

Del testimonio al dato: distant reading y lectura híbrida en la historia oral digital

Sección: Artículo

Recibido: 03/05/2026

Aceptado: 29/06/2026

DOI: 10.46530/virtualis.v18i31.498

From Testimony to Data: Distant Reading and Hybrid Reading in Digital Oral History

Matías Alvarado Leyton

Universidad Gabriela Mistral, Chile

ORCID: 0000-0002-8743-8739

correo: matias.alvarado@ugm.cl

Resumen: Este artículo examina críticamente los límites epistemológicos del *distant reading* en el estudio del testimonio histórico digital. A partir de una investigación teórica de carácter analítico-crítico, se revisan aportes de la historia oral, las humanidades digitales, la crítica de la dataficación y estudios recientes sobre archivos orales digitales, con el propósito de problematizar la transformación del testimonio en dato procesable. El argumento sostiene que las metodologías computacionales no son inadecuadas en sí mismas para analizar corpus testimoniales, pero modifican el estatuto del testimonio al reinscribirlo en un régimen de legibilidad basado en la abstracción, la agregación y la detección de patrones. Si bien esta operación permite identificar recurrencias y relaciones de escala difícilmente visibles mediante la lectura cercana, también puede invisibilizar la densidad narrativa, situada y relacional que constituye el valor historiográfico del testimonio. Frente a esta tensión, se propone una lectura híbrida o estratificada, capaz de articular herramientas computacionales, crítica del corpus e interpretación historiográfica.

Abstract: This article critically examines the epistemological limits of distant reading in the study of digital historical testimony. Based on a theoretical and analytical-critical approach, it reviews contributions from oral history, digital humanities, the critique of datafication, and recent studies on digital oral archives in order to problematize the transformation of testimony into processable data. The article argues that computational methods are not unsuitable in themselves for the analysis of testimonial corpora, but they do modify the status of testimony by reinscribing it within a regime of legibility based on abstraction, aggregation, and pattern detection. Although this operation makes it possible to identify recurrences and large-scale relations that would hardly be visible through close reading alone, it may also obscure the narrative, situated, and relational density that constitutes the historiographical value of testimony. In response to this tension, the article proposes a hybrid or stratified reading approach, capable of articulating computational tools, corpus critique, and historiographical interpretation.

Palabras clave: historia oral; testimonio; distant reading; humanidades digitales; datafificación.

Keywords: oral history; testimony; distant reading; digital humanities; datafication; historical evidence.

Introducción

La historia ha mantenido siempre una relación compleja con sus formas de prueba. Si durante buena parte de su institucionalización moderna el documento escrito ocupó un lugar privilegiado como soporte de verdad, el desarrollo de la historia oral mostró que el conocimiento histórico no podía reducirse a la estabilidad material del archivo ni a la ilusión de una fuente transparente. La voz, la memoria, la experiencia y la subjetividad ingresaron así al campo historiográfico no como simples sustitutos de una ausencia documental, sino como dimensiones constitutivas de la producción de sentido sobre el pasado. El escenario contemporáneo desplaza nuevamente este problema. En la actual ecología digital, marcada por la circulación masiva de textos, imágenes, audios y registros audiovisuales, el testimonio comparece tanto como relato situado de una experiencia como objeto susceptible de inscripción técnica, extracción, clasificación, recombinación y análisis automatizado. La cuestión, por tanto, no consiste solo en preguntarse qué lugar ocupa el testimonio dentro de la historia, sino qué ocurre con su estatuto epistemológico cuando este pasa a formar parte de infraestructuras digitales de almacenamiento, procesamiento y circulación.

Esta transformación debe situarse en el marco más amplio de las sociedades informacionales y de la creciente centralidad de las redes digitales en la organización de la vida social. Como ha señalado Castells (2009), la sociedad red no implica simplemente una ampliación de los canales comunicativos, sino una reorganización profunda de los flujos de información, poder y visibilidad. En ese contexto, las plataformas digitales participan activamente en la producción de las condiciones de legibilidad de los contenidos que alojan, distribuyen y jerarquizan. La voz registrada, el archivo audiovisual, la entrevista digitalizada o el corpus testimonial ya no pueden entenderse al margen de las interfaces que los ordenan, los metadatos que los describen, los algoritmos que los jerarquizan o las herramientas que los vuelven analizables. De este modo, la digitalización amplía el acceso a fuentes históricas y, al mismo tiempo, modifica las operaciones mediante las cuales esas fuentes son seleccionadas, estabilizadas, interpretadas y convertidas en evidencia.

En este punto, la historia oral y las humanidades digitales se encuentran en una zona de contacto especialmente fértil, aunque problemática. La historia oral ha insistido en

que el testimonio debe ser comprendido como una forma de conocimiento situada, relacional y narrativamente mediada. Autores como Portelli (1991), Passerini (1987), Frisch (1990), Thomson (1999) y Thompson y Bornat (2017) mostraron que la fuente oral no vale solo por su capacidad de confirmar hechos, sino también por la forma en que articula memoria, subjetividad, experiencia y significado histórico. Las humanidades digitales, por su parte, han impulsado metodologías orientadas al análisis de grandes corpus culturales, entre ellas el *distant reading*, el modelado de tópicos, el análisis de redes y distintas formas de procesamiento automatizado del lenguaje. Desde la propuesta de Moretti (2013), este giro metodológico ha permitido desplazar la atención desde la lectura intensiva de casos singulares hacia la identificación de patrones, regularidades y estructuras de gran escala. En una línea convergente, Jockers (2013), Underwood (2019) y Piper (2018) han mostrado que los métodos computacionales amplían el volumen de materiales analizables y modifican las preguntas que pueden formularse sobre los corpus culturales.

La literatura reciente sobre archivos orales digitales y análisis computacional de testimonios permite precisar mejor esta zona de cruce. Boyd y Larson (2014) situaron la historia oral digital como un campo en el que acceso, voz, archivo y participación se reorganizan mediante infraestructuras técnicas. Bradley y Puri (2016) advirtieron que la creación de archivos orales digitales abre oportunidades de preservación y consulta, pero también desafíos éticos vinculados con consentimiento, circulación y reutilización de entrevistas. Más recientemente, Pessanha y Salah (2021) exploraron las posibilidades de analizar computacionalmente archivos de historia oral, mientras que Keydar et al. (2026) propusieron un modelo computacional para escuchar testimonios del Holocausto a gran escala, identificando estructuras narrativas típicas y casos divergentes dentro de un corpus testimonial. Estas investigaciones muestran que el análisis computacional del testimonio no constituye una posibilidad meramente especulativa; ya forma parte de las transformaciones metodológicas de los estudios sobre memoria, archivo y violencia histórica. Sin embargo, también evidencian una tensión persistente: cuanto más se amplía la escala del análisis, más necesario se vuelve explicitar las mediaciones que transforman voces situadas en unidades comparables.

