



EL BIG DATA EN SALUD

BIG DATA IN HEALTH

*Estrada Fuentes, Armando**

**Doctor en Medicina, Maestría en Salud Pública, Maestría en Docencia Superior y Licenciatura en Tecnología de Programación y Análisis de Sistemas Computacionales. Email: dr.estrada@yahoo.com*

RESUMEN

El concepto de Big Data se basa en el análisis de un gran volumen de datos. El Big Data trabaja principalmente con datos no estructurados y estructurados.

Las principales características que definen el Big Data pueden resumirse en lo que se conoce como las siete Vs: Volumen, Velocidad, Variedad, Veracidad, Valor, Variabilidad y Visibilidad.

Una tecnología fundamental para el presente y el futuro, eso es el Big Data, porque localiza una enorme cantidad de datos, los extrae, organiza, almacena y analiza, todo esto en tiempo real, lo cual hace que la información se convierta en un activo muy valioso, porque permite predecir y tomar decisiones.

Toda la información en salud para el Big Data proviene de fuentes de información: La web y las redes sociales, los datos de máquina a máquina, los datos biométricos y los datos generados por los usuarios (Expedientes Clínicos Electrónicos).

El Big Data aplicado a la medicina constituye una nueva herramienta de obtención de conocimiento.

Palabras claves: Big Data, Machine Learning, Big Data Analysis

ABSTRAC

The concept of Big Data is based on the analysis of a large volume of data. Big Data works mainly with unstructured and structured data.

The main characteristics that define Big Data can be summarized in what is known as the seven Vs: Volume, Velocity, Variety, Veracity, Value, Variability and Visibility.

A fundamental technology for the present and the future, that is Big Data, because it locates an enormous amount of data, extracts it, organizes, stores and analyzes it, all this in real

time, which makes the information become an asset. very valuable, because it allows you to predict and make decisions.

All the information generated in health for Big Data comes from information sources: the web and social networks, machine-to-machine data, biometric data and data generated by users (Electronic Medical Records).

Big Data applied to medicine constitutes a new tool for obtaining knowledge.

The large amount of information that is processed by computer in the medical environment makes it an ideal field for the use of techniques known as Big Data Analysis (BDA) and Machine Learning (ML), which may allow us to improve our capacity in the future. clinical research and more precisely target the therapies provided to patients.

Keywords: Big Data, Machine Learning, Big Data Analysis

INTRODUCCION

Big data (en español, grandes datos o grandes volúmenes de datos) es un término evolutivo que describe cualquier cantidad voluminosa de datos estructurados y no estructurados que tienen el potencial de ser extraídos para obtener información¹. El concepto de Big Data se basa en el análisis de un gran volumen de conjuntos de datos.

El Big Data trabaja principalmente con datos no estructurados y estructurados. Los datos estructurados son datos con formato o esquema fijo que poseen campos fijos. Por ejemplos: Base de datos, Hoja de cálculo, archivos o ficheros. En cambio, los datos no estructurados son datos sin tipos definidos, sin estructura uniformes, se almacenan principalmente como documentos u objetos. Por ejemplos: Audios, videos, fotografías, emails, artículos, libros, mensajería por WhatsApp, entre otros.

El Big Data es la tecnología capaz de identificar, extraer (de fuentes diversas y dispersas), almacenar, ordenar, gestionar y analizar cantidades masivas de datos que con las aplicaciones tradicionales no se podrían procesar¹.



ANTECEDENTES

Es difícil situar el origen del Big Data. En 1989 el término apareció en un artículo del periodista Erik Larson en el Harper's Magazine y publicado más tarde por The Washington Post. Se refirió a cómo los empresarios cuidaban los datos de los clientes por el bien del consumidor. Sin embargo, estos datos podían ser utilizados en temas y materias mucho más amplios que lo meramente comercial y ahí estaba el peligro para la privacidad de las personas¹.

