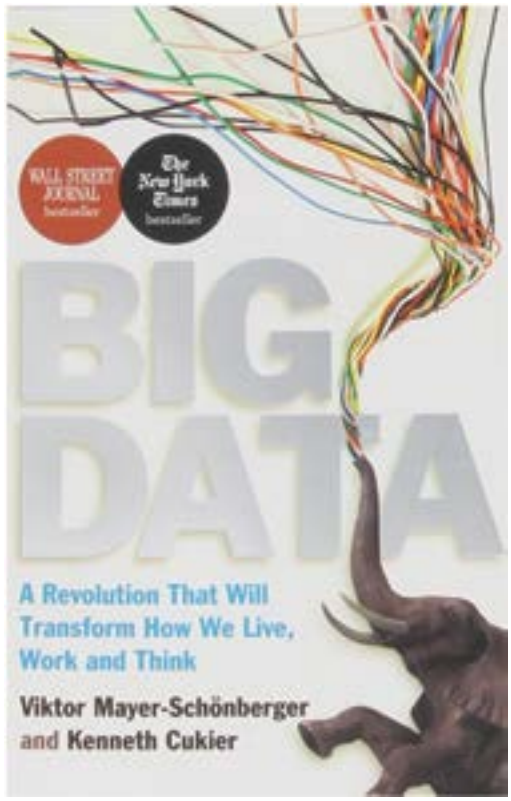
Big data:

la revolución que no debemos ignorar



Reseña

Gerardo Leyva

Un fantasma recorre el mundo: el fantasma de *big data*. Se trata de una revolución que promete incidir de manera importante en la manera como vivimos, trabajamos y pensamos, como reza el subtítulo del libro de Viktor Mayer-Schönberger y Kenneth Cukier. Sirve de mucho tener una referencia clara y sólida de lo que es *big data*. En la actualidad, son muchos los que hablan del tema, pero pocos todavía los que pueden en realidad explicar de qué se trata. Parafraseando a Dan Ariely: "…es como el sexo para los adolescentes: todo mundo habla acerca de ello, nadie realmente sabe cómo hacerlo, todos piensan que todos los demás lo están haciendo y por lo tanto todos aseguran que ellos también lo están haciendo".[1] De *big data* se ha dicho, por ejemplo, que es una gran amenaza para la seguridad y la libertad de los individuos, con tonalidades apocalípticas que anuncian la ominosa llegada del *gran hermano*, imaginado en la distopía de Orwell. Se dice, también, que significa el fin de la teoría y de la ciencia como la conocemos y que nos acerca al mundo indolente y carente de significado imaginado por Huxley. Hay quienes afirman, de igual forma, que *big data* nos solucionará los más diversos aspectos de la vida, de tal manera que será más bien un catalizador del progreso con pocos parangones en la historia de la humanidad. Por supuesto, también están quienes nos advierten de no tomarlo demasiado en serio.[2]

Ante esta diversidad de perspectivas, el libro tiene la virtud de ser una referencia clara y concisa, basada en ejemplos concretos, que dan una visión de conjunto, lúcida y sencilla sobre qué entender por *big data*, cuáles son sus principales dimensiones, sus riesgos y limitaciones. Los autores parten de que los volúmenes de información con los que contamos hoy en día, su variedad y la velocidad con la que crecen nos ponen en una tesitura no sólo de hacer mucho más con más datos, sino de poder hacer cosas radicalmente

1 http://whatsthebigdata.com/2013/06/03/big-data-quotes/

2 Ver la conferencia de Susan Etlinger en TED titulada: *What do we do with all this big data?*, en: https://www.ted.com/talks/susan_etlinger_what_do_we_do_with_all_this_big_data/transcript

diferentes a las que por tradición se han hecho con la información. Sostienen que *big data* "…se refiere a cosas que uno sólo puede hacer a una gran escala y que no pueden realizarse a una escala menor".[3] *Big data* requiere, entonces, de usar toda la información disponible y no sólo una parte, de manera que *N* = todo. El enfoque de *big data* implica que la misma base informativa se puede utilizar para una amplia variedad de propósitos que no podían haber sido siquiera imaginados al momento de generar la información, parte de que ésta se recicla de forma permanente, de modo que la que es desecho para un proceso puede ser la materia prima sustantiva para otro.