La tensión aparece precisamente allí donde estas tradiciones se cruzan. Mientras la historia oral ha construido buena parte de su valor historiográfico sobre la singularidad de la voz, la densidad de la experiencia y la irreductibilidad del relato, el *distant reading* opera mediante procedimientos de abstracción, agregación y modelización que transforman textos o discursos en datos comparables. Esta diferencia remite a un desacuerdo más profundo sobre la naturaleza de la evidencia. Cuando un corpus

testimonial es sometido a métodos de lectura distante, el testimonio deja de ser abordado únicamente como una narración situada y pasa a ser leído como parte de una serie, como recurrencia léxica, como patrón temático o como nodo dentro de una estructura mayor. Esta operación puede abrir posibilidades analíticas relevantes, pero también plantea una pregunta central: ¿qué se gana y qué se pierde cuando la experiencia testimonial es traducida a dato?

El presente artículo examina críticamente esta tensión, analizando los límites epistemológicos del *distant reading* en el estudio del testimonio histórico digital. La hipótesis que orienta el trabajo sostiene que las metodologías computacionales pueden ser pertinentes para el análisis de corpus testimoniales, pero transforman de manera significativa el estatuto del testimonio al reinscribirlo en un régimen de legibilidad basado en la abstracción, la agregación y la detección de patrones. El problema radica en la aplicación no reflexiva de herramientas digitales a fuentes cuya especificidad reside precisamente en su carácter narrativo, situado y mediado. A partir de una revisión teórico-crítica de la historia oral, las humanidades digitales y los estudios sobre cultura digital, se argumenta que el desafío consiste en pensar formas de articulación metodológica capaces de reconocer, al mismo tiempo, la productividad analítica de la escala y los límites epistemológicos de toda reducción del testimonio a dato.

El aporte del artículo se sitúa, entonces, en la formulación de una lectura híbrida o estratificada del testimonio histórico digital. Esta propuesta dialoga con debates previos sobre *distant reading*, macroanálisis, analítica cultural y crítica de los datos humanísticos, pero desplaza la discusión hacia la especificidad de la evidencia testimonial. En lugar de presentar una oposición entre lectura cercana y lectura distante, se propone una secuencia reflexiva entre escalas: exploración computacional del corpus, retorno hermenéutico a los casos, crítica de las mediaciones y reconsideración del patrón desde la densidad histórica de las fuentes. Con ello, se busca contribuir a una reflexión crítica sobre las humanidades digitales desde la disciplina histórica, evitando tanto la desconfianza nostálgica frente a lo computacional como su adopción acrítica como solución metodológica universal.

El artículo se organiza en cuatro secciones principales. En primer lugar, se explicita el método de revisión y análisis conceptual que orienta el trabajo. Luego, el desarrollo se divide en cinco momentos: una reconstrucción del estatuto historiográfico del testimonio en la historia oral; una caracterización del paradigma computacional asociado al *distant reading* y otras metodologías de análisis a gran escala; una problematización de las tensiones epistemológicas que emergen cuando el testimonio

es transformado en dato; una discusión sobre los alcances y potencialidades de dichas herramientas para el trabajo histórico; y, finalmente, una propuesta de lectura híbrida o estratificada que permita articular métodos computacionales e interpretación historiográfica. La sección final sintetiza los resultados conceptuales del argumento y delimita sus proyecciones para una crítica de la evidencia histórica en contextos de dataficación.

Método

El artículo se inscribe en una investigación teórica de carácter analítico-crítico, orientada a examinar las tensiones epistemológicas que emergen cuando el testimonio histórico, tradicionalmente abordado por la historia oral como relato situado de experiencia, es incorporado a regímenes de análisis computacional propios de las humanidades digitales. Se trata de una indagación conceptual destinada a problematizar los supuestos metodológicos que hacen posible dicha incorporación. La opción por un enfoque teórico se justifica porque el problema central del artículo remite al modo en que ciertas operaciones de abstracción, agregación y modelización transforman el estatuto del testimonio como fuente histórica. En esta línea, el trabajo se aproxima a una forma de análisis conceptual y crítico que, como han señalado Booth, Colomb y Williams (2008), permite construir problemas de investigación mediante la revisión, articulación y confrontación de tradiciones bibliográficas distintas. Asimismo, dialoga con la perspectiva de Alvesson y Sköldbberg (2018), para quienes la investigación reflexiva interroga las condiciones interpretativas, teóricas y metodológicas que permiten producir objetos de conocimiento.

La estrategia metodológica se organizó a partir de una revisión crítica de literatura académica especializada, centrada en tres campos de discusión: la historia oral y la teoría del testimonio; las humanidades digitales y las metodologías computacionales aplicadas a corpus culturales; y los estudios recientes sobre dataficación, plataformas e infraestructuras digitales. La selección bibliográfica combinó textos clásicos, necesarios para reconstruir los fundamentos conceptuales de cada debate, con aportes recientes que permiten situar el problema en el marco de las transformaciones digitales contemporáneas. Para ello, se priorizaron publicaciones académicas indexadas, libros publicados por editoriales universitarias o académicas reconocidas y textos ampliamente citados en los debates sobre historia oral, humanidades digitales, dataficación e infraestructuras digitales.

En el caso de la historia oral, se consideraron autores que han problematizado la relación entre testimonio, memoria, subjetividad y verdad histórica, tales como Portelli

(1991), Passerini (1987), Frisch (1990), Grele (1991), Thomson (1999), Perks y Thomson (2016), Abrams (2016), y Thompson y Bornat (2017). En el ámbito de las humanidades digitales, se incorporaron trabajos que han contribuido a definir el giro computacional en las humanidades, entre ellos Moretti (2013), Jockers (2013), Underwood (2019), Piper (2018), Ramsay (2011), Burdick et al. (2012), Berry (2012), Berry y Fagerjord (2017) y Drucker (2011, 2021). A esta discusión se sumaron estudios aplicados a archivos, corpus históricos y testimoniales, como Blevins (2010), Graham et al. (2016), Milligan (2019), Pessanha y Salah (2021) y Keydar et al. (2026). Finalmente, para abordar la dimensión infraestructural y crítica de la cultura digital, se recurrió a aportes sobre dataficación, big data y plataformas, particularmente Kitchin (2014), boyd y Crawford (2012), van Dijck (2014), Gillespie (2014), Plantin et al. (2018), Couldry y Mejias (2019), y D'Ignazio y Klein (2020, 2024).

