-BR e dos demais atores envolvidos na rede não se resume a um convencional “consumerismo de plataforma”. Esse processo de plataformização do ativismo consumerista do SG-BR assumiu contornos específicos devido ao cenário de polarização política no Brasil. Os embates decorrentes

da polarização fazem com que a negociação entre empresas, consumidores e cidadãos ganhe um terceiro elemento, sua filiação ou posicionamento partidário na plataforma.

De modo geral, a análise aqui desenvolvida identificou que grande parte dos perfis que se engajaram no debate do Twitter sobre o SG-BR se situa em comunidades constituídas por apoiadores das pautas reivindicadas pelos SG-BR. Foi possível perceber forte presença de atores da política institucional e de ativistas de diversos movimentos sociais. Além da constituição de comunidades em torno dos perfis de empresas e de meios de comunicação, ganham destaque, no campo de análise, atores da política institucional da esquerda, que tiveram grande alcance de suas mensagens, cenário que reforça a ocorrência de uma certa disputa partidária e ideológica no campo discursivo do conflito, inicialmente restrito à crítica ao mercado.

Deve-se também reconhecer que o trabalho conta com limitações a serem superadas em futuras pesquisas. Como primeira lacuna, destaca-se que a análise foi realizada exclusivamente com base no Twitter, e há grande possibilidade de se chegar a resultados diferentes em outras plataformas, tais como Facebook e Instagram, já que essas possuem regras, arquitetura e públicos diferentes. A segunda lacuna remete aos resultados do relatório coordenado por Santini (2022), no qual se verificou grande probabilidade de mensagens provenientes de “bots”. A pesquisa sobre SG-BR enfatiza que provavelmente boa parte da polarização dos debates contra marcas de empresas no Twitter foi impulsionada artificialmente por “bots” e “trolls”; assim, em grande medida, o cenário da polarização foi superdimensional graças a contas automatizadas ou semiautomatizadas (*social bots*). Deve-se destacar

também que, por opção, o capítulo não focou sua análise na repercussão do ativismo do SG-BR junto das empresas. A análise do comportamento do setor empresarial frente às denúncias públicas do SG-BR, de ativistas do campo da política institucional e de movimentos sociais pode apresentar resultados significativos e merece atenção em futuras pesquisas.

Diante do contexto, um dos grandes desafios do presente trabalho foi abordar um campo de debate relativamente novo nas humanidades digitais no Brasil e na América Latina. Muitas pesquisas são direcionadas para os estudos culturais digitais e o ativismo político *on-line*, mas, em nossa visão, o debate é incipiente se tratando de consumo e ativismo político digital. Portilho (2020), ao analisar as especificidades do consumo político¹⁵ na América Latina, destaca que movimentos sociais e grupos organizados acreditam que a melhor estratégia de “ativismo” seria através do Estado e da consolidação de políticas públicas eficientes, já que o mesmo teria capacidade de regular o “mercado” e atender os interesses dos cidadãos.

A pesquisa aqui desenvolvida considerou as dimensões éticas dos atores descritos na pesquisa, garantindo a privacidade e a proteção dos dados dos usuários envolvidos nesse estudo. O estudo sobre *Sleeping Giants* no Brasil, no campo das Humanidades Digitais, se constitui como uma abordagem importante para compreender as dinâmicas sociais, políticas e culturais que emergem no ambiente digital e para explorar novas estratégias de pesquisa que permitam analisar grandes volumes de dados neste espaço.

15 Vale destacar que Portilho (2020) analisa o consumo político de forma geral; a autora não faz diferenciação entre ativismo *on-line* e *off-line*.

Espera-se que outras pesquisas sobre ativismo no X (antes chamado Twitter) dentro das Humanidades Digitais forneçam análises valiosas em relação ao papel da tecnologia na formação e mobilização de movimentos sociais e sobre a dinâmica da participação cívica e política na era digital. O Twitter se apresenta como uma das principais plataformas de mídia social usadas por ativistas para promover suas causas e mobilizar a opinião pública. A abordagem das Humanidades Digitais observada no *Sleeping Giants* no Brasil pode ajudar a entender como as redes sociais são usadas para organizar e mobilizar movimentos sociais, além de examinar como as narrativas e discursos são moldados e disseminados no ambiente *on-line*.

Por fim, é importante destacar que esta pesquisa não analisou o ativismo de “concorrentes” do SG-BR. Por isso, ressalta-se que, em novembro de 2020, foi criado um perfil do Twitter que se propõe a fazer oposição às pautas apresentadas pelo SB-BR, o Awake Giants Brasil (@awake_giants_br), o qual se define como “União de consumidores na luta contra a censura. Conscientização para o consumo saudável, questionando marcas que aceitam e apoiam a censura dos “dorminhocos”¹⁶. Em 3 de abril de 2022, o perfil contava com 79100 seguidores, quantidade bem menor do que os 489300 do SG-BR, mas significativa para um perfil relativamente recente. Uma análise comparativa e discursiva entre os dois perfis merece novas abordagens, e possivelmente dirá muito sobre o atual cenário de polarização política e sobre o ativismo consumerista de plataforma no Brasil.



16 A proposta de questionar a censura dos “dorminhocos” é uma clara referência ao SG-BR.

Referências bibliográficas

- Adamic, L. A. e Glance, N. (2005). The political blogosphere and the 2004 US election: divided they blog. *Proceedings of the 3rd international workshop on Link discovery* (pp. 36-43), Association for Computing Machinery.
- Alexander, J. C. (1998). Ação Coletiva, Cultura e Sociedade Civil: Secularização, atualização, inversão, revisão e deslocamento do modelo clássico dos movimentos sociais. *Revista Brasileira de Ciências Sociais*. 13(37), 5-31.
- Barberá, P. (2015). How Social Media Reduces Mass Political Polarization. Evidence from Germany, Spain, and the U.S. *Proceedings of the 2015 APSA Annual Meeting*, San Francisco.
- Barnes, J. A. (1954). Class and committees in a Norwegian island parish. *Human relations*, 7(1), 39-58.
- Beck, U. (1997). A reinvenção da política: rumo a uma teoria da modernização reflexiva. In: U. Beck; A. Giddens; S. Lash (Orgs.), *Modernização reflexiva*. Editora da Unesp.
- Bergamo, M. (2020, 26 de dezembro). Revelação de fundadores do Sleeping Giants causa negacionismo e ameaças. *Folha de São Paulo*, Coluna Monica Bergamo. www1.folha.uol.com.br/colunas/monicabergamo/2020/12/revelacao-de-fundadores-do-sleeping-giants-causa-negacionismo-e-ameacas.shtml.
- Biazi, T. (2019). Comunidades #REA: análise de seus rastros no Twitter. *TextoLivre: Linguagem e Tecnologia*. 12(3), 69-92.
- Borges, F. M. (2017). *Consumerismo e consumidores indignados: Ne-tativismo contra as marcas nas redes sociais*. [tese de doutorado em Ciências Sociais inédita]. Pontifícia Universidade Católica de São Paulo (PUC-SP), São Paulo.

- Braun, J. A.; Coakley, J. D.; West, E. (2019). Activism, advertising, and far-right media: The case of Sleeping Giants. *Media and Communication*, 7(4), 68–79. <https://doi.org/10.17645/mac.v7i4.2280>.
- Castells, M. (2018). *Ruptura: a crise da democracia liberal*. Zahar.
- Canclini, N. G. (1999). *Consumidores e Cidadãos: Conflitos multiculturais da globalização*. Editora UFRJ.
- D’Andréa, C. (2020). *Pesquisando plataformas on-line: conceitos e métodos*. EDUFBA.
- Di Felice, M. (2011). *Netativismo: Novos aspectos da opinião pública em contextos digitais*. International Association for Media and Communication Research. IAMCR 2011. Istanbul, 13 a 17 de julho de 2011.
- El País Brasil. (2020, 20 de maio). Movimento expõe empresas do Brasil que financiam, via anúncios, sites de extrema direita e notícias falsas. <https://www.menorca.info/actualidad/internacional/2023/01/10/1860471/brasil-identificados-financiadores-asaltantes-diez-estados.html>
- Filho, J. (2020, 20 de maio). Filipe Martins e a rede de mentiras agonizam: um tuíteiro nocauteou a grana de sites de fake news. *The Intercept Brasil*, Opinião. https://theintercept.com/2020/05/24/fake-news-extrema-direita-filipe-martins-bolsonarismo/?fbclid=IwARovfYhvLmz6CC8XFmf__asThN48RrXFtjYEDivAEcqoqvyqc3jbSGuhwM.
- Fiscarelli, A. M. (2022). Social network analysis for digital humanities. In A. Fickers e J. Tatarinov, *Digital History and Hermeneutics: Between Theory and Practice*. De Gruyter Oldenbourg.
- Freitas, E. T. (2021). Sleeping Giants Brasil contra o ódio e as fake news: Um ativismo consumerista no Instagram. In L. Barbosa,

- F. Portilho, F. Galindo e S. Borges, (Orgs.), *Encontros e caminhos dos estudos do consumo no Brasil*. E-papers.
- Garton, L.; Haythornthwaite, C.; Wellman, B. (1997). Studying on-line social networks. *Journal of computer-mediated communication*, 3(1).
- Gillespie, T. (2010). The politics of “platforms”. *New Media & Society*, 12, 347-364.
- Gomes, J. C.; Pimenta, R. M.; Braga, T. E. N. (2021). *Twitter Monitor: Milestone 4*. [s.l.] Zenodo.
- Gruzd, A., e Roy, J. (2014). Investigating Political Polarization on Twitter: A Canadian Perspective. *Policy and Internet*, 6(1), 28-45.
- Halkier, B. (1999). Consequences of the politicization of consumption: the example of environmentally friendly consumption practices. *Journal of Environmental Policy and Planning*, 1, 25-41.
- Helmond, A. (2019). A plataforma da web. In J. J. Omena, (Ed.), *Métodos Digitais: Teoria-Prática-Crítica*. Livros ICNOVA.
- Hilton, M. (2003). *Consumerism in Twentieth-Century Britain: The search for a historical movement*. Cambridge University Press.
- Hirschman, A. (1983). *De consumidor a cidadão: Atividades privadas e participação na vida pública*. Brasiliense.
- Hypeness. (2020, maio 4). Sleeping Giants: a luta contra as fake news que tira o sono de políticos no Brasil e no Mundo. [Hiper-texto]. (<https://revistaoeste.com/politica/sleeping-giants-divulga-fake-news-e-e-advertido-pelo-twitter/>)
- Joathan, I. e Alves, M. (2020). O Twitter como ferramenta de campanha negativa não oficial: uma análise da campanha eleitoral para a Prefeitura do Rio de Janeiro em 2016. *Galaxia*, 43, 81-98.
- Micheletti, M. (2003). *Political virtue and shopping: Individuals, consumerism and collective action*. Palgrave Macmillan.

- Miller, D., e Slater, D. (2004). Etnografia *On e Off-line*: Cibercafés em Trinidad. *Horizontes Antropológicos*, 10(21), jan./jun, 41-65.
- Moreno, J. L. (1978). *Psicodrama*. Cotrix.
- Nestle, M. (2002). *Food Politics*. University of California Press.
- Omena, J. J. e Amaral, I. (2019). Sistema de Leitura de Redes Digitais Multiplataforma. In J. J. Omena, (Ed.), *Métodos Digitais: Teoria-Prática-Crítica*. Livros ICNOVA.
- Parreiras, C. (2015). *Altporn, corpos, categorias, espaços e redes: um estudo etnográfico sobre pornografia on-line* [tese de doutorado inédita]. Universidade Estadual de Campinas (Unicamp), Campinas.
- Piaia, V.; Matos, E.; Almeida, S.; Dienstbach, D. e Barboza, P. (2021). Breque dos Apps: Uma Análise Temporal de Comunidades e Influenciadores no Debate Público *On-line* no Twitter. *Comunicação e Sociedade*, 39, 57-81.
- Poell, T., Nieborg, D. e Van Dijck, J. (2020). Plataformização. *Revista Fronteiras – estudos midiáticos*, 22(1), 2-10.
- Portilho, F. (2020). Ativismo alimentar e consumo político – Duas gerações de ativismo alimentar no Brasil. *Redes*, 25(2), 12-33.
- Portilho, F. e Barbosa, L. (2016). A adesão à “causa” rural e da agricultura familiar por consumidores e seus movimentos organizados. In F. C. Marques, M. A. Conterato, S. Schneider, *Construção de mercados e agricultura familiar: Desafios para o desenvolvimento rural*. Editora da UFRGS.
- Quimaliños, R. (2020, 17 de maio). O homem que arruinou a extrema direita nos EUA. *O Globo*. <https://brasil.elpais.com/icon/2020-05-17/o-homem-que-arruinou-a-extrema-direita-nos-eua.html>.
- Recuero, R. e Gruzd, A. (2019). Cascatas de *Fake News* Políticas: um estudo de caso no Twitter. *Galaxia*, 41, 31-34. <http://dx.doi.org/10.1590/1982-25542019239035>

- Recuero, R.; Zago, G.; Bastos, M. T. (2014). O Discurso dos #ProtestosBR: análise de conteúdo do Twitter. *Galaxia* (São Paulo, *on-line*), 28, 199-216. <http://dx.doi.org/10.1590/1982-25542014217911>
- Ribeiro, B. G.; Ribeiro, M. H.; Almeida, V.; Meira Jr, W. (2020). Analyzing the “Sleeping Giants” Activism Model in Brazil. *arXiv*. <https://arxiv.org/abs/2105.07523>.
- Santini, R. M. (2022). *Relatório da Rede Sleeping Giants Brasil no Twitter*. <https://bit.ly/3gNlcXK>.
- Sassatelli, R. (2006). Virtue, responsibility and consumer: Framing Critical Consumerism. In J. Brewer, F. Trentmann (Eds.). *Consuming Cultures, Global Perspectives: Historical Trajectories, Transnational Exchanges*. Berg.
- Schwartz, E. (1996). *NetActivism: How citizens use the internet*. O'Reilly.
- Shu, K.; Silva, A.; Wang, S.; Jang, J.; Liu, H. (2017). Fake news detection on social media: A data mining perspective. *arXiv*. <https://arxiv.org/abs/1708.01967>.
- Silva, T. e Stabile, M. (2016). *Monitoramento e pesquisa em mídias sociais: Metodologias, aplicações e inovações*. Uva Limão.
- Smith, M.; Rainie, L.; Himelboim, I.; Shneiderman, B. (2014). *Mapping Twitter Topic Networks: From Polarized Crowds to Community Clusters*. Pew Research Center.
- Stolle, D.; Hooghe, M.; Micheletti, M. (2005). Politics in the supermarket: political consumerism as a form of political participation. *International Political Science Review*, 26 (3), 254-269.
- Trentmann, F. (Ed.). (2006). *The making of the consumer: Knowledge, power and identity in the modern world*. Berg.
- Van, J., Poell, T. e De Wall, M. (2018). *The Platform Society: Public Values in a Connective World*. Oxford University Press.

Venturini, T. e Latour, B. (2019). O tecido social: rastros digitais e métodos quali-quantitativo. In J. J. Omena, (Ed.), *Métodos Digitais: Teoria-Prática-Crítica*. Livros ICNOVA.

***La fiebre del banano:*
una aproximación a
la curaduría de arte
contemporáneo a través
de las humanidades
digitales**

Para citar este capítulo:

<http://doi.org/10.51573/Andes.9789587987362.9789587987379.9789587987867.9>

JUANITA SOLANO ROA
BLANCA SERRANO ORTIZ

Introducción

¿Puede una exposición de arte contemporáneo tener más potencial crítico y enfatizar en la multiplicidad de lecturas de las obras mediante un formato digital? Este capítulo explora esta pregunta a través del análisis del proyecto de humanidades digitales *La fiebre del banano/Banana Craze*¹. Concebido por las historiadoras del arte Blanca Serrano y Juanita Solano, *La fiebre del banano/Banana Craze* se enfoca en el arte contemporáneo de América Latina para narrar la historia del banano, una fruta que ha sido motivo de intervenciones políticas por parte de multinacionales de Estados Unidos en algunas naciones latinoamericanas, objeto de la construcción de estereotipos identitarios y motor de una industria extractiva con grandes repercusiones sobre el medio ambiente y la salud de las personas que trabajan en las plantaciones bananeras.

1 Véase <https://bananacraze.uniandes.edu.co/>

La fiebre del banano/Banana Craze nace inicialmente como un proyecto de curaduría en la web, ideado como un recurso cultural y educativo abierto al público general y especializado. Es decir, como ejercicio de selección y organización de obras de arte en un espacio, añadiendo así capas interpretativas y, por lo tanto, mediando en la lectura de las obras. Sin embargo, al consultar el proyecto hoy en día, es evidente que la intención original se vio rebasada gracias al uso de herramientas digitales, las cuales permitieron expandir las lecturas de las obras exhibidas, la investigación que se continúa llevando a cabo en la actualidad y la divulgación del proyecto. Así, el proyecto puede navegarse de tres maneras complementarias y convergentes: como una exhibición virtual, como un archivo digital y como un proyecto de humanidades digitales públicas. En efecto, la página web que alberga el proyecto cuenta con estas tres secciones: una dedicada a la exposición virtual, que contiene las tres salas principales que dividen la exposición bajo los ejes temáticos de violencias, identidades y ecosistemas (figura 9.1); una sección que permite explorar las obras bajo motores de búsqueda de índole alfabético, cronológico y geográfico (figuras 9.2, 9.3 y 9.4); y un módulo dedicado a la investigación continua, y la difusión y activación del proyecto. Este último se divide en un apartado para textos de investigación, ya sean en formato de blog o artículos académicos que expanden el trabajo presentado en la página web (figura 9.5), una amplia bibliografía sobre el tema de investigación y un repositorio con las grabaciones en video de todos los eventos desarrollados en torno a la plataforma, en colaboración con artistas, cineastas, curadores invitados y la participación del público asistente (figura 9.6). Esta última sección es también donde *La fiebre del banano/Banana*

Craze adquiere valor en términos de proyecto educativo abierto al público general.

Al considerar las obras de arte lugares desde los cuales se produce conocimiento, esto es, como un centro epistemológico en sí mismas, el proyecto plantea que una obra puede ser interpretada desde múltiples perspectivas y áreas de especialización y, por lo mismo, ser vehículo para el desarrollo de líneas de investigación plurales y no lineares, que no se limitan a una única manera de abordar la investigación, sino que reúnen aproximaciones interdisciplinarias. En ese sentido, la curaduría en la web plantea ventajas sobre la curaduría física, dado que permite que una obra de arte esté incluida en varias “salas” temáticas a la vez (en vez de estar limitada a una sola), posibilita que el visitante explore la muestra de acuerdo con sus intereses, mediante la búsqueda alfabética, geográfica, cronológica y temática, y amplía los horizontes de la investigación en curso por medio de la realización de programas públicos virtuales, que expanden las perspectivas del proyecto a través de la participación de distintas voces. Además, el formato digital permite desarrollar textos curatoriales más extensos que lo que habitualmente cabe en la pared de una exposición física, en los cuales se explora la investigación sobre el tema a gran profundidad. Otro de los múltiples aportes de la curaduría en la web consiste en la posibilidad de observar la relación entre las piezas desde múltiples perspectivas, lo que permite establecer nuevas conclusiones sobre el conjunto de obras y no solo a nivel individual.

