



El uso del Big Data en los estudios de opinión pública

Mtra. Erika Tapia Nava

Agosto, 2018

46

**SENADO DE LA REPÚBLICA
INSTITUTO BELISARIO DOMÍNGUEZ**

Comité Directivo

**Senador Manuel Bartlett Díaz
Presidente**

**Senador Daniel Gabriel Ávila Ruíz
Secretario**

**Senador Marlon Berlanga Sánchez
Secretario**

**Secretario Técnico
Onel Ortíz Fragoso**

**Coordinador Ejecutivo de Investigación
Mtro. Juan Pablo Arroyo Ortíz**

Junta Ejecutiva

**Mtro. Alejandro Encinas Nájera
Director General de Investigación Estratégica**

**Mtro. Noel Pérez Benítez
Director General de Finanzas**

**Mtro. Juan Carlos Amador Hernández
Director General de Difusión y Publicaciones**

**Dr. Alejandro Navarro Arrendondo
Director General de Análisis Legislativo**

ÍNDICE

RESUMEN EJECUTIVO.....	4
INTRODUCCIÓN.....	5
1. ¿QUÉ ES EL BIG DATA?.....	7
1.1 Características del Big Data.....	7
1.2 La generación de datos y el Big Data.....	10
1.3 Antecedentes Big Data en la opinión pública.....	13
1.4 Supuestos teóricos metodológicos del uso del Big Data.....	15
1.4.1 Consideraciones teóricas.....	15
1.4.2 Consideraciones metodológicas y epistemológicas.....	17
1.5 Los usos del Big Data. Una aproximación nacional e internacional.....	20
1.5.1 Aplicaciones del Big Data en el gobierno.....	22
1.6.1 Normas y principales obligaciones legales en materia de privacidad en el uso del Big Data.....	25
2. PRINCIPALES USOS EN LA INVESTIGACIÓN PARLAMENTARIA.....	30
2.1 Experiencias de políticas públicas y Congresos que están utilizando Big Data.....	30
2.2 Alcances del Big Data en el proceso legislativo.....	33
2.3 Open data y el análisis legislativo.....	
2.4 Tendencias y patrones de comportamiento en la opinión pública mediante el Big Data.....	43
CONCLUSIONES.....	49
BIBLIOGRAFÍA.....	52

RESUMEN EJECUTIVO

- ⑥ **El Big Data surge como una nueva fase del paradigma intensivo en comunicación e información que emerge a partir de la revolución y el cambio tecnológico iniciado en la década de los años setenta.**
- ⑥ **En la actualidad es aplicado y usado principalmente en el sector empresarial en los Estados Unidos, Reino Unido y algunos países de la Unión Europea.**
- ⑥ **La materialización de los potenciales beneficios para los países en desarrollo como México, requiere de la elaboración de políticas activas y específicas que tengan en cuenta la generación y apropiación de las rentas informacionales, el manejo de la privacidad en cuanto al uso de los datos personales, el desarrollo de capacidades para la creación de valor y la difusión de la información, así como el conocimiento para contribuir a la reducción de las desigualdades socioeconómicas.**
- ⑥ **El uso del Big Data y su implementación en investigación parlamentaria, constituye un área de oportunidad, ya que ofrece la posibilidad de conocer la opinión pública de la ciudadanía de manera más precisa en tiempo real, a un menor costo.**
- ⑥ **Este tipo de datos también pueden constituir un insumo en el diseño de las leyes y la implementación de las políticas públicas, aspectos que constituyen un nuevo nicho que debe ser explorado e impulsado.**
- ⑥ **La información generada en Big Data se podría utilizar para conocer mejor a la opinión pública, así como enviar un “mejor mensaje”. Es decir, se podría conocer de manera precisa al ciudadano, sus características, preferencias, necesidades y deseos. Esto permitirá acercarse más a determinados segmentos de población, optando por la microsegmentación y la geolocalización para aplicar estrategias personalizadas y diferenciadas que permitan presentar propuestas de comunicación a la medida, así como la identificación de necesidades de diversos sectores sociales.**

INTRODUCCIÓN

*Big Data permite a la información interpretarse a sí misma
y adelantarse a nuestras intenciones,
cuánto saben las grandes empresas de nosotros,
y lo que más le preocupa:
lo fácil que está siendo convertir la democracia
en una dictadura de la información,
haciendo de cada ciudadano una burbuja distinta.*

Martin Hilbert, experto en redes digitales

El impulso y desarrollo del Big Data abre diversas oportunidades para el análisis de datos de manera estratégica y para la toma de decisiones inteligentes, a su vez pone de manifiesto la necesidad de generar capacidades competitivas en cada uno de los países que lo promueven o buscan promoverlo.

Desde este marco, Big Data supone un tipo de análisis dinámico que convierte las tendencias de opinión en cifras medibles y cuantificables. Estos datos, si se saben explotar, constituyen insumos básicos y ricos desde la opinión pública para analizar una campaña electoral, comunicar algún mensaje e idea desde el gobierno, conocer el impacto de alguna decisión en la población y orientar o prever algún tipo de comportamiento. Es decir, mediante el análisis con Big Data, se puede llegar al conocimiento de la población de manera oportuna, efectiva y con mayor nivel de precisión en tiempo.

Este contexto de transformación digital no ha sido ajeno a los sistemas políticos, lo cual supone avances y grandes desafíos para las democracias y los parlamentos. Particularmente, el uso del Big Data y su implementación en investigación legislativa constituye un área de oportunidad, ya que ofrece la posibilidad de conocer la opinión pública de la ciudadanía de manera más precisa y a su vez permite analizar las demandas de la población.

Para dar cuenta de ello, el presente reporte de investigación está organizado en dos apartados analíticos. En el primero, denominado “¿Qué es el Big Data?”, se abordan las características, la forma en cómo se generan los datos, algunos antecedentes en la opinión pública, los supuestos teórico metodológicos y su aplicación en asuntos de gobierno y, se aborda el marco normativo en relación a las obligaciones legales en materia de privacidad datos en el uso del Big Data. En el segundo apartado, denominado “Principales usos en la investigación parlamentaria”, se analizan algunas experiencias de políticas públicas y Congresos que están utilizando Big Data y, se da cuenta de los alcances que puede tener el proceso legislativo con esta herramienta. Más adelante, se identifican las tendencias y

patrones de comportamiento en la opinión pública mediante el manejo de datos a gran escala. Al respecto, se observa que las principales tendencias en el uso de Big Data en la investigación parlamentaria permitirían:

- La realización de análisis legislativo y propuestas de ley con un mayor sustento y conocimiento sobre lo que piensa y le interesa a la ciudadanía.
- El aumento de la confianza en el gobierno mediante el conocimiento de las políticas públicas, su ejecución e impacto.
- La construcción colaborativa de propuestas de ley entre el Estado y la sociedad, lo cual tiende a reducir los costos en la implementación y elevar los niveles de cumplimiento y satisfacción, así como de legitimidad con las demandas sociales.
- El fortalecimiento de los mecanismos de participación y colaboración interinstitucional.

Desde esta lógica, se podría utilizar dicha información para enviar un “mejor mensaje” o, conocer la opinión pública y en función de ello orientar el análisis legislativo. Es decir, se podría conocer de manera precisa al ciudadano, sus características, preferencias, necesidades y deseos. Esto permitirá acercarse más a determinados segmentos de población, optando por la microsegmentación y la geolocalización para aplicar estrategias personalizadas y diferenciadas que permitan presentar propuestas de comunicación a la medida y necesidades. Además, los propios datos y su visualización pueden convertirse en material para construir o complementar narrativas, contribuyendo a que ideas complejas puedan ser comunicadas con mayor claridad precisión y de una manera que resulte sencillas y atractivas.

El análisis parlamentario, el conocimiento de la opinión pública y el diseño de las políticas públicas deben considerar de manera necesaria la generación y análisis de datos. Si no se piensa en ellos, se reduce su capacidad de acción y planeación. En este sentido, el Big Data puede constituirse en una herramienta social, como un elemento que revitalice y renueve la acción política a partir del análisis de datos con conocimientos útiles para la acción y la gestión política desde diversas esferas.

1. ¿Qué es el Big Data?

1.1 Características del Big Data

Con el término Big Data se hace referencia a la tendencia del avance de las tecnologías que han cambiado la perspectiva sobre el uso de la información y la toma de decisiones. De acuerdo con López García (2012), desde la presentación del término por McKinsey Global Institute (MGI) en el año 2011, han existido diversos intentos de acotación del concepto. Se ha definido Big Data como el conjunto de datos cuyo tamaño va más allá de la capacidad de captura, almacenado, gestión y análisis de las herramientas tradicionales de base de datos. (Manyika, J y otros, 2011). Por su parte, Gartner define el Big Data como los “activos de información caracterizados por su alto volumen, velocidad y variedad, que demandan soluciones innovadoras y eficientes de procesado para la mejora del conocimiento y toma de decisiones en las organizaciones” (Gartner, 2012:15).

El Big Data supone enormes cantidades de datos que llevaría demasiado tiempo cargarlos en una base de datos relacional para su posterior análisis. Por otro lado, el término hace referencia al tratamiento y análisis de grandes repositorios de datos; estos repositorios varían en tamaño y tipo de fuentes. Asimismo, este concepto se aplica a toda aquella información que no pueda ser procesada por los métodos tradicionales. Según el artículo “*In Perspective*”, elaborado por la organización *Fidelity Worldwide Investment*, Big Data es “el término inglés que designa los conjuntos de datos de gran tamaño y generalmente desestructurados que resultan difíciles de manejar usando las aplicaciones de bases de datos convencionales” (2012: 1).

Esta diversidad supone en esencia una gran heterogeneidad de datos y la representación e interpretación semántica de los mismos. De hecho, esta producción de datos no está directamente lista para ser integrada y ocupada en alguna aplicación. De igual forma, en esta riqueza y variedad de datos se debe contemplar la frecuencia y la disponibilidad de los mismos, así como el tiempo que se dispone para dar respuesta de forma tal que ésta sea oportuna para la toma de decisiones (Malvicino y Yogue, 2015). De acuerdo con el texto, *Fidelity Worldwide Investmen*, el Big Data, tiene las siguientes dimensiones o características:

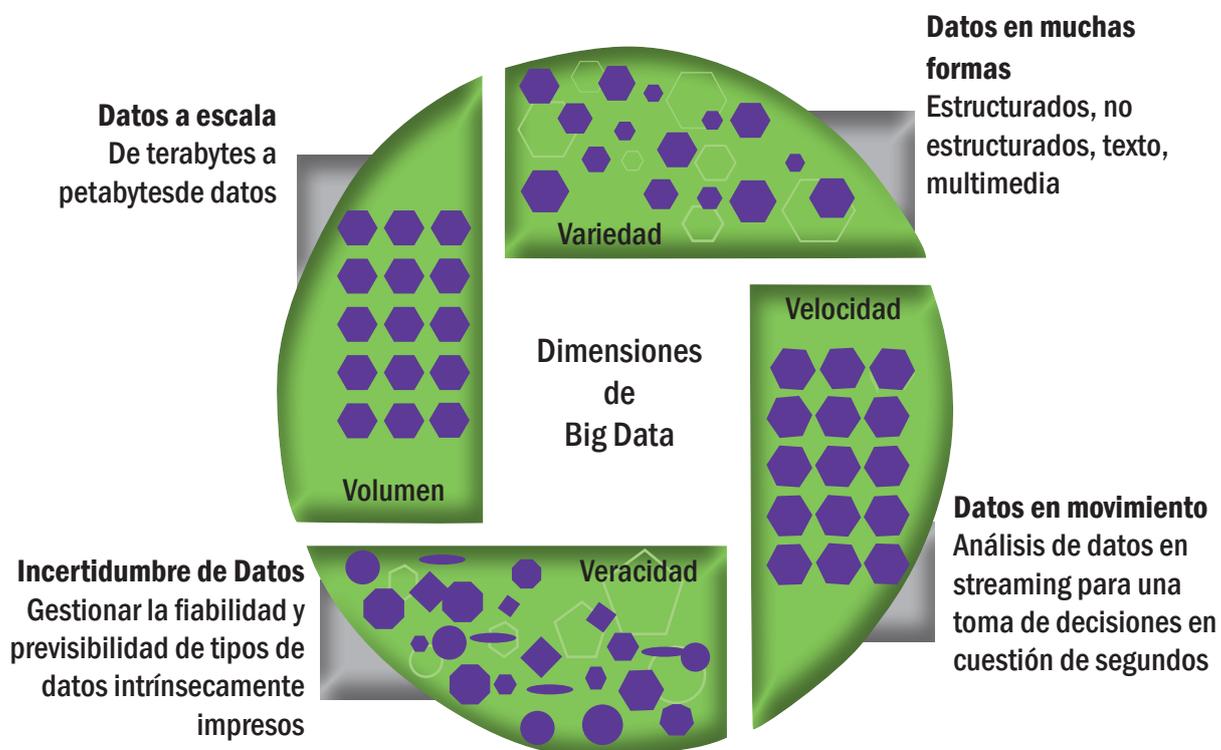
Tabla 1.
Dimensiones o características del Big Data

variedad	Implica una diversidad de tipos, fuentes y datos. La variedad tiene que ver con la complejidad de gestionar la multiplicidad de datos, incluidos los estructurados, semiestructurados y no estructurados. Esta diversidad de datos pueden provenir tanto de información tradicional como no tradicional, propias de una organización, institución o de otras instancias. Esta diversidad de fuentes puede abarcar múltiples fuentes textos, datos, web, tuits, datos de sensores, audios, vídeo y archivos de registro, entre otros.
velocidad	La velocidad con la que se generan, procesan y analizan es dinámica y cada vez continúan aumentando de manera acelerada. La velocidad afecta la latencia: el tiempo de espera entre el momento en el que se crean los datos, el momento en el que se captan y en el que están accesibles. Hoy en día, los datos se generan de forma continua a una velocidad a la que los sistemas tradicionales les resulta imposible captarlos, almacenarlos y analizarlos. Ello cobra principal relevancia cuando hay ciertos datos que deben analizarse en tiempo real para la toma de decisiones.
volumen	Es la cantidad masiva de datos, es la principal característica que se asocia con el Big Data. Los volúmenes de datos continúan aumentando a un ritmo sin precedentes de manera escalar, aspectos que varían en función del sector e incluso de la ubicación geográfica.
veracidad	La calidad de los datos es un factor central. Es importante obtener datos fiables. Ello constituye un elemento importante y un reto fundamental y constante en la producción y generación de análisis con Big Data. Se debe considerar que aún con la aplicación de métodos de limpieza de datos no siempre se puede eliminar la imprevisibilidad inherente de algunos datos: tiempo, economía o futuras decisiones. La necesidad de reconocer y planificar la incertidumbre es una dimensión de Big Data que pretende abordar y debe ser contemplada en los distintos ámbitos donde se aplique este tipo de análisis debido a que algunos tipos de datos son intrínsecamente inciertos, pero a pesar de ello pueden contener información valiosa. ¹

Fuente: Elaboración propia a partir de Michael Schroeck y otros, *Analytics: el uso de Big Data en el mundo real. Cómo las empresas más innovadoras extraen valor de datos inciertos*, IBM Institute for Business Value, 2012.

¹ La incertidumbre se manifiesta de diversas formas, principalmente se genera desconfianza e incertidumbre sobre la información generada por los entornos derivados de redes sociales, ya que no hay certeza acerca de la forma en cómo será su comportamiento futuro y su variabilidad, así como la forma en cómo las personas reaccionarán sobre un determinado tema y el papel que tendrá el mercado, entre otros aspectos. La manera de superar esta limitante es fusionando los datos para dar paso a que éstos sean más precisos mediante el uso de algunas técnicas matemáticas avanzadas.

Gráfica 1.
Dimensiones o características del Big Data



Fuente: Elaboración propia a partir de Michael Schroeck y otros, *Analytics: el uso de Big Data en el mundo real. Cómo las empresas más innovadoras extraen valor de datos inciertos*, IBM Institute for Business Value, 2012.

De igual manera, en el informe *Big data. Avances recientes a nivel internacional y perspectivas para el desarrollo local*, se menciona que Big Data “son datos que exceden la capacidad de procesamiento de los sistemas de base de datos convencionales, caracterizados por un gran volumen, por la gran velocidad con la que se transmiten”, (Malvicino y Yogue, 2015: 26). Además, estos datos por sus propias características no se ajustan a las restricciones de las bases de datos convencionales que tradicionalmente se han usado. Asimismo, se menciona en este texto que la definición de Big Data aceptada a la fecha, se funda en las características y limitaciones que este nuevo paradigma tecnológico informacional impone a las capacidades técnicas y cognitivas actuales, por lo cual el Big Data exige cambios en la forma de procesar los datos, de pensar y de plantear los problemas. En esta línea, siguiendo a Hilbert (2013a), lo crucial de Big Data no es en sí mismo el gran aumento de los datos, sino su análisis para la toma de decisiones inteligente, por lo que prefiere la expresión: “análisis de Big Data” a la de “Big Data”. Esta serie de características del Big Data son fundamentales; sin embargo, según la publicación UN Global Pulse (2012), es necesario incluir la privacidad de datos personales, la facilidad de uso y acceso la información de éstos datos recabados.

1.2 La generación de datos y el Big Data

En la actualidad los datos se generan de una gran diversidad de fuentes: redes sociales, *Smartphones*, *Smart Cities*, las empresas, las denominadas comunicaciones M2M (*machine to machine*), sensores digitales, ya sean de medición eléctrica o de temperatura y sísmicos, entre otros. Al respecto, se espera que este número de datos e información crezca progresivamente y de manera vertiginosa donde a cada instante se genera multitud de información (López García, 2012).

Como menciona el *Código de buenas prácticas en protección de datos para proyectos Big Data*, el volumen es la característica más obvia y recoge el propio nombre de “Big Data”. Se pasa de manejar magnitudes de megabytes, gigabytes, como mucho terabytes, a manejar petabytes de forma cada vez más frecuente: el volumen de datos y variedad ha crecido exponencialmente, tanto por la tipología de datos como por sus fuentes.

