

REVISIÓN DE LITERATURA DE MACRODATOS (*BIG DATA*)

SEBASTIÁN RAMÍREZ ARISTIZÁBAL

UNIVERSIDAD EAFIT

ESCUELA DE ADMINISTRACIÓN

MAESTRÍA EN ADMINISTRACIÓN (MBA)

PEREIRA

AGOSTO DE 2019

REVISIÓN DE LITERATURA DE MACRODATOS (*BIG DATA*)

**Trabajo presentado como requisito parcial para optar al título de magíster en
Administración (MBA)**

SEBASTIÁN RAMÍREZ ARISTIZÁBAL¹

Asesora temática y metodológica: Beatriz Amparo Uribe Ochoa, M. Sc.

**UNIVERSIDAD EAFIT
ESCUELA DE ADMINISTRACIÓN
MAESTRÍA EN ADMINISTRACIÓN (MBA)
PEREIRA**

¹ sebasramirez@hotmail.com

AGOSTO DE 2019

Dedicatoria

A mi esposa Mildred, mi gran amor.

Agradecimientos

A todos mis compañeros y profesores de esta gran experiencia académica. A Carlos Mario, que me dio la oportunidad de ingresar a este programa; a Juan Esteban, que ha tenido siempre una disposición amable para hacer de esta una experiencia extraordinaria.

Muy especialmente a Beatriz Uribe, asesora meticulosa, colaboradora y atenta.

A mi familia en general, por el apoyo en este proceso, en especial a mi hermana Catalina, fuente de conocimiento, admiración e inspiración.

A mi esposa, Mildred, por la paciencia y el apoyo incondicional. Por creer en mí siempre y por darme ánimo y alegría en los momentos más difíciles.

Contenido

Introducción	8
Metodología	10
Aproximación a los orígenes de macrodatos (<i>big data</i>)	11
Estado actual de la cuestión de macrodatos	21
Definición de macrodatos.	22
Características de la información.....	23
Volumen.....	23
Variedad	25
Velocidad	26
Valor	27
Cadena de valor de la información.	28
Registro.	29
Organización.....	29
Análisis	30
Resultado de valor	31
Problemas y retos de macrodatos	31
Cuestiones técnicas.....	31
Cuestiones éticas.....	32
Cuestiones culturales.....	33
Cuestiones epistemológicas.	34
Aplicaciones de macrodatos	35
Retos y tendencias futuras de macrodatos	37
Tipos de organizaciones	38
Segmentación y manejo de clientes	41
Educación de las organizaciones y su talento humano para la cultura de la gestión de la información.....	44
Trabajo interdisciplinar necesario para el abordaje de macrodatos	45

Conclusiones generales de la tendencia de macrodatos.....47

Consideraciones finales.....49

Referencias.....52

Resumen

Macrodatos (*big data*) es un tema que ha cobrado protagonismo en el último tiempo. El término hace referencia a un modelo de gestión de información que se vale del uso de datos voluminosos, variados, veloces y valiosos. Los sistemas de gestión de la información tradicionales no responden a entradas tan masivas y rápidas de información como lo representa la gran cantidad de datos generados hoy en día, gracias a los desarrollos tecnológicos y a la creación de ecosistemas virtuales en los que hay una inmensa cantidad de interacciones. La posibilidad de convertir esa gran cantidad de datos en información valiosa para los procesos de toma de decisiones de las partes interesadas es el gran cambio que macrodatos supone.

Palabras clave: macrodatos, información, volumen, velocidad, valor, registro, organización, análisis, resultado del valor.

Abstract

Big data is a subject that has gained prominence in the last time. The term refers to an information management model that uses voluminous, varied, fast and valuable data. The traditional information management systems do not respond to such massive and rapid inflows of information as is represented by the large amount of data generated today thanks to the technological developments and the creation of virtual ecosystems where there is an immense amount of interactions. The possibility of being able to convert that large amount of data into valuable information for the decision-making processes of the stakeholders is the great paradigm change that big data implies.

Keywords: *big data, information, volume, speed, variety, value, registration, organization, analysis, value result, data.*

Introducción

El presente trabajo trata un tema de actualidad y vigencia como lo es el de *big data* (en adelante macrodatos), que ha sido motivo de atención por parte del ámbito académico en los últimos tiempos y se habla mucho de su importancia, hasta llegar, incluso, a ser llamado la próxima gran revolución en administración. El tema propone una nueva forma de gestión que se genera gracias a los avances tecnológicos de los últimos tiempos y que en el futuro tienden a continuar. Sin embargo, hasta ahora, en esta fase todavía preliminar en la que se encuentra el asunto, se hace necesario seguir reflexionando alrededor del problema hasta que el entendimiento del fenómeno y la capacidad de manejar y sacar provecho de macrodatos sean suficientes. En el estado actual de la cuestión se evidencia que hasta el momento el tema llama la atención, despierta interés, pero todavía no se logra extraer valor significativo en todas las áreas de aplicación posibles de este recurso para el que ya existen medios tecnológicos de gestión; todavía se está en la etapa de identificación y enunciación de problemas alrededor de macrodatos (técnicos y de utilización) y empiezan a surgir los primeros intentos de solución. Por el momento, hay vanguardistas que ya están aplicando esta tecnología, pero todavía se está en etapa de exploración y construcción de modelos replicables.

La intención del trabajo es una revisión de literatura del tema, para lo cual se hizo un estudio de sus orígenes y se terminó con unas consideraciones sobre las tendencias futuras de la cuestión. Para describir el estado actual se utilizó la técnica de matriz conceptual para procesar los conceptos extraídos de la revisión de la literatura. Para seleccionar los textos por considerar en el estudio se eligieron artículos que fueran los más citados respecto al tema. Una vez revisados los noventa y tres artículos preseleccionados, se escogieron cincuenta para ser incluidos en la matriz conceptual del trabajo, al tomar en cuenta el número de

citaciones, el aporte que ofrecían para definir el tema y el año de publicación, porque se seleccionaron los más actuales. El análisis de la matriz arrojó cinco categorías: definición de macrodatos, características de la información, cadena de valor de la información, problemas y retos de los macrodatos y aplicaciones. Así pues, el ejercicio consistió en rastrear estas cuestiones en los artículos seleccionados para condensar en el trabajo el estado actual correspondiente.

Por último, se hizo una revisión sobre el uso, las tendencias y los problemas de los macrodatos.

Al final del ejercicio se observó que es un tema de vanguardia que está siendo motivo de estudio académico porque todavía se están identificando y formulando problemas tales como los aspectos técnicos y la generación de valor a partir de los macrodatos. Ha habido desarrollos tendientes a responder preguntas o a solucionar problemas surgidos del estudio de los macrodatos y en la actualidad se están haciendo usos prácticos que han permitido encontrar todavía más preguntas por responder. Hay un consenso en el estudio del tema acerca del potencial de generación de valor en la gestión de los macrodatos.

Metodología

Para dar cuenta del objetivo general de este ejercicio académico, referente a una revisión de literatura sobre los macrodatos, se propuso rastrear los antecedentes, construir una matriz conceptual de las nociones principales del tema e identificar problemas, usos y futuras tendencias de la cuestión.

En el rastreo de los antecedentes del concepto de macrodatos se encontraron sus antecesores: sistemas de soporte para la toma de decisiones e inteligencia de negocios. En este punto se planteó un recorrido general sobre estos conceptos con el fin de describir sus principales características a grandes rasgos. Esta parte se orientó más a dar cuenta del origen de los macrodatos que al estudio propio y detallado de los mencionados temas.

Más tarde se buscó en las principales publicaciones indexadas sobre el tema de los macrodatos, entendido como la gestión de la información de gran cantidad de datos que los medios tecnológicos actuales ponen a disposición y cómo la gestión de esos datos puede transformarse, a través de técnicas apropiadas de análisis, en información que genere valor para un interés particular. Para dicha búsqueda se usó la base de datos *Scopus*. Allí se seleccionaron los noventa y siete artículos que más se ajustaban a los criterios de selección del trabajo que son: vigencia (fecha de publicación), tema (administración) y pertinencia (mayor número de citas). Después de revisar estos artículos se seleccionaron los cincuenta que más pertinentes resultaron para el ejercicio académico.

Una vez seleccionados los cincuenta artículos se identificaron los principales conceptos referentes a los macrodatos, lo que dio como resultado la postulación de cinco categorías: definición de macrodatos, características de la información, cadena de valor de la información, problemas y retos y aplicaciones. Una vez que fueron identificadas se rastreó la concepción de ellas en cada uno de los artículos considerados con el fin de buscar puntos en común, así como divergencias, que permitieran establecer una discusión y, por último, algunos consensos que ayudaran a establecer el estado actual de la cuestión.

Los hallazgos se transcribieron en el cuerpo del trabajo.

Como último punto, el trabajo concluye con la revisión y la exposición de tendencias futuras del tema según lo encontrado en la revisión bibliográfica efectuada.

Aproximación a los orígenes de macrodatos (*big data*)

Aspectos generales sobre la construcción del concepto: sistemas de soporte para la toma de decisiones (DSS) a macrodatos.

La búsqueda por parte de la ciencia de la gestión de información tiene un origen difícil de establecer; se podría aceptar el famoso postulado cartesiano: *cogito, ergo sum* (pienso, luego existo) que plantea de manera explícita el bucle entre experiencia y reflexión como punto de partida. Sin embargo, esta reflexión cartesiana es precedida por ciclos humanos de acción y reflexión que han buscado hacer uso de la información en su beneficio:

En el sentido de representar más información de la que cualquier humano o grupo de humanos puede comprender, la noción de macrodatos ha existido desde los albores de la conciencia. El mundo y su universo son, para cualquier cosa o persona con sentidos, datos incomprensiblemente grandes. Sin embargo, el uso contemporáneo es distinto, porque marca el surgimiento de la posibilidad de dar sentido a un tesoro inexplicablemente grande de datos grabados, es decir, la promesa de poder utilizarlo de manera significativa, aunque ningún individuo o grupo de individuos pueda comprenderlo. En términos más prosaicos, los macrodatos denotan el momento en que las formas automatizadas de reconocimiento de patrones, conocidas como análisis de datos, pueden alcanzar formas automatizadas de recolección y almacenamiento de ellos (Andrejevic, 2014, p. 1675).

El problema de la gestión de la información se trata en este ejercicio académico en general, con énfasis en el ámbito industrial desde la óptica de la administración. Para ello, en el presente capítulo se hace un recorrido de la evolución del fenómeno, por medio de la revisión de los conceptos de sistemas de soporte para la toma de decisiones (DSS), inteligencia de negocios (BI) –o inteligencia de negocios y analítica (BI&A)– y macrodatos.

En principio, en el contexto industrial, la gestión de la información y la unión con los procesos de toma de decisiones de las empresas se denominó sistemas de

soporte para la toma de decisiones, en adelante DSS². Para revisar este concepto se partió de una definición, tarea difícil ya que como lo exponen Kelemenis y Askounis (2008), «un grave problema de definición es que las palabras tienen una cierta "validez intuitiva"; cualquier sistema que apoye una decisión, de cualquier manera, es un "DSS"»³ (p. 2). La definición en la que se basó el presente trabajo se tomó de uno de los artículos académicos más citados en la materia: “Los sistemas de soporte de decisiones (DSS) son soluciones de tecnología informática que pueden utilizarse para respaldar la toma de decisiones complejas y la resolución de problemas” (Shim, Warkentin, Courtney, Power, Sharda y Carlsson, 2002, p. 111)⁴.

Esta definición es bien amplia, pues, si se revisa bien, se observa que se compone de dos variables: “soluciones de tecnología informática” y “respaldar la toma de decisiones complejas y la resolución de problemas”. En primera instancia, “soluciones de tecnología informática” es un concepto amplio y complejo, porque, en esencia, es cualquier proceso de gestión de la información, que incluye desde aparatos (*hardware*): computadores, celulares, tabletas, etc. hasta procesos de tecnología blanda (*software*): programas, aplicaciones (*apps*), reportes, métricas, indicadores críticos de desempeño (KPI), entre otros aspectos. El otro elemento de la explicación es la parte de “respaldar la toma de decisiones complejas y la resolución de problemas”, bien amplia también, que aplica para cualquier tipo de fenómeno en todo tipo de empresa humana; en este sentido, la característica diferenciadora del concepto es que trate un problema o situación con cierto grado de dificultad, lo que hace sentido: invertir más recursos en temas más complejos e impactantes y menos recursos en temas más simples que no afecten tanto un proyecto; los mismos Shim et al. (2002) plantea esta característica cuando se refiere a una primera clasificación de problemas de gerencia, al dividirlos en dos

² Sigla en inglés de *decision support system*.

³ Traducción no oficial del autor.

⁴ Traducción no oficial del autor.

tipos: el primero es el de los problemas “programados” (todas las tareas programadas rutinarias, estructuradas y de simple resolución) y el segundo el de los “no programados” (contingentes, emergentes y complejos). Después hubo una nueva reclasificación que consta de problemas: estructurados, semiestructurados y no estructurados; el uso de DSS debe contemplar algún tipo de interacción de máquinas de cómputo y la resolución de un problema de naturaleza semiestructurada o no estructurada.

¿De cuáles tipos de problemas se encargan los DSS? Se introdujo a la discusión el concepto de “máquina”. Se aclara que esta palabra se usó en esta reflexión desde la perspectiva de una herramienta importante para la toma de decisiones, con el énfasis en la vertiente administrativa y no en la tecnológica del asunto. De hecho, en el texto antes citado se expone que la investigación científica en torno a este problema se ha dado desde dos intereses o puntos de vista: el desarrollo técnico y tecnológico de máquinas, programas y conexiones que ayuden al proceso de toma de decisiones y el estudio y desarrollo científico del problema de la toma de decisiones organizacionales. En particular, el presente artículo aborda las dos perspectivas. El énfasis y el interés del autor se ubicaron en la segunda perspectiva, de modo específico en lo relacionado con el ser humano y el uso de esta herramienta en la administración de organizaciones y proyectos.

La gestión de la información para la toma de decisiones de valor es un concepto amplio en el sentido que puede ser aplicable a casi cualquier empresa humana y porque provee una gran versatilidad y pragmatismo, puesto que se trata de una herramienta que puede ser usada en beneficio de múltiples proyectos de naturalezas diversas.

Desde la década de 1960, con los primeros estudios de sistemas de soporte de decisión (DSS) hubo varios esfuerzos por introducir modelos de gestión de la información en empresas para toma de decisiones estratégicas; entre otros, cabe mencionar la aparición del sistema MIS (*management information system*) de IBM en abril de 1964. Este sistema proveía a gerentes de grandes organizaciones de

reportes –sobre todo financieros y transaccionales– aunque no permitía todavía una relación interactiva que les suministrara a los gerentes soporte para la toma de decisiones de manera activa (Power, 2007).

En la década de los setenta hubo estudios aplicados sobre la aplicación de los DSS en campos tales como el mercadeo y el mercado de valores. El libro de Keen y Scott Morton (1978) ayudó a poner en primera plana el problema de la gestión de la información:

El objetivo de los primeros desarrolladores de DSS era crear un entorno en el que los tomadores de decisiones humanos y los sistemas basados en tecnología de la información trabajaran en forma interactiva para resolver problemas; El ser humano trata con las partes complejas y no estructuradas del problema, el sistema de información proporciona asistencia automatizando los elementos estructurados de la situación de decisión⁵ (Arnott y Pervan, 2016, p. 68).

Todos los esfuerzos de los sesenta y los ochenta permitieron que hubiera un auge de gestión de la información para toma de decisiones que generaran valor a partir de los ochenta. Con el camino allanado en el tema, empezaron a emerger estudios más especializados, nomenclaturas más específicas, búsquedas de aplicaciones en áreas más delimitadas, especialización del conocimiento y proliferación de producción científica sobre el asunto.

La publicación sobre DSS ha caído en forma significativa con el tiempo. En el período 1990-1994 se publicaron 75.6 documentos DSS por año, en el de 1995-1999 el número aumentó a 84.6 por año y en el de 2000-2003 el número de artículos por año se redujo de manera drástica a 54.7 (Arnott y Pervan, 2016).

⁵ Traducción no oficial del autor.

Si bien la producción científica acerca del concepto llamado DSS tuvo un importante declive, vino después otra noción, inteligencia de negocios (en adelante BI), que es también una solución al problema de la gestión de la información al servicio de la toma de decisiones, en forma específica en el contexto empresarial.

El término inteligencia se encuentra en investigaciones relacionadas con el campo industrial desde los cincuenta (Ghazal, Rabl, Hu, Raab, Poess, Crolotte y Jacobsen, 2013); sin embargo, su gran auge ocurrió partir de los años ochenta y se ha venido sosteniendo en el tiempo hasta la actualidad.

La expresión inteligencia de negocios (BI) también ha sufrido evoluciones y transformaciones desde su aparición; los avances tecnológicos y científicos han permitido que, a partir de los esfuerzos pioneros por abordar este problema, se hayan realizado abordajes cada vez más especializados que en el día de hoy han posibilitado especializar el conocimiento y tener perspectivas cada vez más sofisticadas al respecto, como, por ejemplo la introducción de la variable denominada analítica que llevó el término a evolucionar hacia inteligencia de negocios y analítica (BI&A).

Chen, Chiang y Storey (2012) propusieron tres momentos en la evolución del BI&A:

- BI&A 1.0: un énfasis en la información. Proviene del estudio de manejos de bases de datos, sistemas que se popularizaron en los noventa mediante la utilización de modelos estadísticos desarrollados en los setenta.
- BI&A 2.0: con el auge del internet se comenzó a presentar una oleada en la recolección de datos y la investigación analítica de oportunidades de desarrollo. La interacción en la red de las organizaciones en forma directa con los clientes permitió esta recolección de gran cantidad de datos con grandes velocidades. Este auge se remonta a principios de 2000. “El desarrollo de internet y los métodos para acceder a la red difundieron la mediación informática desde sitios de trabajo

acotados y acciones especializadas hasta la ubicuidad global, tanto en la interfaz institucional como en las esferas íntimas de la experiencia cotidiana”⁶ (Zuboff, 2015, p. 85).

- BI&A 3.0: el auge de BI&A 2.0 fue posible gracias al de internet; en el último tiempo, se evidenció una nueva tendencia gracias a los dispositivos móviles y los usables y la mayor parte de la generación de datos proviene de estos dispositivos. Esta nueva etapa se remonta al año 2011.