Big data —sostienen— implica una nueva filosofía en torno a cómo nos relacionamos con la información. Esto trastoca el concepto de la estadística tal cual como lo conocemos hoy. La forma en la que la aprendemos en las escuelas hasta la actualidad supone un paradigma de datos escasos, derivado de la necesidad de hacer muestras que sean representativas de la población mayor, lo cual, si bien nos permite resolver problemas específicos, nos deja sin la posibilidad de reciclar esa misma información para atender otro tipo de preguntas distintas a aquellas que subyacen al diseño de la muestra y del cuestionario.

Usar *big data* equivale a dejar que los datos hablen. La estadística tradicional es análoga a la fotografía, que implica enfocar el objetivo, apretar el disparador y quedarnos con la imagen de lo que queríamos conocer, resignándonos, sin embargo, a tener una imagen borrosa del resto de los objetos en torno a aquello que quisimos fotografiar en un principio. En cambio, el enfoque de *big data* es análogo a tomar fotografías que capten todos (*N* = todo) los rayos de luz en la primer instancia para que en una segunda seamos nosotros quienes decidamos cuáles destacar y cuáles no, de manera que podemos tener la granularidad que queramos en la definición de los objetos que nos propongamos reflejar a partir de la totalidad de la información. Así, siguiendo la analogía, *big data* nos permitiría una gran cantidad de fotografías distintas a partir de un mismo disparo del obturador.

El tema nos plantea una nueva manera de interactuar con el mundo en la cual, por ejemplo, tendremos que empezar a poner un menor peso en la causalidad como vía para entender nuestro entorno. Contar con información en grandes cantidades sobre la diversidad amplísima de temas, referidos a una suma también tremenda de personas y cosas, nos permite establecer correlaciones útiles que guíen las predicciones que permiten que muchas personas hoy en día tengan ventajas significativas en una diversidad de actividades incluyendo, por supuesto, los negocios. En la actualidad, por medio de *big data* podemos identificar usos fraudulentos de tarjetas de crédito, encontrar los mejores precios para viajar en avión, resolver problemas complejos de planeación y administración urbana, dar seguimiento con gran oportunidad y detalle a la propagación de enfermedades contagiosas, guiar acciones de las fuerzas de seguridad, autorizar o negar créditos, traducir de forma automática textos de los más variados idiomas, seguir el tráfico en las ciudades, así como mejorar el mantenimiento y control de equipos sofisticados de maquinaria, autos y aviones, entre muchas aplicaciones más.

Hoy en día, las compañías telefónicas saben dónde estamos, con quiénes nos comunicamos, hacia dónde nos desplazamos y de quiénes estamos cerca; servicios como el de Amazon revelan nuestras preferencias de compra y, en buena medida, también nuestros intereses intelectuales; *Twitter* publica en qué estamos pensando; *Facebook* hace evidentes nuestras redes sociales, es decir, con quiénes nos relacionamos; *Google* registra qué es lo que buscamos en internet. Incluso, qué tan bien dormimos o si nuestro refrigerador está vacío o qué películas vemos, o el tipo de desperfectos de nuestros automóviles son temas que se están incorporando al universo de *big data*.

