

Big Data: las implicancias de la democratización de los métodos. Oportunidades y desafíos que los dispositivos y datos digitales plantean para la investigación social

Andrés Asenjo Morosetti¹ y
Christian Riveros Pino²

Fecha de recepción: 5 de octubre de 2015

Fecha de aprobación: 12 de mayo de 2016

Resumen

El presente ensayo monográfico intenta brindar una panorámica global del fenómeno del Big Data, entendido como un sistema de producción, procesamiento, sistematización y análisis de datos sociales a gran escala. Actualmente, una multiplicidad de entes se han posicionado como agentes recopiladores de datos sobre conductas y prácticas sociales, gracias a avances tecnológicos que han incrementado las capacidades sistémicas para la obtención y análisis de grandes volúmenes de información. Esto supone una suerte de democratización de los métodos, lo cual plantea una serie de desafíos y oportunidades para los científicos sociales, otrora agentes exclusivos de la producción de lo social. Ante esto, el presente documento ofrece un marco general de análisis, esbozando reflexiones temáticas y estratégicas para la comunidad científica en general y para la sociología en particular. El ensayo se estructura a partir de tres secciones: una introducción, que ofrece un acercamiento global al tema; una sección de discusión, en la que se revisa a grandes rasgos las principales problemáticas y desafíos que presenta el Big Data para la investigación social y los científicos sociales, y una sección de conclusiones, en la que se esbozan posibles alternativas de integración metodológica entre las nuevas formas de investigación social y los métodos tradicionales.

Palabras clave: Big Data, tecnologías de la información, métodos de investigación, investigación social, empirismo.

¹ Sociólogo, Universidad Católica Silva Henríquez, Magíster (c) en Métodos para la Investigación Social Universidad Diego Portales, Santiago, Chile. Contacto: aasenjo@uft.cl trabajando en Universidad Finis Terrae.

² Sociólogo, Universidad Diego Portales, Magíster (c) en Métodos para la Investigación Social Universidad Diego Portales, Santiago – Chile. Contacto: christian.riveros@kronoschile.cl trabajando en KRONOS Consultoría e investigación de mercado.

Big Data: Implications of the democratization of methods.

Opportunities and challenges for social research provided by digital data and devices

Abstract

This monographic essay attempts to provide a comprehensive overview of the phenomenon of Big Data, understood as a system of production, processing, systematization and analysis of large-scale social data. Currently, a multiplicity of entities have positioned themselves as agents collecting data, behaviors and social practices, thanks to technological advances that have increased systemic capacities for collection and analysis of large volumes of information. This is a sort of democratization of methods, which poses a number of challenges and opportunities for social scientists, once exclusive agents of social production. In this way, the present document provides a general framework for analysis outlining thematic and strategic considerations for the scientific community in general and sociology in particular. The essay is structured in three sections: an introduction, which provides a comprehensive approach to the subject; a section of discussion, which reviews in broad terms the main problems and challenges imposed by Big Data for social research and social scientists; and a section of conclusions with possible alternatives for methodological integration among new forms of social research and the traditional methods.

Keywords: Big Data, information technology, research methods, social research, empiricism.

Big Data: as implicações da democratização dos métodos.

Oportunidades e desafios que os dispositivos e dados digitais projetam para a pesquisa social

Resumo

O presente ensaio monográfico tenta oferecer uma visão panorâmica do fenômeno do Big Data, entendido como um sistema de produção, processamento, sistematização e análise dos dados sociais em grande escala. Atualmente, uma multiplicidade de entidades têm se posicionado como agentes coletores de dados sobre comportamentos e práticas sociais, graças aos avanços tecnológicos que aumentaram as capacidades sistêmicas para a obtenção e análise de grandes volumes de informação. Esta é uma espécie de democratização dos métodos, o que coloca uma série de desafios e oportunidades para os cientistas sociais, pois em outros tempos foram os agentes exclusivos de produção social. Diante disso, o presente documento apresenta um quadro geral de análise, delineando reflexões temáticas e estratégicas para a comunidade científica em geral e para a sociologia em particular. O ensaio é estruturado em três seções: uma introdução, que oferece uma abordagem global para o tema; uma seção de discussão,

onde é revista em termos gerais, os principais problemas e desafios que apresenta a Big Data para a investigação social e os cientistas sociais, e uma seção de conclusões onde se esboçam possíveis alternativas de integração metodológica entre as novas formas de investigação social e os métodos tradicionais.

Palavras-chave: Big Data, tecnologias da informação, métodos de pesquisa, pesquisa social, empirismo.

Introducción

La producción de datos sociales ya no constituye una actividad específica o exclusiva de los científicos sociales. Actualmente, entidades gubernamentales, instituciones comerciales y financieras, organismos de seguridad y control, entre otros tipos de organizaciones, son capaces de producir y procesar grandes volúmenes de datos y generar una representación de la realidad en las cuales ellos intervienen. Lo mismo ocurre en términos de las plataformas tecnológicas de interacción y transacción en la web: redes sociales, páginas web y otras entidades que promueven y facilitan el intercambio comercial vía internet, registrando operaciones y transacciones, y generando bases de datos con una extensa cantidad de información. Aún más, el uso extensivo de dispositivos electrónicos interconectados globalmente a la red, permite registrar las acciones de los sujetos poseedores de dichos dispositivos y, ya sea de manera consentida o no consentida por el usuario (ya que este usuario no siempre es consciente de que sus acciones son registradas), las huellas de muchas de sus acciones cotidianas quedan almacenadas, haciéndolas susceptibles de ser rastreadas.

La producción de datos conforma conjuntos de información en una cantidad mucho más amplia, comparada con los volúmenes de información con los que tradicionalmente la investigación social estaba acostumbrada a trabajar. Por lo tanto, consideramos que el fenómeno del Big Data implica una doble dimensión: en efecto, cuando hablamos de Big Data nos referimos tanto a una determinada manera de producir información, como a la extensa cantidad de información susceptible de ser producida por dichos medios de producción de datos.

El uso de dispositivos capaces de registrar las acciones de sus poseedores, tales como teléfonos móviles, tarjetas electrónicas, páginas web, redes sociales, registros y/o formularios administrativos de organismos gubernamentales e instituciones financieras, entre otros, implican el establecimiento de relaciones sociales entre los usuarios y dichos dispositivos; entre los usuarios y los fenómenos que dichos dispositivos representan, y entre usuarios de dispositivos de categorías afines. En este sentido, el presente artículo adopta el prisma de los estudios de la ciencia tecnológica y la teoría del actor red (Latour, 2008; Law, 2004), considerando que la forma de producción de información encarnada por el Big Data a través de los dispositivos que registran datos, la cantidad de datos recolectados y las características de los datos obtenidos a través de dichos medios son capaces de dar cuenta de interacciones que enactan (Mol, 2002), crean o hacen ser cierto un determinado tipo de realidad social. En palabras de Ruppert, Law y Savage (2013), *“digital devices and the data they generate are both the material of social lives and form part of many of the apparatuses for knowing those lives”*³ (p. 4).