La fiebre del banano/Banana Craze es un proyecto colaborativo y pedagógico. Con la página web y las múltiples activaciones que se hacen en los eventos se busca ilustrar a los usuarios sobre la historia y el impacto del monocultivo del banano en América Latina.

<358,



FIGURA 9.1. Captura de pantalla de las tres salas que albergan la exposición dentro del proyecto *La fiebre del banano/Banana Craze*

FUENTE: <https://bananacraze.uniandes.edu.co/>

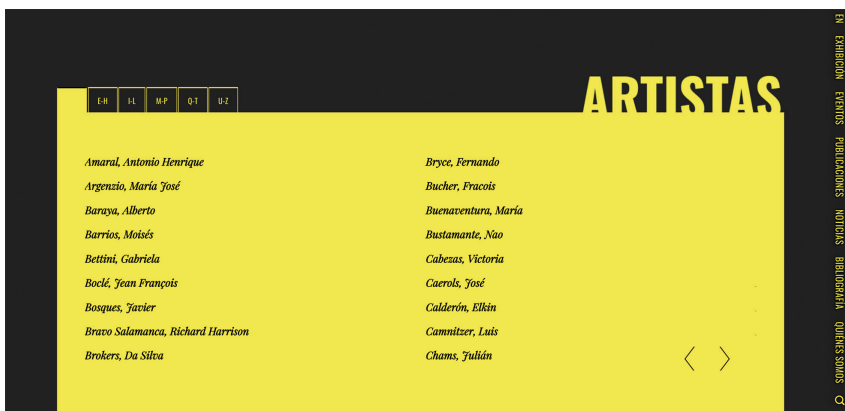


FIGURA 9.2. Captura de pantalla de la sección de búsqueda alfabética por apellidos de los artistas incluidos en *La fiebre del banano/Banana Craze*

FUENTE: <https://bananacraze.uniandes.edu.co/>

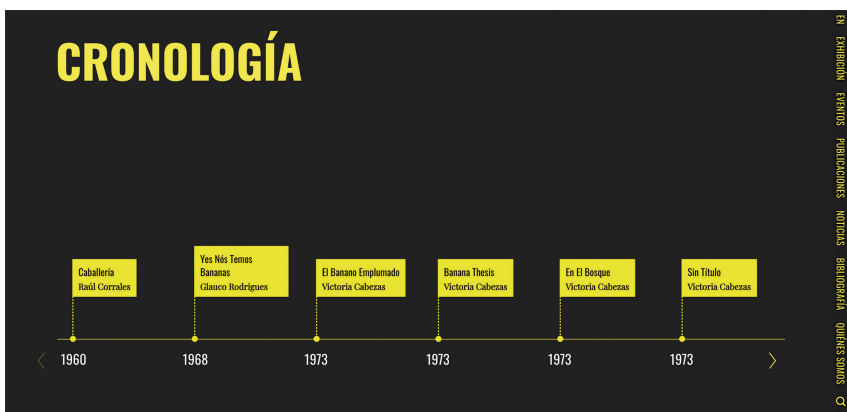


FIGURA 9.3. Captura de pantalla de la sección de búsqueda cronológica por año de realización de las obras incluidas en *La fiebre del banano/Banana Craze*

FUENTE: <https://bananacraze.uniandes.edu.co/>

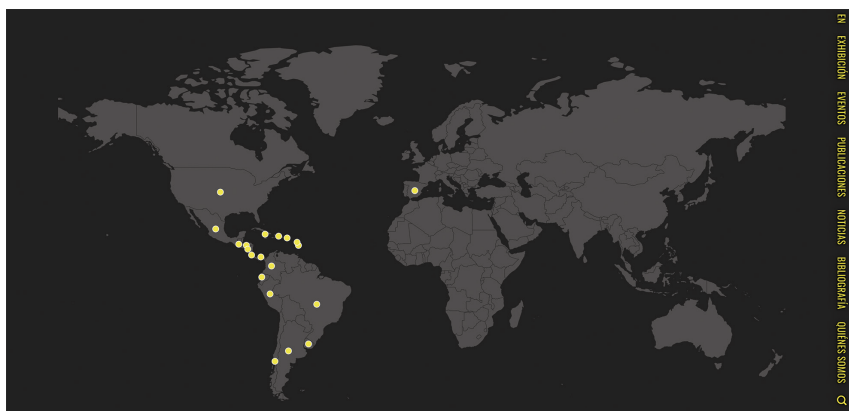


FIGURA 9.4. Captura de pantalla de la sección de búsqueda geográfica por país de proveniencia de los artistas incluidos en *La fiebre del banano/Banana Craze*

FUENTE: <https://bananacraze.uniandes.edu.co/>



FIGURA 9.5. Captura de pantalla de la sección de publicaciones del proyecto *La fiebre del banano/Banana Craze*

FUENTE: <https://bananacraze.uniandes.edu.co/>

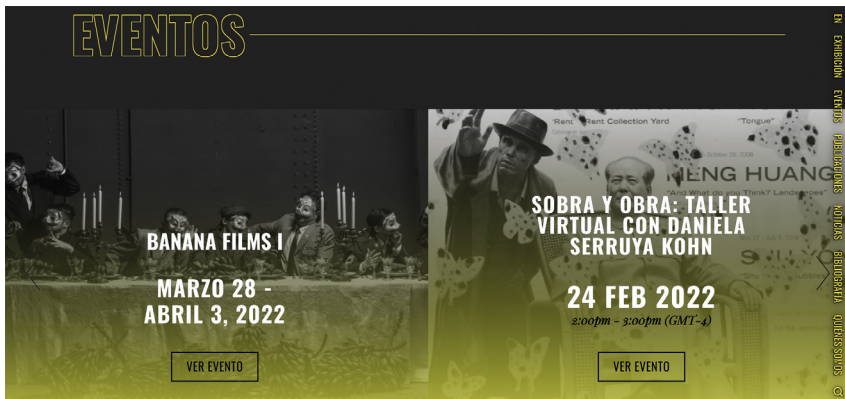


FIGURA 9.6. Captura de pantalla de la sección de eventos del proyecto *La fiebre del banano/Banana Craze*

FUENTE: <https://bananacraze.uniandes.edu.co/>

Otra función pedagógica del proyecto ha sido formar a un grupo de estudiantes que han trabajado con las autoras sobre prácticas y procesos de investigación y creación en las artes. En este capítulo se analizan las prerrogativas y desafíos en la conceptualización y el proceso de construcción de *La fiebre del banano/Banana Craze* teniendo en cuenta cada una de las secciones descritas.

¿Qué investiga el proyecto?

La fiebre del banano/Banana Craze es el primer gran estudio de cómo un recurso natural como el banano ha forjado el pasado y el presente de un continente y cómo este fenómeno se expresa a través de la cultura. Así, reúne más de cien obras de arte contemporáneo latinoamericano en las que el banano es el protagonista. A partir de la fotografía del cubano Raúl Corrales titulada *Caballería*, de 1960, en la que un grupo de hombres cabalga victorioso celebrando la expropiación de las plantaciones de la United Fruit Company por parte del Gobierno revolucionario, la investigación de *La fiebre del banano/Banana Craze* llega hasta el presente y continúa en constante desarrollo. El estudio de estas obras permite comprender desde una perspectiva artística, cultural y filosófica cómo el cultivo masivo del banano contribuyó al aumento de las desigualdades sociales en América Latina, alteró las formas de vida tradicionales locales y transformó el paisaje y el medio ambiente de la región, además de contribuir a la configuración de estereotipos xenófobos, racistas y sexistas sobre sus habitantes (Chapman, 2009; Coleman, 2016; Enloe, 1990; Jenkins, 2000; Klein, 2020; Koepfel, 2008; Soluri, 2006; Striffler y Moberg, 2003).

El banano fue una fruta prácticamente desconocida por fuera de sus espacios de producción hasta finales del siglo XIX, en dramático contraste con la situación actual: hoy en día, no es solamente muy común, sino que también ocupa un lugar destacado en la cultura popular global. Así lo demuestran, por ejemplo, desde el banano adherido a la pared con cinta pegante que Maurizio Cattelan vendió por \$120 000 dólares en Art Basel Miami, en el 2019, hasta el popular *jingle* comercial de Chiquita Banana. Pero ¿cómo ha llegado esta fruta a convertirse en ícono de la violencia, fetiche sexual y símbolo de lo exótico y lo “subdesarrollado”? Este proyecto de investigación busca indagar sobre las políticas sociales, económicas y culturales que el cultivo masivo del banano generó en el continente americano, en particular en América Latina y el Caribe, a través del arte contemporáneo. Las más de cien obras que forman parte de este proyecto desenmascaran y examinan las relaciones de poder y las tensiones establecidas entre el norte y el sur del continente, a partir de los diferentes papeles que desempeñaron cultivadores, exportadores y consumidores. Más que una exhibición virtual, este es un espacio de investigación activo, creciente y en constante desarrollo.

En la actualidad, el banano es la fruta más consumida en el mundo, con una producción estimada de ciento dieciséis millones de toneladas al año. Si bien el mayor consumo de la fruta se da a nivel local, pues forma parte de la dieta básica de los diferentes países que lo producen, su exportación también es masiva y genera alrededor de doce billones de dólares al año (Organización de las Naciones Unidas para la Alimentación y la Agricultura [FAO], 2017). América Latina y el Caribe es la segunda región productora de banano del mundo, después de Asia, pero la región aporta

alrededor del 75 % de las exportaciones mundiales (la producción de banano de Ecuador, por ejemplo, representa el 30 % de la oferta mundial). El banano es la fruta más cultivada en el subcontinente americano, seguida por la piña, el mango, el aguacate y la papaya. Un 80 % de las exportaciones están dirigidas hacia los países desarrollados, principalmente en Europa y Estados Unidos, lo que ilustra la desigual relación de dependencia que examinan numerosas de las obras incluidas en este proyecto (Altendorf, 2019).

Desde mediados del siglo xx, numerosos artistas de Latinoamérica y el Caribe han reflexionado sobre la ubicuidad del banano en la dieta global y han investigado la historia de explotación sufrida por sus naciones como consecuencia de la internacionalización del mercado bananero a partir de fines del siglo xix y hasta el presente. Las bananeras son, por lo tanto, un lugar desde donde es posible reflexionar acerca de la realidad sociopolítica de Latinoamérica; por ello, la representación del banano en las artes visuales latinoamericanas aún prácticas tan diversas entre sí. A pesar de las diferencias existentes entre los contextos y momentos a los que se refieren las obras de este proyecto, estas revelan el legado común que comparten sus artistas. En efecto, desde la relación del banano con el cambio climático, pasando por el impacto del comercio de bananos por parte de compañías extranjeras en la desigualdad social de las comunidades agricultoras, hasta la similitud entre el viaje de muchos migrantes latinoamericanos a países del llamado primer mundo y el transporte de bananos desde los países donde se cultivan hasta los lugares donde se los consume, la comercialización del banano a gran escala toca ejes fundamentales del desarrollo de sus países productores desde comienzos del siglo xx.

Metodología

La metodología de *La fiebre del banano/Banana Craze* resuena con la máxima de Patrick Svensson (2017), quien define las humanidades digitales “como herramienta, como objeto de estudio, como medio de expresión, como laboratorio experimental y como lugar de activismo” (p. 2). De esta manera, este proyecto de investigación articula una exposición, un archivo, un repertorio de publicaciones y una serie de programas públicos, para tejer una red en constante crecimiento en torno al legado cultural del monocultivo del banano en Latinoamérica.

Tras un polémico artículo publicado en *Hyperallergic* en otoño del 2021 sobre cómo el florecimiento de exposiciones virtuales surgidas a raíz de la pandemia habitualmente no ofrecía nada distinto ni mejor a una exposición presencial, Seph Rodney se corrigió a sí mismo seleccionando tres exposiciones virtuales “que sí merecían ser visitadas”. Junto con una muestra del J. Paul Getty Museum y otra de Google Arts and Culture, Rodney seleccionó *La fiebre del banano/Banana Craze*. Rodney (2022) describe el proyecto como

una indagación histórica profunda. Un gran proyecto de investigación que hace uso del medio digital para mostrar imágenes y textos en formas que uno no podría experimentar en la vida real. Uno se asoma a la plataforma como a un abismo repleto de infinidad de maneras para presentar y relacionar la información².

2 Todas las traducciones del inglés fueron realizadas por las autoras de este texto.

En efecto, *La fiebre del banano/Banana Craze* demuestra que la historia del arte digital sirve para crear herramientas de conocimiento totalmente nuevas, al aplicar perspectivas académicas originales combinadas con métodos innovadores de visualización de datos, que permiten comunicar de manera eficaz y sugerente los hallazgos de una investigación (Rodríguez Ortega, 2018). En ese sentido, la historia del arte digital y, en particular para el caso aquí analizado, la curaduría en la web proporcionan la exploración de distintos tipos de fuentes documentales y de bases de datos, que presentados en un formato no digital no podrían abordarse simultáneamente. Es decir, las humanidades digitales posibilitan narrar aquello que de otra forma sería difícil de presentar, abarcar o comprender tanto para el público académico como para el no especializado. Como ocurre en *La fiebre del banano/Banana Craze*, estos proyectos también deben tener un carácter abierto que permita que las herramientas disponibles para el usuario y los datos presentados se sientan parte de un patrimonio común y abierto que cada uno puede consultar y recombinar.

La fiebre del banano/Banana Craze evita la recopilación meramente enciclopédica de imágenes y textos. En cambio, la plataforma explora las numerosas y creativas posibilidades de presentación visual (exposición, mapa, línea del tiempo, etc.) de contenidos y datos que ofrece la tecnología, tanto en el cotejo de datos geográficos y cronológicos de la sección de archivo como en la exposición virtual, donde el encuentro con las obras de arte siempre se desarrolla en relación con otras piezas y contextos y nunca en un recorrido forzosamente lineal ni impuesto previamente al usuario. Así, este proyecto parece responder eficazmente a la cuestión formulada por Paul B. Jaskot (2013): “la pregunta no es qué

puede hacer la historia del arte con lo digital; la pregunta es cuáles son las cuestiones importantes de la historia del arte que pueden abordarse con la ayuda de las herramientas digitales” (p. 140). De esta forma, en *La fiebre del banano/Banana Craze* encontramos, primero, cómo la investigación histórico-artística está en el centro de la plataforma y, después, que el uso de las nuevas tecnologías se articula en torno a esa investigación.

Como se analizará en este texto, a través de los distintos ejes temáticos de la exposición virtual y de los distintos parámetros de búsqueda para cotejar las obras de arte (mapa y línea del tiempo), *La fiebre del banano/Banana Craze* refleja la afirmación de Elijah Meeks (2011) de que “la visualización y la historia espacial no consisten en producir ilustraciones o mapas para comunicar cosas que se han descubierto por otros medios, sino que son un medio de investigación en sí mismo”. Como ha estudiado Isabel Galina (2011), este proyecto aúna metodologías propias de disciplinas como la bibliotecología, las ciencias de la información y las ciencias de la computación, con el fin de alcanzar un resultado único y original en el que el tema de la investigación y las herramientas para analizarlo son inseparables.

Antes de explicar el proyecto en detalle, es importante aclarar que todas las obras de arte incluidas en la plataforma fueron concebidas como piezas físicas y no digitales. Para incluirlas en *La fiebre del banano/Banana Craze*, recurrimos al tradicional recurso de la fotografía que ilustra la obra (como se haría en un catálogo de una exposición), mas no es la obra en sí misma. Esto implicó contactar a cada uno de los artistas, así como a las instituciones que albergan físicamente las obras, para pedir los permisos de reproducción de cada una de ellas. Más adelante profundizaremos sobre los retos

que esto implicó en la construcción de la página web, siguiendo los parámetros legales de derecho de autor en Colombia, país en donde se encuentra alojada la plataforma.

Posibilidades y limitaciones de la curaduría digital

¿Qué es una exhibición virtual? Según Michael Connor (2020), director artístico de la revista *Rhizome*, una exhibición virtual debe considerarse una práctica diferente a una exhibición en un espacio físico (aunque en cierta medida interconectada con esta), en la que la cualidad *performativa* y variable de la cultura nacida digitalmente es un aspecto clave. La exposición digital, ya sea aquella creada directamente para la exhibición de arte digital o como un espacio digital para enmarcar arte que fue concebido de forma física, “expande el trabajo de curaduría en relación con sus modos de producción, presentación y recepción” (Egger y Ackermann, 2020, p. 19). En últimas, y de acuerdo con la propuesta de Marialaura Ghidini (2019), la curaduría digital debe entenderse como un sistema de producción y mediación artística que está atravesado no solo por la curaduría, sino también por las formas de comunicación intertextual, los formatos de publicación y los modos de distribución que permite la tecnología del internet. Autores como Ghidini (2019) establecen un paralelo entre la curaduría digital y el radicalismo de las prácticas curatoriales que surgieron en los años sesenta —como los establecidos por las redes de artistas del arte correo o el arte conceptual—, los cuales forzaron la salida del arte del espacio del museo y la galería:

al idear nuevos formatos para producir y exhibir arte fuera de la galería, la curaduría en la web abarca aspectos que pertenecen a la relación entre la historia de la tecnología y cómo los nuevos modos de distribución y comunicación habilitados por la tecnología cambian la producción de cultura y su acceso. (p. 2)

La fiebre del banano/Banana Craze es un proyecto pensado directamente a partir de un entendimiento del espacio digital, es decir, la estructura de la página, las obras incluidas, los textos desarrollados, la cronología, el mapa y la multiplicidad de lecturas que se pueden hacer tanto de las obras individuales como de las relaciones que se establecen entre todas ellas y sus diversas visualizaciones fueron seleccionadas, organizadas y distribuidas pensando en la experiencia digital del proyecto. La manera en la que los datos se presentan hace que el usuario quiera explorar la base de datos de diversas maneras y que su interacción con la página y la búsqueda de información rebasa la simple “casilla de búsqueda” que usualmente tienen páginas web como las de los museos. Por el contrario, este proyecto invita a explorar las obras desde una multiplicidad de entradas y visualizaciones: *La fiebre del banano/Banana Craze* ofrece lo que Mitchell Whitelaw (2015) ha llamado *interfaz generosa*.

Esta exposición podría catalogarse dentro de lo que se ha definido como *curaduría en la web*, en oposición a la *curaduría en línea*, que busca expandir o reformular lo que se ha presentado ya físicamente (Dietz, 1998). A diferencia de las muestras físicas, las exposiciones virtuales no tienen lugar en un espacio unificado y coherente, como el museo o la galería de arte. Por el contrario,

cada usuario tiene una experiencia diferente en tanto que la muestra se experimenta a través de diversos dispositivos, como celulares, computadores y tabletas de diferentes tamaños, esto es, la experiencia en la pantalla de cada usuario es muy variable. En ese sentido, la exposición en línea no debe plantearse solo como un espacio pictórico navegable, sino que implica una puesta en escena multifacética (Connor, 2020).