Más, mucha más información hace que cambien de forma radical las reglas del juego, pasando de las variaciones acumulativas a los cambios cualitativos. Según cálculos reportados por los autores, para el 2000 sólo una cuarta parte del total de la información almacenada era digital, mientras que para el 2013, menos de 2% era no digital. Ello implica no únicamente que su volumen ha crecido a una velocidad descomunal, sino que también la mayor parte de este crecimiento se

3 Mayer-Schönberger, Viktor and Kenneth Cukier. *Big Data. A Revolution That Will Transform How We Live, Work and Think*. London, John Murray, 2013, p. 6.

ha dado en formato digital; en muchos casos, la información no es estructurada, pero es susceptible de ser explotada. Este incremento de su disponibilidad, junto con las posibilidades que abren las crecientes capacidades de almacenaje y procesamiento de las fuentes digitales, nos ponen en una coyuntura de cambio radical respecto a la manera en que nos relacionamos con los datos y la estadística.

Por tradición, los censos han sido los ejercicios estadísticos más amplios, ya que cubren a la totalidad de la población en estudio; en este sentido, *N* = todo. Sin embargo, pese a los grandes avances en la materia, a la fecha siguen siendo un ejercicio costoso, restringido a una temática necesariamente reducida, que sólo puede desarrollarse con poca frecuencia, de manera que es inviable tener censos, por ejemplo, cada mes, semana o día. En un mundo dominado por el paradigma de datos escasos, lo razonable ha sido recurrir al muestreo como una vía para obtener información por inferencia de la población en general a partir de una muestra representativa. Los autores indican que conforme nos vayamos alejando del paradigma de datos escasos y nos introduzcamos más de lleno en el de *big data*, el muestreo será cada vez menos necesario (si bien es improbable que pronto dejemos de necesitarlo).

Así como en su momento la aparición del muestreo nos abrió nuevas puertas para el conocimiento de la realidad, permitiéndonos realizar investigaciones a costos razonables sobre una amplia variedad de temas y con periodicidad mucho mayor a la de un censo, ahora también *big data* empuja más la línea para armarnos con capacidades analíticas inimaginables hasta hace pocos años. Esto es así porque las encuestas suponen un diseño temático y estadístico que corresponde a un propósito específico, por lo que pueden ser utilizadas de manera adecuada sólo en la medida en que se empleen para responder a preguntas consistentes con ese diseño. En cambio, la información de *big data* puede ser reciclada un número indeterminado de veces para atender temáticas diferentes, además de que ofrece una granularidad constante que permite usar los mismos datos para realizar análisis de dominios diversos con muy distintos niveles de detalle. Ello, de suyo, es posible dada una flexibilidad mayor de forma considerable a la de las encuestas atadas a un diseño conceptual y estadístico, pues aquí se parte en todo momento de la totalidad de la información disponible y no sólo de una fracción de la misma. Así, tal como ha comenzado a experimentarse en el INEGI, la misma base estadística de *big data* (en este caso, tuits) que se utiliza para medir felicidad en un país, se puede reciclar después para revisar los desplazamientos de las personas entre sus ciudades y regiones o para hacer un análisis de expresiones discriminatorias emitidas por algún segmento de la población o para estimar flujos migratorios en la zona fronteriza.

En virtud de que *big data* descansa en la totalidad de la información— o algo muy cercano a ello—, nos permite hacer análisis a distintos niveles sin riesgo de que la *imagen* resulte borrosa. Es muy distinto cuando se parte de datos provenientes de una encuesta probabilística donde, a medida que se hacen más y más cruces de variables, los resultados se tornan crecientemente imprecisos. Más aún, *big data* nos permite conocer el comportamiento natural de las personas, dado que no depende de respuestas a cuestionarios, sino de transacciones afectivas, económicas, informativas, etc., que los individuos hacen como parte de su vida diaria. En buena medida, es por esto que los autores afirman que "…procurarse una encuesta aleatoria en la era del *big data* es como aferrarse a una fusta para caballo en la era del vehículo de motor".[4]

El mundo de las estadísticas oficiales reconoce que tres fuentes agotan la totalidad de los datos disponibles: censos, encuestas y registros administrativos. Sin embargo, pretender que éstas atienden las necesidades de información del mundo de hoy es querer manejar un auto mirando sólo por el espejo retrovisor. *Big data* significa una cuarta y nueva fuente que nos permite explotar, para fines estadísticos, la información digitalizada —estructurada o no—, que constituye cerca de 99% de los datos disponibles y, también, que es la que más rápido está creciendo.