La evolución de la tecnología ha permitido que cada vez se produzcan más datos, con mayor rapidez y que sea viable su almacenamiento, lo que ha generado retos que han exigido que los modelos de análisis para volver útil información tengan que evolucionar también a partir de los nuevos panoramas de generación rápida de información. Los sistemas de gestión de la información que se proponían desde el tiempo del DSS hoy en las empresas han tenido que ser adaptados a las nuevas condiciones de velocidad. Las aplicaciones en áreas funcionales de las organizaciones también representan una novedad con el fin de lograr que haya avances en materia de tránsito, salud, seguridad nacional, comercio electrónico, mercadeo, etc.

Todo este preámbulo que se ha venido describiendo desde los pioneros de DSS hasta llegar a BI&A ha ido de la mano de los desarrollos tecnológicos que han permitido nuevos avances y dado pie a la emergencia de una nueva herramienta, que se apoya en los cimientos establecidos por otras previas pero que buscan responder a las características actuales de la información y, así mismo, a las exigencias de los problemas actuales: macrodatos (BD): “Macrodatos comienza con fuentes autónomas de gran volumen, heterogéneas, con control distribuido y descentralizado, y busca explorar relaciones complejas y en evolución entre los datos”⁷ (Sowmya y Suneetha, 2017, p. 98). Este fenómeno describe la aparición

⁶ Traducción no oficial del autor.

⁷ Traducción no oficial del autor.

de datos masivos y hay ciertos criterios que debe cumplir el fenómeno para considerarse macrodatos (Gandomi y Haider, 2015):

1. Volumen: gran cantidad de datos: los avances en tecnologías de almacenamiento han permitido que cada vez más datos puedan conseguirse y guardarse. “Las tecnologías modernas de uso intensivo de datos y el aumento de los recursos computacionales y de almacenamiento de datos han contribuido en gran medida al desarrollo de la ciencia de los macrodatos”⁸ (Furht, Villanustre, Najafabadi, Khoshgoftaar, Seliya y Wald, 2016, p. 2).
2. Variedad: datos de diferentes tipos: datos estructurados, como bases de datos organizadas, y no estructurados, como audio, video o imagen (Gandomi y Haider, 2015).
3. Valor: información que genere algún tipo de interés para un fin específico, para el caso de un negocio que aporte valor en la cadena o genere algún tipo de ventaja competitiva (Gandomi y Haider, 2015).
4. Velocidad: datos generados de manera muy rápida, incluso accedidos en tiempo real (Gandomi y Haider, 2015).

Los macrodatos son, en comparación con de las herramientas antes descritas, DSS, inteligencia de negocios y macrodatos, la más actual y la que demuestra una tendencia a seguir creciendo en términos de utilización industrial e interés académico. La cadena de valor de macrodatos, según Hu, Wen, Chua y Li (2014) tiene los siguientes elementos: recolección, almacenamiento, recuperación, análisis y utilización de la información y brinda oportunidades de investigación en su totalidad y también en su abordaje parcial en cada uno de sus eslabones; no solo desde una perspectiva tecnológica (avances en *hardware*) sino también desde los puntos de vista de administración y de la utilización de este recurso en las empresas, “¿Por qué deberían los académicos y los profesionales estar interesados en comprender los impactos de los macrodatos? La respuesta simple

⁸ Traducción no oficial del autor.

a esta pregunta crítica es porque los macrodatos tienen el potencial de transformar todo el proceso de negocios”⁹ (Fosso Wamba, Akter, Edwards, Chopin y Gnanzou, 2015, p. 234).

Para el análisis del estado actual de los macrodatos se abordó lo que se ha dado en llamar problemas de los macrodatos:

El análisis de datos es considerablemente más desafiante que limitarse a citar, identificar y comprender datos. Para un análisis efectivo a gran escala, todo esto tiene que suceder de una manera completamente automatizada, lo que requiere que las diferencias en la estructura de los datos y la semántica se expresen en formas que sean comprensibles por medio de la computadora y luego se resuelvan en forma robótica. Incluso para análisis más simples, que dependen de un solo conjunto de datos, sigue existiendo una cuestión importante sobre el diseño adecuado de la base de datos (Parimala, Rajkumar, Ruba y Vijayalakshmi, 2019, p. 2032).

Estos problemas de los BD se remontan a todos los eslabones de la cadena de valor de la información y existen retos para la tecnología (dura y blanda) de recolección de datos, almacenamiento, recuperación, análisis y utilización: “Entre estos desafíos, algunos están causados por las características de los macrodatos, otros por sus modelos y métodos de análisis actuales y otros por las limitaciones de los sistemas disponibles de procesamiento de datos”¹⁰ (Jin, Wha, Cheng y Wang, 2015, p. 62).

⁹ Traducción no oficial del autor.

¹⁰ Traducción no oficial del autor.

Las nuevas formas rápidas de generación de datos a partir de interacciones por medio de las tecnologías emergentes disponibles también suponen un reto para las empresas en la relación con sus partes interesadas:

El contexto en el que se entrega y experimenta el servicio ha cambiado, en muchos aspectos, de manera fundamental. Los avances en tecnología, en especial la de la información, están conduciendo a una proliferación de servicios revolucionarios y cambian la forma en que los clientes se sirven antes, durante y después de la compra”¹¹ (Ostrom, Parasuraman, Bowen, Patrício y Voss, 2015, p. 127).

Estado actual de la cuestión de los macrodatos

El concepto de macrodatos, al ser vanguardista requiere una revisión de literatura de características especiales. Según los lineamientos generalmente aceptados para los ejercicios académicos de información (IS), la investigación de temas novedosos y que representan frontera de conocimiento puede considerar unas cuantas fuentes representativas (Salkind, 2012). En el presente trabajo se sistematizaron cincuenta artículos de investigación en los que se tomó en cuenta, como criterio de validez, la cantidad de veces que han sido citados en publicaciones indexadas; para dicho ejercicio se utilizó la base datos *Scopus*. Además del número de citas, se consideró la materia en la que se inscribe el artículo mediante la selección, no solo de publicaciones “fundacionales” de macrodatos, sino también los macrodatos como tema enmarcado en el desarrollo conceptual de áreas del conocimiento tan variadas como turismo, salud, sociología, geografía, etc. Este método de selección permitió, no únicamente nutrir la discusión desde el punto de vista conceptual, sino observar la noción de macrodatos de manera aplicada y no solo conceptual.

¹¹ Traducción no oficial del autor.

Una vez seleccionados los artículos por considerar, para lo que se tuvo en cuenta lo antes expuesto, se procedió a la construcción de dicha matriz conceptual. Para ello se hizo también una elección de los conceptos encontrados como relevantes y comunes en el corpus conceptual del tema de los macrodatos. Las categorías escogidas fueron: definición de macrodatos, características de la información, cadena de valor de los macrodatos, problemas y retos de los macrodatos y uso de los macrodatos.

Definición de macrodatos

El concepto macrodatos hace referencia a un sistema de gestión de la información, en particular de grandes volúmenes de ella generados en ambientes digitales o ecosistemas de medios digitales: “los macrodatos se utilizan por lo común para referirse a los grandes volúmenes de datos generados y disponibles en internet y en los ecosistemas de medios digitales actuales” (Constantiou y Kallinikos, 2015, p. 1). La particularidad que se encuentra en este concepto es el énfasis que se hace en el volumen de los datos y en la velocidad con la que se generan. Se encuentran en la literatura muchas referencias a cómo la proliferación de aparatos que de alguna manera registran información hacen que estos volúmenes de datos crezcan de manera exponencial.

Los aparatos que registran datos son muy diversos. Existen los ordenadores tradicionales, en los que los usuarios registran datos –por ejemplo, el uso de mensajes de correo electrónico o las compras en el comercio electrónico– o en los que los datos se los solicita al usuario un tercero que los registra –una transacción en un hotel o un envío de un paquete por correo–. Además de estos escenarios típicos, en los que los usuarios registran su propia información o en los que ella la requiere un tercero, y que se observan desde hace tiempo, hay ahora una gran proliferación de aparatos usables (*wereables*) que registran información de forma

automática (Zuboff, 2015). Quizá el ejemplo más evidente sea el caso de los relojes inteligentes que, por lo general, funcionan como un satélite de los teléfonos inteligentes. Estos aparatos registran los datos del usuario en tiempo real, con información como ritmo cardíaco, ubicación geográfica, cuántas llamadas realiza y en promedio cuánto duran, hábitos de sueño, de alimentación, en fin, las posibilidades son casi infinitas y estarían a disposición de los desarrolladores de este tipo de programas. De ahí que, frente a este fenómeno del crecimiento exponencial de la generación de información, emerja también la necesidad de elaborar estrategias para gestionarlo y tratar de sacar valor de él si se entiende que los métodos tradicionales de gestión de información se quedan cortos ante los volúmenes y las velocidades de generación actuales y que en el futuro proyectan un crecimiento aún mayor con el continuo desarrollo de tecnología, como, por ejemplo el internet de las cosas (IoT) (Erevelles, Fukawa y Swayne, 2016). Es evidente, para diversos grupos de interés, el valor que existe en la correcta explotación de estos datos; es ahí cuando el concepto de macrodatos adquiere su relevancia.

En seguida se revisan las gestiones de esas cantidades masivas de información, varios usos posibles de esas conclusiones de valor extraídas de las técnicas de análisis y algunos retos o problemas hallados en la revisión de literatura.

Características de la información

Como antes se mencionó, el factor común y base para hablar en forma rigurosa de macrodatos es el volumen de ellos (Sivarajah, Kamal, Irani y Weerakkody, 2017). Se puede decir que la diferencia entre los estudios anteriores de gestión de información y los macrodatos yace en particular en la naturaleza voluminosa de los datos. Aun cuando los autores llegan a referirse a otras características de los datos, terminan por describir toda la escena de los retos a los que se enfrenta dicho desarrollo para poder traducirse en valor, que ha generado el interés por desarrollar modelos que puedan llegar a abarcar este fenómeno en provecho de los intereses científicos e industriales.

Volumen

En la literatura se encuentra, de maneras explícita e implícita, la noción de que los datos de los macrodatos son voluminosos. Bien se pueden encontrar adjetivos como voluminosos, vastas cantidades o masivos. Por lo común se acepta que el gran tamaño actual de los datos ha hecho que surjan retos frente a su gestión, lo que requiere una reformulación de cada una de las etapas de la cadena de valor de la información (Gandomi y Haider, 2015).

El auge y la democratización del acceso a aparatos “inteligentes” que registran datos, así como la conexión cada vez más permanente a internet, ha facilitado la obtención, el registro y el almacenamiento de información que, por ser generado por el usuario, representa gran valor.

La información en ese estado “crudo” es una gran materia prima para los procesos de decisión, pero requiere ser procesada para convertirse en el producto final: información útil y valiosa para los ejercicios de decisión que se necesitan según el interesado.

Es el volumen precisamente lo que plantea el reto de la transformación de esos volúmenes de datos crudos en información de valor pues se encuentra el obstáculo de que los sistemas de procesamiento de información tradicionales no son efectivos para atender una entrada de datos tan masivos. Hasta aquí, los macrodatos no llegan sino a ser un fenómeno interesante pero no una revolución paradigmática que pueda llevar a la transformación de los modelos tradicionales en los que se mueve cada uno de los particulares interesados en obtener modelos de toma de decisión o de predicción.

Es claro que la literatura al respecto plantea en forma implícita que el volumen en sí mismo hace que la información sea más valiosa, puesto que permite que las muestras de estudios –científicos, de mercado, etnográficos, etc. – sea más grande, lo que refuerza la credibilidad de los hallazgos y reduce el margen de error de las predicciones (Wang, Xu, Fujita y Liu, 2016). Esta promesa

futura de valor ha ayudado a impulsar el interés por buscar formas de aprovechar esos datos, de los que ya es posible, gracias a los avances tecnológicos, obtener registros.

Variedad

El segundo lugar en el podio de los adjetivos más utilizados en la literatura revisada para describir los datos de los macrodatos es variada, que aparece mencionada de alguna forma el 66% de las veces, ya sea como variedad, variada, no estructurada, desorganizada o compleja.

El término variedad se refiere a la naturaleza diversa de los datos que pueden ser obtenidos en formato escrito, de audio, de video, etc. Si se considera esta característica junto con la anterior, volumen, se ve que el panorama para los retos de los macrodatos es cada vez más complejo. Ahora se requieren sistemas que no solo puedan procesar grandes cantidades de información sino tipos diferentes de ella (Jin et al., 2015).

Para un mismo interés, los registros de datos valiosos a los que se puede acceder son diferentes; por ilustrar el argumento, conviene pensar en el caso de un estudio de mercado acerca de la forma en la que un cliente potencial (usuario) puede llegar a expresar su interés o desinterés por consumir un producto que será lanzado puede expresarse por escrito, por medio de audio o en un registro de video. Ahora bien, para terminar de ilustrar la idea, y solo por hablar del caso del registro escrito, este puede expresarse en el papel de un encuestador, mediante la respuesta a un mensaje de correo electrónico o con una publicación en una red social. Con seguridad, la cuestión de fondo en el ejemplo que se plantea es muy simple: interés o desinterés por consumir un producto específico, pero no es posible estimar las cantidades de formas en las que un cliente potencial puede expresar ese interés o desinterés. Si se piensa ahora que, si se reduce un poco más la complejidad del escenario hipotético, los clientes del potencial producto solo serán estudiados y medidos a través del análisis de sus publicaciones en

determinada red social, ello plantearía no solo la pérdida de valor en términos de muestra para la decisión, sino también el hecho de seguir encontrando complejidad en la forma de expresión de cada usuario, sea que se muestre abiertamente desinteresado en probar un producto o que su mensaje tenga un subtexto en el que manifiesta estar desinteresado.

Planteado de esta forma, se trata de un reto enorme procesar información tan diversa y tan voluminosa. Sin embargo, la cuestión no termina ahí, porque aún quedan por considerar un par de características de los datos que siguen complejizando la cuestión.

Velocidad

El término que comparte segundo lugar en con el anterior, Variedad, es velocidad que también aparece mencionado de alguna manera en el 66% que los artículos considerados. Algunas de las formas en los que los autores se refieren a estas características son: velocidad, rápida, generada a gran velocidad, tiempo real, veloz (Daniel, 2015).

Como se explicaba en el apartado de volumen, la proliferación de aparatos con capacidad para obtener datos y la conexión cada vez más extendida a internet, así como generan una cantidad masiva de datos, también permiten que ello ocurra con gran velocidad.

Este fenómeno de la velocidad de la generación de datos plantea algunos problemas interesantes. Por un lado, está la cuestión de la vigencia: ¿cuál información es permitente y cuál no? y ¿cuál es el tiempo de vida de un dato? Poder responder estas preguntas se acerca más a una discusión filosófica sobre la cuestión que una técnica. Pero si pensar en que el modelo de procesamiento de datos, además de ser capaz de tomar grandes volúmenes de datos de diferentes naturalezas (variedad), debe hacerlo de manera muy rápida, tanto las tareas de la captura como la del análisis y la entrega del resultado se complican aún más. En principio, se trataría de sistemas que contengan entradas de datos de manera

permanente y salida de resultados también del mismo tipo para que esta información que resulta y que es valiosa para los procesos de toma de decisiones sea también pertinente, puesto que si no fuera oportuna perdería su criterio de valor.

Valor

El último atributo que se expone en la literatura de los macrodatos es valor; en la literatura revisada se referenció de alguna forma en el 32% de las veces, aunque debe hacerse la salvedad de que es posible que esta característica sea dada por sentada desde que se aborde el tema en un estudio. Se encontró en la revisión que se habla de: valor, datos valiosos, datos que generen valor, datos verídicos (Fosso Wamba et al., 2015).

Como antes se ha mencionado, uno de los valores de estos datos reposa precisamente en su volumen porque constituyen muestras más robustas, pero acá se refiere al valor que tiene cada dato de por sí mismo y no en conjunto. Es decir, el dato es oportuno, es verdadero y sirve para el fin que se requiere. ¿Da luces a conclusiones de valor en un proceso de toma de decisión para un interesado? Hay una condición propia que se plantea en muchas de las explicaciones encontradas sobre el origen de los datos y es que son generados por el “usuario”, lo que le da un peso de valor al dato, es decir, que en muchos de los escenarios las decisiones que se plantea un interesado se toman con entradas generadas por los mismos sujetos a los que se destina la decisión.

En el caso de una farmacéutica que esté decidiendo cuál producto farmacológico desarrollar y tenga acceso a la información de relojes inteligentes de una muestra x de una población y, podrá, de manera directa, rápida y veraz, determinar cuál necesidad tiene esa población foco; en el caso de un diseñador de transporte público que pueda acceder a la información de desplazamiento que registre un aparato con un programa de georreferenciación, como pueda ser un teléfono inteligente, tendrá información más valiosas y fidedigna que la que se pueda

obtener de la boca de viva voz de un ciudadano, se presume que la solución planteada en términos de movilidad sea también más eficaz.

Parte del valor del dato está en esa “pureza” que se obtiene al ser generado por el usuario. Los usos de esto son tan numerosos que tal vez solo la imaginación humana y la ciencia podrían determinar el límite. Es necesario pensar –entre otras aplicaciones posibles– en la respuesta a las necesidades de capacitación de una población, a la atención a individuos en condición de discapacidad, en la optimización de inventarios de una empresa de manufactura, en prevención del crimen en las ciudades o en la optimización de la producción agrícola.

Cadena de valor de la información

En el corpus teórico de la cuestión se encontró un concepto, a veces implícito y en otras ocasiones explícito, al que en la mayoría de los textos se refiere como cadena de valor de los macrodatos.

Dicha cadena de valor se refiere a los elementos de un modelo de gestión de la información y describe la operación de cada uno de ellos y cómo esta operación representa una adición de valor para el producto final:

grandes conjuntos de datos integrados para soportar series y análisis específicos. Todas estas acciones fluirán desde una cadena de valor de datos: un marco para administrar los datos de manera general desde la captura hasta la toma de decisiones y para respaldar a una variedad de partes interesadas y sus tecnologías (Miller y Mork, 2013, p. 57).

Como se ha venido ilustrando, la cantidad masiva de datos por sí solos no pasan de ser un fenómeno llamativo; los macrodatos como tales se refieren a la gestión de esta cantidad voluminosa de datos con la inclusión de un producto final que la transformación de la cadena debe entregar llamado información. A continuación se presenta cómo la cadena transforma esos datos con potencial valioso (entrada) en información de valor para las partes interesadas (salida).