Tabla 2.
Tamaños en Bytes

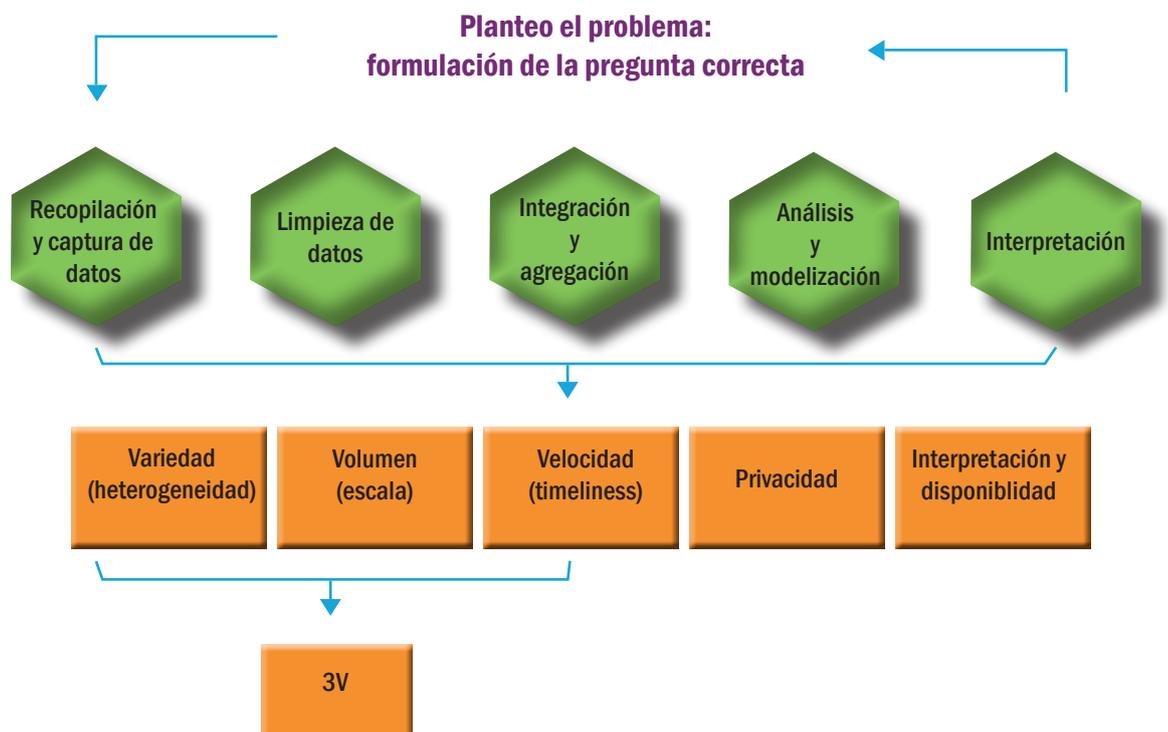
Gigabytes = 1,000,000,000 = 10 ⁹ bytes
Terabytes=1,000,000,000,000 = 10 ¹² bytes
Peta bytes=1,000,000,000,000,000 = 10 ¹⁵ bytes
Exabytes=1,000,000,000,000,000,000 = 10 ¹⁸ bytes
Zettabytes=1,000,000,000,000,000,000,000 = 10 ²¹ bytes
Yottabyte=10 ²⁴ bytes

Fuente: IBM, *Institute for Business Value*, 2012

Se ha pasado de manejar datos estructurados en bases de datos procedentes, en su mayoría, de fuentes internas, a tratar datos estructurados, semiestructurados y desestructurados; de ser datos cuasi estáticos a datos dinámicos o en continuo cambio; de originarse en un número de fuentes limitadas a proceder de personas, máquinas, sensores, etc. Esta variedad y volumen, requieren un tratamiento diferente para poder convertirse en información.

El tiempo es clave, así que la velocidad, captura, movimiento y proceso de los datos se hace a gran velocidad, llegando a ser en tiempo real en algunos casos. Además, algunos autores y organizaciones han añadido nuevas características para definir de forma más precisa al Big Data, por ejemplo, “veracidad” (la calidad de los datos capturados es clave), “variabilidad” (el significado de los datos cambia frecuentemente y se pueden producir inconsistencias que se han de manejar) y “valor” (los ingresos o beneficios del Big Data), (Sáiz y otros, 2015).

Gráfica 2.
Secuencia de análisis de Big Data



Fuente: United Nation, Global Pulse, *Harnessing Big Data for Developemet and humanitarian action*, 2012

De igual forma, se distingue que el Big Data se integra por una serie de datos con una gran diversidad de fuentes (internas y externas) y formatos (variedad), distintos niveles de velocidad, (Sáiz y otros, 2015). De acuerdo con el *Código de buenas prácticas en protección*

de datos para proyectos Big Data, desde un punto de vista de arquitectura y tecnología, se puede estructurar un sistema Big Data en cinco capas principales (Sáiz y otros, 2015: 9):

1. **Capa de fuentes de datos:** en esta capa estarían todos los orígenes de la información, desde bases de datos relacionales hasta cualquier tipo de datos, estructurados o no.
2. **Capa de integración:** aquí se adquieren los datos y se integran en conjuntos con el formato adecuado.
3. **Capa de almacenamiento de datos:** el conjunto de recursos adecuados para el almacenamiento de grandes volúmenes de datos.
4. **Capa de análisis y modelos de computación:** esto incluye diversas herramientas de manejo de datos, que operan sobre los recursos de almacenamiento e incluyen la gestión de los datos y los modelos de programación.
5. **Capa de presentación y aplicación:** incluye las tecnologías de visualización tales como dispositivos móviles, navegadores, etc. Una vez obtenido el conocimiento, éste se puede aplicar en distintos procesos.

Cabe destacar que el impacto inicial del Big Data, se ha producido por la gran cantidad de información creada, la cual tiene que ser almacenada y tiene la propiedad de proceder de diversas fuentes y en consecuencia ser muy heterogénea. Es decir, se usan datos tanto estructurados en forma de bases de datos tradicionales, pero también se maneja todo tipo de actividad producida por los usuarios, como audio, video, imágenes, conversaciones. Ello complejiza y hace muy difícil los procesos de verificación sobre la calidad y algunas veces la información utilizada contiene algunas imprecisiones. Dicha situación se deberá resolver con metodologías y herramientas adecuadas para dicha diversidad de información. Como se menciona en dicho código, el origen de datos es el primer aspecto que debe tenerse en cuenta en la cadena de tratamientos contemplados en un sistema de Big Data. Una parte importante de la complejidad del análisis de estos tratamientos ocurrirá en aquellos casos en que el sistema se nutra de información proveniente de múltiples orígenes.

Principalmente el análisis con Big Data, se ha dado desde el sector empresarial dada su capacidad para influir profundamente en el comercio de una economía integrada a escala global, además de proporcionar soluciones a antiguos retos empresariales. Sin embargo, Big Data está impulsando nuevas formas de transformar procesos, empresas, sectores eco-

nómicos e incluso la propia sociedad. Reconocer lo anterior es importante, ya que Big Data alcanza en la actualidad todos los ámbitos de la vida social: bolsa, climatología, astronomía, salud, política, el mercado, etc. y la cantidad de datos que se genera actualmente en diversas vertientes es abrumadora.

El gran reto del Big Data implica convertir datos imperfectos, complejos y desestructurados en información procesable, que sirva para la toma de decisiones, aunque esto no se da automáticamente. De acuerdo con el informe *Big Data. Avances recientes a nivel internacional y perspectivas para el desarrollo local*, su uso presenta algunos retos; por un lado, los desafíos tecnológicos de infraestructura y servicios –software y hardware–, y, por el otro, la dimensión del tipo de análisis que se requiere (Malvicino y Yogue, 2015). A ello se debe considerar la infraestructura y servicios que dispone para impulsar el análisis con Big Data.

Es necesario identificar que el aprovechamiento de Big Data para el desarrollo social y económico requiere un enfoque multidisciplinario y el diseño de políticas estratégicas que permitan el desarrollo de estudios para la toma de decisiones. Sin embargo, hay que tener en cuenta que los desafíos que se presentan con Big Data, no sólo están referidos al volumen de datos, sino también a su heterogeneidad, a su carácter poco estructurado, a los aspectos de privacidad, a los problemas de procesamiento, a la visualización, entre otros, lo cual complejiza el desarrollo de diversos estudios si no se consideran estos aspectos. A pesar de lo anterior, el Big Data abre la posibilidad de hacer grandes avances en muchas disciplinas científicas, aportes a la innovación y mejoras de productividad y competitividad, además de constituirse en un nicho importante para la realización y planeación de políticas públicas, así como de análisis legislativos (Malvicino y Yogue, 2015).

1.3 Antecedentes Big Data en la opinión pública

El Big Data ofrece oportunidades para alcanzar mejoras en competitividad, productividad y brinda opciones para la mejora en la calidad de vida. Éste surge como una nueva fase del paradigma intensivo en comunicación e información que emerge a partir de la revolución y el cambio tecnológico iniciado en la década de los años setenta.

Un ejemplo interesante de cómo es usado Big Data, es Wal-Mart. Esta gran cadena de supermercados de los Estados Unidos distribuyó con anticipación lo que iban a comprar los ciudadanos al acercarse el huracán Katrina en el año 2005. Esta preparación, se realizó

haciendo un análisis sobre lo que compraban sus clientes en este tipo de eventos, lo que le permitió estar preparada y poder satisfacer ese pico de demanda atípica.

Se puede observar que en el desarrollo y visualización del Big Data como una herramienta útil para identificar la opinión de la población y el electorado, la empresa *Penn and Schoen* en el año de 1994 influyó, para ayudar al Presidente Bill Clinton a mejorar la imagen de su partido, la cual se había deteriorado con las derrotas de ese año. Más adelante en los años siguientes a la presidencia de Bill Clinton, Mark Penn, se convirtió en uno de los consejeros políticos más importantes en la campaña de 1996. En éste ámbito, Penn orientó a Clinton para que se centrara en políticas que atraían a los sectores suburbanos y comenzó a tratar a estas políticas en términos de valores y lo que representaba para estos sectores, más que de economía. De esta forma se identificó que una variable clave en el comportamiento del electorado era el estado civil: si los votantes tenían hijos en el hogar, los padres eran entre 10 y 15 puntos porcentuales más propensos a inclinarse por el voto republicano.

En el año 2008, Penn estableció finalmente 82 nichos de votantes, en su libro *Microtrends: The Small Forces Behind Tomorrow's Big Changes* (2007). De acuerdo con Gutiérrez Rubí las elecciones norteamericanas en el año de 2012, fue cuando el término “Big Data político” cobró actualidad y relevancia, (Gutiérrez Rubí, 2016). Este momento fue dado cuando “un equipo de informáticos, sociólogos, politólogos y matemáticos, denominado el equipo de La Cueva, se dedicó a unificar todas las bases de datos que habían empleado los diversos equipos de campaña demócrata en 2008 y, comenzó a sumar una nueva capa de datos, incluyendo ahora el Big Data Social, con datos conseguidos en redes sociales. El resultado fue: una gran base de datos que incluía listas de donantes, datos recogidos por encuestadores y voluntarios, información y análisis semántico de redes sociales”, entre otros aspectos (Gutiérrez Rubí, 2016: 90). Todo en un solo lugar, para poder analizar la conducta de los votantes y cómo poder impactar de mejor forma con un sólo mensaje. Esto marcó un parteaguas al poder identificar quién dice qué, cuándo lo dice, cómo lo dice, por qué lo dice, qué sienten, y cómo se perciben así mismos frente a un determinado hecho social.

En 2012, el equipo de Barack Obama creó una aplicación en *Facebook*, denominada *Are you in*, que permitía a sus partidarios declararle su apoyo o hacer llegar a sus amigos y amigas un mensaje para que se unieran a él. Pero sin saberlo, bajo esta estructura sucedía algo más: al dar permiso, se sumaban a la base de datos de Obama, aspectos básicos (nombre, fotografía y género), pero también la lista de amigos y cualquier otra información que hubieran dejado abierta para compartir en *Facebook*, incluyendo los mensajes en el muro o sus propios gustos. En ese momento, el equipo de campaña de Obama, podía, por

ejemplo, enviar un mensaje a cada persona que se autoidentificara como demócrata, mujer, casada y mayor de 35 años.

En esta misma línea, el autor menciona que en junio de 2015, con la aprobación del matrimonio igualitario en Estados Unidos, *Facebook* desarrolló una App que pintaba las fotos de perfil con la bandera multicolor, el uso de esta App, se mostraba el respaldo ciudadano a la ley, con la intención de pronosticar el voto progresista o demócrata. Ello permitió identificar algunos patrones de comportamiento sobre el tema.

Desde esta perspectiva, el impulso y desarrollo del Big Data abre diversas oportunidades, pero el análisis de datos masivos para la toma de decisiones inteligentes pone de manifiesto la necesidad de generar capacidades competitivas en cada uno de los países que requieren superar la brecha digital y obtener los beneficios que supone el desarrollo del Big Data, lo cual puede ser aplicado al desarrollo, implementación y seguimiento de aspectos relacionados con temas de opinión pública y políticas públicas y en prácticamente cualquier ámbito o campo de conocimiento.

1.4 Supuestos teóricos del uso del Big Data

1.4.1 Consideraciones teóricas

Como se advierte, Big Data surge “como una nueva fase del paradigma intensivo en información y comunicación que abarca no solo su dimensión tecnológica, sino también una dimensión social, económica, política y cultural” (Malvicino y Yogue, 2015: 2). Desde esta perspectiva, el abordaje de Big Data requiere no solo de insumos técnicos, sino también una instrumentación integral que sea multidisciplinaria e interdisciplinaria. Al igual que con el surgimiento de las tecnologías de la información y la comunicación (TIC´s) se está gestando también en torno al Big Data, una discusión acerca de si nos encontramos frente a un cambio de paradigma tecnológico, una nueva revolución industrial y tecnológica o, si el Big Data no ha hecho más que intensificar la competencia en mercados ya existentes, sin mostrar un impacto real en las mejoras de la productividad en general (Malvicino y Yogue, 2015).

De acuerdo con McKinsey, nos encontramos frente a un punto de inflexión: “en la cúspide de una tremenda ola de innovación, productividad y crecimiento, así como también de nuevas formas de competencia y apropiación de valor” (2011: 2). Si bien, los datos siempre han sido parte del impacto de las nuevas tecnologías, los cambios que Big Data provoca están influyendo en el mercado, lo cual genera nuevas oportunidades de negocios y mejoras en la toma de decisiones a partir del conocimiento de dicha información.

Nos encontramos frente a una revolución industrial de datos precedida desde el comienzo del nuevo siglo por las innovaciones tecnológicas y los dispositivos digitales (McKinsey, 2011; Un Global Pulse, 2012; Naciones Unidas para Impulsar Innovaciones en la Velocidad de Recolección y Análisis de Datos, 2009). Este período se caracteriza por un incremento exponencial en la cantidad y diversidad de datos digitales disponibles en tiempo real, producto de un mayor uso de equipos tecnológicos de mayor capacidad en la vida diaria, lo que permite alcanzar un mayor conocimiento del comportamiento humano (Un Global Pulse, 2012).

Según el informe *Big data. Avances recientes a nivel internacional y perspectivas para el desarrollo local*, las estimaciones que utilizan la cantidad de datos digitales disponibles a nivel mundial pasaron de 150 exabytes en 2005 a 1.200 en 2010, esto es, un crecimiento de ocho veces, y se proyecta un crecimiento del 40% por año, es decir, cuarenta veces el crecimiento de la población mundial. Por lo tanto, no solo se estaría duplicando la cantidad de información digital acumulada cada veinte meses, sino que también dichos datos se estarían haciendo cada vez más jóvenes en términos relativos, al haber una mayor disponibilidad de ello en tiempo real.

En la literatura se menciona que Big Data constituye un fenómeno global que puede llegar a tener un impacto económico real y potencial, que beneficie tanto al sector público como al privado en el aumento de la productividad, la competitividad sectorial y la calidad de vida de la ciudadanía. Sin embargo, se pueden identificar al menos tres paradigmas de pensamiento dentro de la literatura acerca del desarrollo que Big Data está teniendo:

- a.** Los que están a favor de la dinámica que impone este tipo de análisis, ya que lo consideran un cambio positivo para la sociedad. Ello debido a las mejoras en la productividad y a una mayor competitividad; en las respuestas a problemas sociales en áreas como: salud, economía, violencia y la creación de nuevas oportunidades de negocios, entre otros aspectos. (McKinsey, 2011; Brynjolfsson, Hitt y Kim, 2011; Un Global Pulse, 2012; Chen y Zhang, 2014),

b. En el segundo grupo, estaría integrado por los escépticos/pesimistas, que consideran que Big Data es sólo algo de moda que pronto pasará, lo cual no representa en estricto sentido las virtudes y ventajas que se supone se le atribuyen al Big Data, toda vez que se restringen las libertades individuales y la privacidad de los sujetos (Campanario, 2013; The New York Times, 2013; Gordon, 2014).

c. En el tercer grupo se encuentra el sector de los críticos, quienes ven oportunidades, aunque no dejan de cuestionar en alguna medida los alcances del proceso, así como, el impacto social del mismo, ello más allá de los indicadores y resultados socioeconómicos (Pérez y Hilbert, 2010; Manovich, 2012; Boyd y Crawford, 2012).

De acuerdo con Chen y Zhang (2014) el fenómeno Big Data, da origen a un cuarto paradigma científico, debido a los cambios crecientes que las aplicaciones intensivas de datos están provocando en la ciencia. En este nuevo paradigma de acuerdo con Facundo Malvicino y Gabriel Yogue, la viabilidad de este proceso está sujeta a los cambios comunicacionales e informacionales que promueve la digitalización, la convergencia tecnológica y el Internet a distintos niveles: a nivel operativo, se trata del acceso remoto y en tiempo real, se tiene que considerar la información para la toma de decisiones, lo cual como mencionan los autores, le permite un desanclaje territorial y la ausencia de copresencia de los actores sociales en tiempo real. Estos cambios estructurales y territoriales tienen que ver con una mayor y mejor accesibilidad y, conectividad a nivel global (Forte y otros, 2012).