Pero no todo es miel sobre hojuelas, dado que transitar de un mundo de *small data* a uno de *big data* no es algo exento de costos. Uno muy evidente es el tener que trabajar con datos que no surgen de un diseño y que, por lo tanto, son desordenados, con frecuencia

4 *Op. cit.*, p. 31.

inconsistentes y con mucha basura, que es necesario filtrar o hacer a un lado para encontrar lo que se busca. *Big data* es más análogo a un machete que a un bisturí, pero muchas veces la riqueza analítica y la oportunidad y bajo costo de los análisis a partir de esta nueva fuente pueden más que compensar el trabajar con datos plagados de impurezas y el llegar a resultados que pudieran no ser idealmente nítidos. Un ejemplo es el traductor de *Google*, que no se basa en un análisis gramatical de los idiomas que traduce, ni siquiera en la extrapolación de traducciones profesionales y rigurosas hechas por expertos, sino en el entrenamiento de la computadora a partir de la comparación de millones de traducciones de buena, regular y mala calidad, desde las cuales genera modelos estadísticos que le permiten maximizar la probabilidad de que determinada expresión en el idioma A se traduzca de cierta manera en el B. Las traducciones resultantes no son perfectas, pero son útiles y cada vez mejores, además de ser instantáneas y de que se pueden triangular entre muchos idiomas.

La idea es que modelos simples basados en una gran cantidad de datos pueden resultar más eficaces que modelos muy complejos basados en una cantidad relativamente pequeña de información, tal como lo dijera Peter Norving, el gurú de inteligencia artificial de Google. Adentrarse en el mundo de *big data* invita a revisar nuestras ideas en torno a los méritos de la exactitud, dado que la implicación de un dato impreciso es mucho más grave en un mundo de *small data* que en uno de *big data*. Por otro lado, la mejor imagen de conjunto del fenómeno en estudio que deriva de usar N = todo puede más que compensar el costo en imprecisión. Este enfoque es lo que le permite a empresas como Zest Finance hacer dinero dando a los prestamistas información acerca de qué tan probable es que un prestatario pague el dinero que está solicitando. Esto lo hace a partir de grandes volúmenes de información no necesariamente consistente —y de diversa calidad— pero que le permiten hacer estimaciones con balance de precisión/costo razonablemente eficaz y competitivo. La mayor precisión que pudieran tener los cálculos de sus competidores no alcanza a compensar el costo adicional de hacerlo con un enfoque de *small data*.

A fin de cuentas, de lo que se trata es de tener predicciones acertadas y oportunas a un costo razonable. En el mundo de *big data* esos pronósticos no resultan del análisis causal, sino de correlaciones. Las predicciones basadas en ellas están en el corazón de *big data*. Disponer de una gran cantidad de variables a correlacionar permite encontrar patrones que con dificultad se harían evidentes bajo un enfoque basado en intuición experta o en consideraciones teóricas. Esto, por supuesto, también viene con un costo, pues se sabe el *qué* pero no el *por qué*.

Sin embargo, en muchos casos, saber el *qué* puede ser suficiente; por ejemplo, como reportan los autores, una empresa llamada FICO analiza una diversidad de variables, algunas de ellas en apariencia disparatadas, para generar modelos de correlación que ayuden a predecir la probabilidad de que los pacientes tomen sus medicinas, ayudando a identificar a qué pacientes hay que hacerles recordatorios, lo cual obra en favor tanto de los mismos como de los costos de las instituciones proveedoras de salud. Por supuesto, no hay una relación causal clara entre tomarse la medicina y tener un coche, pero si tener automóvil ayuda al pronóstico, con eso es suficiente (para los fines de FICO).

