

---

# Analítica cultural

Lev Manovich

Usaré el término analítica cultural (*Cultural Analytics*) para referirme a la aplicación de métodos matemáticos, computacionales y de visualización de datos en el análisis de objetos, actividades y comportamientos culturales. Mientras que estos métodos también se emplean en los estudios de objetos y archivos culturales históricos, son particularmente pertinentes para la cultura digital contemporánea debido a su escala masiva, su variedad y su velocidad.

## *¿Cómo ver la cultura digital global?*

La organización material de la cultura digital también facilita su análisis computacional. Las expresiones culturales digitales nacen

---

*N. del T.:* En inglés, *Cultural Analytics*. Aunque en castellano encontramos también el término «cultura analítica», fundamentalmente en el entorno empresarial y del marketing, aquí se ha optado por usar «analítica cultural», más respetuoso con el término original y acuñado especialmente en el ámbito de las Humanidades.

como archivos digitales. Las huellas del comportamiento *online* se registran en forma digital. Esto permite el análisis de estas expresiones y trazas utilizando algoritmos y pasando por alto la etapa de digitalización requerida para objetos que no son digitales.

La aplicación de métodos cuantitativos en los análisis culturales comenzaron mucho antes que los ordenadores. Un ejemplo lo constituyen los estudios del poeta y teórico de la literatura Andréi Bely (hijo del prominente matemático) sobre el problema del ritmo en la poesía rusa. Bely usó estadísticas e inventó un sistema gráfico para visualizar estructuras en los poemas. Estos estudios fueron recogidos en su libro *Symbolism* de 1910<sup>1</sup>. Otro ejemplo lo aporta el influyente artículo de 1919 del antropólogo estadounidense Alfred L. Kroeber titulado *On the Principle of Order in Civilization as Exemplified by Changes of Fashion*, en el que presentó un análisis cuantitativo y de interpretación de patrones en la moda femenina del siglo XIX.

A principios de la década de 1990, varios artistas digitales crearon visualizaciones de datos de patrones culturales. En *The Shape of Song* (Martin Wattenberg, 2001) y *History Flow* (Fernanda Viégas y Martin Wattenberg, 2003) se visualizan patrones en composiciones musicales y en páginas de Wikipedia en su dimensión temporal cambiante<sup>2</sup>. La instalación *Listening Post* (Mark Hansen y Ben Rubin, 2001-2002) filtraba en tiempo real discusiones académicas en salas de chat y elaboraba animaciones con los fragmentos de texto que extraía a través de una serie de cientos de pantallas<sup>3</sup>. En *The Top Grossing Film of All Time, 1 x 1* (Jason Salavan, 2000), los fotogramas de la película *Titanic* se redujeron a sus colores más

---

<sup>1</sup> Ver Lynn Gamwell, *Mathematics and Art: A Cultural History* (Princeton University Press, 2015), pp. 169-170.

<sup>2</sup> <http://www.bewitched.com/song.html>, <http://www.bewitched.com/historyflow.html>

<sup>3</sup> <https://frieze.com/article/mark-hansen-and-ben-rubin>

representativos y los pequeños cuadrados en estos colores se dispusieron en una cuadrícula rectangular, haciendo así visible el ritmo visual de la película<sup>4</sup>.

La analítica cultural a gran escala sólo fue posible a mediados de los años 2000. Su desarrollo estuvo motivado por el sorprendente crecimiento de Internet y de las redes sociales en las que cientos de millones de personas y millones de instituciones y empresas comparten contenidos. Esta explosión comienza alrededor de 2004 (Flickr, Facebook, MySpace), se acelera con la adopción global masiva de las redes sociales (2006-2007) y, luego, con los teléfonos con conexión a Internet y con «capacidades mediáticas» (*media capabilities*). En la década de 2010, aumentan las personas de países en desarrollo que se hacen con estos teléfonos y comienzan a crear y compartir contenidos. Por ejemplo, en apenas dos años, de 2013 a 2015, el promedio de personas en 21 países en desarrollo que accedían a Internet aumentó del 45 al 51 por ciento<sup>5</sup>.

