

#07

MATÍAS VALDERRAMA BARRAGÁN

Pontificia Universidad Católica de Chile
Estudiante del Magíster en Sociología
Santiago, Chile

mbvalder@uc.cl

LOS RASTROS DIGITALES DE LO SOCIAL: HACIA UNA AGENDA DE ESTUDIOS DIGI- TALES LATINOAMERICANOS

Digital traces of the social: Towards an agenda of Latin American Digital Studies

Resumen

Las acciones de innumerables actores no sólo están siendo crecientemente mediadas por tecnologías y medios digitales, sino que además son inscritas en largas bases de datos que con métodos tradicionales sería dificultoso de manejar. Estos rastros digitales despliegan lo social en una nueva escala y están siendo recientemente reorientados desde las ciencias sociales. El presente artículo se propone dar breve revista a la literatura en torno a los denominados Big Data y métodos digitales, temas poco explorado hasta el momento en la región y que pueden constituirse como una valiosa fuente de información para la investigación social. Sin embargo, se destaca el peligro de terminar naturalizando la información digital, sobre todo ante la promesa de una mayor automatización y predicción. Desde un enfoque relacional se problematiza que detrás de la data digital existe una vida social con múltiples y heterogéneas entidades interfiriendo con diversos intereses. Esto nos compele a abrir los códigos y afirmar constantemente las limitaciones y “sesgos digitales” que se generan al analizar estos datos, pues darán cuenta de dinámicas sociales y asimetrías de poder empíricamente relevantes. En particular, examinamos a modo de ejemplo lo efímero que puede ser la Web, constituyéndose como un archivo digital dinámico siempre parcial y limitado. Producto de esta limitación, los estudios digitales se han restringido a temporalidades acotadas, cuando la incompletitud y falta de archivación de la Web deben ser problematizadas. Temas como éste, así como las nuevas brechas digitales esbozadas al final del artículo, demandan ser estudiadas críticamente desde Latinoamérica y el Sur global.

Palabras claves

Big Data, métodos digitales, Archivo, sesgo digital, arqueología de medios

Abstract

Digital traces of the social: Towards an agenda of Latin American Digital Studies

The actions of many actors are not only being increasingly mediated by digital technologies, but are also inscribed in long databases that using traditional methods would be difficult to handle. These digital traces display the social in a new scale and are being recently reoriented from the social sciences. This article aims to briefly review the literature on the so-called Big Data and digital methods, themes little explored so far in the region and that can become a valuable source of information for social research. However, the paper emphasizes the dangers of naturalizing digital information, especially by the promise of greater automation and prediction. From a relational approach, it is problematize that behind the digital data there is a social life with multiple and heterogeneous entities interfering with diverse interests. This compels us to constantly open the codes and affirm the limitations and “digital biases” that are generated by analyzing this data, because they will manifest empirically relevant social dynamics and asymmetries of power. In particular, we examine for example the ephemeral it may be the Web, becoming a dynamic, partial and limited digital archive. Because of this ephemerality, the digital studies focused on short temporalities and tried to reduce the digital bias, instead of problematize this situation. Questions like these or the new possible digital gaps outlined in the end, have to be studied critically from the South.

Keywords

Big Data, Digital Methods, Archive, digital bias, Media archeology

01. La ubicuidad de lo digital²

Algo evidente es que la sociedad se encuentra cada vez más mediada por múltiples aparatos digitales. Hacer las compras en el supermercado, vitrinear en tiendas online, ordenar una pizza a domicilio, apagar las luces de una casa a cientos de kilómetros, contar el número de pasos al día, pagar la patente del automóvil, comprar pasajes en avión y luego hacer su *check-in*, salir a andar en bicicleta, subirse al transporte público, etc. todas estas acciones mundanas están siendo crecientemente digitalizadas pasando a ser mediadas por *smartphones*, sensores, geolocalizadores, apps, tarjetas, entre muchos otros aparatos. Es tal el grado de penetración de los nuevos medios y tecnologías digitales que Mike Featherstone (2009) las ha denominado como “medios ubicuos” pues adquieren una existencia fluida y casi omnipresente dentro de la sociedad, la ciudad y la memoria, que los diferencia de los medios de comunicación de masas de carácter más monolíticos.

Pero estos dispositivos no solo median las interacciones en todo momento y en todo lugar, no solo conectan heterogéneos actores según determinados comandos, estrategias y algoritmos; sino que quizás la gran diferencia es que también registran e inscriben esas acciones en largas bases de datos. Ya sea por la decisión explícita de los usuarios involucrados o por la experiencia propia del aparato de automáticamente sentir y

registrar los eventos o intercambios de información, se producen grandes estelas o migas de datos de variadas interacciones diarias. De modo que la ubicuidad de los medios digitales crea un nuevo ambiente en que cada cosa³ alrededor nuestro estaría constantemente transmitiendo y coleccionando información en tiempo real sobre nuestros movimientos, hábitos, preferencias, etc. (Featherstone, 2009). La preponderancia llega a tal, que ya no solo los humanos, sino que los propios aparatos y softwares digitales se constituyen cada vez más en verdaderos arqueólogos del conocimiento (Ernst, 2012), pues van seleccionando qué información, en qué formato y en qué lenguaje guardarlas y archivarlas. Así, mediadores digitales como bots o rastreadores, se manifiestan y entran en escena definiendo qué página debe ir primero en una búsqueda en Google o cómo presentar un hito histórico en Wikipedia o qué temas debieran ser un *trending topic* en Twitter.

Agradecer

En suma, las crecientes interacciones mediadas y registradas van produciendo una verdadera digitalización de la vida social (Marres, 2012) o como le ha llamado más formalmente Robert Latham y Saskia Sassen (2005), una “socio-digitación” en que no solo se convierten técnicamente contenidos análogos a bits de información, sino que más profundamente ciertas lógicas y realidades sociales en lugar de otras, pasan a adquirir

2 El autor desea agradecer los pares evaluadores de la revista por sus comentarios a una primera versión de este artículo. Asimismo, agradecer a las entrañables personas que han ayudado a mejorar y discutir las ideas aquí planteadas. Por último, desea agradecer al Instituto de Sociología por brindar financiamiento parcial en sus estudios de Magíster y al proyecto Fondecyt N° 11140042.

3 Con la llegada del denominado “Internet de las cosas”, inicialmente desarrollado para las tecnologías del hogar pero actualmente siendo adoptado en múltiples lugares como en procesos de industrias, hospitales y transportes, solo hacen que la proyección de la generación y recolección de datos siga creciendo en extremo.

una forma digital que pueden reflejar así como también transformar y reconfigurar lo social. En el proceso, se acumulan grandes cantidades de información sobre la sociedad nunca antes vistas y difíciles de poder procesar con métodos tradicionales, en lo que se ha llamado por Big Data. Parte de la maraña viva, ese movimiento de lo social, pasa a ser inscrito en casi infinitas cadenas de 1s y 0s que generan un particular archivo digital de la realidad social que privilegia ciertas ontologías y oscurece otras a la vez (Bowker, 2014), y que demandan ser estudiadas desde las ciencias sociales de manera urgente.

Esta ha sido la tónica en los últimos años en la academia internacional. Connotadas revistas internacionales han sacado números especiales dedicados a los posibles alcances del Big Data como *First Monday* (Helles y Jensen, 2013), *Policy and Internet* (Margetts y Sutcliffe, 2013), *Journal of Communications* (Parks, 2014), *International Journal of Communication* (Crawford, Miltner y Gray, 2014), e incluso se han creado nuevas revistas dedicadas a temas sobre datos digitales masivos como la revista *Big Data y Society*. Pese a ello, hasta el momento Latinoamérica ha quedado principalmente como un espectador de los acelerados cambios tecnológicos (Trejo, 2013). De hace ya décadas, la discusión en el contexto latinoamericano ha estado dominada por las superación de las brechas digitales de acceso y la desigual conectividad en la región, estudiando las tecnologías digitales como objetos de estudios, dejando desatendidas

los datos que estas producen y las prácticas que ciertos actores están llevando a cabo a la hora de analizarlas y procesarlas. Aun cuando nuestro país ha tenido una acelerada y alta penetración de Internet y presenta un uso intensivo de las redes sociales, la academia aún no se ha volcado a estudiar académicamente los rastros que quedan inscritos digitalmente. Salvo algunos promisorios proyectos en curso y recientes publicaciones, que se han centrado principalmente en análisis de Twitter y la opinión pública (González-Bustamante, 2015; Hilbert et al. 2016; Sajuria y Fábrega, 2016), no se le ha prestado gran importancia desde el ámbito académico al relevante fenómeno social de la producción y re-orientación de datos digitales en Chile.