En el mundo de *small data*, donde la información proviene de un diseño, tener claridad sobre las relaciones causales es una fuente de tranquilidad para el analista, sin embargo, es posible que sea ilusoria en el sentido de que los modelos teóricos de los cuales se parte pueden estar equivocados y los modelos estadísticos que de ellos se derivan podrían no atender a los verdaderos parámetros que subyacen al fenómeno en estudio, como quedó establecido para el caso de la macroeconometría a partir de lo que se conoce como *la crítica de Lucas*.

Una ventaja del *big data* es que, frecuentemente, las fuentes de información son continuas, de manera que las predicciones se pueden hacer con muy alta frecuencia, lo que permite pasar del pronóstico (*forecasting*) al *ahoranóstico* (*now casting*). Considérese, por ejemplo, el caso del *Billion Prices Project* desarrollado por Alberto Cavallo del MIT, que genera índices de precios para diversos países a partir de precios de venta *online* que publican en internet las más diversas empresas comerciales. Mayer Schönberger y Cukier nos cuentan que un robot se encarga de colectar los datos diariamente y de generar los índices de precios que resultan muy

similares a los oficiales en la mayor parte de los casos (con la excepción de Argentina, cuyas estadísticas oficiales de precios han sido objeto de serios cuestionamientos, lo cual permite al autor, también argentino, ofrecer alternativas a las mediciones oficiales). De forma clara, la calidad de los índices oficiales de precios es, en general, mayor que la del ejercicio de Cavallo, aunque él los hace a un costo mucho menor y permite generar información con una frecuencia imposible de replicar por las estadísticas oficiales tradicionales. Con el *Billion Prices Project* (que luego derivó en el emprendimiento *Price Stats*) se puede hacer *now casting* de los índices de precios de muchos países, además de que ofrece un elemento referente de control que antes no existía. De manera similar, las expresiones vertidas en *Twitter* o las búsquedas en *Google* pueden usarse para hacer *now casting* del PIB o de la confianza del consumidor, comenzando así a establecerse vínculos entre *big data* y las estadísticas oficiales.

Parecería que la creciente disponibilidad de información ofrece la posibilidad de encontrar correlaciones que nos permitan pronosticar de manera eficaz casi cualquier variable a partir exclusivamente de criterios empíricos, es decir, sin conocer o pretender conocer las causas del fenómeno. Chris Anderson, el editor en jefe de *Wired Magazine*, se atrevió, incluso, a augurar el *fin de la teoría* sugiriendo que la era de *big data* marcaría el fin de la primacía del método científico como fuente de conocimiento y como herramienta para dominar la realidad. Si bien tal declaración parece exagerada tanto para los autores del libro como para quien esto escribe, el mero hecho de que se haya enunciado nos da una idea del tipo de expectativas que el advenimiento de *big data* ha desatado.

Éste no es un tema de especulación intelectual. Hoy en día, muchas personas se enriquecen a partir de negocios que explotan una gran cantidad de sus potencialidades. La base de éstos es información que tiene su origen en una diversidad de propósitos distintos (comunicarse, ubicarse, divertirse, enterarse, comprar, etc.), en general ajenos al uso estadístico de esos datos. El éxito comercial de *big data* se manifiesta en empresas como Farecast, que ofrece información para elegir el momento más adecuado para comprar un boleto de avión al mejor precio; Fligth Caster, que se dedica a pronosticar los vuelos que serán cancelados; Prismatic, que agrega y jerarquiza el contenido de internet; AirSage, que usa información de millones de teléfonos celulares para proporcionar datos de tráfico en tiempo real; o Sense Networks y Skyhook, que identifican en qué partes de una ciudad hay más actividad nocturna o dónde están y a dónde se dirigen las manifestaciones. Otras como Derwent Capital y MarketPsych ofrecen información extraída de redes sociales convertida en señales para invertir en el mercado de capitales e, incluso, se tienen empresas como Asthmapolis, que usa geolocalización para identificar detonadores ambientales de ataques de asma. A estos emprendimientos recientes —y aún relativamente pequeños— se suman los usos de *big data* que han generado o ahorrado muchos millones de dólares a grandes empresas como Google, UPS, Amazon, Netflix, Target, General Electric, Rolls-Royce y Walmart, así como a una gran cantidad de bancos y empresas financieras y de seguros. Así las cosas, en la medida en que *big data* siga siendo negocio, podemos esperar que la ola continúe creciendo.