Siguiendo esta premisa, la navegabilidad de *La fiebre del banano/Banana Craze* varía según el dispositivo desde el cual se experimenta el proyecto. Mientras que el usuario que entra a través del computador se encuentra en primera instancia con el título del proyecto y un *banner* que rota imágenes de las múltiples obras que forman parte de la muestra, el usuario que accede con su teléfono celular se encuentra con el mismo título, pero sobre un fondo negro, e inmediatamente puede leer una breve descripción de lo que trata el proyecto. Así mismo, la experiencia en el recorrido varía de acuerdo con el dispositivo utilizado para su acceso. En el teléfono móvil la narración es netamente vertical, mientras que en la navegación en la tableta o en el computador se aprovecha la horizontalidad de la pantalla para la creación de la narrativa visual. Esto es en particular visible en la sección de la cronología, donde la línea del tiempo se ve como un todo en el computador, lo que permite acceder directamente a las obras y de forma más aleatoria, según el interés espaciotemporal del visitante; en cambio, en el celular aparece obra por obra, siguiendo la cronología propuesta por fechas de realización de las obras, empezando siempre por la fotografía de Raúl Corrales tomada en 1960.

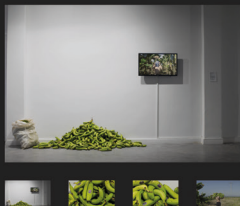
El formato y la experiencia digital de la exposición también ha distanciado a las obras de su experiencia aurática, es decir, de

la huella del autor y de su unicidad como obra de arte atada al aquí y el ahora y, por lo tanto, de una experiencia directa (Benjamin, 2008), lo que permite enfatizar en los contenidos de las piezas. La inmaterialidad de las obras en su versión digital desplaza el valor del objeto hacia una valorización a partir de su proceso y su valor simbólico (Krysa, 2006). De esta manera, lo que se exponen no son obras originales, sino su reproducción fotográfica o en video, las cuales el usuario reconstruye mediante el apoyo de textos explicativos y de las fichas técnicas, que adquieren particular importancia en este formato de exposición, porque permiten imaginar las dimensiones físicas y materiales de las obras (figura 9.7).

Dicho esto, la exhibición digital, diseñada para mostrar obras que fueron pensadas para el espacio físico, limita ciertas lecturas, en especial las materiales y corporales de las piezas, mientras que amplía otras, como los ejes temáticos de interpretación (violencias, ecosistemas e identidades). En otras palabras, y como lo explica Connor (2020), la exhibición en la web implica la *representación* de obras de arte y su objetualidad en una puesta en escena particular (el internet), que las dispone en una relación dinámica entre sí y en un contexto de red más amplio.

Sobre la reproducción de obras de arte mediante reproducciones fotográficas o en video, cabe destacar que un asunto clave en la concepción de este proyecto fue que siempre cumpliera con la ética de la labor digital. Uno de los procesos más laboriosos en la creación de *La fiebre del banano/Banana Craze* fue la solicitud formal —y por escrito— y el pago, cuando así fue requerido, de los derechos de reproducción de las piezas que constituyen el proyecto, tanto a sus autores (artistas) como a sus propietarios (coleccionistas

MILKO DELGADO



ARTIST Milko Delgado

TITLE *Apuntes conceptuales sobre el extractivismo bananero en Barú*

YEAR 2021

ARTISTS COUNTRY OF ORIGIN Panamá

DIMENSIONS Variable dimensions performance and 7:31 video-performance

MEDIUM Installation and video-performance

Credits: Private Collection, The artist would like to acknowledge the community of Puerto Arruñales, Manaca, and Barú, MAC Museo de Arte Contemporáneo de Panamá, Maestros Públicos, Adán Vallejo, and the project collaborators Elicesa Judge and Leydiara Pareda

The Panamanian artist Milko Delgado reflects on the trauma of his country's banana-growing past in his 2021 piece *Apuntes conceptuales sobre el extractivismo bananero en Barú*, comprising a video and an installation. The video shows three actions – in the first, the artist followed a long path through a banana plantation carrying a heavy bunch of bananas on his shoulder, a symbol of Panama's difficulty in shaking off the weight of its neocolonial past linked to the banana industry. In the second action, also in the plantation, Delgado covers his head with a blue plastic bag, similar to those which cover the bunches of fruit to protect them from pests and pesticides. The video shows the artist struggling for air until he cannot bear it any longer and he tears the bag open with his hands in order to be able to breathe normally, finally freed from the asphyxiating climate of the banana plantations. In the third and last action Delgado lies on the ground in the ruins of an abandoned boat on the banana plantation. Blinded by the sun, the artist is partially covered by several bananas which, like him, seem to have given up, as though they are waiting to decompose. The installation that accompanies the work consists of a pile of bananas, each with a sticker similar to those of the big banana companies. Some of the stickers show portraits of the fifteen people (including some of the artist's family members) which Delgado interviewed in Barú during the process of research for the work, including ex-workers from the plantations, activists and community leaders, independent farmers and people with health problems as a result of fumigation with agrochemicals. Other bananas have a black sticker as a symbol of mourning for those who have died due to the pesticides.

ALL ARTISTS

FIGURA 9.7. Captura de una de las fichas de las obras incluidas en *La fiebre del banano/Banana Craze*

FUENTE: <https://bananacraze.uniandes.edu.co/>

privados, museos, entre otros, según las circunstancias de cada caso), incluidos también los permisos para incluir las imágenes en la difusión del proyecto en redes sociales y en el dossier de prensa.

Teniendo en cuenta el contexto latinoamericano, donde los presupuestos para el arte y la cultura son mínimos, las herramientas y el potencial de la curaduría digital son enormes. En cierta medida, este tipo de plataformas cumplen una doble función: son exhibición y catálogo simultáneamente. El espacio virtual permite que los textos curatoriales no se limiten a ensayos cortos explicativos de cada sala, complementados con algunas fichas técnicas extendidas, sino que son la oportunidad para el desarrollo de ensayos más profundos que el visitante puede leer desde la comodidad de su computador, tableta o teléfono móvil. Así mismo, la plataforma virtual puede contar con ensayos y referencias publicados en otros espacios a los que se puede remitir, tal y como sucede en la sección de publicaciones de *La fiebre del banano/Banana Craze*.

Sobre este aspecto, cabe subrayar de nuevo la cuestión de la delicada relación entre el libre acceso a la cultura y la sostenibilidad del proyecto. Desde el inicio, las autoras de *La fiebre del banano/Banana Craze* adquirieron el compromiso ético de que todos los colaboradores (artistas, cineastas, investigadores, etc.) en los eventos públicos de acceso libre y gratuito programados en la plataforma siempre fueran remunerados por su trabajo. Además, también se asumió la responsabilidad de que para este propósito solo establecerían pactos de colaboración con entidades que fomentaran valores éticos afines a los del proyecto. Para cumplir con este objetivo, los eventos públicos de *La fiebre del banano/Banana Craze* recibieron el patrocinio de Equifruit, la mayor empresa de comercio justo de banano de Canadá. De este modo, *La fiebre*

del banano/Banana Craze ha logrado ser un proyecto a coste cero, en el que, si bien sus autoras han trabajado de manera altruista en su elaboración, tampoco les ha supuesto ningún gasto, dado que todos los gastos de producción han estado cubiertos (gracias a la Universidad de los Andes y Banana Link) y además han podido remunerar a todos los colaboradores en los eventos de la plataforma (gracias a Equifruit).

Un recorrido por las galerías virtuales

Como se explicó en la introducción, el apartado de *La fiebre del banano/Banana Craze* dedicado a la exposición está dividido en tres secciones, que evocan las galerías de un centro expositivo y que exploran las siguientes temáticas: violencias, identidades y ecosistemas. La sección de violencias reúne una serie de obras que exploran la larga historia de abusos vivida en el continente americano en relación con el monocultivo del banano, ejercidos principalmente por corporaciones multinacionales extranjeras y los Estados corruptos de las regiones donde se instauraron los enclaves. El establecimiento de las zonas bananeras, en particular en Centro América, el Caribe y los países del norte de Suramérica, nace de una serie de alianzas entre los diferentes países con las corporaciones internacionales, las cuales, a cambio de tierras para cultivar y una serie de beneficios para sus empresas, prometían el desarrollo de infraestructura a nivel local, mediante la construcción de vías, puertos y ferrocarriles, así como una modernización en términos sociales con la construcción de colegios, la electrificación de zonas apartadas, etc. (Striffler y Moberg, 2003). Sin

embargo, dichos procesos de modernización se limitaron al beneficio de los enclaves y no a los sectores más amplios de la sociedad. Además, las intervenciones de estas compañías —en especial la United Fruit Company y la Standard Fruit Company— se extendieron al ámbito político, como en el conocido episodio de apoyo de la CIA al golpe de Estado contra el presidente de Guatemala Juan Jacobo Arbenz, en 1954 (García Ferreira, 2006), o la infame masacre de las bananeras, que tuvo lugar en Colombia en 1928, con el apoyo del Estado en defensa de la United Fruit Company (Herrera Soto y Romero Castañeda, 1979). Pero las violencias vividas por los trabajadores de las compañías y las comunidades locales no se limitaron al ámbito político. En la cotidianidad, los trabajadores de las bananeras trabajan aún hoy sin medidas sanitarias suficientes que protejan sus cuerpos de la exposición a pesticidas y al arduo trabajo manual que llevan a cabo, los salarios son apenas simbólicos y las condiciones laborales rayan en la explotación.

Las obras agrupadas en esta sección reflexionan sobre todos estos tipos de violencia, pero no se limitan al señalamiento de una violencia ejercida sobre los cuerpos humanos, sino también sobre la naturaleza o sobre las diferentes identidades que se desprenden de la intervención directa de las compañías bananeras en los países donde se instalaron. De ahí que el hecho de que se trate de una exposición virtual en vez de presencial enfatice en la multiplicidad de temas que presenta una obra de arte al ponerla en diálogo con otras obras según diversos criterios (temáticos, geográficos, cronológicos, etc.). En ese sentido, el visitante de *La fiebre del banano/ Banana Craze* se encuentra con piezas que están *simultáneamente* en más de una galería, lo que sería imposible en el espacio físico de una exposición. El usuario puede observar esto con claridad

al poner el cursor sobre el ícono de la obra, lo que activa un texto que enuncia los ejes temáticos a los que pertenece. Por ejemplo, la obra *Apuntes conceptuales sobre el extractivismo bananero en Barú* (2021), del artista panameño Milko Delgado, se encuentra tanto en la galería de violencias como en la de ecosistemas. La obra consiste en una videoperformance y una instalación. En la primera, el artista realiza tres acciones que apelan al trabajo físico dentro de las bananeras y, a su vez, reflexionan simbólicamente sobre el peso neocolonial relacionado con la industria de esta fruta. En la instalación, el artista coloca una pila de bananos sobre el suelo, al lado de la videoperformance descrita. Los bananos llevan un adhesivo similar al de las etiquetas que les ponen a las frutas en las empresas bananeras, con la marca y su lugar de origen. Delgado reemplaza estas imágenes y textos con los retratos de quince personas (entre ellas familiares suyos) que el artista entrevistó en el proceso de investigación de la obra, dentro de los que se encuentran antiguos trabajadores de las plantaciones, activistas y líderes comunitarios, agricultores independientes y personas con afectaciones de salud debido a la fumigación con agroquímicos. Otros bananos llevan un adhesivo negro como símbolo de luto por los fallecidos a causa de los pesticidas. Con esto en mente, es evidente que Delgado reflexiona sobre la historia violenta de lo que se vive en las plantaciones a nivel físico, así como sobre el impacto que las fumigaciones y los pesticidas tienen en el medio ambiente.

La sala de ecosistemas se ocupa justamente de esta problemática, aunque no se limita a las afectaciones de los agroquímicos sobre el medio ambiente, sino que también indaga el impacto sobre el territorio local y las implicaciones ecológicas de los monocultivos a gran escala. Por ejemplo, *Frutales Carmen Miranda* (2011), del

artista colombiano Alberto Baraya, hace un estudio del paisaje artificial contemporáneo a través de la apropiación del sistema de clasificación botánica implantado por Linneo en el siglo XVIII, el cual fue adoptado por José Celestino Mutis en la célebre Real Expedición Botánica de la Nueva Granada para estudiar la diversidad de especies americanas. En esta obra, Baraya no solo reflexiona sobre las implicaciones de la poca diversidad de especies de banano que se cultivan actualmente y que han propiciado la epidemia de la enfermedad de Panamá, sino que explora también la construcción de identidades sexistas y tropicalizantes, como es el caso de la actriz y cantante brasileña Carmen Miranda. Por ello, esta obra forma parte tanto de la sección de ecosistemas como de la sala de identidades.

La última galería recoge una serie de obras que exploran la construcción de identidades mediadas por el banano, que van desde una identificación con la nación, hasta el banano entendido como vehículo de identidad cultural, laboral, racial y sexual. Al tratarse del recurso natural más explotado en estos territorios, el banano encarna de una manera casi literal la tierra entendida como nación, como lo ejemplifica la obra *Chiquita* (2013), de la artista ecuatoriana María José Argenzio. En otras obras, esta identificación con la fruta se explora a través de inscripciones sobre la piel del banano, como un tatuaje que se realiza metafóricamente también sobre la piel de los habitantes de los países productores. La relación con la fruta es tan cercana y problemática que, a principios del siglo xx, se acuñó la expresión “república bananera” para hacer referencia a gobiernos corruptos, caóticos y déspotas a partir del referente latinoamericano. Recientemente, este significado ha sido invertido para referirse a los Estados Unidos, a través de piezas

como *Banana flag* (2018-2020), del uruguayo Luis Camnitzer. Artistas de la comunidad *latinx* (migrantes y descendientes de personas latinoamericanas basados en los Estados Unidos) han usado el referente del banano, ya sea como seña de orgullo de una identidad que los une o como una forma de criticar la homogeneización de una comunidad muy diversa.

En últimas, lo que aquí se quiere señalar es que la exposición virtual con salas o galerías organizadas bajo ejes temáticos específicos no solo expande la multiplicidad de perspectivas interpretativas desde las que una obra de arte puede examinarse, sino que también posibilita y abre caminos para múltiples narrativas desde una misma estructura. Así mismo, esta curaduría digital reúne en un único espacio obras que pertenecen a colecciones tanto públicas como privadas, convirtiéndose en un lugar de circulación público de piezas a las que de otra forma sería imposible acceder. En ese sentido, este proyecto se cruza con el campo de las humanidades públicas digitales, al dar acceso a un conjunto de datos dispersos a una comunidad amplia más allá de la academia.

Variabilidad digital

Además de la exposición, *La fiebre del banano/Banana Craze* incluye una serie de herramientas que permiten aproximarse a todas las obras más allá de los ejes temáticos de cada galería (violencias, ecosistemas e identidades). Como expone Lev Manovich en su libro *The language of new media* (2001), la variabilidad y las posibilidades que ofrecen este tipo de plataformas radican en la multiplicidad de accesos posibles a un mismo conjunto de

datos. En el caso de *La fiebre del banano/Banana Craze*, las obras en la plataforma pueden ser investigadas por artista (figura 9.2), siguiendo criterios alfabéticos a través de un listado que incluye a todos los artistas por su apellido; por piezas individuales, mediante un mosaico que muestra el conjunto de las obras que conforman la muestra, a lo largo de los años de producción de las piezas (figura 9.3); o por medio del lugar de origen del artista que las produce (figura 9.4).

La herramienta de la línea del tiempo permite visualizar el año de creación de cada una de las piezas. La obra más temprana del proyecto es de 1960. Se trata de la mencionada fotografía de Raúl Corrales, *Caballería*, que muestra a un grupo de rebeldes a caballo celebrando la expropiación de la última plantación de la United Fruit Company en Cuba tras el comienzo de la Revolución. Entre las obras más recientes del repertorio³ (hay cuatro obras del 2021) está *Plátanos amarillos, after Francisco Oller*, del artista latinx Justin Favela, que reinterpreta una obra icónica de la pintura puertorriqueña de finales del siglo XIX en clave contemporánea, usando la estética de una popular piñata de papel. Mediante la línea del tiempo, el usuario puede analizar cómo han evolucionado los medios plásticos a lo largo de los últimos sesenta años: desde la fotografía documental en blanco y negro, hasta la videoperformance, pasando por la escultura y la instalación; y desde el arte activista (expresado desde el póster propagandístico, por ejemplo), hasta el arte sobre los debates identitarios

3 Este texto fue escrito en el 2022 y sabemos que la base de datos de la plataforma ha crecido desde ese entonces; por tratarse de un trabajo digital activo, puede crecer y cambiar en el tiempo.

(feminismos, racismos, etc.), pasando por el auge de los conceptualismos en relación con las dictaduras en Latinoamérica. Más aún, la cronología en *La fiebre del banano/Banana Craze* ofrece la posibilidad de que el usuario saque sus propias conclusiones respecto a un determinado periodo de la historia. Por ejemplo, en nuestra propia investigación, la variabilidad digital nos ha permitido entender cómo en la primera década del siglo XXI priman las obras que hacen referencia a la identidad racial (véanse las piezas de Miguel Luciano y Jean-Francois Boclé sobre el racismo hacia las personas afrodescendientes en Estados Unidos y el Caribe), en contraposición con piezas más recientes producidas entre los años 2010 y 2020, en las que abunda la preocupación por el medio ambiente (como en las obras de Milko Delgado y Óscar Figuera Chávez acerca de los efectos perniciosos para la salud del uso de pesticidas en las plantaciones).

En el caso de la herramienta del mapa sucede algo similar. El visitante de *La fiebre del banano/Banana Craze* puede navegar por un mapamundi en el que están marcados los países de origen de los artistas participantes en el proyecto. Al clicar en cualquiera de estos países, una nueva pestaña muestra todas las obras producidas en ese contexto, lo cual permite, a través de otro hipervínculo, consultarlas una a una (figura 9.8). De nuevo, esto posibilita que el visitante a la página web desarrolle su propia investigación y saque sus propias conclusiones. Por ejemplo, mediante esta herramienta hemos visto cómo en el caso de Colombia la producción artística relacionada con el banano está especialmente vinculada con la historia de violencias del país (desde la masacre de las bananeras en 1928, hasta el conflicto armado interno que llega hasta el presente), tal como lo reflejan las obras de Doris Salcedo, Sair

García y José Alejandro Restrepo; o cómo en República Dominicana abundan las piezas acerca del influjo del estereotipo de las “repúblicas bananeras” en la construcción de la identidad nacional contemporánea, como en las piezas de Yunior Chiqui Mendoza, Julio Valdez y Radhamés Mejía.

La organización y visualización de información a través de la línea del tiempo y del mapa permite que los patrones y las relaciones entre artistas, lugares y momentos se hagan más visibles. De este modo, la ampliación de los conjuntos de datos que se pueden cotejar posibilita que los usuarios que lo deseen definan un contexto más amplio para sus objetos de estudio, sin dejar de centrarse en la lectura detallada de artistas y objetos individuales. En otras palabras, una de las contribuciones del proyecto es la posibilidad de examinar una gran cantidad de obras bajo parámetros organizados en una diversidad de formas, lo que facilita nuevas conexiones y relaciones entre las piezas. Como lo explicamos, este formato permite establecer nuevas conclusiones sobre las obras como un conjunto y no en términos individuales. Así, el mapa y la línea del tiempo revelan intereses comunes según los lugares de proveniencia de los artistas, por ejemplo, la relación del banano como un elemento identitario en las comunidades diaspóricas que se refleja en las piezas producidas en Estados Unidos o el énfasis que ponen los artistas en una preocupación por el medio ambiente en las obras más recientes. Como lo afirma Pamela Fletcher (2015), citando a Richard White, “La visualización y la historia espacial no se tratan de producir ilustraciones o mapas para comunicar cosas que se han descubierto por otros medios. Es un medio de hacer investigación en sí mismo” (párr. 23).



