

La ética de la información, las nuevas fuentes y metodologías de la historia y las humanidades digitales en el segundo milenio: ‘El diario de Heinrich Witt’ como fuente digital

The ethics of information, new sources, and methodologies in History and Digital Humanities in the second millennium: ‘The Diary of Heinrich Witt’ as a digital source

Alejandro Valdez Tamashiro¹

Pontificia Universidad Católica del Perú

RESUMEN

Las nuevas metodologías aplicadas desde las humanidades digitales permiten la observación de nuevas perspectivas de las fuentes históricas. Desde hace años, se han desarrollado diferentes usos del ámbito digital para el estudio del pasado, y se han encontrado con nuevos problemas éticos sobre la investigación y su impacto en las personas y grupos humanos. El presente artículo trata un caso de estudio en un curso universitario, en donde el diario del comerciante Heinrich Witt y su experiencia en el Perú del siglo XIX son analizados desde la óptica de los mapas de redes sociales y explica cuáles han sido los métodos y limitaciones.

223

¹ Bachiller en Historia por la Pontificia Universidad Católica del Perú (PUCP). Es predocente en cursos de Argumentación e Historia en los Estudios Generales Letras de la PUCP.

Correo: a.valdez@pucp.edu.pe



Palabras clave: humanidades digitales, ética, Heinrich Witt, mapa de redes, Perú, siglo XIX

ABSTRACT

Digital Humanities methodologies enable new perspectives on historical sources. Over the years, various digital applications have emerged for studying the past, raising ethical concerns about research and its impact on individuals and groups. This article presents a case study from a university course, analyzing the diary of merchant Heinrich Witt and his experiences in 19th-century Peru through social network mapping. It explains the methods used and their limitations.

Keywords: Digital Humanities, ethics, Heinrich Witt, network maps, Peru, 19th century

* * *

Las fuentes documentales son la base de la investigación histórica. En el pasado (y en buena parte aún en la actualidad) estas se encontraban, en su mayoría, custodiadas en los archivos de las instituciones públicas y privadas y se conservaban a partir del medio físico en el que se encontraban, el papel. Durante décadas, los historiadores debían consultar estos documentos como base única de las investigaciones y le otorgaban al papel un poder casi absoluto de veracidad sobre los hechos y eventos “reales”. A partir del giro lingüístico y la aceptación académica de fuentes no documentales, es decir, no de papel, se inició un proceso de cambio y ampliación en cuanto a las temáticas de investigación de los historiadores. Áreas como la historia cultural y de las mentalidades, social, de género, etc., aparecieron con fuerza en la historiografía,

por lo que la inclusión de nuevas perspectivas sobre las fuentes fue necesaria.

En su texto sobre las historias de vida, Paul Thompson (1981) menciona que “Los historiadores no son puristas metodológicos, sino *urracas* [énfasis añadido]; ante un problema, *se apoderarán de cualquier evidencia que puedan descubrir y le darán el mejor uso posible* [énfasis añadido]” (p. 118). Esto quiere decir que, debido a la amplitud en el abanico de problemáticas que resolver, los historiadores tuvieron que buscar más allá de los archivos tradicionales para encontrar las fuentes que les permitieran realizar sus investigaciones. Fue en este contexto en que, por ejemplo, la microhistoria y la historia oral se presentaron como formas novedosas de *hacer* la historia, gracias a los avances metodológicos y prácticos en estas áreas.²

En la primera de ellas, Carlo Ginzburg (1992), tomando en cuenta fuentes de corte tradicional (de archivo), buscó la respuesta de un problema general (la cultura popular agraria italiana del s. XVI) a partir de un ejemplo en concreto (el juicio a Menocchio). A partir de la biografía de este personaje, Ginzburg se centró en la reconstrucción de las estructuras sociales y culturales. Esto le permitió al autor encontrar pistas sobre la cultura popular, que de otra manera quedaban ocultas detrás de los silencios de los archivos oficiales. Por otro lado, Willoughby Anderson, en su artículo

² Carlo Ginzburg en *El queso y los gusanos* (1992) utiliza la técnica de la microhistoria, en la que, a partir de una historia particular, es decir, de un personaje o institución, se traza un desarrollo histórico de aquello que lo rodea, pues el individuo está conectado directamente con la sociedad en la que vive. Sobre la historia oral ver Aguirre (2009), Thompson P. (1981) y Thompson H. A. (2013).

“The Presence of the Past: Iconic Moments and the Politics of Interviewing in Birmingham” (2012), realizó una serie de entrevistas a testigos oculares y actores de las protestas de Birmingham de 1963. Para esta autora “desde los años setenta el uso de entrevistas en la historia oral como herramienta para investigar el pasado reciente se ha convertido en una metodología ampliamente aceptada y utilizada” (p. 140, traducción propia).³ Esto les permite a los historiadores acercarse a los sujetos de estudio de una manera que los documentos impresos no permiten y los transforma en agentes históricos, creadores de una fuente nueva. Estas dos formas de *hacer* la historia y la posición frente a ella del historiador (su escritor) trajeron consigo una amplia reflexión sobre las fuentes. Por un lado, se encontraron formas diferentes de ver y analizar los documentos oficiales, mientras que, por el otro, abrieron la posibilidad de que el historiador fuera capaz de crear material histórico en donde él mismo era partícipe del proceso de construcción y elaboración de fuentes históricas.

Esta ampliación de temas y fuentes generó nuevas preguntas, esta vez en relación con la conservación de todos los materiales, tanto nuevos como antiguos, así como de aquellos que tradicionalmente no fueron considerados como dignos de ser custodiados.⁴ Un caso peruano contemporáneo es el de los documentos generados por la Comisión de la Verdad

3 “Since the 1970s the use of oral history interviewing as a tool to research the recent past has become a widespread and well-accepted methodology” (Anderson, 2012, p. 140).

4 En este grupo se enfrentaban las fuentes oficiales y las “informales”. En este segundo grupo se encontraban materiales como posters, afiches, videogramas, panfletos, cancioneros, manuales de cocina, libros de texto escolares, material audiovisual diverso, etc.

y la Reconciliación (CVR), hacia inicios del nuevo milenio, y el problema de su conservación. Carlos Aguirre (2009) mencionó “que una preocupación central de la CVR fue determinar qué destino tendría el acervo documental que ella reuniría durante su funcionamiento” (p. 138). El destino de este fue la creación de un archivo específico en la Defensoría del Pueblo, en donde actualmente se guardan desde actas policiales hasta grabaciones y entrevistas a víctimas del periodo de violencia. Existen dos problemáticas con respecto a estos tipos de fuentes, las orales y las de temática sensible. La primera es sobre la ética del uso y publicación de sus contenidos y la segunda es la conservación y posterior consulta de las mismas.