En el recorrido bibliográfico realizado se encontraron diferentes propuestas de cadena de valor de los macrodatos; si se hace un ejercicio de reducción de la complejidad se propone la agrupación en las siguientes categorías de análisis: registro, organización, análisis y resultado de valor.

Registro

También llamado recolección, captura, almacenamiento, adquisición, generación u observación. Se trata de la captura de la información. Es posiblemente la más evidente de las etapas de la cadena, pero no por eso menos valiosa, pues debe encargarse de no perder capturas y, además, de atender canales de recolección de información tan variada como la naturaleza misma de los datos. Es la etapa inicial que permite la aparición de las demás (Erevelles et al., 2016).

Organización

También referida como extracción, almacenamiento, movilización de los datos, distribución, minería, agrupación, transmisión, manejo, curación, almacenamiento, búsqueda, compartición y transferencia (Miller y Mork, 2013).

Esta fase de la cadena es la que menos veces apareció registrada en la literatura, aunque de igual forma fue referida en muchas ocasiones, con 56% de menciones. Es la parte que viene después de la recolección y su valor reside en estructurar la información conseguida en la etapa anterior. El proceso que aquí opera es valioso porque responde a problemas que se generan a partir de la naturaleza de los datos, por ser voluminosos y variados; en esta etapa de la cadena esa gran cantidad de datos, esos datos tan diferentes entre sí deben ser calificados y ordenados para que la fase siguiente, la del análisis, sea efectiva.

Es importante mencionar que esta clasificación que se hace con fines académicos es un artilugio que pretende, como se ha dicho, reducir la complejidad por medio de un proceso de abstracción de un universo muy rico y enmarañado encontrado en la literatura, puesto que se constató que para muchos autores y textos se vio el paso de organización y análisis –que es el que en seguida se explica– como uno

solo. Se propone esta categorización porque para algunos autores se encuentra especificada pero más aún porque se reconoce el valor que aporta al dar organización a uno datos no estructurados y voluminosos que permiten que a través de los análisis se transformen en información de valor.

Análisis

También nombrado procesamiento o interpretación. Es la operación que transforma los datos estructurados en la etapa anterior en información de valor que puede ser usada en un proceso de toma de decisiones o en un modelo predictivo (Puschmann y Burgess, 2014). Existen numerosas metodologías de análisis de la información y muchas de ellas se examinan con mayor o menor detalle en los textos de macrodatos orientados en mayor medida hacia tecnologías de la información. En términos generales, se observa que la metodología de análisis seleccionada va unida con la naturaleza de la materia interesada en los datos. Se encuentran referencias en la literatura de adaptación de modelos tradicionales a cantidades masivas de datos, lo que se facilita gracias al éxito de la etapa anterior, organización de los datos.

Resultado de valor

También hallado como uso, aplicación, toma de decisión, extracción de conocimiento, valor o visualización. Es la aplicación del resultado de toda la operación de la cadena de valor (Rust y Wang, 2014). Así como el paso que se refiere a la recolección es obvio, este último que se refiere a la aplicación también lo es, pues es el motivo por el que se realiza todo el esfuerzo de la cadena. Aun cuando pueda ser muy obvia esta etapa, apareció referenciada en el 76% de los casos de los artículos incluidos en el presente estudio. Es clave resaltar que no se trata de un simple uso sino de un uso de valor con apoyo en la hipótesis antes planteada, que contempla que el tamaño masivo de los datos propende por un resultado más veraz y valioso.

Problemas y retos de los macrodatos

En cuanto los problemas y los retos que enfrentan los macrodatos, se encontraron cuatro cuestiones principales alrededor de las que reflexionó la literatura; las cuatro categorías de análisis construidas fueron: cuestiones técnicas, cuestiones éticas, cuestiones culturales y cuestiones epistemológicas.

Cuestiones técnicas

Como es lógico por la naturaleza de los artículos considerados para este trabajo, el reto más común en el corpus teórico revisado fue el tema técnico. En esencia, la pregunta en esta categoría versó sobre cómo poder convertir la cantidad masiva de datos a los que se tiene acceso en conclusiones (información) valiosa para procesos de toma de decisiones para las partes interesadas.

Los retos técnicos fueron comunes a lo largo de la cadena de valor que, para efectos del presente trabajo, se estableció como: registro, organización, análisis y resultado de valor, etapas que se unen a las características particulares de los datos: volumen, variedad, velocidad y valor, de modo que genera una cantidad de retos que los desarrolladores de esta corriente deben enfrentar. Se encontraron muchas alusiones, sobre todo en los textos de tecnologías de la información (IT), a los problemas y las posibles soluciones para el almacenamiento, la organización, la recuperación de los datos almacenados y las técnicas de análisis para datos masivos, variados y generados con gran velocidad. Como se ha venido enunciando, los métodos tradicionales son insuficientes. Estos textos no solo enunciaron el problema de adaptar las tecnologías tradicionales a esta nueva clase emergente de datos, sino que algunos llegaron a proponer modelos para enfrentar dicha cuestión. De hecho, se logró identificar que la razón de ser de la mayoría de estos artículos o la pregunta que pretendieron responder (o alrededor de la que reflexionaron) fue justamente la de cómo convertir los macrodatos en información de valor.

Cuestiones éticas

La segunda cuestión más protagónica en la literatura revisada fue la que se preguntó a propósito de las implicaciones éticas de todo el fenómeno de los macrodatos. Se encontró mencionada de alguna manera en el 16% de los textos. La pregunta que acá se planteó es sobre cuál es el límite entre el uso ético de los datos generados por el usuario y cuándo se trata de una violación de la privacidad.

A medida que va transcurriendo el tiempo y este fenómeno va calando en todos los registros de la sociedad, ella misma va respondiendo a las necesidades que van surgiendo y es así como van apareciendo legislaciones que intentan poner límites entre el uso “apropiado” de datos por un tercero, cuáles son los escenarios que deben ser resguardados (incluso por la ley) cómo íntimos o consideraciones sobre todo lo referente al consentimiento informado. En Colombia, por ejemplo, en el 2012 se promulgó la ley del *habeas data*, que alcanzó el rango de ley estatutaria, que fue la ley 1581 del 17 de octubre de 2012, “por la cual se dictan disposiciones generales para la protección de datos personales” (Congreso de Colombia, 2012).

El tema tiene dos versiones: el problema de los usuarios, que ven amenazada su privacidad, y el de los gestores de información, que encuentran barreras para poder acceder a la información. La mediación entre dos intereses se ha intentado discernir con protocolos de autorización de acceso a datos; es común encontrar situaciones en internet en las que un usuario autoriza o no cierto tratamiento de la información, entre otras posibilidades, la de entregarla a terceros. Una vez zanjada esa encrucijada, todavía queda la cuestión sobre el uso apropiado (ético) de los datos a los que el usuario ha autorizado acceso; a este respecto, uno de los artículos incluidos en este estudio se refiere a la cuestión del desbalance de poderes entre aquellos que proveen los datos contra los que demandan los datos y las consecuencias del uso que ellos hagan (Andrejevic, 2014).

Cuestiones culturales

En esta categoría se agruparon las referencias halladas acerca de la aceptación social de una cultura de la información. Preocupaciones como la habituación de los grupos sociales al uso de información relevante en los procesos de toma de decisión, aceptación de cambios organizacionales producto de las conclusiones de la gestión de la información o simplemente convencer a un tomador de decisiones de un grupo de basarse más en información que en la intuición pueden ser unos de los escollos que encuentre la correcta y oportuna implementación de macrodatos (Chen, Preston y Swink, 2015). En algunos pasajes de los textos se encontró referencia a este fenómeno como “falta de adopción de macrodatos”, una especie de rechazo y desconfianza frente a la aceptación de este nuevo paradigma. Argumentos tales como la crítica de la obsesión por la medición y la información (Sharon y Zandbergen, 2017), incompreensión sobre cómo las empresas pueden generar valor a partir de la gestión de la información (Günther, Rezazade Mehrizi, Huysman y Feldberg, 2017), modelos que se replican solo una vez o que solo sirven para una organización o la desconfianza hacia los macrodatos (Braganza, Brooks, Nepelski, Ali y Moro, 2017), cuestionamientos por la privacidad de la información, uso ético de la información (Janssen y van den Hoven, 2015), sobrecarga de información (Opresnik y Taisch, 2015), falta de conceptualización y exceso de datos prácticos (Hargittai, 2015) fueron algunos de los argumentos presentados para justificar la renuencia cultural a la adopción de esta modalidad.

Cuestiones epistemológicas

Por último, se encontraron los argumentos en torno a la preocupación por el rigor epistemológico en la creación de conocimiento nuevo. Las referencias a este tema se pudieron rastrear tan solo en el 6% de los textos seleccionados. Aun así, se creó la categoría de análisis porque de alguna forma está presente la alerta sobre los problemas que pueden surgir y por el interés que todo texto y cada autor académico tienen por el bienestar de la ciencia y el conocimiento.

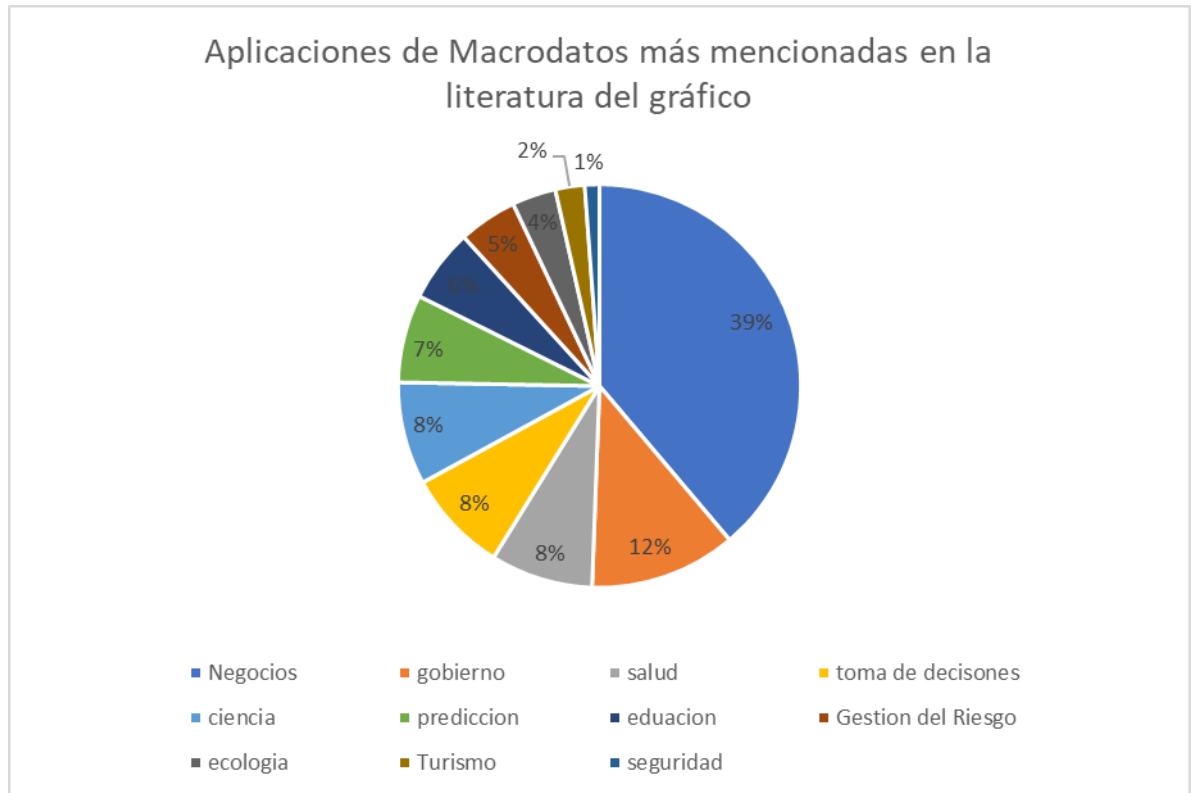
En pasajes anteriores se mencionaba como los macrodatos pueden beneficiar la investigación científica por medio de la robustez de las muestras a las que se puede acceder para un estudio. Además de ello, se reconoció la democratización del acceso a la información por la disponibilidad de información para poderosos y no poderosos, además de la reducción significativa en costos para el acceso. De manera paradójica, ese mismo beneficio es lo que genera una alerta sobre posibles efectos negativos para la ciencia: el imperio de los datos y su robustez pueden generar un desbalance en detrimento de la conceptualización, de modo ello conduciría, incluso, a pensar en el fin de la teoría como tal, que podría llegar a ser reemplazada en su totalidad; este concepto se puede encontrar como un “exceso de inductivismo” que desacredita la creación de ese conocimiento.

Aplicaciones de los macrodatos

En la construcción de la matriz conceptual se incluyó la categoría de “aplicaciones” puesto que fue un elemento común a los textos revisados. Se observó que existe una relación directa entre la materia de estudio del origen del texto y las aplicaciones que se plantean en él. Se pasa a relacionar los hallazgos de la revisión bibliográfica, aunque el tema de aplicaciones de los macrodatos se desarrolla con mayor detalle en el capítulo 3.

A continuación, el gráfico 1 indica, del total de menciones, cuál porcentaje ocurrió de cada tema:

Gráfico 1. Aplicaciones de los macrodatos más mencionadas en la literatura



Fuente: elaboración propia

Como se observa, la aplicación más representativa fue negocios; en este punto es importante aclarar que uno de los filtros utilizados para la búsqueda de artículos fue “negocios”; en esta categoría hay subtemas muy diversos como productividad, optimización de la manufactura y del inventario, manejo de cadenas de suministro, diseño de producto, manejo de clientes, mercadeo, etc.

Después está la categoría gobierno, que gozó de 12% de menciones. Acá entraron todos los temas de diseño urbanístico, transporte, comunicación entre gobierno y ciudadano, etc. Llama la atención que, además de los negocios que generan beneficio económico, se encontró una parte interesada en que se ubique en la esfera de lo público, lo que permite intuir que el futuro de los macrodatos se encuentra asegurado, pues no es solo un de interés particular sino de los gobiernos también.

La medalla de bronce en menciones estuvo compartida en un triple empate entre los conceptos de salud, toma de decisiones y ciencia. Respecto a la salud, abarcó todo el desarrollo de los dispositivos usables que registran datos de salud como ritmo cardíaco, rutinas de ejercicios o ingesta de calorías. Frente a la ciencia, también se ha ilustrado antes cómo la producción de conocimiento se nutre de los macrodatos para construir muestras cada vez más representativas. Por último, lo referente a toma de decisiones fue especial puesto que es transversal a los temas planteados y se encontraron procesos de toma de decisiones en los ámbitos empresariales, gubernamentales, científicos y de salud.

Los demás temas se encuentran en el gráfico, pero no se abordaron en este apartado por ser menos representativos y se dejan a manera de ilustración en la información presentada. En el próximo capítulo se desarrolla de manera más detallada la cuestión del uso de los macrodatos.

Retos y tendencias futuras de los macrodatos

Como se ha venido exponiendo en el presente escrito, el tema de los macrodatos es emergente y de vanguardia; por lo tanto, existe mucho campo de acción para la investigación científica y quedan todavía muchos problemas que abordar. Como se ha venido proponiendo, el tema se puede enfocar desde dos perspectivas: la parte tecnológica del asunto y la administrativa, que fue en la que se hizo énfasis en la revisión. Los problemas futuros y las tendencias son de estas dos naturalezas y la mayoría de las veces mixtos, porque un avance se apoya en otro y así se va enriqueciendo el conocimiento. Según lo rastreado en la literatura de los textos más reconocidos e influyentes acerca de los macrodatos, se proponen las siguientes categorías para discutir las tendencias futuras de este tema: manejo organizacional, segmentación y manejo de clientes, educación de las organizaciones para la cultura de la gestión de la información, trabajo interdisciplinario necesario para el abordaje de los macrodatos y retos tecnológicos por el crecimiento exponencial de la captura de datos.

Tipos de organizaciones

Según Kambatla, Kollias, Kumar y Grama (2014), los siguientes tipos de organizaciones pueden beneficiarse del uso de los macrodatos:

- En los sistemas complejos que no se prestan a modelos intuitivos, el modelado basado en datos y la generación de hipótesis es clave para comprender el comportamiento y las interacciones del sistema. Tales aplicaciones surgen en entornos naturales (las ciencias), en interacciones sociales y en sistemas de ingeniería, entre otros.
- En las empresas comerciales y de negocios, el papel de la analítica de grandes datos es bien reconocido para mejorar la eficiencia y guiar los procesos de decisión.
- En salud y bienestar humano, los análisis de grandes volúmenes de datos ofrecen un tremendo potencial para intervenciones de pronóstico, terapias novedosas y para moldear el estilo de vida y el comportamiento. El análisis de macrodatos también es clave para la eficiencia de costos y la sostenibilidad de la infraestructura de salud.
- En entornos interactivos u orientados hacia los clientes, los análisis de grandes datos guían la interfaz de los sistemas con los consumidores. Los ejemplos de negocios que configuran su entorno operativo para optimizar la experiencia del cliente (y los resultados del negocio) son bien conocidos. Ahora están surgiendo entornos ubicuos, como hogares inteligentes, que se optimizan de manera simultánea para los espacios de vida, así como la huella energética y el costo.
- En los sistemas complejos que se prestan a la configuración y el control, los

análisis permiten la evolución y el diseño controlados (Kambatla et al., 2014, p. ...) ¹².

En la anterior caracterización caben muchos tipos de empresas con diferentes razones sociales y en esencia esta filosofía de la gestión de información para crear valor se puede aplicar a todas las organizaciones que encarguen de solucionar lo que antes se denominaban “problemas complejos”, que son los objetos de estudio de la administración.

Se pudo encontrar en la revisión de las investigaciones de diferentes autores que coincidieron en la utilización de los macrodatos en el manejo de organizaciones: “De manera específica, las investigaciones futuras pueden explorar temas como el liderazgo, la gestión del talento, la tecnología y las herramientas, los ecosistemas de información, la cultura de la empresa, la privacidad de los datos, el valor empresarial y el proceso de toma de decisiones, que tienen un enorme impacto en la implementación de los macrodatos” ¹³ (Fosso Wamba et al., 2015, p. 245).