Explotar *big data* implica reciclar información, con frecuencia, varias veces. Al ser un insumo para crear valor, *big data* también tiene valor. Sin embargo, en la

actualidad no hay una manera única ni un acuerdo de cómo valuar los datos. El hecho de que la misma información pueda ser usada para muy diferentes propósitos no hace más fáciles las cosas. De hecho, sus diversos usos no suelen estar a la vista. Lo que para algunos puede ser un conjunto de datos sin valor, para otros se convierte en una mina de oro. Por supuesto, contar con la información, saber cómo usarla y tener la sensibilidad para identificar cuáles son las aplicaciones relevantes son cosas distintas. Disponer de los datos no es lo único importante. Saber cómo y, sobre todo, qué hacer con ellos son también fuentes importantes de valor. Fue precisamente pensando en las casi infinitas posibilidades que ahora se abren con *big data* que Hal Varían, el economista en jefe de Google, externó su famosa frase en el sentido de que la estadística es la profesión más sexy de principios del siglo XXI.

Sin embargo, Mayer-Schönberger y Cukier consideran que, a medida que el análisis de *big data* se generalice y se estandarice, la importancia relativa de la información en sí misma será cada vez más alta como fuente de valor. Esto es una ventaja en favor de los generadores originales de información y de los intermediarios que puedan integrar datos de distintas fuentes de manera que den pie a usos de mayor alcance que el que permiten las fuentes originales por separado. En todo caso, en la medida en que *big data* se convierta en una fuente de ventaja competitiva para gran cantidad de negocios, la estructura de muchos sectores económicos será redefinida, pero no de una manera uniforme, sino en favor de empresas que ahora ya son muy grandes y disponen de altos volúmenes de información, y recursos humanos y tecnológicos de gran calidad y abundantes, o de empresas pequeñas, flexibles y dinámicas, dejando en desventaja a las de tamaño mediano que carecen de las economías de escala (en datos) y de la adaptabilidad de quienes les flanquean en los rangos de tamaño.

Otra implicación para los negocios es que los expertos comienzan a tener una nueva fuente de competencia en los datos, de maneras distintas a lo hasta ahora conocido. En línea con lo que se plantea en la película *Moneyball* (en la que un joven experto en estadística, pero ignorante del beisbol, desplaza por la vía de mayor rendimiento en términos de juegos ganados a cazadores de talentos *expertos*), *big data* nos invita a reconsiderar nuestros instintos, dado que la analítica que escucha lo que los datos dicen está ocupando con ventaja cada vez más espacios antes considerados materia exclusiva de la experiencia y el juicio humanos. Esto significa, también, que las habilidades que se requerirán en el futuro próximo (y desde hoy) para tener éxito en el mercado de trabajo serán diferentes, con mayor peso en la habilidad para manejar datos de forma hábil y menor peso en el conocimiento de temas específicos. Para los autores, esto sin duda revalorizará el conocimiento de estadística combinado con programación y ciencia de redes y podrá significar un nuevo umbral de funcionalidad económica o el siguiente criterio de alfabetización que privará en las sociedades más prósperas y productivas. En este sentido, los países desarrollados de hoy arrancan en la carrera con ventaja, pues concentran la mayor parte de la información y saben cómo usarla.