1.4.2 Consideraciones metodológicas y epistemológicas

Desde la perspectiva metodológica y epistemológica, el Big Data plantea aquella discusión que desde la teoría sociológica abarca el desarrollo de las TIC´s, el cual tiene que ver con el rol que éstas desarrollan dentro del orden social, más allá de la reflexión sobre los alcances tecnológicos que puedan tener. Esto significa considerar a las nuevas tecnologías de la información no sólo como subproducto de las relaciones sociales o como un factor tecnológico puramente, sino también es necesario considerar su carácter sociológico y las funciones comunicativas que tienen en la actualidad. A partir de lo anterior, se advierten al menos tres aspectos en el uso del Big Data que constituyen una estructura comunicativa históricamente novedosa ligada a tres procesos: digitalización –o informatización de contenidos–, convergencia tecnológica e internet. De esta manera, la digitalización supone la convergencia de distintos soportes técnicos que a su vez implican el robustecimiento y mundialización de la conectividad de internet, lo cual ofrece infraestructura y propagación de la información

dentro de la red. La digitalización se constituye en un medio de comunicación y en un umbral histórico a partir de su masificación, al transformarse en unidades trasladables y transmisibles de información mediante técnicas de binarización (Malvicino y Yogue, 2015).

El segundo, la convergencia tecnológica, se constituye dentro de esta digitalización y se reducen los grados de libertad del desarrollo de técnicas, lo que tiende a un acoplamiento tecnológico que forma una simplificación funcional o, en otros términos, una dependencia en la trayectoria de este desarrollo (Forte *et al* 2012).

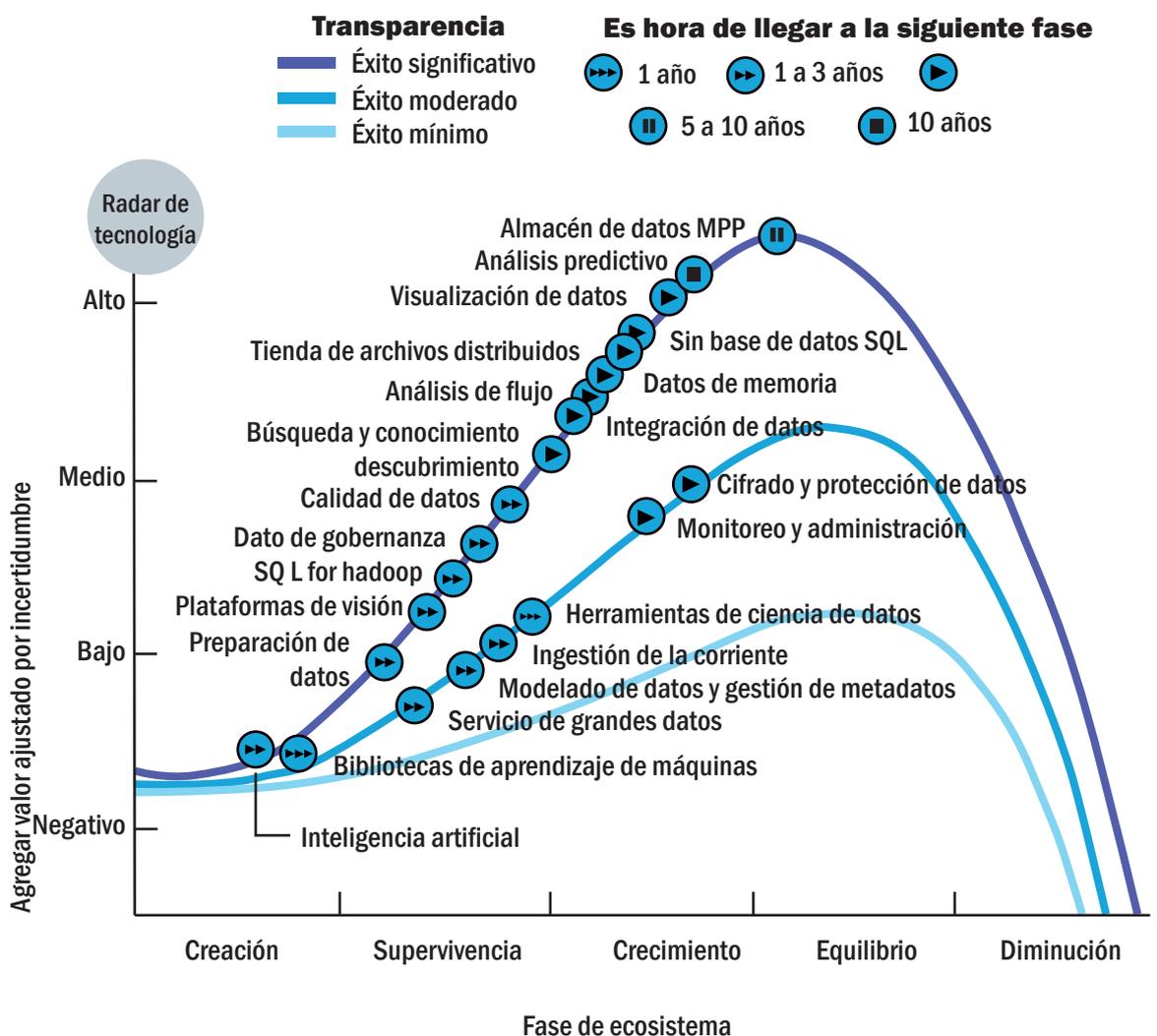
Esta convergencia puede suponer un fenómeno comunicativo porque se desarrolla en un medio de comunicación digital, pero al mismo tiempo condiciona la comunicación por la simplificación funcional, lo cual genera una dependencia mutua y el aumento del riesgo de una disrupción digital y/o comunicacional, generalmente mencionados como brecha digital. Finalmente, el internet ofrece la infraestructura telecomunicativa y su propagación a nivel mundial. Es por ello que la emergencia de un medio de comunicación digital transforma la comunicación y las estructuras de propagación de la misma.

Concretamente en el ámbito epistemológico esta nueva forma de análisis tiene que ver con los cambios en el paradigma del método científico: sobre los límites y la relación entre la teoría, el método, el objeto de estudio y el dato. En este sentido, Manovich (2012) se plantea cómo Big Data puede cambiar el método de investigación en las ciencias sociales y las humanidades, ya que en la medida en que el mundo se vuelve más y más digital, las nuevas técnicas serán necesarias para investigar, analizar y entender esta cantidad de datos. Una primera manifestación de estos cambios es para el autor, la división entre los métodos cuantitativos, comúnmente vinculados a las ciencias sociales, y los métodos hermenéuticos generalmente asociados a las humanidades. Es decir, no sería necesario elegir entre el tamaño de la muestra y la profundidad de los datos porque ya estarían dados desde la propia elaboración del dato.

La creencia que grandes cantidades de datos ofrecen por sí mismas una forma de conocimiento más elevada y una mejor aproximación a la verdad es relativo. Ello renueva la discusión sobre objetividad y subjetividad del método científico, particularmente en las ciencias sociales y las humanidades (Manovich, 2012). Toda vez que la cantidad de datos por sí sola no otorga objetividad absoluta, debido a que estos datos se encuentran sujetos a interpretación de quien elabora el conocimiento sobre el objeto observado. De igual forma, cada una de las decisiones de investigación son permeadas por el investigador. Así que estaríamos hablando de que no existe un dato “totalmente neutro”. En otras palabras, la realidad no se explica por sí sola y, por lo tanto, un empirismo radical es difícil de sostener en la actualidad. De esta manera, un dato es siempre producto de una construcción sobre aquello que es observable; es decir, no existe un dato duro absolutamente puro, objetivo y neutro (Forte, 2015).

Por otro lado, en el análisis de Big Data se debe considerar también los errores que pueda haber en el proceso de recabar los datos masivos, la dificultad de interpretación que se genera por fuentes desconocidas, el nivel de representatividad de datos elegidos, su contextualización o descontextualización, entre otros aspectos. De acuerdo con estos enfoques, la clave del nuevo paradigma en el Big Data es el análisis sistemático de datos masivos previamente para una mejor toma de decisiones.

Gráfica 3.
Trayectoria y tiempo de implementación del Big Data



Fuente: Elaboración a partir de José Manuel Benítez Sánchez, Big Data: Algoritmos, tecnología y aplicaciones, Jornadas de Análisis de Datos Masivos en Empresa, Universidad de Granada, 2013, en *Forrester Research*, <https://www.forrester.com/home/0,6092,1-0,FF.html>, 23 de junio de 2018.

Este proceso podría ser entendido como un paso natural en la evolución de la era de la información o de las sociedades de la información –en términos de Bell, Castells y otros autores– hacia las sociedades del conocimiento. Basadas en la infraestructura digital que permite un vasto incremento de la información, el desafío actual consistiría en convertir la información digital en conocimiento para la toma de decisiones inteligente (Hilbert, 2013a). Por lo tanto, si se acepta la posición que plantea estar en presencia de un cambio de paradigma en la forma de generar la información y transformarla en conocimiento, no debemos dejar de mencionar algunos aspectos estructurales de la creación y apropiación de valor a partir de la información. Por un lado, es necesario destacar que mientras haya más personas trabajando de manera sistemática con un determinado conjunto de datos, mejores y más resultados se pueden obtener.

Sin embargo, es posible que la totalidad de las bondades y ventajas del análisis con Big Data no se hayan dejado ver todavía en los países en desarrollo. Ello implica el desafío adicional de cerrar la brecha tecnológica informacional y la capacitación de los recursos humanos que sigue persistiendo en la actualidad. Asimismo, en el documento *Big data. Avances recientes a nivel internacional y perspectivas para el desarrollo local*, se considera que esta nueva etapa tecnológica es una oportunidad, principalmente para los países en desarrollo, de evitar o moderar el impacto de los *shocks* financieros o climáticos y, reducir el impacto de la volatilidad de los mercados, a partir de la disponibilidad de información en tiempo real que se habilite para dar respuestas oportunas con mayor anticipación.

1.5 Los usos del Big Data. Una aproximación nacional e internacional

En la actualidad el Big Data es aplicado y usado principalmente en el sector empresarial en los Estados Unidos, Reino Unido y algunos países de la Unión Europea. Estas empresas han centrado su atención en analizar datos básicamente estructurados. No obstante, con Big Data, se puede realizar análisis de múltiples tipos de datos, incluida una gran variedad de ellos: análisis de texto como transcripciones de las conversaciones de un centro de atención telefónica o incluso alguna imagen es susceptible de ser analizada desde esta perspectiva. Este tipo de análisis incluye también la capacidad para interpretar y comprender los matices del lenguaje, tales como los sentimientos, el *argot* y las intenciones, entre otros aspectos por complejo que parezca (IBM, Institute for Business Value, 2012).

En la actualidad el Big Data es aplicado y usado principalmente en el sector empresarial en los Estados Unidos, Reino Unido y algunos países de la Unión Europea. Estas empresas han centrado su atención en analizar datos básicamente estructurados. No obstante, con Big Data, se puede realizar análisis de múltiples tipos de datos, incluida una gran variedad de ellos: análisis de texto como transcripciones de las conversaciones de un centro de atención telefónica o incluso alguna imagen es susceptible de ser analizada desde esta perspectiva. Este tipo de análisis incluye también la capacidad para interpretar y comprender los matices del lenguaje, tales como los sentimientos, el *argot* y las intenciones, entre otros aspectos por complejo que parezca (IBM, Institute for Business Value, 2012).

Las primeras iniciativas de Big Data, se centran en obtener conocimientos de fuentes de datos internas nuevas o ya existentes. La mayor parte de los esfuerzos de Big Data se han dirigido a extraer y analizar datos internos de las empresas, lo cual sugiere que algunas empresas están siendo pragmáticas al adoptar Big Data y también existe entre ellas, la necesidad de descubrir lo oculto en los sistemas internos de esas empresas. Los datos internos hasta el día de hoy son los datos más desarrollados y mejor entendidos por las empresas. Estos se han recabado, integrado, estructurado y normalizado a lo largo de años de planificación de los recursos empresariales y gestión de datos maestros, *business intelligence*, así como otras actividades relacionadas, (IBM, Institute for Business Value, 2012).

De acuerdo con esta dinámica, las soluciones de Big Data más eficaces identifican primero los requisitos del negocio, así como la finalidad que persigue; a continuación, adaptan la infraestructura, las fuentes de datos y la analítica a fin de respaldar las oportunidades del negocio. Estas empresas obtienen nuevos conocimientos de fuentes de información internas, tanto de las que ya existían como de las que están disponibles desde hace poco y definen la estrategia tecnológica para realizar un análisis de Big Data y, posteriormente, actualizan progresivamente en proveer infraestructuras a lo largo del tiempo (IBM, Institute for Business Value, 2012).

En la encuesta *Big Data Work*, se señala que la mayor parte de las empresas se encuentra actualmente en las primeras fases del desarrollo de Big Data, básicamente están centradas en comprender los conceptos que se comienzan a utilizar y también, están definiendo la ruta que habrán de tomar para conocer de mejor forma a sus consumidores y por otro, algunas otras empresas están desarrollando o usando las primeras aplicaciones y soluciones sobre su campo de acción (IBM, Institute for Business Value, 2012).²

² Está basado en el Big Data @ Work Survey, llevado a cabo por IBM a mediados de 2012 con 1.144 profesionales procedentes de 95 países y 26 sectores.

De acuerdo con el informe *El uso de Big Data en el mundo real*, las empresas consideran que Big Data proporciona la capacidad para comprender y predecir mejor los comportamientos de los clientes y, al hacerlo, mejorar su experiencia. Transacciones, interacciones multicanal, redes sociales, datos a través de fuentes como las tarjetas de fidelidad y otra información relacionada con los clientes, han aumentado la capacidad de las empresas para crear una imagen completa de las preferencias y demandas de los clientes. A partir de ello, empresas de todo tipo encuentran nuevas formas de interactuar con sus clientes actuales y futuros. Este principio es aplicable al comercio minorista, pero también a las telecomunicaciones, la sanidad, al gobierno, la banca y las finanzas y, al sector de productos al consumidor, donde usuarios finales y ciudadanos están involucrados en interacciones y negociaciones constantes (IBM, Institute for Business Value, 2012).

1.5.1 Aplicaciones del Big Data en el gobierno

Algunas aplicaciones del Big Data en asuntos de gobierno y en el diseño de políticas preventivas son aquellas realizadas en Reino Unido y en los Estados Unidos en el ámbito de la seguridad y las labores policiales: planificación del trabajo policial, la prevención del delito e investigación y, la persecución de presuntos delincuentes. Estas aplicaciones se han realizado de manera pragmática y en diferentes escalas o alcances. En términos de aplicaciones en gran escala, se encuentra el *Modelo de Inteligencia Nacional* (NIM, por sus siglas en inglés) en el cual los Sistemas de Información Geográfica (herramientas de alta intensidad informática), fueron incorporados en el año 2000. Este modelo se creó para lograr la coordinación de responsabilidades, problemas y operaciones en distintos niveles de la jerarquía espacial y en el análisis de actividades policiales.

De igual forma algunos de los sistemas de indicadores del crimen en las Naciones Unidas, la Unión Europea y, en América Latina y el Caribe, han impulsado el *Sistema Regional de Indicadores Estandarizados de Convivencia y Seguridad Ciudadana* -SES, por la Organización de los Estados Americanos, (OEA); ésta última en colaboración con el Banco Interamericano de Desarrollo (BID) (Vilalta Perdomo, 2017).

Por su parte, las Naciones Unidas también cuentan con la Oficina sobre Droga y Crimen, *United Nations office on Drugs and Crimen*, (UNODC), que trabaja y ofrece información estadística sobre corrupción; justicia criminal, reforma carcelaria y prevención del delito; prevención, tratamiento, atención y tráfico de drogas; trata de personas y tráfico ilícito de migrantes; lavado de dinero; y crimen, entre otras. La UNODC contribuye activamente en producir las dos más importantes encuestas: la *Encuesta Internacional de Víctimas del Crimen (International Crime Victims Survey, (ICVS)* y la *Encuesta de Naciones Unidas sobre Tendencias del Crimen y Operaciones de los Sistemas Judiciales, (United Nations Survey of Crime Trends and Operations of Criminal Justice Systems, (CTS)*. La ICVS consiste en una serie de encuestas estandarizadas que miden la experiencia de la ciudadanía con la delincuencia, la policía, la prevención del delito, y el miedo al crimen en un amplio conjunto de países. El financiamiento proviene de organismos internacionales, asimismo la Unión Europea ha participado de forma especial para financiar su aplicación en varios países en vías de desarrollo. Esta información es de utilidad para la detección de grupos de riesgo y la prevención del delito (Vilalta Perdomo, 2017). En el caso de América Latina, se encuentra el Centro de Excelencia para Información Estadística de Gobierno, Seguridad Pública, Victimización y Justicia (INEGI-UNODC), con sede en la ciudad de México. Dicho centro tiene la capacidad, los recursos humanos, la información y el conocimiento práctico para estudiar esta temática desde la perspectiva de Big Data.

Además de la información anterior, la Unión Europea realiza una encuesta de victimización *European Crime and Safety Survey (EU ICS)*, en la mayor parte de los estados miembros, a cargo de diversas instituciones académicas y privadas de prestigio. La encuesta ha tenido desde su creación el propósito de ofrecer estimaciones de victimización que sean viablemente comparativas. A la fecha se han realizado cinco ediciones, con un trabajo de campo a cargo de la compañía Gallup Europa.