En esta misma línea de operaciones interiores de una organización, los mismos autores dicen, a propósito del desempeño organizacional: “El futuro desempeño de la organización está relacionado en forma intrínseca con estas orientaciones, lo que puede garantizar una ventaja competitiva y resultados empresariales difíciles de replicar” ¹⁴ (Fosso Wamba et al., 2015, p. 245); de manera clara acá se alude a ventaja competitiva, pues se refiere a operaciones internas que generan valor y que son difíciles de replicar para la competencia.

Las acciones que generen ventaja competitiva pero que sean replicables con facilidad pierden su capacidad de dar ventaja en poco tiempo porque pasan a ser prácticas estándar de la industria y se convierten en denominadores comunes que

¹² Traducción no oficial del autor.

¹³ Traducción no oficial del autor.

¹⁴ Traducción no oficial del autor.

arrojan resultados similares. Sin embargo, con la capacidad de estos sistemas de macrodatos de analizar gran cantidad de variables, de naturalezas diferentes como lo describe la v de variedad, se obtendrán como resultados realizaciones replicables con dificultad por la complejidad de variables que se interrelacionan en los macrodatos.

Según esta perspectiva futura, las organizaciones tendrán que adaptarse para la gestión de cantidad voluminosa de datos en tiempo real para permanecer competitivas. De esta misma manera, las estructuras organizacionales tendrán que ser flexibles y ágiles. Se requiere replantear las concepciones tradicionales de jerarquías y pensar más en grupos de trabajos que se arman y desarman por proyectos con velocidad y agilidad:

Las corporaciones modernas y sus unidades de negocios estratégicas continuarán perdiendo sus estructuras organizacionales jerárquicas. Las empresas buscan crear entidades comerciales que sean más ágiles, más flexibles y que respondan de mejor modo a un entorno empresarial que cambia con rapidez. Con las reducciones en el personal y en la gerencia media, los altos gerentes y los ejecutivos se involucran de manera más directa en la resolución de problemas, la toma de decisiones y la planificación que en los años ochenta. Las organizaciones ágiles y flexibles también les piden a sus gerentes y a su personal que cambien con frecuencia su enfoque. Por lo tanto, las herramientas de apoyo a las decisiones jugarán un papel más central en este entorno cambiante con rapidez¹⁵ (Shim et al., 2002, p. 121).

Segmentación y manejo de clientes

Una de las formas de generar valor consiste en atender la parte de la cadena de valor que se refiere a los clientes. Desde el punto de vista de los macrodatos, toda

¹⁵ Traducción no oficial del autor.

la información masiva que se puede captar de los clientes sirve para mejorar los niveles de servicios y adaptar las promesas de valor a sus expectativas. Fosso Wamba et al. (2015) describen un caso de uso de gestión de información en un servicio y concluyen:

Las lecciones aprendidas de este estudio de caso en profundidad se aplican no solo al servicio de emergencia, sino también a otros sectores, como el cuidado de la salud, así como a empresas (por ejemplo, multinacionales) con una arquitectura empresarial compleja y múltiples fuentes de datos que les permiten adaptar las demandas de los clientes para lograr una ventaja competitiva en el mercado¹⁶ (p. 245).

De nuevo se hace referencia a lo que en el punto anterior se refería como formas de lograr ventajas competitivas a través del uso de macrodatos, por los aprendizajes valiosos y difíciles de imitar que permite el análisis de información provechosa. Un ejemplo de ello se encuentra en Jin et al. (2015):

No es ningún secreto que en la investigación y las aplicaciones de los macrodatos, la industria está por delante de la academia. Por ejemplo, según lo divulgado por Alibaba en marzo de 2014, su centro de datos ha almacenado más de 100 PB de datos procesados, lo que equivale a 100 millones de películas de alta resolución. Durante el pasado "Día de los solteros" (también conocido como "Día de los dobles 11"), Alibaba obtuvo un valor de 9.300 millones de CNY en este evento de compras, que correspondió a unos 278 millones de clientes. Para este evento anual de compras, Alibaba desarrolló una plataforma de procesamiento de datos en tiempo real llamada Galaxy, que puede manejar 5 millones de transacciones por segundo. La cantidad total de datos que Galaxy puede procesar todos los días es de aproximadamente 2 PB. La industria tiene

¹⁶ Traducción no oficial del autor.

más éxito a este respecto porque tiene dos fuerzas motrices esenciales: en realidad necesita gestionar los macrodatos en tiempo real y tiene los requisitos para hacer un mejor uso de los datos recopilados¹⁷ (p. 63).

Esta parte de la productividad aplicada a los negocios, la generación de valor y el manejo de clientes se encontró como uno de puntos en común de los autores tratados; Pang y Lee (2008), Gandomi y Haider (2015), coinciden en nombrar, además de los citados Jin et al. (2016) y Fosso Wamba et al. (2015), la utilidad de los macrodatos en el manejo de clientes. Pang y Lee (2008) hablan de la minería de datos orientada al análisis de sentimientos –u opiniones– es posible hoy en día gracias a la adopción de procedimiento analíticos de macrodatos, viables desde el punto de vista de los costos, tesis que comparten Gandomi y Haider (2015), que exponen que, a partir de la interpretación de sentimientos expresados a través de redes sociales, se pueden hacer realizaciones significativas que antes, en la época previa a los macrodatos, no hubiera sido posible.

En este mismo orden de ideas, Babu (2012) expone que en la actualidad los clientes exigen respuestas más rápidas y efectivas de parte de los negocios que, para permanecer competitivos, necesariamente deben recurrir a técnicas de gestión de datos (el autor se refiere a inteligencia de negocios) como de macrodatos. Una última referencia a este problema empresarial la hicieron Gandomi y Haider (2015) respecto a la identificación y a uno de los usos aplicados de macrodatos: los análisis predictivos y su generación de valor para una organización:

Para el sistema, el desafío esencial es que un marco de minería de macrodatos debe considerar relaciones complejas entre muestras, modelos y fuentes de datos, junto con sus cambios en evolución con el tiempo y otros posibles factores. Un sistema debe diseñarse con sumo cuidado para

¹⁷ Traducción no oficial del autor.

que los datos no estructurados puedan vincularse a través de sus complejas relaciones para formar patrones útiles y el crecimiento de los volúmenes de datos y las relaciones de elementos deben ayudar a formar patrones legítimos para predecir la tendencia y el futuro”¹⁸ (Sowmya y Suneetha, 2017, p. 105).

Educación de las organizaciones y su talento humano para la cultura de la gestión de la información

Una de las tareas futuras de los macrodatos es la inserción cultural en las empresas y sus personas en el uso de sistemas de gestión de la información. En la revisión de literatura se encontró que se hacía referencia a esta cuestión, además de los posibles caminos futuros para la generación de valor en las empresas y el desarrollo de tecnologías, que el hecho de tener profesionales dispuestos y entrenados para realizar estas gestiones será la vía posible para la materialización de todos los proyectos referentes a macrodatos. Una crítica que se identificó en la construcción del capítulo 1 en la revisión de la literatura a propósito de los sistemas de gestión de la información fue que ellos han servido para tomar decisiones que generan menos esfuerzo –o conservación del esfuerzo– pero no necesariamente “la mejor decisión” o la que más valor genera. Esto se debe al problema paradigmático de la administración, eficiencia versus eficacia, en la que un ser humano decide por una u otra de acuerdo con la interpretación que haga de su entorno, en este caso mediante el uso de técnicas de macrodatos. En el próximo apartado se revisa una salida encontrada en la bibliografía a la gestión de datos masivos para beneficio de los proyectos que los empleen:

¿Cómo pueden los programas académicos de sí seguir satisfaciendo las necesidades de sus estudiantes tradicionales, al mismo tiempo que llegan

¹⁸ Traducción no oficial del autor.

al profesional de TI que necesita nuevas habilidades analíticas? Es posible que se necesite una nueva visión de IS para abordar esta y otras preguntas”¹⁹ (Ghazal et al., 2013, p. 1185).

Trabajo interdisciplinario necesario para el abordaje de los macrodatos

Un gran reto en el uso de macrodatos es la capacidad para procesar los datos, en lo fundamental, por tres de la tres veces antes descritas –velocidad, volumen y variedad–. Los autores del tema concuerdan en que el escollo principal está en poder responder con análisis en tiempo real que generen conclusiones que produzcan valor. Una de las perspectivas que se plantean para la solución de este reto es el desarrollo de sistemas y programas robustos de análisis que soporten cantidades masivas de datos. Otra perspectiva en la del uso de los macrodatos es el abordaje de manera interdisciplinaria. Una de las conclusiones del artículo de Jin et al. (2015) es que, al abordar un proyecto de macrodatos, debe emplearse un sistema integrado de ingeniería.

La tesis que se plantea es abordar la v de variedad a partir de personas con diferentes conocimientos y, de igual manera, por medio en el impacto de las v de volumen y velocidad para poder procesar la información por partes.

Por ejemplo, se necesitan expertos en sistemas de información para proporcionar información sobre cómo se recopilan, almacenan, procesan y recuperan los datos. Se necesitan expertos para garantizar que los problemas correctos se estén realizando de la manera correcta y que los resultados derivados sean relevantes (Waller y Fawcett, 2013). Además, la investigación de calidad de datos emergentes sugiere la necesidad de

¹⁹ Traducción no oficial del autor.

expertos estadísticos y analíticos que conozcan los métodos necesarios para medir, monitorear y controlar la calidad de los datos (Jones-Farmer et al., 2013). Trabajando juntos, los académicos de estas y otras disciplinas pueden emplear las tecnologías y las técnicas correctas para resolver los problemas correctos²⁰ (Hazen, Boone, Ezell y Jones-Farmer, 2014, p. 78).

Esta estrategia de tener expertos con conocimientos se puede unir a otra planteada por Najafabadi (2016), que planteó el uso de muestras. El autor habla de descomponer la cantidad masiva de datos mediante la identificación de partes de información que puedan ser introducidas en un sistema con un algoritmo de aprendizaje profundo, que pueda extraer patrones de la muestra y aplicarla para el análisis automático del resto de la información. Esto podría extenderse a usar la información estructurada como muestra para después poder, con ese patrón, analizar el resto de la información semiestructurada o no estructurada. Uno de los problemas que se identificaron es la dimensión de la muestra, ¿cuál tamaño debería tener la muestra con respecto al total de datos que se pretende analizar? Esta pregunta abre una ruta posible de investigación en el tema de los macrodatos.

Si se aborda la cuestión de construcción de modelos de análisis a partir de trabajo interdisciplinar, se encontró una clasificación de principios que deben tenerse en cuenta para la construcción de programas de macrodatos, según Philip Chen y Zhang (2014), que plantearon que, para poder trabajar con macrodatos, no solo es necesario desarrollar nuevas tecnologías, sino nuevas formas de pensar:

Principio 1. Las buenas arquitecturas y los marcos son necesarios y están en la máxima prioridad.

Principio 2. Admitir una variedad de métodos analíticos. Las aplicaciones de macrodatos a menudo producen tareas complejas que hacen que sea

²⁰ Traducción no oficial del autor.

imposible resolverlas mediante el uso de una o algunas de las disciplinas y los métodos analíticos.

Principio 3. No se ajusta a todos los tamaños. Cuando se trata de análisis de macrodatos, no hay un tamaño que pueda adaptarse a todas las soluciones.

Principio 4. Llevar el análisis a los datos. Dado que el tamaño del conjunto de macrodatos es extremadamente grande, no es aconsejable ni viable recopilar y mover datos a solo uno o varios centros de análisis.

Principio 5. El procesamiento debe ser distribuible para el cálculo en memoria de trabajo.

Principio 6. El almacenamiento de datos debe ser distribuible para el almacenamiento en memoria.

Principio 7. Se necesita coordinación entre el procesamiento y las unidades de datos²¹ (Philip Chen y Zhang, 2014, p. ...).

Conclusiones generales de la tendencia de los macrodatos

Según Philip Chen y Zhang (2014), las técnicas y las tecnologías avanzadas para desarrollar la ciencia de los macrodatos tiene la finalidad de desarrollar métodos científicos más sofisticados para manejar, analizar, visualizar y explotar conocimientos informativos de organizaciones de información diversas, heterogéneas, voluminosas y descentralizadas. El fin es desarrollar la ciencia de los macrodatos que potencie avances sociales y económicos no posibles en el pasado. En las próximas décadas, la competitividad en sectores diversos estará definida por la utilización de macrodatos y así mismo permitirá descubrimientos novedosos que antes, sin estas tecnologías, no hubieran sido posibles. Un cambio de paradigma en la investigación científica está en camino y el futuro gira alrededor de temas tales como: técnicas matemáticas y estadísticas, nuevas

²¹ Traducción no oficial del autor.

herramientas de minería de datos, algoritmos avanzados de aprendizaje automático, así como otras disciplinas analíticas de datos (Philip Chen y Zhang, 2014).

Otra conclusión sobre el futuro del fenómeno es que la cantidad de datos será cada vez mayor pues el aumento en la cantidad de artilugios capturadores de datos (tanto en tipos como en unidades en el mundo) hace que el volumen de datos sea cada vez mayor; así mismo, como aumentan los canales de recolección de datos, también se incrementa de manera significativa la cantidad de datos disponibles, así como la velocidad con la que se producen (Hashem, Yaqoob, Anuar, Mokhtar, Gani y Ullah Khan, 2015).

Por último, se menciona la gran oportunidad que hay en el desarrollo de tecnologías duras y blandas:

La escala, el alcance y la naturaleza (características de la carga de trabajo) de las aplicaciones analíticas de los macrodatos, en forma individual, brindan interesantes perspectivas sobre el diseño y la arquitectura de los futuros sistemas de *hardware* y *software*²² (Kambatla et al., 2014, p. 2562).

Así mismo, Kambatla et al., (2014) hablan del impacto en el *software*. Los sistemas de software, el almacenamiento y la computación deben atender un espacio de problemas bastante grande como resultado de la escala de datos, la naturaleza de las cargas de trabajo y otros requisitos de la aplicación, como la coherencia, la disponibilidad y la tolerancia de partición. Las grandes escalas de datos exigen sistemas de almacenamiento distribuidos altamente escalables, que puedan acomodar y servir grandes volúmenes de datos, con mecanismos eficientes de ingreso y egreso.

Consideraciones finales

²² Traducción no oficial del autor.

Los macrodatos son la respuesta al fenómeno de proliferación exponencial de datos en la actualidad, es decir, la cantidad masiva de datos es una realidad *per se* y los macrodatos son la propuesta de gestión de esos datos que ya han sido descritos como voluminosos, variados, valiosos y veloces. Esa propuesta de gestión se centra en la premisa de poder organizarlos de manera oportuna en información valiosa para las partes interesadas: «Los datos que no están procesados, a menudo se describen como "hechos no organizados" (por ejemplo, Faucher et al., 2008: p. 55), mientras que la información se interpreta y se presenta para informar en un contexto dado»²³ (Whyte, Stasis y Lindkvist, 2016, p. 2).

Así como las técnicas de procesamiento y la propuesta de usos de los macrodatos son una respuesta a la aparición del fenómeno, también lo son otro tipo de consideraciones, como las implicaciones éticas y epistemológicas, que fueron plasmadas en el cuerpo del presente texto.

Por último, como consideración personal del autor, se evidenció que el término macro tiene, tanto un atractivo en el sentido que promete otorgar conclusiones más valiosas por lo tamaños de las muestras, como un significado implícito del que no se puede separar (casi una condición *sine qua non*), que consiste en que trata de manera exclusiva de datos masivos, esto es, si unos datos cumplieran las demás condiciones: velocidad, variedad y valor, pero no con la condición de volumen, no podría hablarse, en el sentido riguroso, de macrodatos:

postura crítica que es esencial, pero a menudo se pierde en medio del clamor en los negocios, la política e incluso en algunos rincones académicos para las "soluciones de Macrodatos" (Crawford et al. 2014). Tal crítica significa preocuparse menos por la "grandeza" de los datos y

²³ Traducción no oficial del autor.

prestar más atención a cómo los datos se ven como "grandes" en relevancia social y valencia normativa²⁴ (Lewis y Westlund, 2014, p. 3).

Esta idea plantea que el prefijo macro no necesariamente se refiere de manera exclusiva al tamaño de los datos sino a la relevancia de ellos; en términos generales, que los datos tengan una importancia macro en un proceso de toma de decisiones hasta convertirse en una cuestión cultural de las partes interesadas. Esta postura ayuda a dar otra perspectiva del tema. La discusión sobre el tamaño de los datos para considerarlos macrodatos se puede encontrar, sobre todo, en los artículos de tecnología de la información, en los que, en lo primordial, se plantean dos argumentos: que el volumen de los datos exceda cierto tamaño, medido en terabytes o petabytes, porque el tamaños difiere entre autores, o que sean datos lo suficientemente voluminosos como para que los sistemas tradicionales de procesamiento de información sean insuficientes y se deban desarrollar nuevos.

Tal vez si se quiere restar peso a la condición de volumen de los datos existan nombres más precisos para el tema, como, por ejemplo, ciencias de datos o ciencias de la decisión, que, como se expuso al principio, venían siendo usadas desde que surgieron los sistemas de soporte de decisión o la inteligencia de negocios, esto es, la posibilidad de hacer toda la gestión de transformación de datos crudos en información valiosa para la toma de decisiones (con inclusión de modelos predictivos), con el fin de atenuar el énfasis protagónico del tamaño de los datos.

²⁴ Traducción no oficial del autor.

