

BIG DATA Y MEMORIA DIGITAL

Claves para su exploración e investigación desde las ciencias sociales

Ignacio Uman

Departamento de Posgrado de la Universidad de Ciencias Empresariales y Sociales (UCES).

ignacio@umancomunica.com.ar

Recibido: 19 de abril de 2018

Aceptado: 20 de junio de 2018

Resumen

Este artículo pretende introducir al lector al análisis del concepto de Grandes Datos (Big Data) desde una perspectiva crítica.

Para desarrollar este recorrido, se recurre a nociones propias de los estudios sociales de la ciencia y la tecnología, estudios de la innovación e historia de la tecnología. Al analizar esta faz socio-técnica, se tienen en cuenta los debates y discusiones que aún están en juego en torno al complejo concepto de Big Data. Se esbozan los principales ejes de análisis teórico-conceptual que puedan situar e introducir al investigador de las ciencias sociales en la problemática de los grandes volúmenes de datos y que tengan en cuenta su abordaje desde múltiples disciplinas, posibilitando entrecruzamientos productivos entre lo social y lo tecnológico. Para ello se exploran algunos casos, proyectos e iniciativas locales donde se ha estudiado e investigado esta temática con profundidad.

De este modo, se propone desarrollar una serie de metodologías para explorar e investigar el fenómeno de Big Data desde las ciencias sociales y humanidades, en búsqueda de una fuerte conexión con otras ciencias tales como las ciencias exactas, la informática y la ingeniería, alcanzando finalmente un enfoque transdisciplinar.

Palabras clave: Big Data, Ciencias Sociales, Memoria Digital.

BIG DATA AND DIGITAL MEMORY

Keys for exploration and research from social sciences

Abstract

This article aims to introduce the reader to the analysis of the Big Data concept from a critical perspective.

To develop this journey we use precise notions from the Social Studies of Science and Technology, the Studies of Innovation and the History of technology. We analyze the problem from this socio-technical point of view, considering the debates and

discussions that are still at stake around the complex concept of Big Data. The outline of theoretical-conceptual analysis can situate and introduce the social researcher into the problem of large volumes of data and large-scale information. It is an approach from multiple disciplines, enabling productive cross-linking between the social field and the technological field. For this reason, we explore some cases, projects and local initiatives where this subject has been studied and depth-researched.

By this way, the article develops a series of methodologies to explore and investigate the Big Data phenomenon from the Social Sciences and Humanities, in search of a strong connection with other sciences such as Exact Sciences, Computer Science and Engineering, finally reaching a transdisciplinary approach.

Keywords: Big Data, Social Sciences, Digital Memory.

En torno al concepto de Big Data

El almacenamiento de información resulta una práctica característica de la humanidad. Desde el propio texto y el antiguo manuscrito, hasta las imágenes y los datos actuales en formato digital, cada tipología de información constituye una tecnología específica de cada momento histórico. En la actualidad, lo que se conoce como Grandes Datos (en adelante Big Data) es una de las tecnologías más importantes para almacenar grandes volúmenes de datos. El término surgió como un concepto recurrente para describir conjuntos de datos cuyo tamaño está más allá de la capacidad de las bases de datos tradicionales para capturar, almacenar, administrar y analizar. Sin embargo, el alcance del término no sólo se refiere a los datos en sí, sino también a un conjunto de tecnologías que capturan, almacenan, administran y analizan, grandes y variables colecciones de datos para resolver problemas complejos.

Big Data surge como nueva fase del paradigma intensivo en comunicación e información, que emerge a partir de la revolución tecnológica iniciada en la década de 1970.

En efecto la primera vez que se registra el uso del concepto es en octubre de 1997, cuando Michael Cox y David Ellsworth publican el artículo “Application-controlled demand paging for out-of-core visualization” (Cox, Ellsworth, 1997). Estos investigadores de la NASA afirmaron que el ritmo de crecimiento de los datos empezaba a ser un problema para los sistemas informáticos, ya que por ese entonces los conjuntos de datos eran lo suficientemente grandes para ocupar las capacidades de la memoria principal, y de los discos locales y remotos. A raíz de la falta de espacio para visualizar, la solución más común sería obtener más recursos para atacar el “problema del Big Data”. Sin embargo, en ese mismo año Michael Lesk publicó el *paper* “How much information is there in the world?” (Lesk, 1997), señalando que hacia el año 2000 el mundo produciría suficiente cinta y disco para almacenar toda la información generada hasta el momento, es decir, varios miles de petabytes. Aquella predicción tenía un significado específico: aunque no sería necesario eliminar información, la mayoría de ella nunca sería consultada.