De este modo, el siguiente artículo busca introducir al lector en torno a las teorizaciones y aplicaciones en las ciencias sociales del denominado Big Data y los métodos digitales. En primer lugar, sin ánimos de ser exhaustivo, se revisan las principales discusiones en torno a este hiperbolizado “diluvio” de datos masivos y cómo estos están siendo re-orientados con novedosos fines investigativos sumamente relevantes para las Ciencias Sociales así como de otros más cuestionables en términos éticos. En segundo lugar, se problematizan algunos de los peligros que se generan al naturalizar la data digital, sugiriendo una aproximación desde un enfoque relacional inspirado fundamentalmente en la teoría del Actor-Red de Bruno Latour (2005), en que se abran los

procesos de producción y procesamiento de estos datos y afirme sus sesgos y limitaciones constantemente, antes que ponerlos entre paréntesis (Marres, 2015). Posteriormente, nos detendremos en enfatizar un particular problema en los estudios digitales asociado al carácter parcial e histórico de estos datos bajo la consideración de un archivo digital. Por último, el artículo de manera general busca alentar a los investigadores de nuestra región a experimentar y contextualizar estos nuevos registros de información, convocando una agenda de estudios digitales desde Latinoamérica, que coloque en crítica las invisibilizaciones y parcialidades históricas que emergen al revisar la vida social de este tipo de datos y sitúe en la discusión las crecientes disparidades entre el Norte y el Sur Global en el procesamiento y análisis de estos datos.

Re-orientando los trazabilidad digital de lo social

Ante la creciente digitalización de los procesos y actividades de todo tipo, se incrementan las oportunidades para las empresas, gobiernos y ciudadanos para redirigir y usar los datos digitales generados en tiempo real para nuevos propósitos e intereses. En particular abre un horizonte rico en oportunidades investigativas pero también variados peligros y problemas éticos para las ciencias sociales que sugieren un gran desafío.

La “Bendición” digital

En 1998 Bruno Latour planteaba en una conferencia de la *Virtual Society?* que con el avance de las tecnologías

digitales se iba a poder abandonar los modelos y abstracciones que nos representan a la sociedad en una escala diferente a la real, para pasar a rastrear la sociedad en su misma dimensión, en una escala 1:1. Los niveles de rastreabilidad incrementarían a tal nivel que por fin íbamos a poder ver como el orden social emerge y se va haciendo en su conjunto (Latour, 1998: 10). Latour insistiría en este punto en *Reensamblar lo Social*, planteando que mientras más se extienda una infraestructura material como los dispositivos inteligentes, computadoras, softwares, etc. más sensible y preciso será nuestro monitoreo de las controversias y por ende, más posible el seguimiento cabal de asociaciones y ensamblajes. En una palabra, la Web al expandirse cada vez más puede ser “convertida en laboratorio mundial” (Latour, 2005: 172).

Estas afirmaciones las desarrollaría posteriormente con Tomasso Venturini en el medialab de Sciences Po, escribiendo una prolífica literatura sobre los rastros digitales (Venturini y Latour, 2010; Venturini, 2010; Latour et al., 2012; Venturini, Jensen y Latour, 2015). Cualquier relación que sea mediada por las tecnologías digitales, toda cosa dicha o hecha o incluso imaginada, dejaría automáticamente tras de sí una huella digital que estará disponible para seguir los hilos de la existencia colectiva. Allí radicaría quizás la promesa más destacable de los medios digitales e Internet, su capacidad de rastrear o hacer trazable lo social a una nueva escala mediante el seguimiento de sus huellas digitales. Como escriben de manera explícita:

Estar interesados en la construcción e fenómenos sociales implica el seguimiento de cada actor involucrado y de cada una de las interacciones entre ellos. Esto era una meta imposible hace solo una docena de años atrás, pero comienza a ser más y más realista al extenderse las tecnologías digitales. [...] El interés en los medios electrónicos radica en el hecho de que cada interacción que pasa a través de ellos deja un rastro que puede ser fácilmente grabado, masivamente almacenado y recuperado a bajo costo. [...] Digital mediation spreads out like a giant roll of carbon paper, offering the social sciences more data than they ever dreamt of. (Venturini y Latour, 2010: 6, cursivas mías).

Para estos autores el diluvio de datos no sería una promesa exagerada sino una verdadera “bendición” actualmente realizable, estaríamos entrando en una era de la abundancia de observaciones y rastros de lo social. Y con este incremento de la trazabilidad de lo social, las ciencias sociales ya no tendrán que seguir añejas divisiones entre estudiar las particulares situaciones micro-sociales o lo macro-estructural, sino que podrá hallar un continuo entre los dos niveles y de manera simultánea. Si antes las ciencias sociales se encontraban obligadas a simular modelos y solventar altos costos para estudiar los fenómenos sociales, ahora con la avalancha de datos digitales podrá aliarse a las ciencias naturales experimentales de manera más fuerte sin nada que envidiarle,

creando nuevas metodologías cuali/cuantitativas inclusive (Venturini, Jensen y Latour, 2015). Por fin se había dado con la manera de recolectar la cantidad y la calidad de información necesaria para romper las tradicionales divisiones y limitaciones de la ciencia social, permitiendo re-visitar viejas teorías y materializar análisis que antes parecía impracticable a gran escala, como los experimentos de mundo pequeño de Stanley Milgram o el estudio monadológico de Gabriel Tarde (Latour, et al. 2012; Gonzalez-Bailon, 2013). Como ha escrito el propio Venturini:

Las interacciones electrónicas se han vuelto tan omnipresentes que ya no pueden ser concebidas como un espacio social independiente. No están más limitadas a un sector en específico, las interacciones digitales están ahora entretrejidas en la tela de la existencia colectiva. Siga los hilos digitales y se desplegará el tejido social (Venturini, 2010: 5).

Nuevos métodos a nuevos datos

Este diluvio de datos nacidos de medios digitales (digitally born) requerirían de nuevas metodologías que re-orienten el sentido de los datos digitales. Los denominados métodos digitales, se pueden enmarcar dentro de este proceso en las ciencias sociales por métodos más abiertos y sensibles a la vitalidad de lo social. Internet, el Big Data, las apps, los tweets, los hipervínculos, los

“me gusta” en Facebook, etc. pueden ser estudiadas no tan solo como cultura online sino como una fuente de información sobre la sociedad en su conjunto, tanto online como offline.

En la literatura se ha concebido de varias maneras estos métodos digitales: una forma bien amplia los entiende como “el uso de tecnologías en línea y digitales para recolectar y analizar datos” (Snee et al. 2015, p. 1). Para Severo y Venturini, los métodos digitales son una “serie de técnicas diseñadas para explorar los rastros de interacciones online como fuente de información de fenómenos sociales” (2015, p. 3). Para otros, más en específico, los métodos digitales excluirían la digitalización de métodos pre-existentes⁴ y serían más bien la reorientación de los dispositivos y plataformas online para la investigación social y política que de otro modo sería *improbable* de lograr debido a su alta automatización (Rogers, 2015). En cualquiera de estas definiciones, diseñar herramientas y softwares para capturar lo inscrito en la Web nos abre un horizonte rico en posibilidades investigativas dentro de las ciencias sociales.

Todo el programa investigativo de los métodos digitales, surge en primer lugar desde la Universidad de Ámsterdam, en la creativa *Digital Methods Initiative* bajo el mando del académico Richard Rogers. Según este autor, dada la creciente evolución de las tecnologías digitales y el registro de grandes cantidades de infor-

mación, se debería apuntar a una reformulación de las preguntas y métodos de investigación de Internet (Rogers, 2013). Siguiendo los planteamientos de la naciente antropología digital británica (Miller y Slater, 2000; Miller y Horst, 2012), se debe dejar atrás las añejas nociones del ciberespacio, como un lugar apartado de la realidad; y posteriormente apuntar a estudiar Internet en sus propios términos.

Me gustaría ayudar a definir una nueva era en la investigación de Internet, una que no se preocupe más de la brecha entre lo real y lo virtual. Su inquietud cambia a los tipos de preguntas por hacer en el estudio de Internet. La Internet se emplea como un sitio de investigación para mucho más que simplemente la cultura en línea. La cuestión ya no es qué parte de la sociedad y la cultura está en línea, sino más bien en cómo diagnosticar el cambio cultural y condiciones de la sociedad a través de Internet. El punto de partida conceptual es el reconocimiento de que la Internet no es sólo un objeto de estudio, sino también una fuente (Rogers, 2013: 21).