A partir de esta perspectiva, es posible advertir que, aunque los terminales en contacto directo con el usuario carezcan de capacidad cognitiva, están dotados de agencia, ya que son capaces de condicionar las prácticas de los sujetos y, en consecuencia, orientar el ensamblaje de los vínculos momentáneos y emergentes a partir de los cuales se compone lo social (Latour, 2008).

El fenómeno del Big Data, los dispositivos que lo constituyen, la forma de producción de datos y la misma información generada mediante los procedimientos del Big Data, son relevantes para el análisis y la investigación social, ya que:

³ “Los dispositivos digitales y los datos que ellos generan son ambos materiales de la vida social y forman parte de los muchos aparatos que permiten conocer acerca de esas vidas”.

- Registran acciones realizadas por sujetos cognoscentes y, de esta forma, dan cuenta de determinados tipos de relaciones sociales.
- El registro de dichos datos está condicionado a que el sujeto cognoscente entable una relación con el dispositivo que registra su acción; por lo tanto, se despliega una relación agente humano - agente material.
- Dichos dispositivos, además, determinan la relación entre individuos y entre individuos y objetos o fenómenos. Por lo que están dotados de agencia.
- Los datos que conforman el Big Data constituyen un insumo para enactar determinados tipos de realidad social.
- En tanto contribuyen a enactar determinadas realidades, resulta necesario indagar qué tipo de realidades quedan invisibilizadas por las prácticas del Big Data y qué implicancias conlleva el ensamblaje de ciertas realidades en desmedro de otras (Mol, 2002).

Además de constituirse en un objeto de investigación para el análisis social, el Big Data ofrece también una oportunidad de innovación metodológica, en tanto sus dispositivos se constituyen en instrumentos para el registro y producción de información relativa al comportamiento humano, y permiten obtener información sobre las acciones de los sujetos en formatos en que la investigación social tradicional escasamente ha podido obtener antes. Sin embargo, esto también reviste la necesidad de reflexionar sobre sus posibles efectos performativos (Callon, 2006).

Por ello, el Big Data sitúa a la ciencia social en escenarios a los cuales no se había visto expuesta antes, frente a los cuales las diversas disciplinas sociales no han logrado elaborar una respuesta clara y/o unificada. Al respecto, las ciencias sociales están tomando dos alternativas para abordar el nuevo escenario planteado por el Big Data. La primera alternativa es una opción resistente, que se empeña en defender las formas tradicionales de investigación social, en oposición a los procedimientos y dispo-

sitivos del Big Data. En contraste, la segunda alternativa aboga por la integración metodológica y la innovación en el estudio de lo social.

La alternativa confrontacional, a nuestro juicio, implica tomar una postura conservadora hacia las nuevas posibilidades que provee este nuevo ensamblaje de prácticas y dispositivos en tanto fenómeno y herramienta de análisis. Esta perspectiva advierte una pugna entre la investigación social tradicional y la investigación social a partir de los dispositivos de Big Data. Al respecto, en el marco de su proyecto de GBCS⁴, Savage (2014) señala:

We are seeing a methodological battle, the outcome of which is currently uncertain, but is likely to end up with greater demarcation and differentiation between the domains of 'old' and 'new' methods. And this politics of method intensifies when new modes of analysis are more directly in competition with the substantive foci of traditional orthodox areas of research⁵ (p. 4).

En la otra vereda, no son pocos los investigadores y comunidades de científicos que apuestan por explorar estas nuevas posibilidades, a partir de novedosas reformulaciones e innovaciones metodológicas que se configuran como una alternativa para asimilar los dispositivos del Big Data e integrarlos como parte del arsenal técnico de los científicos sociales. Ruppert, Law y Savage (2013) señalan:

Several social science research centres and initiatives have taken up the challenge of digital data and methods (...). In the academy, researchers have also adapted social science methods to forms of the digital (...). Our point [is] that a number of

⁴ *The Great British Class Survey*, trabajo empírico efectuado por M. Savage desde la lógica de la producción y data análisis con los modelos de análisis del Big Data.

⁵ "Estamos en presencia de una batalla metodológica, y el resultado actualmente es incierto, pero es probable que termine con una gran demarcación y diferenciación entre los dominios de los antiguos y nuevos métodos. Y esta política de los métodos se intensifica cuando los nuevos modelos de análisis entran en directa competencia con los focos sustantivos de las tradicionales y ortodoxas maneras de hacer investigación".

*initiatives are underway to develop social science methods for compiling and analysing digital data*⁶ (p. 8).

Frente a lo expuesto, el presente documento revisará de manera general los principales problemas implicados en este nuevo escenario, los cuales suponen a la vez determinados desafíos y oportunidades para las ciencias sociales, considerando que el abordaje de los desafíos desplegados por el Big Data es una condición necesaria para aprovechar las posibilidades que esta nueva forma de producción y tratamiento de datos ofrece. Al respecto, el interés subyacente de los investigadores es generar un debate y reflexión con miras a lograr una integración metodológica, mediante la cual los dispositivos que tradicionalmente sirven al Big Data puedan también servir a la ciencia e investigación social, formando parte de su repertorio metodológico.

Big Data y ciencias sociales: los problemas para los científicos de lo social

En el siguiente apartado se esbozan algunos de los principales problemas y desafíos que el fenómeno del Big Data presenta para las ciencias sociales y para los métodos convencionales de investigación social.

Compartir la producción de lo social

Actualmente, múltiples actores son capaces de producir y analizar datos sobre el comportamiento social, e inclusive hacerlo en mayor volumen y rapidez que las ciencias sociales. Esto constituye una amplificación y una diversificación de las formas de conoci-

⁶ “Varios centros de investigación social e iniciativas han tomado el desafío de la producción digital de información y sus métodos (...). En la académica, investigadores han adoptado los métodos de producción social desde lo digital (...). Nuestro punto [es] que una serie de iniciativas está en marcha para desarrollar métodos de las ciencias sociales para la recopilación y análisis de datos digitales”.

miento social, lo cual supone un proceso de democratización de la investigación social, en el sentido en que la interpretación de datos sociales deja de ser potestad de un campo disciplinar específico.

Este proceso de democratización genera una disputa respecto de la hegemonía de ciertos actores sobre la capacidad de ofrecer explicaciones de los fenómenos sociales. Desde esta perspectiva, Savage y Burrows (2014) plantean que:

*We can be certain. Big Data does challenge the predominant authority of sociologists and social scientists more generally to define the nature of social knowledge. It permits a dramatically increased range of other agents to claim the social for their own*⁷ (p. 5).

En efecto, las posibilidades de operación a gran escala, la estructura flexible y dinámica del Big Data respecto de la producción de datos, generan mecanismos y dispositivos que posibilitan la participación de nuevos actores en dicho proceso de producción. Esto, a su vez, permite que la accesibilidad de datos relativos a las relaciones sociales sea más amplia, y que especialistas de diversas áreas conozcan la naturaleza de los datos recolectados y se aventuren a ofrecer explicaciones sobre las realidades sociales que, se supone, son reflejadas desde los datos. De hecho, las redes sociales no requieren un cuestionario técnico o una entrevista semiestructurada para extraer información de sus usuarios. En contraste, el *retail* obtiene información a través de las transacciones diarias; el servicio de impuestos internos a través de las operaciones tributarias; los gobiernos mediante los registros en los programas de beneficios, etc.