El término "Big Data", que abarca la informática y las estadísticas / econometría, probablemente se originó en conversaciones en la mesa del almuerzo en Silicon Graphics Inc. (SGI) a mediados de la década de 1990, en el que John Mashey figuraba prominentemente. Las primeras referencias académicas significativas son posiblemente Weiss e Indurkha (1998) en informática y Diebold (2000) en estadística / econometría. Una nota de investigación no publicada de 2001 por Douglas Laney en Gartner enriqueció el concepto significativamente. Por lo tanto, el término "Big Data" parece razonablemente atribuido a Mashey, Indurkha

y Weiss, Diebold y Laney. Big Data el fenómeno continúa sin cesar, y como disciplina está emergiendo².

Con el nacimiento de la Internet en 1989 se abren los primeros caminos a la generación masiva de datos. De esto deriva la aparición de los primeros sistemas de gestión y almacenamiento de información que nos permite procesar e interpretar el Big Data.

CARACTERÍSTICAS

Para poder estimar lo grande que es el Big Data, empezaremos a describir en el Cuadro No. 1 las diversas medidas de almacenamiento de información digital. Por ejemplo, un megabyte (MB) equivale a 1 millón de bytes y nos permite almacenar un libro de 500 páginas que tuviera sólo texto³.

Con un terabyte, que equivalente a un millón de megabytes, podríamos almacenar 2,767 copias de la Enciclopedia Británica, ó 16,667 horas de música, ó 1,333 horas de videos³.

Se necesitarían casi 5 exabytes, ó 5 millones de terabytes, para almacenar las secuencias del genoma humano de todas las personas del mundo³.

Cuadro N°1 Resumen de medidas de almacenamiento de información digital y sus equivalencias.

Medida	Símbolo	Equivalencias	Equivalencia en Bytes	Equivalencia aproximada	
Bit	b	1 ó 0	1 ó 0	1/8	Sistema binario (1 ó 0)
Byte	B	8 b	10 ⁰ B	1	Un carácter (letra, número o símbolo)
Kilobyte	KB	1024 B	10 ³ B	1024	Una historia muy corta.
Megabyte	MB	1024 KB	10 ⁶ B	1,048,576	Una novela de 500 páginas.
Gigabyte	GB	1024 MB	10 ⁹ B	1,073,741,824	Mil megabytes
Terabyte	TB	1024 GB	10 ¹² B	1,099,511,627,776	2,767 copias de la Enciclopedia Británica
Petabyte	PB	1024 TB	10 ¹⁵ B	1,125,899,906,842,624	1,000 bibliotecas de 5,000 libros
Exabyte	EB	1024 PB	10 ¹⁸ B	1,152,921,504,606,846,976	Mil Petabytes
Zettabyte	ZB	1024 EB	10 ²¹ B	1,180,591,620,717,411,303,424	Mil Exabytes
Yottabyte	YB	1024 ZB	10 ²⁴ B	1,208,925,819,614,629,174,706,176	Mil Zettabytes
Brontobyte	BB	1024 YB	10 ²⁷ B	1,237,940,039,285,380,274,899,124,224	Mil Yottabytes
Geopbyte	GeB	1024 BB	10 ³⁰ B	1,267,650,600,228,229,401,496,703,205,376	Mil Brontobytes
Saganbyte	SB	1024 GeB	10 ³³ B	1,298,074,214,633,706,907,132,624,082,305,024	Mil Geopbytes
Jotabyte	JB	1024 SB	10 ³⁶ B	1,329,227,995,784,915,872,903,807,060,280,344,576	Mil Saganbytes

Fuente: Sistema Internacional de Unidades.

Las principales características que definen el Big Data (Figura No. 1) pueden resumirse en lo que se conoce como las siete Vs: Volumen, Velocidad, Variedad, Veracidad, Valor, Variabilidad y Visibilidad⁴.

1. Volumen: La característica principal que define Big Data es la gran cantidad de volumen de información digital existente que se maneja. En la actualidad, cuando se habla de bases de datos masivas se refiere a magnitudes del orden de petabytes o exabytes⁵.