Y desde esta nueva concepción de Internet como una fuente de información de la sociedad, se han creado códigos para recopilar información de diferentes apps, redes sociales, buscadores, etc. para analizar temas tan variados desde el envejecimiento en Europa (Rogers et al., 2015),

4 Si bien los métodos digitales se alejan hasta cierto punto de la entrevista y la encuesta tradicional (Savage, 2013), no anulan la posibilidad de un componente humano o manual en su ejecución. La distinción entre métodos nativamente digitales y los métodos digitalizados de Rogers (2014) plantea una distancia de un uso más general del concepto en este sentido, que paradójicamente podría replicar divisiones pasadas.

pasando por el calentamiento global bajo el proyecto EMAPS (Venturini et al., 2014) hasta la organización de entidades de resguardo del patrimonio cultural (Severo y Venturini, 2014), entre muchos otros temas.

Pero esta re-orientación de los medios digitales desbordó con creces las pretensiones iniciales y las posibilidades de estudiar lo social mediante el análisis de variados registros digitales siguen incrementando cada día de manera exponencial, encontrándose interesantes aplicaciones en variadas disciplinas. Desde la Economía, autores han encontrado que el incremento del número de visitas a las páginas de Wikipedia de compañías o tópicos económicos puede llegar a anticipar el comportamiento económico de las personas y en particular las caídas en la bolsa (Moat et al., 2013). Otros, mediante la recopilación de datos de sitios de redes sociales, han hallado interesantes conclusiones respecto de los flujos comunicacionales de movimientos sociales a través de Twitter (Hilbert et al., 2016). Analizando las interacciones de 10.1 millones de usuarios en Facebook, otros han encontrado burbujas informativas en que las personas seleccionarían leer aquellos contenidos de sus amigos afines a sus posturas ideológicas (Bakshy, Messig y Adamic, 2015). Desde las recientes “humanidades digitales”, cuestiones que antes parecían poco abordables para hacer análisis cuantitativos, hoy mediante computadoras se pueden realizar métodos y visualizaciones relevantes y de gran alcance. Por ejem-

plo con la digitalización de los libros de la humanidad bajo el proyecto *Ngram* de *Google*, se ha revisado la evolución del positivismo científico (Savage, 2013) o se podido analizar más de 500 billones de palabras de 5 millones de libros digitalizados, para observar cómo ha ido variando el léxico del inglés o ver la prominencia del feminismo en cierto año, entre muchas otras cosas posibles (Michel et al., 2011). Se ha podido analizar alrededor de 35 millones de fotos subidas a Flickr para construir interesantes mapas y relaciones entre sus usuarios que revelan desigualdades a nivel global (Crandall et al. 2009).

No obstante, también han surgido diversos estudios con fuertes cuestionamientos éticos, por ejemplo, han realizado un experimento social masivo manipulando las publicaciones que le aparecían a cerca de 700 mil usuarios de Facebook para analizar la capacidad de las amistades de un usuario en contagiar estados emocionales según el carácter positivo o negativo de sus publicaciones (Kramer, Guillory y Hancock, 2014). Desde el ámbito de la psicología, la huella digital que dejamos en internet para algunos autores permitiría inferir y predecir nuestra personalidad, llegando inclusive a determinar la raza u orientación sexual analizando los “me gusta” en Facebook de una persona (Kosinski et al., 2013).

De esta manera, la capacidad de recopilar grandes cantidades de información digital no solo permite la producción de más información sino que también impli-

ca un desafío epistemológico y sobre todo ético sobre cómo vamos a manejarla y analizarla así como de cuáles deberían ser sus escalas y alcances. “Re-enmarca preguntas claves en torno a la creación del conocimiento, los procesos de investigación, cómo nosotros deberíamos involucrarnos con la información, y la naturaleza y categorización de la realidad” (boyd y Crawford, 2012: 665). El estudio de los rastros digitales nos obliga a una redistribución de nuestras metodologías y experticias a través de los múltiples agentes involucrados así como también de nuestros objetos de estudio de interés (Marres, 2012). Los científicos sociales no pueden pretender estudiar este tipo de registros computacionales de manera solitaria sino que deben abrirse a la multidisciplinariedad (González-Bailon, 2013) y cada vez más nuevas agencias viven produciendo y analizando datos de la sociedad con múltiples intereses detrás sin tener una aproximación científica necesariamente.

02. El peligro de naturalizar la data digital

Ahora bien, pareciera que en un futuro más cercano que lejano, con todo el torrente del Big Data y las tecnologías de la información, será una simple cuestión de la capacidad de nuestras computadoras el estudiar todo lo social de manera mucho más objetiva y “realista” que antes. Ya no nos tendríamos que basar en las volubles respuestas de los encuestados o entrevistados acerca de lo que hacen o no en su vida y por qué. Aparentemente tendríamos acceso a la información sobre lo que hacen de manera directa e inmediata. Esto ha llegado a tal punto que han surgido actores como Chris Anderson, editor en jefe de una de las revistas más conocidas so-

bre computación como lo es la *Wired*⁵, planteando que todo el método científico y todas las teorías sociales están comenzando a quedar obsoletas ante el diluvio de la información digital. Encontrar correlaciones bastará cuando contemos con vastas bases de datos. Los días de complicados modelos explicativos serían cosa del pasado en la Era del Petabyte. Como escribe el propio Anderson en una de sus frases para el bronce:

Este es un mundo en el que las cantidades masivas de datos y matemáticas aplicadas sustituyen a cualquier otro instrumento que pueda ponerse en juego. Desactualizada de todas las teorías de la conducta humana, de la lingüística a la sociología. Olvide la taxonomía, la ontología y la psicología. ¿Quién sabe por qué las personas hacen lo que hacen? El punto es que lo hacen, y podemos realizar un seguimiento y medirlo con una fidelidad sin precedentes. Con datos suficientes, los números hablan por sí mismos (Anderson, 2007).

Como podrán esperar, tuvo un diluvio de críticas. Pero no solo de los sociólogos o psicólogos en defensa de sus disciplinas, sino por una cuestión más de fondo. Es necesario aproximarnos a los datos masivos digitales como un fenómeno socio-cultural, tecnológico y académico pero también como *mitología* (boyd y Crawford, 2012; Crawford, Miltner y Gray, 2014) como así exclamaron en su minuto los teóricos de la escuela de Frankfort ante la Ilustración y el conocimiento científico-técnico. En este caso, se alimentaría la creencia de que los nuevos datos digitales por fin generan un cono-

5 Es interesante constatar que desde *Wired* se han formado los discursos más optimistas y tecno-deterministas sobre lo virtual y el ciberespacio, y ahora nuevamente tiene un rol clave con el nacimiento de los discursos en torno al Big Data y el self-tracking.

cimiento objetivo y que posibilitará anticipar y predecir el comportamiento de las personas. Se habría caído en una arrogancia y exagerada confianza en las capacidades de predicción probabilística utilizando largas bases de datos, descartando la necesidad de explicar los fenómenos sustantivos con solo ver una alta correlación poco intuitiva entre dos variables (Andrejevic, 2014; Crawford, Miltner y Gray, 2014). La data digital sería un sustituto de la recolección y análisis de datos más tradicionales, más que un complemento. Se vuelve imperioso ser constantemente críticos ante aseveraciones bien exageradas como las de Anderson que nos revelan la posibilidad de caer en una gran naturalización de la información digital.

En primer lugar, la cuestión política-económica detrás de estos datos digitales debe ser algo continuamente atendido y revisado. Este nuevo horizonte dado por la creciente digitalización de la vida social no está exento de graves peligros (boyd y Crawford 2012; Bauman y Lyon, 2012, Morozov, 2013; Celorio, 2014 entre muchos otros) y es necesario ponerlos de relieve. Por ejemplo Crawford, Miltner y Gray (2014) nos alertan de que el Big Data seguiría dejándonos con puntos ciegos, sobre todo en lo que no produce rastros digitales. Nuevas exclusiones sociales se generarían entre quienes tienen el acceso y las habilidades digitales y los que no, teniendo implicancias cada vez mayores. Asimismo, la recopilación y uso inescrupuloso de información personal expuesta en Internet por parte de los científicos sociales (en lo que Back (2012) ha llamado por “empirismo intrusivo” por ejemplo), la creciente privatización de las bases de datos y recursos digitales, el control que

tienen las grandes compañías de los datos de sus usuarios en detrimento de las pocas herramientas que tienen estos a su vez para restringir el uso de su información, la consolidación de nuevas jerarquías basadas en nuevas brechas digitales de no solo el acceso a las nuevas tecnologías sino también de la experticia y manipulación de estas bases de datos, la profunda vigilancia a distancia que entidades gubernamentales como la NSA de Estados Unidos están ejerciendo a la sociedad civil, la arbitraria desaparición de ciertos resultados en los principales buscadores, son algunos de los temas que podríamos mencionar someramente.