La mayoría de estas operaciones se realiza sin la necesidad de la intervención de un científico social y, de hecho, “en el área

⁷ “Podemos estar seguros. El Big Data desafía la predominante autoridad de sociólogos y científicos sociales de manera más general para definir la naturaleza del conocimiento social. Lo cual permite un dramático aumento de otros agentes que reclaman lo social para ellos”.

comercial o gubernamental, diferentes expertos tales como ingenieros en computación, diseñadores de softwares o la emergente profesión de científicos de datos [*data scientist*], se están volviendo mediadores más prominentes” (Ruppert, Law y Savage, 2013, p. 13). En esta línea, existen especialistas en la producción de mecanismos sociotécnicos para la generación del conocimiento que, mediante la creación de algoritmos matemáticos, uso de minería de datos y producción de modelos complejos para la toma de decisiones, construyen formas de operar mediante desarrollos computacionales, que automatizan y establecen patrones de búsqueda para el uso y toma de decisiones en la categorización de los análisis. Estas operaciones están brindando un dinamismo inesperado al trabajo con los datos (Ruppert, 2011).

Por lo tanto, es posible aseverar que el campo del científico social, como agente por excelencia para dar cuenta de los datos sociales, se ha diversificado gracias al ingreso de otros actores provenientes de otras disciplinas en este nuevo contexto, quienes reclaman una nueva vocería como agentes legítimos en la producción y análisis de los datos sociales. Así, profesionales con capacidades de gestión en el área del Big Data, como informáticos, ingenieros o físicos, hacen su reclamación en el campo del conocimiento de lo social.

El involucramiento de estos nuevos actores trae aparejadas diversas implicancias para la comprensión de lo social. Una de ellas es el hecho de que los profesionales de otros campos cuentan con un arsenal de herramientas propias de su especialidad y ajenas a las competencias de los científicos sociales. Por lo tanto, el ingreso de una nueva gama de profesionales al campo del análisis social implica también el ingreso de prácticas y dispositivos con el potencial de ser adaptados al mismo, y pueden mostrar eficacia para dar cuenta del acontecer social. En este contexto, mediante la práctica y el trabajo multidisciplinario, los científicos sociales tienen la oportunidad de desarrollar nuevas perspectivas y adquirir competencias tradicionalmente asociadas a otras disciplinas y adaptarlas a su campo de estudio.

También es posible notar la posibilidad de nuevas áreas de estudio en el contexto sociológico. Al respecto, Prydmore y Lyon (2011) realizan un estudio empírico en el cual analizan los programas de marketing de lealtad. Sus conclusiones apuntan a que estos programas, que funcionan generalmente a través de un sistema de puntos y canje de recompensas, se constituyen en dispositivos de vigilancia contemporánea, gracias a los procesos y dispositivos de recolección de datos personales de los consumidores. Los autores señalan que el análisis de estos datos permite observar las prácticas de los consumidores, discernir sobre sus deseos, necesidades y tendencias, para luego orientarlos y modelar así el comportamiento de consumo. Luego, el comportamiento modelado sigue siendo rastreado y orientado permanentemente mediante el despliegue de ofertas y descuentos, especialmente seleccionados según las prácticas de compra detectadas para cada consumidor. Así, los profesionales del marketing van identificando patrones, gustos y preferencias que han resultado de las prácticas de modelamiento. En este panorama, la digitalización de datos y la representación digital del comportamiento del consumidor, mediante el rastreo y análisis de sus prácticas por operaciones de marketing de lealtad, constituye una forma de sociología comercial⁸.

En línea con lo anterior, dado que el estudio del comportamiento social mediante procedimientos de Big Data involucra una serie de dispositivos de uso cotidiano, tales como teléfonos móviles, dispositivos de navegación por internet, diferentes tipos de tarjetas, etc., los cuales no necesariamente se encuentran directamente relacionados al marketing o la economía, según nuestra visión, resulta posible sostener que el Big Data permite la emergencia de una forma distintiva de sociología, una suerte de “data sociología”. Esta nueva forma de hacer sociología vendría marcada por un importante corte físico empírico y estaría fuertemente centrada en las prácticas de captura masiva de datos y co-

⁸ Prydmore y Lyon (2011), tesis apoyada también por otros autores como Ellison y Burrows (2007) y Burrows y Gane (2006).

rrelación de factores o tipos de consumo. Sin embargo, creemos que una data sociología debe complementar –más que intentar explicar por sí misma– la comprensión de los fenómenos sociales. Por ello, bajo nuestro punto de vista, el Big Data se constituye en un conjunto de herramientas susceptibles y deseables de ser incorporadas al quehacer sociológico.

Cambio paradigmático

El modelo convencional del Big Data plantea una serie de cambios para la ciencia en general, dentro de los cuales destaca una dimensión paradigmática. Jym Gray (2009), experto en desarrollo informático, plantea que el Big Data, entendido como la ciencia intensiva de los datos, presenta un nuevo paso de la evolución del método científico. El autor señala que el método científico, en un primer estado o paradigma, resalta un componente empirista; mientras que, en una segunda etapa, de corte más teórico, se orienta sobre la base de modelos y generalizaciones. Un tercer paradigma del método científico viene marcado por la emergencia de una ciencia computacional que simula fenómenos complejos; finalmente, en una última etapa, marcada por el Big Data, prima en el método científico un carácter estadístico exploratorio, constituido por análisis de datos en profundidad, mediante procesos de minería y explotación intensiva de datos. Sobre este cuarto paradigma, Gray (2009) señala:

*Provides an integrating framework that allows the first three to interact and reinforce each other, much like the traditional scientific cycle in which theory offered predictions that could be experimentally tested, and these experiments identified phenomena that required theoretical explanation*⁹ (p. 177).

⁹ “Proporciona un marco integrador que permite a las tres etapas anteriores interactuar y reforzarse mutuamente, al igual que el ciclo científico tradicional, en el cual la teoría ofrece las predicciones que pueden ser probadas experimentalmente, y estos experimentos identifican fenómenos que requieren explicación teórica”.

Al respecto, es posible argumentar que la sinergia de las operaciones científicas de las tres primeras etapas confluyen y dan pie a la emergencia del cuarto paradigma. A su vez, esta cuarta etapa puede entenderse como una fase de integración paradigmática, ya que, lejos de negar los paradigmas anteriores, los incorpora. Desde esta perspectiva, Kitchin (2014) destaca tres características principales que definen al cuarto paradigma del método científico:

- 1) la capacidad de atender datos en grandes volúmenes, de manejarlos y producirlos en altas velocidades, con diversidad y variedad en su estructura y naturaleza;
- 2) hacer exhaustiva la captura a niveles de poblaciones generales, con una alta precisión y detalle, con capacidades relacionales para la homologación de bases de datos, y
- 3) su flexibilidad y capacidad de adaptación a nuevos campos, en conjunto con su rápida capacidad de expansión a nuevos dominios.