2. Velocidad: Otra de las características esenciales de Big Data es la enorme velocidad en la generación, recogida y proceso de la información^{4,5,6}.

3. Variedad: Alude a los diversos tipos y fuentes de información digital. En este sentido, los sistemas Big Data permiten la integración de datos de origen estructurados y no estructurados, así como gráficos, texto, sonido o imágenes^{4,5,6}.

4. Veracidad: El Big Data ha de ser capaz de tratar y analizar inteligentemente el gran volumen de datos con la finalidad de obtener una información verídica y útil que nos permita mejorar nuestra toma de decisiones^{4,5,6}. El Big Data exige no solo que los datos sean muchos, analizados y aprovechados a gran velocidad, de diversas fuentes, sino que estos sean veraces y, por ello, confiables. ^{4,5,6}. La veracidad hace referencia al nivel de fiabilidad asociado a ciertos tipos de datos⁷.
5. Valor: En este punto reside el análisis de la utilidad de la información digital, específicamente en lo relativo a la salud. Como no todos los datos son válidos y pueden utilizarse para tomar decisiones, es importante limpiar los datos y confirmar que son relevantes para el propósito que se desea. El valor hace referencia a los beneficios que se desprenden del uso de Big Data ⁷.
6. Variabilidad: Los datos sueltos no proveen información porque son variables, un mismo comportamiento puede tener distintas causas, puede repetirse en diferentes momentos y tener diferentes significados. Cuando tienes muchos datos, en realidad puedes utilizarlos con muy distintos fines y formatearlos de distintas maneras. No es fácil recoger tantos datos, analizarlos y gestionarlos de la manera más adecuada, así que lo normal es usarlos varias veces. Eso es lo que significa la variabilidad: la opción de utilizar los datos con distintos fines.
7. Visibilidad: Cuando hablamos de visualización nos referimos al modo en el que los datos son presentados. Una vez que los datos son procesados (los datos están en tablas y hojas de cálculo), necesitamos representarlos visualmente de manera que sean legibles y accesibles, para encontrar patrones y claves ocultas en el tema a investigar.

Figura No. 1. Principales características del Big Data.



Tomando lo antedicho en consideración, se advierte que el término Big Data en el campo de la salud guarda relación con vastas cantidades de información compleja relativa a la salud de gran cantidad de individuos y proveniente de múltiples fuentes, pero accesible para menor cantidad de sujetos, en tanto su procesamiento exige el desarrollo de nuevas y más potentes herramientas informáticas. Destaca el Comité que la definición del término “salud” de la OMS debe ser considerada como un “norte normativo” y un enfoque holístico de la problemática sanitaria, razón por la cual Big Data ofrecería una alternativa tecnológica para afrontar globalmente las cuestiones asociadas a la salud⁴.

Una tecnología fundamental para el presente y el futuro, eso es el Big Data, porque localiza una enorme cantidad de datos, los extrae, organiza, almacena y analiza, todo esto en tiempo real, lo cual hace que la información se convierta en un activo muy valioso, porque permite predecir y tomar decisiones. En el ámbito de la salud pública ayuda a anticipar enfermedades, epidemias, a identificar enfermedades raras y mejorar la medicación de pacientes⁷.

LAS FUENTES DE INFORMACIÓN DEL BIG DATA

Podría afirmarse que toda la información generada en salud proviene de alguna de las siguientes tipologías de fuentes de información (Figura No. 2):

1. **La web y las redes sociales:** La generación e interacción de datos de redes sociales como Facebook, Twitter, o LinkedIn, además de la información de sitios web de salud o las aplicaciones de smartphones⁵.
2. **Los datos de máquina a máquina:** La información proveniente de las lecturas de los sensores, medidores y otros dispositivos⁵.
3. **Las grandes transacciones de datos:** Reclamaciones de atención médica y otros registros de facturación cada vez más disponibles en formatos semiestructurados y no estructurados⁵.
4. **Los datos biométricos:** Huellas dactilares, genéticos, escáner de retina, rayos X y otras imágenes médicas, la presión arterial, el pulso y lecturas de oximetría de pulso y otros tipos similares de datos⁵.
5. **Los datos generados por los seres humanos:** Datos no estructurados y semiestructurados, tales como expedientes clínicos electrónicos (Electronic Medical Records), notas de los profesionales sanitarios, correos electrónicos y documentos en papel⁵.