¿Cómo podemos estudiar este universo en expansión de contenidos e interacciones culturales? Por ejemplo, ¿cómo podemos estudiar la fotografía popular contemporánea que aumenta en miles de millones de imágenes cada día? ¿Y las tendencias en el contenido de los mil millones de tableros en Pinterest?<sup>6</sup> ¿Y los patrones de los intereses culturales de los habitantes de los 180 países que constituyen los 250.000 grupos de Meetup y sus 613.000 reuniones mensuales?<sup>7</sup> (Esto es «cultura digital» porque los encuentros son posibles gracias a Internet).

---

<sup>4</sup> <http://www.salavon.com/work/TopGrossingFilmAllTime/>

<sup>5</sup> <http://www.pewglobal.org/2016/02/22/smartphone-ownership-and-internet-usage-continues-to-climb-in-emerging-economies/>

<sup>6</sup> Jordan Novet, «80 percent of Pinterest's traffic comes from mobile devices», 31 de marzo de 2015, <http://venturebeat.com/2015/03/31/80-percent-of-pinterests-traffic-comes-from-mobile-devices/>

<sup>7</sup> <https://www.meetup.com/about/>, última consulta el 26 de octubre de 2016.

A medida que personas en todo el mundo crean, comparten, mezclan (*remix*) e interactúan con billones de artículos de contenido digital diariamente, la escala de la cultura humana se aproxima a la de los sistemas físicos o biológicos naturales. Esto sugiere que las técnicas matemáticas y computacionales, que ya se han utilizado con éxito para estudiar el universo físico y los organismos vivos, también pueden usarse para estudiar culturas. Los ejemplos de tales métodos provienen de la estadística descriptiva e inferencial, el aprendizaje de las máquinas, la simulación, el análisis de redes, la *web science*, y los sistemas dinámicos complejos.

El análisis cuantitativo de grandes muestras de actividad cultural global puede revelar muchas regularidades y patrones que pueden llegar a tener la generalidad de las leyes de la ciencia. Por ejemplo, el análisis de las actividades de los usuarios en las redes sociales a menudo revela la presencia de una distribución estadística denominada «ley potencial» (*power law*)<sup>8</sup>. Al mismo tiempo, podemos utilizar las matemáticas y la computación para satisfacer el objetivo de las «ciencias humanas» (*Geisteswissenschaft*) de comprender e interpretar las diferentes expresiones humanas. Pueden ayudarnos a ver mejor lo que es particular, único y característico de una cultura, en lugar de reducir las expresiones humanas a leyes generales como hacemos en la ciencia. ¿Por qué? Cuando analizamos cuantitativamente un número masivo de expresiones culturales humanas, se hace visible lo singular, lo único. En lugar de ver numerosas expresiones culturales como estereotipos –las instancias de un pequeño número de prototipos–, revelamos un «ADN» cultural único en cada expresión, incluso si algunas de ellas son iguales.

La analítica cultural no reemplaza los métodos existentes de análisis cultural. No podemos simplemente someter «datos cultura-