Estas posibilidades permiten que la producción científica orientada bajo este cuarto paradigma sea capaz de capturar dominios completos para el análisis y, como consecuencia, podría prescindir de teorías, modelos, hipótesis, etc. Además, en este paradigma el trabajo de producción científica se gesta bajo una visión que cree dejar de lado el sesgo y la influencia del científico, promoviendo un acercamiento inductivo desde los datos y valorando la dimensión empírica de los fenómenos.

Al alero de tal desarrollo tecnológico informático se empieza a hablar también de una ciencia social computacional, en la cual se articulan distintas fuerzas que, según Chang, Kauffman y Young (2014) *“include technological changes, the convergence of disciplines, and the availability of new tools and solution approaches for data analytics”*¹⁰ (p. 69). Este nuevo espacio de fuerzas permite

¹⁰ “Incluyen cambios tecnológicos, la convergencia de las disciplinas y la disponibilidad de nuevas herramientas y soluciones con enfoque en el procesamiento de datos”.

manejar información con una hasta ahora inusitada capacidad de amplitud, profundidad y escala de los fenómenos sociales.

En consideración a los diversos procesos y magnitudes de los fenómenos sociales asociados a la investigación y su integración metodológica desde un esquema clásico en las ciencias sociales, se genera, según McGrath (1994), el dilema de las tres puntas, que busca atender la convergencia entre realismo, control (precisión) y generalidad de los fenómenos sociales en la investigación. Este dilema plantea que para la mayoría de las investigaciones se genera la disyuntiva de intentar maximizar el escalamiento de los datos, en relación a los actores a estudiar, los comportamientos que se analizan y su contexto o lugar de producción de los fenómenos.

Debido a que es prácticamente imposible para la investigación la integración de los tres elementos del dilema, por la capacidad material y métodos de análisis asociados desde una visión ortodoxa, es que pese a que se establezcan los más estrictos controles de calidad o estrategias de análisis y generalidad, siempre la realidad supera las posibilidades de acceso al conocimiento social cumpliendo el rigor de estos tres ideales¹¹. Dado que el Big Data y sus dispositivos de acceso, instrumentalización y análisis generan un nuevo contexto de facilidades para la investigación, el dilema planteado proporciona una nueva forma de abordar esta problemática, debido en gran parte a la eficiencia que los nuevos métodos de investigación otorgan, en virtud del fluido y dinámico acceso a los datos, así como a la magnitud y extensión de grupos poblacionales completos. Esto es posible debido a la reducción de costos asociados al uso de nuevas tecnologías y a la convergencia que los distintos métodos de las disciplinas so-

¹¹ Por ejemplo, desde un estudio de características experimentales se incrementa la **precisión** con la cual uno mide los comportamientos y controla las variables; sin embargo se disminuye el componente natural del **realismo** de esa interacción. En este caso, se privilegia la precisión por sobre la realidad. A su vez, se reduce el rango de la población a la cual pueden ser **generalizadas** dichas conclusiones del estudio.

ciales han incorporado a sus procesos de investigación, gracias a las ciencias computacionales y a la internet (Chang, Kauffman y Young, 2014).

Cabe señalar que, pese a que este nuevo paradigma se ve como un ideal en el desarrollo científico, aún el dilema persiste desde aquellas otras realidades que no pueden ser rastreadas o capturadas por estos tipos de dispositivos, por lo cual esta visión y paradigma en su acceso al conocimiento e imputación del campo de lo social está sujeto directamente a las condiciones materiales de producción asociadas a él.

Propiedad de la producción científica

Una de las principales críticas a este nuevo paradigma son los posibles sesgos derivados del entramado político que subyace a su ontología. Estas críticas surgen dado que la producción de datos mediante operaciones y dispositivos de Big Data generalmente obedecen a intereses particulares orientados bajo modelos empresariales o privados, cuyos fines apuntan a generar conocimiento instrumentalizado para la búsqueda de resultados de gestión, con miras a la optimización de su rentabilidad. Al respecto, si bien no es posible aseverar que la investigación social tradicional es neutra o apolítica –y a menudo ha sido impulsada por organismos privados, en términos de la magnitud de las operaciones, de los recursos para la confección de dispositivos, de la capacidad de producción de datos, de la velocidad de procesamiento de datos, etc.–, la brecha en la capacidad de la investigación social público/académica y la investigación empresarial/privada nunca había sido tan amplia.

Dado que muchas veces los datos producidos a partir de procedimientos del Big Data son de tipo privado, el acceso a ellos se realiza bajo una dinámica comercial. Esto implica que el investigador deba incurrir en gastos financieros adicionales para solventar el precio que el poseedor de los datos pueda fijar. Por otra parte, si estos datos son proporcionados sin intercambio mo-

netario, es probable que los mismos se encuentren sujetos a derechos de propiedad que los protegen como un producto patentado. Además, es usual que los propietarios de estos datos solo faciliten la información para realizar investigaciones que vayan en interés de sus propias organizaciones, en las que comúnmente los investigadores deban firmar cláusulas de confidencialidad y no puedan utilizar los datos proporcionados para otros fines que no sean los estipulados por el contratante.

Al respecto, Housley et al. (2014) indican que existen tres factores que hacen difícil que esta relación “investigador/datos privados” se desarrolle de manera sustentable. La primera se relaciona con que algunas investigaciones necesitan más datos que los disponibles de manera pública y libre, lo cual tiene altos costos asociados y el posible avance a desarrollar queda supeditado a la capacidad negociadora de los agentes (arreglos individuales). La segunda, indica que los investigadores quedan a disposición del proveedor de la información, quien eventualmente puede efectuar modificaciones a los acuerdos pactados y afectar el curso de las investigaciones. En tercer lugar, muchas veces los arreglos del uso de los datos impiden que estos sean compartidos con otros investigadores o con la comunidad científica, con lo cual se inhibe la necesidad científica de probar y validar los resultados. Este escenario genera que la integración de bancos de datos producidos por diferentes procedimientos de Big Data sea difícil de lograr.

A lo anterior debe añadirse la diversidad de información y formatos existentes de los diferentes bancos de datos. En efecto, dado que cada institución cuenta con sus propias operaciones, procedimientos y dispositivos de recolección y captura de datos, la unificación y homologación de volúmenes importantes de datos puede resultar trabajosa y costosa, inclusive con la ayuda de softwares específicos de gestión de datos.

En efecto, es posible argumentar que el campo de prácticas del Big Data se presenta atomizado, debido, por una parte, a una

lógica privado/empresarial subyacente, reacia a compartir información y/o a permitir su uso para fines que no benefician sus intereses económicos; y, por otra, a la diversidad de modos y medios de producción operante. Esto complejiza en exceso las tareas de integración de datos en distintos campos de producción, lo cual implica que los análisis sociales realizados a partir de procesos de Big Data como único input, corran el riesgo de resultar reduccionistas y descontextualizados. Al respecto Kitchin (2014) declara:

The result is an analysis that is reductionist, functionalist and ignores the effects of culture, politics, policy, governance and capital (...). These fallacious notions have gained some traction, especially within business circles, because they possess a convenient narrative for the aspirations of knowledge orientated businesses¹² (p. 5).