En este caso el uso del Big Data, supondría grandes sistemas de información integrados por bases de datos provenientes de estadísticas policiales, judiciales, carcelarias, llamadas de emergencia, redes sociales e internet, información de circuito cerrado de televisión (CCTV), datos biométricos, y otros registros administrativos (por ejemplo, el registro vehicular, información tributaria, seguro social, datos censales, etc.). Considerando lo anterior, lo que hace de lo anterior un sistema de información Big Data, es el volumen de información, variedad y velocidad del análisis.

De acuerdo con Carlos Vilalta Perdomo (2017), ello ha contribuido a que los países miembros cuenten con un diagnóstico más preciso de su situación en materia de información criminal, y desarrollen estrategias para el fortalecimiento de las instituciones locales

productoras y usuarias de esos datos. Además, ha servido para que se instalen nuevas metodologías de trabajo común dirigidas a armonizar y validar datos entre las diversas fuentes de registro administrativo dentro de cada país y también entre los distintos países participantes del proyecto. Estos sistemas de información pueden utilizarse para desarrollar políticas públicas de prevención de delito de carácter social, situacional o ambas. Además, pueden utilizarse para labores de estrategia policial predictiva e investigación de delitos.

En suma, se puede señalar que el uso del Big Data, aplicando una metodología adecuada permite mejorar el análisis en una diversidad de aspectos y ámbitos. En el ámbito nacional, se puede observar que el análisis con Big Data puede ser sintetizado en los principales rubros donde ha tenido una mayor incidencia:

a) Venta por internet: es el más comúnmente utilizado y uno de los sectores pioneros del Big Data. Las compras por internet y la publicidad que reciben los distintos tipos de consumidores es cada vez más atinada en relación al contenido y el tipo de la información que se difunde, aunque aparentemente no se haya identificado. Asimismo, en el proceso de compra, son pocos los comercios que no ofrecen productos y servicios complementarios tomando como referencia la experiencia de otros clientes.

b) Venta presencial: análisis de los patrones de compra dentro de la tienda y por cliente. En el caso del cliente, el uso de las tarjetas de fidelización es clave. Este tipo de información ayuda a colocar los productos para maximizar su venta, ofrecer descuentos y productos a los clientes apropiados e identificar compras correlacionadas, entre otros aspectos. De esta manera, cuanto más se conoce cómo compramos, más se puede adaptar la tienda para maximizar sus ventas.

c) El Sector bancario: el Big Data se utiliza en este ámbito para hacer análisis de riesgos en general y en la concesión de préstamos en particular, lucha contra el fraude, personalización de ofertas a clientes, la captación de clientes externos ayudándoles a utilizar su información financiera o a crear servicios con un valor añadido.

d) Industria petrolera: ha sido una de las industrias pioneras en el uso del Big Data, el análisis de la información sísmica y otros datos geológicos ha permitido perforar en los lugares más productivos. En la producción, la monitorización continua mediante múltiples sensores permite maximizar el tiempo de funcionamiento, así como mejorar la seguridad de trabajadores y del entorno.

e) Situaciones de emergencia: Herramienta para la gestión de algunas situaciones de catástrofes como huracanes, sismos, sequías y políticas de desarrollo humano y social.

Se debe tener claro que, si bien el análisis con Big Data crea una oportunidad para que las empresas puedan obtener una ventaja competitiva en el actual mercado digitalizado, permite a su vez transformar la forma en la que interactúan con sus clientes y les prestan servicio. Es probable que no todas las organizaciones adoptarán el mismo enfoque con respecto al desarrollo y la creación de sus capacidades con Big data. Sin embargo, en todos los sectores existe la posibilidad de utilizar las nuevas tecnologías y analíticas de Big Data para mejorar la toma de decisiones y el rendimiento: salud, servicios sociales, educación, mundo académico, etc. En este sentido, el análisis legislativo se puede impulsar con el Big Data como una nueva herramienta para ofrecer información, pero, también, para captar datos de los ciudadanos y conocer su opinión en torno a determinados temas controversiales sobre los cuales se está legislando o se podría legislar instrumentando la toma de decisiones de manera más eficiente.

1.6 Normas y principales obligaciones legales en materia de privacidad en el uso del Big Data

El marco jurídico aplicable a los proyectos de Big Data se compone por un conjunto de normas que lo regulan, aunque ninguna es específica para este tipo de tratamiento. De acuerdo con el *Código de buenas prácticas en la protección de datos para proyectos Big data*, actualmente la legislación a aplicar es la Ley 15/1999 de protección de datos de carácter personal (LOPD) y su reglamento de desarrollo, *Reglamento General de Protección de Datos* (RGPD), (EUGDPR, 2015). Éste se aprobó el 18 de abril de 2016 y se aplicará de manera internacional a partir del 25 de mayo de 2018.

En este mismo documento se menciona que el RGPD, se aplica al tratamiento de datos personales, entendiéndose estos como cualquier información concerniente a personas físicas identificadas o identificables. La identificabilidad, que supone la aplicación de la normativa, se refiere a que una persona pueda ser identificada por un dato o por la combinación de información de diversas fuentes. El análisis de la identificabilidad debe basarse en dos criterios: a) el de la razonabilidad en la disponibilidad de los medios (técnicos, humanos y fuentes de datos) y, b) en la proporcionalidad de los esfuerzos para poder identificar directa o indirectamente a la persona física por parte del responsable del tratamiento o por cualquier otra persona.

Ello se puede obviar cuando no sea posible la identificación de los individuos, o esta requiera de esfuerzos desproporcionados para la identificación de las personas. Además, hay otros tipos de datos tratados que son también analizados con Big Data y no necesaria-

mente son datos personales: tráfico automovilístico, la contaminación o el clima, por lo cual tampoco aplica el dicho reglamento.

En este sentido, si los datos son anónimos, no se considerarán datos de carácter personal, por lo que la normativa tampoco será de aplicación. En este sentido, el anonimato supone que no será posible identificar a la persona con datos o con información de diversas fuentes teniendo en cuenta todos los medios que puedan ser razonablemente utilizados para su identificación (Sáiz y otros, 2015). En contraste, algunos otros casos sí requerirán de manera necesaria el tratamiento de información personal, por ejemplo: redes sociales y transacciones bancarias, entre otros datos donde se pueda ubicar a los sujetos. De acuerdo con ello, cada vez que se pueda identificar a los individuos, se debe considerar la aplicación de la normativa de protección de datos (Sáiz y otros, 2015).

Desde esta perspectiva, además de tener el consentimiento del interesado para el tratamiento de sus datos personales, el artículo 6 del RGPD también establece otros requerimientos: como el tratamiento de datos necesarios para la ejecución de un contrato, cuando haga referencia al cumplimiento de una obligación legal, cuando el tratamiento de los datos tenga por finalidad proteger un interés vital del interesado, o cuando el tratamiento sea necesario para el cumplimiento de una misión realizada en interés público o para la satisfacción del interés legítimo perseguido por el responsable del tratamiento o por un tercero, siempre que no se vulneren los derechos y libertades fundamentales del interesado (Sáiz y otros, 2015), será necesario que la obtención del dato sea libre, explícito y que pueda ser revocado (artículo 7 del RGPD).

En este mismo informe se precisa que en estos casos, se habrá que determinar cuándo el tratamiento es necesario para el cumplimiento de este contrato para la prestación de un servicio, y bien, cuando los datos resultantes de ese servicio se utilizarán para otras finalidades distintas. En algunas otras situaciones el tratamiento que se le pueden dar a los datos es la satisfacción de un interés público o en el ejercicio de poderes públicos. Sin embargo, generalmente esto se lleva a cabo por entidades públicas, aunque en algunos casos la atención de intereses públicos relevantes pudiera asumirse por entidades privadas.

No obstante, el *Reglamento General de Protección de Datos*, precisa que tanto el interés público como los poderes públicos que se ejerzan tienen que estar establecidos en el Derecho de la Unión o de los Estados Miembros. Asimismo, el Reglamento prevé que, para aspectos relacionados con el interés público como la investigación científica, las relacionadas con la atención sanitaria o social, o las relativas a salud pública, los tratamien-

tos serán posibles en las condiciones que determine la legislación europea o nacional, en este caso México. Las normas correspondientes establecerán, además, las garantías necesarias para la protección de los derechos y libertades de los interesados.

Otra de las posibilidades que ofrece la normativa es la que se centra en la satisfacción del interés legítimo perseguido por el responsable del tratamiento, siempre que no se vulneren los derechos y libertades fundamentales del interesado. En estas situaciones será necesario analizar el balance entre el interés legítimo y los derechos y libertades (Sáiz y otros, 2015).

Algunas otras aplicaciones del Big Data tienen una finalidad estadística, cuyo objetivo es obtener datos agregados, que en muchos casos permitan tomar decisiones públicas o de negocio. Esta aplicación supone que el resultado del tratamiento con fines estadísticos no sean datos personales, sino datos agregados, y que este resultado no se utilice para respaldar medidas o decisiones relativas a personas físicas concretas. El RGPD establece que los fines estadísticos no se considerarán incompatibles con los fines iniciales. Asimismo, el tratamiento de los datos implica que aseguren que se aplican medidas técnicas y organizativas para garantizar que no se puede identificar a los interesados.

Esta normativa establece garantías a los titulares para la protección de sus datos, entre los que figura el principio de finalidad, que debe ser determinada, explícita y legítima, para la que se hayan obtenido los datos. Este principio de finalidad puede suponer una de las barreras para los proyectos de Big Data, dado que no siempre se conoce desde el comienzo el alcance del proyecto.

En ese sentido, es necesario precisar que los datos personales no podrán usarse para finalidades incompatibles con aquellas para las que los datos hubieran sido recogidos, lo que no significa que no puedan utilizarse para finalidades diferentes para las que se recogieron (Sáiz y otros, 2015), sino más bien que éstas no deben ser incompatibles. A este respecto el Grupo de Trabajo del Artículo 29, específica en su Dictamen que el principio de finalidad antes mencionado se refiere a lo siguiente:

- Debe existir una relación entre la finalidad original y la finalidad o finalidades ulteriores.
- El tratamiento ulterior debe encontrarse dentro de las expectativas razonables del interesado.

- Debe tenerse en cuenta la naturaleza de los datos objeto de tratamiento y la sensibilidad de los mismos.
- Debe considerarse el impacto que este tratamiento va a tener en los interesados.
- Deben considerarse las medidas de protección que el responsable del tratamiento establece, en particular las medidas técnicas y organizativas: encriptación, seudonimización, separación funcional, transparencia, oposición al tratamiento (Sáiz y otros, 2015: 11).

El artículo 6.4 del RGPD, ha incorporado a la norma estos criterios, haciendo mención expresa al cifrado y la seudonimización dentro de las garantías adecuadas, con lo que serán de aplicación directa al análisis de compatibilidad de todos aquellos tratamientos posteriores que no estén basados en el consentimiento del interesado y, por lo tanto, deberán tenerse en cuenta en la evaluación de impacto que realice el responsable sobre dichos tratamientos.

Otros principios a tener en cuenta son los principios de minimización y conservación de los datos. Éste principio de minimización constituye uno de los elementos más importantes de la normativa de protección de datos aplicables a este entorno. Este principio va intrínsecamente unido al de conservación de los datos por el tiempo que sea necesario para la finalidad que hubieran sido recogidos, (Sáiz y otros, 2015). Este principio es especialmente relevante en el análisis con Big Data, ya que las diferentes fuentes a la que se tiene acceso pueden dar lugar a una recogida masiva de datos que no sean pertinentes para la finalidad del tratamiento. Tampoco deberían conservarse por más tiempo del que sea necesario, periodo tras el cual deberían eliminarse.

Esta normativa establece que se deberán adoptar las medidas técnicas y organizativas apropiadas para garantizar un nivel de seguridad adecuado al riesgo. Para evaluar dicha adecuación, se tendrán en cuenta en particular los riesgos asociados a destrucción, pérdida o alteración, o a la comunicación o acceso no autorizados. Por otra parte, la legislación establece normas específicas para aquellos casos en los que los datos son tratados por diferentes actores. En ese sentido, el RGPD establece que el responsable, es la “persona física o jurídica, autoridad pública, servicio u otro organismo, que determine los fines y medios del tratamiento, y asimismo define al encargado como persona física o jurídica autoridad pública, servicio u otro organismo que trate datos personales por cuenta del responsable del tratamiento” (Sáiz y otros, 2015: 13).

De acuerdo con este reglamento la protección de datos de carácter personal no puede entenderse sin ofrecer a los interesados un conjunto de derechos que les garanticen el control del uso que se hace de sus datos personales por parte de los responsables y encargados de esos tratamientos.

Finalmente, se debe decir que todo interesado tendrá derecho a no ser objeto de una decisión basada únicamente en el tratamiento automatizado, incluida la elaboración de perfiles que produzca efectos jurídicos en él o le afecte significativamente de modo similar. Ello supondrá que el interesado podrá oponerse a dicho tratamiento, pero siempre y cuando: a) la decisión “automatizada” o basada en *profiling* produzca efectos jurídicos que conciernan al interesado (un banco que decide no conceder un préstamo o una hipoteca basándose en determinados perfiles de riesgo) o b) la decisión “automatizada” o basada en *profiling* le afecte significativamente.

Tabla 3.
Conjunto de derechos sobre datos personales por parte de los responsables y encargados de esos tratamientos

- i. Derecho de acceso:** Conocer qué tipo de información personal se conserva y obtener una copia de dicha información.
- ii. Derecho de ratificación:** Solicitar que se actualice o corrija tu información.
- iii. Derecho a la portabilidad de los datos:** Recibir una copia de la información que se te proporcionó de acuerdo con el contrato, de forma que puedas proporcionar dicha información a otra organización.
- iv. Derecho de uso restringido:** Solicitar que se deje de usar tu información personal en ciertos casos, como cuando consideras que tu información personal es incorrecta o se está usando de forma ilegal.
- v. Derecho a objetar:** Objetar el uso de tu información (cuando una parte la está procesando de acuerdo con un interés legítimo) y hacer que eliminen tu información personal.
- vi. Derecho de supresión:** En ciertas circunstancias, también puedes hacer que tu información personal sea suprimida.

Fuente: Elaboración propia, a partir del *Reglamento General de Protección de Datos*, (RGPD), 2016.

2. Principales usos en la investigación parlamentaria

2.1 Experiencias de políticas públicas y Congresos que están utilizando Big Data

Existe una gran oportunidad para crear un sendero de desarrollo que permita reducir las brechas de productividad con los países desarrollados, a partir del diseño de políticas efectivas orientadas a sectores de alto contenido tecnológico e informacional, ya que a la fecha este desarrollo prácticamente se ha dado en los Estados Unidos, Reino Unido y algunos países Unión Europea, en tanto que en América Latina se distingue, aunque de manera pionera, países como Chile, Argentina, Cuba y México. Sin embargo, como se observa para la implementación del Big Data, se requiere superar importantes desafíos para los países en desarrollo.

Frente a este desafío Malvicino y Yogue (2015), plantean las siguientes preguntas en relación a la implementación del Big Data: *¿se encuentra el país en condiciones de aprovechar los beneficios que promete esta nueva fase de sociedades basadas en el conocimiento?* y, si el análisis de grandes volúmenes de datos requiere de una base tecnológico comunicacional, *¿está el país apto para dar respuesta a esta exigencia?*, *¿Cómo impacta el escaso desarrollo local de infraestructura de alta complejidad para el desarrollo de Big Data?*

Estas son algunas preguntas que se deben resolver frente a la posibilidad de impulsar el análisis de Big Data. Sin embargo, aun cuando esto pueda redundar en una mejora en la toma de decisiones en áreas clave como la salud, el empleo, la productividad, la seguridad y el manejo de los recursos naturales, etc. Es probable que la privacidad y los aspectos técnicos y tecnológicos se encuentran débiles en los países en desarrollo por factores, como el atraso tecnológico, la falta de infraestructura y la carencia de personas capacitadas.

Considerando lo anterior, el paradigma de Big Data se desarrolla mediante un proceso de difusión lento y desigual, comprometido por la falta de infraestructura, de capital humano y la ausencia de disponibilidad de recursos económicos e institucionales en los países en desarrollo. Esto abre una nueva brecha digital basada en el análisis de datos para la toma de decisiones inteligentes. En consecuencia, la mejora en el desarrollo económico y social es muy probable que no se dé automáticamente, sino que requiere elaborar estrategias y políticas públicas específicas para tal objetivo acompañado de un marco regulatorio deseable para tal fin.

De igual forma, es necesario dimensionar que el sector público es una gran fuente de datos en rubros tales como: la seguridad social, el sistema legal, los programas sociales, salud, y la educación. Sin embargo, estas fuentes están subutilizadas en los países de América Latina. Se debe indicar que la materialización de los potenciales beneficios para los países en desarrollo requiere la elaboración de políticas activas y específicas que tengan en cuenta la generación y apropiación de rentas informacionales, el manejo de la privacidad en cuanto a los datos personales, el desarrollo de capacidades para la creación de valor y la difusión de la información y el conocimiento para contribuir a la reducción de las desigualdades socioeconómicas.

Este contexto de transformación digital no ha sido ajeno a los sistemas políticos, lo cual supone avances y grandes desafíos para las democracias. Uno de los principales elementos para la democracia es la participación ciudadana que se constituye en una herramienta eficaz para conocer la opinión de la ciudadanía sobre temas claves a través del análisis del Big Data, y con ello lograr que un mayor número de personas se involucren y tengan un mayor conocimiento sobre el quehacer y la situación nacional.

En este marco, se suele hablar de grandes datos (Big Data), en su justa vinculación con los datos abiertos (Open Data), donde la sociedad tiene un papel fundamental como el consultor; y el gobierno, como poseedor de los recursos necesarios para recoger grandes cantidades de datos, así mismo, se visualiza como principal proveedor de información, (Costales Ferrer y Amoroso Fernández, 2016).

Uno de los principales elementos que ha traído como consecuencia el surgimiento y la aplicación de los grandes datos (Big Data) es la transformación de los gobiernos, los cuales han estado condicionados a cambiar la manera de presentarse ante el imparable desarrollo global. Según David Newman “si bien el Big Data hace que las organizaciones sean más inteligentes, el Open Data es lo que conseguirá que realmente se obtengan más beneficios en el entorno altamente competitivo en el que nos encontramos”, (Newman, 2014: 3).

