

Paakat: Revista de Tecnología y Sociedad
e-ISSN: 2007-3607
Universidad de Guadalajara
Sistema de Universidad Virtual
México
paakat@udgvirtual.udg.mx

Año 10, número 19, septiembre 2020-febrero 2021

Repertorio de técnicas digitales para la investigación con contenidos generados en redes sociodigitales

Catalog of digital techniques for research with content generated by social media users

Gabriela Elisa Sued*

<https://orcid.org/0000-0002-4516-678X>

Instituto Tecnológico de Monterrey, México

[Recibido 30/01/2020. Aceptado para su publicación 10/03/2020]
DOI <http://dx.doi.org/10.32870/Pk.a10n19.498>

Resumen

Este trabajo efectúa una clasificación y una descripción sistemática de las técnicas digitales aplicadas al estudio de datos producidos por usuarios en redes sociodigitales. Se identifican los actores institucionales que las producen y promueven, se evalúan las herramientas disponibles y su ámbito de aplicación, y se presentan ejemplos de estudios que las utilizan. Se clasifican las técnicas en función del lugar que ocupan en una secuencia de investigación formada por cuatro etapas: recolección, limpieza, procesamiento y visualización de datos. Asimismo, se enfoca tanto en técnicas de procesamiento de textos como de imágenes; en el primer caso se aborda el análisis cuantitativo de contenidos, el análisis de redes y el análisis de sentimientos, mientras que en el segundo caso se centra en la analítica visual de fotografías y video en línea. También se incluye el análisis cuantitativo de reacciones, como likes y compartidos. En las conclusiones se evalúan de manera crítica los alcances de este estudio, dentro de los que se encuentran: su eficacia para la apreciación de conjuntos extensos de datos producidos en el contexto dinámico y fluido de las redes sociodigitales, y la posibilidad de identificar patrones y recurrencias dentro de estos.

Palabras clave

Medios sociales; recopilación de datos; procesamiento de datos; análisis automático de textos; visualización de datos.

Abstract

This work makes a systematic classification and description of digital techniques applied to the study of data produced by users in social media. The institutional actors that produce and promote them are identified, the available tools and their scope are evaluated, and examples of studies that use them are presented. It Classifies the techniques according to the place they occupy in a research sequence formed by four stages: data collection, cleaning, processing and visualization. It focuses on both textual and image processing techniques. In the first case it addresses textual analytics, network analysis and sentiment analysis, while in the second case it focuses on the visual analysis of photographs and online video. The quantitative analysis of reactions, such as likes and shares, is also addressed. In the conclusions a critical evaluation of its scopes is carried out, among which are its effectiveness for the appreciation of extensive data sets produced in the dynamic and fluid context of social media, and the possibility of identifying patterns and recurrences in them.

Keywords

Electronic media; data collection; data processing; automatic text analysis; data visualization.

Introducción

En la última década ha surgido un fuerte interés por el uso de los contenidos generados por los usuarios de las redes sociodigitales para la investigación social, lo que pone el acento en su doble función: por un lado, constituyen una fuente primaria que permite estudiar las tendencias que la opinión pública expresa, en general, en las diferentes plataformas sociales, además de que permiten comprender la compleja lógica de las plataformas sociales como nuevos medios de comunicación (Rogers, 2009).

En este contexto emergen varios programas de investigación, dentro de los que se destacan la iniciativa de métodos digitales (Rogers, 2013; 2019) y la iniciativa de estudios del software (Manovich, 2016). Ambas se enfocan en el estudio de objetos digitales con métodos digitales, por lo que se ubican en la intersección entre los estudios de medios y las ciencias informáticas.

En relación con análisis anteriores sobre las prácticas y entornos digitales, como la etnografía digital (Hine, 2004; 2015), estas metodologías mueven el foco de estudio de las prácticas de los usuarios a los objetos que estos producen (Marres, 2017). Además de este desplazamiento, los métodos digitales suponen la incorporación de nuevos conceptos, habilidades y técnicas de investigación. Para Rogers (2009), el estudio de objetos digitales generados en la web debe realizarse con métodos y técnicas digitales, relacionado con la lógica del entorno digital; por lo tanto, un cambio de métodos necesita de otro en las técnicas.

La implementación de técnicas digitales para la investigación social y el desplazamiento desde las prácticas hacia los objetos no representan solamente un cambio de procedimientos, sino que significa, por un lado, la posibilidad de entender la lógica de las plataformas sociales y la manera en que estas modelan la expresión pública y la cultura actual (Nieborg y Poell, 2018) y, por otro, la oportunidad de comprender las transformaciones que las

innovaciones técnicas producen en el interior de las ciencias sociales al asumir los retos de estudio y comprensión de nuevos objetos digitales. Para las disciplinas sociales no académicas, como la mercadotecnia, la aplicación de técnicas digitales al análisis de datos no es nueva. Las técnicas analíticas modelan gustos, decisiones y consumos en línea, por lo que es necesario para la investigación académica adoptar nuevos insumos que permitan entender estas nuevas tendencias socioculturales.

Este trabajo efectúa una clasificación y descripción sistemática de las técnicas digitales aplicadas al estudio de datos producidos por los usuarios en las redes sociodigitales, y evalúa las herramientas disponibles y su ámbito de aplicación. Consideraremos técnicas digitales a los modos de hacer que incluyen herramientas digitales o software para el estudio de objetos digitales, sean estos originados en la web o inicialmente analógicos y luego digitalizados (Rogers, 2015). Clasificamos las técnicas por su función en el proceso de investigación: para recolectar, procesar y analizar los datos.

Abrir la caja negra de las técnicas digitales

Con el propósito de sistematizar un panorama heterogéneo y dinámico, es importante distinguir entre técnicas y herramientas. Las técnicas de investigación científica son procedimientos validados por la práctica, orientados generalmente a obtener y transformar información útil para la solución de problemas de conocimiento en las disciplinas científicas (Rojas, 2011); estas operan en un nivel conceptual, mientras que las herramientas trabajan en un nivel concreto.

Respecto a las técnicas digitales, el software es el instrumento o herramienta de la técnica. Las herramientas son programas de software, algoritmos, desarrollados por una persona o un conjunto de estas; se encuentran disponibles de forma pública o privada, dependientes de su contexto de uso o aplicación. Además, las técnicas pueden ser más estables que las herramientas. Una herramienta puede desaparecer, pero otras pueden sustituirla. Generalmente, existe más de una herramienta para una técnica. La selección de la más adecuada deberá responder a un conjunto de parámetros entre los que se encuentran el ámbito de su desarrollo, el tipo de licenciamiento, la complejidad de su interfaz, entre otros.

Al basarse en la escritura de algoritmos complejos que escapan al conocimiento de los científicos sociales, las herramientas digitales son cajas negras: dispositivos que procesan información de entrada y arrojan información de salida, cuyo funcionamiento desconocemos (Latour, 1992). Si como investigadores sociales optamos por emplearlas, será necesario efectuar una delegación de prácticas científicas en terceros, sean estos algoritmos, interfaces

web, software o programadores humanos; por esta razón, Marres y Gerlitz (2016) los denominan métodos de interfaz.

Actualmente, una cantidad de diferentes actores sociales intervienen en la producción de herramientas digitales para el análisis de datos; no todos ellos, aunque sí un número significativo, pertenecen a las comunidades académicas. Existen herramientas provenientes del marketing, la ciencia de datos o, incluso, desarrolladas por las plataformas sociales que pueden usarse para la investigación social. En un sentido inverso, hay emprendimientos de origen académico que se han transformado en empresariales, como el paquete de visualización de datos Tableau, desprendimiento comercial de una investigación llevada a cabo en la Universidad de Stanford (Solomon, 2016).

Estas condiciones de producción distinguen a la investigación digital de los métodos habitualmente usados en las ciencias sociales, en los que el investigador es quien diseña sus instrumentos de recolección y análisis de datos. Una consecuencia de esta apertura es que se debe delegar parte de la investigación a un artefacto construido por terceros. Rogers (2009) denomina reorientación o repurposing al uso de herramientas con fines académicos que no han sido especialmente diseñadas para la investigación. Deberemos considerar que cuanto menos conocimiento tecnológico demande la herramienta, mayor será el conocimiento que estemos delegando sobre ella.