En cuanto a las fuentes orales, los problemas ocurren a la hora de realizar la entrevista. Heather Ann Thompson, mientras investigaba el motín de la prisión neoyorquina Ática de 1971, se enfrentó con la realidad complicada de este tipo de fuente. La autora narró la experiencia de la siguiente manera:

Era este evento el que yo quería entender así que seguí adelante, saqué mis apuntes, encendí mi grabadora digital y me lancé a la entrevista —una pregunta, dos, luego tres. Pero de repente me encontré mirando hacia arriba, mortificada e insegura sobre cómo continuar. Intentando responder mis preguntas, este sobreviviente de Ática se encontró a sí mismo de regreso en septiembre de 1971 y estaba reviviendo ese terrible momento cuando 500 guardias estatales (...) ingresaron a la prisión disparando. *Es seguro que aprendí mucho sobre Ática en esa primera entrevista, pero no había considerado las consecuencias de preguntarle a alguien que revisitará su tormentoso pasado para que pudiera recuperarlo para mis lectores. Mientras el reloj sonaba cerca, solo podía mirar sin*

poder hacer nada mientras el hombre frente a mí lloraba sin control, mostrando una cara con un dolor inimaginable [énfasis propio]. (2013, traducción propia)⁵

El problema de la *revisitación* de las memorias para la investigación de historia reciente a través de las entrevistas personales es un debate ético que afecta (y ha afectado) a cientos de historiadores interesados en el uso de este tipo de fuentes. ¿Es correcto que el historiador, al buscar testimonios, obligue a los entrevistados a recordar y revivir experiencias traumáticas? ¿Cuál es el papel que cumple el historiador en estos momentos? ¿Qué se debe de hacer cuando los entrevistados sufren contando sus experiencias? Estas son algunas de las interrogantes que pueden surgir a la hora de encararse con los testimonios, pues los entrevistados también se enfrentan con su memoria, sus traumas, aquellos recuerdos que buscaron olvidar. En otro orden de cosas, Willoughby Anderson debió enfrentarse a un evento específico, que se imponía en la memoria de sus entrevistados como el único evento importante del Movimiento por los Derechos Civiles. La autora afirmaba incluso que “la influencia abrumadora de los eventos de esa primavera a veces infringía en mi habilidad de conducir

5 It was this event that I wanted to understand, so I plunged ahead, pulling out my notes, starting my digital recorder, and launching into the interview—one question, two, then three. Suddenly, though, I found myself looking up, mortified and completely unsure how to proceed. In trying to respond to my queries, this Attica survivor found himself back in September 1971 and was reliving that terrible moment when over 500 state troopers clothed in garish yellow rain slickers and thick gas masks stormed the prison with guns blazing. To be sure, I learned much about Attica in that first interview, but I had not at all considered the consequences of asking someone to revisit their tortured past so that I might recover it for my readers. As a clock ticked loudly nearby, I could only look on helplessly as the man before me, face now contorted with unimaginable pain, wept uncontrollably. (Thompson, 2013)

las entrevistas" (2012, p. 146). Así, se puede entender que la historia oral y la realización de las entrevistas constituye una lucha por la memoria, en donde el entrevistado y el entrevistador (el historiador) entran en conflicto no solo entre ellos, sino consigo mismos. El entrevistado lucha contra las memorias reprimidas, de eventos traumáticos que marcaron su vida para siempre, y el historiador lucha con las emociones que surgen al escuchar y observar a sus entrevistados quebrarse en llanto o enojarse, pues esta interacción humana es algo a lo que ellos no están acostumbrados.⁶

Con respecto a la publicación de contenidos sensibles, el caso de Susan Reverby es ilustrativo. Esta investigadora norteamericana, mientras revisaba los archivos en búsqueda de información sobre los estudios Tuskegee sobre la sífilis, encontró documentos que involucraban al gobierno de los Estados Unidos con una serie de experimentos humanos conducidos en Guatemala, relacionados con esta enfermedad. La autora, en entrevista con la Public Broadcasting Service (PBS), contó su experiencia de la siguiente manera: "Abrí la caja esperando encontrar más sobre los Tuskegee, pero no había nada sobre ellos. Más bien, todo lo que había en la caja era este material del estudio de Guatemala. Decía muy claramente 'inoculación, sífilis', y yo estaba destrozada por lo que había allí" (2010, traducción propia).⁷

229

6 Para más complicaciones en la realización de las entrevistas orales, ver Aguirre, 2009, pp. 148-154.

7 "And I opened the box expecting to find more on Tuskegee, and there was nothing on Tuskegee in the box, but, in fact, all there was in the box was this material on the Guatemala study. And it said very clearly, inoculation, syphilis, and I was completely floored by what was there" (Reverby, 2010).

Este tipo de contenido, desconocido hasta entonces, provocó una gran polémica que llevó a que el presidente norteamericano de entonces, Barack Obama, interviniéra personalmente y solicitara la investigación de los documentos a través de la Comisión Presidencial para el Estudio de Problemas Bioéticos en el 2010 (Lyons, 2014, pp. 211-212).⁸ Tomando esto en consideración, si el historiador se encontrara con información confidencial y sensible en sus investigaciones, ¿debería de ser publicada? Si es así, ¿quiénes deberían poder acceder a esa información? ¿Los investigadores o el público en general? Estas mismas interrogantes surgieron en la CVR, pues los documentos que generaron eran de carácter sensible y afectaban a todas las personas involucradas, especialmente porque los eventos investigados eran aún recientes en la memoria y realidad colectiva peruana.⁹ En ambos casos, existe un debate acerca de qué hacer con las fuentes, cómo es que deben de ser publicadas y resguardadas y cuál es el impacto que ellas pueden tener en la realidad. Por un lado, el debate llevó al gobierno norteamericano a pedir disculpas públicas e investigar por su cuenta los hechos. Por el otro lado, a partir de una acción del gobierno peruano se estableció una comisión que investigara y estableciera “la verdad” del Conflicto Armado Interno (CAI).¹⁰ Sus conclusiones y los archivos relacionados

8 Para más detalles sobre la polémica, ver Lyons, 2014, pp. 211-212.

9 Aguirre (2009) menciona que “Tratándose de hechos que han ocurrido muy recientemente y que involucran a personas que todavía están vivas, documentos tan sensibles como los que alberga el Centro de Información podrían haber terminado ‘clasificados’ o ‘reservados’ y, por tanto, fuera del alcance de los investigadores” (p. 159).