De acuerdo con el reporte: *Gobierno en internet con una mirada al Big Data*, la apertura de los datos al público es el principal elemento para que el ciudadano se formule interrogantes y criterios de carácter innovador y transformador. Es importante desde esta perspectiva que los gobiernos desarrollen una eficiente política de apertura de datos. Como segundo aspecto deben propiciar la colaboración entre los organismos y el ciudadano, sin dejar de mencionar el eje central que es la incorporación de la tecnología como canal de comunicación; de esta manera, el Estado debe centrarse en promover la participación de los ciudadanos en los procesos de toma de decisiones y diseño de políticas públicas, De igual forma, en este reporte se menciona que es fundamental ubicar el factor tecnológico como

punto medio entre ambas partes con la finalidad de agilizar la gestión de la documentación, la información y los datos generados por cualquiera de las partes, así como viabilizar los flujos de información, el acceso y el intercambio continuo (Costales Ferrer y Amoroso Fernández, 2016).

En la democracia es fundamental realizar dicho proceso mediante la implementación de la herramienta tecnológica Big Data, por tanto, abrir los datos gubernamentales a la ciudadanía, implica gestionar gran cantidad de datos para favorecer la toma de decisiones y facilitar la colaboración y la participación de la sociedad en dicho proceso novedoso. El dar acceso a información del gobierno, abre el paso para que la sociedad en su mayoría, sea participe de las decisiones, ya que generalmente las políticas públicas son comunicadas en sitios estáticos y cuando ya están tomadas.

De igual forma, es necesario que la Administración Pública asuma la aplicación de la gestión y el análisis de gran cantidad de datos para obtener el conocimiento sobre las tendencias y comportamientos, con el objetivo de elaborar las políticas sociales teniendo como eje las necesidades y demandas de la propia ciudadanía. Dicha información se constituye como un bien público y puede convertirse en insumo para diversas actividades y proyectos y, con ello, obtener valor agregado a la información que se recopila. Esto debe lograr que la ciudadanía tenga acceso a los datos en formatos abiertos para lograr aumentar la credibilidad en las instituciones, promoviendo la transparencia y la participación activa del ciudadano (Costales Ferrer y Amoroso Fernández, 2016).

La ventaja principal que ofrece la colaboración entre el ciudadano y la administración pública, se encuentra en lograr un proceso de formulación de políticas públicas de calidad, debido a que la apertura de la información permite que los ciudadanos conozcan los procesos de toma de decisiones, además de las causas y los factores que influyen en una resolución; de esta manera se integra a la ciudadanía durante el proceso y no solo en la resolución final (Naser, 2012).

De igual forma, se pueden extraer patrones de comportamiento al detectar el tipo de funcionamiento o no de servicios públicos, identificar tendencias en la sociedad e integrar resultados obtenidos en sistemas operacionales. La búsqueda, selección y análisis de información relevante puede mejorar también la productividad del país, ya que no se han logrado niveles de eficiencia deseados que garanticen la elevación de los niveles de vida de los ciudadanos.

Lo más acertado es que la gestión pública eficaz y eficiente debe basar sus decisiones en datos objetivos y de calidad. Una nueva faceta del acceso a la información es la

tendencia a publicar datos en formatos abiertos (Open data) para que la sociedad pueda acceder a la información y sea capaz de reutilizarla para diferentes fines, como cruzar bases de datos para generar nueva información o potenciar su utilidad al hacerla más comprensible y dinámica (Sandra Elena y Ana Pichón Rivière, 2014). Por ello, esta debe estar incorporada desde el inicio en el proceso de toma de decisiones (Sandra Elena y Ana Pichón Rivière, 2014).

Promover el debate de leyes de datos abiertos constituye una oportunidad para actualizar y mejorar el acceso a la información que ya se encuentra publicada y permiten especificar los nuevos conjuntos de datos y los registros que se deben publicar para el diseño de estudios parlamentarios y políticas públicas acordes a las necesidades de la población. Incluso ésta información se constituye como un bien público y puede ser un insumo para diversas actividades, proyectos y usos que pueden diferir de los que el gobierno previó cuando liberó la información. Ello podría aumentar la credibilidad en las instituciones, ya que esta se encontraría realmente comprometidas con la transparencia y la rendición de cuentas (Tauberer, 2012).

Esta nueva perspectiva implica promover un cambio cultural orientado a la apertura y transparencia del actuar gubernamental (Open Knowledge Foundation, 2012). Ello con la finalidad de que la ciudadanía diseñe en conjunto políticas de acceso a la información y datos abiertos. Uno de los elementos distintivos del gobierno abierto y de los datos abiertos es la participación real de la sociedad en el diseño e implementación de políticas.

2.2 Alcances del Big Data en el proceso legislativo

En la actualidad los datos son utilizados por las grandes corporaciones para mercadotecnia. Principalmente, se observa que las instituciones científicas, la administración pública y el sector empresarial se encuentran desarticulados. En este escenario se debe identificar el rol del sector público en el desarrollo de nuevos sectores de alta tecnología, considerando por supuesto el papel que puede tener Big Data, así como las posibilidades de impulsarlo de manera eficiente en México y, particularmente para el análisis legislativo. Esto constituye un área de oportunidad para generar nuevas pautas para desarrollar las normas legales que expresen con mayor precisión la voluntad de la ciudadanía con base en el análisis de datos a gran escala (Malvicino y Yoguel, 2016).

El contexto global del desarrollo económico social, adopta un esquema tridimensional que se establece entre la tecnología, el cambio social y las políticas públicas estratégicas (Hilbert, 2013), el cual supone el desarrollo de la sociedad a través de la innovación tecnológica. Hilbert considera que Big Data aplicado a la toma de decisiones tiene que ver con la intención de generar un efecto positivo sobre la eficiencia y la productividad a gran escala, presentándose mayores dificultades para los países en desarrollo. Sobre este punto los gobiernos han comenzado de manera gradual a utilizar el Big Data para mejorar la toma de decisiones y la calidad de los productos y servicios públicos, ya que el uso intensivo de datos genera valor público en muchas dimensiones.

En el informe sobre Big data, *Avances recientes a nivel internacional y perspectivas para el desarrollo local*, (2016), se menciona que las ganancias potenciales en materia de salud con Big Data ascenderían a 300 mil millones de dólares anuales en diez años en los Estados Unidos, lo cual se deriva como resultado de las reducciones en los gastos de atención de la salud, logrado a partir de la reducción de costos por tratamientos médicos poco efectivos, la reducción de efectos adversos provocados por tratamientos, reducción de errores médicos y criterios sesgados hacia intervenciones poco efectivas, la detección de patrones en las patologías y algunos resultados vinculados al efecto adverso de los medicamentos, así como el monitoreo remoto de los pacientes y sistemas personalizados (Malvicino y Yoguel, 2016). Sin embargo, en dicho informe menciona que los principales desafíos de ello son: la digitalización de gran parte de la documentación, las historias clínicas, imágenes, la privacidad de los datos del paciente, la información que pueden tener las empresas y organismos públicos, y el personal capacitado, entre otros aspectos, que dificultan la estandarización de la información.

De acuerdo con esta investigación, el uso de Big Data puede mejorar el monitoreo de la salud pública y la capacidad de respuesta a partir de una base de datos de pacientes y tratamientos que permita una coordinada y temprana detección de enfermedades infecciosas, así como también reducir los tiempos de investigación. De igual forma, el Big Data se ha utilizado para predecir enfermedades, el avance en la digitalización de historias clínicas e imágenes y ofrece una oportunidad para lograr reducciones de costos, detección de irregularidades y eficiencia en los tratamientos médicos. Es importante considerar que el sector salud constituye un amplio espectro que involucra no solo a los prestadores médicos, a la industria farmacéutica y de insumos de salud, sino también sindicatos, programas sociales e incluso políticas laborales, entre otros aspectos, (Malvicino y Yoguel, 2016).

Por su parte Hilbert (2013a) aborda la discusión de la difusión de la información y la competitividad a partir de los casos registrados en la explotación de los recursos naturales. La evidencia sugiere que una mayor disponibilidad de datos, que se recogen en diversos ám-

bitos productivos como el petróleo, el gas y la industria minera, aumenta la transparencia y no afecta la competitividad de la actividad. Otras de las aplicaciones con Big Data son aquellas aplicadas al medio ambiente, la prevención de desastres climáticos y alerta temprana (United Nations University, 2012), cuellos de botella en recursos naturales y aplicaciones concretas de seguimiento del clima y el comportamiento social para comercio minorista.

Otras alternativas que ofrece Big Data de acuerdo a *UN Global Pulse* (2012), es el sector financiero y análisis de riesgo financiero sistémico que recogen digitalmente datos para el estudio del medio ambiente mediante la difusión de sensores; por ejemplo, en los cursos de ríos y las emisiones de humo, en incendios de reservas; bioinformática, astronomía y nuevas tecnologías de secuencias. A este respecto *Un Global Pulse* (2012), menciona aplicaciones de una metodología desarrollada por la *International Food Policy Research Institute*, para detectar volatilidades de precios en los alimentos que pongan en riesgo la seguridad alimentaria de determinados países. También destaca un sistema aplicado a la detección de la amplitud de sismos y terremotos o epidemias, basado en mensajes de Twitter.

Sin embargo, un elemento que podría afectar este desarrollo es la falta de profesionales capacitados en las metodologías y tecnologías de análisis de datos, lo cual impide masificar este tipo de herramientas para apoyar la resolución de problemas y la planificación de políticas públicas. Es decir, no basta con solamente tener información, sino que se requiere poder analizarla y utilizarla en los procesos y programas e iniciativas públicas para que efectivamente se constituya en una fuente valiosa de información.

En este marco es importante tener en cuenta también que las aplicaciones de Big Data no se darán de forma automática. Su implementación no solo está sujeta a la disponibilidad de trabajadores calificados o la previa inversión en las TIC's, sino también a las barreras que puedan imponer el contexto institucional, político y cultural particular de los sectores, así como también los intereses económicos y políticos de los actores involucrados en el país. Asimismo, un tema central en la aplicación de políticas exitosas y en análisis legislativo, será el tratamiento de la privacidad de los datos de las personas para que no haya un daño en su integridad y un descrédito de la herramienta que obstaculice un mejor aprovechamiento de sus beneficios (Malvicino y Yoguel, 2016).

A partir de lo anterior, se observa que Big Data genera una nueva posibilidad de realizar la política y lo público, pero ¿cómo podríamos hacer las cosas de un modo diferente? El texto del *Big Data al Data Thinking* de Antoni Gutiérrez-Rubí, (s/f); señala que en este caso habría que establecer algunos reguladores estatales que aseguren que al menos el sector público y que las compañías tengan igual acceso a esa información para la toma de decisiones, (Gutiérrez-Rubí, s/f).

La apertura de los datos públicos y la posibilidad de su reutilización dan lugar al desarrollo de aplicaciones o soluciones por parte de quienes son usuarios de los servicios públicos. Esto es llamado la inteligencia de las multitudes, el *crowd sourcing* que implica un trabajo colaborativo de distintas instancias: gobierno y ciudadanía, ya que no se puede gestionar lo público de manera eficaz y eficiente, sin un esfuerzo colectivo (Gutiérrez-Rubí, s/f). Entonces conviene entender que tenemos por delante una gran responsabilidad, porque ahora tenemos que definir las instituciones que van a definir el futuro de estas convergencias. Como se advierte, la revolución de los datos ha sacudido industrias enteras como el comercio minorista, el transporte y los servicios financieros, y esta tendencia poco a poco comienza a llegar al gobierno (Diplomado en Big Data para Políticas Públicas, 2017).

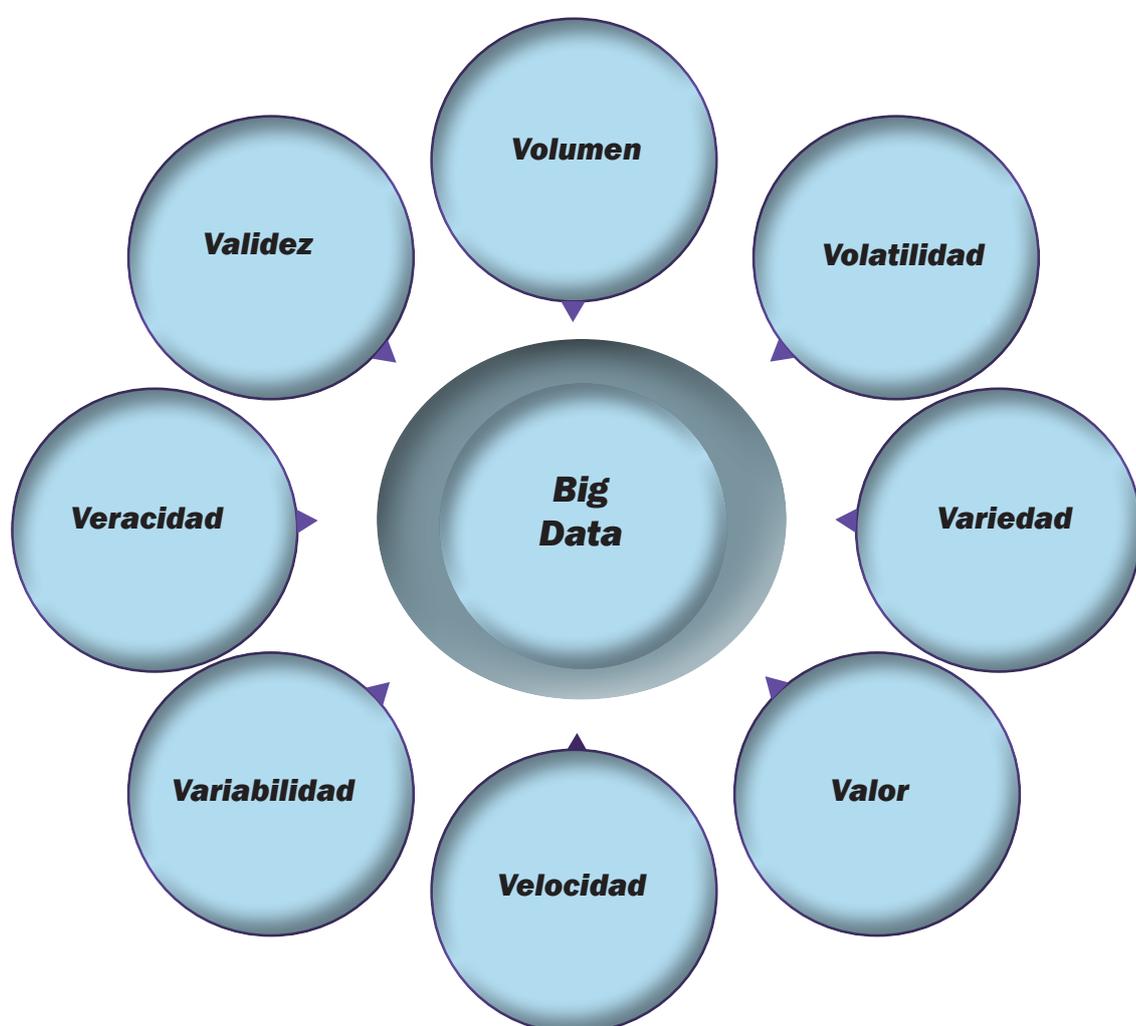
En el estudio *Información para la prevención del delito y la violencia*, se advierte que, en América Latina, se cuenta con indicadores válidos y confiables que muestran cuatro tendencias que requieren una rápida atención en materia de políticas públicas preventivas en el ámbito de la seguridad ciudadana: la violencia homicida, la creciente victimización delictiva, el aumento de la población carcelaria, y el persistente miedo al crimen o sensación de inseguridad (Vilalta, 2017).

En este caso las políticas públicas de prevención de la violencia y la delincuencia requieren indicadores de gestión, procesos, actividades y recursos y el diseño de estrategias que den pautas de acción que puedan ser estratégicas para resolver estos problemas. Estos sistemas de información pueden utilizarse para desarrollar leyes y políticas públicas para la prevención de delito de carácter social, situacional o ambas. Además, pueden utilizarse para labores de estrategia policial predictiva e investigación de delitos (Vilalta, 2017).

Diversos problemas sociales están teniendo una amplia variedad de impactos políticos y sociales en América Latina (Vilalta, 2015). Estos impactos no pueden ser analizados únicamente sobre la base de una sola fuente de información (Lab, 2014) y/o desde una visión espacial, atemporal /o unidimensional del problema. Se requiere información en indicadores útiles desde múltiples perspectivas y con detalle suficiente, para contar así con la capacidad de diagnosticar realidades y percepciones de la ciudadanía y prever tendencias para tomar decisiones acordes con los retos. La información es parte fundamental de las políticas públicas, es su base del conocimiento. El desarrollo de ésta base de conocimiento requiere de distintas acciones: la primera es la sistematización de información que brinde una base empírica a los estudios parlamentarios para el desarrollo de mejores leyes que reflejen el sentir social, así como en la elaboración de las políticas públicas en la materia (Vilalta, 2017).

Los grandes retos de sistemas Big Data en la administración pública y particularmente en el análisis legislativo son la dispersión de fuentes, y el volumen y la variedad de la información disponible. En el uso de información proveniente de fuentes múltiples por ejemplo estadísticas, datos censales georeferenciados sobre una cartografía digital, normalmente no se poseen datos suficientes en los niveles de análisis adecuados que se requieren para realizar diagnósticos para la toma de decisiones.

**Gráfica 4.
Big Data**



Fuente: Elaboración a partir de José Manuel Benítez Sánchez, Big Data: Algoritmos, tecnología y aplicaciones, Jornadas de Análisis de Datos Masivos en Empresa, Universidad de Granada, 2013.