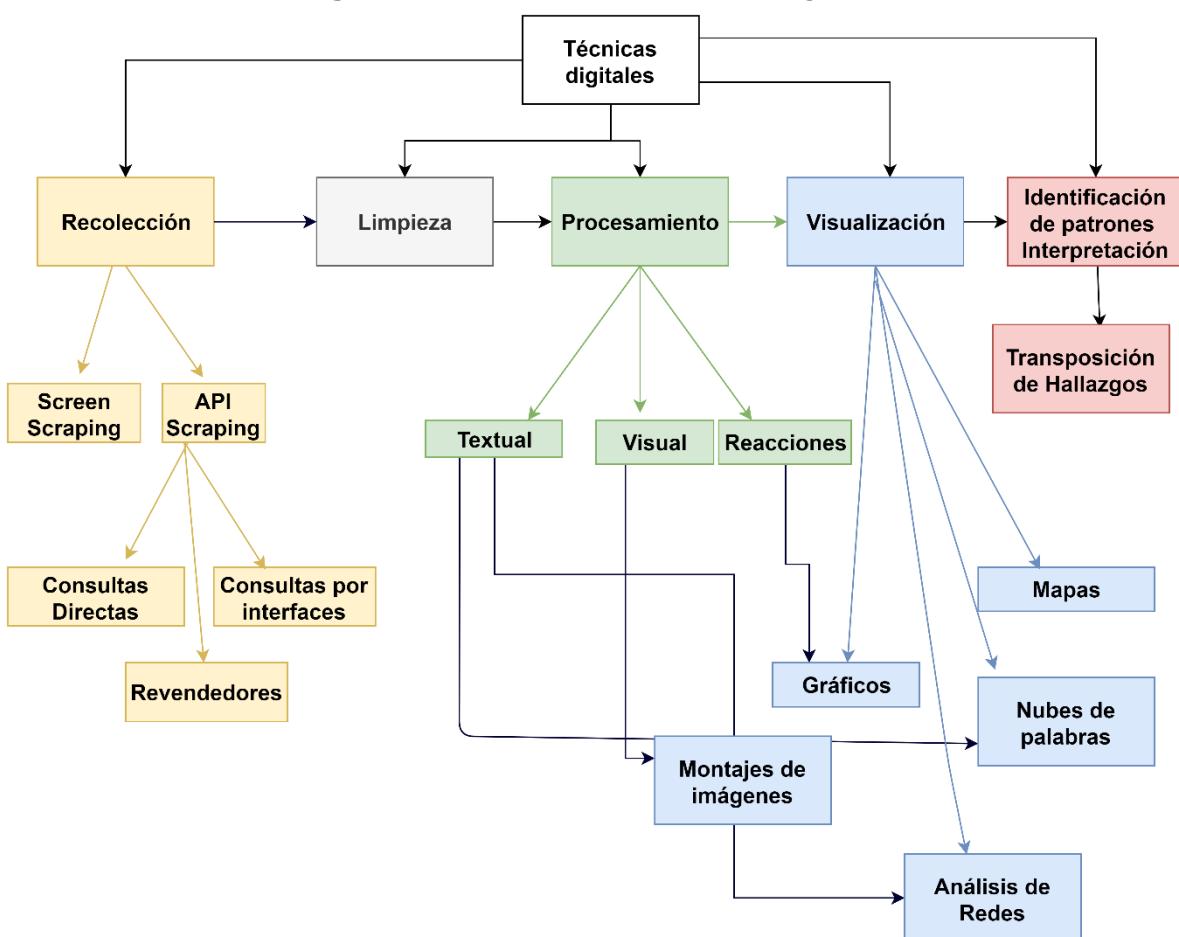
Con el viraje de los estudios sociales y humanísticos hacia los objetos digitales y digitalizados, diferentes centros de investigación universitarios (varios de los cuales renovaron sus denominaciones como Media Lab o laboratorios de medios) desarrollaron instrumentos de recolección, análisis y visualización de datos. Estos fueron puestos, en mayor o menor medida, a disposición de las comunidades académicas. Varias de estas herramientas recolectan, procesan y visualizan datos, y entregan un producto terminado que combina los tres tipos de técnicas; existen otras que entregan datos y metadatos en formatos planos. Los investigadores deberán aplicar después diferentes técnicas e instrumentos para organizar los datos, asignarles sentido y contestar sus preguntas de investigación.

Las técnicas digitales para la investigación social han sido trabajadas exhaustivamente en la literatura en inglés. Se destacan las sistematizaciones sobre minería textual de Moreno y Redondo (2016), las de analítica de datos de Gandomi y Haider (2015) y el manual de métodos de investigación on-line editado por Fielding, Lee y Blank (2016); sin embargo, la literatura en español es escasa. Este trabajo contribuye en esta dirección, al analizar las técnicas en relación con su función en el proceso de investigación.

Una tipología de técnicas digitales

Ante el escenario heterogéneo y cambiante en el que se presentan técnicas y herramientas, no se trata de renunciar a la categorización, sino de trabajar sobre múltiples modos de clasificar instrumentalmente las técnicas digitales. En este caso, nos centraremos en la función de las técnicas en el proceso de investigación, pero existen otras posibilidades: según la materialidad del objeto digital, las características del software que se empleará y la pertenencia del objeto a una plataforma. La figura 1 organiza esta información, según el lugar que ocupan en una secuencia de investigación formada por cuatro etapas secuenciales que abarcan la recolección, la limpieza, el procesamiento y la visualización de datos.

Figura 1. Clasificación de técnicas digitales



Fuente: elaboración propia.

El scraping como técnica de recolección de datos

El *scraping* es una técnica específica del medio digital que permite la recolección automática de datos en línea. Es una de las técnicas más representativas de los métodos digitales actuales, ya que posibilita la investigación basada en datos en

el medio digital (Marres y Weltevrede, 2013). Mediante esta técnica se recuperan datos crudos que, si bien están estructurados en el sitio web al que pertenecían de manera original, necesitarán de una nueva organización para poder ser leídos e interpretados y convertirse en información relevante para la investigación (Carvajal, 2013). Mediante la ejecución de un *software* desarrollado para este fin, y a partir del propósito del estudio, podremos recuperar diferentes tipos de datos: desde los que se encuentren en un informe en PDF para verterlos en una hoja de cálculo, un conjunto de imágenes publicadas en una red social, los tuits que se hayan escrito sobre algún acontecimiento o los metadatos de un conjunto de videos almacenados en YouTube.

Existen tres clases de *scraping*: el *screen and interface scraping*, el *crawler scraping* y el *API scraping* (Elmer, 2015). El primero es el más antiguo; se trata de un método automático usado para extraer datos que han sido diseñados para ser vistos primariamente por humanos. El programa lee la pantalla y simula ser un humano, y colecciona los datos interesantes en listas que pueden ser procesadas de manera automática (Fulton, 2014). Se basa en la extracción de datos del código HTML visualizado en las interfaces de usuario, lo que significa que los datos son formateados y personalizados para usuarios particulares.

Un ejemplo de *screen scraping* es el de la extensión *Bulk Media Downloader*, complemento montado sobre el navegador Firefox. Después de ser instalado, el usuario puede bajar todos los contenidos que ve en su pantalla y almacenarlos en su computadora. Pero como las búsquedas en las plataformas sociales y los sitios web son personalizadas, lo que el usuario recolecta es información sesgada por su experiencia.

El *crawler scraping* extrae la estructura de un sitio web, además de los datos que las bases de datos asociadas al sitio pueden contener. El tercer tipo de *scraping*, habitualmente el más utilizado, es el basado en consultas a las interfaces de programación de aplicaciones, conocidas como API (Application Programming Interfaces). En este caso, los datos se encuentran estructurados en una base de datos en función de los intereses de las plataformas que los almacenan. Por este motivo, deberemos recurrir a las API; estas consisten en un conjunto de algoritmos, funciones y procedimientos que ofrece una plataforma para ser utilizada por otro *software* que solicitará información. Constituyen una herramienta central en el proceso, ya que representan la capacidad de comunicación entre la base de datos de la plataforma que brinda la información y el programa que la recolecta.

Si la plataforma no habilita su API para recolectar información, no podremos obtener sus datos, al menos a través de esta. En general, las plataformas habilitan sus API para que los desarrolladores independientes creen nuevas aplicaciones para los datos, pero no para extraer información a fin de ser analizada; por ejemplo, Facebook habilita sus API para la creación de

videojuegos, encuestas y aplicaciones de publicidad y *marketing*, pero no para la exportación de datos.