Otra de las críticas más recurrentes al Big Data ha sido también por su representatividad. Las bases de datos recopiladas con las nuevas tecnologías digitales no serían necesariamente ni aleatorias ni representativas de las poblaciones estudiadas (boyd y Crawford, 2012; Tufekci, 2014; para el caso chileno véase González-Bustamante, 2015). La data digital contiene a ratos información irrelevante o redundante y adolece en muchos casos de confiabilidad, validez y diferentes formas de ruido, por lo que siempre sería necesario revisar bien la fuente, su proceso de minería de datos y el contexto en que se produjo. Otros autores (Lazer et al. 2014) han llamado la atención ante las pocas posibilidades de replicabilidad y transparencia de los datos en los estudios que están comenzando a utilizar estas fuentes de información. Pero ello podría dejarnos con la idea de que es una mera cuestión metodológica, hacía una mayor sofisticación de las técnicas computacionales, de cómo hacer más sensibles nuestros dispositivos, de refinar las bases de datos, de eliminar los bots, perfeccionar los

algoritmos, etc. en lo que algunos autores han denominado como un aproximación “precavida” a la data digital (Marres, 2015; Marres y Moats, 2015), cuando el tema no acaba ahí.

03. Una postura relacional de la data digital

Tomando en cuenta los ya reconocidos planteamientos de Bruno Latour (2005), podríamos plantear en un segundo momento, una crítica más general sobre el carácter exclusivamente técnico en que se está traduciendo al Big Data como tal. Desde la emergente *Data Science* como en el extremo de la columna de Anderson, hay una clara intencionalidad a la hora de hablar de hechos científicos. Ya no hay declaraciones, cada bit representa acciones realizadas por las personas y con grandes volúmenes de bits de información ya bastará con encontrar correlaciones. Pero como han esbozado de manera explícita boyd y Crawford (2012), las grandes cantidades de información digital no hablan por sí solas sino que requieren de un complejo procesamiento y de bastante interpretación –tanto humana como no-humana, y en ningún caso conseguiremos automáticamente hechos (*facts*), como si de la más grande mediación surgiese un resultado inmediato y objetivo; aparecerán portavoces, grupos y una serie de instrumentos que se entrecruzan e interfieren constantemente⁶. Se ha dicho en otro lugar de manera bien clara y sencilla (Gitelman, 2013) que la data digital en bruto o “en estado de pureza” (*Raw Data*) es un completo oxímoron pues toda *data* ya im-

plica algún grado de registro, procesamiento, almacenamiento o transmisión, etc. que concierne a múltiples y heterogéneos actores. Existe un enredo social detrás de cada bit de información que debe ser descrito antes que dado por sentado bajo un exagerado entusiasmo y por ello se vuelve necesario que las ciencias sociales tengan una mayor actividad y sobre todo sensibilidad ante esa vida social de la información digital (Ruppert et al., 2015).

De lo contrario terminaríamos reactualizando viejas divisiones entre una ciencia social computacional (Lazer et al., 2009) o una física social (Pentland, 2014), aparentemente más objetivas y automáticas y que tiene acceso o se las ingenia para recopilar datos; en contraste con otras disciplinas “subjetivas” que seguirían en la edad de piedra de las encuestas, entrevistas y la arbitrariedad de la declaración humana. Ya en 2007, Mike Savage y Roger Burrows nos alertaban de la futura crisis en las ciencias sociales que comportaría la generación de datos masivos propietarios. Las grandes compañías como Facebook o Twitter⁷ al restringir el acceso a sus API, dificultan su recopilación y usos, lo que conlleva a que solo aquellos que tienen el dinero, los recursos y las redes necesarias puedan procesar e interpretar esa información (boyd y Crawford, 2012; Bruns y Burgess, 2016). Lo cual a su vez genera que los científicos sociales que trabajan para esas compañías tengan una ventaja considerable respecto a los que se enclaustran en la academia, posibilitándose una verda-

6 En este sentido podemos agregar que ningún método digital es completamente “nativo digital”

7 En particular, para una excelente revisión de cómo la API y los términos de servicio de Twitter han ido precarizando la investigación pública y académicos a través de los años, véase Bruns y Burgess (2016).

dera crisis entre una sociología academicista y una de corte comercial (Savage y Burrows, 2007, 2014) y una profunda incapacidad del conocimiento positivista más tradicional con sus métodos estadísticos estandarizados⁸ de lidiar y manejar con la multiplicidad de datos que producen “vivamente” diferentes aparatos digitales (Savage, 2013).

Ante tal sobrecarga de datos con orígenes distintos a los tradicionales y donde el o la cientista social es un actor más dentro de un ecosistema de generadores de información, diversos autores han propuesto indagar en la propia vida social de nuestros métodos (*The Social Life of Methods*). Bajo esta problematización se da cuenta de que ciertos métodos como la encuesta o la entrevista se han afincado históricamente en ciertas disciplinas sin muchas veces indagar en la contingencia de tal adopción. De allí que se apunte, como decíamos anteriormente, a una ciencia social que vaya incluyendo nuevos métodos interdisciplinarios más “vivos”, experimentales y audaces que generan una redistribución de las experticias y técnicas en juego (Back, 2012; Marres, 2012; Savage, 2013; Ruppert, Savage y Law, 2013). El interés en estos nuevos métodos no estaría en sus volúmenes sino en los nuevos modos de registrar y sentir lo social (Ruppert et al., 2015), como los sonidos de la ciudad, las partículas en el aire, sentimientos y afectos, etc. y que están crecientemente posibilitados por las tecnologías digitales⁹.

Por esta razón, estos autores llaman a abrir la caja negra detrás de la producción de datos y a considerar esta producción en términos siempre relacionales. Existen relevantes prácticas sociales detrás de la generación, recopilación, interpretación y análisis de estos datos que deben ser comenzadas a ser objeto de estudio y visibilizadas:

El desafío del Big Data no es simplemente su volumen, sino que trabajar con Big Data crea nuevos problemas, riesgos y vulnerabilidades dada la tendencia a subestimar la vida social de los datos, que no es un fenómeno ni natural ni técnico, sino enactado a través de múltiples y selectivas prácticas sociales y técnicas (Ruppert et al., 2015: 1-2).

Describiendo y analizando esas prácticas socio-técnicas detrás del origen, almacenamiento y procesamiento de estos datos, se pondrán en escena variados actores y determinados intereses a ratos permanecen en una opacidad y extrañeza. Además se podrá vislumbrar previamente posibles peligros a la privacidad, extralimitaciones o usos indebidos de información personal, por ejemplo al dar cuenta sobre el diseño y decisiones políticas en el funcionar de algoritmos y aplicaciones. Sólo así se podrá justificar el análisis de datos masivos, no en términos de volúmenes técnicos y mayor eficiencia, sino en un sentido ético y ciudadano desde su concepción.

8 Se vuelve criticable, no obstante, si acaso con la minería de datos y la creciente automatización de los procesos investigativos, el empirismo en las ciencias sociales más decaer se ha reactualizado en formas aún más positivistas. Creyendo en una neutralidad de las tecnologías digitales, por ejemplo describiendo bases de datos transaccionales como acciones fácticas antes que declaradas, se puede caer fácilmente en una ingenua idea de mayor verdad y objetividad y en una extrapolación de los resultados de manera inapropiada.

9 En este sentido, se vuelve provechoso —más que el peligroso entrecruzamiento de bases de datos— el combinar métodos y complementar el hiperbolizado Big Data con el small data generada mediante investigaciones cualitativas como etnografías, entrevistas, focus groups, etc. Con esto se lograría que los datos adquieran mayor profundidad y sentido para las personas. Véase el caso de Heather Ford (2014).