Carlos Javier Vilalta, (2017) sugiere una ruta ideal de trabajo en la creación y consolidación de un sistema de información para fundamentar el análisis legislativo y las propias políticas públicas de prevención en distintos ámbitos, por lo cual menciona que se debe pasar de:

- El dato al indicador y del indicador de gestión, al indicador de impacto.
- De la multitud y diversidad a la unificación y homologación de indicadores.
- De la segmentación y fragmentación locales y nacionales al encadenamiento regional.
- De la compilación a la difusión pública.
- De la solidez estadística a la confianza ciudadana.
- Finalmente, de la consolidación nacional a la regional.

Estos mayores niveles de información promueven la transparencia y la rendición de cuentas, así como la producción, la recopilación y la disponibilidad de datos en un formato que se pueda utilizar y reutilizar por los diversos actores sociales involucrados, lo cual permitiría tomar decisiones y elevar el nivel de análisis legislativo y, a su vez mejorar, el diseño de las políticas públicas basadas en evidencia. En consecuencia, sino se cuenta con información confiable y completa, no será posible identificar con exactitud las causas y consecuencia de los distintos problemas que se encuentran en la agenda pública (Sandra Elena y Ana Pichón Rivière, 2014).

En este escenario surge la necesidad de generar espacios de trabajo conjunto entre el Congreso de la Unión, el Poder ejecutivo, la sociedad civil, empresarios y académicos para diseñar en conjunto políticas de acceso a la información y datos abiertos desde el esquema de Big Data para obtener datos confiables sobre problemas sociales que ya forman parte de la agenda pública o bien, para lograr que problemas hasta el momento dejados de lado sean visibilizados y considerados por los tomadores de decisiones con la finalidad fortalecer las instituciones y hacer un análisis legislativo más sólido.

2.3 Open data y el análisis legislativo

La posibilidad de conocer y analizar las demandas de la población, como insumo para el diseño de las leyes y la implementación las políticas públicas es un nuevo nicho que deberá ser explorado e impulsado desde un punto de vista estratégico. Es necesario elaborar un plan de seguimiento de la opinión pública para recabar información sobre las distintas demandas de los ciudadanos, evaluar el nivel de conocimiento que tienen sobre las políticas gobierno, y generar información cualitativa, cuantitativa y mixta para el desarrollo de propuestas de ley basadas en información integral de calidad, así como para la implementación de las políticas públicas (Diario El Destape, 20 marzo, 2018).

Un medio para impulsar este desarrollo puede ser el Open Data el cual supone un sistema de datos abiertos. Éstos son datos que pueden ser libremente utilizados, reutilizados y redistribuidos por cualquier persona. Es una tendencia que persigue que determinados datos estén disponibles de forma libre para todo el mundo. Una vez obtenido los mismos datos pueden ser usados, copiados, estudiados, modificados y redistribuidos libremente de varias formas. Por otro lado, esta tendencia también debe identificar las principales políticas que necesitan ser creadas o promovidas por los Estados para la construcción de una infraestructura de información que permita hacer efectivo el ejercicio del derecho a saber (Bertoni y Torres, 2014).

Los datos abiertos tienen información no documental como la geográfica, genoma, compuestos químicos, fórmulas matemáticas y científicas, datos médicos y biodiversidad, entre otras. Estos datos se constituyen en fuentes de datos que históricamente habían estado en el control de organizaciones, públicas o privadas; y cuyo acceso ha estado restringido. En este contexto, se debe señalar que para que un dato sea abierto, tiene que ser accesible y reutilizable, sin exigir permisos específicos, aunque los tipos de reutilización pueden estar controlados mediante una licencia. Asimismo, se debe señalar que los datos tienen que estar disponibles en la Internet sin limitaciones de acceso y tienen que ser ofrecidos en condiciones y en un formato conveniente, que permitan su reutilización, combinación con conjuntos de datos de diferentes orígenes. Lo anterior implica que los datos deben ser preferiblemente procesables por máquinas, en formato no propietario, y no cubierto por licencias que puedan limitar su uso.

Ello puede constituirse como componente esencial del “Gobierno Abierto” y abre un camino para la rendición de cuentas, por lo que los incentivos estatales a la transparencia, deben estar respaldados por mecanismos de reforzamiento. Los formatos y esquemas

abiertos de datos pueden aumentar la credibilidad en las instituciones, promover la transparencia y fomentar la *accountability*. En este sentido, la rendición de cuentas se constituye en un insumo que permite el reconocimiento, la asunción de responsabilidad y la actitud transparente sobre los impactos de las legislaciones, políticas públicas, decisiones y acciones realizadas, así como los productos y desempeño asociados a una determinada institución y organización, (Sáiz, 2016).

Contar con información en formatos públicos, accesibles y reutilizables, coadyuva a la formación ciudadana. A su vez, posibilita que los gobiernos sean más eficaces en la mejora del proceso de formulación de las leyes con un nivel mayor calidad. Esto debido a que la información abierta permite que los ciudadanos conozcan los procesos de toma de decisiones, además de las causas y los factores que influyen en dicha decisión (Sandra Elena y Ana Pichón Rivière, 2014).

Un momento importante clave en la generación del Open Data fue el 30 de septiembre de 2010 en que el Archivo Nacional del Reino Unido, libera una licencia gubernamental de reutilización de los datos generados por este país. Cada vez son más los gobiernos que anuncian y comienzan iniciativas de transparencia y gobierno abierto y que publican sus catálogos de datos. Actualmente, existen proyectos de Open Data desarrollados además de en EEUU, en países como Canadá, Reino Unido, Australia, Nueva Zelanda, Suecia o Noruega y también en España. La Fundación Centro Tecnológico de la Información y la Comunicación (CTIC), dispone de un catálogo que recoge todas las iniciativas de Open Data en el mundo.

Particularmente en América Latina en el informe *Recomendaciones para implementar una política de datos abiertos en el Poder Judicial*, (2014); un ejemplo del uso de Open Data es el caso del Poder Judicial chileno que publica las sentencias, la información presupuestaria y las estadísticas. De igual forma, Argentina es otro país que impulsa el Open Data donde se hizo un esfuerzo por mejorar la transparencia, el Poder Judicial aumentó la publicidad de la información en formatos reutilizables y accesibles. En la actualidad, publica su información estadística en HTML y en data sets, lo que facilita su uso. Asimismo, Uruguay también tiene un esquema de Open Data donde se ha generado un cambio positivo en la forma en la que el Poder Judicial pone a disposición sus datos. Los tres países crearon un ámbito propicio para desarrollar una política de datos abiertos. Sin embargo, la información no cumple aún con todos los requisitos para convertirse en su totalidad en un dato abierto. Al igual que en Chile y Argentina, Uruguay no publica la información en formatos abiertos, sino que utiliza formatos de licencias pagas, lo que limita el acceso a las personas que cuentan con ese programa (Sandra Elena y Ana Pichón Rivière, 2014).

Una experiencia en este ámbito legislativo es el parlamento de Chile, que ha impulsado el desarrollo de un *Plan de Acción del Parlamento abierto 2017-2108*, para el Congreso Nacional por parte del Grupo Bicameral de Transparencia del Congreso Nacional, mediante el cual se acordaron una serie de compromisos para centrar el proceso legislativo en la participación ciudadana (Congreso Nacional de Chile, 06 de diciembre, 2017).

Esta metodología supone el trabajo conjunto entre el Congreso Nacional y organizaciones de la sociedad civil y tienen el apoyo del Programa de Naciones Unidas para el Desarrollo (PNUD). En esta iniciativa participan junto al Grupo Bicameral de Transparencia, organismos como Ciudadano Inteligente, Educación 2020, Fundación Iguales, Instituto Chileno de Derecho y Tecnología, Base Pública, Espacio Público, Corporación Opción, Observatorio Fiscal. Los compromisos adoptados para el impulso de los datos abiertos son los siguientes (Congreso Nacional de Chile, 06 de diciembre, 2017):

- a.** Promover la aprobación de la Ley Orgánica Constitucional del Congreso Nacional (LOC) y elaborar reglamentos sobre la Dirección de Ética y Transparencia Parlamentaria; el de la Comisión Bicameral de Transparencia y el reglamento de Participación Ciudadana.
- b.** Elaborar un registro de participantes en sesiones de comisiones que quedará disponible en una sección de participación ciudadana de los sitios web de la Cámara y el Senado. Ello mediante un formulario electrónico en el que se podrán inscribir organizaciones de ciudadanas y personas jurídicas, ciudadanos y personas interesadas en participar del proceso legislativo.
- c.** Creación de un software para implementar el Registro de Participantes de Comisiones. Esta herramienta -cuyas categorías serán clasificadas por la Biblioteca del Congreso Nacional- permitirá el envío de invitaciones automáticas a las personas u organizaciones del registro de participación en comisiones para que se inscriban y puedan presentar sus puntos de vista por proyecto de ley de acuerdo a su especialización.
- d.** Generar plataformas virtuales de participación ciudadana en el proceso legislativo. Cada Corporación tendrá una plataforma que informará acerca de los contenidos de los distintos proyectos de ley, junto con permitir que las personas voten y envíen sus indicaciones al articulado.
- e.** Perfeccionar el registro de debate en comisiones, mediante la creación de un formato estandarizado que permitirá registrar el resumen del debate y se publicarán en formato do-

cumento y en datos abiertos XML en un plazo no mayor a 48 horas después de concluida cada sesión.

f. Crear de una sección web de participación ciudadana en el proceso legislativo, que agrupe las diferentes instancias y mecanismos para la participación ciudadana, y especificará tanto el propósito de cada una, junto con los criterios de participación y procedimientos.

En México, el Centro de Investigación y Docencia Económicas (CIDE) creó en el año 2016 el primer Laboratorio Nacional de Políticas Públicas (LNPP), con un acervo de casi 6 mil bases de datos sobre estudios jurídicos, opinión pública, economía y administración pública, entre otras áreas. Este proyecto es Financiado por el Consejo Nacional de Ciencia y Tecnología (Conacyt) y forma parte del Programa de Laboratorios Nacionales, este laboratorio será la primera sede externa del INEGI con un laboratorio de microdatos al que podrán acceder investigadores de cualquier institución educativa, gubernamental o de la sociedad, (CONACYT, Comunicado 114/15, 25 de noviembre de 2015).

Algunas de sus funciones de dicho laboratorio son ofrecer guía y asesoría metodológica desde niveles muy elementales como recopilación de información, muestreos, revisiones sistemáticas de literatura, hasta análisis de complejidad y simulaciones con Big Data.

De acuerdo con lo antes expuesto, se puede observar que esta mayor apertura permite ubicar la información y contribuye al desarrollo de la transparencia y rendición de cuentas debido al mayor nivel de información que se da por parte de los distintos actores y el propio control ciudadano (Sandra Elena y Ana Pichón Rivière, 2014). Esta nueva dinámica promueve entre otros aspectos:

- ∞ La realización de análisis legislativo y propuestas de la ley con un mayor sustento y conocimiento sobre lo que piensa y le interesa a la ciudadanía.
- ∞ La innovación y nuevas actividades económicas y productivas.
- ∞ El aumento de la confianza en el gobierno mediante el conocimiento de las políticas públicas, su ejecución e impacto.
- ∞ La construcción colaborativa de políticas entre Estado, sociedad lo cual tiende a reducir los costos en la implementación y elevar los niveles de cumplimiento.
- ∞ El fortalecimiento de los mecanismos de participación y colaboración interinstitucional.

Complementar las políticas de datos abiertos con los sistemas de publicación de la información ya existentes puede ayudar a fortalecer el análisis y seguimiento de las políticas de datos abiertos e identificar nuevos nichos donde puedan desarrollarse políticas de este tipo, justamente los datos abiertos constituyen una oportunidad para desarrollar el análisis legislativo con un mayor número de insumos. No obstante, se debe tener en cuenta tal como menciona Naser (2012), que la sola apertura de datos no cambia por sí misma la realidad, ya que se necesita hacer de manera conjunta un trabajo de sensibilización, concientización, capacitación de los diferentes actores involucrados en la utilización de estos datos.

2.4 Tendencias y patrones de comportamiento en la opinión pública mediante el Big Data

El Big Data supone un tipo de análisis dinámico que convierte las tendencias de opinión en cifras medibles y cuantificables. Estos datos, si se saben explotar, constituyen insumos básicos y ricos desde la opinión pública para analizar una campaña electoral, o para comunicar algún mensaje e idea desde el Gobierno. Sin embargo, frecuentemente estos datos no son bien aprovechados, incluso pueden ser ignorados o sobreestimados. Generalmente las instituciones y la política protagonizada por los partidos tradicionales, siguen gestionando de manera tradicional las políticas públicas y la acción política con un gran desconocimiento de esta realidad, o sólo fijándose en aquellos datos que consideran podrían interesar a la opinión pública o explicar algo sin pensar en un todo o, en nuevas posibilidades de información. Lo cierto es que es necesario desarrollar una cultura del dato para una nueva forma de hacer política y por supuesto para una nueva forma comunicación política la cual sea más consciente, y comprometida basada en datos para la toma de decisiones. Es decir, se requiere pasar del dato a gran escala al pensamiento de los datos que permita conocer e identificar ciertas tendencias (Gutiérrez-Rubí, s/f).

El análisis del Big Data cobró presencia hasta terminar las elecciones norteamericanas de 2012 donde se analizaron las tendencias de opinión de los posibles votantes a partir de una gran base de datos que incluía listas de donantes, datos recogidos por encuestadores y voluntarios, información y análisis semántico de redes sociales, entre otros aspectos que derivaban de los diversos equipos de campaña demócrata en 2008. Todo ello reunido en un único lugar para poder analizar la conducta de los votantes y poder impactar de mejor forma en el electorado con un solo mensaje. De esta forma, se identificó qué saben, qué sienten, por qué, con quién y cómo actúan los distintos sectores de la población. Más adelante con la aplicación de *Are you* de Obama se creó una aplicación en *Facebook* que

permitía a sus partidarios declararle su apoyo o hacer llegar a sus contactos un mensaje para que se uniera o influyera a favor Obama (Gutiérrez-Rubí, s/f).

De acuerdo con Antoni Gutiérrez-Rubí, se crearon perfiles de cada ciudadano que podía votar en el 2012. Elaboraron 16 millones de perfiles, y una vez que clasificaron a cada individuo, empezaron a orientar los mensajes para su reelección. En esto invirtió mil millones de dólares, mucho más que en comerciales de televisión. Y contrató a un grupo para manejar sus mensajes en *Twitter*, *Google*, *Facebook*, y *Craigslist*, entre otros. Este grupo creó los 16 millones de perfiles que les interesaban sobre los votantes indecisos. Sacaron datos de todos lados, incluso tuvieron acceso a las *Setup-Boxes*, que registra ¿Cómo ves la televisión? De esta manera, mandaron mensajes al electorado y le mostraron sólo aquello con lo que iba a estar de acuerdo prácticamente.

Más adelante *Cambridge Analytica* (CA), una empresa londinense fundada en 2013 por Alexander Nix, se dedicó al uso de datos para influir el comportamiento de audiencias. Al respecto, se observa un papel destacado del rol que tuvo la empresa en las últimas elecciones presidenciales de los Estados Unidos y su participación en el *Brexit*, además de haber operado también en otros países. CA usó datos de perfiles de *Facebook* sin el consentimiento de los usuarios mediante un test de personalidad, que operó como disparador para poder recopilar alrededor de 50 millones de muestras que permitieran conocer perfiles políticos de votantes estadounidenses para orientar las preferencias de éstos. En este año Aleksandr Kogan desarrolló un test de personalidad en formato de una aplicación de *Facebook*. Aproximadamente 265 mil usuarios completaron el test, el cual pedía información personal y acceso a la “red de amigos”, “estados” y “me gusta” en la red social. Toda esta información fue vendida a la empresa de Alexander Nix. Con dicha base de datos se obtuvieron los perfiles de aproximadamente 50 millones de usuarios, los cuales fueron manipulados en favor de la campaña de Donald Trump mediante la creación de publicidad seccionada según cada perfil, creando noticias falsas.

Es interesante saber que los usuarios de *Google*, *Facebook* y otras redes sociales gratuitas, no son clientes, como tal, sino que son el insumo del negocio. Es decir, se convierten en el dato mismo. El escándalo político de CA, puso a la luz “un secreto a voces”: el sustento de las plataformas de redes sociales se da a partir de nuestra propia información personal, cuyo valor político y comercial no es valuado en su real magnitud. Algunas empresas como *Cambridge Analytica* se dedican a perfilar, segmentar y clasificar nuestras idiosincrasias y esos datos son empleados en el diseño de estrategias de publicidad y propaganda política que persigue finalidades concretas (García, 2018).

Como se observa, los datos modifican nuestra percepción y conocimiento del entorno social porque nos cambian la perspectiva o la intensidad de la mirada. “Pensamos lo que vemos, como decía Aristóteles. Cambiar la mirada, obtener una nueva visión, es garantizar las bases para encontrar nuevas soluciones a viejos problemas. Conocer más y mejor la causalidad, las relaciones y los ecosistemas de la realidad social que queremos administrar o transformar es imprescindible” (Gutiérrez-Rubí, s/f. p.4).

Conocer el valor que pueden tener las estrategias de comunicación a medida o el implementar campañas que se ejecutan en tiempo real por medio del móvil son justamente las nuevas áreas de oportunidad para la política, la opinión pública y el uso del Big Data. Una oportunidad para la microsegmentación cambia la manera de hacer campañas: de macro a micropolítica orientada con una intencionalidad y dirección a los sujetos que comparten una determinada característica. Estos datos constituyen información, y ésta es poder para discernir como lo menciona Gutiérrez-Rubí, mediante qué mensaje un partido o un candidato puede acercarse más a un simpatizante o votante. Se trata de enviar el mensaje oportuno a la persona oportuna. El objetivo es orientar esa opinión a partir de la elaboración de ciertos contenidos orientados hacia un fin.

Con ello se ha dejado atrás la sobreexposición publicitaria de los medios de comunicación a la precisión de la personalización de los social media. A través de los datos que obtenemos, podemos conseguir impactar más con nuestros mensajes. Con ellos, la política tiene la posibilidad de conocer de manera muy exhaustiva al ciudadano, sus características, preferencias, necesidades y deseos.