Limpieza de la base de datos

La limpieza de las bases de datos es un paso intermedio entre la recolección y el procesamiento de datos. Implica tiempo y esfuerzo, pero es la base para obtener un corpus que pueda manipularse y visualizarse correctamente. Una base de datos limpia posee una estructura de datos clara, confiable y bien organizada, de manera que podamos encontrarle sentido mediante una rápida exploración (Wickham, 2014).

La estructura de datos se corresponde con una organización semántica. En una base de datos limpia cada columna es una variable, cada fila es una observación que contiene uno o más valores, usualmente formados por números o texto, y cada conjunto de observaciones es una tabla. El orden de una base de datos no forma parte de su estructura, pero hace más fácil explorar los datos antes de procesarlos. El orden estándar ubica en primer lugar a las variables fijas y en segundo a las variables numéricas; las variables relacionadas deben ubicarse contiguamente y las filas pueden ordenarse por la primera variable.

Existen problemas comunes en la limpieza de las bases de datos que pueden ser resueltos con los programas usuales de almacenamiento de datos o con software especializado; *Open Refine* (gratuito de Google), es el más usado en esa categoría.

Técnicas de procesamiento: analítica de datos

Una vez que hemos capturado los datos deberemos procesarlos para darles sentido. La ciencia de datos denomina *analítica* a las técnicas usadas para extraer sentido de la información capturada (Gandomi y Haider, 2015). Otros autores del mismo campo prefieren la denominación de minería de datos (Han, Kamber y Pei, 2011). Ambas nociones incluyen el análisis de cualquier tipo de bases de datos, por ejemplo: gubernamentales, comerciales, empresariales y académicas. Aunque no han sido originadas en el campo, estas denominaciones han sido adoptadas por la investigación social. Según el tipo de dato que tratan, se clasifican en analítica textual, de audio, de imagen, de video, de redes sociales y analítica predictiva.

a) Técnicas de analítica textual

La denominación de analítica textual o minería textual proviene de la ciencia de la computación y es una derivación del concepto de minería de datos. Desde la

ciencia de datos, Gandomi y Haider (2015) y Moreno y Redondo (2016) la definen como el descubrimiento de nueva información extraída automáticamente de diferentes fuentes escritas. Para realizar este proceso se emplean programas y algoritmos que efectúan diversas operaciones sobre los textos, a las que los autores denominan *técnicas de analítica textual*. Este procedimiento es descrito como la transformación de información desestructurada y cualitativa a información estructurada y cuantitativa. El desarrollo de procesamiento de lenguaje natural ha avanzado mucho en los últimos años, y el contenido producido por los usuarios en las redes es un área de especial interés para su estudio.

La literatura especializada menciona entre las principales técnicas la extracción de información, el resumen de textos, las respuestas a preguntas formuladas en lenguaje natural (Gandomi y Haider, 2015), el análisis de contenido (Popping, 2016), el análisis de sentimientos (Paltoglou y Thelwall, 2012) y la minería visual de datos (Moreno y Redondo, 2016). De estas, las tres últimas son las empleadas en el análisis de objetos producidos en las plataformas sociales.

Popping (2016) define el análisis de contenido como la reducción de un flujo de texto desestructurado y cualitativo a un conjunto de símbolos manipulables y cuantitativos para realizar una inferencia válida del texto hacia su contexto. Estos símbolos representan, en general, valores de presencia, intensidad y frecuencia. Una de las formas básicas del análisis de contenido es la que analiza con qué frecuencia una palabra aparece dentro de un texto (Robichaud y Blevins, 2011), y es mayormente utilizado en los estudios basados en datos. El recuento de palabras es una técnica simple, pero parece brindar resultados relevantes en el caso del análisis de los objetos textuales de la web social. Esto puede deberse a que los posteos en las redes sociales suelen ser breves, simples y carentes de complejidades discursivas (Paltoglou y Thelwall, 2012). Underwood (2013) considera que la técnica es simple y efectiva; si bien se centra en la unidad más simple del lenguaje, la palabra, esta es por sí misma un elemento lingüístico de suficiente complejidad.

Moreno y Redondo (2016) destacan la capacidad de la minería visual para colocar grandes cantidades de datos en una jerarquía visual, un mapa o un gráfico que puede ser inspeccionado de modo interactivo. Esto puede ser útil para explorar un gran número de documentos y relaciones entre tópicos y temas. Un ejemplo de esta categoría son las nubes de palabras, los gráficos de árbol de temas o las redes de palabras. En este sentido, la categoría no es completamente autónoma, ya que es la expresión visual de otro tipo de minería textual: las nubes de palabras manifiestan visualmente el análisis de frecuencia de palabras, así como los árboles temáticos lo son del análisis de temas relacionados, y los gráficos de red implican la expresión entre al menos dos entidades que entablan entre sí alguna relación.

Voyant-Tools¹ es una de las herramientas de minería visual más famosas. Fue desarrollada desde el año 2003 por Stefan Sinclair y Geoffrey Rockwell (de la Universidad de Alberta y McGill, respectivamente); es gratuita y de código libre. Está orientada a los análisis académicos del ámbito de las humanidades digitales (Voyant Tools, 2018). Con los años, ha incorporado nuevos módulos cuyas funcionalidades complejizan la inicial de análisis de frecuencia de palabras. El entorno completo actual incluye análisis de tópicos, gráficos de red y burbujas, análisis de frases, entre otras aplicaciones.

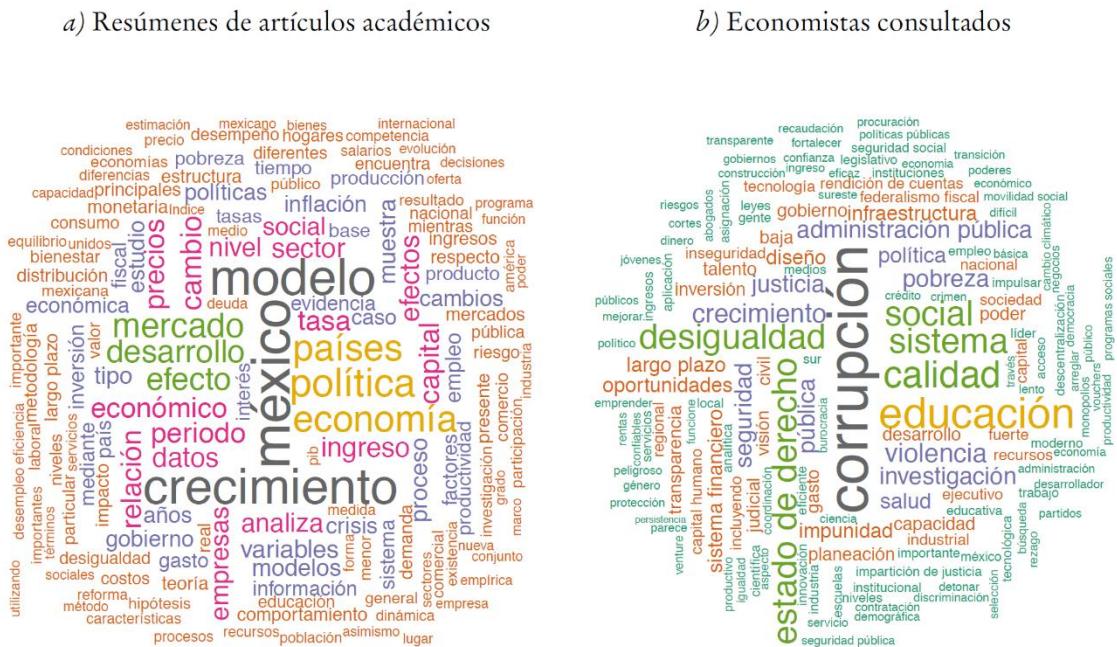
La herramienta comienza a ser empleada en diversas áreas de la investigación social, no solamente en la que corresponde al análisis de medios sociales. En el área de la economía, por ejemplo, Campos-Vázquez y López-Araiza (2018) la utilizan para comparar un conjunto de palabras que son usadas frecuentemente, las cuales son extraídas de artículos publicados en las principales revistas de economía en México, que contienen nociones utilizadas por un grupo de economistas entrevistados acerca de las prioridades de la investigación para la economía mexicana (ver figura 2).