Según señala el autor, los análisis sociales basados únicamente a partir de procedimientos de Big Data tienden a empaacar al individuo en su dimensión práctica y terminan ofreciendo una versión comprimida del consumidor, cliente o usuario. El proceso de empaque del consumidor es descrito por Ariztía (2013) en un estudio empírico realizado en diferentes agencias de publicidad. En su investigación, Ariztía detalla cómo los consumidores son enactados por los profesionales del marketing mediante sucesivos procesos de evaluación y cualificación, que reducen al consumidor a un conjunto de cualidades altamente específicas, especialmente seleccionadas para servir a los objetivos de la campaña publicitaria. Luego, una versión empacada e idealizada del sujeto es presentada a la audiencia como un consumidor real en las campañas que son emitidas al aire.

¹² “El resultado es un análisis que es reduccionista, funcionalista e ignora los efectos de la cultura, los políticos, la política, la gobernabilidad y el capital (...). Estas nociones son falsas y han ganado atracción, especialmente en círculos empresariales, porque ellos poseen una narrativa convincente de sus aspiraciones orientadas por los negocios”.

En conclusión, la falta de integración de diversos bancos de datos de diferentes campos, la alta especificidad de cada sistema de recolección de datos y la centralidad de lo empírico de los análisis sociales realizados bajo las directrices de los procedimientos del Big Data pueden derivar en análisis sin sustrato ni contexto social de anclaje.

Hegemonía de lo empírico, reduccionismo y descontextualización del sujeto

En *The coming crisis of empirical sociology*, Savage y Burrows (2007) plantearon que se avecinaba una crisis en el campo empírico de la sociología, ya que se advertía la emergencia de nuevas prácticas asociadas al modelo económico capitalista que orientaban nuevas formas de investigación, las cuales alteraban dramáticamente los procesos tradicionales de producción de datos y de gestión del conocimiento de lo social. En la misma línea, Rupert, Law y Savage (2013) citan a Thrift (2005), quien señala que “*new methods became fully enmeshed in the circuits of ‘knowing capitalism’, in which the systematic gathering of information about costumers, clients, employees and competitors became routine to corporate strategy*”¹³ (p. 9).

El capitalismo cognoscente (Thrift, 2005) representa una nueva forma de capitalismo que ha hecho del conocimiento el centro de la operación de los actores económicos, y una de sus principales características es su habilidad para gatillar cambios en el mismo sistema capitalista, incrementando su capacidad de adaptación mediante la incorporación e integración de las críticas que le son imputadas a su propia operación, como resultado de una actividad reflexiva. Un buen ejemplo de la acción del “*knowing capitalism*” es la estrategia corporativa de manufacturar y comer-

¹³ “Los nuevos métodos se imbricaron en los circuitos del capitalismo cognoscente, en el cual la recolección sistemática de información sobre los servicios, clientes, empleados y competidores se convirtió en una rutina en las estrategias corporativas”.

cializar *merchandising* de objetos que históricamente simbolizaron una crítica al mismo sistema capitalista, como artículos del activista político Ernesto “Che” Guevara.

En dicho marco, el Big Data se presenta profundamente imbricado con el *knowing capitalism*, y se constituye como un instrumental cada vez más utilizado por esta nueva forma de capitalismo. Esto permite observar el problema empírico de las ciencias sociales señalado por Savage y Burrows (2007), dado que el sujeto, en tanto objeto de conocimiento, queda reducido a sus dimensiones prácticas. En efecto, el tipo de análisis empleado por la investigación social bajo procedimientos de Big Data, enmarcada en el capitalismo cognoscente, descontextualiza al sujeto en cuanto lo empaca en acciones altamente específicas, visualizadas a partir de registros transaccionales y/o contables. Por lo tanto, el conocimiento sobre los sujetos queda restringido a las acciones que los dispositivos de registro han logrado capturar.

Como consecuencia, estos tipos de análisis tienden a comprimir la opinión del sujeto en los contextos de producción de esa información. Los niveles de gestión de la información, en la operación dinámica de captura y recolección de datos, se manejan “en el nivel individual sobre lo que la gente es, hace y dice” (Couldry y Powell, 2014, p. 1).

Esto aplica inclusive a las redes sociales, ya que la visibilidad o pérdida de anonimato frente a la multiplicidad de agencias vinculadas a través de dispositivos como Facebook, Twitter u otras plataformas de interacción, ejercen un poder agente sobre el sujeto opinante. En ese sentido, las prácticas y opiniones realizadas a través de redes sociales –es decir, publicadas– consisten en una *performance* a través de la cual el sujeto proyecta una serie de cualidades a las cuales se encuentra vinculado, pero en ningún caso puede considerarse como *la* identidad del sujeto (Goffman, 1971). Al respecto, Mol (2002), en su estudio empírico sobre la práctica médica, señala: “*the identity people*

perform is not deep, it is a mere performance"¹⁴ (p. 35). En contraste, los dispositivos tradicionales de investigación social, como la encuesta, entrevista o *focus group*, ofrecen un nivel de anonimato y confidencialidad mayor que las redes sociales, por lo cual, en el momento de recolección de datos, la agencia es distribuida de manera diferente, lo que posibilita la visibilización de mixturas más profundas.

A pesar de que los dispositivos de Big Data inscritos en redes de interacción múltiple (redes sociales) ofrecen un espacio mayor para la expresión de la subjetividad, en general la obtención de los insumos para el análisis social en los procedimientos de Big Data queda sujeta a las posibilidades estructurales del sistema de producción de datos, ya sea mediante los registros de prácticas de consumo con medios digitales de pago, antecedentes y perfiles de datos personales en bases de datos comerciales, o las opiniones vertidas en 140 caracteres de Twitter.

Lo anterior conduce a una forma de conocimiento social orientado desde una visión pragmática, en la cual, según Kitchin (2014), *"the sort of person you are is largely determined by your social context, so if I can see some of your behaviors, I can infer the rest, just by comparing you to the people in your crowd"*¹⁵ (p. 8). Tras esto, la recolección de datos empíricos a través de dispositivos de Big Data articulan una nueva definición de la naturaleza del conocimiento social. Ante esto, Ruppert (2011) señala:

Once joined up the different categories and measurements of subjects assembled in each database –biographical, biometric and transactional– can then be compared, combined and

¹⁴ "La identidad performada de la gente no es profunda, es una mera puesta en escena".

¹⁵ "La clase de persona que eres está determinada en gran medida por el contexto social, por lo que, si puedo ver alguno de sus comportamientos, puedo inferir el resto, comparándote con la gente de la multitud".

*reassembled in myriad ways to produce population metrics*¹⁶
(p. 219).