Referencias

- Andrejevic, M. (2014). Big data, big questions| The big data divide. *International Journal of Communication*, 8(1), 1673-1689. Recuperado de <https://ijoc.org/index.php/ijoc/article/view/2161>
- Arnott, D., & Pervan, G. (2016). A critical analysis of decision support systems research. *Formulating Research Methods for Information Systems: Volume 2*, 127-168. doi: 10.1057/9781137509888_5
- Babu, K. V. S. . J. (2012). Business Intelligence: Concepts, Components, Techniques and Benefits. SSRN Electronic Journal. doi:10.2139/ssrn.2150581
- Braganza, A., Brooks, L., Nepelski, D., Ali, M., & Moro, R. (2017). Resource management in big data initiatives: processes and dynamic capabilities. *Journal of Business Research*, 70, 328-337. doi: 10.1016/j.jbusres.2016.08.006
- Chen, H., Chiang, R. H. L., & Storey, V. C. (2012). Business intelligence and analytics: From big data to big impact. *MIS Quarterly: Management Information Systems*, 36(4), 1165-1188.
- Chen, D. Q., Preston, D. S., & Swink, M. (2015). How the use of big data analytics affects value creation in supply chain management. *Journal of Management Information Systems*, 32(4), 4-39. doi: 10.1080/07421222.2015.1138364
- Congreso de la República (2012). *Ley estatutaria 1581, de 17 de octubre de 2012, por la cual se dictan disposiciones generales para la protección de datos personales*. Bogotá: Congreso de la República. Recuperado de http://www.secretariassenado.gov.co/senado/basedoc/ley_1581_2012.html

- Constantiou, I. D., & Kallinikos, J. (2015). New games, new rules: big data and the changing context of strategy. *Journal of Information Technology*, 30(1), 44-57. doi: 10.1057/jit.2014.17
- Daniel, B. (2015). Big data and analytics in higher education: opportunities and challenges. *British Journal of Educational Technology*, 46(5), 904-920. doi: 10.1111/bjet.12230
- Erevelles, S., Fukawa, N., & Swayne, L. (2016). Big data consumer analytics and the transformation of marketing. *Journal of Business Research*, 69(2), 897-904. doi: 10.1016/j.jbusres.2015.07.001
- Fosso Wamba, S., Akter, S., Edwards, A., Chopin, G., & Gnanzou, D. (2015). How “big data” can make big impact: findings from a systematic review and a longitudinal case study. *International Journal of Production Economics*, 165, 234-246. doi: 10.1016/j.ijpe.2014.12.031
- Furht, B., Villanustre, F., Najafabadi, M. M., Khoshgoftaar, T. M., Seliya, N., & Wald, R. (2016). Deep learning techniques in big data analytics. *Big Data Technologies and Applications*, 133-156. doi: 10.1007/978-3-319-44550-2_5
- Gandomi, A., & Haider, M. (2015). Beyond the hype: big data concepts, methods, and analytics. *International Journal of Information Management*, 35(2), 137-144. doi: 10.1016/j.ijinfomgt.2014.10.007
- Ghazal, A., Rabl, T., Hu, M., Raab, F., Poess, M., Crolotte, A., & Jacobsen, H.-A. (2013). BigBench: towards an industry standard benchmark for big data analytics. *Proceedings of the 2013 International Conference on Management of Data-SIGMOD'13*, 36(4), 1197-1208. doi: 10.1145/2463676.2463712
- Günther, W. A., Rezazade Mehrizi, M. H., Huysman, M., & Feldberg, F. (2017). Debating big data: a literature review on realizing value from big data. *Journal of Strategic Information Systems*, 26(3), 191-209. doi: 10.1016/j.jsis.2017.07.003

- Hargittai, E. (2015). Is bigger always better? Potential biases of big data derived from social network sites. *Annals of the American Academy of Political and Social Science*, 659(1), 63-76. doi: 10.1177/0002716215570866
- Hashem, I. A. T., Yaqoob, I., Anuar, N. B., Mokhtar, S., Gani, A., & Ullah Khan, S. (2015). The rise of “big data” on cloud computing: review and open research issues. *Information Systems*, 47, 98-115. doi: 10.1016/j.is.2014.07.006
- Hazen, B. T., Boone, C. A., Ezell, J. D., & Jones-Farmer, L. A. (2014). Data quality for data science, predictive analytics, and big data in supply chain management: an introduction to the problem and suggestions for research and applications. *International Journal of Production Economics*, 154, 72-80. doi: 10.1016/j.ijpe.2014.04.018
- Hu, H., Wen, Y., Chua, T.-S., & Li, X. (2014). Toward scalable systems for big data analytics: a technology tutorial. *IEEE Access*, 2, 652-687. doi: 10.1109/access.2014.2332453
- Janssen, M., & van den Hoven, J. (2015). Big and open linked data (BOLD) in government: a challenge to transparency and privacy? *Government Information Quarterly*, 32(4), 363-368. doi: 10.1016/j.giq.2015.11.007
- Jin, X., Wah, B. W., Cheng, X., & Wang, Y. (2015). Significance and challenges of big data research. *Big Data Research*, 2(2), 59-64. doi: 10.1016/j.bdr.2015.01.006
- Kambatla, K., Kollias, G., Kumar, V., & Grama, A. (2014). Trends in big data analytics. *Journal of Parallel and Distributed Computing*, 74(7), 2561-2573. doi: 10.1016/j.jpdc.2014.01.003
- Keen, P. G. W., & M. S. Scott Morton (1978). *Decision support systems: an organizational perspective*. Boston, MA: Addison-Wesley.
- Kelemenis, A. M., & Askounis, D. T. (2008). A coherent framework for the development of a human resource decision support system. *Proceedings of*

the 4th IEEE International Conference on Management of Innovation and Technology, ICMIT, 4(4), 228-233. doi: 10.1109/ICMIT.2008.4654367

Lewis, S. C., & Westlund, O. (2014). Big data and journalism. *Digital Journalism, 3(3), 447-466. doi: 10.1080/21670811.2014.976418*

Miller, H. G., & Mork, P. (2013). From data to decisions: a value chain for big data. *IT Professional, 15(1), 57-59. doi: 10.1109/MITP.2013.11*

Najafabadi, M., Villanustre, F., Khoshgoftaar, T., Seliya, N., Wald, R., and Muharemagic, E., Deep learning applications and challenges in big data analytics. *Springer*. Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0>)

Opresnik, D., & Taisch, M. (2015). The value of big data in servitization. *International Journal of Production Economics, 165, 174-184. doi: 10.1016/j.ijpe.2014.12.036*

Ostrom, A. L., Parasuraman, A., Bowen, D. E., Patrício, L., & Voss, C. A. (2015). Service research priorities in a rapidly changing context. *Journal of Service Research, 18(2), 127-159. doi: 10.1177/1094670515576315*

Pang, B., Lee, L., (2008), "Opinion Mining and Sentiment Analysis", *Foundations and Trends in Information Retrieval: Vol. 2: No. 1–2, pp 1-135.*
<http://dx.doi.org/10.1561/1500000011>

Parimala, K., Rajkumar, G., Ruba, A., & Vijayalakshmi, S. (2019). Challenges and opportunities with big data. *International Journal of Scientific Research in Computer Science and Engineering, 5(5), 16-20. doi: 10.26438/ijsrcse/v5i5.1620*

Philip Chen, C. L., & Zhang, C. Y. (2014). Data-intensive applications, challenges, techniques and technologies: a survey on big data. *Information Sciences, 275, 314-347. doi: 10.1016/j.ins.2014.01.015*

- Power, D.J. (2014). Using “Big Data” for analytics and decision support. *Journal of Decision Systems*.
- Puschmann, C., & Burgess, J. (2014). Big data, big questions| Metaphors of big data. *International Journal of Communication*, 8, 1690-1709. Recuperado de <https://ijoc.org/index.php/ijoc/article/view/2169/1162>
- Rust, R. T., & Huang, M.-H. (2014). The service revolution and the transformation of marketing science. *Marketing Science*, 33(2), 206-221. doi: 10.1287/mksc.2013.0836
- Salkind, N. J. (2012). *Exploring research*, 8^a ed. Boston, MA: Pearson.
Recuperado de http://dinus.ac.id/repository/docs/ajar/Neil_J._Salkind_2012_-_Exploring_Research_.pdf
- Sharon, T., & Zandbergen, D. (2017). From data fetishism to quantifying selves: self-tracking practices and the other values of data. *New Media and Society*, 19(11), 1695-1709. doi: 10.1177/1461444816636090
- Shim, J. P., Warkentin, M., Courtney, J. F., Power, D. J., Sharda, R., & Carlsson, C. (2002). Past, present, and future of decision support technology. *Decision Support Systems*, 33(2), 111-126. doi: 10.1016/S0167-9236(01)00139-7
- Sivarajah, U., Kamal, M. M., Irani, Z., & Weerakkody, V. (2017). Critical analysis of big data challenges and analytical methods. *Journal of Business Research*, 70, 263-286. doi: 10.1016/j.jbusres.2016.08.001
- Sowmya, R., & Suneetha, K. R. (2017). Data mining with big data. *Proceedings of 2017 11th International Conference on Intelligent Systems and Control, ISCO 2017*, 26(1), 246-250. doi: 10.1109/ISCO.2017.7855990
- Waller, M. A., & Fawcett, S. E. (2013). Data science, predictive analytics, and big data: a revolution that will transform supply chain design and management. *Journal of Business Logistics*, 34(2), 77–84.

Wang, H., Xu, Z., Fujita, H., & Liu, S. (2016). Towards felicitous decision making: an overview on challenges and trends of big data. *Information Sciences*, 367-368, 747-765. doi: 10.1016/j.ins.2016.07.007

Whyte, J., Stasis, A., & Lindkvist, C. (2016). Managing change in the delivery of complex projects: configuration management, asset information and “big data.” *International Journal of Project Management*, 34(2), 339-351. doi: 10.1016/j.ijproman.2015.02.006

Zuboff, S. (2015). Big other: surveillance capitalism and the prospects of an information civilization. *Journal of Information Technology*, 30(1), 75-89. doi: 10.1057/jit.2015.5

6. Anexo

Matriz conceptual de macrodatos									
Nº	Título del artículo	Autor(es)	Año	<i>Abstract</i>	Definición de macrodatos	Características de la información	Cadena de valor de la información	Problemas y retos de los macrodatos	Aplicaciones

1	Data Mining with Big Data	Chen, H., Chiang, R.H.L., & Storey, V.C.	2012	<p><i>Business intelligence and analytics (BI&A) has emerged as an important area of study for both practitioners and researchers, reflecting the magnitude and impact of data-related problems to be solved in contemporary business organizations. This introduction to the MIS Quarterly Special Issue on Business Intelligence Research first provides a framework that identifies the evolution, applications, and emerging research areas of BI&A. BI&A 1.0, BI&A 2.0, and BI&A 3.0 are defined and described in terms of their key characteristics and capabilities. Current research in BI&A is analyzed and challenges and opportunities associated with BI&A research and education are identified. We also report a bibliometric study of critical BI&A publications, researchers, and research topics based on more than a decade of related academic and industry publications. Finally, the six articles that comprise this special issue are introduced and characterized in terms of the proposed BI&A research framework</i></p>	La recolección de información y su uso permitirán diferentes aplicaciones en la industria que generen valor	Volumen, variedad, velocidad	Recolección, extracción, análisis	Velocidad para producir resultado como fruto del análisis de información que pueda usarse con éxito en las empresas. Capacidad para procesar grandes volúmenes de datos. Educación de las futuras generaciones en técnicas de gestión de la información	Comercio electrónico, inteligencia de mercado, gobierno, salud, seguridad
---	---------------------------	--	------	---	---	------------------------------	-----------------------------------	---	---

2	Beyond the hype: Big data concepts, methods, and analytics	Wu, X., Zhu, X., Wu, G.- Q., Ding, W.	2014	<p><i>Big Data concern large-volume, complex, growing data sets with multiple, autonomous sources. With the fast development of networking, data storage, and the data collection capacity, Big Data are now rapidly expanding in all science and engineering domains, including physical, biological and biomedical sciences. This paper presents a HACE theorem that characterizes the features of the Big Data revolution, and proposes a Big Data processing model, from the data mining perspective. This data-driven model involves demand-driven aggregation of information sources, mining and analysis, user interest modeling, and security and privacy considerations. We analyze the challenging issues in the data-driven model and also in the Big Data revolution</i></p>	Grandes volúmenes de datos representados por "dimensionalidades" heterogéneas y diversas (sic)	Volumen, variedad	Recolección, almacenamiento, movilización de los datos, análisis	<p>Exploración de grandes volúmenes de datos y extracción de información útil para futuras acciones. Análisis en tiempo real. Capacidad de almacenamiento de datos. Transporte de la información de un lugar a otro. Barreras de privacidad de la información, características complejas y dinámicas de los datos</p>	Biotecnología, modelos predictivos, toma de decisiones
---	---	---	------	--	--	-------------------	--	---	--

3	Challenges and opportunities with big data	Gandomi, A., Haider, M.	2015	<p><i>Size is the first, and at times, the only dimension that leaps out at the mention of big data. This paper attempts to offer a broader definition of big data that captures its other unique and defining characteristics. The rapid evolution and adoption of big data by industry has leapfrogged the discourse to popular outlets, forcing the academic press to catch up. Academic journals in numerous disciplines, which will benefit from a relevant discussion of big data, have yet to cover the topic. This paper presents a consolidated description of big data by integrating definitions from practitioners and academics. The paper's primary focus is on the analytic methods used for big data. A particular distinguishing feature of this paper is its focus on analytics related to unstructured data, which constitute 95% of big data. This paper highlights the need to develop appropriate and efficient analytical methods to leverage massive volumes of heterogeneous data in unstructured text, audio, and video formats. This paper also reinforces the need to devise new tools for predictive analytics for structured big data. The statistical methods in practice were devised to infer from sample data. The heterogeneity, noise, and the massive size of structured big data calls for developing computationally efficient algorithms that may avoid big data pitfalls, such as spurious correlation</i></p>	Cantidad masiva de datos complejos y variables que requieren técnicas para ser gestionados con el fin de obtener productos que generen aprendizajes valiosos	Volumen, velocidad, veracidad (los macrodatos deben trabajar con datos inciertos, como los sentimientos variables expresados a través de una red social), variedad (complejidad), valor	Captura, almacenamiento, distribución, manejo y análisis de la información	Análisis de información no estructurada. Correlaciones espurias entre datos. Análisis de información en tiempo real	Análisis de texto, audio, video, redes sociales y de tipo predictivo
---	--	-------------------------	------	---	--	---	--	---	--

4	<p><i>How 'big data' can make big impact: Findings from a systematic review and a longitudinal case study</i></p>	<p>Labrinidis, A., Jagdish, H.V.</p>	<p>2012</p>	<p><i>The promise of data-driven decision-making is now being recognized broadly, and there is growing enthusiasm for the notion of "Big Data," including the recent announcement from the White House about new funding initiatives across different agencies, that target research for Big Data. While the promise of Big Data is real - for example, it is estimated that Google alone contributed 54 billion dollars to the US economy in 2009 - there is no clear consensus on what is Big Data. In fact, there have been many controversial statements about Big Data, such as "Size is the only thing that matters." In this panel we will try to explore the controversies and debunk the myths surrounding Big Data</i></p>	<p>Cantidad masiva de datos y sistemas que sirven para gestionar (almacenar, recolectar y analizar) dichos datos</p>	<p>Voluminosa, rápida, heterogénea</p>	<p>Recolección, almacenamiento, análisis, uso</p>	<p>Heterogeneidad de los datos, incompleta, escala, rapidez, privacidad y complejidad</p>	<p>Diseño de filtros en la etapa de adquisición de datos</p>
---	---	--------------------------------------	-------------	--	--	--	---	---	--

5	Business intelligence from big data to big impact	Fosso Wamba, S., Akter, S., Edwards, A., Chopin, G., Gnanzou, D.	2015	<p><i>Big data has the potential to revolutionize the art of management. Despite the high operational and strategic impacts, there is a paucity of empirical research to assess the business value of big data. Drawing on a systematic review and case study findings, this paper presents an interpretive framework that analyzes the definitional perspectives and the applications of big data. The paper also provides a general taxonomy that helps broaden the understanding of big data and its role in capturing business value. The synthesis of the diverse concepts within the literature on big data provides deeper insights into achieving value through big data strategy and implementation</i></p>	<p>Datos voluminosos, heterogéneos, recolectados con gran velocidad y cuya gestión produce valor en las organizaciones</p>	<p>Volumen, velocidad, variedad, valor y veracidad</p>	<p>Almacenamiento, análisis</p>	<p>Políticas de datos, tecnología y técnicas, cambio organizacional y talento, acceso a los datos, estructuras industriales</p>	<p>Creación de valor, experimentación, exposición a la variabilidad y mejoramiento del desempeño; segmentación de poblaciones para personalizar soluciones, toma de decisiones, innovación en nuevos modelos de negocio, productos y servicios. Usos en industrias tecnológica, de servicios, salud, gobierno, comercio minorista, educación, ecología y manufactura</p>
---	---	--	------	--	--	--	---------------------------------	---	--

6	Social network analytics – Challenges in topic discovery, data collection, and data preparation	Stieglitz, S., Mirbabaie, M., Ross, B., Neuberger, C.	2018	<p>Since an ever-increasing part of the population makes use of social network in their day-to-day lives, redes social network is being analysed in many different disciplines. The social network analytics process involves four distinct steps, data discovery, collection, preparation, and analysis. While there is a great deal of literature on the challenges and difficulties involving specific data analysis methods, there hardly exists research on the stages of data discovery, collection, and preparation. To address this gap, we conducted an extended and structured literature analysis through which we identified challenges addressed and solutions proposed. The literature search revealed that the volume of data was most often cited as a challenge by researchers. In contrast, other categories have received less attention. Based on the results of the literature search, we discuss the most important challenges for researchers and present potential solutions. The findings are used to extend an existing framework on social network analytics. The article provides benefits for researchers and practitioners who wish to collect and analyse social network data</p>	Sistema de información voluminosa	Volumen	Adquisición, minería, análisis	Volumen	Investigación a través de volúmenes amplios de datos recolectados a través de canales de redes sociales
---	---	--	------	--	-----------------------------------	---------	--------------------------------	---------	---

7	Big data emerging technology: Insights into innovative environment for online learning resources	Huda, M., Maseleno, A., Atmotiyoso, P., Siregar, M., Ahmad, R., Jasmi, K.A., Muhamad, N.H.N., Mustari, M.I., Basiron, B.	2018	<p><i>Attempts to adopt the network data massively from the social network refer to provide the particular means in extracting the value from information space such as message, conversation, transaction and others, where the sources of structured data come from enterprise resources data and sources of unstructured data come from audio and video. It can be achieved to expand the process of extracting the value from social network to pattern the data sources to fulfil the organization goal. This paper aims to reveal the way of big data approach in extracting data value from data complexity involving variety and velocity into the volume. This study was conducted using contents analysis by reviewing some literatures in peer-reviewed journals, chapters, books and proceedings in developing prototype using data analytics associated from the topic, users and time analytics. The findings reveal that big data emerging technology with analytic process provides particular advantages to transform the pattern of information fitted into the innovative environment of online learning resources (OLR) to enhance in developing the learning resources. Both prototype and model of data extraction value could be enhanced to facilitate the learning environment in supporting implementations with ease and convenience. This study is expected to contribute to improve the learning environment and outcomes with performance and achievement by enhancing students' learning process development to provide online resources in higher education context</i></p>	Macrodatos implica modelos de análisis para extraer valor con el fin de dar estructura a una cantidad masiva de datos no estructurados generados con altas velocidades.	Voluminosa, generada con gran velocidad, compleja y no estructurada	Organización, minería de datos, análisis de datos	Gan cantidad de datos no estructurados generados a gran velocidad; el reto es organizarlos (darles estructura) de manera que se pueda extraer valor de ellos	Educación
---	--	--	------	---	---	---	---	--	-----------

