

Analítica Cultural 101

Una introducción a los estudios de la cultura en la era de los datos masivos

Sofía Alamo y Gino Cingolani

1. Introducción

El pasaje del paradigma de información escasa al paradigma de información abundante cambia las reglas del juego de los investigadores de la cultura, y no únicamente en términos de acceso. En el pasado los investigadores se veían obligados a trasladarse hasta las bibliotecas o bien a hacerse con el objeto cultural en soporte material o una representación más o menos fidedigna del mismo para poder estudiarlo. Hoy internet y los medios sociales nos garantizan acceso y conectividad pero también nos traen nuevos problemas: pareciera ser que el desafío actual es entender cómo hacer un recorte significativo de toda esa masa de producciones culturales que dan vueltas a lo largo y a lo ancho de todo el planeta para responder preguntas.

Actualmente son más de 3 mil millones los usuarios de Internet, se envían más de 200 mil millones de e-mails, se realizan casi 4 mil millones de búsquedas en Google, se twitteen 600 millones de tweets, se miran 10 mil millones de videos en YouTube, se suben 60 millones de fotos a Instagram y Facebook, el país más poblado del mundo, tiene 1600 millones de usuarios.¹ A cada paso, con nuestros dispositivos móviles, desde nuestras computadoras en nuestra casa y nuestro trabajo generamos datos sobre nuestra ubicación, nuestras películas favoritas en Netflix, y canciones más escuchadas en Spotify.

¹ Información consultada de <http://www.internetlivestats.com/>, última revisión: martes 26 de marzo de 2016.

Nótese que no es sólo información lo que corre por estos medios: también es producción cultural. Estos datos nos hablan de quiénes somos como sociedad. Y allí es donde vamos a poner el acento para definir ese abordaje. Nos interesará indagar de qué manera esta nueva dimensión social que es el software da forma a nuestro mundo, cómo se produce conocimiento a partir de estas ingentes cantidades de datos, de qué forma las interfaces y plataformas digitales redefinen y reinventan la estética mediática contemporánea. Las grandes preguntas que nos vamos a hacer estarán relacionadas con comprender las nuevas formas que adoptan las tecnologías de control, comunicación, representación, interacción y poder al ser atravesadas por la dimensión del software.

Hasta ahora, el estudio de lo social y lo cultural dependía de dos tipos de datos: *datos superficiales* sobre muchos (abordados por ciencias como la sociología a través de recortes estadísticos) y *datos profundos* sobre pocos (objeto de estudio de la psicología y la antropología a partir de metodologías como la descripción densa y la lectura cercana).

El mencionado ascenso de los medios sociales, junto con las herramientas computacionales que pueden procesar velozmente cantidades masivas de datos, hacen posible, fundamentalmente, un nuevo acercamiento para el estudio de los seres humanos y la sociedad. Una de las variantes, enmarcada dentro de las llamadas humanidades digitales es la analítica cultural. La propuesta de esta perspectiva metodológica es estudiar las prácticas tecnoculturales de los usuarios de medios sociales pero sin tener que renunciar a la escala de las prácticas. Lo que busca son recurrencias y repeticiones de elementos, denominadas también patrones, que delineen una tendencia factible de ser interpretada (cfr. Manovich, 2009)

2. Campo – definición

Para hablar de analítica cultural es importante destacar dos movimientos fundamentales en la historia técnica de la humanidad: por un lado, la posibilidad de digitalizar productos culturales (es decir, convertir registros y producciones analógicas a un código binario que nos permite hacer procesamiento y cálculo) y por otro lado, la capacidad de resguardar esta información de forma masiva

a partir de la baja de los costos de almacenamiento de archivos para luego poder analizarla y utilizarla. Estos dos movimientos permiten realizar un proceso minucioso, preciso y en tiempo real de algo que antes solo era posible de manera muy rudimentaria a través de etnografías y encuestas sistematizadas, comprender fenómenos a partir de las ingentes cantidades de información que se generan en medios sociales (Tanto tradicionales como digitales).

A esto es lo que llamamos “big data” y no es notable únicamente debido a su tamaño en términos de cantidad de datos sino por cómo éstos están relacionados entre sí. Gracias a ello, ingenieros en sistemas, investigadores y científicos de la computación se ocuparon de modernizar las herramientas para poder analizarlos, dando así origen a las nuevas tecnologías de procesamiento y análisis de la información. ¿De qué modo estas herramientas de procesamiento y visualización de la información nos permitirán construir herramientas teórico metodológicas que nos permitan comprender la cultura global y conectada del siglo XXI?