10 El Decreto Supremo N° 065-2001-PCM fundó oficialmente la Comisión de la Verdad y la Reconciliación a partir de la Resolución Suprema N° 314-2000-JUS, la cual organizaba un grupo de investigación interinstitucional.

con ellas, en particular los conservados en el Lugar de la Memoria, la Tolerancia y la Inclusión Social (LUM), continúan siendo debatidos y atacados desde varios espacios, académicos y populares (como las redes sociales).¹¹

Por otro lado, en el Perú existe una crisis sustancial sobre las condiciones de conservación y protección de los documentos históricos, la que se ilustra en el caso del Archivo General de la Nación, en el que no solo los materiales, sino también la infraestructura y la situación legal del terreno amenazan al acervo documental que custodian.¹² Esta dramática situación ha estado presente desde hace varias décadas. Rafael Varón escribió en 1986 que los archivos se encontraban en un estado deplorable de conservación, en lugares inadecuados y con diversas falencias que:

11 Una de las últimas polémicas sobre el LUM fue su clausura temporal por la Municipalidad de Miraflores. Ver:

<https://idehpucp.pucp.edu.pe/entrevistas/iris-jave-se-puede-determinar-como-es-que-se-pone-el-foco-mediante-una-medida-administrativa-solamente-con-el-lum-porque-ciertamente-les-representa-una-amenaza-para-lo-que/>

<https://www.infobae.com/peru/2023/03/28/lum-el-dia-en-que-rafael-lopez-aliaga-propuso-cederlo-a-las-ffaa/>

<https://ojopublico.com/derechos-humanos/el-lum-es-el-unico-museo-clausurado-por-un-certificado-ocho-anos>

[https://larepublica.pe/cultural/2023/03/29/lum-el-lugar-de-la-memoria-cerrado-por-la-intolerancia-ministerio-de-cultura-municipalidad-de-miraflores-renovacion-popular-349508 .](https://larepublica.pe/cultural/2023/03/29/lum-el-lugar-de-la-memoria-cerrado-por-la-intolerancia-ministerio-de-cultura-municipalidad-de-miraflores-renovacion-popular-349508)

12 Ver: <https://larepublica.pe/sociedad/2022/05/24/poder-judicial-archivo-general-de-la-nacion-pide-mas-tiempo-y-recursos-ante-proximo-desalojo-palacio-de-justicia>

<https://larepublica.pe/expectaculos/2019/07/28/un-ministerio-a-la-deriva>

<https://www.uarm.edu.pe/noticia/articulo-qu-hacer-con-el-archivo-general-de-la-nacion-agn/>

en nuestro medio se convierte[n] en el mal menor. La falta de instrumentos de investigación cede paso en la conciencia de aquellos pocos responsables que están vinculados directamente al Patrimonio Documental de la Nación, a la destrucción y desaparición definitiva de documentos de insustituible valor histórico. (p. 80)

Es en este sentido que los historiadores se preguntan, ¿qué es lo que sucederá con las fuentes documentales que, como se ha mencionado anteriormente, constituyen la base esencial para la investigación? El propio Varón se preguntaba lo mismo y mencionaba que “Ante tanta adversidad, no debe perderse la esperanza de poder contener algún día la destrucción irreparable de nuestro Patrimonio Documental” (p. 92). En los años recientes, se ha encontrado y aplicado una solución. Mediante el uso de la tecnología y del internet se ha empezado a digitalizar una gran cantidad de información,¹³ a manera de imágenes, texto, gráficos, etc., y se ha garantizado así su conservación *ad infinitum*.¹⁴ Por ejemplo, la página web Fuentes Históricas del Perú reúne una gran cantidad de documentos digitalizados en otros repositorios (como Google Books) y los centraliza en un solo lugar, con lo que facilita su búsqueda y el acceso de los investigadores.¹⁵ Así como este

- 13 La pandemia del COVID-19 impidió el acceso físico a los archivos y bibliotecas, por lo que estas instituciones iniciaron (o continuaron) los esfuerzos de digitalización y preservación de sus materiales de manera casi forzada, pues de lo contrario estaban ante la posibilidad de cerrar sus puertas de manera permanente.
- 14 Teóricamente, mientras el repositorio virtual no desaparezca, el documento se conservará en formato digital para siempre.
- 15 Este emprendimiento de humanidades digitales fue fundado por un grupo de estudiantes de la Universidad Nacional Mayor de San Marcos hace algunos años y se encuentra en constante actualización de datos e interfaz de usuario. Ver: <https://fuenteshistoricaselperu.com/>

emprendimiento peruano, existen cientos de miles de repositorios nacionales e internacionales, buscadores y archivos en línea, los cuales pueden ser fácilmente accesibles con solo apretar un botón en el navegador de internet y que contienen millones de documentos digitalizados, así como videografaciones, cartas, afiches, grabaciones, cintas, música, etc. Todas estas fuentes, antes en peligro de desaparecer, poco a poco están encontrando un medio nuevo en la virtualidad, confiable y con bajas posibilidades de perderse para siempre.