04. Afirmando el carácter parcial de los rastros digitales

A continuación queremos examinar desde esta aproximación relacional de los datos digitales una particular parcialidad de los rastros digitales que muchas veces se tiende a subestimar o es pasada por alto desde la literatura que respecta a la historicidad de la web y de lo volátil que resulta ser la información online. Comúnmente se plantea que cada acción mediando dispositivos digitales quedara inscrita y almacenada automáticamente sin prestar atención a los movimientos y transformaciones que se producen en la vida social de esos datos y medios y tecnologías digitales que los producen. Dicho de otro modo, los rastros digitales no son “dejados” de manera pasiva y neutral, ni mucho menos son estáticos, son una producción social y se enmarcan en historias que son relevante de indagar. Así, el archivo digital tiene una serie de parcialidades que se pondrán de manifiesto al utilizar estos métodos digitales y que pueden limitar los alcances de esta gran “bendición”.

Como han planteado Schnider y Foot (2004: 115), la existencia de la Web es “una mixtura única entre lo efímero y lo permanente”. Por un lado, los contenidos en la Web requieren una cierta persistencia en línea para ser transmitidos y una vez que comienzan a ser compartidos en Internet se vuelve difícil lograr borrarlos. Pero a la vez los rastros digitales pueden ser profundamente efímeros, durando relativamente poco y siendo muy fáci-

les de manipular, pues dependen de una intencionalidad tanto humana como no-humana de mantener en línea esos trazos, ya sea por actualizar un sitio web, para bajar una información equivocada o para censurar contenidos contrarios. Es por esto que una de las principales preguntas que se pueden hacer dentro de los estudios de la Web, como plantean los autores, es por las formas de acciones comunicativas que son inscritas en la Web y cómo van cambiando en el tiempo (Schnider y Foot, 2004: 116).

Si vamos a cuestiones más concretas, desde el lado de lo efímero uno podría preguntarse por qué un medio digital como el buscador de Google es sumamente reacio a borrar de sus resultados ciertos links, mientras que en otros medios digitales algunos tienen el poder de decidir qué ocultar, borrar y modificar y cuándo hacerlo, de manera sencilla y rápida. Un dueño de una página web de algún político o una importante compañía a la orden de un clic puede bajar información relevante para evaluar su actuar y puede establecer diferentes restricciones mediante los archivos *robot.txt* para bloquear el rastreo de esa información por parte de terceros. Y a su vez, desde el lado de lo que –infamemente- permanece por años, resulta sumamente necesario el llamado derecho al olvido –aunque difícil de respetar- cuando se trata de los individuos y actores con menor poder. Por ejemplo, la subida y difusión de fotos y videos sexuales muy personales de antiguas parejas a modo de venganza o las tácticas para seleccionar personal que involucran búsquedas en Internet sobre los postulantes, plantean una serie de problemáticas que reflejan cómo el poder distin-

gue y selecciona qué debe ser viral para escandalizar, que debe ser guardado para futuros escándalos y que debe ser “apagado” para evitar otros cuantos más. De este modo, por un lado en la Web se difunde y circula mucha más información de lo que algunos desearían; y por otro, se modifica, suprime y desaparece mucha información que otros desearían que se conservase y se mantenga. De esta manera, la Web no se constituye como un archivo bajo de la idea tradicional e idealizada de una biblioteca que conserva el conocimiento que merece estar disponible para las futuras generaciones de la humanidad. Es más bien un archivo descentrado, inestable y sumamente contingente que se transforma constantemente por estrategias y relaciones de poder finalmente.

La llamada *Media Archeology*, inspirada en las nociones de Archivo y la arqueología del conocimiento de Foucault (1969), ha analizado este aspecto de Internet. Wolfgang Ernst, uno de los exponentes de esta corriente, ha resaltado que Internet no debe ser considerado como ese tradicional archivo estable y estático, cualitativamente selectivo según la calidad de los contenidos y bien establecido en lugares físicos acotados. Más bien, Internet es un archivo completamente dinámico, de constante actualización y adición en tiempo real, basado en la acumulación de información y un marcado “formato-centrismo” pues priman los softwares y protocolos de transmisión más que el medio material o hardware. Como decíamos, los archivos digitales tienen algún grado de perduración pero a la vez son efímeros y de un tiempo sumamente breve o micro-temporales como plantea

Ernst (2013). Esto obliga a concebir la Web como algo que reúne información, actores y entidades de manera momentánea y fugaz por lo que “ostensiblemente el más grande archivo digital, el Internet, es de hecho una colección o una ensamblaje” (Ernst, 2013: 85). En otras palabras, según Ernst, Internet sería una suerte de trans-archivo, dinámico e inquieto que es elaborado tanto por humanos como por las máquinas (los crawler, bots, etc. son prueba viva de ello), que nunca está acabado, siempre está siendo afectado y transformado bajo un sistema de comandos (su *arché*), no tiene un tiempo frío, ordenado y estabilizado. De modo que al estudiar los rastros digitales de lo social desde esta sensibilidad arqueológica, no se nos confiere un regreso de lo muerto o ya pasado —o pisado— a nuestro presente sino más bien una presencia retrasada (Ernst, 2013) o en un régimen temporal diferente pues es perdurada en una memoria tecnológica parcial. Como ha escrito Mike Featherstone:

El archivo digital no debería ser visto solamente como parte de la manía contemporánea por ‘grabar y almacenar’ facilitada por las tecnologías digitales, sino que proveen un archivo fluido, procesual, dinámico, en que la topología de los documentos puede ser reconfigurada una y otra vez [...] El énfasis se mueve a rastros archivísticos fluidos y complejos, en que el archivo no puede proveer un acceso directo al pasado, sino solo una refiguración textual de él (2006: 596).

Cualquier objeto de estudio, abordado por tecnologías y medios digitales, se revela entonces construido de manera híbrida y jamás como “lo que es”, sino como lo que quedó de este a posteriori, luego del despliegue de una serie de tácticas por mantener online, de (in)visibilizar o borrar contenidos y materiales digitales de ese objeto. Nuevamente nos encontramos con que los rastros digitales no aparecen “en estado libre” sino que sólo podemos llegar a ellos luego de su toque con el poder (Foucault, 1999). Lo que permanece en línea, esos rastros que despliegan lo social a juicio de Venturini y Latour (2010), tiene una historia política detrás que de no afirmarla estaríamos visualizando limitadas visiones de mundo y reproduciendo sus parcialidades. Una mirada más arqueológica a los medios, consistiría aplicar entonces una suerte de *ingeniería en reversa epistémica* (Ernst, 2013) que abra no solo la caja negra sino también las herramientas que usamos en ese proceso de apertura, pues “los medios no son solo objetos sino también sujetos (“autores”) de la arqueología medial” (Ernst, 2013: 58). Por lo que los médium también tienen injerencia política en la forma en cómo se registran y se archivan nuestro propio conocimiento de la realidad. Y en este caso nos compele a no sólo hablar de la bendición de los rastros digitales de lo social, sino dar cuenta de las parcialidades y de estas autorías de los medios tecnológicos en ellas, de las transformaciones que han sufrido esos rastros en el maleable y cambiante archivo digital.

Pero justamente debido a lo volátil, dinámico y efímero de Internet, los estudios digitales continuamente se han centrado en temporalidades específicas y bien reduci-

das. Por las limitaciones de los medios y a las capacidades de manejo de la data, se hace difícil una visión histórica más amplia o longitudinal. Por ejemplo la actual API de Twitter sólo permite recopilar tweets de menos de una semana de antigüedad, por lo que se termina prefiriendo recopilar información de *tweets* de lapsos de tiempo de semanas o días –e incluso así se alcanzan bases de datos de gran volumen. Como han planteado Boyd y Crawford por su parte:

Twitter y Facebook son ejemplos de fuentes de Big Data que ofrecen funciones de archivación y búsqueda bien pobres. Seguidamente, los investigadores tienden a focalizarse en algo en el presente o en el pasado inmediato –siguiendo las reacciones a una elección, el final de un show de TV, o un desastre natural- por la pura dificultad o imposibilidad de acceder a datos más antiguos (2012: 666).

Pero aquí la respuesta no debiese ser resignarse a estudiar lo más contemporáneamente polémico, aquel viral que circula en vivo y en directo, sino más bien en la crítica de por qué esto es así y abrir o visibilizar las relaciones e intereses detrás de tales restricciones y parcialidades del archivo de estas redes sociales. Así, considerando los rastros y huellas digitales como acontecimientos socio-técnicos, que adquieren una existencia remanente en una memoria o archivo digital sumamente dinámico y volátil; la descripción arqueológica de qué cosas permanecen a la vista, accesibles a la orden de un clic, y que cosas han sido transformadas

o han dejado de estar “disponibles” para nuestro país, región o cultura, nos permitirán encontrar elementos relevantes y sumamente políticos del desenvolvimiento –también arqueológico muchas veces- de Internet. Esta imparcialidad del archivo digital, entonces, no solo implica una búsqueda por mejores instrumentos de recopilación de datos digitales sino también la afirmación de que hay terceros y tácticas de poder que están condicionando la propia investigación de estos rastros. Siguiendo con el caso de Twitter, diversas empresas han generado un lucrativo negocio con Twitter, permitiendo recopilar tweets más antiguos a cambio de sumas de dinero, replicando las divisiones que ya comentaban Savage y Burrows (2007). De esta manera, el problema técnico de la parcialidad del archivo, pasa a ser una cuestión más que técnica.