Desde un punto de vista económico y organizativo, las principales características de Big Data son las denominadas tres V: volumen, variedad y velocidad. (Mayer-Schönberger,

2013). Además del volumen de información obtenida, existe una gran variedad de datos representados de diversas maneras, por ejemplo, a partir de dispositivos móviles, audio, video, sistemas de GPS, sensores digitales, etc., de tal forma que las aplicaciones que analizan estos datos requieren que la velocidad de respuesta sea suficientemente rápida para lograr obtener la información correcta en el momento preciso. Al mismo tiempo, la información generada en un entorno de Big Data lleva a una mayor incertidumbre sobre los datos recolectados, por lo que se requiere una mayor calidad sobre el dato obtenido, la fuente y su disponibilidad a futuro. En este aspecto, la veracidad o verosimilitud del dato (lo que algunos referentes en la temática proponen como la 4V) también resultaría fundamental en la cadena de valor de la información.

Memoria digital y gran volumen como aproximaciones para entender Big Data

En la vida cotidiana, los sujetos usuarios de las tecnologías dejan atrás rastros virtuales de cada una de sus acciones: desde las personas con las que se relacionan, sus perfiles en redes sociales, los lugares a donde van, las etiquetas y comentarios de sus gustos y preferencias, los productos que consumen, entre muchas otras variables. Estas pistas digitales narran la vida y el comportamiento humano (en comunidad) con mayor precisión que cualquier otro medio, registrando la actividad como realmente ocurrió.

Según un informe reciente de *IBM Big Data & Analytics Hub* (2015) el 90% de toda la información existente en el mundo fue creada entre 2013 y 2015. Esto significa que en el último minuto que tomó leer los primeros párrafos de este artículo, más de 100 nuevas horas de video se subieron a YouTube, se escribieron casi 1 millón de mensajes de Twitter y se compartieron en Facebook varios millones de publicaciones. El mismo reporte estimó que para el año 2016 se esperaban tener 19900 millones de aparatos conectados, lo que dio un promedio de 2,5 aparatos por persona en el planeta.

Entonces, ¿cuánta información ha creado la humanidad y cuáles son sus consecuencias? Un estudio sin precedentes publicado por la revista *Science* (Hilbert, Lopez, 2011) sostiene que la cantidad estimada de información generada por la humanidad hasta 2007 fue de 295 exabytes, duplicándose en 2011 a 600 exabytes, o lo que es lo mismo a un trillón de bytes (capacidad de un millón de computadoras de escritorio actuales). El informe también indica que la tecnología digital domina sobre la analógica puesto que desde 2007, el 99,9% de la información generada resulta en formato digital, o al contrario, que solo el 0,1% de la información del planeta está en papel. En sintonía con este informe, Eric Schmidt, director general de Google, afirmó públicamente que la humanidad había creado hasta 2003 una cantidad equivalente a 5 exabytes, añadiendo que desde 2011 esa cifra se generaba en apenas dos días.

Las predicciones en torno a Big Data formuladas por *IBM Big Data & Analytics Hub* resultan no menos impactantes:

- La cantidad de datos que la humanidad ha almacenado hasta 2007 es de 295 exabytes. Esto sería 30 veces el número de insectos en la Tierra hoy en día.
- En 2011 esta cantidad se duplicó a 600 exabytes, o lo que es lo mismo a un trillón de bytes. Sería la capacidad que pueden contener un millón de computadoras de escritorio actuales.

- 35 zettabytes de datos serán creados para 2020, lo que significa un crecimiento de 300 veces respecto al año 2005. Equivaldría a ocupar tantos DVDs apilados desde la tierra hasta a alcanzar la mitad del trayecto a Marte.
- Para 2020 se producirán 2,3 trillones de gigabytes por día. Esto equivaldría a 2 trillones de películas con calidad de televisión.
- Se calcula que dentro de dos años, 6 mil millones de personas tendrán celulares, lo cual daría un promedio de casi 1 celular por habitante (0,85) estimando una población mundial de 7 mil millones de personas.