Sin embargo, ante la inundación de datos que la digitalización y los repositorios digitales generan surge una nueva duda: ¿cómo aproximarse a ellos? Aquí, la tecnología nuevamente brinda su apoyo elemental a la hora de consolidar, organizar y analizar la información de las fuentes digitales. Las humanidades digitales y la historia digital nacen de esta manera, como una serie de metodologías necesarias para abarcar este tsunami de datos, analizarlos y entenderlos efectivamente. Para Gary Hall (2001):

Las humanidades digitales pueden ser generalmente entendidas como aquellas que abarcan *todas las actividades académicas de las humanidades que involucran la escritura sobre medios digitales y tecnología y que se enlazan con los procesos de producción, práctica y análisis de medios digitales* [énfasis propio]. Estas actividades pueden incluir el desarrollo de nuevas teorías sobre los medios, la creación de archivos y literatura electrónica interactiva, la construcción de bases de datos y wikis en línea, la producción de galerías de arte y museos virtuales, o la exploración de cómo las diferentes tecnologías modifican la enseñanza y la investigación. (p. 1, traducción propia)¹⁶

16 The digital humanities can be broadly understood as embracing all those

De esta manera, las humanidades digitales permiten una serie de ventajas al emplear herramientas tecnológicas contemporáneas en los niveles metodológico y analítico. Guldí y Armitage (2014) las relacionan con la aparición de la *big data*, corpus en donde se encuentra “desde la decodificación del genoma humano a las billones de palabras de los reportes oficiales producidos en masa por las oficinas gubernamentales” (p. 88). El análisis de esta *big data* necesita ser realizada a través de herramientas tecnológicas y programas computacionales, pues la cantidad de información no es físicamente manejable por un ser humano. Por otra parte, la historia digital:

(...) engloba diversas prácticas históricas, como los esfuerzos por la digitalización en archivos, bibliotecas y museos, la investigación asistida por computadoras, la enseñanza en línea y la diseminación pública y profesional del conocimiento histórico, así como la investigación de ‘lo digital’, las tecnologías computacionales y digitales. (Paju, Oiva y Fridlund, 2020, pp. 3-4, traducción propia)¹⁷

Ambas constituyen parte del entramado metodológico que los historiadores utilizan para tratar a la *big data* y las fuentes

scholarly activities in the humanities that involve writing about digital media and technology, and being engaged in processes of digital media production, practice and analysis. Such activities may include developing new media theory, creating interactive electronic archives and literature, building online databases and wikis, producing virtual art galleries and museums, or exploring how various technologies reshape teaching and research. (Hall, 2001, p. 1)

17 “(...) encompasses diverse historical practices, such as digitization efforts at archives, libraries and museums, computer-assisted research, web-based teaching and professional and public dissemination of historical knowledge, as well as research on the history of ‘the digital’, computers and digital technologies” (Alegre y Alvarado, 2002, p. 1).

que se encuentran accesibles en línea. Es en el cruce de la investigación asistida por computación, de la construcción de base de datos y de la búsqueda de cómo la tecnología influencia al quehacer del historiador donde se introduce el curso de Ética Profesional del Historiador (1HIS09-5005) dictado en la Facultad de Letras y Ciencias Humanas de la Pontificia Universidad Católica del Perú (PUCP). Este curso fue realizado en el ciclo 2023-1 por los doctores Magally Alegre y Patricio Alvarado, y tuvo como finalidad, aparte de la discusión ética sobre el quehacer del historiador como profesional, la realización de un trabajo de campo enfocado en:

la dimensión ética del acceso a la información, tanto a nivel de conservación de la memoria histórica y su acceso a través de medios digitales, como desde las discusiones de la privacidad de datos, el uso de big data para la investigación histórica y los cuestionamientos que surgen desde las herramientas de inteligencia artificial para la creación académica. (2023, p. 1)

Para esto, se enfocó en *El diario de Heinrich Witt*, editado por Ulrich Mücke y publicado por la editorial Brill en el 2016. Este diario narra la extensa vida de un comerciante de Altona, Holstein, y sus viajes, experiencias y relaciones con una serie de personajes de la época. Este texto es particularmente importante para la historia peruana, pues Witt vivió varias décadas en el país y el estilo de narración tipo diario brinda a los investigadores un recuento en primera persona de la vida social del siglo XIX.

235

Este corpus documental, de 10 volúmenes y más de 7000 páginas de extensión, se encuentra disponible en línea, como un texto de libre acceso. El proyecto del curso buscó, a partir de esta fuente, desarrollar herramientas digitales que permi-

tieran la comprensión y análisis de las redes de este comerciante alemán y algunas figuras peruanas importantes. Este objetivo inicial fue fijado a través de una consulta previa a los estudiantes del curso, quienes debatieron sobre los posibles temas de trabajo a partir de una rápida revisión del índice.¹⁸ En la discusión, se planteó la construcción de redes a partir del índice, específicamente a partir de la cantidad de menciones sobre un personaje en particular y la revisión contextual de personajes dentro del texto. Cada uno de estos pasos necesitó la atención y las habilidades digitales de los alumnos para la elaboración de una base de datos.

Tomando como punto de partida el índice del libro, a cada estudiante se le asignó una cantidad de páginas (cinco estudiantes en total), la cual debía de ser revisada manualmente para determinar las cantidades de menciones de los nombres, así como discriminar lugares y “personajes históricos” de las “otras” personas.¹⁹ Para ello se empleó el programa de hojas de cálculo Microsoft Excel 2013, debido a su aparente simpleza. Sin embargo, esta *simpleza* fue engañoso, pues la conversión del texto del PDF al archivo de Excel no fue sencilla. Se encontró la solución copiando y pegando las diversas

236

18 El índice de *El diario de Witt* contiene nombres propios de personas y lugares que son mencionados en este documento. Este está organizado por número de repeticiones, en qué página-volumen se encuentran mencionados y algunos detalles adicionales para ayudar a la comprensión del lector. Se encuentra en orden alfabético y no distingue entre personas y lugares.

19 Se determinó que, para el establecimiento de las redes sociales de Witt, se debía de discriminar los lugares de la base de datos. Asimismo, los personajes “históricos”, como por ejemplo Jesús [Jesucristo], Shakespeare, Aristóteles etc., no fueron incluidos debido a que, claramente, no se encuentran relacionados con Witt al no haber sido sus contemporáneos. Por razones semejantes también se excluyó a la realeza y la nobleza.

entradas del índice desde el visualizador de PDF de Google Chrome hacia la hoja de cálculo de Excel, mas no desde el lector de PDF Adobe Reader, pues de este modo se pegaba todo el texto en una sola celda y no como celdas separadas. Una vez realizado este procedimiento, se encontró un nuevo problema. Las celdas contenían tanto los nombres como los números de volumen y de página, lo que dificultaba la revisión del contenido.