---

<sup>8</sup> <http://www.nature.com/articles/srep01783>

les» al análisis computacional o a la visualización de datos y esperar que aparezcan ideas interesantes. Para producir resultados útiles en lugar de seguir reinventando viejas preguntas y respuestas, la investigación de la analítica cultural debe basarse en los paradigmas y las teorías desarrolladas en los estudios de las culturas en los últimos 150 años. Las disciplinas relevantes incluyen la antropología cultural, la sociología de la cultura, las humanidades académicas (historia y teoría de la literatura, arte, cultura visual, música, cine, *media*, juegos, arquitectura y diseño), los estudios culturales, los *media studies*, las ciencias de la comunicación, los estudios digitales y los *software studies*. Esto no implica que la analítica cultural sólo esté interesada en probar, validar o rechazar hipótesis existentes desarrolladas previamente por estas disciplinas. El desarrollo de nuevos conceptos y teorías necesarias para el análisis cuantitativo de la cultura en general, y de la cultura digital en particular, es una parte importante de la analítica cultural. Del mismo modo, en lugar de utilizar únicamente métodos matemáticos, computacionales y de visualización de datos ya existentes, la analítica cultural puede estimular el desarrollo de nuevos métodos apropiados para el análisis de cualquier tema. Uno de estos ejemplos lo constituyen las técnicas de visualización desarrolladas por Damon Crockett en nuestro laboratorio para comparar patrones en grandes colecciones de imágenes<sup>9</sup>.

### *Ejemplos de analítica cultural*

Hoy en día, el trabajo y las herramientas de la analítica cultural se pueden encontrar en muchos lugares. Incluyen publicaciones en

---

<sup>9</sup> Damon Crockett, «Direct visualization techniques for the analysis of image data: the slice histogram and the growing entourage plot», en *International Journal for Digital Art History*, no. 2 (2016), <http://journals.ub.uni-heidelberg.de/index.php/dah/article/view/33529/27217>

revistas de ciencia y humanidades, código *GitHub* y depósitos de datos, grandes proyectos institucionales centrados en reunir a largo plazo registros digitales de muchas colecciones separadas, como *europaena.eu*, e instalaciones de arte de corta duración en museos y espacios públicos. La mayoría de las redes sociales más populares incluyen un panel de control que permite a los usuarios ver los patrones de su actividad.

En términos de cantidad, la mayoría de las publicaciones y los artículos relevantes provienen de la informática. Algunos de estos investigadores identifican este tipo de trabajo usando el término «computación social» (*social computing*)<sup>10</sup>. Otros pertenecen a diferentes campos de la informática, tales como la computación multimedia, la visión artificial (*computer vision*), la recuperación de información musical, el procesamiento del lenguaje natural, la *web science* y el aprendizaje automático. El número de publicaciones y artículos que analizan los contenidos culturales y las interacciones supera ya los cientos de miles.

Muchas publicaciones científicas justifican su investigación haciendo referencia a aplicaciones industriales ya existentes, por ejemplo, sistemas de búsqueda o recomendación de contenidos *online*. El supuesto general es que la informática creará mejores algoritmos así como otras tecnologías informáticas útiles para la industria y otras organizaciones. Sin embargo, mirando muchos de los ejemplos de estos artículos, se hace evidente que en realidad están haciendo preguntas típicas de las humanidades o de los *media studies* aplicadas a los medios contemporáneos, pero a una escala

---

<sup>10</sup> Ver los programas de los congresos sobre estos temas, por ejemplo, la *19th ACM Conference on Computer-Supported Cooperative Work and Social Computing*, San Francisco, del 27 de febrero al 2 de marzo de 2016 (último acceso el 29 de julio de 2016, <http://cscw.acm.org/2016/submit/>) o la *24th International World Wide web Conference*, Florencia, Italia, del 18 al 22 de mayo de 2015. (Última consulta el 29 de julio de 2016, <http://www.www2015.it/accepted-papers/>).

mucho mayor. Consideremos, por ejemplo, estas publicaciones: «Quantifying Visual Preferences around the World» y «What We Instagram: A First Analysis of Instagram Photo Content and User Types»<sup>11</sup>. El primer estudio analiza las preferencias mundiales para el diseño de páginas web a través de 2,4 millones de calificaciones de 40.000 personas de 179 países. Obviamente, el estudio de la estética y el diseño tradicionalmente ha formado parte de las humanidades. El segundo estudio analiza los temas más frecuentes de las fotos de Instagram, una cuestión que se puede comparar con los estudios de historia del arte sobre los géneros en la pintura holandesa del siglo XVII.