Siguiendo con estas premisas, en 2009 la revista *Wired* (Anderson, 2008) publicó un artículo en el que se presentaba el impacto del aluvión de datos recientes y de los efectos en cuanto a ya no requerirse el método científico y las teorías. En este artículo, *Wired* anunció que este era el “principio de la era del petabyte”. De esta forma, gradualmente el concepto abandonó su nicho y se volvió una idea de alcance masivo que “revolucionó” el estado de la tecnología. Desmitificando la propuesta de *Wired* y como se verá a lo largo del artículo, al analizar Big Data es fundamental no reducir el fenómeno al tecnologicismo (determinación de las características tecnológicas por sobre otras características) o al sociologicismo (las tecnologías de la información y las comunicaciones como variable dependiente del sistema social) sino que se busca pensar una interacción dinámica, fluída y compleja entre la tecnología y sus variables sociales.

Discusiones productivas desde un enfoque sociotécnico

¿Cómo comprender a Big Data desde una mirada integral? Big Data no es solamente una técnica o tecnología abstracta que funciona de manera autónoma sino que se encuentra inserta en un régimen sociotécnico. Existen sistemas sociotécnicos, entendidos como el conjunto de elementos (artefactos, conocimiento, capital, trabajo, mercados) que permiten cumplir con funciones sociales como el transporte, la comunicación o la educación (Geels, 2004: 897). Se trata de mediaciones técnicas y artefactos que también conforman el entramado social (Latour, 2007, 2008). Estos recursos no pueden actuar de forma autónoma sino que se requiere de actores, grupos sociales tales como investigadores, productores, empresarios, usuarios, y autoridades públicas, entre otros para que funcionen, es decir se creen, comercialicen y usen.

Para enriquecer esta visión, las autoras Danah Boyd y Kate Crawford (2012) consideran a Big Data como “un fenómeno cultural, tecnológico y académico”, que reside en la interacción entre la tecnología (capacidad computacional más precisión de los algoritmos), el análisis para identificar patrones y facilitar la toma de decisiones, y la mitología o creencia de que grandes cantidades de datos ofrecen una forma de conocimiento más elevada y una mejor aproximación a la verdad. Como todo nuevo fenómeno sociotécnico, Big Data despierta posiciones tecnofílicas (utopía) y tecnofóbicas (catástrofe) sobre la nueva etapa. La visión optimista sugiere que Big Data permitirá resolver graves problemas sociales en las áreas de salud, terrorismo o inseguridad, cambio climático, etc.; mientras que la visión pesimista alerta sobre el advenimiento de una era de vigilancia, con violación a la privacidad de las personas, menos libertades civiles y mayor control corporativo. Sin embargo, Boyd y Crawford destacan que las posiciones extremas oscurecen la discusión sobre los cambios vertiginosos que este nuevo fenómeno protagoniza en un contexto de incertidumbre: las

decisiones que se tomen en el surgimiento de esta etapa condicionarán su desarrollo futuro.

Teniendo en cuenta la creciente capacidad de recolección de datos y de algoritmos capaces de determinar patrones de comportamiento humano, es necesario discutir qué sistemas están motorizando estas prácticas y cuáles los regulan (el escándalo reciente de *Cambridge Analytica* por haber obtenido en Estados Unidos datos de 50 millones de usuarios en Facebook sin permiso, pone en discusión justamente estas premisas que naturalmente conjugan la omnipotente regulación algorítmica con los intereses de empresas y gobiernos).

Para Boyd y Crawford, el fenómeno Big Data adopta una jerarquía semejante al fordismo: el nuevo fenómeno emerge como un sistema de conocimiento que está cambiando incluso los mismos objetos de conocimiento, la epistemología y la ética, y plantea una transformación de la teoría y el aprendizaje a partir de nuevas herramientas.

En este sentido, la disponibilidad de grandes cantidades de datos renueva la discusión sobre objetividad y subjetividad del método científico, particularmente en las ciencias sociales y las humanidades. La cantidad de datos por sí sola no otorga objetividad absoluta, ya que estos se encuentran sujetos a interpretación, del mismo modo en que el objeto observado y las decisiones de investigación se basan en la subjetividad del investigador. Big Data no se explica por sí solo y los métodos de interpretación están sujetos a debates filosóficos. Esto es así aun sin considerar los errores que puedan tener los datos masivos, la dificultad de interpretación que se genera por fuentes desconocidas, el problema del sesgo o bien la (poca) representatividad del conjunto de datos elegido y su descontextualización, entre otras cuestiones.