El procedimiento de análisis consistió en identificar, sistematizar y comparar las categorías centrales que organizan estos campos de discusión. En particular, se trabajó con cuatro ejes conceptuales: fuente, evidencia, mediación y escala. Estas categorías fueron seleccionadas porque permiten ubicar el punto de contacto entre las tres tradiciones revisadas. La historia oral obliga a reconsiderar qué cuenta como fuente y en qué condiciones una voz puede ser tratada como material historiográfico. La teoría de la evidencia permite examinar cómo un relato situado se transforma en prueba histórica. La noción de mediación conecta la entrevista, la memoria, el archivo, la transcripción, la plataforma y el modelo computacional. Finalmente, la escala permite formular la tensión entre lectura cercana y lectura distante, es decir, entre la singularidad narrativa del testimonio y su integración a corpus extensos.

El eje de la fuente permitió examinar cómo la historia oral amplió la noción de fuente histórica al incorporar voces, memorias y experiencias que no siempre encontraban inscripción suficiente en el archivo escrito. El eje de la evidencia permitió interrogar qué se entiende por prueba cuando el testimonio deja de ser leído únicamente como relato singular y comienza a ser tratado como parte de un corpus susceptible de análisis automatizado. El eje de la mediación hizo posible atender a las operaciones narrativas, técnicas, institucionales y algorítmicas que intervienen en la producción, circulación y lectura del testimonio. El eje de la escala permitió problematizar la distancia entre la lectura cercana, propia de una tradición hermenéutica atenta a la singularidad del relato, y la lectura distante, orientada a identificar regularidades en grandes conjuntos de datos. La comparación entre estos ejes permitió reconstruir el punto exacto en que los supuestos de la historia oral y de las humanidades digitales entran en tensión.

Desde el punto de vista analítico, el artículo adopta una lógica de contraste entre tradiciones metodológicas. Por una parte, se reconstruye la manera en que la historia oral ha entendido el testimonio como una forma de conocimiento situada, dialógica y narrativamente mediada. Por otra, se examina cómo las humanidades digitales han desarrollado procedimientos de análisis que transforman textos, discursos y registros culturales en datos comparables, modelables y visualizables. Este contraste permite mostrar que la tensión entre historia oral y *distant reading* remite a una diferencia profunda sobre la naturaleza de la evidencia histórica. En este punto, resultan especialmente relevantes las advertencias de Drucker (2011, 2021) sobre la necesidad de pensar los datos humanísticos como construcciones interpretativas, así como las observaciones de Boyd y Crawford (2012) respecto de las promesas y riesgos epistemológicos asociados al análisis masivo de datos.

A partir de esta operación comparativa, el artículo ofrece una reflexión metodológica orientada a delimitar las condiciones bajo las cuales el análisis computacional de testimonios puede ser historiográficamente productivo. Las categorías de *distant reading*, dataficación, mediación y lectura híbrida permiten preguntar qué ocurre cuando una fuente cuyo valor depende de la voz, la experiencia, la memoria y la situación de enunciación es sometida a procedimientos de abstracción algorítmica. En este sentido, el método seguido sostiene una doble operación: reconoce las posibilidades que ofrecen las metodologías computacionales para detectar patrones, recurrencias y estructuras invisibles a la lectura individual, y advierte que esas mismas herramientas pueden empobrecer el análisis si no se articulan con una reflexión crítica sobre la especificidad testimonial de los materiales estudiados. Así, el artículo se sitúa dentro de una perspectiva reflexiva de las humanidades digitales, entendidas no solo como un conjunto de técnicas, sino como un campo de disputa sobre las formas contemporáneas de producir, validar y comunicar conocimiento histórico.

Desarrollo

El desarrollo del artículo se organiza en cinco apartados porque la tensión entre historia oral y metodologías computacionales compromete distintas dimensiones del trabajo historiográfico. Cuando el testimonio histórico es incorporado a regímenes de lectura distante, modelización o procesamiento automatizado, se transforma la escala del análisis, pero también la forma en que se define la fuente, se reconoce la evidencia y se justifica la interpretación. Por ello, la división propuesta distingue cinco momentos del argumento. En primer lugar, se reconstruye el estatuto del testimonio en la historia oral, atendiendo a su carácter narrativo, situado y relacional. En segundo lugar, se examina el paradigma computacional de las humanidades digitales,

especialmente en torno al *distant reading*, la dataficación y la abstracción del texto. En tercer lugar, se analiza el punto de fricción entre ambos campos: la conversión del testimonio en dato. En cuarto lugar, se consideran los alcances efectivos de las metodologías computacionales, evitando reducir el argumento a una crítica defensiva frente a lo digital. Finalmente, se propone una lectura híbrida o estratificada del testimonio, capaz de articular escala, interpretación y mediación. Esta estructura permite avanzar desde la reconstrucción conceptual hacia la problematización epistemológica y, desde allí, hacia una propuesta metodológica situada dentro de las humanidades digitales, pero atenta a las exigencias propias del conocimiento histórico.

El testimonio como forma de conocimiento histórico

La historia oral supuso una de las intervenciones más significativas en la reconsideración del estatuto de la fuente histórica durante el siglo XX. Su importancia radicó en ampliar el repertorio documental disponible para el historiador y en cuestionar la jerarquía moderna que había situado al documento escrito, institucionalmente archivado y materialmente estable, como forma privilegiada de acceso al pasado. En este sentido, la fuente oral fue incorporada como una forma específica de conocimiento histórico, capaz de dar cuenta de memorias, experiencias y subjetividades que no siempre encontraban lugar en los registros oficiales. Como ha señalado Grele (1991), la entrevista de historia oral constituye una narrativa conversacional producida en una situación relacional concreta. La voz que allí aparece es el resultado de una interacción situada entre entrevistador, entrevistado, memoria, lenguaje y contexto social. Esta advertencia impide reducir el testimonio a una simple fuente de datos verificables y obliga a considerarlo como una forma de producción de sentido.

Desde esta perspectiva, el aporte de la historia oral consistió en desplazar el problema de la verdad testimonial desde la mera correspondencia factual hacia una comprensión más compleja de la memoria, la narración y la experiencia. Portelli (1991) formuló este punto con particular claridad al sostener que el valor de las fuentes orales descansa tanto en aquello que permiten saber sobre los hechos como en lo que revelan sobre sus significados. Una fuente oral puede equivocarse en una fecha, alterar una secuencia o condensar acontecimientos distintos y, aun así, ofrecer una verdad histórica relevante sobre expectativas, traumas, identidades o formas de conciencia colectiva. De manera semejante, Passerini (1987) mostró que los silencios, las omisiones y las reformulaciones del recuerdo pueden ser leídos como huellas de procesos históricos más amplios, donde la experiencia individual se encuentra atravesada por marcos

culturales, políticos y afectivos. La subjetividad deja así de figurar como un obstáculo para el conocimiento histórico y pasa a ser una de sus dimensiones constitutivas.

