

# Estudio comparativo de plataformas y herramientas de tratamiento de sistemas Big Data en Telemedicina

Y. Noval García<sup>1</sup>, K. Espinosa Acereda<sup>2</sup>

<sup>1</sup> Universidad del País Vasco, Bilbao, España, ynoval001@ikasle.ehu.eus

<sup>2</sup> Ingeniería de Comunicaciones, Universidad del País Vasco, Bilbao, España, koldo.espinosa@ehu.eus

## Resumen

*En el presente artículo vamos a analizar qué es el Big Data y las posibilidades que ofrece en Telemedicina, tratando de explicar ambos conceptos y su relación. Asimismo, estudiaremos diferentes herramientas de Big Data para el sector sanitario que existen en la actualidad, lo que nos dará una visión de cómo se encuentra ahora mismo de desarrollado, a qué áreas está enfocado y a qué público va dirigido. Además, queremos observar las limitaciones que tiene y cómo se pueden sortear. Observaremos que el Big Data abre ante nosotros nuevos horizontes para la medicina. La información que hasta ahora sólo nos indicaba cómo habían sido las cosas en el pasado, puede ayudarnos a predecir comportamientos, a encontrar nuevos patrones y, en definitiva, a indicarnos nuevas maneras de ver el futuro.*

## 1. Introducción

En 2008 llamó la atención una noticia relacionada con Google y su capacidad de predecir la gripe. El estudio se basaba en la comparativa de cifras de afectados por la gripe en los inviernos anteriores y las búsquedas realizadas en Google en esas zonas. Se observó que allí donde había epidemia de gripe, las palabras y expresiones más utilizadas eran 45, de las cuales la mayoría eran términos médicos relacionados con la enfermedad [1].

Asumieron que si había una relación directa entre las búsquedas en Google y las zonas más afectadas por la gripe en años anteriores, la habría también ese invierno. Así es como *Google Flu Trends* fue capaz de anticipar y analizar en tiempo real dónde la gripe estaba incidiendo más y cómo era la evolución, minuto a minuto, dejando, de esta manera, muy atrás las estadísticas del Centro de Control de Enfermedades (CDC) del país, que tarda unas dos semanas en procesar los resultados de las encuestas hechas en los diferentes centros de salud.

Este es un ejemplo llamativo del uso que se puede hacer de la información y de la capacidad que tienen los datos, muchas veces poco relevantes, para tras juntarse, analizarse y relacionarse unos con otros, obtener conclusiones o razonamientos.

Este artículo trata de explicar qué es el Big Data y cómo esa información y su análisis pueden ser utilizados en medicina. Para ello, se repasan varias herramientas y se realiza una comparativa entre ellas, basándonos en las diferentes áreas del sector sanitario que cubren y en cómo lo hacen, o hasta dónde llega su alcance y las limitaciones que tienen.

## 2. Estado del arte

### 2.1. Telemedicina

La Telemedicina es la atención médica o el intercambio de información médica mediante el uso de las telecomunicaciones o sistemas de información, es decir, a distancia. Esto incluye no sólo prestación de la asistencia médica (consulta, diagnóstico y tratamiento), también la enseñanza y el intercambio de datos médicos [2].

El no encontrarse presencialmente en un lugar para diagnosticar a alguien, el poder monitorizar a un paciente sin que éste se encuentre hospitalizado, el poder aprender y practicar un procedimiento como si estuviera el paciente, el intercambio de información desde diferentes partes del mundo, todas estas posibilidades son hoy reales gracias a la Telemedicina.

### 2.2. Big Data

Existen diferentes definiciones sobre Big Data, pero la más aceptada es que llamamos Big Data a la gestión y análisis de enormes volúmenes de datos, que se producen y procesan a gran velocidad (incluso en tiempo real) y cuya variedad de formatos es enorme, desde una publicación en una red social, hasta una señal GPS. En lo que se conoce como las “3 v” del Big Data [3,4].

Otro concepto a tener en cuenta es el análisis predictivo que forma parte del Big Data, pues consiste en obtener mediante el análisis de grandes cantidades de datos, nuevas relaciones, patrones o tendencias que nos ayuden a predecir el comportamiento futuro de aquello que queremos conocer [5,6].

En el sector sanitario cada día se recoge y almacena más información, desde investigaciones con nuevos fármacos o terapias, hasta los historiales médicos de los pacientes. El poder centralizar esa información, analizarla y monitorizarla en tiempo real nos supondría un avance significativo en muchos campos, pues nos permitiría ver correlaciones que a simple vista no se observan. Para ello, pacientes, profesionales, población, instituciones, todo debe estar conectado.