En el área de la administración del conocimiento, Cortés (2018) la emplea para examinar comparativamente los enunciados de misión de numerosas universidades alrededor del mundo. Desde las humanidades digitales, Lacalle y Vilar (2019) la aplican a un corpus de revistas académicas sobre literatura a fin de identificar las temáticas más estudiadas. En el análisis de medios sociales, Sued (2018) la usa para identificar redes semánticas entre las palabras más frecuentes que se utilizan en las descripciones de un conjunto de fotografías publicadas en Instagram.

Paltoglou y Thelwall (2012) denominan *análisis de sentimientos* al proceso automático que determina si un segmento de texto contiene expresiones emocionales o subjetivas, si hay polaridad positiva o negativa. Creada en la década de 1990 para procesar las valoraciones que clientes y críticos realizaban a productos de venta en línea, en los últimos años ganó mayor importancia, a partir del interés en el análisis de los contenidos publicados en blogs y redes sociales.

En el contexto de las plataformas sociales, el análisis de sentimientos se orienta a identificar si las interacciones contienen expresiones de estados de ánimo y si expresan sentimientos positivos (como entusiasmo o alegría) o negativos (como desacuerdo o ironía). Se basa en clasificadores de léxico que estiman el nivel de valencia emocional en orden de realizar una predicción. Los clasificadores se almacenan en diccionarios que pueden ser producidos automáticamente o creados por los usuarios. El *software* lee textos y usa un algoritmo para producir una medida estimada del contenido de sentimientos.

Figura 2. Comparación entre palabras más frecuentes en artículos académicos y economistas consultados



Fuente: Campos-Vázquez y López-Araiza (2018). Imagen reproducida con autorización del autor y la editora.

El principal problema de su aplicación a los contenidos de las redes sociales es que las interacciones, principalmente en los temas políticos y controversiales, suelen contener ironía y sarcasmo. Estas formas son difíciles de identificar por los algoritmos (Thelwall, 2017), por lo que su uso no controlado podría llevar a resultados erróneos. Otro problema es el reconocimiento del idioma, ya que la mayoría de ellas se desarrollan para el inglés, lo que presenta un problema al reconocimiento automático de las estructuras en español y otros idiomas.

b) Análisis visual de redes

Visualizar e interpretar fenómenos como redes es parte de las culturas digitales actuales, estructuradas como espacios de interacción donde la información fluye y se comparte sin orden ni jerarquías aparentes. A la estrategia general de investigación diseñada sobre la base de estructuras de red se le denomina *análisis visual de red* (AVR). Los gráficos de red develan relaciones entre elementos que aparentemente se encuentran dispersos (Venturini, Jacomy y Pereira, 2015).

El AVR se centra en dos rubros esenciales: los nodos y las aristas. Los primeros son las entidades que se relacionan entre sí; las segundas son los elementos que construyen las relaciones. Además, pueden caracterizarse según su peso o importancia en la red. Los nodos que establecen más relaciones con

los demás, por ejemplo, los más retuiteados por otros, pueden verse más grandes, mientras que las aristas, como los tuits más retuiteados pueden verse más gruesas (Goldbeck, 2013).

Venturini, Jacomy y Pereira (2015) identifican tres elementos básicos para la interpretación de gráficos de red: la posición de los nodos, el tamaño y el color. Esta ubicación es asignada por alguno de los algoritmos de espacialización incluidos en la interfaz de usuario. Los más comúnmente usados son los de vector-fuerza, que organizan los nodos en el espacio y minimizan los cruces de aristas, colocando cerca a los nodos que poseen más conexiones entre sí, lejos de los que poseen menos conexiones, y en la periferia del gráfico a los desconectados. Así la mayoría de los gráficos revelarán regiones en las que muchos nodos se ensamblan y otras casi vacías. Estas proximidades y lejanías que construyen zonas de desigual densidad son reveladas de manera automática por los algoritmos de espacialización.

Las zonas de mayor aglomeración constituyen los *clústeres* de la red. Las espacializaciones, opuestas a los clústeres, se llaman hoyos estructurales o *structural holes*; su presencia indica desconexión entre clústeres, mientras que su ausencia puede ser interpretada como signo de una oposición. El tamaño del clúster se define de acuerdo con los nodos que contiene, y su densidad a partir de la cantidad de aristas que vinculan a los nodos. Cuanto más denso es un clúster, habrá mayor cohesión entre sus miembros. Los clústeres son más apretados cuando contienen muchas aristas y más sueltos cuando contienen pocas; en estos últimos, lo interesante no es lo que une a los nodos, sino lo que los separa de los clústeres apretados y añade distancias entre ellos. Además de los clústeres, los nodos y las aristas también pueden tener diferente tamaño; este depende de la cantidad de aristas que generan o reciben; los mayores se consideran las “autoridades” de la red.

Los colores representan operaciones de clasificación de nodos, las comunidades, clústeres o agrupamientos entre nodos que pueden encontrarse en una red, la cual es determinada por el algoritmo de modularidad. Cuanto más se acerca a 1 el coeficiente de modularidad, más estructura de comunidades contiene la red (Blondel *et al.*, 2008). Existen varios algoritmos embebidos en la interfaz de Gephi que miden las relaciones entre nodos y enlaces; aquí hemos planteado los esenciales para producir e interpretar gráficos de red.

En el estudio de Twitter, el AVR se ha usado especialmente para modelizar las conversaciones entre usuarios a través de retuits y menciones; se consideran los primeros como nodos y los segundos como aristas. Al aplicar el algoritmo de centralidad se identifican los usuarios centrales de la red, quienes pueden ser tanto los que más tuitean como los que son más retuiteados.

El AVR hace visible quiénes son los que más conversan en las redes, cuánto lo hacen y con qué otros usuarios lo hacen; lo que no identifica es qué se conversa. Si la discusión deriva en una disputa de sentidos y de influencias sobre las agendas mediáticas, se vuelve necesario saber no solamente quiénes son los que ejercen la influencia y en qué direcciones se esparcen, sino también qué sentidos se construyen.

En la investigación sobre plataformas sociales, varios estudios de caso han usado AVR. En la investigación social tradicional se había aplicado al estudio de los movimientos sociales no virtuales, y después se incorporó en los que suceden en línea. Tremayne (2014) analiza el movimiento *Occupy Wall Street* ocurrido en Nueva York en 2011. En México, Monterde *et al.* (2015) lo aplican al caso mexicano #yosoy132; Reguillo (2017) lo emplea para reflexionar sobre los movimientos sociales 2.0; Pedraza y Cano (2019) lo usan para realizar un mapeo de organizaciones feministas a partir de los datos recolectados en Facebook; con el formato novedoso del laboratorio de medios, el Signa_Lab² del ITESO lo utiliza para estudiar la conversación pública sobre temas de la coyuntura política.

Otra manera de implementar el AVR es aplicándolo a hashtags concurrentes (es cuando dos o más hashtags se encuentran en una misma publicación). Si se entiende que estos son macrotérminos que cristalizan y sintetizan los discursos, el análisis de redes se combina, en este caso, con el análisis textual de contenido, y agrega un componente que permite obtener la variedad de temas que se tratan en un conjunto de publicaciones (Borra & Rieder, 2014).

El uso de hashtags concurrentes es más habitual en Instagram que en Twitter. Sued (2018) lo usa para graficar una red entre hashtags que concurren de un conjunto de 400 fotografías de la ciudad de México publicadas en Instagram, lo que permitió determinar, mediante la concurrencia, diferentes clústeres temáticos para las imágenes (figura 3).