Esta reconsideración del testimonio también alteró la relación entre autoridad historiográfica y experiencia vivida. Frisch (1990) propuso la noción de autoridad compartida para dar cuenta del carácter necesariamente negociado de la producción de sentido en la historia oral. El historiador no extrae pasivamente información de un sujeto que recuerda, pero tampoco puede disolver su responsabilidad interpretativa en la voz del entrevistado. Entre ambos se configura un espacio de producción histórica en el que memoria, escucha, edición, archivo y análisis participan conjuntamente. Thomson (1999) reforzó esta perspectiva al mostrar que las fuentes orales poseen un valor empírico y subjetivo a la vez: permiten acceder a acontecimientos, pero también a los modos en que esos acontecimientos son recordados, resignificados y narrados en el tiempo. Abrams (2016), en una línea convergente, ha insistido en que la historia oral debe entenderse como un encuentro narrativo en el que el pasado es constantemente rearticulado desde el presente. Esto complejiza su valor como fuente, porque obliga a analizar lo dicho y las condiciones en que algo puede ser recordado, formulado, escuchado y reconocido como significativo.

De este modo, la historia oral preparó una crítica decisiva a toda concepción sustancial de la fuente histórica. El testimonio vale por hacer visible la historicidad de toda mediación. Como han señalado Perks y Thomson (2016), el trabajo con fuentes orales exige atender simultáneamente a la producción, conservación, circulación y reutilización de las entrevistas. Esto resulta especialmente importante para el problema que aquí interesa, porque muestra que la historia oral nunca fue simplemente una metodología de recolección de voces, sino también una teoría práctica de la mediación testimonial. Su aporte más profundo consistió en ampliar el archivo y en mostrar que la prueba histórica depende de operaciones de inscripción, interpretación y validación. Precisamente por ello, el encuentro entre historia oral y humanidades digitales resulta tan productivo como problemático: ambas trabajan con mediaciones, aunque las conceptualizan de maneras distintas.

El paradigma computacional y la abstracción del texto

Las humanidades digitales han introducido una transformación metodológica significativa al proponer formas de análisis orientadas a grandes volúmenes de información cultural. En este marco, el *distant reading* formulado por Moretti (2013) ocupa un lugar central, no solo por la influencia que ha tenido en los estudios literarios, sino porque expresa un desplazamiento más amplio en las formas de lectura y

conocimiento: frente a la imposibilidad de leerlo todo de manera intensiva, se propone analizar patrones, series, géneros, frecuencias y estructuras de gran escala. La operación redefine la relación entre lectura y distancia. Allí donde la tradición hermenéutica había privilegiado la cercanía con el texto singular, el *distant reading* sostiene que ciertos fenómenos solo se vuelven visibles al aumentar la escala y renunciar parcialmente a la lectura directa. En esta misma línea, Jockers (2013) propuso el concepto de macroanálisis para referirse al uso de métodos computacionales capaces de detectar regularidades en corpus extensos, mientras que Underwood (2019) mostró cómo los modelos estadísticos pueden contribuir a formular nuevas preguntas sobre cambio histórico, género, recepción y periodización.

Este giro computacional no debe ser entendido como una simple incorporación instrumental de software o herramientas de procesamiento de texto. Implica una modificación de las condiciones bajo las cuales los objetos culturales se vuelven analizables. Para que un texto pueda ser procesado computacionalmente, debe ser transformado en datos: tokens, frecuencias, vectores, matrices, tópicos, redes o unidades comparables. Como ha señalado Kitchin (2014), la dataficación consiste precisamente en traducir aspectos del mundo social y cultural a formatos cuantificables, procesables y susceptibles de análisis algorítmico. En el campo de las humanidades, esta operación permite identificar patrones que la lectura individual difícilmente podría advertir, pero también implica una forma específica de abstracción. El texto deja de ser abordado inicialmente como una unidad densa de sentido y pasa a ser descompuesto en elementos formalizables. Piper (2018) ha insistido en que esta operación puede abrir preguntas nuevas sobre la cultura escrita, siempre que se reconozca que toda lectura computacional supone decisiones previas sobre qué contar, cómo clasificar y qué relaciones considerar relevantes.

La abstracción es una de las condiciones de posibilidad del paradigma computacional. Ramsay (2011) planteó que el uso de algoritmos en humanidades debe entenderse como una forma de lectura transformativa, en la medida en que el procedimiento computacional reorganiza el objeto y produce una nueva configuración interpretativa. En un sentido semejante, Drucker (2011) ha advertido que los datos humanísticos no son datos “dados”, sino construcciones interpretativas que dependen de decisiones de modelado, representación y visualización. Esta advertencia resulta fundamental para el problema del testimonio, porque impide asumir que un corpus testimonial digitalizado existe simplemente como materia neutral a la espera de ser procesada. Antes de todo análisis computacional hay operaciones de selección, transcripción, codificación, normalización y clasificación. En consecuencia, las humanidades digitales desplazan la mediación interpretativa hacia distintos momentos del proceso

investigativo: la construcción del corpus, la definición de variables, el diseño del modelo y la interpretación de resultados.

El debate sobre lectura cercana y lectura distante ha dado lugar, además, a propuestas metodológicas intermedias. En los estudios literarios, el macroanálisis no elimina la interpretación de casos, sino que reorganiza el acceso a ellos a partir de patrones cuantitativos (Jockers, 2013). En la historia digital, el macroscopio del historiador permite explorar grandes masas documentales sin renunciar a la contextualización histórica de los hallazgos (Graham et al., 2016). En la analítica cultural, Manovich (2020) ha planteado que la visualización y el análisis de grandes conjuntos de objetos culturales requieren articular escala, interfaz e interpretación. Berry (2012) y Berry y Fagerjord (2017) han insistido, por su parte, en que las humanidades digitales constituyen una forma de conocimiento crítico sobre la cultura digital y no únicamente un repertorio de técnicas. Estas propuestas resultan relevantes para pensar el testimonio histórico, porque muestran que la oposición entre lectura cercana y distante puede ser reemplazada por procedimientos de ida y vuelta entre corpus, patrón, caso e interpretación.