Otro ejemplo lo constituye un influyente artículo titulado «What is Twitter, a Social Network or a News Media?»<sup>12</sup>. Publicado en 2010, ha sido citado 4.271 veces<sup>13</sup>. Este artículo ofreció el primer análisis a gran escala de la red social Twitter con 106 millones de tweets de 41,7 millones de usuarios. El estudio examinó los *trending topics*, explorando «en qué categorías se clasifican los *trending topics*, cuánto duran y cuántos usuarios participan». Este análisis puede ser visto como la actualización del trabajo clásico en el

---

<sup>11</sup> Katharina Reinecke *et al.*, «Quantifying Visual Preferences Around the World», en *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (New York: ACM, 2014): 11-20. Última consulta el 29 de julio de 2016, <http://www.eecs.harvard.edu/~kgajos/papers/2014/reinecke14visual.pdf>; Yuheng Hu *et al.*, «What We Instagram: A First Analysis of Instagram Photo Content and User Types», en *Proceedings of the 8th International AAAI Conference on Weblogs and Social Media* (ICWSM, 2014). Última consulta el 29 de julio de 2016, <http://rakaposhi.eas.asu.edu/instagram-icwsm.pdf>

<sup>12</sup> Haewoon Kwak *et al.*, «What is Twitter, a Social Network or a News Media?» en *Proceedings of the 19th International World Wide Web (WWW) Conference* (ACM, 2014): 591-600. Última consulta el 29 de julio de 2016, <http://www.eecs.wsu.edu/~assefaw/CptS580-06/papers/2010-www-twitter.pdf>.

<sup>13</sup> *Google Scholar*, «Haewoon Kwak». Última consulta el 29 de julio de 2016, <https://scholar.google.com/citations?user=M6i3Be0AAAAJ&hl=en>

campo de la comunicación, que se remonta a la obra pionera de Paul F. Lazarsfeld y sus colegas a finales de la década de 1930, cuando contaban manualmente los temas de las emisiones radiofónicas<sup>14</sup>. La gran diferencia es que, en los años treinta, las emisiones fueron creadas por un número limitado de estaciones profesionales que recurrían a un número pequeño de géneros, mientras que Twitter puede tener numerosos temas con diferentes niveles de generalidad, duración y cobertura geográfica. Al mismo tiempo, dado que Twitter y otros servicios de *microblogging* representan una nueva forma mediática como la pintura al óleo, los libros impresos y la fotografía tiempo atrás, estudiar la especificidad de Twitter como medio también debería ser un tema para las humanidades.

Las publicaciones científicas más relevantes normalmente sólo analizan grandes muestras de contenido digital *online*. En humanidades, muchos estudiosos han analizado textos literarios históricos usando una variedad de métodos computacionales y algoritmos desarrollados originalmente en la informática (especialmente en el campo de la recuperación de información) y en el análisis de redes. Por ejemplo, en «Mapping Mutable Genres in Structural Complex Volumes», el erudito literario Ted Underwood y sus colaboradores han analizado 469.200 volúmenes de la Trust Digital Library<sup>15</sup>. Los investigadores también aplicaron tales métodos a otros tipos de textos como los periódicos del siglo XIX<sup>16</sup>.

---

<sup>14</sup> Paul F. Lazarsfeld; Frank N. Stanton (eds.), *Radio Research, 1941* (New York: Duel, Sloan and Pearce, 1942).

<sup>15</sup> Ted Underwood *et al.*, «Mapping Mutable Genres in Structurally Complex Volumes», en *Proceedings of the 2015 IEEE Conference on Big Data* (IEEE, 2013). Última consulta el 29 de julio de 2016, <http://arxiv.org/abs/1309.3323>

<sup>16</sup> Ver, por ejemplo: David A. Smith *et al.*, «Infectious Texts: Modeling Text Reuse in Nineteenth-century Newspapers», en *Proceedings of 2015 IEEE Conference on Big Data* (IEEE, 2013): 84-94. Última consulta el 29 de julio de 2016, <http://www.ccs.neu.edu/home/dasmith/infect-bighum-2013.pdf>