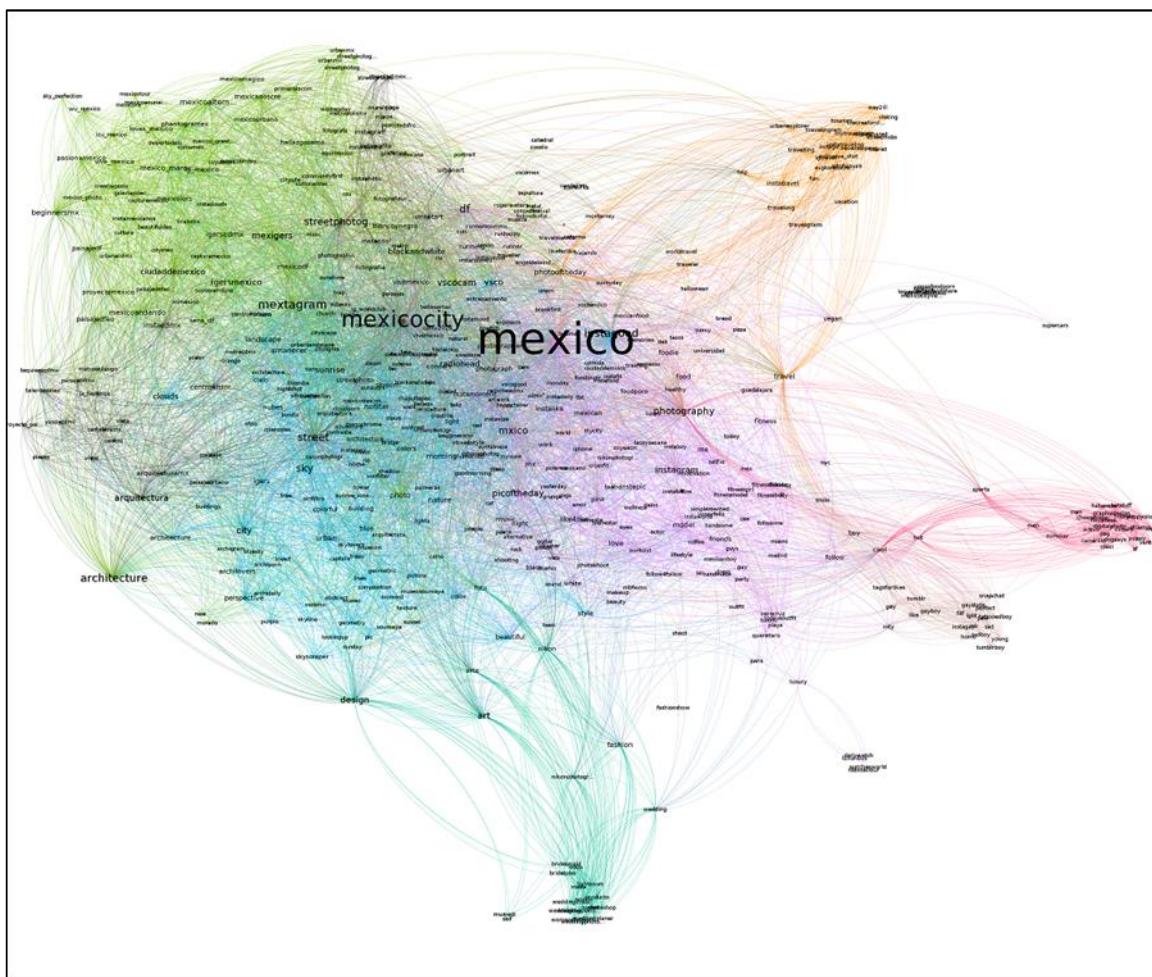
c) Técnicas de analítica visual

A pesar de que las plataformas sociales están repletas de imágenes, la cultura visual no ha sido un objeto central para los estudios digitales, en especial sobre lo que hace al experimentar y difundir metodologías relacionadas con el estudio de temáticas y estilos vinculados a plataformas (Highfield y Leaver, 2016).

El concepto de analítica visual, renombrado por Manovich (2009) como analítica cultural, refiere al estudio de las imágenes creadas y subidas por individuos en plataformas de fotografía compartida como Instagram, Flickr, Pinterest, Tumblr e, incluso Facebook y Twitter (Niederer y Taudin, 2015). La publicación masiva de imágenes que ocurre en estas plataformas representa un

cambio de escala en la cultura visual contemporánea y una resignificación tanto del concepto de fotografía como de sus modos de producción, circulación y consumo (Fontcuberta, 2011). La analítica visual estudia imágenes estáticas y dinámicas, siempre a gran escala. La cantidad de objetos estudiados es importante porque permite identificar patrones y regularidades con mayor facilidad.

Figura 3. Red de cohashtags en #cdmx



Fuente: Sued (2018).

Nota: clústeres por color. Verde claro: ciudades globales. Rosado: moda, estilos de vida, arte y decoración. Verde oscuro: arte urbano, graffiti y turismo local. Celeste: comidas y bebidas. Naranja: turismo, viajes y colectivos de Instagram globales. Turquesa: turismo local y colectivos de Instagram locales.

El programa de procesamiento de imágenes Image-J es el instrumento usado por la fase técnica de la analítica cultural. Image-J es de código abierto y fue desarrollado por el Instituto Nacional de Salud de Estados Unidos, readaptado para su uso en el ámbito de las humanidades digitales y el estudio de medios mediante la extensión Image Plot. El programa actúa en dos fases: en la primera realiza una medición de valores de brillo, saturación y tonalidad

de las imágenes; en la segunda reconstruye las imágenes en un montaje o collage, las transforma en miniaturas y las organiza según la medida o variable que el usuario elija.³

En estos montajes o collages se pueden identificar patrones estéticos de color, brillo y saturación. Aunque existen modos de producir montajes de acuerdo con categorías de contenido previamente determinadas, los alcances de la propuesta se limitan a la identificación de patrones estéticos y, en ocasiones, temporales. Otras herramientas de uso más simple pueden ofrecer prestaciones similares, por ejemplo Image Sorter.⁴

Existe un artículo paradigmático sobre el uso de la analítica cultural en el estudio de medios sociales escrito por Hochman y Manovich (2013), en el que se comparan los valores visuales de millones de fotografías correspondientes a trece ciudades, publicadas en la plataforma Instagram. Un estudio más reciente es el de Pearce *et al.* (2018), quienes emplean en su análisis multiplataformas del contenido generado por usuarios sobre el cambio climático.

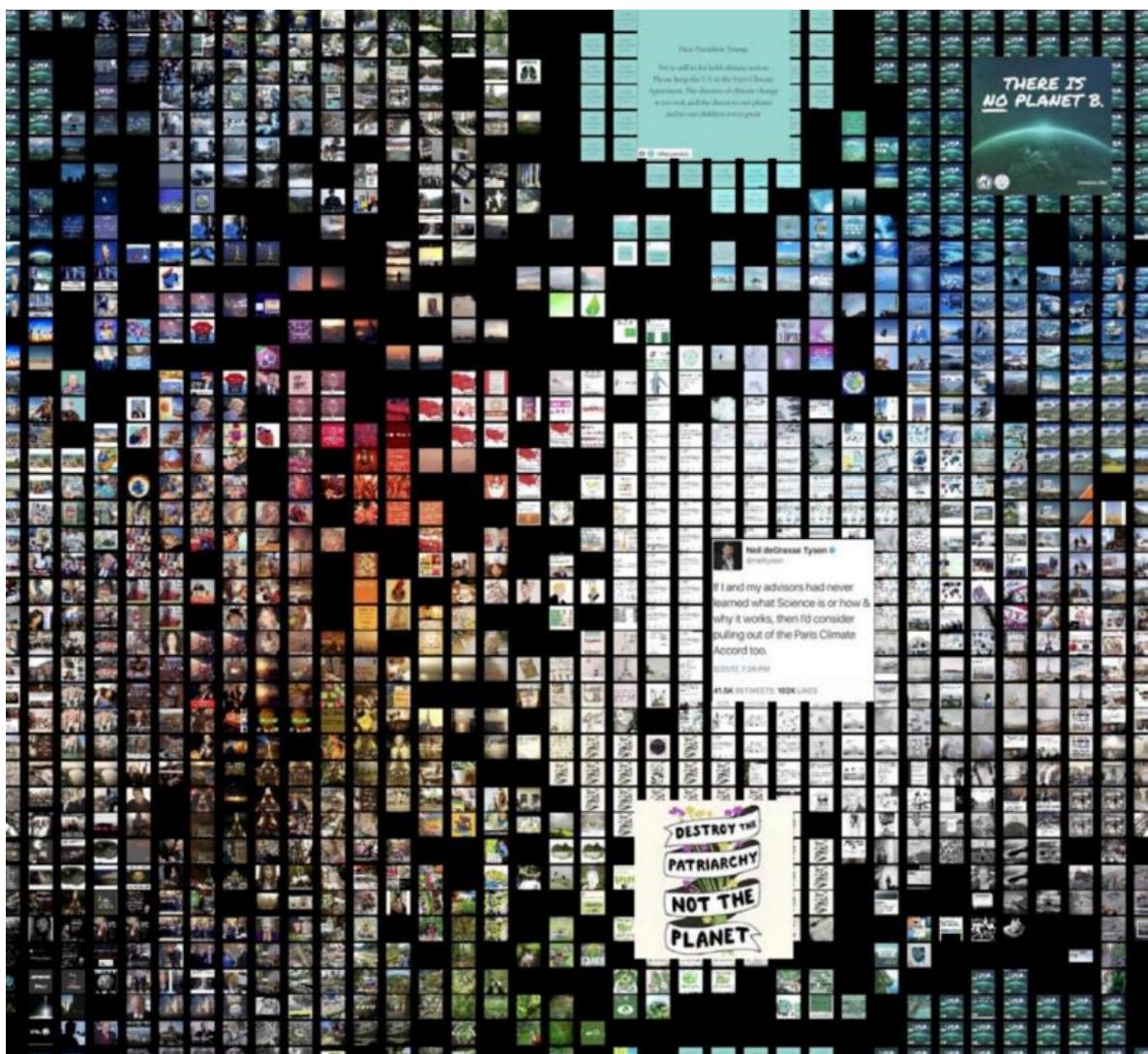
La figura 4 muestra un montaje fotográfico realizado con la herramienta Image Sorter (Berthel, 2006), compuesto por 1 203 imágenes recolectadas del hashtag #parisagreement de Instagram. En el montaje, las imágenes se organizan por color, lo que permite al investigador identificar clústeres homogéneos; por ejemplo, el grupo de imágenes blancas representa a las científicas, mientras que el grupo donde predomina el rojo representa a las de la investigación empírica que alertan sobre los peligros del cambio climático. Las imágenes de mayor tamaño son las más compartidas en varias plataformas.