Por su parte, Chen y Zhang (2014) plantean que el fenómeno Big Data, que da lugar a la ciencia de datos, se presenta como “un cuarto paradigma científico”, debido a los cambios que las crecientes aplicaciones intensivas en datos están provocando en la ciencia. Los otros tres paradigmas serían la ciencia empírica (explicación de la naturaleza basada en experiencia empírica), la ciencia teórica (Leyes de Newton y Kepler) y la ciencia computacional (simulación científica).

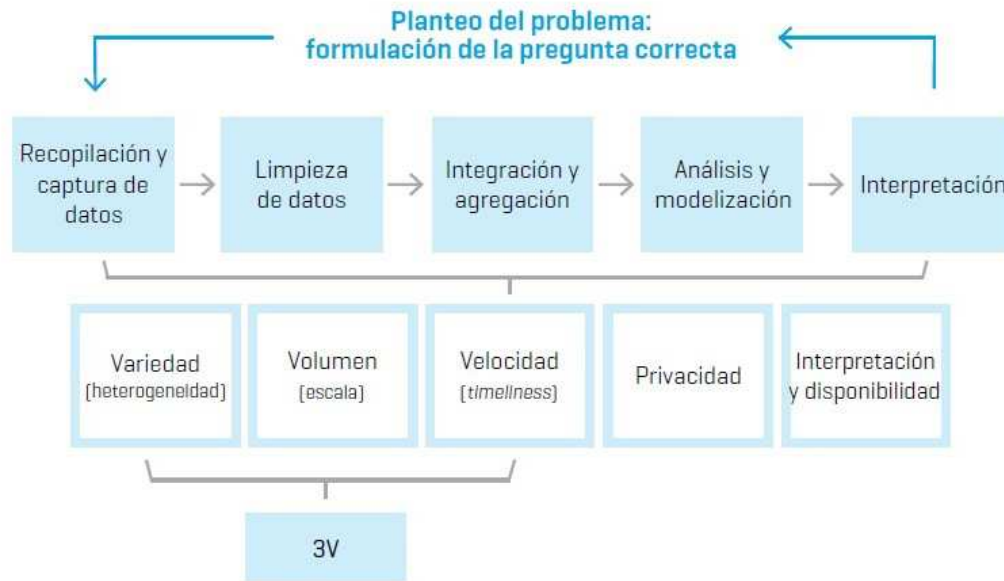
Si bien esta podría ser una visión tecno-optimista, desde un enfoque más global es necesario preguntarse sobre cómo se utiliza y se interpreta la información masiva que se recolecta. En este sentido, la realidad no se explica por sí sola y, por lo tanto, un empirismo radical no es sostenible, ya que forma parte de un sistema cerrado. De esta manera, un dato ya es una construcción sobre aquello que es observable; es decir, no existe un dato crudo absolutamente puro y objetivo (Forte, 2015). Una mayor cantidad de datos no necesariamente significa mayor calidad; es decir, los datos no quedan exentos de la necesidad de mecanismos de validación de los resultados.

Esquema conceptual para Grandes Datos

Siguiendo a Hilbert (2013), lo crucial de Big Data no es el gran aumento de los datos en sí mismo, sino su análisis para la toma de decisiones inteligentes, por lo que prefiere la expresión “análisis de Big Data” a la de “Big Data”. En este sentido, Boyd y Crawford (2012) consideran que la novedad de Big Data, más que el volumen, se encuentra en las capacidades de búsqueda y agregación de grandes cantidades de datos relacionados (*cross-reference data sets*). Por otro lado, UN Global Pulse sugiere que el análisis de

Big Data involucra múltiples fases que plantean un desafío en sí mismo y proponen un esquema lineal para caracterizar todo el sistema, al que se agrega aquí un proceso de retroalimentación:

Figura 1: Secuencia de análisis de Big Data



Arriba: las principales etapas en el análisis de Big Data. **Abajo:** las características de Big Data que hacen de esta tarea un desafío.