Los autores detrás de los *Digital Methods Initiative* no han estado al margen de esta discusión sobre lo temporal y efímero de los archivos digitales y han buscado hacerse cargo de esto. El llamado es a seguir los medios digitales “a fin de no ser sorprendidos por los medios inestables (por ejemplo, la desaparición de una API) o lo efímero del contenido (cuando un sitio se actualiza o es sacado)” (Rogers et al. 2015: 31). Por lo que el carácter desordenado, efímero e inestable de Internet y en particular la cambiante vida de las páginas web¹⁰, no debe ser considerado como un sesgo imprevisto a eliminar o poner entre paréntesis, sino que debe ser algo a explic(it)ar pues es lo propio de todo medium. “El dinamismo de los datos en línea, lejos de ser solo un problema (deterioro, incompletitud), puede a su vez

ser desplegado para fines analíticos en la investigación en línea” (Marres y Weltverde, 2013: 14). Como ha propuesto recientemente Noortje Marres a raíz de las constantes manipulaciones y ruidos que generan los bots en bases de datos de Twitter, seguimos aquí una *postura afirmativa* de estos sesgos digitales:

En vez de tratar los sesgos digitales como un fenómeno negativo a poner entre paréntesis, debemos pues desarrollar tácticas metodológicas y empíricas que permitan responder la pregunta de cómo los dispositivos digitales participan en el enactamiento de controversias y la formación de temas (Marres, 2015: 23).

Los medios tecnológicos, como buenos mediadores, no son nunca neutrales y que queden en línea ciertos rastros digitales y otros desaparezcan a lo largo de la vida de la data digital, nos dicen algo que debe ser descrito. El siempre parcial y dinámico enredo de entidades, como bots, algoritmos, usuarios, corporaciones, plataformas, etc. no solo es producido por los temas a investigar, sino que a su vez los moldean y direccionan a ciertas rutas y no otras. Los hipervínculos, los tweets, los likes, etc. así como sus herramientas de análisis, no son neutrales a la hora de estructurar y valorar el contenido digital (Marres, 2015). Los medios digitales, de esta forma, nos representan de manera incompleta y retrasada temporalmente nuestro objeto de estudio, pero esas limitaciones, como por ejemplo la inestabilidad y desaparición de cierta información en la Web, no debe ser considerados como algo ajeno, circunstancial o ex-

10 Por ejemplo el trabajo que realiza la Internet Archive de guardar todo tipo de objetos de la web, y en particular el de poder registrar la vida social de sitios web mediante la herramienta WayBack Machine es sumamente útil para la investigación social digital que busque una aproximación histórica y longitudinal.

terno, sino como una verdadera dinámica sumamente política de visibilizar o invisibilizar el objeto analizado. ¿Qué implicancias tiene que cierto contenido haya sido sacado de un sitio? ¿A quién beneficia que aparezca cierto actor como el más relevante según los rastros que han perdurado hasta hoy? ¿Qué rastros fueron trastocados, manipulados o borrados? ¿Qué agencia tiene el dispositivo inscriptor en el análisis? Son preguntas que le agregan complejidad y mayor consistencia a la investigación digital.

Con cada metodología e instrumentos podremos capturar y sentir ciertas cosas y no otras, hay opacidades en cada representación. Las propias limitaciones y defectos de estos nuevos métodos digitales, así como de otros, serán resultado de las dinámicas sociales y desigualdades que producen esa información después de todo. No todo pasa por dispositivos digitales, y no necesariamente toda mediación deja una huella digital permanente ni es procesada de manera fácil y económica. Que ciertos sitios web desaparezcan o cambien de un día para otro, que Google muestre ciertos resultados en base a nuestros historiales, que las discusiones públicas en Twitter sean muchos redundantes *retweets* y *bots* repitiendo mensajes programados, que cierta data geolocalizada muestre mapas sesgadamente hacia el Norte, que cientos de artículos Wikipedia sobre países de Latinoamérica y África sean modificados por personas en Europa y USA y principalmente por hombres, etc. nos hablan de las grandes y siempre presentes deficiencias que tendrán los datos que recopilemos de estas tecnologías digitales —así como de cualquier tipo de dato. Pero

estos “defectos”, “sesgos” o “ruidos” también nos dicen algo y dan cuenta de que no son saberes dados sino que tienen su vida social, su historia política, un enredo de actores involucrados detrás, que es necesario indagar y afirmar antes que poner entre paréntesis, a la hora de hacer cualquier investigación social digital.

05. Latinoamérica digital: yendo más allá de la conectividad

Como vimos, el Big Data, la data transaccional, las tecnologías de *self-tracking*, las aplicaciones *Smart* y todos los dispositivos digitales ubicuos que median, inscriben y archivan de algún modo la vida social contemporánea, ciertamente acumulan rastros de lo social en niveles y magnitudes novedosas que pueden iluminar nuevos fenómenos sociales que de otro modo serían improbables de poder capturar o sentir. Sin embargo, no se debe caer en un exagerado entusiasmo, la información digital no debe ser entendida como dada, estabilizada y meramente técnica. No son puros 1s y 0s que son estructurados algorítmicamente de manera neutral, sino que se merecen algo más pues tienen una vitalidad de suyo propia. En la activa producción de datos digital se involucran y conectan a heterogéneas entidades tanto humanas como no-humanas (Latour, 2005; Ruppert et al. 2015) interfiriendo o agregando capas de complejidad al análisis. De modo que permiten una particular trazabilidad de lo social con sus propias limitantes, pero que pueden perfectamente complementarse con otros métodos como entrevistas, etnografías o encuestas, etc.

Es perentorio que el Sur y en particular Latinoamérica se haga parte de los debates que están suscitando estas tecnologías y fuentes digitales de información en las ciencias sociales principalmente cargados a casos de Estados Unidos y Europa. Más aún cuando las principales plataformas, algoritmos, servidores, redes sociales, “hubs”, etc. se mantienen y desarrollan en el Norte. Se están radicalizando asimetrías entre países del Sur y los conglomerados transnacionales que dominan gran parte del uso cotidiano de Internet y establecen nuevas relaciones de dependencia y control (Arora, 2016). Es urgente indagar en las nuevas formas de dominación que están ejerciendo los controladores de estas nuevas infraestructuras digitales, bloqueando ciertos contenidos en algunos países, vulnerando los consentimientos de los usuarios, restringiendo el uso académico de datos, haciendo más efímeros ciertos rastros y otros más permanentes, etc. boyd y Crawford (2012) relevaban la creciente separación entre quienes poseen las bases de datos (*data rich*) y los que no (*data poor*), y muchas de esas bases de datos están siendo recopiladas y guardadas celosamente por empresas y gobiernos. Nuevas diferencias de poder se están radicalizando entre quienes generan los datos digitales y quienes los coleccionan, almacenan y procesan en lo que Andrejevic (2014) ha llamado por la brecha de Big Data. No todos los actores tendrán las mismas herramientas y conocimientos para hacer una minería y procesamiento de los datos digitales. Se requieren potentes computadoras y el desarrollo de algoritmos que aunque se programen en código abierto, seguirán siendo utilizados por una población reducida y experta en la materia, bajo un particular

lenguaje y persiguiendo determinados intereses. Si bien Andrejevic plantea esta división entre los usuarios individuales, productores de los datos, respecto de los grandes recolectores de datos como compañías, gobiernos o universidades, es evidente que no solo se puede hallar estas diferencias a nivel individual sino también a nivel regional.

