



**Nuevas oportunidades  
de negocio para emprender  
basadas en *big data***



Andalucía Emprende, Fundación Pública Andaluza  
**CONSEJERÍA DE CONOCIMIENTO, INVESTIGACIÓN Y UNIVERSIDAD**



## **Colección**

Estudios de Investigación

## **Edición y dirección**

Área de Conocimiento y Estrategia

Dirección de Cultura Emprendedora, Innovación y Conocimiento

**Andalucía Emprende, Fundación Pública Andaluza**

**Consejería de Conocimiento, Investigación y Universidad**

**JUNTA DE ANDALUCÍA**

## **Realizado por:**

E3 Evolución Pymes S.L.

## **Consultores**

Saturnino Jiménez Aguilar

Natalia Palomino González

Andalucía Emprende, Fundación Pública Andaluza, no se hace responsable de las opiniones recogidas en este documento, siendo autores los responsables de los análisis y resultados obtenidos del trabajo desarrollado.

Finalizado en febrero de 2018

## INDICE

<b>BLOQUE I: MARCO CONCEPTUAL DEL BIG DATA .....</b>	<b>5</b>
1.1. Introducción.....	5
1.2. Nacimiento y evolución del big data.....	7
1.3. Principales características.....	8
1.4. Mapa de tecnologías.....	9
1.4.1. Tipos de información .....	9
1.4.2. Conceptos tecnológicos relacionados .....	11
1.4.3. Ecosistema del big data (landscape 2017).....	12
1.4.4. Arquitectura Lambda.....	17
1.4.5. Tecnologías y herramientas .....	18
1.5. Mapa de servicios: Ventajas del big data para las empresas .....	200
<b>BLOQUE II: ANALISIS Y DIAGNÓSTICO DE LOS NEGOCIOS Y EXPERIENCIAS EMPRESARIALES EN BIG DATA.....</b>	<b>22</b>
2.1. Mapa de sectores y actividades para sacar partido al big data.....	22
2.2. Estado y evolución del big data .....	25
2.2.1. Contexto internacional.....	25
2.2.2. Contexto nacional.....	27
2.2.3. Contexto andaluz .....	29
2.3. Identificación de experiencias empresariales con big data.....	31
2.3.1. Contexto internacional.....	31
2.3.2. Contexto nacional.....	35
2.3.3. Contexto andaluz .....	40
<b>BLOQUE III: INNOVAR Y EMPRENDER CON BIG DATA.....</b>	<b>43</b>
3.1. Análisis DAFO .....	43
3.2. Factores para emprender con el uso de big data.....	46
3.2.1. Factores tecnológicos .....	48
3.2.2. Capacitación y formación.....	50
3.2.3. Información y conocimiento acerca del big data.....	52
3.2.4. Ciberseguridad.....	54
3.3. Nuevos nichos de mercado: el enfoque sectorial.....	54
3.4. Experiencias internacionales para fomentar el uso del big data .....	61



## **BLOQUE IV: ANÁLISIS DEL BIG DATA EN ANDALUCÍA ..... 67**

---

<b>4.1. Experiencias empresariales en el uso de big data: casos de éxito en Andalucía.....</b>	<b>67</b>
<b>4.2. Nuevos nichos de mercado para el uso del big data en Andalucía: energías renovables, salud, sector agro y turismo. ....</b>	<b>68</b>
<b>4.2.1. Sector Agro .....</b>	<b>69</b>
<b>4.2.2. Sector de energías renovables.....</b>	<b>70</b>
<b>4.2.3. Sector salud.....</b>	<b>71</b>
<b>4.2.4. Sector Turístico.....</b>	<b>72</b>
<b>4.3. Propuestas de actuaciones a desarrollar.....</b>	<b>75</b>
<b>4.3.1. Acciones de comunicación, divulgación y formación.....</b>	<b>76</b>
<b>4.3.2. Financiación.....</b>	<b>77</b>
<b>4.3.3. Desarrollo sectorial .....</b>	<b>78</b>
<b>4.3.4. Conclusiones:.....</b>	<b>79</b>

## **ANEXO 1: RESULTADOS DE ENCUESTAS Y ENTREVISTAS. .... 80**

## **ANEXO 2: BIBLIOGRAFIA ..... 83**

## **ANEXO 3: GLOSARIO ..... 87**

---



## BLOQUE I: MARCO CONCEPTUAL DEL BIG DATA

### 1.1. Introducción

En los últimos años se asiste a una auténtica revolución tecnológica que está cambiando la manera de actuar y relacionarse con el mundo. El nacimiento de las nuevas tecnologías ha convertido al ser humano en un sujeto activo, que necesita interactuar con su entorno. Esto, unido a la globalización, transforma el concepto hombre-sociedad.