Fuente: CIECTI

Aunque la etapa de análisis y modelización (*analysis/modelling*) suele ser la que mayor atención atrae según los autores, no se obtienen demasiados resultados si se descuidan el resto de las etapas y el contexto en el cual se realiza el análisis de Big Data. La fase de filtrado y limpieza de los datos puede demandar hasta el 80% del trabajo de análisis. Por ende, si se acepta esta estimación y se considera que una vez que el dato es manipulado se pierde su objetividad intrínseca, prácticamente no hay objetividad posible. Por lo tanto, quien tiene la propiedad de los datos (o los medios/recursos para apropiárselos) y quien realiza el análisis es determinante en la construcción de sentido. Esto agrega otro desafío en el análisis de Big Data: no sólo plantea un problema técnico de hardware y software, sino que también requiere técnicos y profesionales capacitados para llevar a cabo esta tarea.

Existe una gran variedad de técnicas y tecnologías para agregar, manipular, analizar y visualizar Big Data, desarrolladas a partir de diferentes campos como la estadística, las ciencias de la computación, la matemática aplicada y la economía, a partir de un enfoque flexible y multidisciplinario. La Comisión Europea (2013) elaboró de este modo la lista de técnicas, tecnologías y herramientas:

Figura 2: Diferentes tipos de técnicas, tecnologías y herramientas de visualización de Big Data



Fuente: CIECTI

Ventajas y desventajas para el investigador

Para el investigador, el uso de técnicas de análisis de grandes volúmenes de datos presenta diferentes pros y contras.

Provee información complementaria a la de las fuentes de datos tradicionales, pudiendo responder a las preguntas de investigación desde otra perspectiva. Muchas de estas fuentes tienen una cobertura global, posibilitando abordar estudios comparativos entre ciudades o entre países.

Los datos masivos geolocalizados tienen una muy alta resolución espacial y temporal: espacial, porque cada dato está localizado por sus coordenadas geográficas y no agregado espacialmente; temporal, porque para cada dato se almacena el momento en que se genera (año, mes, día, hora, minuto y segundo), de forma que se dispone de información siempre actual y se pueden realizar estudios evolutivos y monitorizar procesos. Algunas de estas fuentes de datos son gratuitas y se pueden descargar directamente desde Internet utilizando la correspondiente API.

Pero no todas son ventajas. Algunas de estas fuentes de datos masivos son de difícil acceso por diversos motivos. Puede suceder que las organizaciones y empresas que los generan no estén dispuestas a compartirlos con los investigadores, por el valor (costo) de los datos, o bien por diversas políticas de privacidad y/o protección de los datos.

Otro problema deriva de la dificultad de procesar los datos masivos, ya que su volumen excede las capacidades de los gestores de bases de datos convencionales.

Además este tipo de datos presenta el inconveniente de que generalmente son sesgados (lo que sucede en muchos casos cuando se trabaja con bases de datos de Internet o segmentos de usuarios de las redes sociales), lo que obliga a utilizar otras fuentes de forma complementaria para intentar compensar el sesgo. De este modo, el análisis con Big Data tiene un carácter exploratorio sin que siempre sea posible determinar estadísticamente los márgenes de error y los niveles de significación.

A modo de cierre. Big Data desde un abordaje transdisciplinario. Casos locales.

En una entrevista reciente en *The Clinic Online*¹ el Dr. en Comunicación alemán Martin Hilbert decía que el Big Data convirtió a las ciencias sociales en la ciencia más rica en datos: “Antes tenías que negociar hasta con diplomáticos para que te prestaran una base de datos de cien filas por cien columnas. Y en las universidades hacían experimentos con quince alumnos de pregrado, que necesitaban créditos extra para pasar el ramo, todos blancos, todos de 18 años, y decían ‘miren, así funciona la psicología humana’”. Desde su punto de vista, gracias al enorme flujo de datos que proveen los dispositivos digitales, las ciencias “blandas” podrían finalmente alcanzar la densidad de las “duras” a fuerza de datos.