En este punto, el paradigma computacional muestra tanto su potencia como su límite. Su potencia reside en que permite formular preguntas de escala, detectar regularidades discursivas, comparar conjuntos amplios de materiales y hacer visibles relaciones que permanecen opacas para la lectura cercana. Su límite aparece cuando esa capacidad de abstracción se confunde con una forma superior de objetividad. Como advirtieron Boyd y Crawford (2012), el entusiasmo por el *big data* suele apoyarse en la idea de que grandes cantidades de datos permiten acceder de manera más directa o completa a la realidad social. Sin embargo, los datos son producidos desde infraestructuras, categorías, dispositivos y finalidades específicas. Esta crítica no invalida las metodologías computacionales; exige situarlas. En el caso de la historia, y especialmente en el estudio del testimonio, la pregunta decisiva apunta al tipo de transformación que sufre la fuente cuando se vuelve procesable bajo esa lógica.

Del testimonio al dato: tensiones epistemológicas

La tensión principal entre historia oral y *distant reading* aparece cuando el testimonio, entendido por la primera como relato situado de experiencia, es transformado por el segundo en una unidad de análisis agregable. Esta operación afecta la forma misma en que el testimonio puede ser leído y valorado. En la historia oral, el testimonio incluye el modo en que recuerda, narra, calla, organiza temporalmente la experiencia y produce sentido desde una posición determinada. Cuando ese testimonio pasa a formar

parte de un corpus computacional, tiende a ser fragmentado en unidades comparables: palabras, temas, recurrencias, entidades, emociones, relaciones o secuencias. Esta traducción puede ser metodológicamente productiva, aunque introduce una pérdida parcial de densidad contextual. El problema surge cuando la especificidad del testimonio queda subsumida bajo una lógica que privilegia la regularidad sobre la singularidad y la estructura sobre la situación de enunciación.

En términos epistemológicos, esta tensión remite a la diferencia entre experiencia y patrón. La historia oral ha insistido en que el testimonio contiene una dimensión experiencial inseparable de su forma narrativa. Ricoeur (2004) señalaba que el testimonio implica siempre una relación entre memoria, relato, confianza y reconocimiento; quien testimonia comunica información y se presenta como alguien que estuvo allí, que recuerda y que pide ser creído. Esta estructura de confianza, incluso cuando debe ser sometida a crítica, forma parte de la naturaleza misma del testimonio. El análisis computacional tiende a operar sobre materiales desprendidos de esa escena testimonial, especialmente cuando trabaja sobre transcripciones, metadatos o corpus normalizados. En ese tránsito, la voz puede convertirse en texto, el texto en dato y el dato en patrón. Cada paso permite nuevas operaciones analíticas, pero también desplaza algo de la situación original de enunciación.

Esta dificultad se intensifica cuando se considera que los testimonios digitales circulan como objetos inscritos en plataformas, repositorios, motores de búsqueda y sistemas de recomendación. Van Dijck (2014) ha mostrado que la dataficación es también una ideología que tiende a convertir la vida social en datos explotables y calculables. En una línea cercana, Couldry y Mejias (2019) han propuesto pensar el colonialismo de datos como una nueva forma de apropiación de la experiencia humana mediante infraestructuras digitales. Aunque estas discusiones no se refieren exclusivamente al testimonio histórico, resultan útiles para advertir que la conversión de voces y relatos en datos no es una operación neutral. Cuando un testimonio es digitalizado, etiquetado, transcrito, indexado o procesado, ingresa en un régimen de legibilidad que define qué aspectos de su contenido serán recuperables, comparables o visibles. Lo que no se ajusta a esas categorías puede quedar relegado a un margen analítico.

La dimensión infraestructural de este proceso exige una atención particular. Gillespie (2014) ha mostrado que los algoritmos no solo ordenan información, sino que participan en la producción de relevancia pública, visibilidad y acceso. Plantin et al. (2018), al vincular los estudios de infraestructura con los estudios de plataformas, han señalado que las plataformas digitales operan como infraestructuras culturales que estabilizan formas específicas de circulación, dependencia y control. Aplicada al

testimonio, esta discusión permite comprender que la entrevista digitalizada no se vuelve analizable únicamente por estar disponible en línea. Su legibilidad depende de bases de datos, sistemas de indexación, estándares de metadatos, criterios de segmentación, motores de búsqueda, protocolos de acceso y modelos de procesamiento. En consecuencia, el análisis computacional del testimonio debe incorporar una crítica de la infraestructura que vuelve posible la lectura distante.

De ahí que la crítica al *distant reading* aplicado al testimonio deba formularse como advertencia frente a la pérdida de mediaciones. El testimonio es una cadena de producción, inscripción, edición, circulación y validación. Esa cadena incluye al sujeto que recuerda, la situación de entrevista, el dispositivo de registro, la transcripción, el archivo, la interfaz, los criterios de búsqueda y los marcos interpretativos desde los cuales será leído. Cuando el análisis computacional trabaja únicamente sobre el producto final —por ejemplo, una transcripción limpia o un corpus lematizado— corre el riesgo de invisibilizar las operaciones que hicieron posible ese material. Como advierte Drucker (2021), toda práctica digital humanística requiere atender a los procesos de modelado que convierten los materiales culturales en objetos analizables. En el caso del testimonio, esta advertencia es especialmente relevante, porque el sentido reside en el contenido verbal y en la situación relacional y técnica que lo produce.

La tensión puede formularse del siguiente modo: el *distant reading* permite ver el testimonio como parte de una serie, mientras que la historia oral recuerda que cada testimonio es también un acontecimiento narrativo singular. Entre ambas operaciones existe una diferencia epistemológica que debe ser explicitada. Si se trabaja con testimonios como datos, es necesario reconocer que se ha producido una traducción; y toda traducción conserva, transforma y pierde algo del objeto inicial. La pregunta central apunta a las condiciones bajo las cuales ese análisis sigue siendo historiográficamente responsable. Esto exige abandonar tanto la confianza ingenua en el dato como la defensa nostálgica de una oralidad pura. El testimonio siempre estuvo mediado; lo que cambia en el presente digital es la densidad, escala y opacidad de esas mediaciones.

Alcances y potencialidades del análisis computacional

Reconocer los límites epistemológicos del *distant reading* permite valorar con mayor precisión sus aportes al estudio histórico. En el caso de corpus testimoniales amplios, las herramientas de análisis textual pueden identificar recurrencias temáticas, desplazamientos léxicos, cambios de énfasis, asociaciones semánticas o patrones de

enunciación que serían difíciles de percibir mediante lectura cercana exclusivamente. Milligan (2019) ha mostrado que la abundancia documental propia de la era digital ha obligado a los historiadores a desarrollar nuevas formas de lectura, búsqueda y selección, pues el problema ya no es solo la escasez de fuentes, sino también su exceso. En este sentido, el análisis computacional puede funcionar como una estrategia de orientación dentro de archivos extensos, permitiendo detectar zonas de interés, contrastar hipótesis preliminares o formular preguntas que luego deberán ser abordadas mediante interpretación cualitativa.