8	From data fetishism to quantifying selves: Self-tracking practices and the other values of data	Sharon, T., Zandbergen, D.	2017	<p><i>This article foregrounds the ways in which members of the Quantified Self ascribe value and meaning to the data they generate in self-tracking practices. We argue that the widespread idea that what draws self-trackers to numerical data is its perceived power of truth and objectivity—a so-called “data fetishism”—is limiting. Using an ethnographic approach, we describe three ways in which self-trackers attribute meaning to their data-gathering practices which escape this data fetishist critique: self-tracking as a practice of mindfulness, as a means of resistance against social norms, and as a communicative and narrative aid. In light of this active engagement with data, we suggest that it makes more sense to view these practitioners as “quantifying selves.” We also suggest that such fine-grained accounts of the appeal that data can have, beyond its allure of objectivity, are necessary if we are to achieve a fuller understanding of Big Data culture</i></p>	Información cuantitativa generada que puede usarse para generarles valor a los individuos	Información rápida generada a partir de mediciones que permiten hacer evaluaciones que generen valor	Adquisición, análisis, uso	Crítica de la obsesión por la medición y la información	Autocrecimiento
---	---	-------------------------------	------	--	---	--	----------------------------	---	-----------------

9	<p><i>Debating big data: A literature review on realizing value from big data</i></p>	<p>Günther, W.A., Rezazade Mehrizi, M.H., Huysman, M., Feldberg, F.</p>	<p>2017</p>	<p><i>Big data has been considered to be a breakthrough technological development over recent years. Notwithstanding, we have as yet limited understanding of how organizations translate its potential into actual social and economic value. We conduct an in-depth systematic review of IS literature on the topic and identify six debates central to how organizations realize value from big data, at different levels of analysis. Based on this review, we identify two socio-technical features of big data that influence value realization: portability and interconnectivity. We argue that, in practice, organizations need to continuously realign work practices, organizational models, and stakeholder interests in order to reap the benefits from big data. We synthesize the findings by means of an integrated model</i></p>	<p>Problema moderno considerado un gran avance tecnológico pero en el que aún no está muy claro cómo genera valor</p>	<p>Información son datos recolectado que representan potencialmente algún valor, debe ser organizada y analizada para identificar una forma de generar valor social y económico.</p>	<p>Recolección, análisis, aplicación</p>	<p>No está claro todavía como se puede generar valor social y económico a partir de la recolección de datos masivos</p>	<p>Uso empresariales y sociales</p>
---	---	---	-------------	---	---	--	--	---	-------------------------------------

10	Management challenges in creating value from business analytics	Vidgen, R., Shaw, S., Grant, D.B.	2017	<p><i>The popularity of big data and business analytics has increased tremendously in the last decade and a key challenge for organizations is in understanding how to leverage them to create business value. However, while the literature acknowledges the importance of these topics little work has addressed them from the organization's point of view. This paper investigates the challenges faced by organizational managers seeking to become more data and information-driven in order to create value. Empirical research comprised a mixed methods approach using (1) a Delphi study with practitioners through various forums and (2) interviews with business analytics managers in three case organizations. The case studies reinforced the Delphi findings and highlighted several challenge focal areas: organizations need a clear data and analytics strategy, the right people to effect a data-driven cultural change, and to consider data and information ethics when using data for competitive advantage. Further, becoming data-driven is not merely a technical issue and demands that organizations firstly organize their business analytics departments to comprise business analysts, data scientists, and IT personnel, and secondly align that business analytics capability with their business strategy in order to tackle the analytics challenge in a systemic and joined-up way. As a result, this paper presents a business analytics ecosystem for organizations that contributes to the body of scholarly knowledge by identifying key business areas and functions to address to achieve this transformation</i></p>	Los macrodatos son la generación, la organización y los usos de cantidades masivas de información	Masiva	Generación, almacenamiento, análisis	Cómo las empresas pueden generar valor a partir de la gestión de la información	Usos industriales y empresariales
----	---	-----------------------------------	------	---	---	--------	--------------------------------------	---	-----------------------------------

11	A Big Data Analytics Method for Tourist Behaviour Analysis	Miah, S.J., Vu, H.Q., Gammack, J., McGrath, M.	2017	<p><i>Big data generated across social network sites have created numerous opportunities for bringing more insights to decision-makers. Few studies on big data analytics, however, have demonstrated the support for strategic decision-making. Moreover, a formal method for analysing social network-generated big data for decision support is yet to be developed, particularly in the tourism sector. Using a design science research approach, this study aims to design and evaluate a 'big data analytics' method to support strategic decision-making in tourism destination management. Using geotagged photos uploaded by tourists to the photo-sharing social network site, Flickr, the applicability of the method in assisting destination management organisations to analyse and predict tourist behavioural patterns at specific destinations is shown, using Melbourne, Australia, as a representative case. Utility was confirmed using both another destination and directly with stakeholder audiences. The developed artefact demonstrates a method for analysing unstructured big data to enhance strategic decision making within a real problem domain. The proposed method is generic, and its applicability to other big data streams is discussed</i></p>	Datos generados a través de canales de recolección que tienen el potencial de generar valor	Voluminosa	Recolección, agrupación, análisis, uso	Falta de modelos de análisis que generen valor	Turismo
----	--	--	------	--	---	------------	--	--	---------

12	Big Data Analytics for Physical Internet-based intelligent manufacturing shop floors	Zhong, R.Y., Xu, C., Chen, C., Huang, G.Q.	2017	<p><i>Physical Internet (PI, π) has been widely used for transforming and upgrading the logistics and supply chain management worldwide. This study extends the PI concept into manufacturing shop floors where typical logistics resources are converted into smart manufacturing objects (SMOs) using Internet of Things (IoT) and wireless technologies to create a RFID-enabled intelligent shop floor environment. In such PI-based environment, enormous RFID data could be captured and collected. This study introduces a Big Data Analytics for RFID logistics data by defining different behaviors of SMOs. Several findings are significant. It is observed that task weight is primarily considered in the logistics decision-making in this case. Additionally, the highest residence time occurs in a buffer with the value of 12.17 (unit of time) which is 40.57% of the total delivery time. That implies the high work-in-progress inventory level in this buffer. Key findings and observations are generated into managerial implications, which are useful for various users to make logistics decisions under PI-enabled intelligent shop floors</i></p>	Gran cantidad de datos generados a través de dispositivos que pueden ayudar en la planeación y los manejos de procesos productivos.	Masiva, desorganizada, generada en tiempo real	Generación, transmisión, análisis, aplicación	Información voluminosa desorganizada que debe procesarse para generar valor	Planeación de producción, manejo de inventarios, cadenas de abastecimiento
----	--	--	------	---	---	--	---	---	--

13	Big data: Dimensions, evolution, impacts, and challenges	Lee, I.	2017	<p><i>Big data represents a new technology paradigm for data that are generated at high velocity and high volume, and with high variety. Big data is envisioned as a game changer capable of revolutionizing the way businesses operate in many industries. This article introduces an integrated view of big data, traces the evolution of big data over the past 20 years, and discusses data analytics essential for processing various structured and unstructured data. This article illustrates the application of data analytics using merchant review data. The impacts of big data on key business performances are then evaluated. Finally, six technical and managerial challenges are discussed</i></p>	Datos variados y voluminosos generados con gran velocidad.	Voluminosa, variada, veloz	Generación, almacenamiento, análisis, uso	Análisis de datos estructurados y no estructurados para generar valor	Empresarial, industrial, gerencial
----	---	---------	------	---	--	----------------------------	---	---	------------------------------------

14	Toward sustainability: using big data to explore the decisive attributes of supply chain risks and uncertainties	Wu, K.-J., Liao, C.-J., Tseng, M.-L., Lim, M.K., Hu, J., Tan, K.	2017	<p>Rapid market changes aimed at sustainability have led to supply chain risks and uncertainties in the Taiwanese light-emitting diode industry. These risks and uncertainties can be captured by social network, quantitative and qualitative data (referred to herein as big data), but the industry has been unable to manage this information boom to respond to customer needs. These various types of data have their own characteristics that affect decision making about developing firm capabilities. This study aggregates the various data to undertake an extensive investigation of supply chain risks and uncertainties. Specifically, this study proposes using the fuzzy and grey Delphi methods to identify a set of reliable attributes and, based on these attributes, transforming big data to a manageable scale to consider their impacts. Subsequently, both the fuzzy and grey Decision Making Trial and Evaluation Laboratories applied to determine the causal relationships for supply chain risks and uncertainties. The results reveal that capacity and operations have greater influence than other supply chain attributes and that risks stemming from triggering events are difficult to diagnose and control. The implications, conclusions and findings are addressed</p>	Información cuantitativa y cualitativa que puede ayudar a controlar riesgo	Voluminosa, valiosa	Generación, adquisición, uso predictivo	Extracción de valor de los datos generados y recolectados	Riesgo, incertidumbre, cadena de abastecimiento
----	--	--	------	--	--	---------------------	---	---	---

15	An optimization model for green supply chain management by using a big data analytic approach	Zhao, R., Liu, Y., Zhang, N., Huang, T.	2017	<p><i>This paper presents a multi-objective optimization model for a green supply chain management scheme that minimizes the inherent risk occurred by hazardous materials, associated carbon emission and economic cost. The model related parameters are capitalized on a big data analysis. Three scenarios are proposed to improve green supply chain management. The first scenario divides optimization into three options: the first involves minimizing risk and then dealing with carbon emissions (and thus economic cost); the second minimizes both risk and carbon emissions first, with the ultimate goal of minimizing overall cost; and the third option attempts to minimize risk, carbon emissions, and economic cost simultaneously. This paper provides a case study to verify the optimization model. Finally, the limitations of this research and approach are discussed to lay a foundation for further improvement</i></p>	Conjunto de datos que, al ser analizados, pueden generar información de valor	Volumen, valor	Recolección, análisis, uso de valor	Uso del valor obtenido de los datos	Planeación de la producción para reducir huella de carbono y costos
----	---	---	------	---	---	----------------	-------------------------------------	-------------------------------------	---

16	<p><i>Environmental efficiency analysis of China's regional industry: a data envelopment analysis (DEA) based approach</i></p>	<p>Chen, L., Jia, G.</p>	<p>2017</p>	<p><i>China's economic has got remarkable achievements during the past 30 years, which has brought tremendous pressure on natural resources and serious environmental pollution at the same time. In this paper, we aim to make environmental efficiency analysis of China's regional industry. Data envelopment analysis (DEA) method is incorporated and the slacks-based measure (SBM) model considering undesirable outputs is introduced to measure the environmental efficiency of different regions. Big data theory is introduced in the collection and selection of the input and output data of the regions. Further, we evaluate the environmental efficiencies of China's industry using data from 2008 to 2012. The results show that apart from several developed provinces, the environmental efficiencies of China's industry are generally low. And the environmental efficiencies of the regions did not show any increasing trend through the past 5 years. Additionally, lager differences exist in environmental efficiencies between the regions in China. Finally, we suggest the Chinese government to focus on the low environmental efficiencies and the unbalanced development of its regional industry</i></p>	<p>Adquisición, manejo, análisis y almacenamiento de datos que, por su naturaleza masiva, no puede hacerse con sistema convencionales</p>	<p>Volumen, velocidad, variedad, valor</p>	<p>Adquisición, almacenamiento, manejo, análisis</p>	<p>Extracción de conclusiones de valor de la gestión de los macrodatos</p>	<p>Afectación ambiental por producción</p>
----	--	--------------------------	-------------	--	---	--	--	--	--

17	Big data and predictive analytics for supply chain and organizational performance	Gunasekaran, A., Papadopoulos, T., Dubey, R., Wamba, S.F., Childe, S.J., Hazen, B., Akter, S.	2017	<p><i>Scholars acknowledge the importance of big data and predictive analytics (BDPA) in achieving business value and firm performance. However, the impact of BDPA assimilation on supply chain (SCP) and organizational performance (OP) has not been thoroughly investigated. To address this gap, this paper draws on resource-based view. It conceptualizes assimilation as a three stage process (acceptance, routinization, and assimilation) and identifies the influence of resources (connectivity and information sharing) under the mediation effect of top management commitment on big data assimilation (capability), SCP and OP. The findings suggest that connectivity and information sharing under the mediation effect of top management commitment are positively related to BDPA acceptance, which is positively related to BDPA assimilation under the mediation effect of BDPA routinization, and positively related to SCP and OP. Limitations and future research directions are provided</i></p>	Técnicas destinadas a manejar gran cantidad de datos	Volumen, velocidad, variedad	Recolección, análisis, uso	Utilización de la información	Empresarial, industrial, gerencial, análisis predictivo, cadena de abastecimiento, planeación
----	---	---	------	---	--	------------------------------	----------------------------	-------------------------------	---

18	Critical analysis of Big Data challenges and analytical methods	Sivarajah, U., Kamal, M.M., Irani, Z., Weerakkody, V.	2017	<p><i>Big Data (BD), with their potential to ascertain valued insights for enhanced decision-making process, have recently attracted substantial interest from both academics and practitioners. Big Data Analytics (BDA) is increasingly becoming a trending practice that many organizations are adopting with the purpose of constructing valuable information from BD. The analytics process, including the deployment and use of BDA tools, is seen by organizations as a tool to improve operational efficiency though it has strategic potential, drive new revenue streams and gain competitive advantages over business rivals. However, there are different types of analytic applications to consider. Therefore, prior to hasty use and buying costly BD tools, there is a need for organizations to first understand the BDA landscape. Given the significant nature of the BD and BDA, this paper presents a state-of-the-art review that presents a holistic view of the BD challenges and BDA methods theorized/proposed/employed by organizations to help others understand this landscape with the objective of making robust investment decisions. In doing so, systematically analysing and synthesizing the extant research published on BD and BDA area. More specifically, the authors seek to answer the following two principal questions: Q1 – What are the different types of BD challenges theorized/proposed/confronted by organizations? and Q2 – What are the different types of BDA methods theorized/proposed/employed to overcome BD challenges?. This systematic literature review (SLR) is carried out through observing and understanding the past trends and extant patterns/themes in the BDA</i></p>	Inteligencia colectiva generada a partir de compartir, en lo primordial en ambientes tecnológicos	Volumen, velocidad, variedad	Generación, análisis, uso	Análisis de la información para generar valor	Eficiencia, nuevos negocios, ventajas competitivas
----	---	---	------	--	---	------------------------------	---------------------------	---	--

research area, evaluating contributions, summarizing knowledge, thereby identifying limitations, implications and potential further research avenues to support the academic community in exploring research themes/patterns. Thus, to trace the implementation of BD strategies, a profiling method is employed to analyze articles (published in English-speaking peer-reviewed journals between 1996 and 2015) extracted from the Scopus database. The analysis presented in this paper has identified relevant BD research studies that have contributed both conceptually and empirically to the expansion and accrual of intellectual wealth to the BDA in technology and organizational resource management discipline

19	Resource management in big data initiatives: Processes and dynamic capabilities	Braganza, A., Brooks, L., Nepelski, D., Ali, M., Moro, R.	2017	<p><i>Effective management of organizational resources in big data initiatives is of growing importance. Although academic and popular literatures contain many examples of big data initiatives, very few are repeated in the same organization. This suggests either big data delivers benefits once only per organization or senior managers are reluctant to commit resources to big data on a sustained basis. This paper makes three contributions to the Special Issue's theme of enhancing organizational resource management. One is to establish an archetype business process for big data initiatives. The second contribution directs attention to creating a dynamic capability with big data initiatives. The third identifies drawbacks of resource based theory (RBT) and it's underpinning assumptions in the context of big data. The paper discusses lessons learnt and draws out implications for practice and business research. The paper's intellectual and practical contributions are based on an in-depth case study of the European ICT Poles of Excellence (EIPE) big data initiative and evidence from the extant literature</i></p>	Fenómeno global que tiene potencial de generar valor a partir de la gestión de cantidad masiva de datos	Volumen, velocidad, valor	Recolección, análisis, uso	Modelos que se replican solo una vez o que solo sirven para una organización o porque hay desconfianza hacia los macrodatos	Empresarial, industrial, gerencial
----	---	---	------	--	---	---------------------------	----------------------------	---	------------------------------------

20	Exploring adaptive teaching competencies in big data era	Huda, M., Maselena, A., Shahrill, M., Jasmi, K.A., Mustari, I., Basiron, B.	2017	<p><i>In the last decade, the adoption of digital tool to support educational process has been emerged among the universities around the world. In big data era, the demand to utilize it in adaptive teaching should be considered to enable the teaching performance especially in accessing the resources. This paper aims to explore the framework model as a way for teachers in adapting big data to help their teaching performance especially in accessing the resources. The literature review was conducted from peer review journals, books and conferences. The findings reveal that process and management skills should be engaged into adaptive teachings competencies. It included commitment in planning, time management, and technology skills. This study is expected to contribute in strengthening teaching performances in the application guideline in the big data era to support assessing the multi-channels of sources of knowledge to extract new insights of value in exploring the adaptive teaching competencies</i></p>	Plataforma que colecciona, relaciona y analiza grandes volúmenes de datos	Volumen, variedad	Recolección, organización, análisis	Utilización de los modelos de macrodatos en un ámbito académico para el aprendizaje	Académico, aprendizaje
----	--	---	------	--	---	-------------------	-------------------------------------	---	------------------------