En *El oficio de científico. Ciencia de la ciencia y reflexividad*, Pierre Bourdieu (2003) se propone des-idealizar las rutinas de la comunidad científica, para que se comprendan mejor “los mecanismos sociales que orientan la práctica científica y convertir (a los investigadores) en dueños y señores no sólo de la naturaleza de su práctica sino también del mundo social en el que se produce”. Según Bourdieu ese aprendizaje es muy valioso para investigadores jóvenes porque los científicos no politizados y sin visión empresarial corren el riesgo de quedar marginados delante de las grandes maquinarias científicas promovidas por empresas o instituciones públicas.

De este modo, en las nuevas ciencias sociales se podría luchar por el monopolio de la producción de la verdad, hoy en manos de las ciencias duras, y trabajar también en el uso de políticas públicas con análisis de grandes volúmenes de datos (no sirve más la encuesta a 100 estudiantes de una universidad para comprobar hipótesis en una tesis doctoral). Las condiciones históricas determinan el hecho científico. Y hoy estas condiciones consistirán en la manipulación y el análisis de grandes volúmenes de datos. El propio Hilbert ha expuesto cómo ha cambiado la situación de Estados Unidos, en torno a las políticas públicas e investigación con Big Data: “no teníamos datos y por eso no funcionaban las políticas públicas; y de la noche a la mañana, el 95% de los sujetos que estudiamos pasó a tener un sensor de sí mismo 24 horas al día”.²

En el caso de la Argentina algunas (pocas) organizaciones ya están trabajando conjuntamente para estudiar el fenómeno, generar oportunidades para superar la brecha digital y elaborar políticas de investigación en sectores estratégicos donde se pueda aplicar Big Data.

El Centro Interdisciplinario de Estudios en Ciencia, Tecnología e Innovación³ (asociación civil integrada por la Universidad Nacional de Quilmes y la Facultad Latinoamericana de Ciencias Sociales, FLACSO Argentina) ha hecho diversos aportes de análisis en base a Big Data, donde se cruzan enfoques sociales con enfoques tecnológicos y productivos del país; así como también existen estudios de consultoras privadas con aportes de políticas públicas.

¹ Véase The Clinic Online: <http://www.theclinic.cl/2017/01/19/martin-hilbert-experto-redes-digitales-obama-trump-usaron-big-data-lavar-cerebros/> (recuperado el 1 de abril de 2018).

² Ibidem.

³ Para mayor información: <http://www.ciecti.org.ar/>

La Fundación Sadosky,⁴ dependiente del Ministerio de Ciencia, Tecnología e Innovación Productiva, cuenta con un Programa de Ciencia de Datos cuyo objetivo es “contribuir a que Argentina se convierta en líder regional en la temática de Grandes Datos (Big Data), considerada clave para la autonomía tecnológica, el desarrollo económico y social y la competitividad”. Uno de los ejemplos de acercamiento entre las ciencias duras y las ciencias blandas, ha sido la organización del curso Análisis Computacional de Textos y Redes⁵ (Instituto de Cálculo, 2017).

Por su parte, existen otros grupos de investigación y discusión en el país que piensan el fenómeno de los datos, la informática y las redes sociales desde una perspectiva transversal, multi y transdisciplinaria. Algunos ejemplos son el Centro Latinoamericano de Formación Interdisciplinaria (CELFÍ-DATOS),⁶ el grupo de Sociofísica argentina (Facultad de Ciencias Exactas y Naturales-UBA)⁷ y el grupo Antropocaos (Facultad de Filosofía y Letras-UBA).⁸

En varios de estos sectores e instituciones se está discutiendo la posibilidad de que los estudios y análisis de grandes volúmenes de datos sean un fenómeno transversal a diversas ciencias y disciplinas, tanto duras como blandas (aun cuando ésta sigue siendo una distinción poco feliz en la comunidad científico-académica).⁹

Apéndice: ¿Big data y la Memoria Digital nos hacen aún más inteligentes?

En un relato que Jorge Luis Borges escribió en 1942 se narra la historia de Ireneo Funes, el memorioso¹⁰. Según cuenta Borges, Funes era un peón de campo nacido en Fray Bentos, Uruguay. Una trágica caída de un caballo lo confinó a un pobre lecho en un rancho de la campiña oriental. Fue entonces cuando se produjo en Funes un cambio sustancial. A raíz del accidente desarrolló una cualidad sorprendente y prodigiosa: la perfecta memoria.