En el ámbito de la historia reciente, la memoria y los archivos testimoniales, esta posibilidad resulta particularmente relevante. La digitalización de colecciones orales ha ampliado las condiciones de acceso y preservación de entrevistas, y ha modificado las formas de búsqueda, segmentación, indexación y reutilización de esos materiales, planteando nuevas oportunidades metodológicas y desafíos éticos para la historia oral digital (Boyd y Larson, 2014; Bradley y Puri, 2016). Grandes colecciones de entrevistas, testimonios judiciales, archivos de derechos humanos, repositorios audiovisuales o memorias digitalizadas pueden contener patrones discursivos que exceden la percepción individual del investigador. Un ejemplo especialmente ilustrativo lo ofrecen los archivos orales digitalizados sobre violencia política, genocidio o experiencias traumáticas, como el Visual History Archive de la USC Shoah Foundation, que reúne más de 60 000 testimonios audiovisuales de sobrevivientes y testigos del Holocausto y otros genocidios, y que permite buscar en la colección mediante términos, temas, metadatos y segmentos específicos de las entrevistas (USC Shoah Foundation, s. f.).

En este tipo de corpus, las herramientas de búsqueda avanzada, etiquetado temático, análisis de frecuencia, modelado de tópicos, análisis de entidades o procesamiento multimodal pueden ayudar a identificar zonas discursivas que luego requieren una lectura cercana e históricamente situada. Pessanha y Salah (2021) han mostrado que los archivos de historia oral pueden ser explorados computacionalmente mediante procedimientos capaces de reconocer patrones, relaciones y dimensiones narrativas difíciles de observar mediante lectura individual. Keydar et al. (2026), por su parte, desarrollaron un modelo computacional de escucha aplicado a testimonios del Holocausto que busca identificar estructuras narrativas típicas y desviaciones significativas dentro de un conjunto amplio de entrevistas. La relevancia metodológica de estos trabajos no reside únicamente en la aplicación de técnicas computacionales, sino en la formulación de un problema propiamente testimonial: cómo analizar patrones sin borrar la singularidad de las voces que componen el corpus. En un sentido más amplio, el modelado de tópicos y otras formas de análisis textual han sido

utilizados en historia para explorar grandes masas documentales, identificar agrupamientos temáticos emergentes y abrir preguntas que posteriormente deben ser interpretadas desde marcos historiográficos específicos (Blevins, 2010; Graham et al., 2016; Milligan, 2019; Yang et al., 2011).

La utilidad del análisis computacional también reside en su capacidad para desplazar ciertas intuiciones historiográficas. Allí donde la lectura cercana puede privilegiar casos particularmente conmovedores, excepcionales o narrativamente densos, los métodos de escala pueden mostrar recurrencias menos visibles, silencios estructurales o distribuciones desiguales dentro de un conjunto documental. En este sentido, el *distant reading* puede contribuir a descentrar la autoridad del caso ejemplar y obligar a considerar patrones colectivos. Esto es especialmente significativo para el estudio del testimonio, porque permite preguntar no solo qué dice una voz singular, sino cómo ciertas formas de decir, recordar o callar se distribuyen socialmente. Como han señalado D'Ignazio y Klein (2020, 2024), el trabajo computacional en humanidades debe atender también a aquello que los datos no registran fácilmente: ausencias, sesgos, exclusiones y formas de desigualdad incorporadas en los propios archivos. Así, el análisis de escala puede ser una herramienta crítica, siempre que no confunda frecuencia con importancia ni visibilidad cuantitativa con relevancia histórica.

Por ello, el potencial del *distant reading* aplicado al testimonio debe entenderse como complementario y no sustitutivo. Su valor radica en reorganizar el campo de preguntas posibles. Puede ayudar a identificar patrones de memoria, comparar corpus testimoniales de distintos periodos, mapear transformaciones discursivas o detectar recurrencias afectivas y temáticas. Pero esos resultados requieren siempre una vuelta a la lectura cercana, a la contextualización histórica y a la crítica de las condiciones de producción del corpus. Como ha sostenido Manovich (2020), la analítica cultural permite estudiar grandes conjuntos de objetos culturales, pero su productividad depende de la capacidad de articular visualización, cuantificación e interpretación. En el caso del testimonio histórico, ningún gráfico, modelo o red agota el sentido de la fuente. La herramienta abre una pregunta; el trabajo historiográfico reconstruye mediaciones, contextos y significados.

Hacia una lectura híbrida del testimonio

A partir de lo anterior, se vuelve necesario pensar una forma de lectura híbrida o estratificada del testimonio histórico digital. Esta propuesta parte de una premisa básica: la lectura cercana y la lectura distante iluminan dimensiones distintas de los corpus testimoniales contemporáneos. La primera permite atender a la densidad

narrativa, afectiva y contextual de la experiencia, pero puede quedar limitada frente a archivos masivos o colecciones extensas. La segunda permite detectar patrones de gran escala, aunque corre el riesgo de abstraer el testimonio de sus condiciones situadas de producción. Una lectura híbrida organiza un tránsito reflexivo entre escalas. En primer lugar, el análisis computacional puede operar como una fase exploratoria destinada a identificar recurrencias, agrupamientos o anomalías dentro del corpus. En segundo lugar, esas zonas de interés deben ser sometidas a lectura cercana, atendiendo a los contextos de enunciación, a las formas narrativas y a las condiciones históricas de producción. Finalmente, los hallazgos deben ser devueltos al conjunto para evaluar si el caso singular permite matizar, discutir o complejizar el patrón inicialmente identificado.

Esta lectura estratificada dialoga con propuestas metodológicas desarrolladas en otros campos disciplinares. En los estudios literarios, el *distant reading* ha sido discutido como una ampliación de la lectura, no como su cancelación (Moretti, 2013; Jockers, 2013; Underwood, 2019). En la historia digital, el trabajo con grandes corpus ha sido pensado como una estrategia para navegar la abundancia documental y producir nuevas preguntas interpretativas (Graham et al., 2016; Milligan, 2019). En las humanidades digitales críticas, Drucker (2011, 2021), Berry (2012) y Berry y Fagerjord (2017) han insistido en que las tecnologías de análisis no deben separarse de los supuestos epistemológicos que las organizan. La lectura híbrida propuesta en este artículo recoge esas discusiones, pero las reorienta hacia el problema testimonial: la necesidad de articular patrón y experiencia, escala y voz, modelo y mediación.