Figura No. 2. Fuentes de información del Big Data.



EL BIG DATA ANALYTICS Y EL MACHINE LEARNING

La gran cantidad de información que se procesa informáticamente en el entorno de la medicina la convierte en un campo ideal para el empleo de técnicas conocidas como Big Data Analysis (BDA) y Machine Learning (ML), que pueden permitir en el futuro mejorar nuestra capacidad de investigación clínica y dirigir de manera más precisa las terapias que se proporciona a los pacientes.

Big Data Analytics

El Big Data Analytics se refiere al almacenamiento, administración y análisis de grandes volúmenes de datos a través de métodos estadísticos o científicos para descubrir relaciones entre los datos. Recientemente se ha aplicado al campo de la salud como la medicina personalizada, registros de salud, estancias y readmisiones de pacientes y biomedicina. Como ejemplo, se aplica el modelo propuesto en el caso del monitoreo a distancia de la salud de un paciente con problemas del corazón, como una base para su implementación real en un trabajo futuro⁸.

Los campos de la salud que se pueden beneficiar del Big Data Analytics son^{5,8}:

1. Medicina personalizada: Aunque actualmente se han desarrollado muchos medicamentos para distintas enfermedades, se requiere tener diagnósticos óptimos que permitan determinar que enfermedad tiene un paciente y proporcionar un tratamiento adecuado^{5,8}.

2. Registros de salud: Estos registros se refieren a los historiales de salud de pacientes, fechas de ingreso a hospitales, enfermedades adquiridas, padecimientos, tratamientos a los que se ha sometido, notas médicas, resultados clínicos, entre otros datos. Aunque los centros de atención de salud en el mundo no cuentan con un estándar único para los registros, es importante analizar la información de distintos hospitales^{5,8}.
3. Estancias y readmisión de pacientes: Muchos pacientes son internados, tratados u operados en hospitales, sin embargo, luego de haber sido dados de alta, retornan a los pocos días, esto representa un costo para los pacientes, hospital y aseguradoras. La tecnología de Big Data Analytics podrían proveer técnicas de análisis de datos para minimizar la reincidencia hospitalaria^{5,8}.
4. Biomedicina: Gracias al secuenciado del genoma humano, ahora es posible analizar enfermedades para comprender su estructura a nivel celular y molecular, para generar fármacos más efectivos^{5,8}.

Machine Learning

El Big Data emplea herramientas diferentes en el ámbito de la inteligencia artificial, como el Machine Learning (aprendizaje por las máquinas) o las redes neuronales.

El concepto de Machine Learning (ML)⁹ o «aprendizaje máquina» data de mediados del siglo XX, y se definió ya en un artículo de Samuel de 1959¹⁷ como un apartado de la inteligencia artificial que usa técnicas estadísticas y algoritmos computacionales para proporcionar a los ordenadores la capacidad de «aprender», es decir, mejorar sus resultados en una tarea específica tras procesar datos en suficiente cantidad y sin unas instrucciones explícitas externas (y por tanto potencialmente sesgadas) proporcionadas por el programador. El ámbito del ML está estrechamente relacionado con otros campos, como la simulación y el modelado, la optimización de sistemas y la estadística. En todos ellos se emplean de manera intensiva técnicas matemáticas comunes que requieren entrenamiento específico.

El Machine Learning o aprendizaje automático es un concepto derivado como respuesta al fenómeno de Big Data, el cual se refiere al diseño de programas o algoritmos que pueden aprender reglas a partir de datos, adaptarse a cambios y mejorar el rendimiento con la experiencia.