21	Factors influencing big data decision-making quality	Janssen, M., van der Voort, H., Wahyudi, A.	2017	<p><i>Organizations are looking for ways to harness the power of big data (BD) to improve their decision making. Despite its significance the effects of BD on decision-making quality has been given scant attention in the literature. In this paper factors influencing decision-making based on BD are identified using a case study. BD is collected from different sources that have various data qualities and are processed by various organizational entities resulting in the creation of a big data chain. The veracity (manipulation, noise), variety (heterogeneity of data) and velocity (constantly changing data sources) amplified by the size of big data calls for relational and contractual governance mechanisms to ensure BD quality and being able to contextualize data. The case study reveals that taking advantage of big data is an evolutionary process in which the gradually understanding of the potential of big data and the routinization of processes plays a crucial role</i></p>	Conjunto de datos voluminosos, variados y veloces difíciles de analizar con los métodos tradicionales	Volumen, velocidad, variedad	Generación, recolección, procesamiento	Uso de macrodatos para toma de decisiones	Procesos de decisión
----	--	---	------	---	---	------------------------------	--	---	----------------------

22	Toward the development of a big data analytics capability	Gupta, M., George, J.F.	2016	<p><i>The era of big data has begun such that organizations in all industries have been heavily investing in big data initiatives. We know from prior studies that investments alone do not generate competitive advantage; instead, firms need to create capabilities that rival firms find hard to match. Drawing on the resource-based theory of the firm and recent work in big data, this study (1) identifies various resources that in combination build a big data analytics (BDA) capability, (2) creates an instrument to measure BDA capability of the firm, and (3) tests the relationship between BDA capability and firm performance. Results empirically validate the proposed theoretical framework of this study and provide evidence that BDA capability leads to superior firm performance</i></p>	Sistemas de gestión de la información basados en grandes volúmenes de datos	Volumen	Recolección, análisis, utilización capaz	Uso capaz de la información que se organiza en los sistemas de gestión de la información	Organizacional, competitividad
----	---	-------------------------	------	---	---	---------	--	--	--------------------------------

23	How to improve firm performance using big data analytics capability and business strategy alignment?	Akter, S., Wamba, S.F., Gunasekaran, A., Dubey, R., Childe, S.J.	2016	<p><i>The recent interest in big data has led many companies to develop big data analytics capability (BDAC) in order to enhance firm performance (FPER). However, BDAC pays off for some companies but not for others. It appears that very few have achieved a big impact through big data. To address this challenge, this study proposes a BDAC model drawing on the resource-based theory (RBT) and the entanglement view of sociomaterialism. The findings show BDAC as a hierarchical model, which consists of three primary dimensions (i.e., management, technology, and talent capability) and 11 subdimensions (i.e., planning, investment, coordination, control, connectivity, compatibility, modularity, technology management knowledge, technical knowledge, business knowledge and relational knowledge). The findings from two Delphi studies and 152 online surveys of business analysts in the U.S. confirm the value of the entanglement conceptualization of the higher-order BDAC model and its impact on FPER. The results also illuminate the significant moderating impact of analytics capability–business strategy alignment on the BDAC–FPER relationship</i></p>	Cantidades masivas de datos que soportan diferentes tipos de toma de decisiones	Voluminosa, variada, veloz, valor	Generación, almacenamiento, análisis, uso de valor	Capacidad de transformar los sistemas de gestión de macrodatos en ventajas competitivas y valor	Organizacional, competitividad, desempeño empresarial
----	--	--	------	--	---	-----------------------------------	--	---	---

24	Big data reduction framework for value creation in sustainable enterprises	Rehman, M.H.U., Chang, V., Batool, A., Wah, T.Y.	2016	<p>Value creation is a major sustainability factor for enterprises, in addition to profit maximization and revenue generation. Modern enterprises collect big data from various inbound and outbound data sources. The inbound data sources handle data generated from the results of business operations, such as manufacturing, supply chain management, marketing, and human resource management, among others. Outbound data sources handle customer-generated data which are acquired directly or indirectly from customers, market analysis, surveys, product reviews, and transactional histories. However, cloud service utilization costs increase because of big data analytics and value creation activities for enterprises and customers. This article presents a novel concept of big data reduction at the customer end in which early data reduction operations are performed to achieve multiple objectives, such as (a) lowering the service utilization cost, (b) enhancing the trust between customers and enterprises, (c) preserving privacy of customers, (d) enabling secure data sharing, and (e) delegating data sharing control to customers. We also propose a framework for early data reduction at customer end and present a business model for end-to-end data reduction in enterprise applications. The article further presents a business model canvas and maps the future application areas with its nine components. Finally, the article discusses the technology adoption challenges for value creation through big data reduction in enterprise applications</p>	Conjunto de datos estructurados, no estructurados y semi estructurados acumulados a partir de fuentes heterogéneas	Voluminosa, variada	Recolección, análisis, uso	Extracción de valor de los análisis de los sistemas de gestión de la información	Empresarial, competitivo, creación de valor, sostenibilidad
----	--	--	------	--	--	---------------------	----------------------------	--	---

25	Marketing analytics for data-rich environments	Wedel, M., Kannan, P.K.	2016	<p><i>The authors provide a critical examination of marketing analytics methods by tracing their historical development, examining their applications to structured and unstructured data generated within or external to a firm, and reviewing their potential to support marketing decisions. The authors identify directions for new analytical research methods, addressing (1) analytics for optimizing marketing-mix spending in a data-rich environment, (2) analytics for personalization, and (3) analytics in the context of customers' privacy and data security. They review the implications for organizations that intend to implement big data analytics. Finally, turning to the future, the authors identify trends that will shape marketing analytics as a discipline as well as marketing analytics education</i></p>	Recolección, organización, análisis y uso de cantidades vastas de información valiosa	Volumen	Recolección, análisis, uso	Capacidad de analizar los datos recolectado para generar valor y ventaja	Aumento de valor, mejoramiento de la experiencia del cliente, incremento en satisfacción y lealtad
----	--	----------------------------	------	---	---	---------	----------------------------	--	--

26	Towards felicitous decision making: An overview on challenges and trends of Big Data	Wang, H., Xu, Z., Fujita, H., Liu, S.	2016	<p><i>The era of Big Data has arrived along with large volume, complex and growing data generated by many distinct sources. Nowadays, nearly every aspect of the modern society is impacted by Big Data, involving medical, health care, business, management and government. It has been receiving growing attention of researches from many disciplines including natural sciences, life sciences, engineering and even art & humanities. It also leads to new research paradigms and ways of thinking on the path of development. Lots of developed and under-developing tools improve our ability to make more felicitous decisions than what we have made ever before. This paper presents an overview on Big Data including four issues, namely: (i) concepts, characteristics and processing paradigms of Big Data; (ii) the state-of-the-art techniques for decision making in Big Data; (iii) felicitous decision making applications of Big Data in social science; and (iv) the current challenges of Big Data as well as possible future directions</i></p>	Complejidad en la captura y el manejo de datos voluminosos	Volumen, velocidad, variedad	Captura, curación, análisis, visualización, toma de decisiones	Transformación de la gestión de los macrodatos en tomas de decisiones valiosas (felices)	Creación de valor a través del mejoramiento de cadena de generación de valor con la integración y la explotación de información. Sector médico, cuidado de la salud, negocios, gerencia y gobierno. Ciencias naturales, ciencias de la vida, ingeniería, arte e investigación
----	--	---------------------------------------	------	---	--	------------------------------	--	--	---

27	An empirical study of the rise of big data in business scholarship	Frizzo-Barker, J., Chow-White, P.A., Mozafari, M., Ha, D.	2016	<p><i>Big data has captured the interests of scholars across many disciplines over the last half a decade. Business scholars have increasingly turned their attention to the impact of this emerging phenomenon. Despite the rise in attention, our understanding of what big data is and what it means for organizations and institutional actors remains uncertain. In this study, we conduct a systematic review on "big data" across business scholarship over the past six years (2009-2014). We analyzed 219 peer-reviewed academic papers from 152 journals from the most comprehensive business literature database. We conducted the systematic review both quantitatively and qualitatively using the data analysis software NVivo10. Our results reveal several key insights about the scholarly investigation of big data, including its top benefits and challenges. Overall, we found that big data remains a fragmented, early-stage domain of research in terms of theoretical grounding, methodological diversity and empirically oriented work. These challenges serve to improve our understanding of the state of big data in contemporary research, and to further prompt scholars and decision-makers to advance future research in the most productive manner</i></p>	Predicciones, conexiones y relaciones entre cantidades vastas de datos	Volumen, velocidad, variedad	Captura, organización, análisis	Transformación de los macrodatos en información útil y valiosa	Negocios, gobiernos, educación, iniciativas sociales. Manejo del cambio, generación de valor
----	--	---	------	--	--	------------------------------	---------------------------------	--	--

28	<p><i>Managing change in the delivery of complex projects: Configuration management, asset information and 'big data'</i></p>	<p>Whyte, J., Stasis, A., Lindkvist, C.</p>	<p>2016</p>	<p><i>As we enter an era of 'big data', asset information is becoming a deliverable of complex projects. Prior research suggests digital technologies enable rapid, flexible forms of project organizing. This research analyses practices of managing change in Airbus, CERN and Crossrail, through desk-based review, interviews, visits and a cross-case workshop. These organizations deliver complex projects, rely on digital technologies to manage large data-sets; and use configuration management, a systems engineering approach with mid-20th century origins, to establish and maintain integrity. In them, configuration management has become more, rather than less, important. Asset information is structured, with change managed through digital systems, using relatively hierarchical, asynchronous and sequential processes. The paper contributes by uncovering limits to flexibility in complex projects where integrity is important. Challenges of managing change are discussed, considering the evolving nature of configuration management; potential use of analytics on complex projects; and implications for research and practice</i></p>	<p>Sistemas de agrupación de información</p>	<p>Voluminosa, valiosa</p>	<p>Recolección, organización, decisión</p>	<p>Organización de los datos en información</p>	<p>Manejo de proyectos, toma de decisiones</p>
----	---	---	-------------	---	--	----------------------------	--	---	--

29	Big Data consumer analytics and the transformation of marketing	Erevelles, S., Fukawa, N., Swayne, L.	2016	<p>Consumer analytics is at the epicenter of a Big Data revolution. Technology helps capture rich and plentiful data on consumer phenomena in real time. Thus, unprecedented volume, velocity, and variety of primary data, Big Data, are available from individual consumers. To better understand the impact of Big Data on various marketing activities, enabling firms to better exploit its benefits, a conceptual framework that builds on resource-based theory is proposed. Three resources-physical, human, and organizational capital-moderate the following: (1) the process of collecting and storing evidence of consumer activity as Big Data, (2) the process of extracting consumer insight from Big Data, and (3) the process of utilizing consumer insight to enhance dynamic/adaptive capabilities. Furthermore, unique resource requirements for firms to benefit from Big Data are discussed</p>	Extracción de información escondida a través de interpretación	Volumen, velocidad, variedad	Recolección, organización, interpretación, valor	Traducción de la recolección de cantidades masivas de datos en ventajas competitivas y valor	Mercadeo
----	---	---------------------------------------	------	---	--	------------------------------	--	--	----------

30	Predicting the performance of online consumer reviews: A sentiment mining approach to big data analytics	Salehan, M., Kim, D.J.	2016	<p><i>Although online consumer reviews (OCRs) have helped consumers to know about the strengths and weaknesses of different products and find the ones that best suit their needs, they introduce a challenge for businesses to analyze them because of their volume, variety, velocity and veracity. This research investigates the predictors of readership and helpfulness of OCR using a sentiment mining approach for big data analytics. Our findings show that reviews with higher levels of positive sentiment in the title receive more readerships. Sentimental reviews with neutral polarity in the text are also perceived to be more helpful. The length and longevity of a review positively influence both its readership and helpfulness. Because the current methods used for sorting OCR may bias both their readership and helpfulness, the approach used in this study can be adopted by online vendors to develop scalable automated systems for sorting and classification of big OCR data which will benefit both vendors and consumers</i></p>	Sistemas de información voluminosa que generan valor cuyo uso es un reto	Volumen, variedad, velocidad y veracidad	Recolección, análisis, uso	Generación de valor a partir de los macrodatos	Negocios, experiencia de cliente, minería de sentimientos, predicción
----	--	------------------------	------	--	--	--	----------------------------	--	---

31	<p><i>Big Data for Development: A Review of Promises and Challenges</i></p>	Hilbert, M.	2016	<p><i>The article uses a conceptual framework to review empirical evidence and some 180 articles related to the opportunities and threats of Big Data Analytics for international development. The advent of Big Data delivers a cost-effective prospect for improved decision-making in critical development areas such as healthcare, economic productivity and security. At the same time, the well-known caveats of the Big Data debate, such as privacy concerns and human resource scarcity, are aggravated in developing countries by long-standing structural shortages in the areas of infrastructure, economic resources and institutions. The result is a new kind of digital divide: a divide in the use of data-based knowledge to inform intelligent decision-making. The article systematically reviews several available policy options in terms of fostering opportunities and minimising risks</i></p>	<p>Cantidad masiva de datos generados a través de canales que puede ser organizada para extraer información valiosa</p>	<p>Velocidad, volumen, variedad</p>	<p>Recolección, almacenamiento</p>	<p>Incapacidad de analizar la cantidad masiva de datos</p>	<p>Desarrollo, toma de decisiones, productividad, banca, inversión, manufactura</p>
----	---	-------------	------	--	---	-------------------------------------	------------------------------------	--	---

32	Big data in manufacturing: a systematic mapping study	O'Donovan, P., Leahy, K., Bruton, K., O'Sullivan, D.T.J.	2015	<p><i>The manufacturing industry is currently in the midst of a data-driven revolution, which promises to transform traditional manufacturing facilities in to highly optimised smart manufacturing facilities. These smart facilities are focused on creating manufacturing intelligence from real-time data to support accurate and timely decision-making that can have a positive impact across the entire organisation. To realise these efficiencies emerging technologies such as Internet of Things (IoT) and Cyber Physical Systems (CPS) will be embedded in physical processes to measure and monitor real-time data from across the factory, which will ultimately give rise to unprecedented levels of data production. Therefore, manufacturing facilities must be able to manage the demands of exponential increase in data production, as well as possessing the analytical techniques needed to extract meaning from these large datasets. More specifically, organisations must be able to work with big data technologies to meet the demands of smart manufacturing. However, as big data is a relatively new phenomenon and potential applications to manufacturing activities are wide-reaching and diverse, there has been an obvious lack of secondary research undertaken in the area. Without secondary research, it is difficult for researchers to identify gaps in the field, as well as aligning their work with other researchers to develop strong research themes. In this study, we use the formal research methodology of systematic mapping to provide a breadth-first review of big data technologies in manufacturing</i></p>	Conjunto masivo de datos que a través del análisis se puede convertir in información para la planeación	Volumen, velocidad	Recolección, almacenamiento, análisis, uso	Cómo usar tecnología de macrodatos para la manufactura	Manufactura, planeación, automatización
----	---	--	------	--	---	--------------------	--	--	---

33	A novel social network competitive analytics framework with sentiment benchmarks	He, W., Wu, H., Yan, G., Akula, V., Shen, J.	2015	<p><i>In today's competitive business environment, there is a strong need for businesses to collect, monitor, and analyze user-generated data on their own and on their competitors' social network sites, such as Facebook, Twitter, and blogs. To achieve a competitive advantage, it is often necessary to listen to and understand what customers are saying about competitors' products and services. Current social network analytics frameworks do not provide benchmarks that allow businesses to compare customer sentiment on social network to easily understand where businesses are doing well and where they need to improve. In this paper, we present a social network competitive analytics framework with sentiment benchmarks that can be used to glean industry-specific marketing intelligence. Based on the idea of the proposed framework, new social network competitive analytics with sentiment benchmarks can be developed to enhance marketing intelligence and to identify specific actionable areas in which businesses are leading and lagging to further improve their customers' experience using customer opinions gleaned from social network. Guided by the proposed framework, an innovative business-driven social network competitive analytics tool named VOZIQ is developed. We use VOZIQ to analyze tweets associated with five large retail sector companies and to generate meaningful business insight reports</i></p>	Técnicas de gestión de información valiosas para tomas de decisiones	Veloz, voluminosa, valiosa (generada por el usuario)	Recolección, análisis, decisión	Conversión de la información extraída en decisiones valiosas	Mercadeo, ventas, competencia
----	--	--	------	---	--	--	---------------------------------	--	-------------------------------

34	How the use of big data analytics affects value creation in supply chain management	Chen, D.Q., Preston, D.S., Swink, M.	2015	<p><i>Despite numerous testimonials of first movers, the underlying mechanisms of organizations big data analytics (BDA) usage deserves close investigation. Our study addresses two essential research questions: (1) How does organizational BDA usage affect value creation? and (2) What are key antecedents of organizational-level BDA usage? We draw on dynamic capabilities theory to conceptualize BDA use as a unique information processing capability that brings competitive advantage to organizations. Furthermore, we employ the technology-organization-environment (TOE) framework to identify and theorize paths via which factors influence the actual usage of BDA. Survey data collected from 161 U.S.-based companies show that: organizational-level BDA usage affects organizational value creation; the degree to which BDA usage influences such creation is moderated by environmental dynamism; technological factors directly influence organizational BDA usage; and organizational and environmental factors indirectly influence organizational BDA usage through top management support. Collectively, these findings provide a theory-based understanding of the impacts and antecedents of organizational BDA usage, while also providing guidance regarding what managers should expect from usage of this rapidly emerging technology</i></p>	Información voluminosa, veloz, variada estructurada de naturalezas estructurada y no estructurada	Veloz, voluminosa, variada	Recolección de datos, organización, análisis, decisión, creación de valor	Falta de adopción organizacional de tecnologías de macrodatos.	Decisión, procesos, negocios, productividad, manejo de cadena de abastecimiento, crecimiento empresarial
----	---	--	------	---	---	----------------------------	---	--	--

35	<p><i>Big Data and Journalism: Epistemology, expertise, economics, and ethics</i></p>	<p>Lewis, S.C., Westlund, O.</p>	<p>2015</p>	<p><i>Big data is a social, cultural, and technological phenomenon—a complex amalgamation of digital data abundance, emerging analytic techniques, mythology about data-driven insights, and growing critique about the overall consequences of big-data practices for democracy and society. While media and communication scholars have begun to examine and theorize about big data in the context of media and public life broadly, what are the particular implications for journalism? This article introduces and applies four conceptual lenses—epistemology, expertise, economics, and ethics—to explore both contemporary and potential applications of big data for the professional logic and industrial production of journalism. These distinct yet inter-related conceptual approaches reveal how journalists and news media organizations are seeking to make sense of, act upon, and derive value from big data during a time of exploration in algorithms, computation, and quantification. In all, the developments of big data potentially have great meaning for journalism’s ways of knowing (epistemology) and doing (expertise), as well as its negotiation of value (economics) and values (ethics). Ultimately, this article outlines future directions for journalism studies research in the context of big data</i></p>	<p>Grandes volúmenes de información variada que, a través del análisis, puede generar valor</p>	<p>Voluminosa, variada, veloz</p>	<p>Recolección, organización, análisis, uso</p>	<p>Privacidad de la información</p>	<p>Periodismo, economía, toma de decisiones</p>
----	---	----------------------------------	-------------	--	---	-----------------------------------	---	-------------------------------------	---