Esta lectura estratificada exige tratar los datos humanísticos como construcciones interpretativas y no como evidencias inmediatas. Drucker (2011) propuso distinguir entre datos y “capta” para subrayar que aquello con lo que trabajan las humanidades no es simplemente recibido, sino seleccionado, construido y capturado desde una perspectiva. Esta idea resulta particularmente pertinente para el estudio del testimonio, porque obliga a reconocer que todo corpus testimonial digital es el resultado de decisiones previas: qué voces fueron registradas, quién las entrevistó, en qué condiciones, cómo fueron transcritas, qué fragmentos fueron preservados, qué metadatos se añadieron y qué aspectos quedaron fuera del registro. En este sentido, una metodología híbrida debe incorporar una crítica del corpus. Antes de analizar patrones, es necesario preguntar cómo se produjo el conjunto documental, qué ausencias contiene, qué sesgos hereda y qué formas de experiencia resultan más o menos legibles dentro de su arquitectura.

La lectura híbrida también permite recuperar, en otro nivel, una de las intuiciones fundamentales de la historia oral: el testimonio es contenido y relación. En contextos digitales, esta relación se vuelve más extensa y más compleja, porque involucra al entrevistador y al entrevistado, y también a archivos, plataformas, interfaces, algoritmos, usuarios y comunidades interpretativas. Latour (2005) ha insistido en que los fenómenos sociales deben ser comprendidos como redes de asociaciones heterogéneas, y esta perspectiva resulta útil para pensar el testimonio digital como un objeto distribuido entre actores humanos y no humanos. Una entrevista conservada en un archivo digital, un testimonio alojado en una plataforma audiovisual o una memoria publicada en redes sociales no son fuentes aisladas, sino nodos dentro de sistemas de circulación y validación. Por eso, una lectura híbrida debe atender simultáneamente a la voz, al corpus y a la infraestructura que vuelve esa voz accesible, clasificable y analizable.

En este sentido, el concepto de testimonio críticamente mediado puede funcionar como principio articulador. Pensar el testimonio de esta manera implica reconocer que su valor historiográfico reside en la reconstrucción de las mediaciones que lo hacen posible. Frente al entusiasmo tecnófilo, esta perspectiva advierte que ningún método computacional puede sustituir la crítica histórica de la fuente. Frente a la desconfianza humanista, recuerda que las herramientas digitales pueden ampliar el campo de lo visible y abrir preguntas nuevas sobre archivos extensos, discursos colectivos y formas de circulación. La cuestión decisiva consiste en comprender cómo se relacionan experiencia y patrón. El testimonio puede ser leído como relato singular y como parte de una serie; como voz situada y como dato construido; como memoria individual y como fragmento de una ecología digital más amplia. La tarea del historiador consiste, precisamente, en hacer explícitos esos tránsitos.

De este modo, una lectura híbrida debe entenderse como una posición epistemológica. Su punto de partida es que toda fuente histórica es mediada, pero que no todas las mediaciones tienen los mismos efectos. Las mediaciones narrativas de la memoria no son equivalentes a las mediaciones algorítmicas de una plataforma; la transcripción de una entrevista no produce el mismo tipo de transformación que su lematización o su incorporación a un modelo de tópicos. Por ello, el análisis debe distinguir niveles y no simplemente acumular técnicas. En última instancia, el *distant reading* obliga a explicitar con mayor rigor las condiciones bajo las cuales una voz se convierte en fuente, una fuente en corpus, un corpus en dato y un dato en evidencia. Allí radica la productividad de esta tensión para las humanidades digitales y para la disciplina histórica: en la construcción de una crítica más compleja de la evidencia en la era digital.

Conclusiones

El análisis desarrollado permite sostener que la relación entre historia oral y metodologías computacionales no debe ser entendida como una oposición simple entre interpretación humanística y técnica digital, ni como una disputa entre cercanía y distancia. La cuestión afecta el modo en que la disciplina histórica define sus fuentes, construye evidencia y justifica sus procedimientos de validación. La historia oral mostró, desde sus formulaciones más influyentes, que el testimonio no podía ser reducido a un depósito de información factual, pues su valor reside también en la memoria, la subjetividad, la narración, el silencio y la experiencia situada. Las humanidades digitales, por su parte, han introducido formas de análisis capaces de operar sobre grandes corpus, identificar patrones de escala y abrir preguntas que difícilmente podrían formularse desde la lectura individual. El problema central consiste en comprender qué ocurre cuando una forma de evidencia construida históricamente desde la singularidad de la voz es reinscrita en procedimientos de abstracción, agregación y modelización.

En este sentido, el artículo ha sostenido que el *distant reading* puede contribuir al estudio del testimonio histórico, aunque transforma de manera significativa su estatuto epistemológico. Al convertir el testimonio en parte de un corpus procesable, lo desplaza desde la escena de enunciación hacia un régimen de legibilidad basado en datos, recurrencias, tópicos, redes o patrones. Esta operación es productiva en la medida en que permite detectar regularidades, contrastar conjuntos amplios de fuentes y visibilizar estructuras que permanecen opacas a la lectura cercana. Sin embargo, también puede empobrecer el análisis si se olvida que el testimonio es contenido verbal, acontecimiento narrativo, relación social, inscripción técnica y mediación histórica. Allí se encuentra el núcleo de la tensión examinada: el paso del testimonio al dato debe entenderse como una traducción que modifica parcialmente aquello que vuelve historiográficamente valiosa a la fuente testimonial.

La propuesta de una lectura híbrida o estratificada responde a esta dificultad mediante una articulación reflexiva entre escalas. Las herramientas computacionales pueden funcionar como instancias exploratorias, capaces de orientar la mirada hacia patrones, anomalías, recurrencias o zonas de densidad dentro de corpus extensos. No obstante, esos resultados requieren ser reinsertados en una crítica historiográfica de la fuente, atenta a las condiciones de producción, transcripción, archivo, circulación y validación del testimonio. En este sentido, el concepto de testimonio críticamente mediado

permite comprender que el valor histórico de una fuente reside en la reconstrucción de la cadena de mediaciones que la hace posible como evidencia.

Desde esta perspectiva, la discusión ofrece elementos relevantes para repensar algunas prácticas historiográficas vinculadas con archivos digitalizados, plataformas audiovisuales, repositorios testimoniales, sistemas de búsqueda, modelos de lenguaje e infraestructuras algorítmicas. La digitalización aumenta la disponibilidad de fuentes y modifica las operaciones mediante las cuales esas fuentes son seleccionadas, clasificadas, jerarquizadas y convertidas en objeto de análisis. Por ello, una historiografía atenta al presente debe interrogar críticamente los supuestos de sus herramientas digitales. El desafío consiste en comprender de qué manera esas metodologías reconfiguran las preguntas, los objetos y las formas de prueba de la disciplina.