Este potencial del Big Data supondría la llegada de la Medicina 5P que consiste en una medicina Personalizada, Predictiva, Preventiva, Participativa y Poblacional [7].

### 3. Herramientas de Big Data en Telemedicina

En este apartado vamos a focalizarnos en algunas de las herramientas de tratamiento de sistemas Big Data que existen actualmente para Telemedicina y en qué consiste cada una, para qué se utiliza y cómo ha ayudado a mejorar lo que existía anteriormente. Se ha tratado de abarcar todas las áreas del sector sanitario. Tras ello, podremos realizar un estudio comparativo sobre las mismas.

#### 3.1. PatientsLikeMe

Es, al mismo tiempo, una red de pacientes y una plataforma web de investigación que alberga los datos y tratamientos de más de 500.000 personas con más de 2.700 enfermedades diferentes [8].

Sus fundadores, los hermanos Jamie y Benjamin Heywood, trataban de poner en marcha una plataforma de intercambio de información sobre pacientes con ELA, donde pudieran hablar de sus tratamientos y contar sus historias. Se ha convertido en mucho más que eso. Además de recoger toda la información que los pacientes van intercambiando, la clasifica en diferentes grupos, de tal manera que se puede conocer el tratamiento y evolución de un paciente a través del tiempo.

En el momento en que la gente empezó a compartir toda esta información, se dieron cuenta del poder que esto tenía y de las amplias posibilidades si llegaba a las manos adecuadas, y así desarrollaron su modelo de negocio, vendiendo esta información a diferentes compañías, desde farmacéuticas a centros de investigación de alguna enfermedad o universidades. El objetivo es crear la mayor red de pacientes del mundo, pues esta información anónima supone una enorme fuente de datos en tiempo real para cualquier estudio que se quiera realizar, datos que, de otra manera, costaría mucho dinero y esfuerzo poder reunir.

#### 3.2. MedISys

Se trata de una herramienta de monitorización que busca, analiza y monitoriza alrededor de 900 webs especializadas en salud y más de 20.000 páginas relacionadas con unos 7.000 portales generales y 20 de noticias; todo esto en 70 idiomas. La idea es identificar rápidamente cualquier amenaza para la salud pública, desde enfermedades a plagas, pasando por sustancias psicoactivas [9,10].

Además del análisis, muestra la información en tiempo real, a través de informes, disponibles en 42 idiomas, que pueden mostrar diferentes temas o fuentes de información según el usuario considere oportuno.

#### 3.3. EuResist

La enfermedad del VIH tiene, actualmente, diferentes tratamientos que conllevan el uso de numerosos fármacos. Se pretende adaptar la medicación al paciente, pues las respuestas de unos y otros son muy diferentes. Por ello, resulta muy complicado acertar a la primera con el tratamiento más adecuado para cada enfermo. Existe además la dificultad añadida de que el cuerpo se va

haciendo, poco a poco, inmune a la medicación, por lo que cada cierto tiempo es necesario cambiarla [11].

EuResist surgió por la necesidad de poder determinar el mejor tratamiento contra el VIH para cada paciente. Se vio la posibilidad de que el Big Data pudiera ayudar a la toma de esta decisión y así mejorar la calidad de vida del enfermo desde el principio.

Lo primero que se llevó a cabo fue la unión de bases de datos de VIH de Alemania, Italia y Suecia con un total de más de 62.000 pacientes. Después, con la ayuda de la Unión Europea y bajo el nombre del proyecto *The EuResist Network GEIE* se desarrolló el sistema capaz de analizar esa información y decidir qué tratamiento es el más adecuado. En la actualidad, el porcentaje de aciertos es superior al 77% de los casos [12].

#### 3.4. WuXi NextCODE

Se trata de una compañía que utiliza Big Data en información genómica para ayudar a realizar investigaciones que mejoren la salud de las personas [13]. Cuentan con una plataforma capaz de organizar, analizar y compartir esa información, y es la única con información a nivel de población y no de personas.

#### 3.5. Predixion software

Se trata de un software que se instala en diferentes aparatos médicos, desde ultrasonidos a medidores de glucosa, y que monitoriza en tiempo real el estado de los mismos. Además, utiliza la tecnología Big Data para analizar ese estado y predecir cualquier anomalía que pueda afectar al estado de la máquina [14].