Básicamente se podría definir como la capacidad que tienen las computadoras para aprender y llegar a conclusiones; aportando soluciones a problemas concretos y generando

conocimiento a partir de la información proporcionada por el Big Data. El “cerebro” de estos softwares de aprendizaje artificial o Machine Learning es el algoritmo, que es una secuencia o serie de instrucciones que dan las pautas para estructurar los datos en forma de modelos que permitan al sistema “pensar” y aportar soluciones o emitir un diagnóstico o previsión.

Se puede dividir el Machine Learning en dos grandes grupos¹⁰:

- El aprendizaje supervisado o predictivo, en donde la máquina aprende no sólo de los propios datos finales (inputs) sino que es posible darle modelos o datos adicionales ya categorizados (outputs) para que el aprendizaje sea mucho más fiable.
- El aprendizaje no supervisado o descriptivo en el que sólo se dan los inputs a la máquina para que encuentre patrones interesantes a partir de los datos.

El abanico de aplicaciones y utilidades que las técnicas de Machine Learning tienen en el ámbito de la salud es cada vez más amplio. Estas son algunas de las más destacables¹¹:

- Ofrecer a los profesionales y usuarios del sistema sanitario herramientas avanzadas de soporte al proceso de toma de decisiones.
- Proporcionar tanto a los estudiantes como a los profesionales de la rama sanitaria información puntual, actualizada y a tiempo real sobre procedimientos o temas concretos, así como ofrecer la resolución de dudas o problemas que puedan surgir en la práctica diaria.
- Potenciar y agilizar procesos clave en la gestión sanitaria tanto en temas de logística como en los relacionados con la atención al paciente.
- Favorecer el diagnóstico de precoz de un buen número de patologías (el campo de la oncología es uno de los que más se pueden beneficiar de estas técnicas, según los expertos).
- Permitir la detección no invasiva de algunos problemas de salud.
- Predecir de forma precisa el riesgo que existe de desarrollar una determinada enfermedad, facilitando el camino hacia la obtención de biomarcadores, por ejemplo.
- Valorar anticipadamente los efectos que el tratamiento va a tener en cada paciente.
- Hacer posible una atención médica y una medicina cada vez más personalizada.

- Ahorrar costes y aumentar la productividad de los profesionales.
- Facilitar la innovación y el desarrollo de nuevos productos y servicios.

LA INVESTIGACIÓN Y EL BIG DATA

Los datos útiles para la investigación en salud pública y en epidemiología proceden habitualmente de fuentes diseñadas para la investigación o bien de fuentes secundarias, como las historias clínicas, pruebas de laboratorio, censo de población, registros de enfermedades, etc¹².

Se concluye que las buenas prácticas clínica en la investigación en salud pública y en epidemiología no han de ser diferentes para las investigaciones que usen Big Data. Por tanto, la división entre la investigación con Big Data y la investigación tradicional no parece pertinente¹².

La investigación que aplique análisis de datos masivos requiere sistemas de supercomputación y nubes que no deberían externalizarse. El Comité de Ética de la Investigación (CEI) debe exigir que el cruce entre bases de datos de distinta índole, entre las que se encuentren historias clínicas, se efectúe en el perímetro de la institución sanitaria y no fuera¹³.

Las autoridades competentes deberían habilitar los correspondientes sistemas informáticos para que la exploración de datos sea viable y segura; y para que todo personal, que trate datos, cumpla con el deber de la confidencialidad¹³.

Sin duda, la investigación clínica es una de las aéreas en las que el análisis de datos a gran escala promete tener un mayor impacto. Por ejemplo, revelar patrones en la expresión genética de pacientes permitiría elucidar los mecanismos a través de los cuales ciertas enfermedades actúan. Determinar que estructuras moleculares tienen correlaciones fuertes con efectos fisiológicos podría tener un gran impacto en el desarrollo de nuevos fármacos. Y agrupar una cantidad masiva de casos y controles permitiría validar conclusiones clínicas obtenidas a partir de estudios con un número reducido de participantes¹⁴.

RETOS

Algunos de los principales retos de este sector son: requiere inversiones muy altas, hacen falta profesionales calificados, generar confianza en las personas, considerar que no toda la información recopilada es cierta, debido a que las personas suelen mentir en su vida real y en internet también lo hacen¹.