Excel tiene una función llamada *Texto en columnas* dentro de la pestaña de *Datos*. Esta función, como se ve en los anexos, permite la separación del contenido de las celdas individuales según parámetros establecidos.²⁰ Utilizando esta función, se separaron los nombres del resto de números²¹ y se crearon dos columnas. Para separar los números, y así individualizar cada mención específica, se volvió a emplear la misma función, pero con parámetros diferentes.²² De esta manera, se obtuvo una celda de nombre y una cantidad delimitada de celdas con el número de página al lado derecho. Empero, los problemas no acabaron ahí. En el índice, los volúmenes están incluidos como caracteres adicionales en la línea de página, separados solamente por dos puntos (:). Esta disposición, una vez trasladada al archivo de Excel, supuso otra complicación, puesto que la separación de las menciones en

20 Ver en el anexo las figuras 1-4 para una explicación más detallada sobre el uso de esta función y los parámetros utilizados.

21 El formato de pegado empleado creó un carácter invisible y único (a manera de tabulación) entre el nombre y el inicio de los números. Se insertó este carácter en la función y esto permitió separar ambas partes de las celdas en dos: una celda con el nombre y otra celda con todos los números (menciones).

22 En este caso, se insertó como valor de separación la coma (,) entre los caracteres de la celda.

celdas específicas no diferenciaba los números de página de los números de volumen en donde se encontraban tales menciones, o incluso era difícil de apreciar cuáles páginas eran parte de qué volumen. Se estableció, entonces, una ayuda visual mediante la función *Formato condicional*, con la que se determinó que todas las celdas que tuvieran dos puntos (:) fueran resaltadas con color rojo.²³ Así, todas las celdas no resaltadas quedaron al lado derecho de las que sí indicaban los números de páginas y sus respectivos volúmenes. Esto se realizó así porque si se utilizaba la función *Texto en columnas* para separar los caracteres que incluían los dos puntos (:), las nuevas separaciones reescribían las celdas ya existentes y sería difícil diferenciar entre los números de volúmenes y los números de página en el rango entre el 1 y el 10.

Una vez tabulada la tabla, se realizaron dos ejercicios adicionales. El primero de ellos fue la inclusión de un valor de *Sexo* de cada personaje. Para esto se revisaron manualmente las 262 entradas, se revisaron los nombres de cada una y se determinó el sexo de cada personaje según el primer nombre y los sufijos “Mr.” o “Mrs.”. Luego, se determinaron las tres entradas con mayor cantidad de menciones y se procedió a investigarlas para la redacción de sus biografías. Este segundo ejercicio fue el más complicado, pues las tres primeras menciones correspondían a la esposa de Witt, María, y a dos personas de apellido Schutte. Debido a la poca información sobre estos dos últimos, se les cambió por Charles Semper y Melchor Velarde,²⁴ a los que se supuso como personajes más

23 Ver en el anexo las figuras 5-7 para una explicación más detallada.

24 El cambio se realizó teniendo en cuenta la nacionalidad de los personajes, la cantidad de menciones y la cercanía que podían tener con Witt. Charles Semper fue la excepción pues este fue un comerciante alemán, amigo

sencillos de trabajar. A pesar de ello, no fue así. No había información de rápido acceso sobre ellos porque no fueron grandes personalidades públicas. Para poder encontrar datos pertinentes se consultó, en primera instancia, el propio diario de Witt. En segundo lugar, se buscó en diccionarios biográficos y geográficos peruanos.

Para el caso de Melchor Velarde, de nacionalidad peruana, se encontró información indirecta a través de la entrada de su hijo, Jorge Velarde. El resto de información sobre él se extrajo del propio diario. Por otro lado, Charles Semper fue investigado únicamente a partir de la información contenida en el libro.²⁵ Se empleó, como fuente adicional, un recorte de periódico anexado por Witt sobre el fallecimiento de Semper.²⁶ Todos estos ejercicios, incluyendo la tabla tabulada, permitieron a los miembros del curso de Humanidades Di-

de Witt desde su infancia. Melchor Velarde fue un primo de María Sierra, la esposa de Witt, con quien compartió muchas anécdotas a lo largo del diario. Ambos personajes constituyen el cuarto y quinto personaje con más menciones en la tabla, por lo que se les dio prioridad. Al ser Velarde de nacionalidad peruana, se le tomó en consideración por la relativa facilidad con la que se podían encontrar datos relevantes sobre él.

- 25 El caso de Charles Semper fue más complejo de desarrollar debido a su nacionalidad y a la ambigüedad sobre algunos datos de su vida que el propio Witt no aclaraba. Por ejemplo, su nacionalidad, fecha y lugar de nacimiento no son claras, pero sí que a Witt le gustaba jugar ajedrez con Semper en cada una de sus reuniones. Estos datos fueron inferidos a partir de pequeñas referencias sobre él.
- 26 Para la lectura de este recorte de periódico alemán, se utilizó la herramienta digital Transkribus para el reconocimiento por OCR del texto y su conversión a un formato digitalizado (una transcripción). Se emplearon los servicios de traducción de la propia herramienta y de Google Translate para entender el texto del recorte (de alemán a inglés). Esto último es una manera de contrastar las traducciones independientes y aumentar el nivel de precisión de la traducción automática. Vea <https://readcoop.eu/transkribus/old-german-scripts/>

gitales (HD) establecer redes entre el propio Witt y los demás personajes mediante el uso de herramientas digitales, como Python.²⁷ Para generar las redes, se tomaron en cuenta las conexiones que podían establecerse a partir de la distancia entre las palabras, ya se encontraran en la misma página o no lo estuviesen, un procedimiento que redujo el margen de error del sistema Python.²⁸ A partir de esto, se generó una serie de imágenes (representaciones visuales) de las redes entre los personajes y gráficos representativos de los datos. Estas visualizaciones permitieron una comprensión más clara acerca de las relaciones de un comerciante alemán, la sociedad peruana y los procesos sociales del siglo XIX.²⁹

El análisis del diario, como se ha dado a entender, se realizó a partir del trabajo conjunto entre los estudiantes del curso de Ética como de HD y fue un reto para ambos grupos. Por un lado, se encuentra la situación descrita anteriormente para los

27 Se determinó que, para facilitar el análisis inicial de los datos, se agrupaban los volúmenes de la siguiente manera: 1-2, 3-4, 5-6, 7-9 y el 10 por su cuenta.