La tecnología *Predixion* permite acceder al estado de cada aparato desde una pantalla centralizada, de manera que en todo momento podemos ver si funciona correctamente. De esta manera, se permite planificar las intervenciones médicas en función de información real sobre las máquinas disponibles. También permite observar con antelación si existe algún síntoma de problema, pudiendo darle solución incluso antes de que ocurra, y conocer el rendimiento del aparato.

#### 3.6. Practice Fusion

Es una compañía estadounidense que nació para ofrecer a centros de salud y hospitales la gestión y el almacenamiento de los historiales médicos de los pacientes. Actualmente, es la mayor plataforma de este tipo en EEUU, contando con más de 38 millones de historiales que representan alrededor del 6.7% del total. También cuenta con la integración de este sistema en numerosas áreas como laboratorios o centros de procesado de imágenes [15].

La plataforma almacena los historiales y analiza los datos que hay en ellos, creando alertas o recomendaciones en tiempo real que ayudan a los doctores en la toma de decisiones. Ofrecen para ello una interfaz de comunicación con los profesionales sencilla e intuitiva, teniendo otra, también, para los pacientes. Ambos pueden hacer un seguimiento de las visitas, tratamientos, enfermedades, etc. que han ido teniendo.

### 3.7. Verily

Es una compañía que desarrolla herramientas para recoger y organizar información relacionada con la salud [16]. Cuentan con numerosos proyectos relacionados con Big Data, entre los que podemos destacar *Onduo*, una herramienta colaborativa entre *Sanofi* y *Verily* que busca mejorar la calidad de vida de los pacientes con diabetes tipo 2.

### 3.8. IBM Watson Health

Se trata de un sistema capaz de *entender* todos los tipos de información como lenguaje natural, imágenes, texto, incluso sin estructura, o audio. Además, es capaz de *razonar* y dar recomendaciones, basadas no sólo en la evidencia, también en el todo o personalidad de la persona que lo pide. Es también capaz de aprender automáticamente, a través de numerosos algoritmos, y puede interactuar con seres humanos.

Pensando en lo útil que serían todas estas capacidades en otras disciplinas como la salud, se empezó a colaborar con diferentes universidades con el objetivo de aprovechar *Watson* para recolectar y analizar toda la información que generamos como pacientes y profesionales para poder ofrecer mejores soluciones, incluso en tiempo real [17].

El propósito de la colaboración entre *Watson* y diferentes instituciones investigadoras sobre el cáncer fue mejorar y acelerar la capacidad de los profesionales para emitir un diagnóstico. Esta nueva manera de procesar la información permite reducir los tiempos de semanas a minutos, con lo que eso conlleva para la salud y el bienestar de los pacientes [18].

## 4. Comparativa

### 4.1. Áreas del sector sanitario

La gestión de hospitales y centros de salud cubre un amplio espectro, desde los procesos internos del hospital, pasando por la optimización de recursos hasta los sistemas informáticos. Debido a la mayor demanda sanitaria y a la falta de más recursos que anteriormente, la gestión sanitaria se ha vuelto un reto aun mayor. Hemos tratado en el apartado anterior dos herramientas de Big Data que forman parte de esta área, *Practice Fusion* y *Predixion software*. Estas herramientas pertenecen a diferentes áreas del proceso de gestión y actúan de manera independiente; sin embargo, ambas forman parte de un todo llamado gestión sanitaria. Durante esta investigación, se ha echado en falta el desarrollo de un plan de implantación de Big Data, donde se identifiquen las diferentes áreas y se vayan abordando una a una, haciendo que cada pieza forme parte de un todo.

El diagnóstico y la ayuda en la toma de decisiones cuentan con numerosas herramientas de Big Data, *PatientsLikeMe*, *Practice Fusion*, *WuXi NextCODE* e *IBM Watson Health*.

Las investigaciones médicas y farmacología pueden hacer uso de tres herramientas, *PatientsLikeMe*, *WuXi NextCODE* e *IBM Watson Health*.

La prevención de enfermedades gracias a los hábitos saludables o la identificación de poblaciones de riesgo, cuenta con *Practice Fusion* y *Verily* como herramientas que aportan valor.

Herramientas como *MedISys* permiten identificar y monitorizar situaciones de alerta médica. Como cualquier herramienta basada en crowdsourcing, la veracidad de la información y su utilidad deberán ser contrastadas para poder garantizar la utilidad de la información final inferida. Para ello, resulta fundamental disponer de información de fuentes oficiales, por ejemplo, de las agencias de salud de los propios estados miembro.