Con todas sus ventajas, *big data* entraña también riesgos importantes para la privacidad y las libertades individuales. No se trata de un simple incremento cuantitativo, sino de un cambio radical en las reglas del juego. Esto implica que las leyes y reglas que hoy existen para proteger la privacidad ya no la salvaguardan en las nuevas circunstancias. Con *big data*, por ejemplo, no tiene sentido que los individuos den su consentimiento para tal o cual uso de la información que entregan, porque muchos de los usos ni siquiera han sido concebidos en la mente de sus autores al momento en que los datos son captados de origen. No es posible, entonces, dar consentimiento de algo que se ignora. Por otra parte, es poco práctico que Google u otro usuario regrese con cada individuo a pedirle su autorización cada vez que quiere hacer un nuevo uso de información a éste referida. Incluso, el pedir que se borre o ignore la información de alguien en particular puede terminar poniendo en ese individuo los reflectores que, en principio, quería evitar (como pudimos observar en un litigio reciente contra Google en México). Por otra parte, el concepto de confidencialidad vía anonimización también queda en entredicho ante la posibilidad de combinar datos de distintas fuentes que terminen por revelar la identidad que se busca proteger. Un nuevo signo de nuestros tiempos es que empresas, gobiernos, organizaciones legales e ilegales —e, incluso, individuos— pueden saber mucho más sobre cada uno de nosotros de lo que quisiéramos y que, a partir de ello, están en condiciones

de predecir aspectos de nuestras vidas que pueden usar con ventaja en nuestra contra.

Ante estas circunstancias, los autores proponen nuevas reglas del juego que salvaguarden el derecho de los individuos a decidir ante lo que pudiera predecirse de su conducta y que hagan responsables a los usuarios de la información por las consecuencias de los usos que hagan de la misma. Asimismo, sugieren la creación de auditores de información, a los que llaman *algoritmistas*, que jueguen el papel de defensores de los derechos de la población proveedora de la información original. "Estos nuevos profesionales serían expertos en las áreas de informática, matemáticas y estadística, y actuarían como revisores de los análisis y predicciones de *big data*. Los *algoritmistas* tomarían un juramento de imparcialidad y confidencialidad, similarmente a como los contadores y algunos otros profesionales hacen ahora. Ellos podrían evaluar la selección de las fuentes de datos, la selección de las herramientas analíticas y predictivas, incluyendo los algoritmos y modelos, y la interpretación de los resultados. En el caso de una disputa, ellos podrán tener acceso a los algoritmos, enfoques estadísticos y conjuntos de datos que produjeron una cierta decisión."[5] Sin embargo, no deja de causar cierto nerviosismo y escalofrío el que los autores propongan algo tan vago y lejano para atender a un riesgo tan actual, inminente y ominoso.

Lejos de ser una moda tecnológica, *big data* es algo que llegó para quedarse. Se trata de una verdadera revolución que está afectando muchas esferas de nuestras vidas. Es una fuerza de tal dimensión que no tiene sentido oponerle resistencia, sino buscar la mejor manera de adaptarse a ella y sacarle el máximo provecho. Todo parece indicar que los países más desarrollados están en mejores condiciones para sacar mayor ventaja de esta revolución, ampliando con ello las brechas de desarrollo en el mundo. Sin embargo, el hecho de estar ante un cambio tan grande implica también la posibilidad de tomar acciones decididas en naciones e instituciones menos desarrolladas que les permitan montarse en la cresta de la ola y acortar distancias. De manera específica para las oficinas nacionales de estadística, estamos ante un reto no definido con suficiencia todavía, pero que se antoja sin precedentes y crítico, pues pudieran perder relevancia ante la aparición de fuentes alternativas de información útil, más barata, diversa y oportuna basadas en *big data*. En este sentido, es fundamental tener presente que la misión de este tipo de instituciones es la de ofrecer al público información para el mejor conocimiento de la realidad y la eficaz toma de decisiones, no la generación de encuestas, censos y registros administrativos. Es algo similar a lo que ocurrió con el Pony Express con la llegada del ferrocarril que unía al este con el oeste de los Estados Unidos de América: resultaba fundamental que entendiera que su papel no era el tener los mejores caballos y los más diestros y audaces jinetes, sino comunicar a personas distantes unas de otras.

5 *Op. cit.*, p. 180.