Además de la analítica visual, Rose (2016) menciona otra técnica importante para el análisis de grandes cantidades de imágenes: el análisis de visual contenido, que posee una base cuantitativa desarrollada originalmente para el análisis de textos verbales. Consiste en la cuantificación de un universo de fotografías que puede ser clasificado en categorías a partir de un procedimiento de etiquetado manual, en el cual el analista es quien determina, de acuerdo con sus preguntas de investigación, las categorías correspondientes, y después se encarga de etiquetarlas y registrar el procedimiento en un archivo que, en una segunda etapa, pueda ser presentado como un análisis claro de frecuencia de categorías.

Los avances realizados en los últimos tres años en el campo de la visión computacional han permitido el desarrollo y la comercialización de programas de clasificación automática de contenidos (Sightengine, 2018). El propósito de esto es que las computadoras puedan entender el mundo al extraer información útil de las señales digitales para llevar a cabo razonamientos complejos (Wang, Komodakis y Paragios, 2013). Los sistemas de visión computacional aplicada al reconocimiento de imágenes se basan en algoritmos que emplean la técnica de

deep learning para el reconocimiento de imágenes. Tres principales empresas tecnológicas como Google, Amazon y Microsoft han desarrollado Vision APIS a las que se puede obtener un acceso gratuito por tiempo y funcionalidades limitadas; sin embargo, el manejo de la interfaz aún es complejo para grandes colecciones de datos (Vitale, 2015).

Figura 4. Montaje fotográfico: 1 203 imágenes publicadas en Instagram, #parisagreement



Fuente: Pearce *et al.* (2018). (Diseño de imagen: Federica Bardelli, Carlo de Gaetano y Michele Mauri, reproducida con permiso de autores y editores.).

El análisis del video *on-line* es una posibilidad interesante para la comprensión de la cultura participativa y sus vínculos con las prácticas sociales y culturales, principalmente porque la API de YouTube se encuentra disponible para la recolección de información, lo que posibilita llevar a cabo una analítica de datos y metadatos. En un trabajo en el que se analiza cuantitativa y cualitativamente la cultura participativa que se construye en esta plataforma,

Burgess y Green (2018) diseñan una metodología mixta para el video *on-line*. Los autores realizan una lectura distante basada en un corpus de 4 320 videos recolectados durante tres meses de 2007.

La metodología no consiste en mirar los videos uno a uno como en una lectura cercana; por el contrario, la propuesta radica en la construcción de una base de datos con los metadatos de los videos. A partir de su análisis, los autores realizan una interpretación de las prácticas de las culturas participativas en YouTube, lo que sitúa su mirada en una tipificación de los productores de los videos, así como en las diferentes maneras en las que los usuarios expresan sus reacciones ante estos, y distinguen entre las acciones de mirar, comentar y compartir un video.

d) Métricas de reacciones

Las métricas de reacciones se definen como las maneras de medir la recepción y la circulación de una publicación en las plataformas sociales. Se basan en los indicadores de recepción de los usuarios, como los *like*, comentarios, compartidos, clics a una dirección web o la cantidad de visualizaciones de un contenido. Estos objetos se presentan, generalmente, con un valor numérico que puede asociarse a una reacción afectiva (como en el caso de Facebook), y reciben, por lo tanto, un tratamiento cuantitativo que responde a la lógica de selección de publicaciones de las plataformas; en esta no sería relevante la pregunta sobre quiénes vieron o comentaron determinado contenido, sino cuántas personas lo vieron, a cuántas les gustó, o cuántas veces se compartió.

La literatura proveniente del *marketing* y los negocios digitales se refiere a las reacciones con el término inglés *engagement*; que, en estas áreas, no es tanto signo de compromiso del usuario como un indicador de éxito de una estrategia de comunicación en línea. En este sentido, el máximo compromiso implica que el usuario se involucre con la publicación de un modo que afecte su vida fuera de línea; por ejemplo, que decida visitar un lugar turístico a partir de las imágenes que vio en una red social; y un nivel mínimo es que solamente haya visto la fotografía.

La investigación en métodos digitales entiende que la métrica de reacciones es valiosa para comprender las interacciones en los medios sociales, pero propone una reorientación que consiste en identificar qué nivel de interés social presentan los usuarios sobre un asunto determinado y cómo se expresa y construye ese interés en términos de conversaciones y conexiones entre usuarios.

Rogers (2018) destaca el aspecto de productividad social de las redes en relación con la construcción de redes de influencia que justifican la necesidad de estudiarlas en términos de *engagement*. En línea con la propuesta general

de los métodos digitales, entiende la publicación de contenidos sociales como un espacio donde se expresan diferentes asuntos sociales. Este aspecto se destaca en el uso de las redes para el activismo social al influenciar agendas o, en términos políticos, al incidir sobre el pensamiento de los ciudadanos.

Gerlitz y Helmond (2013) realizan un análisis de reacciones a partir de los intercambios simbólicos que suceden entre los usuarios y las plataformas que las llevan a formular una economía del *like*; mientras que para los primeros el *like* tiene un valor afectivo, para las segundas presenta un valor comercializable en cuestión de marketing y consumo personalizado. Por lo tanto, las plataformas promueven el uso del *like*. De acuerdo con la lógica de esta economía, las estrategias de publicación de los usuarios se orientan a lograr visibilidad y reacciones, en una manera de operar que se reproduce en las plataformas sociales.

La cantidad de reacciones que recibe un *post* puede recolectarse mediante la técnica de *scraping* de API o, si se trata de un objeto digital como un artículo periodístico puesto a circular en redes sociales, pueden emplearse herramientas de monitoreo social; por ejemplo, Crowdtangle⁵ permite monitorear las reacciones de los enlaces que circulan en Facebook y Twitter.

Técnicas de visualización de datos

La producción de una visualización consiste en codificar información en un conjunto de elementos básicos como el tamaño, la forma, el color y la posición de cada uno de los ítems que la forman (Cairo, 2018). El autor identifica tres elementos básicos en una visualización: un marco, una o más codificaciones visuales y anotaciones. Las codificaciones visuales son los elementos más importantes y también los más difíciles de utilizar, estos pueden ser: el alto y el ancho de los elementos, su posición, los colores y sus variaciones tonales, el área, el grosor de las líneas, entre otros.

A los efectos de las técnicas revisadas en este artículo podemos considerar que las etapas de procesamiento y visualización se combinan en varias de ellas. Esto es lo que sucede con las nubes de palabras que produce la minería de textos, con los montajes fotográficos que resultan de la analítica visual y con los grafos que elabora el AVR. En otros casos existe un *software* que visualiza datos, por lo general cuantitativos (pero no exclusivamente) en gráficos. Hay diferentes tipos: gráficos de barras, de líneas, de pastel, de área, diagramas de Gantt, entre otros. Los mapas que incluyen información extra geográfica y se producen con fines específicos también se consideran gráficos.