Algunos investigadores han presentado estudios sobre *datasets* históricos de *media* visuales y de audio que utilizan creativamente métodos de la informática tomados de los campos del procesamiento de imágenes, la visión artificial (*computer vision*) y la recuperación de información musical (MIR). Los ejemplos incluyen artículos como «Toward Automated Discovery of Artistic Influence»<sup>17</sup>, «Measuring the Evolution of Contemporary Western Popular Music»<sup>18</sup> y «Quicker, Faster, Darker: Changes in Hollywood Film over 75 Years»<sup>19</sup>. El primero presenta un modelo matemático para el descubrimiento automático de la influencia entre artistas. El segundo investiga los cambios en la música popular usando un *dataset* de 500.000 canciones. El tercero analiza los cambios graduales en una forma cinematográfica ejemplificada por 160 películas creadas entre 1935 y 2010.

La analítica cultural no sólo involucra a los investigadores que hacen análisis en sus laboratorios y publican los correspondientes artículos, sino también a la creación de interfaces web interactivas que permiten al público investigar tendencias en *datasets* culturales. Uno de los proyectos prominentes en este sentido es *N-Gram Viewer*, creado en 2010 por los científicos de Google Jon Orwant y Will Brockman siguiendo el prototipo de dos estudiantes de doctorado de Harvard en biología y matemáticas aplicadas<sup>20</sup>. Un visitante de esta web puede introducir una serie de palabras o frases,

---

<sup>17</sup> Babak Saleh *et al.*, «Toward Automated Discovery of Artistic Influence», en *Multimedia Tools and Applications* 75, no. 7 (2016): 3565-3591, doi: 10.1007/s11042-014-2193-x

<sup>18</sup> Joan Serrà *et al.*, «Measuring the Evolution of Contemporary Western Popular Music», en *Nature Scientific Reports* 2, article no. 521 (2012), doi:10.1038/srep00521

<sup>19</sup> James E. Cutting *et al.*, «Quicker, Faster, Darker: Changes in Hollywood Film over 75 Years», en *i-Perception* 2, no. 6 (2011): 569-576, doi: 10.1068/i0441aap.

<sup>20</sup> <https://books.google.com/ngrams>

y al instante ver gráficos que comparan las frecuencias de las apariciones de estas palabras a través de millones de libros publicados a lo largo de varios siglos. En mi propio laboratorio (*lab.cultural.analytics.info*) hemos creado dos proyectos que permiten a los visitantes visualizar tendencias en colecciones de imágenes extraídas de las redes sociales. *Selficity* (2014-2015) permite comparar patrones en miles de autorretratos de Instagram (*selfies*) compartidos en seis ciudades del mundo. En otro proyecto titulado *On Broadway* (2014) creamos una interfaz para la exploración del contenido y los rastros de las actividades de la gente en la calle Broadway en Manhattan. Los datos incluyen 660.000 fotos de Instagram geocodificadas, ocho millones de *check-ins* en Foursquare y 22 millones de recogidas y paradas de taxi.

### *Abordar la cultura: comportamientos versus objetos culturales*

Los enfoques modernos sobre el estudio de la cultura pueden agruparse en dos categorías. El primero se centra en las normas, convenciones, valores, comunicaciones y comportamientos personales e interpersonales. La antropología cultural ejemplifica mejor este paradigma. El segundo se centra en los objetos culturales creados por individuos o grupos. Este paradigma encuentra un ejemplo en campos de las humanidades académicas como los estudios literarios. El ámbito de la comunicación y los estudios culturales a veces combinan ambos enfoques: aquí, los investigadores analizan los mensajes culturales, las formas en que influyen en las personas, las maneras en que las personas pueden apropiárselos para construir sus significados y las actividades de sus productores.