Sin embargo, de acuerdo con Gutiérrez-Rubí, no es suficiente tener los datos. Hay que analizarlos y clasificarlos si queremos poder aprovechar toda la información que poseemos. “Vivimos en un mundo cambiante, donde la afluencia de datos es incesante. Cada minuto se comparten 280.000 tuits, se envían 240 millones de mensajes de correo electrónico, se hacen 4 millones de búsquedas en Google, se envían 350.00 fotos con *WhatsApp* y, justo al acabar de escribir estas cifras ya han quedado obsoletas” (s/f.: 7). El estudio de *Nature* muestra como cada individuo tiene sus propias huellas mediante movilidad la cuál es única, así como las huellas digitales que sirven para conocer la identidad de cada uno de los individuos. Los investigadores encontraron después de analizar cerca de año y medio de datos generados por los teléfonos móviles de un millón y medio de personas, que sólo serían necesarios cuatro puntos espacio-temporales para identificar el trazo de movilidad único del 95 por ciento de los individuos (Nature, 2016).

Martin Hilbert (2017), experto en redes digitales menciona que *Facebook* vale billones de dólares por la información y está dentro de las diez empresas del mundo tasadas a un precio más alto porque es proveedoras de información. Ahora teniendo entre 100 y 250 *likes* de un individuo en *Facebook*, se puede predecir su orientación sexual, su origen étnico, sus opiniones religiosas y políticas, su nivel de inteligencia y de felicidad, si usas drogas, si sus papás son separados o no. Con 150 *likes*, los algoritmos pueden predecir el resultado de tu test de personalidad mejor que tu pareja. Y con 250 *likes*, mejor que tú mismo. De acuerdo con Hilbert, por medio del Big Data tenemos tantos datos y tanta capacidad de procesarlos, así como de identificar correlaciones, que podemos hacer a la sociedad muy predecible y justamente cuando puedes predecir, puedes programar.

De igual forma contrasta el especialista en redes digitales que antes tenías que negociar con diplomáticos para que te prestaran una base pequeña de datos y menciona también que en las universidades hacían experimentos con apenas 15 alumnos de pregrado, que necesitaban créditos extra para pasar el grado, todos se seleccionaban compartiendo algunas características: blancos, de 18 años, estudiantes, etcétera y, bajo ese supuesto consideraban que así funciona la psicología humana y en consecuencia, se podía identificar las tendencias de comportamiento de la población, lo que hacía que al final, las decisiones que se tomaban no eran las correctas y en consecuencia las políticas públicas no funcionaban como se hubiera esperado. En contraste, de la noche a la mañana, el 95% de los sujetos que estudiamos pasó a tener un sensor de sí mismo 24 horas al día. Hoy a través de la tecnología sabemos dónde están las personas, pero también sabemos qué compran, qué comen, cuándo duermen, cuáles son sus amigos, sus ideas políticas y su vida social entre otras cosas. El gran cambio es que estamos conociendo a la sociedad como nunca antes y podemos hacer predicciones con un nivel científico que antes no era posible hacerlo.

Esta es una de las características de esta tecnología social: que relaciona personas construyendo comunidades de intereses, pero que también las empodera, anima y promueve la acción autónoma y creativa de los usuarios. Por ejemplo: los teléfonos inteligentes ya son parte de nuestra vida cotidiana. La gente habla directamente entre ella e interactúa mediante *WhatsApp*, de forma inmediata en cualquier lugar y hora. En este nuevo escenario, los partidos y los gobiernos deben aprender ver el teléfono celular como una nueva herramienta para ofrecer la información, pero también, para captar datos de los ciudadanos, como mínimo su número telefónico. El uso intensivo de *WhatsApp*, se convierte, entonces en una vía y una oportunidad para aproximar a personas con intereses comunes a través de la tecnología y resolver necesidades reales a partir de la propia iniciativa individual. Con ello se abre la reflexión derivada sobre el alcance y uso de la información, así como de los datos

públicos generados. De esta manera, el Congreso puede usar dicha información para enviarnos un mejor mensaje o conocer la opinión de la ciudadanía sobre algún tema que está en la agenda pública y en función de ello, realizar el análisis legislativo para un proyecto de ley que refleje la opinión pública de la ciudadanía.

De acuerdo con Gutiérrez-Rubí gracias al análisis y al cruce del Big Data, ahora los partidos políticos y los gobiernos pueden conocer de manera más precisa al ciudadano, sus características, preferencias, necesidades y deseos. Esto les permitirá acercarse mejor a determinados segmentos de población optando por la microsegmentación y la geolocalización para aplicar estrategias muy personalizadas y diferenciadas, presentar propuestas y estrategias de comunicación a medida. Esta segmentación se ha usado en las campañas electorales de todo el mundo para adaptar los mensajes del candidato a cada colectivo determinado, con el fin de conseguir una aproximación más efectiva. El objetivo es que el electorado se sienta escuchado y próximo a un partido, así como a quienes son y serán sus representantes. El análisis legislativo y el diseño de las políticas debe considerar de manera necesaria los datos, si no piensa en ellos reduce su capacidad de acción y planeación. La política que los desprecia es irresponsable, pero también es irresponsable la política que sólo se basa en ellos sin considerar otros aspectos. En este sentido, sugiere dicho autor que el Big Data puede constituirse como una herramienta social y como un elemento que puede revitalizar y renovar también la acción política además de la economía a partir del análisis de datos con conocimientos útiles para la acción y la gestión política, lo cual significa ir del *Big Data al Data Thinking* de la nueva política basada en el conocimiento de la opinión pública (Gutiérrez-Rubí, s/f).

En consecuencia, los partidos y gobiernos ya no están en disposición de actuar con base a una sola línea de acción y conseguir audiencia completa porque los públicos a los que están dirigidos son diversos y heterogéneos. Deben ofrecer a los votantes y a los ciudadanos contenidos útiles y de calidad, personalizado y segmentado a cada público, y a su vez debe ser disponible en multidispositivos para que pueda ser fácilmente compartido a través de una gran variedad de canales.

La gran cantidad de datos segmentados de que podemos disponer abre una oportunidad para comprender al público objetivo y ayudar a diseñar el contenido de las leyes de manera más adecuada donde éste se sienta reflejado y conecte de manera más adecuada con sus intereses. Además, los propios datos y su visualización pueden convertirse en material para construir narrativas, contribuyendo a que ideas complejas puedan ser comunicadas con mayor claridad, precisión y de una manera que resulte atractiva. En la actualidad la

tecnología ya no es una elección, sino una obligación para interactuar con los ciudadanos a partir del conocimiento de la opinión pública. Esta realidad cambia las ecuaciones de los formatos, las tipologías y los resultados de las relaciones entre usuarios-ciudadanos.

El conocimiento de la opinión pública por medio de los datos mediatizados y de las *APP* permite también, la movilización electoral, ya que puede favorecer prácticas políticas de intervención dirigida, por ejemplo: para tener contactos con los electores puerta a puerta, para la convocatoria de acciones, o para la participación electoral. Aspectos que permiten movilizar el voto hacia cierta dirección o conocer mejor a los votantes y sus necesidades.

Se está produciendo un cambio en la comprensión y aceptación de una determinada praxis y en la participación en la gestión política e institucional mediante el conocimiento y seguimiento de dicha información; es decir, el conocimiento de estos datos, y la gestión, monitorización, uso e interpretación de éstos por parte de la ciudadanía produce en ella el efecto de poder controlar, vigilar y fiscalizar lo que los políticos hacen o dicen. Ello muestra una tendencia creciente en la sociedad donde los datos son cada vez más abundantes y donde entenderlos va a ser básico no sólo para comprender a nuestra ciudadanía o votantes, sino también para llegar a ellos de una forma más rápida, más efectiva y con un mayor impacto mediante un determinado mensaje o bien para conocer su opinión y orientar el análisis legislativo en alguna dirección en particular.

De esta forma, podemos identificar la opinión de los ciudadanos sobre diversos temas sociales y sobre las principales políticas que necesitan ser creadas o promovidas o legisladas por los Estados (Bertoni y Torres, 2014). La confianza de los ciudadanos es clave para que pueda existir un despliegue de las potencialidades de las herramientas analíticas; sin embargo, para que se produzca esa confianza, es necesario que las personas estén convencidas de que se ha tomado en serio su derecho a la privacidad y a la protección de datos; que se han evaluado los riesgos para su privacidad desde el inicio del proyecto y que se han puesto en marcha las medidas necesarias para eliminarlos o mitigarlos; y que en todo momento los tratamientos de datos se ajustan a la normativa de protección de datos en vigor (Sáiz, 2016).

Conclusiones

El desarrollo y aprovechamiento del análisis con Big Data para impulsar el desarrollo social y económico requiere un enfoque multidisciplinario y el diseño de políticas estratégicas que permitan el desarrollo de estudios para la toma de decisiones, sean estos en el ámbito de la opinión pública o sobre cualquier otro aspecto de la investigación parlamentaria. El Big Data abre la posibilidad de realizar avances en muchas áreas de la administración pública, disciplinas científicas, así como para la opinión pública, aportes a la innovación y mejoras de productividad y competitividad y, por supuesto, se constituye en la actualidad en un nicho importante para realización y planeación de políticas públicas y análisis legislativos.

En este sentido, se debe considerar que los retos que se presentan con Big Data no solo están referidos al volumen de datos, sino también a su heterogeneidad, en su carácter poco estructurado, los aspectos de privacidad, los problemas de procesamiento y a la visualización de datos entre otros, lo cual complejiza el desarrollo de diversos estudios si no se consideran estos aspectos. Desde esta perspectiva, es importante reflexionar sobre los grandes desafíos que implica en el análisis de Big Data, ya que su implementación no sólo representa un aspecto técnico en relación al hardware y software, sino que también requiere de técnicos y profesionales capacitados que puedan implementar este tipo de análisis al interior de las instituciones. Esto implica una demanda de científicos de datos que utilicen distintas técnicas para extraer patrones de comportamiento derivados de grandes bases de datos al combinar métodos estadísticos y procesos de *machine learning*. Estas técnicas incluyen *association rule learning*, *cluster analysis*, clasificación, algún tipo de análisis de regresión y la elaboración de algoritmos que, gradualmente vayan perfeccionándose, así como del empleo de computadoras que desarrollen comportamientos basados en datos empíricos y redes neuronales inspiradas en estructuras de redes neuronales biológicas (*Deep Learning*), que funcionan de manera muy similar al cerebro, para buscar patrones en el comportamiento de los datos y la optimización de los mismos. De igual forma en este proceso, es central tener una nube de cómputo, ya que los recursos de computación requieren ser altamente escalables y se ofrecen como un servicio mediante una red.

La planificación de estas funciones es necesario para el análisis con Big Data, sin embargo en la actualidad se carece de ellas, lo que puede representar una fuerte limitante en su impulso. Por otro lado, para el análisis Big Data se requiere también crear una base de información que respalde el volumen, la variedad y la velocidad de datos que crecerán rápidamente y a su vez que esta información esté integrada, sea escalable, extensible y segura.

La integración de esta información es fundamental para cualquier esfuerzo analítico, ya que la capacidad de conectar los datos es otro de los grandes retos que se tienen que enfrentar y planificar. De igual forma es importante que las distintas iniciativas que se emprendan con Big Data, centren sus iniciativas en ámbitos que puedan proporcionar el máximo valor agregado al dato y previo a ello, identificar y comprender verdaderamente las necesidades previendo posibles comportamientos futuros. Este valor agregado puede derivarse a través de interacciones más oportunas, informadas o pertinentes y también pueden surgir a medida que las operaciones son capaces de mejorar la experiencia de dichas interacciones.

En este caso, la estrategia y los requisitos de Big Data dentro de una iniciativa que busca impulsar este análisis tiene que armonizar las necesidades de los usuarios o grupos a los que está dirigido mediante una guía u hoja de ruta que oriente a la institución mediante de un enfoque práctico para desarrollar su propio valor intrínseco. De igual forma, en su implementación es importante desarrollar funcionalidades analíticas sobre la base de prioridades, toda vez que ello implica hacer frente a una creciente variedad de herramientas analíticas, al mismo tiempo que se ha de sobrellevar una escasez de habilidades analíticas.

Lo mencionado implica invertir para adquirir herramientas como desarrollar habilidades tecnológicas en propio personal. Desde esta perspectiva es probable que surjan nuevos roles y modelos de trayectorias profesionales con el desarrollo de habilidades analíticas y funcionales. En este caso es central educar al personal e investigadores para explorar, crear una base para la acción y, continuar ampliando sus conocimientos centrándose en aquellos casos donde el uso donde Big data proporciona una ventaja competitiva.

Ello nos lleva a recordar que el análisis con Big Data surge como una herramienta potente y novedosa, pero a su vez surgen algunas interrogantes sobre sus posibles usos, alcances e implicaciones que puede tener. Es decir, se debe tener presente que estos sean lícitos y cuenten con el respaldo legal, además de que no generen abusos en sus usos basados en el valor económico de los datos personales considerados en la actualidad como: “el petróleo del siglo XXI”. En este caso, los datos en sí mismos no se deben considerar de manera discriminatoria, excluyendo a los sectores minoritarios mostrando una clara tendencia ideologizante en base a los resultados.

En este sentido, surgen algunas preocupaciones sobre su posible potencial en su uso con sectores y poblacionales vulnerables: niños, ancianos o colectivos tradicionalmente marginados, por lo que es necesario establecer garantías adecuadas en todos los ámbitos para el uso de datos de manera adecuada. Lo anterior nos lleva a reflexionar *¿Hasta dónde queremos y se puede llegar como sociedad de la información? ¿Qué beneficios nos aportará esta nueva forma de explotación de la información?*

de datos a cualquier precio? ¿Debemos plantearnos aspectos éticos sobre el uso que le dará a la información? y ¿Qué tipo de decisiones se pueden o deben tomar sobre cientos de miles de individuos con base a esa información recabada a gran escala?.

Es relevante que los interesados sean conocedores del impacto que los diferentes tratamientos pueden tener sobre su privacidad. Por otro lado, se debe tener presente los riesgos y amenazas que vienen aparejados con los aspectos técnicos y de seguridad. Así, la anonimización, debe ser considerada teniendo en cuenta el riesgo de reidentificación de la información, lo que obliga a realizar un análisis del riesgo para eliminar o minimizarlo, haciendo uso de técnicas de cifrado.

La generación e integración de datos requieren un esquema de gobernanza que garantice la obtención de datos de calidad. Desafortunadamente, la incapacidad para conectar datos almacenados ha sido uno de los grandes retos frente al volumen, variabilidad, validez y veracidad de éstos desde su origen. Es probable que Big Data genere su valor hasta que se utilice y supere algunos retos principalmente en la sistematización y organización de los datos, así como en la generación de una plataforma sólida, la cual permita obtener indicadores estandarizados derivados de una diversidad de fuentes.

En síntesis, algunas ventajas que se pueden derivar del uso del Big Data aplicado al conocimiento de la opinión pública y la investigación parlamentaria son las siguientes:

- a.** Decisiones más inteligentes, a partir de aprovechar nuevas fuentes de datos para que se pueda eficientar y mejorar el análisis legislativo.
- b.** Decisiones más rápidas y oportunas. Ello a partir de la captura y análisis de datos en tiempo real o de manera rápida.
- c.** Decisiones que marquen la diferencia en el conocimiento de la opinión pública para la elaboración de determinados contenidos más precisos orientados a un sector poblacional.

El diseño de las políticas, el análisis legislativo y el conocimiento de la opinión pública debe considerar de manera necesaria los datos. Sino piensa en ellos, reduce su capacidad de acción y planeación. En este sentido, el Big Data puede constituirse en una herramienta social como un elemento que puede revitalizar y renovar la forma de hacer política a partir del análisis de datos que deriven en conocimientos útiles para la acción y la gestión desde diversas esferas de lo social.