36	A scalable framework for spatiotemporal analysis of location-based social network data	Cao, G., Wang, S., Hwang, M., Padmanabhan, A., Zhang, Z., Soltani, K.	2015	<p><i>In the past several years, social network (e.g., Twitter and Facebook) has experienced a spectacular rise in popularity and has become a ubiquitous location for discourse, content sharing and social networking. With the widespread adoption of mobile devices and location-based services, social network typically allows users to share the whereabouts of daily activities (e.g., check-ins and taking photos), thus strengthening the role of social network as a proxy for understanding human behaviors and complex social dynamics in geographic spaces. Unlike conventional spatiotemporal data, this new modality of data is dynamic, massive, and typically represented in a stream of unstructured media (e.g., texts and photos), which pose fundamental representation, modeling and computational challenges to conventional spatiotemporal analysis and geographic information science. In this paper, we describe a scalable computational framework to harness massive location-based social network data for efficient and systematic spatiotemporal data analysis. Within this framework, the concept of space-time trajectories (or paths) is applied to represent activity profiles of social network users. A hierarchical spatiotemporal data model, namely a spatiotemporal data cube model, is developed based on collections of space-time trajectories to represent the collective dynamics of social network users across aggregation boundaries at multiple spatiotemporal scales. The framework is implemented based upon a public data stream of Twitter feeds posted on the continent of North America. To demonstrate the advantages and performance of this framework, an interactive flow mapping interface (including both single-source and</i></p>	Cantidad masiva de datos generados que pueden utilizarse para extraer conclusiones valiosas	Volumen, velocidad	Recolección, almacenamiento, análisis, conclusiones	Usos de los datos para obtener conclusiones de valor	Geografía, estudios etnográficos, salud
----	--	---	------	--	---	--------------------	---	--	---

multiple-source flow mapping) is developed to allow real-time and interactive visual exploration of movement dynamics in massive location-based social network data at multiple scales

37	Big data and its epistemology	Frické, M.	2015	<p><i>The article considers whether Big Data, in the form of data-driven science, will enable the discovery, or appraisal, of universal scientific theories, instrumentalist tools, or inductive inferences. It points out, initially, that such aspirations are similar to the now-discredited inductivist approach to science. On the positive side, Big Data may permit larger sample sizes, cheaper and more extensive testing of theories, and the continuous assessment of theories. On the negative side, data-driven science encourages passive data collection, as opposed to experimentation and testing, and hornswoggling ("unsound statistical fiddling"). The roles of theory and data in inductive algorithms, statistical modeling, and scientific discoveries are analyzed, and it is argued that theory is needed at every turn. Data-driven science is a chimera</i></p>	Ciencia orientada desde los datos	Volumen, velocidad	Observación, organización, análisis, teorización	El inductivismo como epistemología desacreditada en la creación de conocimiento	Investigación, ciencia
----	-------------------------------	------------	------	---	-----------------------------------	--------------------	--	---	------------------------

38	<p><i>Big and Open Linked Data (BOLD) in government: A challenge to transparency and privacy?</i></p>	<p>Janssen, M., van den Hoven, J.</p>	<p>2015</p>	<p><i>Big and Open Linked Data (BOLD) results in new opportunities and have the potential to transform government and its interactions with the public. BOLD provides the opportunity to analyze the behavior of individuals, increase control, and reduce privacy. At the same time BOLD can be used to create an open and transparent government. Transparency and privacy are considered as important societal and democratic values that are needed to inform citizens and let them participate in democratic processes. Practices in these areas are changing with the rise of BOLD. Although intuitively appealing, the concepts of transparency and privacy have many interpretations and are difficult to conceptualize, which makes it often hard to implement them. Transparency and privacy should be conceptualized as complex, non-dichotomous constructs interrelated with other factors. Only by conceptualizing these values in this way, the nature and impact of BOLD on privacy and transparency can be understood, and their levels can be balanced with security, safety, openness and other socially-desirable values</i></p>	<p>Grandes volúmenes de datos de una variedad de fuentes que necesitan ser procesados</p>	<p>Volumen, variedad, valor</p>	<p>Recolección, procesamiento, uso</p>	<p>Privacidad de la información</p>	<p>Gobierno</p>
----	---	---------------------------------------	-------------	---	---	---------------------------------	--	-------------------------------------	-----------------

39	<p>Reflections on societal and business model transformation arising from digitization and big data analytics: A research agenda</p>	<p>Loebbecke, C., Picot, A.</p>	<p>2015</p>	<p><i>In the era of accelerating digitization and advanced big data analytics, harnessing quality data for designing and delivering state-of-the-art services will enable innovative business models and management approaches (Boyd and Crawford, 2012; Brynjolfsson and McAfee, 2014) and yield an array of consequences. Among other consequences, digitization and big data analytics reshape business models and impact employment amongst knowledge workers - just as automation did for manufacturing workers. This Viewpoint paper considers the mechanisms underlying how digitization and big data analytics drive the transformation of business and society and outlines the potential effects of digitization and big data analytics on employment - especially in the context of cognitive tasks. Its aim is to outline a critical research agenda to explore and conceptualize evident changes in business models and society arising from these technological advances</i></p>	<p>Grandes volúmenes de datos no estructurados que se mueven con rapidez</p>	<p>Volumen, velocidad</p>	<p>Recolección, análisis, valor</p>	<p>Incapacidad de las organizaciones de extraer valor de la gestión de macrodatos</p>	<p>Modelos de negocios, empleo, transformación social</p>
----	--	---------------------------------	-------------	--	--	---------------------------	-------------------------------------	---	---

40	<p><i>Big Data and analytics in higher education: Opportunities and challenges</i></p>	Daniel, B.	2015	<p><i>Institutions of higher education are operating in an increasingly complex and competitive environment. This paper identifies contemporary challenges facing institutions of higher education worldwide and explores the potential of Big Data in addressing these challenges. The paper then outlines a number of opportunities and challenges associated with the implementation of Big Data in the context of higher education. The paper concludes by outlining future directions relating to the development and implementation of an institutional project on Big Data</i></p>	Análisis de datos para la toma de decisiones informadas	Volumen	Recolección, análisis, extracción de conocimiento	Capacidad de análisis de grandes volúmenes de datos	Negocios, gobierno, salud, educación
----	--	------------	------	---	---	---------	---	---	--------------------------------------

41	Simulation optimization: A review and exploration in the new era of cloud computing and big data	Xu, J., Huang, E., Chen, C.-H., Lee, L.H.	2015	<p>Recent advances in simulation optimization research and explosive growth in computing power have made it possible to optimize complex stochastic systems that are otherwise intractable. In the first part of this paper, we classify simulation optimization techniques into four categories based on how the search is conducted. We provide tutorial expositions on representative methods from each category, with a focus in recent developments, and compare the strengths and limitations of each category. In the second part of this paper, we review applications of simulation optimization in various contexts, with detailed discussions on health care, logistics, and manufacturing systems. Finally, we explore the potential of simulation optimization in the new era. Specifically, we discuss how simulation optimization can benefit from cloud computing and high-performance computing, its integration with big data analytics, and the value of simulation optimization to help address challenges in engineering design of complex systems</p>	Sistema de información compuesto por grandes cantidades de datos de los que se puede extraer valor en sistemas predictivos	Voluminosa, valiosa	Recolección, organización, análisis, uso	Conversión en información valiosa de los macrodatos recolectados	Modelos predictivos, simulaciones
----	--	---	------	---	--	---------------------	--	--	-----------------------------------

42	Editorial Essay: What Is Organizational Research For?	Davis, G.F.	2015	<p><i>Organizational research is guided by standards of what journals will publish and what gets rewarded in scholarly careers. This system can promote novelty rather than truth and impact rather than coherence. The advent of big data, combined with our current system of scholarly career incentives, is likely to yield a high volume of novel papers with sophisticated econometrics and no obvious prospect of cumulative knowledge development. Moreover, changes in the world of organizations are not being met with changes in how and for whom organizational research is done. It is time for a dialogue on who and what organizational research is for and how that should shape our practice</i></p>	Sistemas de datos voluminosos que, al ser analizados, revelan información	Voluminosa, veloz, valiosa, variada	Recolección, análisis, uso	Privacidad de la información, uso ético de la información	investigación organizacional
----	--	-------------	------	--	---	-------------------------------------	----------------------------	---	------------------------------

43	<p><i>Harvesting big data to enhance supply chain innovation capabilities: An analytic infrastructure based on deduction graph</i></p>	<p>Tan, K.H., Zhan, Y., Ji, G., Ye, F., Chang, C.</p>	<p>2015</p>	<p><i>Today, firms can access to big data (tweets, videos, click streams, and other unstructured sources) to extract new ideas or understanding about their products, customers, and markets. Thus, managers increasingly view data as an important driver of innovation and a significant source of value creation and competitive advantage. To get the most out of the big data (in combination with a firm's existing data), a more sophisticated way of handling, managing, analysing and interpreting data is necessary. However, there is a lack of data analytics techniques to assist firms to capture the potential of innovation afforded by data and to gain competitive advantage. This research aims to address this gap by developing and testing an analytic infrastructure based on the deduction graph technique. The proposed approach provides an analytic infrastructure for firms to incorporate their own competence sets with other firms. Case studies results indicate that the proposed data analytic approach enable firms to utilise big data to gain competitive advantage by enhancing their supply chain innovation capabilities</i></p>	<p>Cantidad enorme de datos variados generados con gran velocidad con la que los sistemas de información tradicional no pueden lidiar</p>	<p>Volumen, velocidad, variedad</p>	<p>Recolección, organización, uso</p>	<p>Tamaño de los datos, velocidad en que son generados, insuficiencia de sistemas de análisis para convertir los datos en información valiosa</p>	<p>Mercado, diseño de productos, negocios, ventaja competitiva, cadena de abastecimiento</p>
----	--	---	-------------	--	---	-------------------------------------	---------------------------------------	---	--

44	The value of big data in servitization	Opresnik, D., Taisch, M.	2015	<p>Abstract Servitization has become a pervasive business strategy among manufacturers, enabling them to undergird their competitive advantage. However, it has at least one weakness. While it is used worldwide also in economies with lower production costs, services in manufacturing are slowly becoming commoditized and will become a necessary, though not sufficient, condition for reaching an above average competitive advantage. Consequently, in this article we propose a new basis for competitive advantage for manufacturing enterprises called a Big Data Strategy in servitization. We scrutinize how manufacturers can exploit the opportunity arising from combined Big Data and servitization. Therefore, the concept of a Big Data Strategy framework in servitization is proposed. The findings are benchmarked against established frameworks in the Big Data and servitization literature. Its impact on competitive advantage is assessed through three theoretical perspectives that increase the validity of the results. The main finding is that, through the proposed strategy, new revenue streams can be created, while opening the possibility to decrease prices for product-services. Through the proposed strategy manufacturers can differentiate themselves from the ones that are already servitizing. This article introduces the possibility of influencing the most important of the five "Vs" in Big Data-Value, in addition to the other four "Vs" - Volume, Variety, Velocity and Verification. As in regards to servitization, the article adds a third layer of added value - "information", beside the two existing ones: product and service. The results have strategic implications for managers</p>	Gran cantidad de datos generados y recolectados a través de diferentes canales	Volumen, velocidad, variedad, veracidad, valor	Generación, recolección, almacenamiento, procesamiento, uso, valor	Sobrecarga de información	Generación de valor, diseño de servicios para clientes
----	--	--------------------------	------	---	--	--	--	---------------------------	--

45	<p><i>Is Bigger Always Better? Potential Biases of Big Data Derived from Social Network Sites</i></p>	Hargittai, E.	2015	<p><i>This article discusses methodological challenges of using big data that rely on specific sites and services as their sampling frames, focusing on social network sites in particular. It draws on survey data to show that people do not select into the use of such sites randomly. Instead, use is biased in certain ways yielding samples that limit the generalizability of findings. Results show that age, gender, race/ethnicity, socioeconomic status, online experiences, and Internet skills all influence the social network sites people use and thus where traces of their behavior show up. This has implications for the types of conclusions one can draw from data derived from users of specific sites. The article ends by noting how big data studies can address the shortcomings that result from biased sampling frames</i></p>	<p>Cantidad masiva de datos generados a través de diferentes canales, como redes sociales, de los que se puede extraer información.</p>	<p>Veloz, voluminosa, variada</p>	<p>Recolección, análisis, uso</p>	<p>Privacidad de la información, uso ético de la información, falta de conceptualización y exceso de datos prácticos</p>	<p>Investigación, análisis de comportamiento, ciencias sociales</p>
----	---	---------------	------	--	---	-----------------------------------	-----------------------------------	--	---

46	New games, new rules: Big data and the changing context of strategy	Constantiou, I.D., Kallinikos, J.	2015	<p><i>Big data and the mechanisms by which it is produced and disseminated introduce important changes in the ways information is generated and made relevant for organizations. Big data often represents miscellaneous records of the whereabouts of large and shifting online crowds. It is frequently agnostic, in the sense of being produced for generic purposes or purposes different from those sought by big data crunching. It is based on varying formats and modes of communication (e.g., texts, image and sound), raising severe problems of semiotic translation and meaning compatibility. Crucially, the usefulness of big data rests on their steady updatability, a condition that reduces the time span within which this data is useful or relevant. Jointly, these attributes challenge established rules of strategy making as these are manifested in the canons of procuring structured information of lasting value that addresses specific and long-term organizational objectives. The developments underlying big data thus seem to carry important implications for strategy making, and the data and information practices with which strategy has been associated. We conclude by placing the understanding of these changes within the wider social and institutional context of longstanding data practices and the significance they carry for management and organizations</i></p>	Grandes volúmenes de datos generados a través de los ecosistemas digitales	Volumen, variedad	Recolección, análisis, uso	Transformación de los macrodatos en información útil	Sistemas de información (IS), gerencia, ciencias sociales, competitividad, negocios
----	---	-----------------------------------	------	---	--	-------------------	----------------------------	--	---

47	Challenges and solutions for marketing in a digital era	Eeflang, P.S.H., Verhoef, P.C., Dahlström, P., Freundt, T.	2015	<p><i>Internet usage continues to explode across the world with digital becoming an increasingly important source of competitive advantage in both B2C and B2B marketing. A great deal of attention has been focused on the tremendous opportunities digital marketing presents, with little attention on the real challenges companies are facing going digital. In this study, we present these challenges based on results of a survey among a convenience sample of 777 marketing executives around the globe. The results reveal that filling "talent gaps", adjusting the "organizational design", and implementing "actionable metrics" are the biggest improvement opportunities for companies across sectors</i></p>	Conjunto de datos voluminosos y complejos que no pueden ser analizados con métodos tradicionales	Voluminosa, variada	Captura, curación, almacenamiento, búsqueda, compartición, transferencia, análisis, visualización	Procesamiento de datos tan voluminosos	Manejo de clientes, diseño de producto, estrategia, mercadeo
----	---	--	------	---	--	---------------------	---	--	--

48	Metaphors of big data	Puschmann, C., Burgess, J.	2014	<p><i>Metaphors are a common instrument of human cognition, activated when seeking to make sense of novel and abstract phenomena. In this article we assess some of the values and assumptions encoded in the framing of the term big data, drawing on the framework of conceptual metaphor. We first discuss the terms data and big data and the meanings historically attached to them by different usage communities and then proceed with a discourse analysis of Internet news items about big data. We conclude by characterizing two recurrent framings of the concept: as a natural force to be controlled and as a resource to be consumed</i></p>	Conjunto de datos recolectados que pueden ser transformados en información valiosa a partir de análisis	Volumen, velocidad, variedad	Generación, organización, uso	Control de la generación y uso de los macrodatos	Negocios, análisis, modelos predictivos
----	-----------------------	----------------------------	------	---	---	------------------------------	-------------------------------	--	---

49	<i>The big data divide</i>	Andrejevic, M.	2014	<p><i>This article extends the notion of a "big data divide" to describe the asymmetric relationship between those who collect, store, and mine large quantities of data, and those whom data collection targets. It argues that this key distinction highlights differential access to ways of thinking about and using data that potentially exacerbate power imbalances in the digital era. Drawing on original survey and interview findings about public attitudes toward collection and use of personal information, it maintains that the inability to anticipate the potential uses of such data is a defining attribute of data-mining processes, and thus of the forms of sorting and targeting that result from them</i></p>	Cantidad masiva de datos generados y recolectados a través de diferentes canales	Volumen, velocidad, variedad	Generación, recolección, análisis	Desbalance de poderes entre los que generan datos y los que los recolectan	Investigación
----	----------------------------	----------------	------	---	--	------------------------------	-----------------------------------	--	---------------

50	Using 'Big Data' for analytics and decision support	Power, D.J.	2014	<p>People and the computers they use are generating large amounts of varied data. The phenomenon of capturing and trying to use all of the semi-structured and unstructured data has been called by vendors and bloggers 'Big Data'.</p> <p>Organizations can capture and store data of many types from almost any source but capturing and storing data only adds value when it has a useful purpose. Big Data must be used to provide input to analytics and decision support capabilities if it is to create real value for organizations. Some bloggers, industry leaders and academics have become disillusioned by the term Big Data. It is a marketing term and not a technical term. More descriptive terms like unstructured data, process data and machine data are more useful for information technology (IT) professionals. Researchers need to study, and document use cases that explain how specific, novel data, so-called Big Data, can be used to support decision-making</p>	Conjunto de datos tan voluminosos que los sistemas tradicionales de análisis no los pueden procesar	Volumen, velocidad, variedad	Generación, captura, análisis, uso	Transformación de los macrodatos en información útil y valiosa	Negocios, investigación, desarrollo
----	---	-------------	------	--	---	------------------------------	------------------------------------	--	-------------------------------------

Fuente: elaboración propia