FIGURA 9.8. Captura de pantalla de pestaña que lista las obras producidas en Colombia

FUENTE: <https://bananacraze.uniandes.edu.co/>

Entre la pedagogía y lo público

Dos elementos clave de *La fiebre del banano/Banana Craze* son su carácter pedagógico y su dimensión pública. El carácter pedagógico se puede extrapolar a dos niveles. Primero, la participación activa en calidad de pasantes de un grupo de estudiantes de la Universidad de los Andes, quienes apoyaron en la investigación sobre las obras y en la consecución de los derechos de autor de las imágenes y fotografías de las piezas. Los estudiantes además escribieron textos cortos para la sección del blog, que desarrollaron con el acompañamiento de las coautoras del proyecto. Los ensayos del

blog, todos escritos en español, están firmados por cada uno de los estudiantes. Esto no solo amplía el número de voces que participan en *La fiebre del banano/Banana Craze*, sino que demuestra la índole horizontal de la investigación, al darle espacio y autoría a los alumnos de la universidad. El blog forma parte de la sección de publicaciones de la plataforma (figura 9.5). Dividida en dos apartados, blog y artículos académicos, las publicaciones expanden el tema de investigación del proyecto, ceñido a la ubicuidad del banano en el contexto latinoamericano, para conectarlo con otros territorios y momentos de la historia del arte.

El segundo aspecto pedagógico del proyecto se relaciona con la intención de la plataforma: educar al público general sobre el impacto del monocultivo del banano en América Latina. Sin embargo, esto no se logra simplemente mediante el montaje de la página web, sino a través de una serie de eventos públicos, todos *online*, los cuales han posibilitado la participación de una audiencia amplia que se puede conectar y aprender desde cualquier lugar del mundo. El registro de todos estos eventos se encuentra en la sección de eventos de la página, que ha permitido establecer vínculos entre el proyecto y sus autoras y otras instituciones y pensadores (artistas, cineastas, académicos, entre otros) (figura 9.6). Se trata de la sección más activa desde que salió la plataforma, dado que aquí se publican los videos y la información relevante de cada uno de los eventos acaecidos desde el lanzamiento del proyecto. Estos eventos son de acceso libre y gratuito; de igual forma, los videos que quedan como registro, testimonio y archivo de cada uno de estos encuentros virtuales se encuentran disponibles de manera libre y gratuita para usuarios de cualquier lugar del mundo. Los eventos hasta la fecha incluyen conversatorios con artistas incluidos en la

exposición virtual, talleres impartidos por artistas, ponencias académicas por parte de profesores universitarios y ciclos de cine documental acerca de las bananeras en Latinoamérica. Además, las autoras de *La fiebre del banano/Banana Craze* han sido invitadas a presentar el proyecto en instituciones internacionales, incluidas las siguientes universidades: University of the Arts Philadelphia, Bard College, Universidad del Norte, Universidad Complutense y University of Essex. En ese sentido, *La fiebre del banano/Banana Craze* puede considerarse un proyecto en el que se entrecruzan la curaduría en la web y las humanidades públicas digitales; en otras palabras, involucra otros públicos más allá del académico (Schroeder, 2017), y diseña y establece una relación con una audiencia amplia, teniendo en cuenta sus intereses y necesidades (Rodríguez Ortega, 2016).

La realización de numerosos eventos por parte de las autoras del proyecto (diez en total, durante los nueve meses que lleva la plataforma activa; seis propios y cuatro por invitación) demuestra que *La fiebre del banano/Banana Craze* es un proyecto de investigación a largo plazo, que no se limita ni se agota en la exposición, sino que está en continuo crecimiento y en constante búsqueda de participantes externos que amplíen las lecturas que posibilita. Así, la investigación es actualizada, renovada y expandida de forma permanente, con multitud de lecturas nuevas que la enriquecen y la relacionan con otros campos, más allá de la historia del arte latinoamericano. Es importante notar que el proyecto crece no solo por la participación de los colaboradores invitados a presentar en los programas públicos, sino también por la interacción fundamental del público, que interviene activamente en los turnos de preguntas, debates y conversaciones que se generan después de cada presentación.

Retos y desafíos en el diseño de *La fiebre del banano/Banana Craze*

Si bien es cierto que una curaduría es en muchos aspectos un trabajo colectivo entre artistas, museógrafos, diseñadores, instaladores y curadores, el trabajo en curaduría digital y humanidades públicas digitales implica una aproximación colectiva más amplia, que se aleja de la investigación tradicional. El método lineal de investigación académica que sigue más o menos los pasos de investigar, analizar, escribir y publicar se reemplaza por un proceso iterativo, más bien circular, que conlleva no solo investigar y escribir, sino también hacer pruebas de la plataforma, revisar prototipos y corregir sobre estos (Baca y Helmreich, 2017). Para el caso de una exhibición digital como *La fiebre del banano/Banana Craze*, se requirió de una estrategia basada en un equipo de colaboradores que tenían diferentes experticias.

En primera instancia, estaban las curadoras, quienes conceptualizaron la exposición, investigaron sobre el tema, eligieron las obras y escribieron sobre cada una de ellas; a partir de esto generaron una red de conexiones entre las piezas elegidas por medio de los ejes temáticos. En segundo lugar, se contó con grupo de estudiantes que apoyaron tanto en el proceso de investigación como en la consecución de los permisos de reproducción de las imágenes. Este es uno de los retos más grandes para hacer una exhibición digital de arte, pues implica conseguir los permisos morales y patrimoniales de cada una de las obras que forman parte de la muestra. En una exhibición de más de cien piezas, este es uno de los desafíos más importantes. Paralelo al trabajo en contenidos y autorizaciones legales, se trabajó con un equipo de

diseñadores externo a la universidad, el cual se encargó de traducir la investigación a un lenguaje visual atractivo y claro para el público general en el espacio del internet⁴. La paleta de colores elegida y el lenguaje gráfico de la página web parten del color del banano en su estado maduro (amarillo y negro) y lo adaptan a una estética contemporánea, con el fin de atraer la atención con colores como el amarillo neón y el gris oscuro. Es importante señalar que en ningún momento se utiliza la iconografía de la fruta como elemento gráfico dentro de la web, más allá de las imágenes de las obras. Con el equipo de diseñadores se discutieron y plantearon las diferentes maneras en las que se podía acceder a los contenidos y cómo podían ser visualizados de diferentes formas: geográfica, cronológica, alfabética o temáticamente, para el caso de las obras, por ejemplo.

Como primera fase del desarrollo de la página se hizo la propuesta gráfica, que integraba todos los puntos discutidos anteriormente. Seguido a esto, se montó la plataforma en un *host* de la Universidad de los Andes, utilizando una plantilla de WordPress, la cual fue adaptada a las necesidades del proyecto con base en el diseño gráfico que se propuso. A la plantilla se le sumaron algunos *plugins*; por ejemplo, Interactive Geo Maps en la sección del mapa y Cool Timeline en la cronología, que sirven como apoyo para el desarrollo de los contenidos y formas de visualización conceptualizadas por las autoras. Una vez terminado el proceso de diseño, tanto gráfico como de programación, las autoras tuvieron unas

4 Los diseñadores fueron remunerados con una beca que se obtuvo para el desarrollo de un proyecto de humanidades digitales, otorgada por el Centro de Investigación y Creación de la Universidad de Los Andes.

sesiones de entrenamiento con el equipo diseñador para aprender a manejar los contenidos desde WordPress de manera independiente. También cuentan con una serie de recursos, como instructivos y videos, para hacer la autogestión del proyecto sostenible a largo plazo. Además, se realiza un mantenimiento anual a la página con el equipo de diseño, con el fin de mantenerla actualizada y en funcionamiento sin ningún contratiempo.

Teniendo en cuenta la tensa relación que establece la investigación entre el norte y el sur global frente al monocultivo del banano, es fundamental que la página sea bilingüe, de tal forma que las reflexiones que se generan mediante la investigación sean accesibles tanto para el público angloparlante como para el hispano. De esta forma, se requirió de la colaboración de un grupo de traductores, quienes se encargaron de traducir todos los textos al inglés⁵. La revisión de todos los aspectos de la página, así como el progresivo crecimiento de sus contenidos, está ahora a cargo de las curadoras. El poder continuar con el aumento de contenidos y la autogestión del proyecto digital fue un factor fundamental desde el principio.

Quizá un elemento importante que no se tuvo en cuenta y que resulta primordial en los proyectos de humanidades digitales tiene que ver con la forma en la que la plataforma seguirá vigente a pesar de los cambios tecnológicos que suceden tan rápidamente en el ámbito digital. ¿Qué compromisos se deben hacer a la hora de pensar en el archivo de la investigación y su permanencia

5 La traducción de todos los textos en *La fiebre del banano/Banana Craze* se realizó gracias al trabajo voluntario del banco de traductores de la organización sin ánimo de lucro Banana Link, que se dedica a la difusión y el fomento del comercio justo del banano.

como una plataforma vigente en la red? ¿Cómo volver sostenible en términos tecnológicos la plataforma a largo plazo, sin comprometer componentes de diseño y contenido? Estas son preguntas que quedan por resolver en este caso de estudio.

***La fiebre del banano/Banana Craze* y las humanidades digitales en América Latina**

Entre los crecientes repositorios, archivos y bases de datos digitales en América Latina, donde cada vez es más frecuente encontrar también exposiciones virtuales y asociaciones gremiales de historiadores, antropólogos, sociólogos, etc., *La fiebre del banano/Banana Craze* contribuye al campo de las humanidades digitales al aportar un formato híbrido entre la exposición virtual, el archivo digital y las humanidades públicas digitales. Esta combinación de elementos en una misma plataforma (y las posibilidades que generan) ya supone una aportación tecnológica poco habitual. Algunos referentes comparables que se lanzaron más o menos al mismo tiempo que *La fiebre del banano/Banana Craze* son: la exposición en línea *Pensar todo de nuevo* (2020)⁶, curada por Andrea Giunta para la web de la galería Rolf Art, en Argentina, y el proyecto *Entre-ríos* (2020), gestionado por Lisa Blackmore, Emilio Chapela y Diego Chocano⁷.

6 Véase <https://rolfart.com.ar/viewing-rooms/pensar-todo-de-nuevo/>

7 Véase <https://entre-rios.net/>

Pensar todo de nuevo consiste en una exposición virtual que explora obras de artistas latinoamericanos representados por la galería, que examinan la crisis del presente a través de obras del pasado. La muestra responde a la pandemia global que inició en marzo del 2020, y la selección de imágenes de obras de arte y los textos curatoriales se complementan con una serie de programas públicos en video, los cuales se realizaban cada vez que se abría un nuevo capítulo de la muestra. La exposición se narró en cuatro capítulos, en una interfaz que intercala las obras con textos breves de la curadora y otros invitados. A diferencia de *La fiebre del banano/Banana Craze*, esta exposición se activó por solo unos meses y el recorrido que presenta solo se puede navegar en forma lineal, siguiendo los cuatro ejes propuestos por la curadora, y no hay manera de reorganizar las piezas o de buscarlas bajo otros parámetros.

Entre-ríos, por su parte, es una plataforma web multilingüe que examina la relación entre los cuerpos de agua y los cuerpos humanos a través de la obra de artistas contemporáneos latinoamericanos, charlas virtuales y activaciones *in situ*. Al igual que *La fiebre del banano/Banana Craze*, el proyecto está articulado bajo unos ejes curatoriales que los gestores del proyecto llaman *canales*. Además, los proyectos incluidos en cada uno de los canales se articulan con los otros ejes del proyecto, el cual incluye un mapa para buscar por geolocalización los lugares de producción de las obras, textos, charlas y publicaciones incluidas en la plataforma, una mediateca y diferentes formas de búsqueda mediante etiquetas que pueden usarse entrando a la sección del archivo o en la búsqueda general. Este proyecto es uno de los más cercanos a *La fiebre del banano/Banana Craze*, pues involucra una multiplicidad de voces y utiliza la curaduría en la web

y las humanidades públicas digitales para entablar un diálogo horizontal, más allá de la academia, sobre temas actuales que involucran la relación entre arte contemporáneo y medio ambiente. No obstante, *Entre-ríos* se distingue de *La fiebre del banano/Banana Craze* en que la navegación no jerarquizada parece estar orientada hacia la producción de contenidos culturales para ser consumidos por un público especializado, mientras que *La fiebre del banano/Banana Craze* plantea una estructura y una serie de recursos (como la bibliografía) con fines didácticos, dirigidos a un público interesado en la historia cultural, sin ser necesariamente consumidor habitual de arte contemporáneo.

Además, a diferencia de *Entre-ríos*, *La fiebre del banano/Banana Craze* se enfoca en una problemática específicamente latinoamericana, cargando así de sentido la categoría de “arte latinoamericano”, utilizada de forma frecuente para agrupar prácticas y contextos demasiado diversos entre sí. *La fiebre del banano/Banana Craze* identifica un problema netamente latinoamericano que afecta a todos los países de la región: desde aquellas naciones que sufrieron el intervencionismo imperialista de la United Fruit Company, hasta las regiones asoladas por los perniciosos efectos medioambientales y de la salud pública del monocultivo, pasando por la frecuente asociación entre la identidad latinoamericana y el banano a raíz fundamentalmente del estereotipo xenófobo, racista y clasista de las “repúblicas bananeras”. No hay lugar de América Latina que no se vea afectado por el peso cultural del banano.

La fiebre del banano/Banana Craze reúne varias de las funciones incluidas en estos ejemplos de la curaduría en la web, lo que da la posibilidad de una triple fórmula que combina la exposición virtual, la búsqueda por parámetros en la base de datos y los

programas públicos en video, lo cual convierte a *La fiebre del banano/Banana Craze* en un proyecto de investigación en curso que alberga distintos repositorios de información (la exposición, la bibliografía, las publicaciones, los videos, etc.), en aras de ofrecer la mayor y más variada cantidad de materiales sobre un mismo tema (en este caso, el impacto del monocultivo del banano en el arte de Latinoamérica).

No obstante, temática y metodológicamente el proyecto también ofrece parámetros y lecturas muy novedosos, dado que se ocupa de cuestiones poco estudiadas en el ámbito general de la historia del arte latinoamericano. Para empezar, *La fiebre del banano/Banana Craze* es una plataforma bilingüe español-inglés, además de acceso libre y gratuito. Esto supone que todos los contenidos no solo están disponibles para cualquiera con una conexión a internet, sino que participan del circuito global anglófono de los estudios culturales, cuya red e impacto, como es sabido, superan a aquellos en otras lenguas.

Segundo, *La fiebre del banano/Banana Craze* incluye obras de arte que superan los límites del canon de la historia del arte de la región: la plataforma cuenta tanto con artistas reconocidos y de trayectoria establecida como con artistas emergentes, incluye a artistas de países insuficientemente estudiados hasta ahora (por ejemplo, de Centroamérica y Ecuador) y contempla la producción artística realizada por migrantes y en la diáspora, lo que amplía los límites del estudio meramente geográfico de la creación cultural. Más aún, los ejes temáticos de la exposición virtual y las herramientas de búsqueda (cronológica y geográfica) de la sección de archivo permiten establecer diálogos entre

las preocupaciones de artistas de contextos, momentos y lugares muy diferentes.

Por último, *La fiebre del banano/Banana Craze* está abierto y en constante contacto con otras disciplinas culturales más allá de las bellas artes. En especial a través de los programas públicos, el proyecto contempla la creación vernácula, la cultura popular, la literatura, el cine, el *artivismo* ecologista y la historia sociopolítica como elementos complementarios relacionados entre sí y que se enriquecen mutuamente. De esta manera, la plataforma incrementa de forma exponencial los públicos a los que se dirige, quedando a disposición de creadores culturales, investigadores, profesores, periodistas, amantes de la cultura, público general, entre otros.

Conclusión

La fiebre del banano/Banana Craze ha tenido una gran acogida tanto en el ámbito académico como en la cultura digital más divulgativa. En el primer año desde su lanzamiento, ha recibido más de once mil visitas, de más de cuarenta y cinco países, con una media de ochocientas visitas al mes. La plataforma ha realizado más de una docena de eventos públicos en línea que han recibido más de cuatrocientos asistentes. El proyecto ha sido reseñado en la prensa internacional en más de veinte medios de comunicación (incluidos diarios, revistas especializadas, televisión y radio) y ha sido galardonado con el premio Garfinkel Prize for Digital Humanities, de la American Studies Association, y con el tercer puesto del Digital Humanities Award.

Como demuestra *La fiebre del banano/Banana Craze*, el propósito de las humanidades digitales no es comunicar investigaciones que se han desarrollado por otros medios a través de herramientas digitales. Las humanidades digitales son un medio de investigación en sí mismo, porque generan preguntas que de otro modo no se plantearían, revelan relaciones históricas que de otra manera pasarían desapercibidas y socavan, o corroboran, las historias sobre las que construimos nuestras propias versiones del pasado. Por ejemplo, mediante el uso de la curaduría en la web se establecieron relaciones entre las obras que no eran visibles inicialmente. Así, se visualizó el reciente interés ecocrítico de algunos artistas, en contraposición a las reflexiones de carácter más político de las obras producidas en las décadas de las dictaduras latinoamericanas. También fue posible incluir una obra en más de una “sala” de forma simultánea, aprovechando así la multiplicidad de capas de lectura que las obras artísticas ofrecen, algo que sería imposible en una curaduría presencial. Además, la curaduría en la web permitió la inclusión de un número muy grande de obras provenientes de cualquier lugar del mundo, algo que en una curaduría física se habría tenido que limitar por los costos de envío, seguros y espacio de exhibición. Y, finalmente, el hecho de que el proyecto forme parte del espacio público de la web hace que cualquier persona con internet, en cualquier lugar del mundo, tenga acceso a la muestra.

La activación de la página mediante los diferentes eventos que se han llevado a cabo, de carácter gratuito y *online*, ha hecho de este proyecto un ejemplo de humanidades públicas digitales, al llevar la producción académica a un público amplio y variado, además de convertirse así en una plataforma colaborativa y de carácter pedagógico. A través de una presentación no lineal-argumental

de la información en la exposición virtual, la posibilidad de que el usuario filtre los datos según sus intereses y la constante adición de nuevos contenidos en la sección de publicaciones y eventos, *La fiebre del banano/Banana Craze* permite que la investigación sobre el legado cultural de este monocultivo en América Latina sea construida de forma colaborativa y continúe creciendo de manera orgánica, como un recurso educativo y cultural al acceso de todos.