28 Por ejemplo, si se encontrara el nombre de María de Witt en la última línea de la página “X” y Melchor Velarde en la primera línea de la página “X+1” el sistema, sin el parámetro *Distancia de palabras*, consideraría que ambos personajes no se encuentran conectados entre sí debido a que se encuentran en páginas diferentes. Eso constituye una limitación sustancial, pues cortaría el diario de una manera antinatural y reduciría la eficiencia del análisis automático.

29 Los resultados preliminares, a partir de las imágenes generadas, mostraron la existencia de una serie de relaciones sociales caracterizadas por una fuerte presencia masculina. Sin embargo, las mujeres tenían una considerable presencia en las zonas más cercanas a Witt. Esto llama la atención pues las mujeres, si bien no son el grupo más representativo en las imágenes y gráficos, no se quedan atrás y, en promedio, tienen una tasa de menciones y relaciones más alta que los hombres, pues existen múltiples nombres que son mencionados una cantidad ínfima de veces.

alumnos de Ética del Historiador, pues estos no contaban con los conocimientos necesarios para la realización de las redes y el uso del programa Python. Por otro lado, si bien los estudiantes de HD tenían estas habilidades, no tenían la manera de contextualizar las redes ni volver inteligibles los resultados.³⁰ Asimismo, hubo un problema con su *hardware*. El Dr. Vera, durante una de las sesiones de clase, explicó que debido a que los estudiantes y él mismo debían de utilizar sus propias computadoras (laptops personales) el programa Python tardaba aproximadamente tres horas en procesar y generar las imágenes de las redes solo para un grupo de volúmenes de *El diario de Witt* y que, si existía un error en el programa, este procedimiento debía de ser repetido múltiples veces.

Por cuenta propia, este análisis de redes resultaría imposible de realizar debido a la gran cantidad de datos, relaciones y niveles. Aquí es donde las humanidades digitales y las herramientas de análisis digital (en este caso, computacional) facilitan los procedimientos. La actividad realizada en el curso fue la de crear una base de datos que contuviera tres categorías principales: nombre, sexo y cantidad de menciones. Esta base de datos, insertada en el programa de programación Python, adquiere un valor mayor, pues facilita la visualización de las redes y, en conjunto con el contraste de fuentes cualitativas, el análisis complejo de las relaciones humanas.³¹ El objetivo

30 Este fue un tema de discusión durante las sesiones conjuntas. El Dr. Vera habló de la interdisciplinariedad y del cómo es necesario un trabajo en conjunto, utilizando las habilidades especializadas de ambos grupos para generar productos más amplios y complejos, supliendo las limitaciones de los otros.

31 El procesamiento de datos partió no solo desde la programación *per se* de un programa computacional, sino también se empleó Chat GPT para el análisis de los textos escritos y la relación entre los nombres y el texto en el que se situaba la mención.

de los estudiantes, y por lo tanto del curso, fue desarrollar esta nueva fuente de carácter digital en un lenguaje reconocible por la computadora y más accesible a la vista.³² Este ejercicio constituye una parte esencial de lo que significan las humanidades digitales y la historia digital y su relación con las nuevas generaciones de historiadores.

El historiador francés Emmanuel Le Roy Ladurie escribió en 1979 que la relación entre el historiador y las computadoras era una actividad compleja, larga y repetitiva,³³ pero que resultaba esencial para el futuro, pues “(...) uno puede aventurarse a profetizar, mirando hacia la historia cuantitativa de los ochenta, que, al menos en esta área, el historiador del mañana tendrá que ser capaz de programar una computadora para poder sobrevivir” (1979, p. 6, traducción propia).³⁴

32 La base de datos permite una revisión más corta y directa de los personajes mencionados en *El diario de Witt* sin la necesidad de revisar la totalidad de sus páginas. Si bien se trata de una fuente digital, puesto que se encuentra digitalizado y con la capacidad de emplear OCR en él, la base de datos resume, tabula y organiza una categoría específica de los datos, por lo que facilita el trabajo del investigador.

33 En el primer capítulo explica algunos de los procedimientos que los historiadores debían de realizar en la década de los setenta para poder utilizar las computadoras de su época. La creación de las bases de datos se realizaba de manera física en tarjetas perforadas, las cuales contenían pocos datos y debían de ser producidas en masa por los usuarios. Las computadoras personales (PC) de la época podrían haber facilitado estos procedimientos, pero la capacidad de procesamiento y almacenamiento de datos seguía siendo limitada. Por ejemplo, la RAM de la IBM 5100 (1975) era de *hasta* 64 Kb. La memoria de un iPhone 14 (2023) es de 6 GB, es decir, 6 000 000 de Kb (o 6 291 456 Kb en binario). Es decir, hay una diferencia de 1:98,304 y esto es solo con un teléfono celular. Las PC pueden tener mucho más RAM y, por lo tanto, una mayor capacidad de procesamiento (por ejemplo, la PC del autor de este ensayo tiene 32 GB).

34 “(...) one can venture to prophesy, looking forward to the quantitative history of the 1980's, that in this area at least, tomorrow's historian will

Le Roy Ladurie tuvo razón al considerar a la computadora como una herramienta básica para el historiador del futuro. En la actualidad, las investigaciones no solo se escriben en computadoras, sino que las fuentes mismas se encuentran en línea. Los historiadores no pueden prescindir del mundo digital ni de lo que buscan las humanidades digitales en términos de fuentes, temáticas y herramientas, puesto que con la gran cantidad de información actualmente disponible llegan sobrecargarse de datos. Los alumnos del curso de Ética han tomado el primer paso hacia las humanidades digitales. Se han familiarizado con un programa computacional y, al encontrarse frente a un obstáculo, han encontrado soluciones desde el ámbito digital. Estos ejercicios permitirán, ahora y en el futuro, acercar a las nuevas generaciones de historiadores de la PUCP a ese mundo, uno que se moviliza y avanza sin detenerse ni esperar a nadie. Le Roy Ladurie lo advirtió en los años setenta y, efectivamente, han ocurrido grandes desarrollos tecnológicos desde esa época.³⁵ Las nuevas generaciones que se imbuyen (y han imbuido) de todas estas novedades deben estar atentas y saber cómo utilizarlas para obtener de ellas el máximo provecho y, de este modo, ampliar sus habilidades y el alcance de sus investigaciones en el mundo globalizado de hoy y no, como predecía Le Roy Ladurie, quedarse atrás y no poder sobrevivir.