Según comentaba Borges, el presente era rico y nítido y también las memorias más antiguas y triviales. Al igual que cualquier computadora de última generación y de máxima capacidad de almacenamiento, Funes podía recordarlo todo, era infalible. “Más recuerdos tengo yo solo que los que habrán tenido todos los hombres desde el que mundo es el mundo”, dirá. Y agregará “mi memoria, señor, es como vaciadero de basuras”.

⁴ La Fundación Dr. Manuel Sadosky es una institución público privada cuyo objetivo es favorecer la articulación entre el sistema científico-tecnológico y la estructura productiva en todo lo referido a la temática de las Tecnologías de la Información y la Comunicación (TIC). Creada a través del Decreto Nro. 678/09 del Poder Ejecutivo Nacional, la Fundación es presidida por el ministro de Ciencia, Tecnología e Innovación Productiva: <http://www.fundacionsadosky.org.ar/category/programa-ciencia-de-datos/>

⁵ Publicación del curso: <http://www.fundacionsadosky.org.ar/analisis-computacional-de-textos-y-redes-en-ciencias-sociales/>

⁶ <http://www.celfi.gob.ar/celfi-datos/sobre-el-celfi>

⁷ Véase <http://sociofisicaargentina.blogspot.com.ar/> y <http://nexciencia.exactas.uba.ar/sociofisica-ciencias-sociales-computacionales-sofia-maxi-san-miguel>

⁸ <http://www.antropocaos.com.ar/>

⁹ Otro aporte interesante en este aspecto resulta el de Walter Sosa Escudero, economista e investigador de la Universidad de San Andrés, quien en diversos trabajos analiza la relación entre Big Data, Estadística, Economía y Sociedad.

Véase: <https://waltersosa.weebly.com/uploads/2/2/1/8/22189288/bigdataexactas.pdf>

¹⁰ Véase http://biblio3.url.edu.gt/Libros/borges/el_memorioso.pdf (recuperado el 15 de marzo de 2018).

Si bien la metáfora del síndrome de Funes (Rey Lennon, Bartoli Piñero, 2008: 44-49) es ilustrativa para pensar que hoy cualquier grupo u organización cuenta con métodos y tecnologías de Big Data y Aprendizaje Automático (*machine learning*) cuya capacidad de memoria es histórica, comparativamente hablando con etapas anteriores de la humanidad, no por ese motivo Big Data vuelve a las personas, organizaciones y comunidades más capacitadas para pensar e interpretar.

Como ejemplo concreto, puede tomarse esa representación un tanto siniestra de HAL, la “computadora pensante” que ideara Clarke y que Stanley Kubrick lleva al cine en *2001, Odisea en el espacio*, como mito de la computadora inteligente donde las técnicas de inteligencia artificial estarían por encima de la capacidad del hombre. En términos de McLuhan (1985), se cree que el sujeto cuenta con una extensión de la capacidad pensante en vez de una extensión de la memoria humana. No obstante, esta “prótesis” cognitiva sigue siendo algo muy valioso y que aprovechado de manera correcta y rigurosa puede ampliar los diversos horizontes de conocimiento.

Ya narraba Borges que Funes también intentó ordenar la información contenida en su extensa memoria. En 1886 Funes “había discurrido un sistema original de numeración y en muy pocos días había rebasado el veinticuatro mil”. Su disparatado principio era suplantarlo cada número por una palabra y un signo. Sin embargo, la tarea era interminable e inútil. En este sentido, de ese interminable caudal de información Funes era incapaz de concebir ideas generales y su inagotable memoria particular le impedía captar los conceptos universales.

Es por ello de que a pesar de que Big Data permite a individuos, organizaciones, universidades, empresas y comunidades la obtención, análisis, procesamiento y visualización de la información, en formas inéditas y resultados de amplia validez técnica; ello no alcanza ya que se requiere métodos de planificación, comprensión, relacionamiento y cierta sistematización de los datos obtenidos, logrando entender (al fin y al cabo) por qué el software y/o las tecnologías utilizadas en la investigación arrojaron uno u otro resultado. En síntesis, más allá de la innovación en el uso de cualquier técnica de Big Data e inteligencia artificial, la capacidad de interpretación humana sigue siendo esencial.