Referencias

- Altendorf, A. (2019). Bananas and major tropical fruits in Latin America and the Caribbean: The significance of the region to world Supply. En Organización de las Naciones Unidas para la Alimentación y la Agricultura (FAO) (Ed.), *Food Outlook: Biannual Report on Global Food Markets* (pp. 73-76). FAO.
- Baca, M. y Helmreich, A. (2017, 15 de febrero). Introducing three digital art history case studies. *Getty Iris Blog*. https://blogs.getty.edu/iris/dah_baca_helmreich/
- Banana Link. (s.f.). *All about bananas*. <https://www.bananalink.org.uk/all-about-bananas/>
- Benjamin, W. (2008). The work of art in the age of mechanical reproduction. En M. Jennings, B. Doherty y T. Levin (Eds.), *The work of art in the age of its technological reproducibility, and other writings on media* (pp. 19-55). Massachusetts Institute of Technology Press.

- Chapman, P. (2009). *Bananas: How the United Fruit Company shaped the world*. Canongate.
- Coleman, K. (2016). *A camera in the Garden of Eden: The self-forging of a banana republic*. University of Texas Press.
- Connor, M. (2020, 13 de mayo). Curating online exhibitions. Part 1: Performance, variability, objecthood. *Rhizome*. <https://rhizome.org/editorial/2020/may/13/curating-online-exhibitions-pt-1/>
- Dietz, S. (1998). Curating (on) the web: The museum in an interface culture. *Museums and the Web*. https://www.museumsandtheweb.com/mw98/papers/dietz/dietz_curatingtheweb.html
- Egger, B. y Ackermann, J. (2020, julio). Meta-curating: Online exhibitions questioning curatorial practices in the postdigital age. *History of Digital Art*, 5, 19-32. <https://doi.org/10.11588/dah.2020.5.72123>
- Enloe, C. (1990). *Bananas, beaches & bases: Making feminist sense of international politics*. University of California Press.
- Fletcher, P. (2015, 18 de junio). Reflections on digital art history. *Re-Views: Field Editors' Reflections*, <https://doi.org/10.3202/caa.reviews.2015.73>
- Galina, I. (2011). ¿Qué son las humanidades digitales? *Revista Digital Universitaria*, 12(7), 3-9.
- García Ferreira, R. (2006). La CIA y el exilio de Jacobo Árbenz. *Perfiles Latinoamericanos*, 13(28), 59-82. http://www.scielo.org.mx/scielo.php?script=sci_arttext&pid=S0188-76532006000200003
- Ghidini, M. (2019). Curating on the web: The evolution of platforms as spaces for producing and disseminating web-based art. *Arts*, 8(78), 2-27. <https://doi.org/10.3390/arts8030078>

- Herrera Soto, R. y Romero Castañeda, R. (1979). *La zona bananera del Magdalena: historia y léxico*. Imprenta Patriótica del Instituto Caro y Cuervo.
- Jaskot, P. (2013). Review of *Debates in the Digital Humanities*. *Visual Resources*, 29 (1-2), 140.
- Jenkins, V. (2020). *Bananas: An American history*. Smithsonian Books.
- Klein, S. (2020). *The fruits of empire: Art, food, and the politics of race in the age of American expansion*. University of California Press.
- Koepfel, D. (2008). *Banana: The Fate of the Fruit that Changed the World*. Plume.
- Krysa, J. (2006). *Curating immateriality: The work of the curator in the age of network systems*. Autonomedia.
- Manovich, L. (2011). *The language of new media*. The MIT Press.
- Meeks, E. (2011). More networks in the humanities or did books have DNA? *Digital Humanities Specialist*. <https://dhs.stanford.edu/visualization/more-networks/>.
- Rodney, S. (2022). The Getty, Google Arts, & Banana Craze create three online shows worth a visit. *Hyperallergic*. <https://hyperallergic.com/703153/three-online-shows-worth-a-visit/>
- Rodríguez Ortega, N. (2016, 11 de marzo). Las humanidades digitales: Un marco para la reflexión crítica sobre la cultura. *Dixit*. <https://dixit.iarthislab.eu/las-humanidades-digitales-una-marco-de-reflexion-critica-sobre-la-cultura-i-cshdsur/>
- Rodríguez Ortega, N. (2018, noviembre). Five central concepts to think of digital humanities as a new digital humanism project.

- Artnodes, (22). <https://raco.cat/index.php/Artnodes/article/view/n22-rodriguez>
- Schroeder, R. (2017, 9 de mayo). What is public humanities? *Day of Public Humanities*. <https://dayofph.wordpress.com/what-is-public-humanities/>
- Soluri, J. (2006). *Banana Cultures: Agriculture, consumption, and environmental change in Honduras and the United States*. University of Texas Press.
- Striffler, S. y Moberg, M. (2003). *Banana wars: Power, production and history in the Americas*. Duke University Press.
- Svensson, P. (2017). *Envisioning the digital humanities: The alliance of digital humanities organizations*. <http://www.digitalhumanities.org/dhq/vol/6/1/000112/000112.html>
- Whitelaw, M. (2015). Generous interfaces for digital cultural collections. *Digital Humanities Quarterly*, 9(1). <http://www.digital-humanities.org/dhq/vol/9/1/000205/000205.html>

Movimientos sociales y humanidades digitales: el estallido del 18 de octubre en Chile

Para citar este capítulo:

<http://doi.org/10.51573/Andes.9789587987362.9789587987379.9789587987867.10>

DAVID BORDONABA-PLOU

Introducción

{ La participación política se ha entendido en la literatura especializada principalmente en relación con los procesos electorales (Gallego, 2014), es decir, se centra en cuestiones como los patrones específicos de voto de ciertos grupos o en los porcentajes de abstención. Sin embargo, como Dalton *et al.* (2003) subrayan, las formas de participación política han incrementado durante las últimas décadas. Estas han ayudado a potenciar la participación y, en muchos sentidos, han hecho las sociedades actuales más justas e igualitarias. Entre estas nuevas formas de participación política pueden destacarse las siguientes: el activismo y la participación política en las redes (van Laer y van Aelst, 2009; Margetts *et al.*, 2016); los “presupuestos participativos”¹ (Baiocchi y Ganuzza, 2017); la firma de peticiones (Dalton, 2017);

1 Esta traducción y las que aparecen a continuación son propias.

el “consumo político” (Micheletti, 2002; Stole y Micheletti, 2013; Zorell, 2019) o “consumocracia” (Novo, 2014), y la participación en manifestaciones y protestas (Dalton, 2017; Ortiz, 2015).

Durante la última década ha habido una gran cantidad de movimientos sociales en el mundo. La Primavera Árabe en el 2010, el movimiento Occupy Wall Street en Estados Unidos en el 2011 o el 15M en España en el 2012 son solo algunos de ellos. Latinoamérica no ha sido una excepción, ya que pueden constatare movimientos de este tipo en países como México, Venezuela, Ecuador o Uruguay (Almeida y Cordero, 2015; Munck, 2020). En este trabajo me centraré en uno de estos movimientos, el estallido social del 18 de octubre de 2019 (18O) en Chile². El 18O fue una serie de manifestaciones multitudinarias y disturbios que se originaron en la capital, Santiago de Chile, pero rápidamente se extendieron a lo largo de todo el país. El detonante del estallido fue el aumento del precio del transporte público, lo que provocó que miles de estudiantes llevaran a cabo actos de evasión masiva, lo cual a su vez llevó a que se iniciaran las primeras movilizaciones masivas. Sin embargo, el 18O solo puede entenderse como la culminación de una serie de movimientos sociales previos producidos por el creciente malestar y la sensación de desprotección que sentía la ciudadanía. Jiménez-Yañez (2020) señala los siguientes movimientos como precedentes del 18O: “Mochilazo (2001), la Revolución de los Pingüinos (2006), Marcha de los Paraguas (2011); Niunamenos (2016); Ola Feminista o Mayo feminista (2018) y #EvasiónMasiva (2019)” (p. 952). Casi un mes después, luego de producirse grandes pérdidas materiales y, lo que es peor, humanas, el 15 de

2 Para un análisis más detallado, véase Aguilera (2002).

noviembre las distintas fuerzas políticas llegaron a un acuerdo por la paz social, cuyo objetivo era iniciar un proceso constituyente para poder poner fin a las manifestaciones.

En el caso chileno, al igual que en el resto de Latinoamérica, los especialistas de disciplinas como la ciencia política o la filosofía política han estudiado los movimientos sociales desde perspectivas como la recesión democrática (Bianchi *et al.*, 2015), la desigualdad social (Justino y Martorano, 2016) o la crisis medioambiental (Cartagena, 2015). Aparte de estos, la cuestión de la participación ciudadana ha sido uno de los temas centrales en la discusión. Sin embargo, se cuenta con poca evidencia de la influencia que estos movimientos tienen en la clase política dirigente.

El objetivo del presente trabajo es plantear que las humanidades digitales (HD) nos proporcionan excelentes herramientas para estudiar los movimientos sociales, uno de los temas más discutidos en ciencias políticas. Esta adecuación de las HD ha sido defendida de manera poco frecuente por dos motivos principales. Por un lado, las perspectivas y herramientas propias de las HD no han sido especialmente usadas por las ciencias políticas, no solo en contextos angloparlantes, sino también en América Latina. Es cierto que las ciencias políticas, al igual que las HD, trabajan con datos y emplean estadística en los análisis (Thomas, 2005; Gerring, 2012). No obstante, el empleo de corpus, una metodología central y definitoria de las HD (Hockey, 2004; Ide, 2004; Hugues *et al.*, 2016; Le Deuff, 2018), no suele ser representativa en los trabajos que analizan movimientos sociales³. El propósito de este

3 Muchos de los trabajos sobre movimientos sociales emplean una cantidad considerable de datos, por ejemplo, resultados electorales y votaciones de otro

trabajo es mostrar que los métodos de corpus no solo son una metodología viable para el estudio de los movimientos sociales, sino que nos ofrecen una perspectiva privilegiada para analizar ciertas características de este tipo de movimientos.

Por otro lado, la perspectiva política no ha sido especialmente determinante en muchos trabajos de HD. Como Liu (2012) afirma, “aunque los humanistas digitales desarrollan herramientas, datos y metadatos de forma crítica, [...] rara vez extienden su crítica al registro completo de la sociedad, la economía, la política o la cultura” (p. 491). Sin embargo, esta tendencia ha ido cambiando en los últimos años. Así, por ejemplo, varios trabajos denuncian el anglocentrismo presente en las HD y las desventajas e injusticias que este genera en el contexto académico para todas las personas de habla no inglesa (Fiormente, 2012; Galina, 2013, 2014; Mahony, 2018). En una línea similar, el presente trabajo defiende la importancia de prestar atención al impacto que nuestra investigación en HD pueda tener para la sociedad y la política.

Como se ha comentado, en este texto me centraré en los movimientos sociales; más concretamente, en la repercusión que estos tienen en la clase política, al investigar su influencia en los debates parlamentarios que acontecen con posterioridad a la fecha en la que un movimiento social tuvo lugar. Tomaré como

.....
 tipo, número de protestas en distintos lapsos de tiempo, número de conversaciones entre el Gobierno y los principales actores del movimiento, o número de víctimas y violaciones de los derechos humanos (Almeida y Cordero, 2015). Sin embargo, como puede apreciarse, ninguno de estos datos proviene de corpus lingüísticos.

caso de estudio el estallido social del 18O en Chile. Las razones de esta elección son de diverso tipo. Primero, porque los debates parlamentarios del Congreso de los Diputados de la República de Chile son perfectamente accesibles. Segundo, por la importancia y repercusión que ha tenido a nivel mundial el 18O, sobre todo por el proceso constituyente que aconteció después del estallido, el cual tuvo como resultado la redacción de una nueva Constitución. Para ello, primero compilé dos corpus con los debates parlamentarios acontecidos en la Cámara de Diputados de la República de Chile cinco meses antes y cinco meses después del inicio del 18O, y luego evalué en qué medida la Cámara se hizo eco de las demandas del movimiento popular.

El plan para el capítulo es el siguiente. La primera sección presenta una taxonomía de las distintas formas de participación política, para así determinar qué tipo de participación política fue el 18O. La segunda sección describe los materiales y métodos usados en el trabajo para llevar a cabo los análisis sobre el caso de estudio. La tercera sección presenta los resultados de los análisis, mostrando la pertinencia de los métodos de corpus para estudiar el aspecto examinado en el trabajo, esto es, la interacción entre distintas formas de participación política. La cuarta sección discute los resultados de los análisis, examinando en qué medida las demandas de la ciudadanía fueron centrales en los debates parlamentarios posteriores al estallido social y, por consiguiente, si la clase política dirigente escuchó las demandas expresadas por la ciudadanía durante el estallido. Finalmente, en la última sección expongo algunas conclusiones.

Tipos de participación ciudadana

A lo largo de los años se han propuesto distintas taxonomías de participación política (Verba y Nie, 1972; Teorell *et al.*, 2007). En este trabajo sigo la propuesta de Ekman y Amnå (2012), por ser la más exhaustiva y comprensiva de todas ellas, ya que distingue tanto formas de participación política manifiestas —cuyas acciones son fácilmente distinguibles— como formas de participación política no manifiestas —relacionadas con las actitudes de la ciudadanía y cuyas acciones no pueden distinguirse de manera clara, por ejemplo, tener interés por temas políticos o llevar un estilo de vida comprometido con ciertas ideas políticas—. Dentro de los tipos de participación política manifiesta, Ekman y Amnå (2012) distinguen los siguientes: (1) *participación política formal individual*, por ejemplo, votar en unas elecciones o postularse para un cargo; (2) *participación política formal colectiva*, por ejemplo, ser miembro de un partido político o de un sindicato; (3) *participación política extraparlamentaria legal individual*, por ejemplo, la firma de peticiones o el boicot a un producto; (4) *participación política extraparlamentaria legal colectiva*, por ejemplo, formar parte de una huelga o un movimiento social; (5) *participación política extraparlamentaria ilegal individual*, por ejemplo, un ataque personal a una tienda de venta de pieles o la desobediencia civil individual; y (6) *participación política extraparlamentaria ilegal colectiva*, por ejemplo, las protestas violentas o los ataques de grupos de liberación animal a laboratorios donde se experimenta con animales.

De acuerdo con dicha clasificación, el estallido social del 18O en Chile fue una participación política manifiesta

extraparlamentaria colectiva, en la que se usaron acciones tanto legales como ilegales. Es decir, la participación política fue en general por medios legales —participación en protestas y manifestaciones pacíficas—, pero en ocasiones se convirtieron en protestas violentas, donde se emplearon medios ilegales, como quemas de estaciones de metro, ataques contra oficinas públicas y mobiliario urbano, o saqueos a supermercados. El *Informe anual del Observatorio de Conflictos 2020* del Centro de Estudios de Conflicto y Cohesión Social (COES) detalla que el 40 % de las tácticas usadas fueron pacíficas, el 34 % fueron tácticas disruptivas no violentas (huelgas, paros o tomas de plazas) y el 26 %, tácticas violentas. Así, durante el estallido se utilizaron diversas formas de protesta, lo cual “hace del estallido un fenómeno ambiguo y complejo que evade condenas o elogios fáciles” (Joignant *et al.*, 2020, p. 12).

Sin embargo, el objetivo de este trabajo no es evaluar las formas de participación política que se produjeron en el 18O, sino estudiar en qué medida la clase política dirigente escuchó las demandas del estallido. Para ello, primero constataré que los distintos trabajos que analizan la interacción entre formas de participación política en Chile se han centrado en la interacción entre los movimientos sociales y otros aspectos del sistema democrático, como los partidos políticos o la intención de voto. Por ejemplo, von Bülow y Bidegain (2015) analizan la relación entre las protestas estudiantiles en Chile entre el 2005 y el 2013, y los partidos políticos del panorama de aquel momento. Los autores concluyen que, a diferencia de países como Bolivia, Ecuador o Brasil, donde varios partidos políticos de centroizquierda se han originado en movimientos sociales, en Chile el movimiento estudiantil no derivó en alianzas con los partidos políticos.

Asimismo, existen trabajos que han estudiado el estallido social en Chile. Miranda *et al.* (2022) investigaron la interacción entre dos tipos de participación política: la electoral y la movilización social. En este sentido, y de acuerdo con la taxonomía detallada en Ekman y Amnå (2012), Miranda *et al.* (2022) estudiaron la interacción entre la participación política formal individual y la participación política extraparlamentaria colectiva. En concreto, su investigación defiende una vinculación causal entre la baja participación electoral y el aumento de la participación en protestas que desembocó en el estallido social. Según los autores, la sociedad chilena “se había caracterizado por canalizar sus demandas a través de mecanismos institucionalizados” (p. 46), pero, ya que la estructura institucional heredada de la dictadura estaba diseñada para proteger el *statu quo*, la ciudadanía fue perdiendo progresivamente su fe en las instituciones y en la participación electoral. Por esta razón, la población civil comenzó a explorar nuevas formas de “hacer oír sus demandas de manera más directa, ante la inviabilidad de otros canales de articulación y promoción de sus expectativas” (p. 50).

En una línea similar, el presente trabajo estudia la interacción entre una forma de participación política extraparlamentaria colectiva, el estallido social, y una forma de participación política formal colectiva, los debates parlamentarios. Para ello, examinaré en qué medida el 18O influyó en los debates parlamentarios que acontecieron después, al analizar si el parlamento se hizo eco de las demandas de la ciudadanía. Miranda *et al.* (2022) constatan que una de las principales razones por las que la gente se moviliza es porque quieren que sus demandas sean escuchadas por la clase dirigente:

Los sucesos de fines de 2019 en Chile ponen en evidencia que quienes votan esperan de alguna forma que sus demandas sean canalizadas y que cuando así no ocurre salen a protestar, entendiendo la protesta como una forma válida de comunicación con el sistema político. (p. 55)

En este sentido, puede decirse que una de las pretensiones de los movimientos sociales es establecer una vía de comunicación entre quienes participan en dicho movimiento y las personas que forman parte de la clase política dirigente. Este trabajo se cuestiona si el 18O permitió establecer dicho canal comunicativo y, en caso afirmativo, cuál fue el impacto y alcance de este. La cuarta sección detalla los análisis llevados a cabo para estudiar esta cuestión. Sin embargo, antes presentaré los métodos y materiales específicos usados en el trabajo.

Aplicando las herramientas de las humanidades digitales a los movimientos sociales

Las humanidades siempre han girado, en mayor o menor medida, en torno al estudio de los textos escritos. La llegada de las HD ha supuesto un salto cuantitativo y cualitativo en muchos aspectos relacionados con el manejo y análisis de este tipo de textos, por ejemplo, frente al almacenamiento digital, la anotación o el análisis textual. En este trabajo me centro en los corpus, es decir, en “datos lingüísticos que ocurren de manera natural, almacenados en formato electrónico, diseñados para ser representativos de determinados

tipos de texto y analizados con ayuda de herramientas informáticas” (Nesi, 2016, p. 206), así como en los métodos propios de la lingüística de corpus. Aunque dar una definición consensuada de *lingüística de corpus* es imposible⁴, podemos definirla “como el tratamiento de un conjunto de textos legibles por máquina que se considera una base adecuada para estudiar un conjunto específico de cuestiones de investigación” (McEnery y Hardie, 2012, p. 1).