La literatura latinoamericana sobre Internet y las tecnologías digitales, según Trejo (2013) estaría marcada por diferencias generacionales principalmente. La generación actual serían “inmigrantes digitales” en que han vivenciado de cerca el paso de un mundo análogo a uno digital, considerando los dispositivos digitales como objetos de estudio por sí mismos y a Internet como una arena especial y clausurada, por ejemplo exacerbando la distinción de una cibercultura y contrastándola con el mundo análogo y los medios de comunicación tradicionales o pre-existentes. De modo que la pregunta central se vuelve en como abrir el acceso a que todos y todas ingresen a este ciberespacio. Por ende, bajo esta metáfora espacial de Internet, la discusión en América Latina sobre los medios y tecnologías digitales se ha focalizado en las brechas de acceso: la falta de acceso a Internet en zonas rurales, las exclusiones económicas de poblaciones más pobres, brechas de género, o más recientemente en las diferencias de habilidades digitales. En esta línea, metas como las que se ha propuesto el Gobierno de Chile de la conectividad universal, buscando generar el acceso a todas las personas a redes de alta velocidad y calidad, se vuelven problemáticas sí tal conectividad universal comporta el

ensanchamiento del dominio y control de los datos por parte de empresas y conglomerados transnacionales con base en *Silicon Valley*, y que más aún, hacen opacos usos de los datos digitales y difícilmente disputables desde el Sur. La pregunta no debiese ser por quiénes no están pudiendo producir rastros digitales y cómo hacer para que los produzcan -mejorando la conectividad entre otras cosas- sino también bajo qué plataformas, qué servicios y hardware, qué procesos de tratamiento de datos, bajo que regímenes de propiedad y marcos regulatorios. Sin este tipo de problematizaciones, “la superación de la brecha digital representa exacerbar la brecha del Big Data” (Andrejevic, 2014: 1686). Dicho de otro modo, la promesa de una total conectividad – así como de otras nuevas promesas tecnológicas como las *Smart Cities* o *Internet of Things*- puede generar fructíferos beneficios para los contextos del Sur, pero a la vez pueden desembocar mayormente, según los modos en que ésta se lleve a cabo, en un crecimiento de las diferencias de poder y una reproducción de dependencias tecnológicas, potenciando un legado colonial y la sensación de falta de empoderamiento en los usuarios ante actores transnacionales tan grandes y ubicuos (Andrejevic, 2014; Arora, 2016).

Por lo que es necesario impregnar los estudios latinoamericanos de un enfoque “nativamente digital” en su sentido fuerte¹¹, re-apropiándonos y contextualizando los planteamientos de Richard Rogers (2013) a nuestra realidad regional. Que consideremos los medios y tecnologías digitales (Internet, redes sociales, plataformas, sitios web, videojuegos, comercio electrónico, datos

transaccionales, etc.) ya no como creadores de espacios, arenas o territorios aislados, autoclausurados, políticamente neutrales e irreales (Miller y Slater, 2000; Miller y Horst, 2012), sino que *afirmarlos* como posibles fuentes de información o indicadores de comportamientos de nuestras sociedades, con sus propias limitaciones y parcialidades político-económicas que deben ser descritas y problematizadas (Marres, 2015). Si la novedad de las tecnologías computacionales implicaba atender a sus usos y “efectos” *sobre* la sociedad, en lo que Trejo (2006) esquematizó bajo el concepto de cibernociología; hoy podemos apuntar a una sociología digital que estudie los variados fenómenos de nuestras culturas latinoamericanas incluyendo estos dispositivos digitales en el repertorio metodológico, explorando *con* Internet la materialización de las diversas dinámicas e inequidades sociales que nos caracterizan, así como también re-posicionando a actores que de otro modo seguirán sido invisibilizados en la producción social de datos digitales (Rogers, 2013; Rogers, et al., 2015).

Finalmente, con la provocación de la data digital y desde un enfoque arqueológico, podemos ser más conscientes de sus limitaciones, exageraciones y peligros. Notar que existen múltiples vidas en nuestros métodos y fuentes de información, y que dependen de nuestras capacidades e instrumentos técnicos de cada época así como de relaciones de poder. Poner de manifiesto la vida social de los datos quizás los dejará como más imprecisos y parciales, pero a la vez se harán sensible las texturas de las inequidades y asimetrías de nuestras sociedades. Por todo esto, abrir un agenda de estudios

11 Entendiendo este concepto no como una mera distinción generacional por superarse en el tiempo, sino como una forma de aproximación epistemológica a lo digital, que lo conciba antes que un territorio o terra incognita, como ensamblajes de datos, humanos, algoritmos, dispositivos, etc. que generan información sobre entidades híbridas y que a su vez las performan y afectan dialógicamente (Vease Lupton, 2016).

digitales que tengan un carácter latinoamericano y a la vez un alcance global (Trejo, 2013), con una aproximación crítica e interdisciplinaria, y que re-orienta y se apropie de Internet y los múltiples dispositivos digitales para indagar los fenómenos sociales que resulten de interés para la propia región, se vuelve un desafío complejo pero sugerente. Mas sumamente necesario, toda vez que aún falta por examinar empíricamente cómo en el Sur la creciente generación de datos masivos digitales, de manera más o menos visible, hacen trazable lo social, ensamblando a diferentes entidades y no otras, y representando a específicos individuos, regiones, hemisferios, historias y realidades.

06. Referencias bibliográficas

- Anderson, C. (23 de Junio, 2008). "The end of theory". *Wired*, 16. Recuperado de: http://archive.wired.com/science/discoveries/magazine/16-07/pb_theory
- Andrejevic, M. (2014). "The Big Data Divide". *International Journal of Communication* 8, 1673–1689. Recuperado el 10 de Junio de 2016 de: <http://ijoc.org/index.php/ijoc/article/view/2161>
- Arora, P. (2016). "The Bottom of the Data Pyramid: Big Data and the Global South". *International Journal of Communication* 10, 1681–1699.
- Back, Les (2012). "Live sociology: social research and its futures". *The Sociological Review*, 60(1), 18–39. DOI: 10.1111/j.1467-954X.2012.02115.x
- Bakshy, E., Messing, S., Adamic, L. (2015). "Exposure to ideologically diverse news and opinion on Facebook". *Science* 348, (6239), pp. 1130-1132. DOI: 10.1126/science.aaa1160
- Bauman, Z. y Lyon, D. (2013). *Vigilancia Liquida*. Buenos Aires: Editorial Paidós.
- boyd, D. & Crawford, K. (2012). "Critical Question for Big Data". *Information, Communication & Society*. 15(5), pp. 662-679, DOI: 10.1080/1369118X.2012.678878
- Bowker, G.C. (2014). "The theory/Data thing". *International Journal of Communication* 8, 1795–1799. Recuperado el 12 de Junio de 2016 de: <http://ijoc.org/index.php/ijoc/article/view/2190>
- Bruns, A. y Burgess, J. (2016). "Methodological innovation in precarious spaces: the case of Twitter". En Helene Snee et al. (eds.), *Digital Methods for Social Science. An Interdisciplinary guide to research innovation*. Nueva York, NY: Palgrave MacMillan
- Celorio, M. (2011). *Internet y Dominación. Hacia una sociología de la nueva espacialidad*. Ciudad de México: Plaza y Valdés Editores.
- Crandall, D., Backstrom, L., Huttenlocher, D. y Kleinberg, J. (2009). "Mapping the World's photos". *World Wide Web Conference*, Abril 20-24.
- Ernst, W. (2013). *Digital Memory and the Archive*. Minneapolis, MN: University of Minnesota Press.
- Featherstone, M. (2006). "Archive. Special Issue Problematizing Global Knowledge". *Theory, Culture y Society*, 23 (2-3), 591-596.
- Featherstone, M. (2009). "Ubiquitous Media. An Introduction". *Theory, Culture y Society*, Vol 26 (2-3), 1-22.
- Ford, H. (2014). Big Data and Small: Collaborations between ethnographers and data scientists. *Big Data y Society*, DOI: 10.1177/2053951714544337
- Foucault, M. (1969). *La Arqueología del Saber*. Buenos Aires: Editorial Siglo XXI.
- Foucault, M. (1999 [1994]). *Estrategias de poder*. [Trad. y Edición de Julia Varela y Fernando Alvarez Uría] Buenos Aires: Paidós.