Bibliografía

- Anderson, C. (2009). “The End of Theory: The Data Deluge Makes the Scientific Method Obsolete”, *Wired Magazine*. Disponible: <https://www.wired.com/2008/06/pb-theory/> (recuperado el 10 de abril de 2018).
- Bourdieu, P. (2003). *El oficio de científico. Ciencia de la ciencia y reflexividad* Barcelona, Anagrama.
- Boyd, D.; Crawford, K. (2012). “Critical Questions for Big Data”. *Information, Communication & Society*, 15:5, pp. 662-679.
- Chen, C. L. P. y Zhang C. (2014). “Data-Intensive Applications, Challenges, Techniques and Technologies: A Survey on Big Data”, *Information Sciences*, Nº275, Macau (China): Department of Computer and Information Science, Faculty of Science and Technology, University of Macau, pp. 314-347. <http://www.cs.unibo.it/~montesi/CBD/Articoli/SurveyBigData2.pdf>

- Cox, M.; Ellsworth, D. (1997). "Application-Controlled Demand Paging for Out-of-Core Visualization". *Report NAS-97-010*. NASA Ames.
- Forte, M.Á. (2015). *Modernidad: tiempo, forma y sentido*, Buenos Aires: Eudeba.
- Geels, F. (2004). "From sectoral systems of innovation to sociotechnical systems. Insights about dynamics and change from sociology and institutional theory." *Research Policy*, vol. 33, pp. 897–920.
- Hilbert, M.; Lopez, P. (2011). "The World's Technological Capacity to Store, Communicate, and Compute Information". *Science* 332, 60.
- Hilbert, M. (2013), "Big Data for Development: From Information to Knowledge Societies", Washington, United Nations eclac. Disponible en <<http://ssrn.com/abstract=2205145>>.
- Latour, B. (2007). *Nunca fuimos modernos: ensayo de antropología simétrica*. Siglo XXI de España Editores.
- Latour, B. (2008). *Reensamblar lo social: una introducción a la teoría del actor-red*. Madrid: Manantial.
- Lesk, M. (1997). "How Much Information Is There In the World?". Disponible: <http://www.lesk.com/mlesk/ksg97/ksg.html>. (recuperado el 10 de abril de 2018).
- Malvicino, F. Yoguel, G. Big Data. (2017). "Avances recientes a nivel internacional y perspectivas para el desarrollo local". Disponible: <http://www.cieci.org.ar/wp-content/uploads/2017/07/DT3-BigData-avances-y-perspectivas-de-desarrollo-local.pdf> (recuperado el 4 de diciembre de 2017).
- Mayer-Schönberger, V. (2013). *Big Data. La revolución de los datos masivos*. Madrid: Turner Publicaciones.
- McLuhan, M. (1985). *Guerra y paz en la Aldea Global*. Barcelona: Planeta-Agostini.
- OECD (2013). "Exploring Data-Driven Innovation as a New Source of Growth: Mapping the Policy Issues Raised by 'Big Data'", *oecd Digital Economy Papers*, N° 222, junio, oecd Publishing. Disponible en <<http://dx.doi.org/10.1787/5k47zw3fcp43-en>>.
- Rey Lennon, F. y Bartoli Piñero, J. (2008). *Reflexiones sobre el management de la comunicación*. Buenos Aires: La Crujía.
- Un Global Pulse (2012). "White paper: Big Data por Development: Opportunities & Challenges". Documento creado a partir de conversaciones mantenidas entre miembros de universidades de Estados Unidos (Massachusetts Institute of Technology, University of California at Berkeley, Stanford, entre otras) y empresas tecnológicas líderes (Yahoo!, Google, Microsoft, IBM, HP). <<http://www.unglobalpulse.org/projects/BigDataforDevelopment>>.

Fuentes electrónicas

BIBLIOTECAS LANDIVARIANAS

http://biblio3.url.edu.gt/Libros/borges/el_memorioso.pdf (recuperado el 15 de marzo de 2018).

IBM BIG DATA HUB <http://www.ibmbigdatahub.com/> (recuperado el 15 de enero de 2017).

CLEI 2017 <http://www.clei201746jaiio.sadio.org.ar/sites/default/files/Mem/STS/STS-10.pdf> (recuperado el 10 de enero de 2018).

THE CLINIC <http://www.theclinic.cl/2017/01/19/martin-hilbert-experto-redes-digitales-obama-trump-usaron-big-data-lavar-cerebros/> (recuperado el 1 de abril de 2018).