Las herramientas de Big Data Analysis y Machine Learning suponen una gran oportunidad para mejorar la gestión estratégica de las unidades, el manejo de casos clínicos concretos y la investigación clínica. Sin embargo, para poder aprovechar esta nueva metodología necesitamos evolucionar incorporando nuevos recursos humanos (personal especializado con conocimientos clínicos y entrenamiento en inteligencia artificial) y tecnológicos. Además, debemos ser capaces de armonizar los condicionantes de privacidad y seguridad de los datos de nuestros pacientes con la posibilidad de utilizar grandes bases de datos clínicos de manera eficiente¹¹.

¿Qué importancia tiene este hecho desde el punto de vista de la Medicina? Está claro que no podemos esperar que los médicos se conviertan en expertos en Big Data, al igual que no pueden ser consumados estadísticos. Pero, lo mismo que en los grupos de investigación de cierto tamaño se dispone de bioestadísticos, en un futuro inmediato será preciso contar con otro tipo de expertos versados en el manejo y análisis de Big Data. Y, lo que es más importante los médicos deberán entenderlos principios de este nuevo tipo de análisis de una forma semejante a como hoy tienen conocimientos básicos de estadística; podrán así orientar el diseño de los estudios de investigación y, lo que es clave, entender la validez metodológica de los resultados de esos estudios, propios o ajenos. Todo ello aconseja que se comiencen a introducir este tipo de enseñanzas en la formación curricular de los futuros médicos, que sin duda se van a enfrentar a una práctica de la medicina muy diferente de la actual.

PRONÓSTICOS POSITIVOS

La gran mayoría de los agentes que participan en las estructuras de los servicios de salud reconocen que el análisis del Big Data puede ofrecer nuevas posibilidades en la elaboración de modelos predictivos, patrones de comportamiento, el descubrimiento de nuevas necesidades, reducir riesgos, así como proveer servicios más personalizados, todo ello en tiempo real y teniendo en cuenta toda la información relevante¹⁵.

El mejor cuidado de la salud, impulsado por el tratamiento global y sistemático de los datos, debiera plantearse desde tres perspectivas: la individual de cada paciente, la de las comunidades de pacientes y, por último, la de la sociedad en su conjunto¹⁵.

La computación en la nube, posibilitará en los próximos años contar con la capacidad de cálculo para procesamiento de datos con la complejidad a nivel de Big Data a costos accesibles,

ya que las aplicaciones se configurarán y portarán una sola vez, desplegándose en base a necesidad. La computación de altas prestaciones será un servicio más que podrán utilizar los profesionales de las ciencias de la salud para beneficio de pacientes y la obtención de nuevos descubrimientos¹⁶.

La computación cuántica está a la vuelta de la esquina y la historia del Big Data sigue avanzando. La capacidad creciente de acumulación y análisis de datos aplicado a la medicina constituye una nueva herramienta de obtención de conocimiento que, básicamente, supone substituir la técnica del muestreo por el análisis de todos los datos de un conjunto, o al menos de una abrumadora mayoría, esperando encontrar relaciones de coincidencia en vez de relaciones de causalidad. Esto va a suponer una serie de cambios a distintos niveles. En primer lugar, a nivel conceptual, pues el principio de causalidad está sólidamente arraigado en la mentalidad biomédica; pero además va a requerir que nos familiaricemos con los métodos de análisis de Big Data, para contribuir al diseño de estudios y entender de forma crítica sus resultados. Por otra parte, actualizará el debate entre privacidad de datos y bien común, e incluso puede que aconseje rediseñarlos actuales criterios curriculares. Pero, en definitiva, nos debe obligar a interesarnos en un campo con un gran futuro, para que en su desarrollo posterior en el campo de la Medicina estemos presentes como agentes, y no como receptores pasivos, de lo que tecnologías de este tipo nos ofrecen¹⁷.

CONCLUSIONES

El concepto de Big Data se basa en el análisis de un gran volumen de datos.