Podemos decir que, todas las herramientas forman parte de este nuevo conocimiento, más generalizado y al alcance de una mayoría. Aun así, seguimos sufriendo la descentralización del mismo, vemos que hay herramientas muy concretas que proporcionan información sobre enfermedades concretas, haciendo que los profesionales deban utilizar todas y cada una de ellas.

Durante la investigación hemos echado en falta herramientas de Big Data que ayuden en la monitorización de pacientes y enfermedades crónicas. Sabemos que existen numerosas herramientas de monitorización, la mayoría basadas en lo que hoy llamamos Internet de las Cosas, es decir, en dar a los pacientes autonomía para estar en sus casas y poner los aparatos necesarios, bien en su cuerpo, bien en su entorno, para poder monitorizarles, siendo estos capaces de avisar a los profesionales si algo ocurre. Toda esa información sobre el paciente y su entorno que resultaría tan útil para entender su evolución y las causas de lo que le ocurre, no está siendo analizada, y ahí es donde entra en juego el máximo potencial de la monitorización, porque no sólo puedes conocer el estado del paciente y anticiparte a que algo ocurra, puedes utilizar toda esa información para *aprender*, para tomar decisiones y para acceder a nuevo conocimiento.

### 4.2. Público

La mayoría de las herramientas están pensadas para profesionales de la salud, y abarcan diferentes grupos como Instituciones públicas, Gestores de hospitales o centros de salud, Investigadores o Profesionales médicos. Algunas de estas herramientas, como *Practice Fusion* y *Verily*, tienen una interfaz de usuario, permitiendo que el propio paciente vea su evolución, gestione sus citas o se le hagan recomendaciones sobre hábitos de vida saludables. Otras, como *PatientsLikeMe*, están pensadas directamente para los pacientes.

### 4.3. Almacenamiento de la información

Al hablar de todas estas herramientas, la información que contienen y analizan, surge la pregunta de ¿dónde está esa información? Y todas ellas tienen, casualmente, la misma respuesta, *en la nube*. Con ello nos referimos a que se encuentran almacenados en centros de datos, accesibles desde cualquier parte y cualquier ordenador. Estos centros de datos pueden ser utilizados en exclusiva por la compañía, o pueden ser compartidos. En ambos casos, la seguridad está garantizada y únicamente el que debe

acceder a ello, puede hacerlo. Esto ofrece una capacidad mayor de almacenamiento y grandes recursos [19,20].

## 5. Conclusiones

Podemos afirmar que la evolución del sector sanitario pasa por unirse a las nuevas tecnologías como Big Data. Esta tecnología ofrece la posibilidad de procesar y analizar grandes volúmenes de información en tiempo real, obteniendo de ella un potencial hasta ahora inimaginable. Además, ofrece la posibilidad de, no sólo utilizar la información almacenada en la institución de salud, sino también toda aquella que se encuentre en otras herramientas o en la web, como noticias o artículos.

Si añadimos el análisis predictivo a las posibilidades de Big Data, nuevos horizontes se abren ante nosotros. La información que hasta ahora sólo nos indicaba cómo habían sido las cosas en el pasado, puede ayudar a predecir comportamientos, a sacar nuevos patrones o correlaciones, e indicarnos nuevas pistas para el futuro.

Las herramientas utilizadas actualmente en Telemedicina son diversas y abarcan casi todas las áreas de la salud: unas son capaces de monitorizar la salud pública; otras ayudan a la gestión de instituciones sanitarias; las hay que ofrecen nuevas posibilidades a las investigaciones científicas e incluso las que ayudan a prevenir enfermedades y a los profesionales en el diagnóstico. Las posibilidades son inmensas y todas ellas colaboran, desde diferentes aspectos y funciones, a la mejora del conocimiento, acercándonos a la sanidad del futuro, personalizada, predictiva, preventiva, participativa y poblacional.

Asimismo, se deben afrontar una serie de retos. La adopción de estos nuevos sistemas y la relación entre ellos conllevan una inversión inicial de capital, no siempre al alcance, y suponen nuevas maneras de trabajar, pues no sólo profesionales sanitarios van a estar envueltos en estas tecnologías. También debemos tener en cuenta la dificultad de discernir entre la información veraz y la que no lo es, algo siempre complicado al trabajar con volúmenes y variedad de datos tan elevados. Por último, se debe asegurar la privacidad y confidencialidad de la información, requisitos indispensables al manejar información personal. En definitiva, se necesita una planificación a largo plazo, para poder ir consiguiendo poco a poco los objetivos; se necesita ofrecer información y formación a los diferentes agentes involucrados; se necesita desarrollar la tecnología de seguridad necesaria y adaptar las expectativas a la realidad de esta nueva tecnología.