Por lo tanto, es posible argumentar que el tipo de análisis social empleado por el capitalismo cognoscente, mediante dispositivos de Big Data como sus instrumentos de inquisición, gatillan una crisis en el campo empírico de las ciencias sociales en dos sentidos: en primer lugar, la potente maquinaria de captura de información, generación de resultados y constante actualización de datos, hace ver los procedimientos tradicionales de recolección de información de las ciencias sociales como algo obsoleto; en segundo lugar, esta nueva forma de recolección y operación de datos desafía la autoridad de los científicos sociales para emitir juicios sobre lo social.

Dado que los resultados de los análisis realizados bajo el prisma del capitalismo cognoscente, a la luz de las herramientas facilitadas por el Big Data, presentan a un sujeto descontextualizado y lo reducen a un ente emisor de acciones registrables, los dispositivos del Big Data proveen versiones de la realidad que no logran ser enactadas adecuadamente, dado que no encuentran un correlato en la diversidad de interacciones de la vida cotidiana de los sujetos. En otras palabras, el dato práctico, el registro unitario de la acción y la agregación aditiva de estos actos no garantizan el ensamblaje de una realidad adecuadamente enactada. En este sentido, si bien las ciencias sociales siempre han producido versiones reducidas o empacadas de los fenómenos sociales (Law, 2004), en la mayoría de los casos al menos consideran la presencia o vinculación de mixturas subyacentes a partir del cual emergen los fenómenos. En contraste, los procedimientos de ensamblaje del Big Data tienden a omitir la existencia de esas mixturas y empacan a un sujeto aislado y descontextualizado,

¹⁶ “Una vez unidas las diferentes categorías y las mediciones de los sujetos ensamblados en cada base de datos –biográfica, biométrica y transaccional–, entonces se pueden comparar, combinando y reensamblándose en miles de formas que producen métricas de análisis de la población”.

considerando que lo social es constitutivo de la agregación de datos empíricos registrados por los usuarios en cada una de sus plataformas o dispositivos.

A modo de ejemplo, resulta interesante el trabajo empírico de Sunderland y Denny (2011), quienes, a partir de una metodología de estudio de caso, examinaron las estrategias de segmentación de consumidores realizadas por las empresas mediante la recopilación del Big Data. En su investigación, las autoras demuestran cómo el consumidor objetivo hacia el cual se orientaba la empresa que constituía su caso de estudio, definido por la misma compañía a través de la colección de datos digitales, era un empaque tan específico y reducido que “no fue posible encontrar en el campo los consumidores que la compañía pensó que había identificado en las bases de datos” (Sunderland y Denny, 2011, p. 12). Esto permite argumentar que el correlato entre el consumidor digital, enactado por procedimientos de Big Data, y el consumidor “real”, que vive su vida cotidiana en una serie de relaciones diversas y heterogéneas, suele ser débil. Este desencaje entre el consumidor empaquetado y el consumidor real se debe a la excesiva reducción de atributos a la cual son constreñidos los consumidores, los cuales son aislados y comprimidos en un paquete de cualidades altamente específicas, definidas por los mismos *data scientists* en sus operaciones de producción de conocimiento.

En conclusión, la estructura de los dispositivos que sirven al Big Data muestra un sujeto reducido a la unidad de sus actos, lo cual permite argumentar que en este modo de producción de conocimiento subyace una visión individualista sobre la sociedad, sentada en una premisa micro del orden social. Según lo anterior, un análisis social basado únicamente en datos recopilados a través de los dispositivos que sirven al Big Data, sin arreglo teórico o triangulación con otras técnicas de investigación, tiende a comprender “lo social” como una agregación de individuos y una adición de sus acciones individuales. Por lo tanto, considerando tanto las potencialidades del conocimiento producido a través de las prácticas del Big Data como sus eventuales riesgos, se ad-

vierte la conveniencia de triangular sus prácticas y dispositivos con el arsenal de la ciencia social en general y de la sociología en particular.

Sinergia entre sociología y Big Data

La creencia de que es posible explicar las relaciones sociales a partir de algoritmos matemáticos que contribuyen a registrar transacciones, compras, datos geográficos, opiniones vertidas en redes sociales, entre otro tipo de información, radica en el supuesto de que la capacidad explicativa que poseerían estos datos empíricos recolectados a través de los dispositivos del Big Data sería suficiente para dar cuenta de los fenómenos sociales.

Dicha premisa, junto con prescindir de los científicos sociales para ofrecer explicaciones sobre lo social, prescinde además de los elementos teóricos sobre los cuales se realiza la investigación social tradicional. Ante esto, autores como Chris Anderson, citado por Kitchin (2014), apuntan a que el Big Data representa el fin de la teoría social y un triunfo del pragmatismo hacia lo empírico:

*Ensemble approaches signal a new era of knowledge production characterized by 'the end of theory'. In a provocative piece, Anderson (2008) argues that 'the data deluge makes the scientific method obsolete'; that the patterns and relationships contained within Big Data inherently produce meaningful and insightful knowledge about complex phenomena*¹⁷ (p. 3).

Según señala el autor, el Big Data supone el fin de la teoría, debido a la capacidad explicativa que poseen los datos recolectados, los cuales provienen desde lo empírico y logran dar cuenta

¹⁷ "El Big Data, las nuevas maneras de análisis de datos y los enfoques de ensamblajes son señal de una nueva forma de producir conocimiento, caracterizada por el fin de la teoría. En un artículo provocador, Anderson (2008) argumenta que 'el diluvio de datos hace al método científico obsoleto', los patrones y relaciones contenidas junto al Big Data inherentemente producen conocimiento significativo y detallado sobre fenómenos complejos".

sin especulaciones y alusiones heurísticas a “cómo es la realidad”. Ante esto, el ejercicio creativo desde lo teórico ve mermado su crédito para dar explicaciones satisfactorias respecto de lo social frente a los científicos de datos, quienes, a partir de su información empírica, ofrecen nuevas explicaciones sobre “lo que la sociedad realmente hace”.

No obstante, a partir de lo mencionado en los apartados anteriores, el ingreso de estos nuevos actores –científicos de datos– y dispositivos al campo de la investigación social puede generar una proliferación especulativa sobre explicaciones espurias de lo social, carentes de sustento y desconectadas del campo en el cual los sujetos se desenvuelven. Aquello tiene implicancias políticas, puesto que las explicaciones que se ofrecen sobre los fenómenos sociales contribuyen a enactar cierto tipo de realidad en desmedro de otras realidades posibles.

Siguiendo las ideas de Law (2004), dar cuenta adecuadamente de la realidad social implica tanto considerar las presencias –es decir, los elementos visibles ensamblados en la explicación ofrecida– como las ausencias manifiestas, y también las ausencias por otredad. Esta consideración es necesaria, dado que las relaciones entre los elementos incluidos, los excluidos deliberadamente y los excluidos inconscientemente de la explicación ofrecida por el analista, influyen en cómo se articulan los vínculos que ensamblan la realidad implicada en la explicación ofrecida por el analista, que ha resultado enactada a través del proceso de investigación y es presentada como los resultados del estudio.