BIBLIOGRAFIA

- Acosta, J. (2016). *Generando valor con analytics. Performance Improvement: Mejor resultado. Mayor valor.* Recuperado de [http://www.hollandlaw.nl/Publication/vwLUAssets/valor_con_analytics/\\$FILE/EY-generando-valor-analytics.pdf](http://www.hollandlaw.nl/Publication/vwLUAssets/valor_con_analytics/$FILE/EY-generando-valor-analytics.pdf)
- AIS Group. (2016). *Siete pasos para poner en marcha una estrategia de Big Data orientada a las ventas.* Recuperado de: <http://tecnologiapara-tuempresa.ituser.es/estrategias/2016/05/siete-pasos-para-poner-en-marcha-una-estrategia-de-big-data-orientada-a-las-ventas>
- Amoroso, Y; Costales, D. (2016). Big Data: una herramienta para la administración pública. *Ciencias de la información*, Vol. 47 (3).
- Angelis, C. (2016). *La opinión pública entre la razón y el control social. Una actualización en la era del Big Data.* En AVATARES de la comunicación y la cultura. N° 11. ISSN 1853-5925.
- BBC Mundo. (2016). *¿Cuánto dinero gana Facebook contigo y como lo hace?* Redacción BBCMundo. Recuperado de: <http://www.bbc.com/mundo/noticias-37871331>
- BBC Mundo. (2017). *La página en la que puedes ver cómo Facebook te analiza para enviarte anuncios personalizados.* En NEWS MUNDO. Recuperado de: <http://www.bbc.com/mundo/noticias-39745436>
- BBC Mundo. (2018). *Con un test de Facebook, Cambridge Analytica obtuvo información de millones de usuarios.* Recuperado de: <https://www.animalpolitico.com/2018/03/como-un-test-de-personalidad-de-facebook-le-sirvio-a-cambridge-analytica-para-recolectar-informacion-privada-de-millones-de-usuarios-sin-que-lo-supieran/>
- Belbis J.I. (2015). Estudio de caso. Apertura Legislativa en el Cono Sur. ¿Y los datos? Iniciativa Latinoamericana por los Datos Abiertos. Recuperado de: <http://www.opendataresearch.org/dl/symposium2015/odrs2015-paper6.pdf>

- Big Data. (s.f) Del Bit.. al Big Data. Recuperado de: http://cdn2.hubspot.net/hub/239039/file-359994269-pdf/docs/PowerData_-_Del_bit%E2%80%A6_Al_Big_Data.pdf
- Cavagna. L (S.F). Big Data es el presente, el futuro es como rentabilizar los datos. Documento SAP. Recuperado de: <http://www.computing.es/siteresources/files/789/51.pdf>
- CONICYT. (2014). *Datos científicos abiertos. La ciencia la hacemos entre todos*. Programa de información científica. Recuperado de <http://datoscientificos.cl/files/manual-2014.pdf>
- CONICYT. (2014). La ciencia se hace entre todos. Creative CommonsAtribución-Compartir Igual 4.0 Internacional. Programa de información científica. Recuperado de: <http://datoscientificos.cl/>
- Costales, D; y Amoroso, Y. (2016). *Gobierno en internet con una mirada al Big Data*. La Habana, Cuba: Universidad de las Ciencias Informáticas.
- DATA CENTER MARKET. (2016). *La mayoría de las empresas condiciona su futuro a la adopción del Big Data. AIS Group propone una hoja de ruta para la adopción del Big Data orientado a las ventas*. Recuperado de: <http://www.datacentermarket.es/tendencias-tic/noticias/1089087032809/mayoria-empresas-condiciona-futuro.1.html>
- Datos abiertos. (2018). Wikipedia, La enciclopedia libre. Recuperado de: https://es.wikipedia.org/w/index.php?title=Datos_abiertos&oldid=108710794
- De Bustos, Y. (2017). *El Open Data como motor de innovación*. En COMPUTING. Recuperado de: <http://www.computing.es/analytics/opinion/1097263046201/open-data-motor-de-innovacion.1.html%E2%0A%20EI%20Open%20Data%20como%20motor%20de%20innovaci%C3%B3n>

- Dezhi. A, Guangli. W, Jun. L, Shengcai. Z y Yan. L (s.f). Key Techniques of Public Opinion Mining Based on Big Data. Recuperado de: <http://ijssst.info/Vol-17/No-18/paper1.pdf>
- Digital Tech Institute. (2017). *RECOMENDACIONES PARA INICIAR UN PROYECTO BIG DATA EMPRESARIAL*. Recuperado de: <https://www.digitaltechinstitute.com/recomendaciones-proyecto-big-data-empresa/>
- Domínguez, N. (2018). Por qué no hay asesores científicos en el parlamento. *El país*. Recuperado de https://elpais.com/elpais/2018/04/27/ciencia/1524855597_720944.html
- Elías, C. (2016). *Del Social Data al Big Data: evolución de la comunicación política y las campañas electorales en la Era Digital*. Recuperado de http://www.revistalatinacs.org/15SLCS/2016_libro/053_Elias.pdf
- Ferre, A; Sánchez, E. (2013). *Open data, big data: ¿hacia dónde nos dirigimos?*. Anuario ThinkEPI, v, 7. pp.150-156.
- García, A. (2018). *Cambridge Analytica, el Big Data y su influencia en las elecciones*. Recuperado de <http://www.celag.org/cambridge-analytica-el-big-data-y-su-influencia-en-las-elecciones/>
- Grautoff. M (2018). Big Data electoral. *El economista*. Recuperado de: <https://www.eleconomista.com.mx/opinion/big-data-electoral-20180502-0124.html>
- Guama, S.C. (2015). *Estudio comparativo de métodos existentes para integrar la información estructurada y no estructurada de una industria enfocado en la generación de conocimiento, desde la perspectiva de una solución integral del Big Data*. Tesis de maestría, Universidad de las Américas. Recuperado de [http://dspace.udla.edu.ec/bitstream/33000/3385/1/UDLA-EC-TMGSTI-2015-16\(S\).pdf](http://dspace.udla.edu.ec/bitstream/33000/3385/1/UDLA-EC-TMGSTI-2015-16(S).pdf)

- Gutiérrez, A. (S.F). *POLÍTICA: del Big Data al data thinking*. En ACOP Papers #2. Recuperado de: <https://compolitica.com/wp-content/uploads/ACOPPapersN%C2%BA2.pdf>
- Herrera Ávila Fernando, (24/05/2017) Senador Fernando Herrera Ávila, en conferencia Big Data, inteligencia artificial. Recuperado de <http://www.pan.senado.gob.mx/2017/05/senador-fernando-herrera-avila-en-conferencia-big-data-inteligencia-artificial/>
- Hopenhayn, D. (2017). *Martin Hilbert, experto en redes digitales: “Obama y Trump usaron el Big Data para lavar cerebros”* Hopenhayn. *big data, Estados Unidos*. Recuperado de: <http://www.theclinic.cl/2017/01/19/martin-hilbert-experto-redes-digitales-obama-trump-usaron-big-data-lavar-cerebros/>
- Ibáñez Adolfo. (2017). *Diplomado big data para políticas públicas*. Recuperado de: <https://gobierno.uai.cl/diplomado/diplomado-big-data-politicas-publicas/>
- IE. (2017). *Big data: ruta crítica para desarrollar nuevas oportunidades de negocio*. Recuperado de <https://www.ie.edu/es/relaciones-corporativas/insights/big-data-ruta-critica-desarrollar-nuevas-oportunidades-negocio/>
- IMPARK. (2014). *Estrategias de fidelización: cómo y cuánto contribuyen a las ventas de una marca*. Recuperado de: <http://ipmark.com/estrategias-de-fidelizacion-cmo-y-cunto-contribuyen-a-las-ventas-de-una-marca/>
- Interactive. (2016). *¿Qué aporta el Big Data a las campañas políticas?* Recuperado de: <http://interactivadigital.com/que-aporta-el-big-data-a-las-campanas-politicas/>
- Jeff. (2018). *Social Big Data and Big Data what’s the connection?* Recuperado de: <http://chinetekstrategy.com/blog/2017/12/28/social-listening-big-data/>

- Jeff. (2018). Social Listening: How public opinion monitoring applies to election campaigns? Recuperado de: <http://chinetekstrategy.com/blog/2018/02/05/social-listening-public-opinion-monitoring/>
- Johnson, E. (2014). Hype or Reality? How supply chain convergence and cloud impact global trade management usage. Recuperado de https://www.dropbox.com/home/SERVICIO%20SOCIAL/Bibliograf%C3%ADa%20BIG_data_toda/Informaci%C3%B3n%204?preview=datos+en++nube.pdf
- Kosinki, M. (2017). Ponencia magistral. Big Data inteligencia artificial y el futuro de la democracia. *Aportes al debate parlamentario*, No.4. Conferencia llevada a cabo en el Instituto Belisario Domínguez. Senado de la República, México.
- Larrea, I. (2017). Big Data y Open Data ¿Qué son y por qué importan?. En linked in. Recuperado de: <https://es.linkedin.com/pulse/big-data-y-open-qu%C3%A9-son-por-qu%C3%A9-importan-izaskun-larrea>
- Leites, P. (2017). Campañas políticas en tiempos del Big Data. En: la VOZ. Recuperado de: https://www.dropbox.com/referrer_cleansing_redirect?hmac=2PXh%2BhrU50s5Ptd1nR%2F808Ajf3mVnQnUV9SsP49BoFM%3D&url=http%3A%2F%2Fwww.lavoz.com.ar%2F-mundo%2Fcampanas-politicas-en-tiempos-de-big-data
- Leyva, G. (2015). Big data: la revolución que no debemos ignorar. *Realidad, datos y espacio. Revista internacional de estadística y geografía*, 6 (2), pp. 73-79.
- López, D. (2013). *Análisis de las posibilidades de uso de Big Data en las organizaciones*. Recuperado de <https://repositorio.unican.es/xmlui/bitstream/handle/10902/4528/TFM%20-%20David%20L%C3%B3pez%20Garc%C3%ADaS.pdf?sequence=1>

- Magnani, E. (2017). Big Data y política. *El poder de los algoritmos*. En: REVISTA NUEVA SOCIEDAD, No. 269. PP. 45-55. Recuperado de: http://nuso.org/media/articles/downloads/2.TC_Magnani_269.pdf
- Moneo, A. (2014). *Los motores del open data en América Latina*. En Banco de Desarrollo Mundial (BID). Recuperado de: <https://blogs.iadb.org/abierto-al-publico/2014/09/11/los-motores-del-open-data-en-america-latina/>
- UAI. (2018). *Diplomado Big Data para políticas públicas*. Escuela de gobierno. Recuperado de http://marketing.uai.cl/forms/folletos/gobierno/DBD_2018.pdf
- OBS. SCHOOL. B (S. F). El Director de Proyectos de Big Data, primeros pasos. Recuperado de: <https://www.obs-edu.com/int/blog-investigacion/sistemas/el-director-de-proyectos-de-big-data-primeros-pasos>
- ONU Derechos Humanos. (2018). *México experto de Naciones Unidas y de la CIDH exhortan a proteger la independencia de los medios en nuevo proyecto de ley sobre publicidad oficial*. Ginebra Washington. Recuperado de https://www.hchr.org.mx/index.php?option=com_k2&view=item&id=1106:mexico-expertos-de-naciones-unidas-y-de-la-cidh-exhortan-a-proteger-la-independencia-de-los-medios-en-nuevo-proyecto-de-ley-sobre-publicidad-oficial&Itemid=266
- Oracle. (2012). *From Overload to Impact: An Industry Scorecard on Big Data Business Challenges*. Recuperado de <http://www.oracle.com/us/industries/oracle-industries-scorecard-1692968.pdf>
- Oracle. M. (2014). Oracle Value Change execution. Reinventamos la Excelencia en Logística. Recuperado de: <https://www.oracle.com/index.html>
- Oracle. (2014). Transportation Management Systems. Recuperado de: https://www.logisticsmgmt.com/wp_content/oracle_wp_transmangmt_0814b.pdf

- Oracle. (2016). *Porqué los líderes de la cadena de suministro están migrando a la nube*. Recuperado de https://www.dropbox.com/home/SERVICIO%20SOCIAL/Bibliograf%C3%ADa%20BIG_data_toda/Informaci%C3%B3n%202?preview=Oracle-Why-Supply-Chain-Leaders.pdf
- Osorio, J; y Zuleta, C. (2016). *Estado del arte sobre frameworks de infraestructura para big data*. Recuperado de: <http://repositorio.utp.edu.co/dspace/bitstream/handle/11059/7219/00416083.pdf?sequence=1>
- Peña, M. (2018). *El Gobierno creó la unidad de “opinión pública” para manejar big data*. En: el destape. Recuperado de: <https://www.eldesatapeweb.com/el-gobierno-creo-la-unidad-opinion-publica-manejar-big-data-n41140>
- Pichón, A Elena, S; y. (2014). *Recomendaciones para implementar una política de datos abiertos en el Poder Judicial*. Documento de Políticas Públicas/ Recomendación N° 133. Buenos Aires: CIPPEC.
- Power Data. (S.f). *El gobierno de datos eficaz. La guía para minimizar errores y alcanzar objetivos de gobernanza de dato*. en: Recuperado de: https://cdn2.hubspot.net/hubfs/239039/docs/Ebook-Gobierno-Datos-Eficaz_1.pdf?t=1431528417133-
- Power Data. (S.f). *Master Data Management: Cómo crear una visión única de cliente Mejora la experiencia de cliente gracias a la gestión de tus datos maestros*. Recuperado de: <https://blog.powerdata.es/el-valor-de-la-gestion-de-datos/mejora-las-relaciones-con-los-clientes-con-un-master-data-management>
- Pymes y autónomos. (2014). *Que es Big Data y como se implementa en una empresa*. Recuperado de: <https://www.pymesyaautonomos.com/inspiracionparatunegocio/que-es-big-data-y-como-se-implementa-en-una-empresa>
- Quiroz.U. A. (2013). *Ontología de parlamentarios chilenos autores: modelación y aplicación*. Universidad Tecnológica Metropolitana. Santiago, Chile. Recuperado de: <https://www.bcn.cl/obtienearchivo?id=documentos/10221.1/40432/1/252336.pdf>

- Ransbotham, S; y Kiron, D. (2018). *Using Analytics to Improve Customer Engagement*. MIT Sloan Management Review.
- Reiventing Higher Education. (2017). 'Big data': ruta crítica para desarrollar nuevas oportunidades de negocio. Recuperado de: <https://www.ie.edu/es/relaciones-corporativas/insights/big-data-ruta-critica-desarrollar-nuevas-oportunidades-negocio/>
- Rodríguez, P; Palomino, N; y Mondaca, J. (2017). *El uso de datos masivos y sus técnicas analíticas para el diseño e implementación de políticas públicas en Latinoamérica y el Caribe*. Chile: BID. Recuperado de <https://publications.iadb.org/bitstream/handle/11319/8485/El-uso-de-datos-masivos-y-sus-tecnicas-analiticas-para-el-diseño-e-implementación-de-políticas-públicas-en-Latinoamérica-y-el-Caribe-resumen-de-políticas.PDF?sequence=3>
- Schroeck, M; Shockley, R; Smart, J; Romero, D; y Tufano, P. (S.F). *Analytics: el uso de Big Data en el mundo real Cómo las empresas más innovadoras extraen valor de datos inciertos*. IBM, Global Business Services Business Analytics and Optimisation.
- Sinnaps. (s.f). *¿Qué es el método de la ruta crítica?*. Recuperado de: <https://www.sinnaps.com/blog-gestion-proyectos/metodo-de-la-ruta-critica>
- Scrollini, F. (2017). *Precisamos políticas de datos abiertos*. En *Investigación, Opinión*. Recuperado de: <https://idatosabiertos.org/precisamos-politicas-de-datos-abiertos/>
- Soares, S. (2016). *Una Plataforma Completa para el Gobierno de Big Data, la Gestión de Datos y la Analítica*. en Information Asset, llc.
- Tamaris. A. (2016). *Aplican Big Data para políticas públicas*. En: REFORMA. Recuperado de: <https://www.reforma.com/aplicacioneslibre/articulo/default.aspx?id=747937&md5=4f10a435e0d6d27ae86461e6ca1eb4c6&ta=0dfdbac11765226904c16cb9ad>
- Tantum. (s.f). *¿Puede Big Data mejorar la toma de decisiones estratégicas?* Recuperado de <http://www.tantum.com/argentina/big-data-y-toma-de-decisiones-estrategicas/>

- Tomar, L; Guichenev, W; Kyarisiima, H; y Zimani, T. (2016). *Big Data in the public sector. Selected Applications and Lessons Learned*. IDB.
- Torres. N. (2014). Hacia una política integral de gestión de la información pública. Todo lo que siempre quisimos saber sobre archivos (y nunca nos animamos a preguntarle al acceso a la información). Buenos Aires, Argentina. Departamento de Diseño de la Universidad de Palermo. Recuperado de: https://datos.bcn.cl/es/obtienearchivo?id=documentos/10221.1/44372/1/Hacia_una_politica_integral-kk.pdf
- Velasco, P; y Violler, P. (2016). *Información financiera y discriminación laboral en Chile: un caso de estudio sobre Big Data*. Recuperado de <https://www.derechosdigitales.org/wp-content/uploads/big-data-informe.pdf>
- Vidal, M. (2017). *Big Data. Inteligencia Artificial, Política*. Campaña electoral, inteligencia artificial y Big Data ¿influye la tecnología en el voto?. Recuperado de: <https://www.marcvidal.net/blog/2017/12/5/campaa-electoral-inteligencia-artificial-y-big-data-influye-la-tecnologia-en-el-voto#>
- Vidal, M. (2012). *Del efecto viral al efecto vital. El camino desde la emisión hacia la propagación*. Recuperado de: <https://www.marcvidal.net/blog/2012/04/el-efecto-viral.html>
- Vilalta, C. (2017). *Información para la prevención del delito y la violencia*. Preparado para la División de Innovación para Servir al Ciudadano. Banco Interamericano de Desarrollo (BID).
- Zaforas, M. (2017). *Open Data, o cómo los datos están cambiando el mundo*. En Blog Tecnología para Negocios. Recuperado de: <https://www.paradigmadigital.com/techbiz/open-data-los-datos-estan-cambiando-mundo/>

DIRECCIÓN GENERAL DE ANÁLISIS LEGISLATIVO

Dr. Alejandro Navarro Arredondo
Director General

Mtra. Gabriela Ponce Sernícharo
Investigadora

Mtro. Cornelio Martínez López
Investigador

Dr. Itzkuauhtli Benedicto Zamora Saenz
Investigador

Mtra. Irma del Rosario Kánter Coronel
Investigadora

Mtro. Christian Uziel García Reyes
Investigador

Dra. Carla Angélica Gómez Macfarland
Investigadora

Mtra. Lorena Vázquez Correa
Investigadora

Lic. María Cristina Sánchez Ramírez
Investigadora

Mtro. Israel Palazuelos Covarrubias
Investigador

Dr. Juan Pablo Aguirre Quezada
Investigador

Denise Velázquez Mora
Diseño Editorial

Serie: Cuaderno de Investigación. No. 46
“El uso del Big Data en los estudios de opinión pública”

Elaborado por: Mtra. Erika Tapia Nava

Cómo citar este documento:

Tapia Nava, Erika "El uso del Big Data en los estudios de opinión pública"
Cuaderno de Investigación No. 46,
Instituto Belisario Domínguez, Senado de la República, Ciudad de México, 2018, 62p.

Las opiniones expresadas en este documento son de exclusiva responsabilidad de su autora o autor y no reflejan, necesariamente, los puntos de vista del Instituto Belisario Domínguez o del Senado de la República.

Cuadernos de Investigación es un trabajo académico cuyo objetivo es apoyar el trabajo parlamentario.



[IBD Senado](#)



[@IBD Senado](#)



www.senado.gob.mx

Donceles, No. 14, Centro
Histórico, C.P. 06020, Delg.
Cuauhtémoc, Ciudad de México