Se cuenta con varias herramientas para la elaboración de gráficos. En general, pueden desarrollarse con las herramientas gráficas de las planillas de

cálculo comunes. Además de estas, el paquete comercial de Tableau Software⁶ es utilizado frecuentemente, así como la aplicación RawGraphs,⁷ diseñada por Density Design Lab del Politécnico de Milán; funciona en línea, es de código abierto, gratuita y empleada en el ámbito académico. Combina los tipos de gráfico estándar con estilizaciones basadas en el diseño gráfico de cada una de ellas.

Otro emprendimiento académico es Wrangler,⁸ diseñado por el grupo de visualización de la Universidad de Stanford, utilizado en el área de periodismo de datos; también se encuentra disponible en línea, no es de código abierto y generó un emprendimiento comercial de visualización llamado Trifacta. En cuanto a la producción de gráficos basados en modelos, la extensión para navegadores Draw.io⁹ es gratuita y de uso simple.

Conclusiones. Alcance y limitaciones de las técnicas digitales

Este artículo revisó un repertorio de técnicas digitales para el estudio de contenidos generados por los usuarios en las redes sociodigitales. Consideradas en su producción, diferentes ámbitos, actores académicos y no académicos han desarrollado técnicas y herramientas para el tratamiento de datos producidos en estas redes. Si las técnicas y herramientas provienen del *marketing* digital o el entorno empresarial, pueden ser reorientadas hacia la investigación académica. En el sentido inverso, herramientas digitales que se generaron en el entorno universitario pueden instalarse después en contextos extra académicos; sin embargo, por ser el *software* su materialidad constituyente, es necesario saber que estas herramientas funcionan como cajas negras: por un lado, el investigador, quien no es el creador sino el usuario, delega en ellas una parte del conocimiento sobre su funcionamiento y, por otro, debe limitarse a tratar la información que la herramienta recolecta.

Las técnicas digitales reciben diferentes modos de clasificación: según el lugar que ocupan en el proceso de investigación, en las técnicas de recolección, de limpieza, de procesamiento y de visualización; además, pueden distinguirse según el tipo de objeto a analizar: texto, imagen, video o reacciones. En la etapa de recolección de información, la técnica de *scraping* permite recolectar y estructurar conjuntos de datos desestructurados y formar un corpus de objetos que luego puede tratarse con una o más técnicas. Durante el procesamiento de la información, según la bibliografía analizada y los casos considerados como ejemplos de uso en este trabajo, las técnicas digitales presentadas resultan eficientes para explorar de modo cuantitativo conjuntos extensos de datos. Estas técnicas se vuelven necesarias para visibilizar las variaciones y recurrencias en los conjuntos de datos, permiten identificar patrones y tendencias a los que puede asignarse una interpretación.

Si la función fundamental de una técnica es operar sobre un conjunto de datos para extraer información de ellos, las técnicas presentadas son eficientes; aunque no todas poseen la misma capacidad de acceso y uso. De las presentadas en este trabajo, las técnicas de minería textual y las de analítica textual pueden llevarse a cabo mediante herramientas de uso simple que no necesitan mayor capacitación para los investigadores; otras, como el análisis visual de redes o la visión computada, requieren de entrenamientos específicos en el primer caso, y de trabajos interdisciplinarios con la ciencia de datos, en el segundo. Futuros trabajos podrán aportar, mediante diversos métodos al conocimiento acerca de los tipos de técnicas digitales y sus contextos de investigación en Latinoamérica.

Referencias

- Berthel, K. U. (2006) *Image Sorter v. 2.01* [software de computadora en disco]. Berlín: Mensch-Maschine-Kommunikation.
- Blondel, V. D.; Guillaume, J.-L.; Lambiotte, R. & Lefebvre, E. (2008). Fast unfolding of communities in large networks. *Journal of Statistical Mechanics: Theory and Experiment*, 2008(10), 10008. <https://doi.org/10.1088/1742-5468/2008/10/P10008>
- Borra, E. & Rieder, B. (2014). Programmed method: Developing a toolset for capturing and analyzing tweets. *Aslib Journal of Information Management*, 66(3), 262-278. <https://doi.org/10.1108/AJIM-09-2013-0094>
- Burgess, J. & Green, J. (2018). *YouTube: Online video and participatory culture. Second Edition*. Cambridge: Polity Press
- Cairo, A. (2018). The elements of visualization. Trabajo presentado en Data Visualization for Storytelling and Discovery. Curso en línea de la Fundación Knight. Recuperado de: https://journalismcourses.org/courses/DE0618/Module1_Video2_Presentation.pdf
- Campos-Vázquez, R., & López-Araiza, S. E. (2018). El estatus de la ciencia económica en México. *El Trimestre Económico*, 85(340), 683-700. <https://doi.org/10.20430/ete.v85i340.771>
- Carvajal, R. (2013). ¿Qué es el scraping y cómo hacerlo bien? *Manual de Periodismo de Datos Iberoamericano* Recuperado de: <http://manual.periodismodedatos.org/rigoberto-carvajal.php>
- Cortés Sánchez, J. D. (2018). Mission statements of universities worldwide: Text mining and visualization. *Intangible Capital*, 14(4), 584-603. <https://doi.org/10.3926/ic.1258>
- Elmer, G. (2015). Scraping the first person, en *Compromised Data: From Social Media to Big Data* (pp. 112-125). New York: Bloomsbury Academic.
- Fielding, N. G.; Lee, R. M. & Blank, G. (2016). *The SAGE Handbook of Online Research Methods*. SAGE.
- Fontcuberta, J. (2011). Por un manifiesto postfotográfico. *La Vanguardia*. Recuperado de: <http://www.lavanguardia.com/cultura/20110511/54152218372/por-un-manifiesto-posfotografico.html>
- Fulton, K. (2014). Screen scraping: How to stop the internet's invisible data leeches. *TechRadar*. Recuperado de:

- <http://www.techradar.com/news/internet/web/screen-scraping-how-to-stop-the-internet-s-invisible-data-leaches-1214404>
- Gandomi, A. & Haider, M. (2015). Beyond the hype: Big data concepts, methods, and analytics. *International Journal of Information Management*, 35(2), 137–144. <https://doi.org/10.1016/j.ijinfomgt.2014.10.007>
- Gerlitz, C. & Helmond, A. (2013). The like economy: Social buttons and the data-intensive web. *New Media & Society*. <https://doi.org/10.1177/1461444812472322>
- Goldbeck, J. (2013). *Analyzing the Social Web* (1st Edition). Morgan Kaufmann. Recuperado de: <https://www.elsevier.com/books/analyzing-the-social-web/golbeck/978-0-12-405531-5>
- Han, J.; Kamber, M. & Pei, J. (2011). *Data Mining: Concepts and Techniques*. Waltham, United States of America: Morgan Kauffman/Elsevier.
- Highfield, T. & Leaver, T. (2016). Instagrammatics and digital methods: Studying visual social media, from selfies and GIFs to memes and emoji. *Communication Research and Practice*, 2(1), 47-62. <https://doi.org/10.1080/22041451.2016.1155332>
- Hine, C. (2004). *Etnografía virtual*. Barcelona: Editorial UOC.
- Hine, C. (2015). *Ethnography for the Internet: Embedded, embodied and everyday*. London: Bloomsbury Academic, an imprint of Bloomsbury.
- Hochman, N. & Manovich, L. (2013). Zooming into an Instagram City: Reading the local through social media. *First Monday*, 18(7). <https://doi.org/10.5210/fm.v18i7.4711>
- Lacalle, J. M. y Vilar, M. A. (2019). Estudios literarios y lectura distante: Un primer acercamiento a la actualidad de la investigación en las revistas académicas argentinas. *Anclajes*, 23(1), 19-40.
- Latour, B. (1992). *Ciencia en acción: Cómo seguir a los científicos e ingenieros a través de la sociedad*. Barcelona: Labor.
- Manovich, L. (2009). Cultural Analytics: Visualizing Patterns in the era of more media. Sitio Web de Lev Manovich http://manovich.net/content/04-projects/063-cultural-analytics-visualizing-cultural-patterns/60_article_2009.pdf
- Manovich, L. (2016). The Science of Culture? Social Computing, Digital Humanities and Cultural Analytics. *Journal of Cultural Analytics*. <https://doi.org/10.22148/16.004>
- Marres, N. & Gerlitz, C. (2016). Interface Methods: Renegotiating Relations between Digital Social Research, STS and Sociology. *The Sociological Review*, 64(1), 21-46. <https://doi.org/10.1111/1467-954X.12314>
- Marres, N. & Weltevrede, E. (2013). Scraping the Social? Issues in live social research. *Journal of Cultural Economy*, 6(3), 313-335.
- Marres, N. (2017). *Digital Sociology: The Reinvention of Social Research*. Cambridge, United Kingdom: Polity Press.
- Monterde Mateo, A.; Carrillo Martin, R.; Esteve del Valle, M. & Aragón Aragón, P. (2015). #YoSoy132: ¿Un nuevo paradigma en la política mexicana? *IN3 Working Paper Series*. Recuperado de: <http://in3-working-paper-series.uoc.edu/in3/es/index.php/in3-working-paper-series/article/view/2066.html>
- Moreno, A. & Redondo, T. (2016). Text Analytics: the convergence of Big Data and Artificial Intelligence. *International Journal of Interactive Multimedia and Artificial Intelligence*, 3(Special Issue on Big Data and AI), 57-64. <https://doi.org/10.9781/ijimai.2016.369>
- Nieborg, D. B. & Poell, T. (2018). The platformization of cultural production: Theorizing the contingent cultural commodity. *New Media & Society*, 20(11), 4275-4292. <https://doi.org/10.1177/1461444818769694>