Finalmente, este trabajo propone situar el debate sobre *distant reading* y testimonio en el horizonte de una crítica de la evidencia histórica en contextos de dataficación. Si el testimonio siempre fue una fuente mediada, la era digital intensifica y vuelve más opacas esas mediaciones. La tarea del historiador consiste en hacer visibles los tránsitos que conectan experiencia, relato, archivo, corpus, modelo e interpretación. En esa reconstrucción se juega la posibilidad de unas humanidades digitales críticas: capaces de aprovechar la potencia analítica de la escala, pero sin renunciar a la densidad ética, narrativa y epistemológica del testimonio. Así, la tensión entre historia oral y *distant reading* abre una agenda de investigación necesaria para pensar cómo se produce, circula y valida el conocimiento histórico en la cultura digital contemporánea.

Obras citadas

- Abrams, L. (2016). *Oral history theory* (2nd ed.). Routledge.
- Alvesson, M., y Sköldbberg, K. (2018). *Reflexive methodology: New vistas for qualitative research* (3rd ed.). SAGE.
- Berry, D. M. (Ed.). (2012). *Understanding digital humanities*. Palgrave Macmillan.
- Berry, D. M., y Fagerjord, A. (2017). *Digital humanities: Knowledge and critique in a digital age*. Polity Press.
- Blevins, C. (2010). Topic modeling Martha Ballard's diary. *Cameron Blevins*. <http://www.cameronblevins.org/posts/topic-modeling-martha-ballards-diary/>
- Booth, W. C., Colomb, G. G., y Williams, J. M. (2008). *The craft of research* (3rd ed.). University of Chicago Press.
- boyd, d., y Crawford, K. (2012). Critical questions for big data: Provocations for a cultural, technological, and scholarly phenomenon. *Information, Communication & Society*, 15(5), 662-679. <https://doi.org/10.1080/1369118X.2012.678878>
- Boyd, D. A., y Larson, M. A. (Eds.). (2014). *Oral history and digital humanities: Voice, access, and engagement*. Palgrave Macmillan. <https://doi.org/10.1057/9781137322029>
- Bradley, K., y Puri, A. (2016). Creating an oral history archive: Digital opportunities and ethical issues. *Australian Historical Studies*, 47(1), 75-91. <https://doi.org/10.1080/1031461X.2015.1122072>
- Burdick, A., Drucker, J., Lunenfeld, P., Presner, T., y Schnapp, J. (2012). *Digital Humanities*. MIT Press.
- Castells, M. (2009). *Communication power*. Oxford University Press.
- Couldry, N., y Mejias, U. A. (2019). *The costs of connection: How data is colonizing human life and appropriating it for capitalism*. Stanford University Press.
- D'Ignazio, C., y Klein, L. F. (2020). *Data feminism*. MIT Press.
- D'Ignazio, C., y Klein, L. F. (2024). Data feminism for AI. En *FACCT '24: Proceedings of the 2024 ACM Conference on Fairness, Accountability, and Transparency* (pp. 100-112). <https://doi.org/10.1145/3630106.365854>
- Drucker, J. (2011). Humanities approaches to graphical display. *Digital Humanities Quarterly*, 5(1).
- Drucker, J. (2021). *The digital humanities coursebook: An introduction to digital methods for research and scholarship*. Routledge.
- Frisch, M. H. (1990). *A shared authority: Essays on the craft and meaning of oral and public history*. State University of New York Press.

- Gillespie, T. (2014). The relevance of algorithms. En T. Gillespie, P. J. Boczkowski, y K. A. Foot (Eds.), *Media technologies: Essays on communication, materiality, and society* (pp. 167-194). MIT Press.
- Graham, S., Milligan, I., y Weingart, S. (2016). *Exploring big historical data: The historian's microscope*. Imperial College Press.
- Grele, R. J. (1991). *Envelopes of sound: The art of oral history* (2nd ed.). Praeger.
- Jockers, M. L. (2013). *Macroanalysis: Digital methods and literary history*. University of Illinois Press.
- Keydar, R., Pinchevski, A., Ifergan, M., y Abend, O. (2026). The testimony of the multitude: Toward a computational model of listening to Holocaust testimony. *Holocaust and Genocide Studies*. <https://doi.org/10.1093/hgs/dcag006>
- Kitchin, R. (2014). *The data revolution: Big data, open data, data infrastructures and their consequences*. SAGE.
- Latour, B. (2005). *Reassembling the social: An introduction to actor-network-theory*. Oxford University Press.
- Manovich, L. (2020). *Cultural analytics*. MIT Press.
- Milligan, I. (2019). *History in the age of abundance? How the web is transforming historical research*. McGill-Queen's University Press.
- Moretti, F. (2013). *Distant reading*. Verso.
- Passerini, L. (1987). *Fascism in popular memory: The cultural experience of the Turin working class*. Cambridge University Press.
- Perks, R., y Thomson, A. (Eds.). (2016). *The oral history reader* (3rd ed.). Routledge.
- Pessanha, F., y Salah, A. A. (2021). A computational look at oral history archives. *ACM Journal on Computing and Cultural Heritage*, 15(1), Article 10. <https://doi.org/10.1145/3477605>
- Piper, A. (2018). *Enumerations: Data and literary study*. University of Chicago Press.
- Plantin, J.-C., Lagoze, C., Edwards, P. N., y Sandvig, C. (2018). Infrastructure studies meet platform studies in the age of Google and Facebook. *New Media & Society*, 20(1), 293-310. <https://doi.org/10.1177/1461444816661553>
- Portelli, A. (1991). *The death of Luigi Trastulli and other stories: Form and meaning in oral history*. State University of New York Press.
- Ramsay, S. (2011). *Reading machines: Toward an algorithmic criticism*. University of Illinois Press.
- Ricoeur, P. (2004). *Memory, history, forgetting*. University of Chicago Press.
- Thompson, P., y Bornat, J. (2017). *The voice of the past: Oral history* (4th ed.). Oxford University Press.
- Thomson, A. (1999). Making the most of memories: The empirical and subjective value of oral history. *Transactions of the Royal Historical Society*, 9, 291-301. <https://doi.org/10.2307/3679406>

- Underwood, T. (2019). *Distant horizons: Digital evidence and literary change*. University of Chicago Press.
- USC Shoah Foundation. (s. f.). *Visual History Archive*. Recuperado el 4 de mayo de 2026, de <https://vha.usc.edu/home>
- van Dijck, J. (2014). Datafication, dataism and dataveillance: Big data between scientific paradigm and ideology. *Surveillance & Society*, 12(2), 197-208. <https://doi.org/10.24908/ss.v12i2.4776>
- Yang, T.-I., Torget, A. J., y Mihalcea, R. (2011). Topic modeling on historical newspapers. In *Proceedings of the 5th ACL-HLT Workshop on Language Technology for Cultural Heritage, Social Sciences, and Humanities* (pp. 96-104).