Como se comentó en la introducción, uno de los objetivos del trabajo es mostrar que los métodos de corpus suponen una metodología no solo viable, sino adecuada para estudiar algunas características de los movimientos sociales. Sin embargo, alguien podría preguntarse si los métodos de corpus son estrictamente necesarios para analizar los movimientos sociales, es decir, si no podríamos llevar a cabo el estudio usando análisis de texto tradicionales —lecturas detalladas de extractos de aquellos textos que son centrales para entender aquello que estudiamos—. La respuesta es no, porque el uso de métodos de corpus genera ciertas ventajas que los métodos tradicionales no tienen (Bluhm, 2013, 2016), por ejemplo, una mayor cantidad de evidencia considerada o una mayor replicabilidad. Además, el uso de métodos de corpus no implica exclusividad, en el sentido de que usar métodos de corpus es perfectamente compatible con métodos como el análisis cualitativo (Bordonaba-Plou, 2022a).

En el caso que nos ocupa, y como se mostrará en la siguiente sección, los métodos de corpus nos permitirán examinar la

4 Hay quienes defienden que la lingüística de corpus es una disciplina propia de la lingüística (Stubbs, 1993; Tognini-Bonelli, 2001; Teubert, 2004). Sin embargo, también hay quienes consideran que la lingüística de corpus es un conjunto de métodos que pueden aplicarse a distintas disciplinas (McEnery y Wilson, 2001; Parodi, 2008; McEnery y Hardie, 2012).

interacción entre dos formas de participación política: los movimientos sociales y los debates parlamentarios. Más concretamente, los métodos de corpus nos ofrecen la posibilidad de examinar cuál fue la repercusión que tuvo un movimiento social como el 18O en Chile en los debates de la clase política dirigente, y constatar en qué medida las demandas de la ciudadanía fueron atendidas.

Para estudiar esto, usaré dos corpus de debates parlamentarios, que constan de los diarios de sesiones de la Cámara de Diputados de la República de Chile. Las actas o diarios de sesiones han sido usados en un amplio abanico de disciplinas pertenecientes a las humanidades y las ciencias sociales, por ejemplo, historia, sociología, análisis crítico del discurso o filosofía, así como para abordar una amplia gama de temas: *big data* e historia digital (Pančur y Šorn, 2016), islamofobia (Cheng, 2015), dimensiones del conocimiento relevante en las interacciones parlamentarias (van Dijk, 2002), racismo (Augoustinos y Every, 2007) o desacuerdo (Bordonaba-Plou, 2022b; Bordonaba-Plou, 2022c).

Además, los corpus de debates parlamentarios ya han sido usados en las investigaciones relacionadas con las HD (Guldi, 2019). En este campo, gran parte de los esfuerzos se han centrado en mejorar la accesibilidad e interoperabilidad de los corpus de debates parlamentarios, por ejemplo, trabajando en los procesos de documentación o desarrollando estándares de anotación comunes que faciliten el análisis de las distintas características de los discursos parlamentarios. Como Fišer y Lenardič (2018) señalan, “el desarrollo de herramientas analíticas integradoras y ricamente anotadas es necesario para conseguir una mejor comprensión de las especificidades de los debates parlamentarios” (p. 2). Es especialmente relevante a este respecto Common Language Resources

and Technology Infrastructure (CLARIN)⁵, una infraestructura europea de investigación orientada a las humanidades y las ciencias sociales, que ofrece treinta y un corpus parlamentarios de distintos países⁶. Dentro de CLARIN, merece una mención particular la iniciativa Parla CLARIN —un esquema *text encoding initiative* (TEI)⁷ para anotar corpus de actas parlamentarias, cuyo objetivo es la creación, el enriquecimiento y la utilización de corpus parlamentarios (Erjavec y Pančur, 2021; Erjavec *et al.*, 2022)—, así como la serie de workshops CLARIN LREC (Fišer *et al.*, 2018, 2020).

Al ser una estructura de investigación europea, CLARIN está centrada en los debates del Parlamento Europeo y en los debates parlamentarios de países como Reino Unido, Noruega, República Checa, Dinamarca, Estonia, Finlandia, Francia, Alemania, Lituania, Portugal, Eslovenia, Suecia, España, Bulgaria, Croacia, Italia, Islandia, Turquía, entre otros. Al haber desarrollado estándares para la producción, anotación y documentación que permiten estudiar de manera sistemática los debates parlamentarios, la posibilidad de implementar dichos estándares en países no europeos es evidente e inmediata, al menos inicialmente. En el caso que nos ocupa, los debates parlamentarios de la República de Chile, en principio parece posible usar las iniciativas descritas para producir corpus anotados de los debates parlamentarios, ya que la accesibilidad a los recursos y la estructura de los materiales no son en ningún sentido distintas a las de los países europeos. Así, los académicos que estudian los discursos parlamentarios de la

5 Véase <https://www.clarin.eu/>.

6 Véase <https://www.clarin.eu/resource-families/parliamentary-corpora>.

7 Véase <https://tei-c.org/>.

Cámara de Diputados de la República de Chile, o que se sirven de estos en sus estudios, se beneficiarían enormemente si pudieran contar con este tipo de corpus. Sin embargo, la realidad es que queda mucho por realizar para extrapolar estas iniciativas a Chile, en concreto, y a Latinoamérica, en general.

El presente trabajo no está centrado en el desarrollo de los procesos de producción, anotación y documentación de los corpus de debates parlamentarios y en su posible aplicación en el marco de las HD en Latinoamérica. Su objetivo es más modesto: pretende mostrar cómo otras herramientas usadas en las HD, las cuales son accesibles para investigadores individuales y pequeños grupos de investigación, nos pueden servir para estudiar algunas características de los discursos de este tipo de corpus. Mostraré cómo las expresiones regulares y algunos métodos de la lingüística de corpus, como los análisis de colocaciones, nos permiten examinar la influencia de los movimientos sociales en los debates parlamentarios. En este sentido, esta investigación revela que el análisis y estudio de los debates parlamentarios es posible aunque no exista una infraestructura de investigación desarrollada. La implementación de infraestructuras como CLARIN en América Latina sigue siendo una necesidad, por lo que este trabajo expone la viabilidad de algunos métodos para el estudio de los debates parlamentarios hasta el momento en el que dichas infraestructuras se desarrollen.

Para esta investigación compilé dos corpus. El primero está compuesto con los diarios de sesiones de los cinco meses anteriores al inicio del estallido, desde la sesión 29.^a de la legislatura 367.^a, celebrada el martes 28 de mayo del 2019, hasta la sesión 86.^a de la legislatura 367.^a, del jueves 10 de octubre del 2019. El corpus tiene

un tamaño total de 2 417 902 palabras. El segundo se compone de los diarios de sesiones de los cinco meses posteriores al inicio del estallido, desde la sesión 93.^a de la legislatura 367.^a, celebrada el martes 22 de octubre del 2019, hasta la sesión 4.^a de la legislatura 368.^a, del miércoles 18 de marzo del 2020. El corpus tiene un tamaño total de 2 774 056 palabras.

El proceso de compilación y edición de los corpus se dividió en los siguientes pasos. Primero, *formación de los corpus*. Los diarios de sesiones pueden descargarse manualmente⁸ o de manera automática, por ejemplo, usando un programa de código abierto como RStudio⁹, por medio de los paquetes XML¹⁰ y RCurl¹¹, que permiten manejar las páginas HTML y las direcciones URL, respectivamente.

Segundo, *edición de los corpus*. Como los archivos digitales estaban en formato PDF, tuve que convertirlos en archivos TXT, usando el paquete pdftools¹² de RStudio. Luego, eliminé la información no deseada, por ejemplo, los encabezados o los números de página; para ello, usé expresiones regulares¹³. Las expresiones regulares son una notación estándar para caracterizar y buscar cadenas de texto; una cadena de texto es “cualquier secuencia de caracteres alfanuméricos (letras, números, espacios, tabulaciones y signos de puntuación)” (Jurafsky y Martin, 2008, p. 18).

8 Véase https://www.camara.cl/legislacion/sesiones_sala/sesiones_sala.aspx.

9 Véase <https://www.rstudio.com/>.

10 Véase <https://cran.r-project.org/web/packages/XML/index.html>.

11 Véase <https://cran.r-project.org/web/packages/RCurl/index.html>.

12 Véase <https://cran.r-project.org/web/packages/pdftools/index.html>.

13 En el apéndice A se detallan todas las expresiones regulares usadas para editar los dos corpus.

Una vez compilados y editados los corpus, llevé a cabo un *análisis lingüístico de los debates parlamentarios*. Así, realicé dos tipos de análisis. Primero, un análisis de los cincuenta sustantivos más frecuentes. Segundo, un análisis de las colocaciones del término “ciudadanía”. Las colocaciones son palabras “que aparecen con frecuencia dentro de la vecindad de otra palabra, normalmente más a menudo de lo que esperaríamos que las dos palabras aparecieran juntas debido al azar” (Baker *et al.*, 2013, p. 36). Dado que el trabajo pretende examinar si las demandas de la ciudadanía fueron escuchadas por la clase política, uno de los puntos centrales fue determinar cuáles habían sido las demandas que más se habían repetido en cada uno de los dos periodos. Por esta razón, examiné las colocaciones del término “ciudadanía”, para comprobar qué demandas se tuvieron en cuenta en cada uno de los dos periodos. Para llevar a cabo el primer análisis, usé la herramienta Wordlist de Sketch Engine¹⁴; para el segundo usé la herramienta Sketch Word del mismo *software*.

En definitiva, este proceso permite identificar qué demandas de la ciudadanía se nombraban en el Parlamento antes y después del 180, así como observar los cambios en la centralidad de estas demandas.

Analizando la interacción entre el 180 y los debates parlamentarios

Comenzaré presentando los resultados de los dos análisis en el corpus preestallido, es decir, el corpus compuesto por los diarios de

14 Véase <https://www.sketchengine.eu/>.

sesiones de la Cámara de Diputados de la República de Chile acontecidos cinco meses antes del inicio del estallido social; luego, presentaré los resultados de los dos análisis en el corpus posestallido.

Entre los sustantivos más frecuentes en el corpus preestallido, encontramos términos relacionados con la forma propia en la que tienen lugar los debates parlamentarios; por ejemplo, “diputado”, “don”, “señor”, “proyecto”, “ley”, “presidente”, “número”, “términos”, “comisión”, “año”, “artículo”, “ministro”, “palabra”, “fecha” o “informe” aparecen entre los cincuenta sustantivos más frecuentes. Sin embargo, también encontramos varios que no pertenecen al registro propio de los debates parlamentarios (tabla 10.1)¹⁵.

TABLA 10.1. Sustantivos más frecuentes en el corpus preestallido

TÉRMINO	POSICIÓN	FRECUENCIA ABSOLUTA	FRECUENCIA RELATIVA *
Comuna	7	7055	29,18
Persona	10	4860	20,10
Chile	25	3135	12,96
País	30	2849	11,78
Forma	32	2752	11,38

Continúa...

15 Los términos que aparecen en la tabla son lemas, es decir, las formas canónicas “a las que se reduce todo un PARADIGMA flexivo y que se toma como representante de todas las variantes morfológicas de la palabra” (Alcaraz y Martínez, 1997, p. 321). Por ejemplo, el término “empresa” recoge ocurrencias tanto de “empresa” como de “empresas”.

Continúa...

Salud	33	2726	11,27
Agua	34	2707	11,19
Empresa	35	2701	11,17
Estado	37	2567	10,62
Medida	38	2499	10,33
Materia	39	2486	10,28
Sistema	42	2425	10,03
Sector	45	2383	9,85

* La frecuencia relativa es igual a la frecuencia absoluta dividida por el número de palabras del corpus; el resultado suele multiplicarse por una base normalizadora (10 000 o 1 000 000). En este caso, la base normalizadora es 10.000.

FUENTE: elaboración propia.

En el examen de colocaciones de “ciudadanía” son especialmente significativos dos conjuntos de términos: los verbos y los términos que aparecen en la construcción “(...) de la ciudadanía”. Por un lado, entre los verbos que coocurren con más frecuencia con “ciudadanía”, cabe destacar “elegir”, con un log dice¹⁶ de 9.3. Este verbo nos permite identificar algunas demandas en este periodo, en particular la necesidad de hacer reformas del sistema tributario y el apoyo a las pymes. Asimismo, el verbo “esperar”, con

16 El log dice es una medida estadística que expresa “la tendencia de dos palabras a coocurrir en relación con la frecuencia de estas palabras en el corpus”; es una “medida estandarizada que opera en una escala con un valor máximo fijo de 14” (Gablasova *et al.*, 2017, p. 164).

un log dice de 8.5, refiere a aquello que la ciudadanía espera de sus representantes, concretamente mayor seguridad pública y derogar la Ley Reservada del Cobre, que tiene como objetivo destinar parte de los ingresos por la venta de cobre a la adquisición de material armamentístico para la defensa nacional.

Por otro lado, entre las colocaciones de la construcción “(...) de la ciudadanía”, existen varias que se refieren a demandas específicas: “expectativa”, “exigencia”, “requerimiento” y “demanda”. Al examinar los casos de “requerimiento”, puede observarse que todos se asocian con el debate sobre un proyecto de ley para aprobar la convocatoria de plebiscitos comunales. Los casos de “expectativa” muestran que las demandas tienen que ver con el sistema de salud y el precio de la energía. Los casos de “exigencia” permiten identificar dos demandas: penas más duras para los delitos de abuso sexual a menores y una menor privatización de los recursos hídricos. Finalmente, los casos de “demanda” están relacionados con la seguridad pública y la transparencia en las instituciones.

A continuación, se detallan los análisis realizados en el corpus posestallido, es decir, aquel compuesto por los diarios de sesiones de la Cámara de Diputados de la República de Chile acontecidos cinco meses después del inicio del estallido social. Empezaremos por el análisis de los cincuenta sustantivos más frecuentes. Al igual que en el caso anterior, aparecen muchos términos relacionados con el vocabulario propio de los debates parlamentarios; sin embargo, también hay otros sustantivos no relacionados con los debates parlamentarios (tabla 10.2).

De nuevo, al igual que ocurría en el corpus preestallido, dos conjuntos de términos son en particular relevantes para el análisis de colocaciones: los verbos y los términos que aparecen en la

TABLA 10.2. Sustantivos más frecuentes en el corpus posestallido

TÉRMINO	POSICIÓN	FRECUENCIA ABSOLUTA	FRECUENCIA RELATIVA
Comuna	7	5661	20,40
País	12	4701	14,67
Chile	21	3888	14,01
Persona	23	3806	13,72
Salud	29	3358	12,10
Acuerdo	31	3188	11,49
Estado	33	3127	11,27
Gobierno	35	3096	11,16
Medida	37	3013	10,86
Sistema	38	2907	10,48
Región	41	2796	10,08
Trabajo	47	2515	9,07
Empresa	49	2465	8,88

FUENTE: elaboración propia.

construcción “(...) de la ciudadanía”. Entre los verbos que coocurren con “ciudadanía” hay varios que refieren de manera directa a demandas concretas. Por ejemplo, “demandar”, con un log dice de 11, muestra pertinentes las siguientes demandas: la reforma tributaria, mejores pensiones, mejorar el sistema de salud, disminuir el número de diputados y senadores, la concesión de subsidios para rentas bajas, el reajuste de los salarios, el voto obligatorio o las rebajas en el transporte público. “Querer”, con un log dice de 10, permite identificar las siguientes demandas: mejoras en el sistema de salud, mayor dignidad y solidaridad, mayor participación de la ciudadanía, mejoras en las pensiones y reforma constitucional. “Reclamar”, con un log dice de 9.6, alude al deseo de la ciudadanía de conseguir una menor precarización del trabajo, un mayor apoyo a las pymes, una mayor participación de la ciudadanía, la convocatoria a plebiscitos comunales, mejoras en el sistema de salud, y respeto a las costumbres y formas de vida. “Pedir”, con un log dice de 9.1, refiere a las siguientes demandas: acabar con el monopolio de las instituciones crediticias, mejoras de las pensiones, reforma constitucional, mayor seguridad pública, mejoras en la calidad de vida, mayor dignidad y solidaridad, mayor transparencia de las instituciones, terminar con la desigualdad y subir el salario mínimo. “Esperar”, con un log dice de 8.9, permite identificar las siguientes demandas: apoyo a las pymes, aumento del salario mínimo, reforma de la prueba de acceso a la universidad, el reconocimiento de errores por parte de la clase política, atender la angustia social, aumento de las pensiones, reforma tributaria y transversalidad en el Parlamento. Finalmente, “exigir”, con un log dice de 8.7, se usa para apuntar que la ciudadanía demanda mejoras en la calidad de vida, disminuir el número de diputados y

senadores, mejoras en el sistema de salud, mayor dignidad y participación, y una mayor transparencia de las instituciones.

De las colocaciones de la construcción “(...) de la ciudadanía”, dos términos permiten identificar solicitudes concretas. Por un lado, “demanda”, con un log dice de 11.4, refiere a la reforma del sistema tributario, el establecimiento de la jornada laboral de cuarenta horas, fijar una retribución para personas cuidadoras, mejorar las pensiones, mejorar el sistema sanitario, reformar la Constitución y mejorar la transversalidad en el Parlamento. Por otro lado, el término “necesidad” alude a la idea de paridad en el Parlamento y a la reforma del sistema sanitario.

Discusión

Aunque el análisis de los cincuenta sustantivos más frecuentes revela que muchos de estos términos aparecen en ambos corpus, nos permite concluir la existencia de ciertos cambios en los debates parlamentarios antes y después del 180. De especial relevancia son los términos que aparecen solo en uno de los corpus. Por ejemplo, “agua” solo está en el corpus preestallido, lo cual indica que los debates respecto a la gestión de los recursos hídricos fueron habituales en los meses anteriores al estallido. El término “acuerdo” solo aparece en el corpus posestallido, lo que probablemente revela la necesidad de llegar a acuerdos en el Parlamento ante las demandas de la sociedad en su conjunto. También es importante señalar la presencia del término “salud” en ambos corpus, lo que indica la centralidad de los debates sobre el sistema de salud en Chile, siendo algo más relevante en el periodo posterior

al estallido, ya que la frecuencia relativa es algo más elevada en aquel corpus.

El análisis de colocaciones permite extraer dos conclusiones. Primero, el número total de demandas es muy superior en el corpus posestallido: mientras en el corpus preestallido pueden identificarse diez demandas, en el corpus posestallido se revelan setenta y tres. Esta diferencia se refleja en el número de colocaciones que señalan explícitamente demandas de la ciudadanía en cada corpus, así como en el número de ocurrencias de dichas colocaciones. Por ejemplo, mientras que en el corpus preestallido aparecen dos verbos, “elegir” (2 ocurrencias) y “esperar” (2), en el corpus posestallido aparecen seis verbos, “demandar” (10), “querer” (13), “reclamar” (4), “pedir” (15), “esperar” (16) y “exigir” (9).

Segundo, el número de demandas distintas también se incrementó durante dicho periodo. Los análisis muestran que las demandas que el Parlamento atribuye a la ciudadanía durante el periodo anterior al estallido son de carácter económico (reformas del sistema tributario, apoyo a las pymes, derogación de la Ley Reservada del Cobre y disminución del precio de la energía eléctrica), político/institucional (convocatoria a plebiscitos comunales y mayor transparencia de las instituciones) y social (mayor seguridad pública, mejoras en el sistema de salud y mayores penas para delitos de abuso sexual a menores). Casi todas estas demandas vuelven a aparecer en el corpus posestallido. Las únicas que no se repiten son la derogación de la Ley Reservada del Cobre, la disminución del precio de la energía eléctrica y mayores penas para delitos de abuso sexual a menores. Además, pueden distinguirse hasta veintiún nuevas demandas (tabla 10.3).