have to be able to program a computer in order to survive". Le Roy Ladurie, 1979, p. 6)

35 Ver la nota número 33, sobre las diferencias tecnológicas entre el pasado y el presente.

Referencias

- Aguirre, C. (2009). ¿De quién son estas memorias? El archivo de la Comisión de la Verdad y Reconciliación del Perú. *Jahrbuch für Geschichte Lateinamerikas*, 46(1), 135-166.
- Alegre, M., y Alvarado, P. (2023). *Sílabo de ética profesional del historiador*. Lima: PUCP.
- Anderson, W. (2012). The presence of the past: Iconic moments and the politics of interviewing in Birmingham. En C. Bond Potter y R. C. Romano (Eds.), *Doing recent history: On privacy, copyright, video games, institutional review boards, activist scholarship, and history that talks back* (pp. 139-154). University of Georgia Press.
- Aurell, J. (2018). Los efectos del giro lingüístico en la historiografía reciente. *Rilce. Revista de Filología Hispánica*, 20(1), 1-16. <https://doi.org/10.15581/008.20.26710>
- Bertaux, D. (Ed.). (1981). *Biography and society: The life history approach in the social sciences*. Sage Publications.
- Crawford, K. (2021). *Atlas of AI: Power, politics, and the planetary costs of artificial intelligence*. Yale University Press.
-
- Cuartas-Restrepo, J. M. (2017). Humanidades digitales, dejarlas ser. *Revista Colombiana de Educación*, (72), 65-78. *SciELO*, http://www.scielo.org.co/scielo.php?script=sci_arttext&pid=S0120-39162017000100003&lang=es
- Gayol, V., y Melo, J. (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>

- Guldi, J., y Armitage, D. (2014). *The history manifesto*. Cambridge University Press.
- Guy, P. (1996). Historia oral. En J. L. Gil Aristu (Trad.), *Formas de hacer historia*. Alianza Editorial.
- Ginzburg, C. (1992). *The cheese and the worms: The cosmos of a sixteenth-century miller* (Trad. John y Anne Tedeschi). The Johns Hopkins University Press.
- Hall, G. (2011). The digital humanities beyond computing: A postscript. *CultureMachine*, 12, 1-11. <https://culturemachine.net/the-digital-humanities-beyond-computing/>
- Lerma, A., Gutierrez-Vasques, X., Priani, E., & Salmi, H. (2022). Underlying sentiments in 1867: A study of news flows on the execution of Emperor Maximilian I of Mexico in digitized newspaper corpora. *Digital Humanities Quarter*, 16(4), s/p. <http://www.digitalhumanities.org/dhq/vol/16/4/000649/000649.html>
- Le Roy Ladurie, E. (1979). *The territory of the historian* (Trad. Ben y Sian Reynolds). University of Chicago Press.
- Lyons, B. (2014). History, ethics and the Presidential Commission on Research in Guatemala. *Public Health Ethics*, 7(3), 211-224. <https://doi.org/10.1093/phe/phu019>
- Mücke, U. (2017). El diario de Heinrich Witt y la historia del Perú en el siglo XIX. *Histórica*, 41(1), 171-186. <https://doi.org/10.18800/historica.201701.006>
- Mücke, U. (Ed.). (2016). *The Diary of Heinrich Witt (10 vols.)*. Brill. <https://brill.com/edcollbook-oa/title/32515?language=en>
- Paju, P., Oiva, M., y Fridlund, M. (2020). Digital and distant histories: Emergent approaches within the new digital history. En M. Fridlund, M. Oiva, y P. Paju

- (Eds.), *Digital histories: Emergent approaches within the new digital history* (pp. 3-18). Helsinki University Press. <https://doi.org/10.33134/HUP-5-1>
- Reverby, S. (2009). *Examining Tuskegee: The infamous syphilis study and its legacy*. University of North Carolina Press.
- Reverby, S. (2010). Researcher 'Floored' by discovery of intentional infections in Guatemala / Entrevistada por Ray Suarez. *Public Broadcast Service (EE. UU.)*. <https://www.pbs.org/newshour/show/researcher-floored-by-discovery-of-intentional-infections-in-guatemala>
- Thompson, H. A. (2013, 1 de octubre). The art of history: Writing the perilously recent past: The historian's dilemma. *Perspectives on History*. <https://www.historians.org/perspectives-article/writing-the-perilously-recent-past-the-historians-dilemma-october-2013/>
- Thompson, P. (1981). Historia de vida y análisis del cambio social (Trad. Victoria Schusheim). En D. Bertaux (Ed.), *Biography and society: The life history approach in the social sciences* (pp. 289-306). Sage Publications - Studies in International Sociology.
- Varón, R. (1986). Preservación y promoción del patrimonio documental de la Nación. En J. M. Ossio (Ed.), *Patrimonio cultural del Perú: Balance y perspectivas* (pp. 75-92). Fomciencias - Asociación Peruana para el Fomento de las Ciencias Sociales.

Anexo

En este anexo, todas las figuras son capturas de pantalla tomadas del *software* de hojas de cálculo Microsoft Excel 2013.

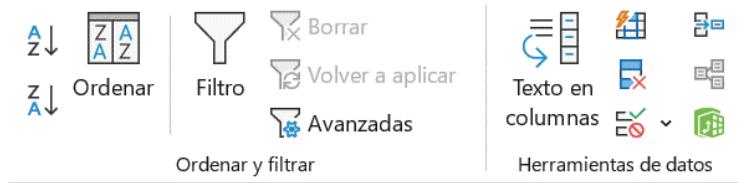
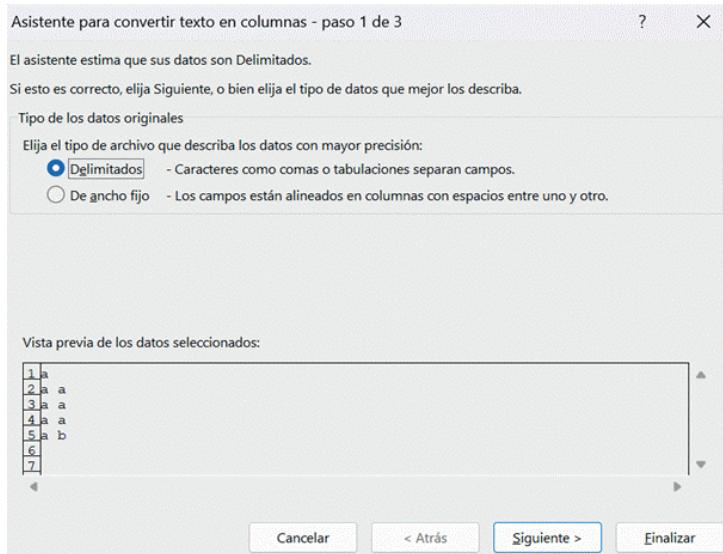


Figura 1. Localización de la función Texto en columnas en la interfaz de Excel. En la captura pantalla de un archivo Excel, se observa la función *Texto en columnas* (*TeC*) en la pestaña de *Datos de la interfaz*.