La investigación analítica cultural puede beneficiarse de ambas visiones de la cultura: como comportamientos, convenciones y va-

lores, y como objetos y mensajes. El estudio de ambos tipos de datos a escala requiere igualmente métodos computacionales y nuevas tecnologías de detección. Para estudiar los comportamientos y mensajes *online*, necesitamos recopilar y analizar cuantitativamente los vínculos entre las personas en las redes sociales, sus mensajes, comentarios, *re-shares* y otras actividades. El estudio de los comportamientos físicos se beneficia de estas nuevas tecnologías que capturan los registros y los rastros de estos comportamientos a escala: fotografía por satélite, GPS, sensores, aplicaciones móviles, etc.

Sin embargo, el uso de tales tecnologías para la investigación cultural también trae consigo los problemas relacionados con la privacidad y el acceso a los datos. La anonimización necesaria de los datos de las redes sociales antes del análisis dificulta la interpretación de estos datos. Por ejemplo, si conociéramos la edad, el género, la etnia, el lugar de residencia, la educación y otros detalles sobre los autores de los *posts* de las redes sociales que podemos recopilar *online*, esto nos ayudaría a interpretar las tendencias que podemos extraer, con algoritmos, de las colecciones de tales publicaciones. Esta es sólo una de las razones por las que el uso de métodos de menor escala como la observación, las encuestas, la «lectura densa» (*thick reading*) y la «lectura atenta» (*close reading*) continúan siendo muy beneficiosos. Ser capaz de observar a gente concreta en entornos de la vida real o leer manualmente textos, da a los investigadores el contexto más amplio que a menudo falta si empezamos con datos culturales y de comportamiento capturados automáticamente.

### *La analítica cultural y la industria tecnológica*

En una historia de los estudios modernos sobre la cultura, la analítica cultural representaría una nueva «tercera etapa» posible.

A finales del siglo XIX, los antropólogos de la cultura Edward Tylor y Franz Boas articularon el concepto moderno de cultura como los símbolos, las formas expresivas y las creencias presentes en todas las sociedades y grupos humanos. Esto llevó a los estudios comparativos de las sociedades tradicionales a pequeña escala todavía existentes en todo el mundo (y que se encontraban a menudo en las colonias de los países occidentales). Esta perspectiva también enfatizó que todas las culturas son distintas y que la opinión entonces popular de que algunas culturas son superiores a otras es incorrecta. Boas utilizó esta nueva perspectiva de la antropología cultural para combatir la discriminación existente en América contra los inmigrantes, los negros y las personas revolucionarias.

En el siglo XX emergen nuevas disciplinas y paradigmas para estudiar las culturas de masas contemporáneas en los países occidentales, y más tarde en todo el mundo. Estas culturas de masas son posibles por las nuevas tecnologías de comunicación de masas como la radio, los registros de audio y la televisión. Los ejemplos más destacados son la teoría crítica alemana (década de 1920), la teoría de la comunicación (década de 1950), los estudios culturales británicos (años sesenta) y los americanos (años ochenta), los *Film y TV Studies* (años setenta), y los *Game Studies* (1990). Los investigadores que trabajan en estos campos y paradigmas estudian las industrias de los medios y la comunicación de masas en las sociedades contemporáneas.

Los procesos de globalización después del final de la Guerra Fría (1991) y el desarrollo de redes planetarias como la *World Wide Web* (1990) crearon una nueva cultura digital global a principios del siglo XXI. La analítica cultural es un enfoque de estudio que corresponde a esta nueva etapa. Si la gente hoy usa herramientas de *software* (aplicaciones de edición de fotos, servicios de blogs, etc.) para crear objetos digitales, la analítica cultural también utiliza *software* para analizar grandes muestras de tales objetos. Si las

empresas tecnológicas responsables de plataformas e interfaces digitales como Amazon, Netflix, Spotify, YouTube, Facebook, etc., utilizan la ciencia de datos (*data science*) para analizar, mostrar y recomendar contenidos, la analítica cultural también se basa en la ciencia de datos para estudiar los patrones de los objetos culturales y la interacción de las personas con estos mismos objetos.