A pesar de los retos, la rentabilidad a largo plazo está asegurada, se calcula que más de 100 billones de € [21] se podrían ahorrar sólo en mejora de la eficiencia de procesos administrativos. Es la única manera de poder seguir teniendo un sistema sostenible a largo plazo.

## Referencias

- [1] Artículo en el periódico la Vanguardia: <http://www.lavanguardia.com/vida/20081123/53585155233/google-ya-sabe-que-tendras-la-gripe.html> (Consultado: Agosto 2017).
- [2] Artículo en la página web de World Health Organization: [http://apps.who.int/iris/bitstream/10665/194008/1/EB99\\_30\\_spa.pdf](http://apps.who.int/iris/bitstream/10665/194008/1/EB99_30_spa.pdf) (Consultado: Agosto 2017).
- [3] McAfee A, Brynjolfsson E. Big Data: The Management Revolution. *Harvard Business Review*. Octubre 2012
- [4] Luo et al. Big Data Application in Biomedical Research and Health Care: A Literature Review. *Biomedical Informatics Insights*, 2016:8 1–10.
- [5] Kruse, C. S., Goswamy, R., Raval, Y., & Marawi, S. Challenges and Opportunities of Big Data in Health Care: A Systematic Review. *JMIR Med Inform*. 2016; 4(4).
- [6] Artículo en la página web de Machine learning mastery: <https://machinelearningmastery.com/gentle-introduction-to-predictive-modeling/> (Consultado: Agosto 2017).
- [7] Blog de Deusto: <https://blogs.deusto.es/bigdata/la-medicina-5p-sanidad-y-big-data/> (Consultada: Agosto 2017).
- [8] Página web de PatientsLikeMe: <https://www.patientslikeme.com> (Consultada: Agosto 2017).
- [9] Entrada en la página web de Medical Information System: <https://ec.europa.eu/jrc/en/scientific-tool/medical-information-system> (Consultada: Agosto 2017).
- [10] Rortais, A., Belyaeva, J., Gemo M., Van der Goot, E., Linge, J.P. MediSys: An early-warning system for the detection of (re-)emerging food and feed-borne hazards. *Food Research International*. Elsevier. Volume 43, Issue 5, June 2010, Pages 1553-1556.
- [11] Página web de EuResist: <https://www.euresist.org> (Consultada: Agosto 2017).
- [12] Artículo en la página web de European Commission: [https://ec.europa.eu/research/sme-techweb/pdf/success\\_stories\\_2013.pdf](https://ec.europa.eu/research/sme-techweb/pdf/success_stories_2013.pdf) (Consultado: Agosto 2017).
- [13] Página web de WuXi NextCODE: <https://www.wuxinextcode.com/about-us/> (Consultada: Agosto 2017).
- [14] Página web de Predixion software: <https://greenwavesystems.com> (Consultada: Agosto 2017).
- [15] Página web de Practice Fusion: <http://www.practicefusion.com> (Consultada: Agosto 2017).
- [16] Página web de Verily: <https://verily.com> (Consultada: Agosto 2017).
- [17] Artículo en la página web de IBM: <http://www-03.ibm.com/press/us/en/pressrelease/46748.wss> (Consultado: Agosto 2017).
- [18] Artículo en la página web de Mediatrends: <https://www.mediatrends.es/a/119232/ibm-watson-que-es-ia-inteligencia-artificial/> (Consultado: Agosto 2017).
- [19] Artículo en la página web del Instituto de ingeniería del conocimiento: <http://www.iic.uam.es/innovacion/big-data-la-nube-servicios-cloud/> (Consultado: Agosto 2017).
- [20] Vatsalan, D, Sehili, Z, Christen, P, Rahm, E. Privacy-Preserving Record Linkage for Big Data: Current Approaches and Research Challenges. *Handbook of Big Data Technologies*. 2017
- [21] Artículo en la página web de McKinsey: <http://www.mckinsey.com/business-functions/digital-mckinsey/our-insights/big-data-the-next-frontier-for-innovation> (Consultado: Agosto 2017).