La analítica cultural es uno de los abordajes que se encuentra bajo el campo de las Humanidades Digitales y los Estudios del Software. Podemos definirla como el uso de métodos computacionales para el estudio de grandes cantidades de datos culturalmente significativos. Es decir, por un lado implica estudiar la dimensión que el software introduce en nuestra cultura a través de algoritmos, plataformas e interfaces y cómo media en los abordajes de las humanidades, la crítica cultural y las ciencias sociales; y a su vez apunta a la investigación y desarrollo de nuevos métodos para la creación de recursos que nos permitan analizar esas prácticas y hacernos nuevas preguntas.

Tanto las Humanidades Digitales como los Estudios del Software se vienen desarrollando hace más de veinte años y no sólo se dedican al estudio de productos culturales de los medios sociales. Una gran heterogeneidad de estudios se enmarcan dentro de estas perspectivas, algunas de ellas hacen más énfasis en los procesos históricos de las herramientas computacionales, mientras que otros se centran en discusiones teóricas sobre las modificaciones en los métodos de investigación a partir de la utilización de estas tecnologías.

3. Referente principal: Lev Manovich

Lev Manovich dirige la [Software Studies Initiative](#), un laboratorio de investigación y diseño donde despliega investigaciones y proyectos bajo esta perspectiva. Gracias a los aportes de un equipo transdisciplinario de colaboradores con competencias variadas (análisis de datos, visualización de información, diseño web, diseño de medios, historia del arte, ciencias de la computación y estudios del software), combinan métodos de investigación y tecnologías que vienen de estos diferentes campos para realizar proyectos que van desde la publicación de papers académicos, pasando por la visualización de grandes cantidades de datos, instalaciones interactivas hasta el desarrollo de software de código abierto para llevarlas a cabo. Tiene dos sedes: una en la City University de Nueva York (CUNY) y otra en la Universidad de Los Angeles, California (UCLA).

Entre los productos culturales que analizaron se encuentran las tapas de la revista Times (1923-2009); las pinturas de Vincent van Gogh, Piet Mondrian, y Mark Rothko; 20,000 fotografías del Museo de Arte Moderno de Nueva York (MoMA); un millón de páginas de publicaciones de manga; la comparación visual de más de dos millones de imágenes de Instagram de 13 ciudades, que les permitió llevar a cabo diferentes proyectos como [Phototrails](#), [Lo excepcional y lo cotidiano: 144 horas en Kyiv](#) y [On Broadway](#), entre otros; 30 mil piezas subidas a la comunidad online artística DeviantArt. Actualmente se encuentran analizando un data set de 265 millones de imágenes twitteadas durante el 2011 y el 2014. Realizaremos una descripción de dos de los mencionados proyectos con el fin de dar a conocer el proceso de análisis que estos investigadores proponen.

Un clásico ejemplo de sus investigaciones es el análisis que realizaron de 883 publicaciones de manga (historietas japonesas o historietas que siguen el estilo de las historietas japonesas) que sumaban un total de 1.074.790 de páginas. Este corpus de imágenes fue descargado masivamente de One Manga, un sitio de que reunía una gran cantidad de mangas digitalizados y traducidos al inglés. La pregunta general que motivó este estudio fue la posibilidad de identificar la variabilidad de los fenómenos, artefactos y procesos propios de la cultura visual. Este tipo de reflexiones nos invitan a pensar si la cultura producida de manera independiente o artesanal es más diversa que la industrial, si se ven

afectadas por la irrupción de las herramientas digitales y si es cierto que la globalización disminuye esa diversidad. En particular, querían analizar cuáles eran las variantes estilísticas, de género, de audiencia, según su país de origen, utilizando las etiquetas que los mismos fans habían atribuido a cada uno de los títulos descargados.

A través de softwares desarrollados por el mismo laboratorio, lograron procesar esas páginas convertidas en imágenes que fueron mensuradas y luego transformadas en una gran base de datos posible de ser visualizada. Cabe destacar que mientras la mayoría de los softwares de visualización representan la información a través objetos como puntos, líneas, barras, el software desarrollado por este laboratorio permite realizar visualizaciones mostrando las imágenes propias del corpus permitiendo en simultáneo explorar tanto los metadatos (fechas de carga, filtros usados, las coordenadas espaciales) y los patrones creados por su contenido.