Por ende, la invisibilización de determinadas realidades en favor de otras implica un trasfondo político. Esta situación, a su vez, plantea un dilema ético que no puede ser obviado, dada la capacidad performativa de los resultados de las investigaciones sociales, por cuanto contribuyen a un conocimiento que moviliza a la acción y aporta información para la toma de decisiones.

En tal panorama, es posible aseverar que una mirada complementaria entre procedimientos de Big Data, conocimiento socio-

lógico y trabajo multidisciplinario puede nutrir ambos campos científicos. Mientras la sociología advierte a los *data scientists* sobre las implicaciones del reduccionismo y la descontextualización del sujeto, y provee orientación en la interpretación de las prácticas para evitar reducciones espurias, el sociólogo se nutre de un conocimiento nuevo, accede a bancos de datos de naturaleza distinta a los cuales solía trabajar tradicionalmente y adquiere herramientas para desarrollar nuevas competencias que lo conviertan en un *data sociologist*.

En la actualidad el *open data movement* está movilizando a distintos agentes en pos de democratizar el espacio y producción de los datos de manera pública. Un ejemplo de esto es el proyecto COSMOS¹⁸, gestado desde un enfoque multidisciplinario, como una propuesta en que las ciencias sociales y otras disciplinas abogan por el uso extensivo y a gran escala de la información pública disponible. Un ejemplo de sus actividades de análisis se observó durante los Juegos Olímpicos del año 2012. Con ocasión de la competencia de los 10.000 metros planos, en la cual el atleta local Mohamed Farah pasó de estar en los últimos lugares a ganar la medalla de oro, el proyecto logró dar cuenta de datos emitidos en vivo y en directo por los fanáticos del deporte que eran usuarios de la red twitter en el Reino Unido, y que emitieron comentarios en dicha red social mientras se desarrollaba la competencia. El análisis de los comentarios logró evaluar los sentimientos positivos/negativos y sus variaciones durante los hitos ocurridos en el transcurso de la competencia y sus relaciones a niveles de variables con el género.

Tal como el proyecto COSMOS, crecientemente se están desarrollando iniciativas de integración metodológica en diversas materias, lo cual abre las posibilidades de gestión del conocimiento guiadas por principios de carácter participativo, con orientaciones hacia el libre acceso, habilitación para el uso públi-

¹⁸ Collaborative Online Social Media Observatory: <https://www.cs.cf.ac.uk/cosmos/>

co y la modificación de datos *online*. Al mismo tiempo, las redes sociales y plataformas web generan datos públicos susceptibles de ser recolectados mediante procedimientos sistemáticos. Otra alternativa es que, a partir de la experiencia observada, los científicos sociales adquieran o desarrollen sistemas y dispositivos de tipo Big Data, que permitan recolectar información relativa a las prácticas de los individuos en grandes volúmenes. Los softwares de recolección de tendencias en redes sociales constituyen un ejemplo de esta última posibilidad.

Conclusiones

Quienes hablan de Big Data, a menudo suelen recalcar lo fructífero de este conjunto de prácticas y dispositivos para la producción del conocimiento social; sin embargo, con frecuencia se omite lo valioso que puede ser el conocimiento proveniente de la sociología para optimizar la producción de conocimiento del Big Data. En esta línea, el presente documento ha revisado una serie de implicancias que el fenómeno del Big Data conlleva para la investigación social, el cual se expresa en la emergencia de un nuevo paradigma, que abre un nuevo espectro de posibilidades y oportunidades para el fortalecimiento y la innovación en la investigación social, pero que a su vez, plantea importantes desafíos para las ciencias sociales, en tanto viene a disputar el campo de producción e interpretación de datos sobre el comportamiento humano, donde las ciencias sociales tradicionalmente gozaban de una situación hegemónica.

El Big Data como una nueva gama de posibilidades de desarrollo científico social, además de ser un fenómeno de interés para la sociología en tanto objeto de estudio e indagación sociológica de sus capacidades agentes, genera posibilidades de análisis pocas veces vistas en investigación social, lo cual se expresa en al menos tres términos:

- Disponibilidad de extensos volúmenes de datos.

- Accesos a datos sobre diversos tipos de interacciones sociales, con la posibilidad de acceder a información difícil de obtener mediante los métodos de recolección de información tradicionales.
- Información sobre lo que los individuos efectivamente hacen (no solo sobre lo que dicen que hacen).

Entre los desafíos para la ciencia social, observamos la restricción al acceso y uso de los datos; el elevado nivel técnico que supone la operación de los dispositivos que sirven al Big Data; la pérdida de exclusividad de la producción de lo social, y dilemas éticos asociados a la producción y uso de información.

El acceso a dichos datos no siempre es expedito para el investigador, dadas las lógicas comerciales que gobiernan la producción y análisis de Big Data. Además, los contenidos orientados a la adquisición de competencias para poner en marcha procedimientos de Big Data, tales como programación, informática y manejo de paquetes estadísticos *ad hoc* al volumen de información esperada, suelen estar fuera de los programas de formación que dictan las escuelas de sociología en la actualidad, lo cual supone barreras importantes para la integración metodológica entre ciencias sociales y Big Data, ya que muchas veces los investigadores no cuentan con las habilidades técnicas idóneas para el profundo manejo computacional que requiere el uso y análisis junto al Big Data.

Como conclusión, se advierte que, pese a las amplias posibilidades de conocimiento y lo provechosos que pueden resultar los datos generados por los dispositivos que sirven al Big Data en manos de los científicos sociales, las limitantes de esta forma de conocimiento, en tanto estructura privada de producción de datos, condicionan un despliegue extensivo y colaborativo de las investigaciones. Esto dificulta la validación de dichos resultados al interior de las comunidades científicas, la difusión del conocimiento generado para contribuir a la toma de decisiones y el desarrollo de competencias de los científicos sociales, las cuales son aristas fundamentales para el desarrollo de la ciencia social.

Por otra parte, la proliferación de datos producidos por los dispositivos del Big Data ha facilitado el ingreso de nuevos agentes, quienes desafían la hegemonía de los científicos sociales en dos frentes: en la potestad de producir datos sobre las relaciones sociales, por un lado, y en la hegemonía de la interpretación de datos y la autoridad sobre emisión de juicios sobre “lo social”, por otro.

Tanto la forma de producción de datos que supone el Big Data como la emisión de juicios sobre lo social plantean dilemas éticos para la investigación social. El primero implica un problema ético en términos de transparencia respecto del sujeto investigado, puesto que los individuos no siempre están conscientes de que sus datos son registrados y, por otra parte, muchas veces se utilizan los datos de sus acciones sin su consentimiento. El segundo punto, en tanto, supone el riesgo de que agentes que emiten juicios sobre lo social, desprovistos de herramientas conceptuales para explicar los fenómenos sociales que intentan explicar, pueden contribuir a visibilizar ciertas realidades que resultan enactadas en desmedro de otras. Aquello implica un dilema ético, puesto que determinadas realidades pueden favorecer ciertos intereses y, por lo tanto, pueden ser fortalecidas mientras se invisibilizan otras. Ante este último punto, la solidez brindada por el acervo teórico de las disciplinas ligadas al área social se alza como una de las principales ventajas de los científicos sociales frente a los científicos de datos.