Cuando las disciplinas académicas del siglo XX estudiaban los productos de las industrias culturales como la música popular, las películas, los programas de televisión y los juegos, no dependían directamente de estas industrias. Sus objetos de estudio eran a menudo programas y productos singulares disponibles a través de la radio, los discos, la televisión, las salas de cine, las librerías y los quioscos, los mismos productos consumidos por los particulares. Para recopilar y analizar múltiples programas, los académicos utilizaron la misma tecnología de consumo que los espectadores y oyentes normales. Por ejemplo, entre 1980 y 1990 cada estudio de los *Television Studies* utilizaba un aparato de vídeo para grabar su propia colección de programas. Cuando yo mismo enseñaba «arte de los nuevos medios» en la década de 1990 y la primera parte de los años de 2000, conté con mi propia colección de cientos de cintas de vídeo y CD-ROM acumulados a lo largo de los años. (Esta era mi posesión más valiosa en ese momento).

Sin embargo, la analítica cultural depende del acceso a un gran número de objetos digitales y de datos sobre sus creadores, usuarios y sus interacciones entre sí y estos objetos. La mayoría de las plataformas digitales de gran tamaño proporcionan estos datos a través de APIs. Este término es una abreviatura de *Application Programming Interface*, un conjunto de comandos utilizados para la interacción con el servicio digital a través de un programa informático. Las APIs para las redes sociales se hicieron comunes en la segunda mitad de los años 2000 porque estas empresas querían estimular el desarrollo de aplicaciones de terceros para poder interactuar con

el contenido compartido en estas redes. Las APIs también permiten incrustar el contenido compartido de una plataforma en otra y el intercambio automático de la información del usuario entre plataformas. Por ejemplo, en lugar de publicar directamente en cada una de sus redes sociales, usted puede utilizar servicios como Buffer o Hootsuite para publicar automáticamente en todas estas redes. Estos servicios se comunican con las redes a través de APIs, lo que les permite publicar en su nombre.

Aunque su propósito original era el uso automático en aplicaciones de terceros y otras plataformas de redes sociales, las APIs también permiten la descarga gratuita en su ordenador de grandes volúmenes de contenido compartido por muchos usuarios. Dado que estas APIs eran abiertas para que todo el mundo las utilizase, los investigadores interesados en el análisis cualitativo de grandes muestras de contenido y en la actividad de los usuarios en las redes sociales se aprovecharon de la situación, lo que llevó a un rápido crecimiento de la investigación y a los cientos de miles de publicaciones y artículos que ya mencioné. Si estas APIs gratuitas no existieran, nuevos campos como el de la computación social y la ciencia social computacional (*computational social science*) no se habrían desarrollado en absoluto, o lo habrían hecho a un ritmo mucho más lento.

Si las empresas decidiesen limitar significativamente el acceso a sus APIs (como hizo Instagram el 1 de junio de 2016), o eliminarlas completamente, el efecto sería muy significativo en estos campos. Por fortuna, aún tenemos otros métodos para obtener grandes cantidades de datos *online* como el *Web Scraping*. Este método ya fue utilizado por los emergentes análisis de redes en la segunda mitad de los años noventa. De hecho, cuando me di cuenta de las posibilidades de la analítica cultural en 2005, así es como yo imaginaba obtener grandes datos culturales de la web. Otro método es el rastreo web utilizado por los motores de bús-

queda para localizar e indexar el contenido de miles de millones de páginas web.