Midiendo las características propias de las imágenes como colores, saturación, brillo, tamaño, etc., identificaron patrones de repetición y variación, concluyendo que si bien el universo de los manga es heterogéneo, algunas variables se ven más presentes que otras y en ningún caso la variabilidad es infinita. Quizás el aprendizaje más rico que arrojó esta investigación fue el carácter descriptivo de la variación de géneros, estilos y su relación con la audiencia a la que estaban destinadas a partir del uso de los metadatos generados por los usuarios que subieron los mangas al sitio. Nuevamente, los criterios de la categorización cambiaron, en este caso las categorías no fueron determinadas por un software (como sucedió al procesar las imágenes), sino que fueron establecidas por los fans sin seguir una taxonomía previa, sino inventando una *folksonomía* propia.

Por otro lado, el proyecto *Phototrails* fue uno de los más ambiciosos y quizás donde pudieron plasmar con mayor detenimiento su metodología y su preocupación por el nuevo paradigma de investigación que estaban configurando. El trabajo estuvo centrado en describir ese proceso y comprender de qué forma la interfaz de una plataforma particular de medios sociales, como es Instagram, media entre las experiencias de los usuarios a la hora de producir, compartir e interactuar con el contenido que ellos mismos generan.

En la investigación que titularon “Haciendo zoom en una ciudad Instagram”,

examinaron esas cuestiones a través del análisis de la popular aplicación. En primer lugar, investigaron las *affordances*² (*Término en inglés que define aquellas características del software que nos permiten interactuar con él*) proporcionadas por la interfaz de Instagram y las formas en que esta y sus herramientas estructuran la comprensión y el uso de los usuarios sobre el “medio Instagram”. Luego, compararon las identidades visuales de 13 ciudades del mundo usando 2,3 millones de fotos de Instagram de esas ciudades. Realizaron visualizaciones espacio-temporales, de tonalidad de color y de contenido para mostrar cómo estas unidades de cultura pueden ofrecer acercamientos sociales, políticos y culturales a las actividades de las personas en determinados lugares y períodos de tiempo.

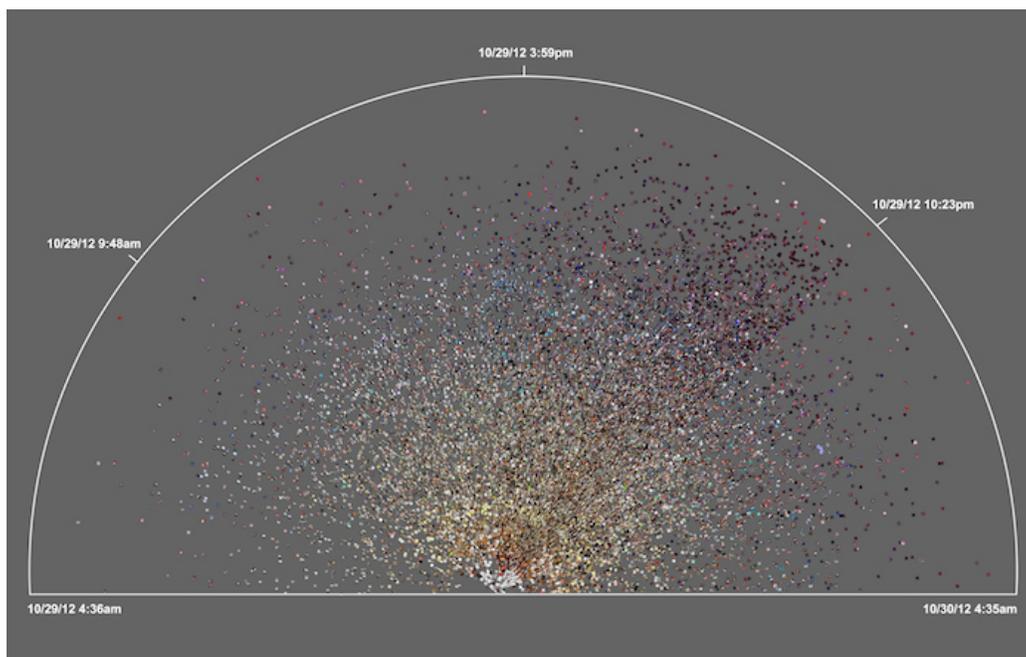


Fig. 1: Una visualización radial de puntos mostrando 23.581 fotografías subidas a Instagram en el área de Brooklyn, New York durante el huracán Sandy (29-30 de Noviembre de 2012). La distancia de la fotografía en relación al centro (radio) corresponde con su tonalidad de color promedio, el ángulo de la fotografía (Su posición alrededor del perímetro) corresponde con el momento de subida. Nótese la línea demarcatoria que revela el momento en el que se corta el suministro eléctrico de la región e indica la intensidad de la experiencia compartida (Un descenso dramático de la cantidad de fotografías y sus colores más oscuros a la derecha de la línea) Versión en alta definición en: phototrails.net. (Manovich 2013, Figura 5)