TABLA 10.3. Atribución de demandas a la ciudadanía por parte del Parlamento antes y después del 18O

DEMANDAS CIUDADANAS PREESTALLIDO	DEMANDAS CIUDADANAS POSESTALLIDO
Reforma del sistema tributario	Reforma del sistema tributario (3 ocurrencias en el corpus)
Apoyo a las pymes	Mejores pensiones (9 ocurrencias en el corpus)
Mayor seguridad pública (2 ocurrencias en el corpus)	Mejoras en el sistema de salud (16 ocurrencias en el corpus)
Derogación de la ley reservada del cobre	Disminuir el número de diputados y senadores (2 ocurrencias en el corpus)
Convocatoria a plebiscitos comunales	Subsidio para rentas bajas
Mejoras en el sistema de salud	Subida de los salarios
Disminución del precio de la energía eléctrica	Voto obligatorio (2 ocurrencias en el corpus)
Mayores penas para delitos de abuso sexual a menores	Rebajas en el transporte público
Mayor transparencia de las instituciones	Trabajos menos precarios
	Apoyo a las pymes (2 ocurrencias en el corpus)

Continúa...

Continúa...

DEMANDAS CIUDADANAS PREESTALLIDO	DEMANDAS CIUDADANAS POSESTALLIDO
	Convocatoria a plebiscitos comunales
	Reforma de las instituciones crediticias
	Reforma constitucional (6 ocurrencias en el corpus)
	Mayor seguridad pública
	Mejoras en la calidad de vida (4 ocurrencias en el corpus)
	Mayor dignidad (3 ocurrencias en el corpus)
	Mayor transparencia de las instituciones (2 ocurrencias en el corpus)
	Terminar con la desigualdad (3 ocurrencias en el corpus)
	Salario mínimo (2 ocurrencias en el corpus)
	Reforma de la prueba de acceso a la universidad
	Reconocimiento de errores por parte de la clase política
	Atender la angustia social

Continúa...

Continúa...

DEMANDAS CIUDADANAS PREESTALLIDO	DEMANDAS CIUDADANAS POSESTALLIDO
	Transversalidad en el Parlamento (2 ocurrencias en el corpus)
	Mayor participación de la ciudadanía (3 ocurrencias en el corpus)
	Respeto a las costumbres y formas de vida
	Jornada laboral de cuarenta horas
	Retribución a personas cuidadoras (2 ocurrencias en el corpus)
	Paridad en el Parlamento

FUENTE: elaboración propia.

En definitiva, los resultados muestran que los discursos de la clase política en los debates parlamentarios posestallido están más centrados en las demandas, expectativas y necesidades de la ciudadanía. Sin embargo, podemos preguntarnos si las demandas atribuidas durante este periodo representan las solicitudes reales de la ciudadanía.

El *Informe anual del Observatorio de Conflictos 2020* del COES apunta que “sólo el 25 % de las acciones durante el estallido exhibieron demandas específicas” (Joignant *et al.*, 2020). Esto tiene que ver, entre otros factores, “un sesgo del instrumento basado en la revisión de prensa” (p. 14), es decir, el informe usa

metodologías de recolección de datos basadas en informes de prensa y esto puede diluir la heterogeneidad de las demandas, ya que “reportar protestas para un diario obliga a trabajar rápido, sin mucha capacidad de análisis ni de profundización” (Allain *et al.*, 2020). Sin embargo, el informe señala que las demandas centradas en aspectos materiales, en particular la salud, la vivienda y las pensiones, fueron las demandas más específicas del estallido (Joignant *et al.*, 2020).

De las demandas posestallido identificadas por medio del análisis de colocaciones, las tres más frecuentes son, de mayor a menor, mejoras en el sistema de salud, mejores pensiones y reforma constitucional. Las demandas por un mejor sistema sanitario pasan de una aparición en el corpus preestallido a dieciséis en el corpus posestallido, mientras que las demandas por mejores pensiones pasan de cero en el corpus preestallido a nueve en el corpus posestallido. Además, el análisis de los cincuenta sustantivos más frecuentes confirma esta tendencia. El término “pensión” es el 249.º sustantivo más usual en el corpus preestallido, con una frecuencia absoluta de 762 y una frecuencia relativa de 3,15. Sin embargo, ocupa el puesto 85.º en el corpus posestallido, con una frecuencia absoluta de 1863 y una frecuencia relativa de 6,71. Asimismo, “salud” pasa de ser el 33.º en el corpus preestallido, con una frecuencia absoluta de 2726 y una frecuencia relativa de 11,27, a ser el 29.º en el corpus posestallido, con una frecuencia absoluta de 3358 y una frecuencia relativa de 12,10.

No obstante, en el caso de “vivienda” ocurre a la inversa: es el 149.º sustantivo más usado en el corpus preestallido, con una frecuencia absoluta de 1132 y una frecuencia relativa de 4,68, y el 445.º

en el corpus posestallido, con una frecuencia absoluta de 717 y una frecuencia relativa de 2,58.

Así, los debates en torno a la salud y las pensiones adquirieron más protagonismo después del 18O. Pero la demanda de vivienda digna, una de las tres demandas centrales del estallido, identificadas en el *Informe anual del Observatorio de Conflictos 2020* del COES, fue completamente olvidada, ya que el término aparece en medida mucha menor en los debates parlamentarios posteriores al estallido.

Como se ha comentado, la tercera demanda más frecuente en el corpus posestallido es la reforma constitucional. Esto anticiparía de alguna manera la evolución de los sucesos políticos de los meses posteriores, ya que la ciudadanía chilena fue convocada a un plebiscito nacional en el 2020 para votar si estaba de acuerdo con la formación de una asamblea constituyente que redactara una nueva Constitución. Sin embargo, la reforma constitucional no es considerada una demanda central del estallido en el *Informe anual del Observatorio de Conflictos 2020* del COES. Según este, las demandas por un cambio constitucional aumentaron durante el estallido social respecto a etapas anteriores, pero hay registros de ellas en solo el 3,5% de las protestas. Es decir, no era una demanda central de las distintas protestas que se produjeron durante el 18O, lo cual puede llevar a pensar que no fueron la principal causa del inicio del proceso constituyente que se originó *a posteriori*. Como Joignant *et al.* (2020) sugieren, “la inmensa mayoría de las acciones contenciosas durante el estallido *no* levantó explícitamente el posible resultado del principal acuerdo inmediato entre los partidos políticos para enfrentar las movilizaciones: el inicio de un proceso de eventual cambio constitucional” (p. 15).

En definitiva, es posible que una de las tres demandas que con más frecuencia se atribuye a la ciudadanía en los debates parlamentarios posestallido sea la reforma constitucional, pero, como apunta el informe, esta no era una de las demandas principales de la ciudadanía. Algunas personas en el Parlamento expresan esto de manera explícita. Por ejemplo, la diputada Joanna Pérez, debatiendo el proyecto de ley de presupuestos del sector público para el 2020, perteneciente a la sesión 105.^a de la legislatura 367.^a, del jueves 14 de noviembre del 2019, dijo lo siguiente:

Más allá de que todos queremos una nueva Constitución, no podemos desviarnos de los temas urgentes que demanda la ciudadanía. Claramente, el foco de atención y nuestro esfuerzo deben estar puestos en las pensiones y en la salud, y si es necesario el endeudamiento, hay que hacerlo por el país.

En síntesis, los resultados de los análisis muestran que las demandas expresadas durante el 18O por la ciudadanía llegaron a oídos del Parlamento chileno y, en algunos casos, fueron el foco de muchos debates que tuvieron lugar los meses posteriores al estallido. Concretamente, mejorar el sistema de salud y las pensiones, dos de las demandas centrales durante el estallido, fueron las más frecuentes; sin embargo, la demanda por una vivienda digna, otra demanda central, prácticamente desapareció en los debates posestallido. Además, los resultados muestran que la reforma constitucional, quizá una de las que más atención atrajo durante los meses posteriores, aunque atribuida por el Parlamento a la ciudadanía, era propia de la clase política. En este sentido, puede decirse que el estallido social constituyó una vía de comunicación

imperfecta entre la ciudadanía y el Parlamento, ya que, por un lado, permitió que solo algunas de las demandas centrales del estallido fueran debatidas en el Parlamento y, por otro, una demanda que no era central para la ciudadanía pasó a ocupar el centro de atención en los meses posteriores al estallido.

Conclusiones

Los movimientos sociales son una forma de participación política cada vez más relevante en el siglo XXI. La literatura sobre movimientos sociales suele usar metodología propia de las ciencias sociales, por ejemplo, cuestionarios y entrevistas, para investigar las distintas características de los movimientos sociales. En este capítulo he mostrado que también podemos usar las herramientas de las HD para investigar estos movimientos. En particular, he usado métodos de corpus para investigar la interacción entre dos formas de participación política: los movimientos sociales, una forma de participación política extraparlamentaria colectiva, y los debates parlamentarios, una forma de participación política formal colectiva. Para ilustrar cómo estos métodos pueden usarse para investigar la interacción entre estas dos formas de participación política, he analizado la interacción entre el 18O en Chile y los debates parlamentarios que tuvieron lugar en los meses posteriores al estallido. Así mismo, he investigado en qué medida el Parlamento chileno escuchó las demandas expresadas por la ciudadanía durante el 18O, al examinar los cambios que hubo en los debates parlamentarios antes y después del estallido.

Los análisis revelan un cambio en la consideración de las demandas de la ciudadanía en los debates parlamentarios posteriores al estallido. Dos de las demandas centrales del estallido, mejorar el sistema de salud y las pensiones, ocuparon un lugar central en los debates parlamentarios posestallido. Sin embargo, otra de las demandas centrales del estallido, la demanda de vivienda, no estuvo presente en dichos debates y, de este modo, quedó desatendida. Además, la reforma constitucional, una de las solicitudes que más atención ha atraído en los medios de comunicación de muchos países, aunque no era propia del estallido, también fue central en los debates posestallido. En este sentido, el presente trabajo permite concluir que el 18O representó un medio para establecer una vía de comunicación imperfecta entre la ciudadanía y la clase política chilena.

Para concluir, solo queda subrayar que el proceso de análisis detallado en el presente trabajo representa una de las posibles maneras en que pueden aplicarse las metodologías de análisis textual propias de las HD para investigar los movimientos sociales o la interacción entre distintas formas de participación política. La metodología usada y detallada en este trabajo es una manera eficiente y fácilmente reproducible de analizar cómo y en qué medida son consideradas las demandas de la ciudadanía en los debates parlamentarios. No obstante, para que esto sea posible, es necesario el trabajo previo de consulta y encuestas a la ciudadanía, con el fin de saber cuáles son dichas demandas, es decir, la pluralidad metodológica y el trabajo conjunto entre las ciencias sociales y las humanidades son necesarios para investigar fenómenos complejos como los movimientos sociales. En definitiva, las HD nos ofrecen

diversas maneras de investigar fenómenos humanos complejos y profundizar en el significado y alcance de estos.



Referencias

- Aguilera, C. (2022). “Chile despertó”: Los sentidos políticos en la revuelta de octubre. *Polis. Revista Latinoamericana*, 21(61), 10-31.
- Alcaraz, E. y Martínez, M. A. (1997). *Diccionario de lingüística moderna*. Ariel.
- Allain, M., Delamaza, G., Maillet, A. y Rivas, R. (2020, 31 de octubre). Demandas, organizaciones y violencias: perspectivas para entender la revuelta de 2019. *Voces COES*. <https://coes.cl/opinion-demandas-organizaciones-y-violencias-perspectivas-para-entender-la-revuelta-de-2019/>.
- Almeida, P. y Cordero, A. (Eds.) (2015). *Handbook of social movements across Latin America*. Springer.
- Augoustinos, M. y Every, D. (2007). Contemporary racist discourse: Taboos against racism and racist accusations. En V. Weatherall, B. M. Watson y C. Gallois (Eds.), *Language, discourse and social psychology* (pp. 233-254). Palgrave Macmillan.
- Baiocchi, G. y Ganuzza, E. (2017). *Popular democracy. The paradox of participation*. Stanford University Press.
- Baker, P., Gabrielatos, C. y McEnery, T. (2013). *Discourse analysis and media studies: The representation of Islam in the British press*. Cambridge University Press.

- Bianchi, M., Leó, C. y Perin, A. (2016). Transformaciones de la participación política en América Latina. *Revista Eletrônica de Ciência Política*, 7(2), 46-66.
- Bluhm, R. (2013). Don't ask, look! Linguistic corpora as a tool for conceptual analysis. En M. Hoeltje, T. Spitzley y W. Spohn, (Eds.), *Was dürfen wir glauben? Was sollen wir tun? Sektionsbeiträge des achten internationalen Kongresses der Gesellschaft für Analytische Philosophie* (pp. 7-15). e.V. DuEPublico.
- Bluhm, R. (2016). Corpus analysis in philosophy. En M. Hinton (Ed.), *Evidence, experiment and argument in linguistics and philosophy of language* (pp. 91-109). Peter Lang.
- Bordonaba-Plou, D. (2022a). Bucles virtuosos: Una defensa del uso de métodos cualitativos y cuantitativos en filosofía del lenguaje. *ALPHA: Revista de Artes, Letras y Filosofía*, 1(54), 59-70.
- Bordonaba-Plou, D. (2022b). Desacuerdos estratégicos y dinámicas de conflicto antagónicas. *Isegoría*, 66, e25.
- Bordonaba-Plou, D. (2022c). Análisis de la existencia fáctica de los desacuerdos profundos no epistémicos en los debates parlamentarios. *Cuadernos de Filosofía*, 40, 93-116.
- Cartagena, R. E. (2015). Environmental conflicts and social movements in postwar El Salvador. En P. Almeida y A. Cordero Ulate (Eds.), *Handbook of social movements across Latin America* (pp. 237-254). Springer.
- Cheng, J. E. (2015). Islamophobia, muslimophobia or racism? Parliamentary discourses on Islam and Muslims in debates on the minaret ban in Switzerland. *Discourse & Society*, 26(5), 562-586.
- Dalton, R. J. (2017). *The participation gap. Social status and political inequality*. Oxford University Press.

- Dalton, R. J., Scarrow, S. E. y Cain, B. E. (2003). New forms of democracy? Reform and transformation of democratic institutions. En B. E. Cain, R. J. Dalton y S. E. Scarrow (Eds.), *Democracy transformed. Expanding political opportunities in advanced industrial democracies* (pp. 1-20). Oxford University Press.
- Ekman, J. y Amnå, E. (2012). Political participation and civic engagement: Towards a new typology. *Human Affairs*, 22(3), 283-300.
- Erjavec, T. y Pančur, A. (2021). The Parla-CLARIN recommendations for encoding corpora of parliamentary proceedings. *Journal of the Text Encoding Initiative*, 14. <https://doi.org/10.4000/jtei.4133>.
- Erjavec, T., Ogrodniczuk, M., Osenova, P., et al. (2023). The ParlaMint corpora of parliamentary proceedings. *Lang Resources & Evaluation*, 57, 415-448. <https://doi.org/10.1007/s10579-021-09574-0>.
- Fiormonte, D. (2012). Towards a cultural critique of the digital humanities. *Historical Social Research*, 37(3), 59-76.
- Fišer, D., Eskevich, M. y De Jong, F. (Eds.). (2020). *Proceedings of the Second Parla-CLARIN Workshop*. European Language Resources Association (ELRA). <https://www.aclweb.org/anthology/2020.parlaclarin-1.0>.
- Fišer, D. y Lenardič, J. (2018). CLARIN Corpora for Parliamentary Discourse Research. En D. Fišer, M. Eskevich y F. de Jong (Eds.), *Proceedings of the LREC2018 Workshop Parla-CLARIN: Creating and using parliamentary corpora* (pp. 2-7). European Language Resources Association (ELRA). http://lrec-conf.org/workshops/lrec2018/W2/summaries/14_W2.html.
- Fišer, D., Lenardič, J. y Erjavec, T. (2018). CLARIN's key resource families. En *Proceedings of the Eleventh International Conference*

- on *Language Resources and Evaluation (LREC 2018)*. European Language Resources Association (ELRA). <https://www.aclweb.org/anthology/L18-1210>.
- Gablasova, D., Brezina, V. y McEnery, T. (2017). Collocations in corpus-based language learning research: Identifying, comparing, and interpreting the evidence. *Language Learning*, 67(S1), 155-179.
- Galina, I. (2013, 19 de julio). Is there anybody out there? Building a global digital humanities community. *Red Humanidades Digitales*. <http://humanidadesdigitales.net/blog/2013/07/19/is-there-anybody-out-there-building-a-global-digital-humanities-community/>.
- Galina, I. (2014). Geographical and linguistic diversity in the digital humanities. *Literary and Linguistic Computing*, 29(3), 307-316.
- Gallego, A. (2014). *Unequal political participation worldwide*. Cambridge University Press.
- Gerring, J. (2012). *Social science methodology: A critical framework*. Cambridge University Press.
- Guldi, J. (2019). Parliament's debates about infrastructure: An exercise in using dynamic topic models to synthesize historical change. *Technology and Culture*, 60(1), 1-33.
- Hockey, S. (2004). The history of humanities computing. En S. Schreibman, R. Siemens y J. Unsworth (Eds.), *A Companion to Digital Humanities* (pp. 1-19). Wiley Blackwell.
- Hugues, L., Constantopoulos, P. y Dallas, C. (2016). Digital methods in the humanities: Understanding and describing their use across the disciplines. En S. Schreibman, R. Siemens y J. Unsworth (Eds.), *A new companion to digital humanities* (pp. 150-170). Wiley Blackwell.

- Ide, N. (2004). Preparation and analysis of linguistic corpora. En S. Schreibman, R. Siemens y J. Unsworth (Eds.), *A companion to digital humanities* (pp. 289-305). Wiley Blackwell.
- Jiménez-Yáñez, C. (2020). #Chiledespertó: causas del estallido social en Chile. *Revista Mexicana de Sociología*, 82(4), 949-957.
- Joignant, A., Garretón, M., Somma, N. y Campos, T. (Eds.). (2020). *Informe anual Observatorio de Conflictos 2020*. Centro de Estudios de Conflicto y Cohesión Social (COES). <https://coes.cl/wp-content/uploads/Informe-Anual-Observatorio-de-Conflictos-2020-COES.pdf>.
- Jurafsky, D. y Martin, J. H. (2008). *Speech and language processing. An introduction to natural language processing, computational linguistics, and speech recognition*. Prentice Hall.
- Justino, P. y Martorano, B. (2016). Inequality, distributive beliefs and protests: A recent story from Latin America. *HiCN Working Papers*, (218). <https://ideas.repec.org/p/hic/wpaper/218.html>
- Le Deuff, O. (2018). *Digital humanities: History and development*. Wiley Blackwell.
- Liu, A. (2012). Where is cultural criticism in the digital humanities? En M. K. Gold (Ed.), *Debates in the digital humanities* (pp. 490-509). University of Minnesota Press.
- Mahony, S. (2018). Cultural diversity and the digital humanities. *Fudan Journal of the Humanities and Social Sciences*, 11, 371-388.
- Margetts, H., John, P., Hale, S. y Yasseri, T. (2016). *Political turbulence: How social media shape collective action*. Princeton University Press.
- McEnery, T. y Hardie, A. (2012). *Corpus linguistics: Method, theory and practice*. Cambridge University Press.