El Big Data trabaja principalmente con datos no estructurados y estructurados.

Las principales características que definen el Big Data pueden resumirse en lo que se conoce como las siete Vs: Volumen, Velocidad, Variedad, Veracidad, Valor, Variabilidad y Visibilidad.

La información que se recopila en salud proviene de fuentes de información como la web y las redes sociales, los datos de máquina a máquina, las grandes transacciones de datos, los datos biométricos y los datos generados por los seres humanos (Expedientes Clínicos Electrónicos).

Las técnicas conocidas como Big Data Analysis (BDA) y Machine Learning (ML) pueden permitir en el futuro mejorar nuestra capacidad de investigación clínica y dirigir de manera más precisa las terapias a los pacientes.



Uno de los principales retos, para poder aprovechar la nueva metodología del Big Data, es la formación nuevos recursos humanos (personal especializado con conocimientos clínicos y entrenamiento en inteligencia artificial) y tecnológicos.

La computación cuántica y el Big Data aplicado a la medicina constituirán una nueva herramienta de ayuda en el manejo de las patologías y en la obtención de nuevos conocimientos.

REFERENCIAS

1. Leal, Silvia (2015) E-Renovarse o morir: 7 Tendencias tecnológicas para convertirte en un líder digital. LID Editorial.
2. Diebold, F. X. (2012). On the Origin (s) and Development of the Term 'Big Data'.
3. Pulido Cañabate, E. (2016). Big data: ¿Solución o problema? Encuentros Multidisciplinares.
4. Pucheta, L. L. (2017). Big Data y su impacto en el ámbito de la salud. Reporte del Comité Internacional de Bioética (UNESCO).
5. Peña-López, I. (2017). Big Data en salud digital.
6. Menasalvas, E., Gonzalo, C., & RODRÍGUEZ-GONZÁLEZ, A. (2017). Big Data en Salud: retos y oportunidades. Economía Industrial (Ejemplar dedicado a: Nuevas tecnologías digitales), 405, 87-97.
7. Puyol Moreno, J. (2014). Una aproximación a Big Data= An approach to Big Data.
8. Verduzco Reyes, G., Bautista Thompson, E., Ruiz Vanoye, J. A., & Fuentes Penna, A. (2017). Modelos de tecnologías del Big Data Analytics y su aplicación en salud.
9. Monleón-Getino, A. (2015). El impacto del Big-data en la Sociedad de la Información. Significado y utilidad. Historia y comunicación social, 20(2), 427.
10. Arcila-Calderón, C., Barbosa-Caro, E., & Cabezuelo-Lorenzo, F. (2016). Técnicas big data: análisis de textos a gran escala para la investigación científica y periodística. El profesional de la información, 25(4), 623-631.
11. Reiz, A. N., de la Hoz, M. A., & García, M. S. (2019). Big Data Analysis y Machine Learning en medicina intensiva. Medicina Intensiva, 43(7), 416-426.
12. Pérez, G. (2016). Peligros del uso de los big data en la investigación en salud pública y en epidemiología. Gaceta Sanitaria, 30(1), 66-68.
13. Lecuona, I. D. (2019). Evaluación de los aspectos metodológicos, éticos, legales y sociales de proyectos de investigación en salud con datos masivos (big data). Gaceta Sanitaria, 32, 576-578.
14. Andía, M. E., Arrieta, C., & Long, C. A. S. (2019). Una guía conceptual para usar y entender Big Data en la investigación clínica. Revista Médica Clínica Las Condes, 30(1), 83-94.
15. Martínez Sesmero, J. M. (2015). "Big Data"; aplicación y utilidad para el sistema sanitario. Farmacia Hospitalaria, 39(2), 69-70.
16. Zablah, I., & Urquía-Osorio, H. (2016). Las Ciencias de la Salud y la revolución del BigData. Revista del Cuerpo Médico del HNAAA, 9(2), 88-89.
17. Med, N. C., & de Teresa Galván, E. (2019). Big Data y Medicina.