*La analítica cultural, la ciencia social computacional  
y las humanidades digitales*

La analítica cultural guarda paralelismos con dos disciplinas que también emergen en la segunda mitad de la década de 2000 y que aprovechan la disponibilidad de datos sobre el comportamiento de las personas *online* o grandes colecciones de artefactos históricos digitalizados. La primera es la «ciencia social computacional» antes mencionada y la segunda son las «humanidades digitales». La ciencia social computacional utiliza datos a gran escala de las redes sociales para abordar cuestiones formuladas en sociología, economía y otras ciencias sociales. Aquí, el ámbito de estudio es la sociedad contemporánea como «sociedades digitales», es decir, los comportamientos de las personas en las redes sociales. Los investigadores en humanidades digitales realizan análisis cuantitativos de objetos culturales creados por autores profesionales, por ejemplo, los escritores de ficción del siglo XIX. Esto es consecuente con las humanidades académicas del siglo XX que han estudiado y enseñado los cánones de las obras «importantes» e ignorado a la cultura popular. Por lo tanto, si la primera estudia exclusivamente el comportamiento *online* contemporáneo de miles de millones de personas corrientes, estas últimas estudian los productos creados en el pasado por un pequeño número de creadores profesionales de períodos históricos anteriores.

La analítica cultural no tiene tales limitaciones. ¿Por qué? Porque es una caja de herramientas de métodos y técnicas ya existentes (y aún por inventar), en lugar de ser una disciplina académica propiamente dicha. Por lo tanto, no está limitada por las tradicio-

nes y los objetivos de las humanidades, las ciencias sociales o la informática. Como ya demuestran muchos estudios, proyectos e interfaces de analítica cultural, es igualmente interesante el uso de técnicas computacionales, matemáticas y de visualización con los datos que representan el presente y el pasado, los comportamientos culturales y de los objetos, las actividades de los creadores ocasionales y las de los profesionales.

Así, la analítica cultural cubre un territorio más amplio que las humanidades digitales porque incluye el análisis de la cultura digital contemporánea creada por millones de creadores no profesionales. Otra diferencia clave es que la comprensión de la cultura en la analítica cultural incluye la perspectiva de la antropología, al contrario de lo que hacen las humanidades sólo académicas. En esta perspectiva, la cultura es una parte esencial de toda sociedad humana en todos los períodos históricos, frente a la reducción de considerar sólo lo mejor de la humanidad creado (o seleccionado) por las élites. Aunque en las últimas décadas las humanidades académicas han comenzado a ir más allá de esta visión estrecha de la cultura formulada por Matthew Arnold en 1869, en la práctica esto significa a menudo sustituir algunos textos y autores del canon por otros, siempre y cuando se preserve su pequeño tamaño.

Pero la analítica cultural tampoco puede reducirse a la ciencia social computacional, aunque a menudo utilicen los mismos tipos de datos en sus estudios. El enfoque de la analítica cultural se centra en el estudio de las formas humanas simbólicas y expresivas, y no sólo en el de los patrones sociales. El arte visual, la música, la danza, la arquitectura, el mito, el lenguaje, los adornos, los símbolos y el uso del ritmo, surgieron mucho antes en la evolución humana que los mercados económicos, las organizaciones burocráticas, las prisiones, los hospitales, las carreras profesionales y otros fenómenos estudiados por los sociólogos modernos. De hecho, sólo cuando, a mediados del siglo XIX, las ciudades alcan-



zan cierto tamaño y el individuo se separa completamente (o «se aliena») de otros debido al debilitamiento de la religión y otros vínculos, la «sociedad» como un objeto separado de estudio se vuelve posible. Por lo tanto, si sólo vamos a utilizar los grandes datos de las culturas digitales en red globales para estudiar lo social más que lo cultural, nos negaremos la oportunidad de entender la humanidad en la contemporaneidad. Nos negaremos a entender lo que nos hace humanos hoy en día frente al pasado, a las formas y los patrones específicos de la cultura humana, mediados y modelados por plataformas digitales, redes e interfaces digitales.

L. M.

Traducción: *Óscar Gómez*

---

Este texto se enmarca dentro del Proyecto I+D+I «El periodista como historiador del presente. Análisis del documento en las nuevas formas de la información» (Ref.: CSO2014-55527-P), del Ministerio de Economía y Competitividad.