- McEnery, T. y Wilson, A. (2001). *Corpus linguistics: An introduction*. Edinburgh University Press.
- Micheletti, M. (2002). Consumer choice as political participation. *Statsvetenskaplig Tidskrift*, 105(3), 218-234.
- Miranda Leibe, L., Palanza, V. y Sánchez Staniak, F. (2022). ¿Voto versus protesta? La movilización como mecanismo de participación ciudadana. *Polis: Revista Latinoamericana*, 21(61), 32-58.
- Munck, R. (2020). Social movements in Latin America: Paradigms, people, and politics. *Latin American Perspectives*, 47(4), 20-39.
- Nesi, H. (2016). Corpus studies in EAP. En K. Hyland y P. Shaw (Eds.), *Routledge Handbook of English for Academic Purposes* (pp. 206-217). Routledge.
- Novo, A. (2014). “Consumocracia”: El consumo político como forma de participación de la ciudadanía. *Política y Sociedad*, 51(1), 121-146.
- Ortiz, D. G. (2015). State repression and mobilization in Latin America. En P. Almeida y A. Cordero Ulate (Eds.), *Handbook of social movements across Latin America* (pp. 43-59). Springer.
- Pančur, A. y Šorn, M. (2016). Smart big data: Use of Slovenian parliamentary papers in digital history. *Contributions to Contemporary History*, 56(3), 130-146.
- Parodi, G. (2008). Lingüística de corpus: Una introducción al ámbito. *Revista de Lingüística Teórica y Aplicada*, 46(1), 93-119.
- Stole, D. y Micheletti, M. (2013). *Political consumerism. Global responsibility in action*. Cambridge University Press.
- Stubbs, M. (1993). British traditions in text analysis. From Firth to Sinclair. En M. Baker, G. Francis y E. Tognini-Bonelli (Eds.), *Text and technology: In honour of John Sinclair* (pp. 1-36). John Benjamins Publishing.

- Teorell, J., Torcal, M. y Montero, J. R. (2007). Political participation: Mapping the terrain. En J. W. van Deth, J. R. Montero y A. Westholm (Eds.), *Citizenship and involvement in European democracies: A comparative analysis* (pp. 334-357). Routledge.
- Teubert, W. (2004). Language and corpus linguistics. En M. A. K. Halliday, W. Teubert, C. Yallop y A. Čermáková (Eds.), *Lexicology and corpus linguistics* (pp. 73-112). Continuum.
- Thomas, G. (2005). The qualitative foundations of political science methodology. *Perspectives on Politics*, 3(4), 855-866.
- Tognini-Bonelli, E. (2001). *Corpus linguistics at work*. John Benjamins Publishing.
- Van Dijk, T. A. (2002). Knowledge in parliamentary debates. *Journal of Language and Politics*, 2, 93-109.
- Van Laer, J. y Van Aelst, P. (2009). Cyber-protest and civil society: The internet and action repertoires of social movements. En Y. Jewkes e Y. Majid (Eds.), *Handbook on internet crime* (pp. 230-254). Routledge.
- Verba, S. y Nie, N. H. (1972). *Participation in America: Political democracy and social equality*. Harper & Row.
- Von Bülow, M. y Bidegain, G. (2015). It takes two to tango: Students, political parties and protest in Chile (2005-2013). En P. Almeida y A. Cordero Ulate (Eds.), *Handbook of social movements across Latin America* (pp. 179-194). Springer.
- Zorell, C. V. (2019). *Varieties of political consumerism. From boycotting to buycotting*. Palgrave Macmillan.

Apéndice A

Para limpiar los encabezados de las páginas pares:

- `([a-z+)-\n\"n\"[o-9]+\s+CÁMARA DE DIPUTADOS\n([a-z+|[o-9]+)` sustituir por `\1\2`
- `([a-z+)\n\"n\"[o-9]+\s+CÁMARA DE DIPUTADOS\n([a-z+|[o-9+|[\^w\s])` sustituir por `\1\2`
- `([a-z+)(\:|\,|;|[o-9+))\n\"n\"[o-9]+\s+CÁMARA DE DIPUTADOS\n([a-z+|[o-9]+)` sustituir por `\1\2\3`
- `\n\"n\"[o-9]+\s+CÁMARA DE DIPUTADOS\n` sustituir por nada
- `\n\"n\"[o-9]+\s+CÁMARA DE DIPUTADOS` sustituir por nada

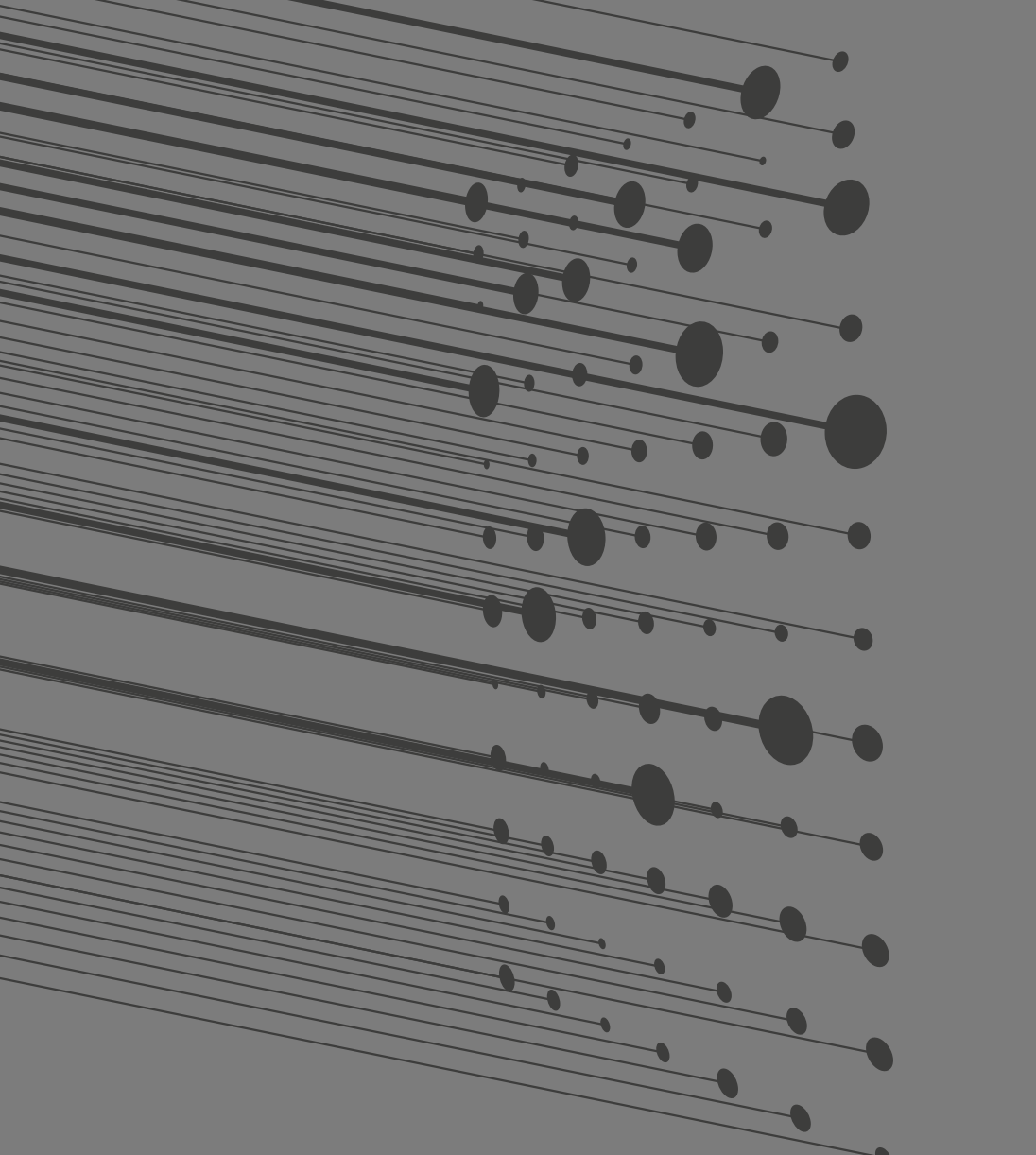
Para limpiar los encabezados de las páginas impares:

- `([a-z+)-\n\"n\"[o-9]+\s+SESIÓN\s[o-9]+ª, EN ([A-Z+|MIÉRCOLES) [o-9]+ DE [A-Z]+ DE [o-9]+\s+[o-9]+\n([a-z+|[o-9]+)` sustituir por `\1\3`
- `([a-z+)\n\"n\"[o-9]+\s+SESIÓN\s[o-9]+ª, EN ([A-Z+|MIÉRCOLES) [o-9]+ DE [A-Z]+ DE [o-9]+\s+[o-9]+\n([a-z+|[o-9+|[\^w\s])` sustituir por `\1\3`
- `([a-z+)(\:|\,|;|[o-9+))\n\"n\"[o-9]+\s+SESIÓN\s[o-9]+ª, EN ([A-Z+|MIÉRCOLES) [o-9]+ DE [A-Z]+ DE [o-9]+\s+[o-9]+\n([a-z+|[o-9]+)` sustituir por `\1\2\4`
- `(á|é|í|ó|ú)\n\"n\"[o-9]+\s+SESIÓN\s[o-9]+ª, EN ([A-Z+|MIÉRCOLES) [o-9]+ DE [A-Z]+ DE [o-9]+\s+[o-9]+\n([a-z+|[o-9]+)` sustituir por `\1\3`

- `(\.)\n\"\\n\"\\s+SESIÓN\s[o-9]+ª, EN ([A-Z]+ |MIÉRCOLES) [o-9]+ DE [A-Z]+ DE [o-9]+\s+[o-9]+\n` sustituir por `\1\n`
- `\"\\n\"\\s+SESIÓN\s[o-9]+ª, EN ([A-Z]+ |MIÉRCOLES) [o-9]+ DE [A-Z]+ DE [o-9]+\s+[o-9]+\n` sustituir por nada

Para limpiar los saltos de línea con palabra partida por guion en el resto del texto:

- `([a-z]+)-\n([a-z]+)` sustituir por `\1\2`
- `([a-z]+)-\n(ñ)` sustituir por `\1\2`
- `(á|é|í|ó|ú)-\n([a-z]|[o-9]+)` sustituir por `\1\2`
- `[:blank:]+-o-\n` sustituir por nada



SOBRE LOS AUTORES

MARIA JOSÉ AFANADOR LLACH es profesora asociada en la Facultad de Artes y Humanidades de la Universidad de los Andes, Colombia, donde coordina la Maestría en Humanidades Digitales desde el 2018. Es editora de *Programming Historian* en español desde el 2016 y miembro del comité base de la Red Colombiana de Humanidades Digitales. Sus temas de investigación son la historia del siglo XVIII, la transición entre Colonia y República, la historia digital, los archivos digitales, las narrativas cartográficas y la brecha digital.

ALBERTO ALCÁNTARA es arqueólogo por la Escuela Nacional de Antropología e Historia y maestrando en Historia por la Universidad Iberoamericana, ambas en la Ciudad de México. Especializado en análisis de radiocarbono para época prehispánica, historia de la ciencia de los siglos XIX y XX, historia digital y divulgación e interpretación del patrimonio. Creador de los proyectos BetoASaber y ArqueoData.

ADRIANA ÁLVAREZ SÁNCHEZ es doctora en Historia, profesora de la Licenciatura en Historia de la Facultad de Filosofía y Letras de la Universidad Nacional Autónoma de México y miembro de la Red de Humanidades Digitales. Ha colaborado con Miriam Peña en proyectos de formación digital, a través de los cuales han impartido talleres para docentes.

JUAN ÁLVAREZ UMBARILA es candidato a Doctor en Estudios Comparados por la Universidad de Lisboa, Portugal. Literato, editor y crítico cultural.

DAVID BORDONABA-PLOU es profesor ayudante doctor en la Universidad Complutense de Madrid, España. Es doctor en Filosofía por la Universidad de Granada. Su investigación se centra en la filosofía experimental del lenguaje, el papel de las intuiciones, las humanidades digitales y cuestiones políticas como el análisis de los debates parlamentarios o la polarización.

ISABEL GALINA RUSSELL es investigadora en el Instituto de Investigaciones Bibliográficas de la Universidad Nacional Autónoma de México. Sus líneas de trabajo están relacionadas con las humanidades digitales, el patrimonio documental digital, la preservación digital, y repositorios y bibliotecas digitales. Es miembro fundador de la Red de Humanidades Digitales.

SUEMI HIGUCHI é doutora em Linguística pela Pontifícia Universidade Católica do Rio de Janeiro e Pesquisadora na Escola de Ciências Sociais da Fundação Getulio Vargas, Centro de Pesquisa e Documentação de História Contemporânea do Brasil.

JULIANA MARQUES é doutora em Sociologia pelo Instituto de Estudos Sociais e Políticos da Universidade do Estado do Rio de Janeiro e Mestre em Sociologia e Antropologia pela Universidade Federal do Rio de Janeiro. É Coordenadora de Pesquisa na Escola de Ciências Sociais da Fundação Getulio Vargas, Centro de Pesquisa e Documentação de História Contemporânea do Brasil.

JIMMY MEDEIROS é doutor em Políticas Públicas, Estratégias e Desenvolvimento pela Universidade Federal do Rio de Janeiro. Coordena o Ensino de Graduação e é Professor de Pós-Graduação

na Escola de Ciências Sociais da Fundação Getulio Vargas, Centro de Pesquisa e Documentação de História Contemporânea do Brasil.

GUILLERMO MORALES es filólogo, bibliotecario y publicista digital; egresado de la Facultad de Filosofía y Letras y del Instituto de Investigaciones Bibliotecológicas y de la Información, de la Universidad Nacional Autónoma de México. Sus áreas de interés son los estudios históricos sobre el libro, y las bibliotecas y su relación con el mundo digital.

DANIEL COELHO DE OLIVEIRA é doutor em Ciências Sociais pelo Programa de Pós-Graduação de Ciências Sociais em Desenvolvimento, Agricultura e Sociedade da Universidade Federal Rural do Rio de Janeiro. Professor no Departamento de Ciências Sociais da Universidade Estadual de Montes Claros (Unimontes), Brasil. Professor permanente do Programa de Pós-graduação em Desenvolvimento Social e do Mestrado Associado da Universidade Federal de Minas Gerais e a Unimontes em Sociedade, Ambiente e Território.

MIRIAM PEÑA PIMENTEL actualmente se desempeña como investigadora en el Instituto de Investigaciones Bibliográficas de la Universidad Nacional Autónoma de México, como especialista en humanidades digitales y estudios literarios, con particular interés en el desarrollo de modelos para la publicación enriquecida y en pedagogía digital para la formación de recursos humanos en las humanidades digitales.

ERNESTO PRIANI SAISÓ es profesor en la carrera en Filosofía de la Facultad de Filosofía y Letras de la Universidad Nacional

Autónoma de México. Desde el 2004 comenzó a interesarse por las humanidades digitales como metodología para el estudio de la historia de la filosofía; actualmente dirige el proyecto Biblioteca Digital del Nuevo Pensamiento Hispánico (bdpn.unam.mx), que inició en el 2008. Participó en la fundación de la Red de Humanidades Digitales, de la cual fue presidente entre el 2018 y el 2024.

NICOLÁS QUIROGA es doctor en historia, investigador del Consejo Nacional de Investigaciones Científicas y Técnicas (Conicet), Argentina, y profesor en la Facultad de Humanidades de la Universidad Nacional de Mar del Plata. Actualmente es director del Centro de Estudios Históricos de la misma universidad. Director del grupo de investigación Movimientos Sociales y Sistemas Políticos en la Argentina Moderna. Investiga sobre los inicios de la web en Argentina.

GIMENA DEL RIO RIANDE es investigadora independiente del Instituto de Investigaciones Bibliográficas y Crítica Textual del Consejo Nacional de Investigaciones Científicas y Técnicas (Conicet), donde dirige el HD Lab, primer laboratorio de investigación argentino en humanidades digitales. Es profesora titular de Nuevas Tecnologías aplicadas a la Historia en la Universidad del Salvador, y dirige la Diplomatura en Humanidades Digitales en la Universidad de Ciencias Empresariales y Sociales. Es miembro fundador de la Asociación Argentina de Humanidades Digitales. Se interesa por la edición filológica digital de textos y el uso de tecnologías de código abierto y sostenibles como forma de democratizar el conocimiento en humanidades digitales.

SERGIO RODRÍGUEZ GÓMEZ es candidato a doctor en Comunicación, Lenguajes e Información por la Pontificia Universidad Javeriana, Bogotá, Colombia. Artista y semiólogo.

BLANCA SERRANO ORTIZ DE SOLÓRZANO es directora de proyectos en el Institute for Studies on Latin American Art y cocuradora de *La fiebre del banano/Banana Craze*. Doctora en Historia del Arte por el Instituto de Bellas Artes de la Universidad de Nueva York.

JUANITA SOLANO ROA es profesora asistente del Departamento de Historia del Arte de la Universidad de los Andes, Bogotá, y cocuradora de *La fiebre del banano/Banana Craze*. Doctora en Historia del Arte por el Instituto de Bellas Artes de la Universidad de Nueva York.

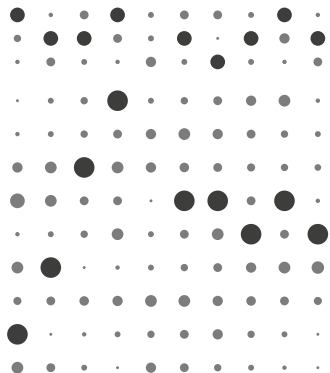
RAFFAELE VIGLIANTI es desarrollador sénior de *software* de investigación en el Instituto de Tecnología en Humanidades de la Universidad de Maryland. Su investigación se centra en las humanidades digitales y el estudio y edición de textos, donde *texto* incluye notación musical. Así mismo, investiga prácticas nuevas y eficientes para modelar y publicar fuentes textuales como recursos digitales innovadores y sostenibles.

{ Este libro inaugura la Colección Humanidades
Digitales de la Facultad de Artes
y Humanidades de la Universidad de los Andes.

Fue compuesto en **Rogliano**, tipografía
diseñada por el chileno Rodrigo López Fuentes,
y **Capitolina**, diseñada por el brasileño
Christopher Hammerschmidt.

Las ilustraciones de la portada y las portadillas
representan con puntos los títulos, usando el
sistema de las tarjetas perforadas utilizadas en los
años sesenta y setenta para ingresar información
e instrucciones a computadoras.

} ;



H o l a m u n d o



Todos los libros de Ediciones Uniandes
a un clic de distancia

Conoce nuestra página web



Escanea el código o visita
ediciones.uniandes.edu.co



Ediciones Uniandes
Vicerrectoría de Investigación y Creación