- Niederer, S. & Taudin Chabot, R. (2015). Deconstructing the cloud: Responses to Big Data phenomena from social sciences, humanities and the arts. *Big Data and Society*, 2(2). <http://journals.sagepub.com/doi/full/10.1177/2053951715594635>
- Paltoglou, G. & Thelwall, M. (2012). Twitter, MySpace, Digg: Unsupervised Sentiment Analysis in Social Media. *ACM Trans. Intell. Syst. Technol.*, 3(4), 1-19. <https://doi.org/10.1145/2337542.2337551>
- Pearce, W.; Özkula, S. M.; Greene, A. K.; Teeling, L.; Bansard, J. S.; Omena, J. J. & Rabello, E. T. (2018). Visual cross-platform analysis: Digital methods to research social media images. *Information, Communication & Society*. <https://doi.org/10.1080/1369118X.2018.1486871>
- Pedraza Bucio, C. I. & Cano Rodríguez, C. A. (2019). Resistencias sumergidas. Cartografía de la tecnopolítica feminista en México. *Teknokultura. Revista de Cultura Digital y Movimientos Sociales*, 16(2), 197-212. <https://doi.org/10.5209/tekn.64163>
- Popping, R. (2016). Online Tools for Content Analysis, en *The Sage Handbook of Online Research Methods* (pp. 329-343). London, New York: SAGE.
- Reguillo Cruz, R. (2017). *Paisajes insurrectos: Jóvenes, redes y revueltas en el otoño civilizatorio*. Barcelona: NED Ediciones.
- Robichaud, A. & Blevins, C. (2011). 4: Content-Based Analysis » Tooling Up for Digital Humanities. *Tooling Up Digital Humanities*. Recuperado de: http://toolingup.stanford.edu/?page_id=205
- Rogers, R. (2009). *The End of Virtual. Digital Methods*. Amsterdam: Vossiuspers UvA. Recuperado de: http://www.govcom.org/rogers_oratie.pdf
- Rogers, R. (2013). *Digital methods*. Boston: MIT Press.
- Rogers, R. (2015). Digital Methods for Web Research, en *Emerging Trends in the Social and Behavioral Sciences*. <https://doi.org/10.1002/9781118900772.etrds0076>
- Rogers, R. (2018). Otherwise Engaged: Social Media from Vanity Metrics to Critical Analytics. *International Journal of Communication*, 12(0), 23. Recuperado de: <https://dare.uva.nl/search?identifier=e7a7c11b-b199-4d7c-a9cb-fdf1dd74d493>
- Rogers, R. (2019). *Doing Digital Methods*. London: SAGE.
- Rojas Crotte, I. R. R. (2011). Elementos para el diseño de técnicas de investigación: una propuesta de definiciones y procedimientos en la investigación científica. *Tiempo de educar*, 12(4), 277-297.
- Rose, G. (2016). *Visual methodologies: An introduction to researching with visual materials* (4ta. edición). Londres: SAGE.
- Sightengine. (2018). Benchmarking Google Vision, Amazon Rekognition, Microsoft Azure on Image Moderation. Medium. Recuperado de: <https://medium.com/sightengine/benchmarking-google-vision-amazon-rekognition-microsoft-azure-on-image-moderation-73909739b8b4>
- Solomon, B. (2016). How Tableau Built A \$3 Billion Data Empire on Top of Beautiful Charts. *Forbes*. Recuperado de: <https://www.forbes.com/sites/briansolomon/2016/05/04/how-tableau-built-a-3-billion-data-empire-on-top-of-beautiful-charts/>
- Sued, G. (2018). Métodos digitales para el estudio de la fotografía compartida. Una aproximación distante a tres ciudades iberoamericanas en Instagram. *Empiria. Revista de metodología de ciencias sociales*, 0(40), 15-39. <https://doi.org/10.5944/empiria.40.2018.22009>
- Thelwall, M. (2017). Heart and soul: Sentiment strength detection in the social web with SentiStrength. En Hlys, J (ed) *Cyberemotions: Collective emotions in cyberspace* (pp. 437-445). Recuperado de <http://sentistrength.wlv.ac.uk/documentation/SentiStrengthChapter.pdf>

- Tremayne, M. (2014). Anatomy of Protest in the Digital Era: A Network Analysis of Twitter and Occupy Wall Street. *Social Movement Studies*, 13(1), 110-126. <https://doi.org/10.1080/14742837.2013.830969>
- Underwood, T. (2013). Wordcounts are amazing. *The Stone and the Shell*. Recuperado de: <https://tedunderwood.com/2013/02/20/wordcounts-are-amazing/>
- Venturini, T.; Jacomy, M. & Pereira, D. (2015). Visual Network Analysis. Recuperado de: [www.tommasoventurini.it http://www.tommasoventurini.it/wp/wp-content/uploads/2014/08/Venturini-Jacomy_Visual-Network-Analysis_WorkingPaper.pdf](http://www.tommasoventurini.it/wp/wp-content/uploads/2014/08/Venturini-Jacomy_Visual-Network-Analysis_WorkingPaper.pdf)
- Vitale, A. (3 de diciembre 2015). What Google Cloud Vision API means for Deep Learning Startups. *Medium*. Recuperado de: <https://medium.com/google-cloud/what-google-cloud-vision-api-means-for-deep-learning-startups-cd39226922e5>
- Voyant Tools. (2018). En Wikipedia. *Wikipedia*. Recuperado de: https://en.wikipedia.org/w/index.php?title=Voyant_Tools&oldid=845947425
- Wang, C.; Komodakis, N. & Paragios, N. (2013). Markov Random Field modeling, inference & learning in computer vision & image understanding: A survey. *Computer Vision and Image Understanding*, 117(11), 1610-1627. <https://doi.org/10.1016/j.cviu.2013.07.004>
- Wickham, H. (2014). Tidy Data. *Journal of Statistical Software*, 59(10), 1-23.

Este artículo es de acceso abierto. Los usuarios pueden leer, descargar, distribuir, imprimir y enlazar al texto completo, siempre y cuando sea sin fines de lucro y se cite la fuente.

CÓMO CITAR ESTE ARTÍCULO:

- Sued, G. E. (2020). Repertorio de técnicas digitales para la investigación con contenidos generados en redes sociodigitales. *Paakat: Revista de Tecnología y Sociedad*, 10(19). <http://dx.doi.org/10.32870/Pk.a10n19.498>
-

* Gabriela Elisa Sued. Es especialista en cultura digital. Doctora en Humanidades por el Tecnológico de Monterrey. Correo electrónico: gabriela.sued@gmail.com

¹ Ver <http://voyant-tools.org/>

² Ver <https://signalab.iteso.mx/>

³ Tutorial sobre Image Plot en <http://lab.softwarestudies.com/p/imageplot.html>

⁴ Ver <http://www.tucows.com/preview/510399/ImageSorter>

⁵ Ver <https://apps.crowdtangle.com/chrome-extension>

⁶ Ver <https://tableausoftware>

⁷ Ver <https://rawgraphs.io/>

⁸ Ver <https://www.trifacta.com/start-wrangling/>

⁹ Ver <https://www.draw.io/>