- Gitelman, Lisa (2013). *“Raw Data” Is an Oxymoron*. Cambridge, MA: The MIT Press.
- Graham, M. (9 de Marzo, 2012) “Big data and the end of theory?”. En *The Guardian*. Recuperado el 23 de Junio de 2015 de: <http://www.theguardian.com/news/datablog/2012/mar/09/big-data-theory>
- González-Bailon, S. (2013). “Social Science in the Era of Big Data”. *Policy y Internet*, 5 (2), 147-160. DOI: 10.1002/1944-2866.POI328.
- González-Bustamante, B. (2015) “Evaluando Twitter como indicador de opinión pública: una mirada al arribo de Bachelet a la presidencial chilena 2013”. *Revista SAAP*, 9(1), 119-141.
- Helles, R., y Jensen, K. B. (2013). “Introduction to the special issue “making data—big data and beyond.” *First Monday*, 18(10). Recuperado de: <http://firstmonday.org/article/view/4860/3748>
- Hilbert, M., Vázquez, J., Halpern, D., Valenzuela, S., Arriagada, E. (2016). “One Step, Two Step, Network Step? Complementary Perspectives on Communication Flows in Twittered Citizen Protests”. *Social Science Computer Review*, Abril 12, DOI: 10.1177/0894439316639561
- Kramer, A., Guillory, J., Hancock, J. T. (2014). Experimental evidence of massive-scale emotional contagion through social networks. *PNAS* 111 (24), 8788–8790.
- Kosinski, M., Stillwell, D., y Graepel, T. (2013). “Private traits and attributes are predictable from digital records of human behavior”. *PNAS*, 110, 5802–5805.
- Latour, B. (1998). “Thought experiments in social science: from the social contract to virtual society”, *Ist Virtual Society? Annual Public Lecture*, [En Línea] Recopilado de: <http://www.artefaktum.hu/it/Latour.htm>
- Latour, B. (2005). *Reensamblar lo social. Una introducción a la teoría del actor-red*. Buenos Aires: Manantial.
- Latour, B., Jensen, P., Venturini, T., Grauwin, S. y Boullier, D. (2012). “The whole is always smaller than its parts’ –a digital test of Gabriel Tarde’s monads”. *The British Journal of Sociology*, 63(4), 590-615.
- Lazer, D., Pentland, A., Adamic, L., Aral, S., Barabási, A., Brewer, D., Christakis, N., Contractor, N., Fowler, J., Gutmann, M., Jebara, T., King, G., Macy, M., Roy, D. y Van Alstyne, M. (2009). “Computational Social Science”. *Science*, 323, 721.
- Latham, R., y Sassen, S. (2005). *Digital Formations: IT and New Architectures in the Global Realm*. Princeton, Nueva Jersey: Princeton University Press.
- Lupton, D. (2016). “Digital companion species and eating data: Implications for theorising digital data–human assemblages”. *Big Data y Society* 3 (1), DOI: 10.1177/2053951715619947
- Margets, H. y Sutcliffe, D. (2013). “Editorial. Addressing the Policy Challenges and Opportunities of “Big Data””. *Policy y Internet*, 5 (2), 139-146.
- Marres, N. (2012a). “The Redistribution of Methods”. *The Sociological Review*, Vol 60 (1), 139-165. DOI: 10.1111/j.1467-954X.2012.02121.x

Marres, N. (2012b). *Material Participation: Technology, the Envi-ronment and Everyday Publics*. Londres: Palgrave Macmillan.

Marres, N. (2015). "Why Map Issues? On Controversy Analysis as a Digital Method". *Science, Technology, and Human Values*, 40 (5), 1-32. DOI: 10.1177/0162243915574602.

Marres, N. y Rogers, R. (2005). "Recipe for Tracing the Fate of Issues and their publics on the Web". En Latour, B. y Wiebel, P. (Eds) *Making Things Public. Atmospheres of democracy*. Cambridge, MA: MIT Press.

Marres, N. y Moats, D. (2015). "Mapping Controversies with Social Media: The Case for Symmetry". *Social Media + Society*, Julio-Diciembre, pp. 1-17.

Marres, N. y Weltevrede, E. (2013) Scraping the social?, *Journal of Cultural Economy*, 6(3), pp. 313-335, DOI: 10.1080/17530350.2013.772070

Michel, J., Shen, Y. K., Presser Aiden, A., Veres, A., Gray, M. K., The Google Books Team, Pickett, J. P., Hoiberg, D., Clancy, D., Norvig, P., Orwant, J., Pinker, S., Nowak, M. A., Lieberman Aiden, E. (2011). "Quantitative Analysis of Culture Using Millions of Digitized Books". *Science*, 331, pp. 176- 182.

Miller, D. y Horst, H. (Eds.) (2012). *Digital Anthropology*. Nueva York: Berg Publishers.

Miller, D. y Slater, D. (Eds.) (2000). *The Internet: An ethnographic approach*. New York: Berg Publishers.

Moat, H. S., Curme, C., Avakian, A., Kenett, D. Y., Stanley, H. E., Preis, T. (2013). "Quantifying Wikipedia usage patterns before stock market moves". *Scientific Reports* 3: 1801, SOI: 10.1038/srep01801.

Morozov, E. (2013). *To Save Everything, Click Here: The Folly of Technological Solutionism*. Nueva York, NY: Public Affairs

Pentland, A. (2014). *Social Physics: How Social Networks can make us smarter*. Nueva York, NY: The Penguin Press

Rogers, R. (2013). *Digital Methods*. Cambridge, MA: MIT Press.

Rogers, R. (2014). "Political Research in the Digital Age". *International Public Policy Review* 8(2), pp. 73-84.

Rogers, R., Sanchez-Querubín, N. y Kil, A. (2015). *Issue Mapping for an Ageing Europe*. Amsterdam: Amsterdam University Press.

Ruppert, E., Savage, M. y Law, L. (2013). "Reassembling Social Science Methods: The challenge of Digital Devices". *Theory, Culture y Society*, 30 (4), pp. 22-46.

Ruppert, E. Harvey, P., Lury, C., Mackenzie, A., McNally, R., Baker, S. y Lewis, C. (2015). Socializing Big data: from concept to practice. *CRESC Working Paper Series No. 138*. The University of Manchester and the Open University.

Sajuria, J. y Fábrega, J. (2016). "Do we need polls? Why Twitter will not replace opinion surveys, but can

- complement them”. En Helene Snee et al. (eds.), *Digital Methods for Social Science*. Nueva York, NY: Palgrave MacMillan
- Savage, M. (2013). “The ‘Social Life of Methods’: A Critical Introduction”. *Theory, Culture y Society*. 30 (4), 3-21. DOI: 10.1177/0263276413486160
- Savage, M. y Burrows, R. (2007). “The coming crisis of empirical sociology”, *Sociology*, (41), 885–899.
- Savage, M. and Burrows, R. (2014). “After the crisis? Big Data and the methodological challenges of empirical sociology”. *Big Data y Society*. 1(1) pp.1-6.
- Schneider, S. y Foot, K. (2004). “The web as an object of study”. *New Media y Society*, 6(1), 114-122. DOI: 10.1177/1461444804039912
- Trejo, R. (2006). “Cibersociología, territorio de encuentro y en construcción”. Ponencia para el Encuentro Nacional de la Asociación Mexicana de Investigadores de la Comunicación. Morelia, Michoacán. Recuperado el 14 de Mayo de 2016 de: <https://rtrejo.files.wordpress.com/2012/04/cibersociologc3ada-amic-abril-06.pdf>
- Trejo, R. (2013). The Study of the Internet in Latin America. Achievements, Challenges, Futures. En Valdivia, A. (Ed.) *The International Encyclopedia of Media Studies. Volume VI: Media Studies Futures*. Oxford: Blackwell Publishing Ltd
- Tufekci, Z. (2014). “Big Questions for Social Media Big Data: Representativeness, Validity and Other Methodological Pitfalls”. *Proceedings of the Eighth International AAAI Conference on Weblogs and Social Media*, Palo Alto, CA: The AAAI Press.
- Venturini, T. (2010). “Building on faults: how to represent controversies with digital methods”. *Public Understanding of Science* 20(10), 1–17.
- Venturini, T. y Latour, B. (2010). “The Social Fabric: Digital Traces and Quali-quantitative Methods”. En *Proceedings of Future En Seine 2009*. Paris: Editions Future en Seine.
- Venturini, T., Jensen, P. Latour, B. (2015) “Fill in the Gap. A New Alliance for Social and Natural Sciences”. *Journal of Artificial Societies and Social Simulation*, 18 (2) 11.
- Venturini, T., Meunier, A., Munk, A., Rogers, R., Borra, E., Rieder, B., Bounegru, L., Sanchez-Querubin, N., Ciuccarelli, P., Mauri, M., Azzi, M., Ciminieri, D., Uboldi, G., Gerry, P., Kitcher, H., Schon, R., Kaltenbrunner, A., Laniado, D., Fleischhauer, M. (2014). “Climaps by EMAPS in 2 pages (a summary for policy makers and busy people)”. *Social Science Research Network*. Disponible en SSRN: <http://ssrn.com/abstract=2532946> o <http://dx.doi.org/10.2139/ssrn.2532946>