247

Figura 2. Primer paso para aplicar la función *TeC*. En la captura de pantalla, se aprecia el primer paso para la realización de la función *TeC*. La opción a escoger es *Delimitados*. Esto permitirá introducir opciones más complejas, mientras que *De ancho fijo* es una separación determinada por una distancia específica.

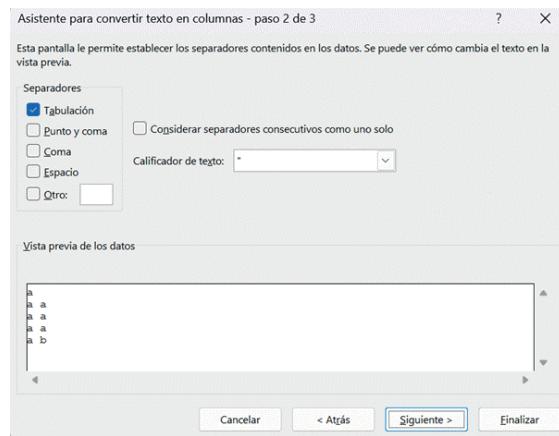


Figura 3. Segundo paso para aplicar la función TeC. En la captura de pantalla se observa el segundo paso para la realización de la función *TeC*: establecer parámetros de separación de las columnas. Para el trabajo se emplearon los separadores *Coma*, *Espacio* y *Otro*. En este último, se colocó el carácter invisible mencionado anteriormente y los dos puntos (:).

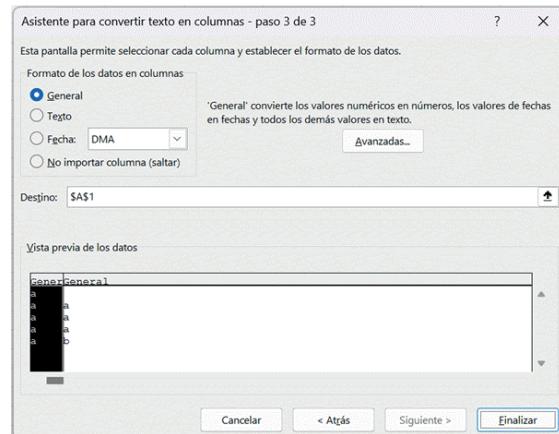


Figura 4. Tercer paso para aplicar la función TeC. En la captura de pantalla se muestra el tercer y último paso de la función *TeC*. Aquí se establece el formato del *output* (datos de salida o resultados). El formato escogido fue el de *Texto*, pues de lo contrario el archivo Excel entendía las celdas con dos puntos (:) como una división matemática o una hora. La vista previa permite observar cómo es que se vería el *output* final.



Figura 5. Localización de la opción Formato condicional. En la captura de pantalla se muestra la opción de *Formato condicional* (FC) en la pestaña *Inicio*. Esta función permite la aplicación de un tipo de formato de celda específico para una selección y que está determinada por una serie de parámetros, los que combinan varios aspectos que podrían estar presentes en el archivo Excel. Ver figura 6 para más detalles.

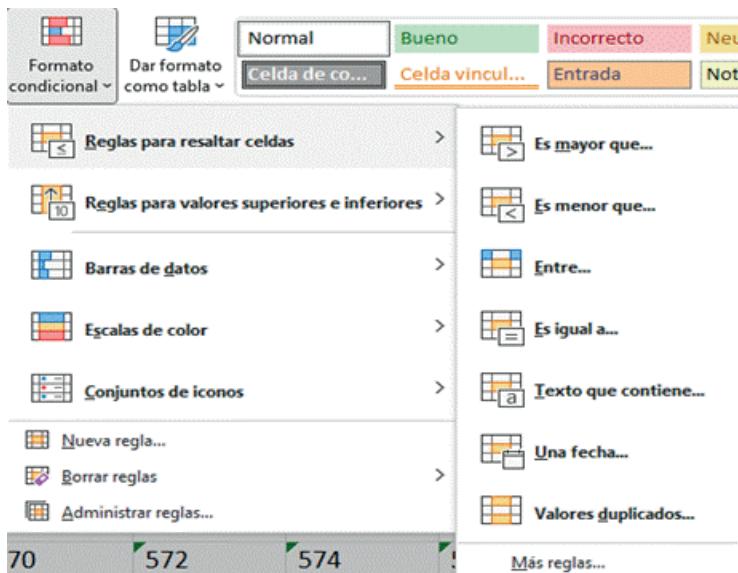


Figura 6. Opciones de Formato condicional. En la captura de pantalla se muestran las diferentes opciones de FC. Para el presente trabajo se utilizó *Reglas para resaltar celdas* y *Texto que contiene...*

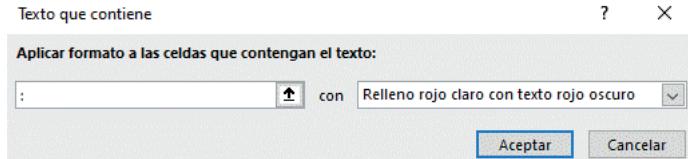


Figura 7. Parámetros para la opción “Texto contiene” en la opción “Formato condicional”. En la captura de pantalla se especifican los parámetros que definirán la condición para la aplicación del formato. Primero, se estableció que todas las celdas que contuvieran dos puntos (:) fueron resaltadas, para así determinar los números de volumen y las posiciones relativas de estos en la tabla de Excel.

* * *

Recibido: 26 de agosto de 2024

Aceptado: 3 de octubre de 2024