El escenario descrito muestra el establecimiento de una relación asimétrica entre los científicos sociales tradicionales y los nuevos productores de los datos sociales, cuya forma de relacionarse es necesario auspiciar, en miras de traducir los desafíos presentados por el Big Data en oportunidades para la ciencia social.

Al respecto, nuestra visión apunta a que la respuesta más plausible para que la ciencia social pueda abordar los desafíos impuestos por este nuevo paradigma es el fomento del trabajo mancomunado y multidisciplinario que incorpore informáticos, diseñadores de softwares y programadores, y que permita que la

ciencia social incorpore, reformule y adapte sus formas de investigación a las técnicas proporcionadas por los dispositivos de Big Data, al mismo tiempo que se reformula a sí misma para lograr un adecuado acoplamiento y adaptación al nuevo escenario que la emergencia del Big Data ha desplegado en el escenario de la producción de lo social.

El proceso de adaptación de la ciencia social a este nuevo escenario pasa por un cambio en las prácticas y aspectos formativos. En efecto, si la ciencia se comprende como un resultado práctico de un conjunto de operaciones y esfuerzos organizativos (Latour, 2008), una serie de modificaciones en el ensamblaje de los vínculos que enlazan esas operaciones puede permitir la reducción de la asimetría. En este sentido, la adquisición de competencias técnicas *ad hoc* al Big Data por parte de los científicos sociales se configura como punto de inflexión para la reducción de esta asimetría. Por lo tanto, gran parte de las prácticas que debieran evaluarse en miras de comenzar a producir *social data scientists* debiera centrarse en las aulas, en el aprendizaje experiencial y en el trabajo mancomunado interdisciplinario.

En el caso específico de la sociología, el presente documento plantea que el balance de relaciones decantaría en la emergencia de una nueva forma de sociología: una “data sociología”. Este nuevo campo, vendría a “sociologizar” de manera incluyente la experiencia de otros campos (informática, ingeniería, etc.) y también de otras áreas de su propio campo (teorías y métodos convencionales).

En nuestra mirada, el Big Data y por consiguiente la data sociología se vuelven menos solventes si reclaman para sí la exclusividad de la producción de lo social y, además, difícilmente podrían lograrla, dada su tendencia al reduccionismo y descontextualización. Por lo tanto, estas actividades no vienen a reemplazar la forma tradicional de investigación social, sino a enriquecerla. En efecto, la interacción constante entre la combinación de saberes multidisciplinarios, el uso de distintos métodos, la triangulación de diferentes técnicas y bases teóricas guías, son todas actividades

necesarias para operar una red de producción de conocimiento sobre lo social que logre captar e incorporar aquellas mixturas que, hasta el momento, son invisibilizadas en muchas ocasiones por los mecanismos y dispositivos ligados al Big Data.

Referencias

- Ariztía, T. (2013). Unpacking insight: how consumers are qualified by advertising agencies. *Journal of consumer culture*, 15(2), 143-162.
- Burrows, R. & Gane, N. (2006). Geodemographics, software and class. *Sociology*, 40, 793-812.
- Callon, M. (2006). What does it mean to say that economics is performative? *CSI Working Paper Series*, 5. Recuperado en mayo de 2016 de: <https://halshs.archives-ouvertes.fr/halshs-00091596/document>
- Chang, R., Kauffman, R. & Young, K. (2014). Understanding the paradigm shift to computational social science in the presence of big data. *Decision Support Systems*, 63, 67-80.
- Couldry, N. & Powell, A. (2014). Big data from the bottom up. *Big Data & Society*, 2. Recuperado en mayo de 2016 de: <http://bds.sagepub.com/content/spbds/1/2/2053951714539277.full.pdf>
- Ellison, N. & Burrows, R. (2007). New spaces of (Dis)engagement? Social politics, urban technologies and the rezoning of the city. *Housing studies*, 22, 295-312.
- Gofman, E. (2001). *La presentación de la persona en la vida cotidiana* (1ª ed. en esp., 3ª reimp.). Buenos Aires: Amorrortu Editores.
- Gray, J. (2009). The fourth paradigm. Microsoft Corporation, USA. Jim Gray on escience: A Transformed Scientific Method, xvii-xxxi; Jim Gray's Fourth Paradigm and the Construction of the Scientific Record, 177-183.
- Housley, W., Procter, R., Edwards, A., Burnap, P., Williams, M., Sloan, L., Rana, O., Morgan, J., Voss, A. & Greenhil, A. (2014). Big Data and the sociological imagination: A collaborative response. *Big Data & Society*, Aug, 2. Recuperado en mayo de 2016 de: <http://bds.sagepub.com/content/spbds/1/2/2053951714545135.full.pdf>.

- Kitchin, R. (2014). Big Data, new epistemologies and paradigm shifts. *Big Data & Society*, Jun, 1. Recuperado en mayo de 2016 de: https://people.cs.kuleuven.be/~bettina.berendt/teaching/ViennaDH15/kitchin_2014.pdf
- Latour, B. (2008). *Reensamblar lo social, una introducción a la teoría del actor-red*. Buenos Aires: Ediciones Manantial.
- Law, J. (2004). *After method: mess in social science research*. Londres: Routledge.
- Mcgrath, J. (1994). *Methodology matters: Doing Research in the Behavioral and Social Sciences*. USA: University of Illinois.
- Mol, A. (2002). *The Body Multiple: Ontology in Medical Practice*. USA: Duke University Press.
- Prydmore, J. & Lyon, D. (2011). Marketing as surveillance: Assembling Consumer as Brands. In D. Zwick & J. Cayla (eds.). *Inside Marketing: Practices, Ideologies, Devices* (pp. 15-136). New York, USA: Oxford University Press.
- Ruppert, E. (2011). Population objects: Interpassive Subjects. *Sociology* 45(2), 218-233. Recuperado en mayo de 2016 de: <http://soc.sagepub.com/content/45/2/218.full.pdf+html>
- Ruppert, E., Law, E. & Savage, M. (2013). Reassembling Social Science Methods: the challenge of digital devices. *Theory, Culture & Society*, 30(4), 22-46. Recuperado en mayo de 2016 de: <https://research.gold.ac.uk/7978/1/DigitalDevices-DistCopyJSept2013.pdf>
- Savage, M. & Burrows, R. (2007). The coming crisis of empirical sociology. *Sociology*, 41(5), 885-899.
- Savage, M. & Burrows, R. (2014). After the crisis? Big Data and the methodological challenges of empirical sociology. *Big Data & Society*, Jun, 1(1). Recuperado en mayo de 2016 de: <http://bds.sagepub.com/content/spbds/1/1/2053951714540280.full.pdf>
- Sunderland, L. & Denny, R. (2011). Consumer segmentation in Practice: An Ethnographic. In D. Zwick & J. Cayla (eds.). *Inside Marketing: Practices, Ideologies, Devices* (pp.61-137). New York. USA: Oxford University Press.
- Thrift, N. (2005). *Knowing capitalism*. California USA: SAGE Publications Ltda.