² En 1988 Donald Norman utilizó el término para referirse a esas posibilidades de acción que son inmediatamente percibidas por el usuario.

4. Algunas críticas

Si bien a esta perspectiva metodológica se la ha criticado por considerar sus interpretaciones superficiales (en parte por la relación que existe entre el procesamiento y visualización de datos y la propuesta de “lectura distante” de Franco Moretti) cabe destacar que en verdad, lo que se está señalando es justamente la necesaria combinación del “estudio de patrones” con “lecturas cercanas” de actores particulares.

De esta manera, la herramienta principal de esta perspectiva es, entonces, una lectura multi esalar. Para llevarla a cabo, se utilizan técnicas especiales de visualización (gráficos radiales de puntos y montajes de imágenes) que muestran todas las imágenes en grandes conjuntos organizados por metadatos y/o propiedades visuales. A su vez, se continúa mirando muy detalladamente a las interfaces, herramientas y alcances del software (como fue en el caso de Instagram) que permiten la práctica de los medios sociales (Manovich y Hochman, 2013).

Debemos señalar, sin embargo, que muchos de los estudios realizados a partir de esta línea hicieron mayor foco en el desarrollo de software para el procesamiento y análisis de información que para responder a las grandes preguntas que nos fueron prometidas. Es nuestro desafío como investigadores indagar sobre la posibilidad de alcanzar una mayor profundidad para poder complejizar el campo.

Para ello, debemos superar los enfoques dualistas y maniqueos que separan al mundo científico anglosajón (cuantitativo, empirista, poco crítico) del entorno académico latinoamericano (cualitativo, especulativo, a veces hiper-crítico). Sin necesidad de caer la disolución de las fronteras en el marco de la globalización, Jesús Martín Barbero nos propone aprender a hablar el lenguaje de la ciencia anglosajona sin renunciar a una visión propia (Scolari, 2010).

5. Conclusión

Lo más relevante que introduce la analítica cultural en tanto perspectiva metodológica es la apuesta crecientemente e intensificada hacia el trabajo co-

laborativo que se encuentra en las antípodas de la investigación tradicional: centrada en un tema, en una cabeza, en una tradición de pensamiento y en una carrera profesional individual (Piscitelli, 2012).

Al trabajar colaborativamente el académico digital se ve obligado a pensar los estudios de manera transdisciplinaria con el fin de apropiarse la web como parte constitutiva de cualquier emprendimiento y poner en cuestionamiento algo impensado para la academia tradicional que es pensar/hacer en épocas de la web masiva.

A esto es a lo que se refiere Berry (2011) cuando afirma que el giro computacional, digital y masivo pone de manifiesto la forma en que los medios producen cambios epistémicos.

Referencias:

BERRY, D., (2011): "Digital Humanities: First, Second and Third Waves" en catedradatos.com.ar. Traducción de Alonso e Ibañez para la Cátedra de Datos Piscitelli, Carrera de Ciencias de la Comunicación, Facultad de Ciencias Sociales, Universidad de Buenos Aires.

MANOVICH, L. (2009). "Cultural Analytics: Visualizing Cultural Patterns in the Era of 'More Media'". Recuperado de: manovich.net.

MANOVICH, L. (2012): "From Readers to Pattern Recognition", enero 2012, recuperado de: thecreatorsproject.vice.com.

MANOVICH, L., NADAV, H. (2013): "Zooming into an Instagram City", First Monday, 1 de julio de 2013. Recuperado de: firstmonday.org.

MANOVICH, L., NADAV, H. *Software Studies Initiative* (2016). Recuperado de: lab.softwarestudies.com.

PISCITELLI, A. (2012) "Las "Digital Humanities" y como pensamos en la era de la analítica cultural". Recuperado de: www.filosofitis.com.ar.

SCOLARI, C. (2010): "Jesús Martín Barbero entre la ecología y las interfaces". Recuperado de: hipermediaciones.com.