La sociedad offline era lo que imperaba hasta el momento, pero gracias al nacimiento y proliferación de Internet, la sociedad online ha ganado terreno. Castells, M. (2001) denomina a esta metamorfosis de la sociedad, “sociabilidad online” y la define a continuación:

Lo más interesante es la idea de que son comunidades personales, comunidades de personas basadas en los intereses individuales y en las afinidades y valores de las personas. Es decir, en la medida en que se desarrollan en nuestras sociedades proyectos individuales, proyectos de dar sentido a la vida a partir de lo que yo soy y quiero ser, Internet permite esa conexión saltando por encima de los límites físicos de lo cotidiano, tanto en el lugar de residencia como en el lugar de trabajo y genera, por tanto, redes de afinidades.<sup>1</sup>

Efecto directo de esto es la explosión de información de amplio espectro. Esta transformación de la información ha sido denominada por García Carrero, J. (2013) como “infoexplosión” (explosión de la información junto a la situación económica y de negocio que obligan a las organizaciones a redefinir sus estrategias de gestión de la información)


En la ‘era post-PC` la información ha crecido de manera exponencial; también ha cambiado la ubicuidad del acceso, la generación y la gestión de los contenidos. El acceso se produce a través de cualquier dispositivo, desde cualquier lugar y en cualquier momento y como expone García Cantero, J. (2012): “tenemos la capacidad de generar, reproducir y etiquetar contenido en la palma de nuestra mano y de distribuirlo globalmente con el acceso ubicuo y permanente a la Red”.

El nacimiento y potencial de la nube es otro de los aspectos a tener en cuenta para contextualizar el big data. Al igual que la información se genera y se accede a ella desde cualquier parte, también se almacena y gestiona desde cualquier lugar. Hoy la mayoría de la información es no estructurada y una gran parte de la misma es gestada por los usuarios finales.

Ante tal magnitud de información -solamente en los dos últimos años se ha generado el 90% de los datos que existen en el mundo- las empresas y organizaciones públicas y privadas tienen como reto reflexionar sobre qué papel juega la información en sus entidades, proporcionando estrategias que favorezcan la gestión de su actividad. Ya sea en la medicina, en las multinacionales o en la Política. El manejo inteligente del gran volumen de datos precisa del big data para transformar la información en conocimiento y el conocimiento en decisiones acertadas.

---

<sup>1</sup> Castells, M. (2001): Lección inaugural del programa de doctorado sobre la sociedad de la información y el conocimiento (UOC). <http://tecnologiaedu.us.es/cuestionario/bibliovir/106.pdf>.



Después de que se nombrara el término big data por primera vez en un artículo de los investigadores de la NASA, Michael Cox y David Ellsworth en 1997, han sido numerosas las definiciones que se han dado del concepto.

La Search Data Center define el big data como “un término evolutivo que describe cualquier cantidad voluminosa de datos estructurados, semiestructurados y no estructurados que tienen el potencial de ser extraídos para obtener información”.<sup>2</sup>

Por su parte, Gartner lo describe como “un activo de información de gran volumen, alta velocidad y/o alta variedad que exige formas innovadoras y rentables de procesamiento de información que permiten una mejor comprensión, toma de decisiones y automatización de procesos”.<sup>3</sup>

José Carlos López López, director de operaciones de IMC Group, lo define a continuación:

“Denominamos big data a la gestión y análisis de enormes volúmenes de datos que no pueden ser tratados de manera convencional, ya que superan los límites y capacidades de las herramientas de software habitualmente utilizadas para la captura, gestión y procesamiento de datos. Dicho concepto engloba infraestructuras, tecnologías y servicios que han sido creados para dar solución al procesamiento de enormes conjuntos de datos estructurados, no estructurados o semi-estructurados (mensajes en redes sociales, señales de móvil, archivos de audio, sensores, imágenes digitales, datos de formularios, emails, datos de encuestas, logs etc.) que pueden provenir de sensores, micrófonos, cámaras, escáneres médicos e imágenes”.<sup>4</sup>

Mike Gualtieri, vicepresidente y principal analista de Forrester, determina que el big data es “la solución al crecimiento exponencial de los datos, en el momento en que se hace difícil su administración con respecto al almacenamiento, procesamiento y acceso”.<sup>5</sup>

En 2012, IBM Institute for Business Value llevó a cabo un estudio, con la colaboración de la Escuela de Negocios Saïd en la Universidad de Oxford, para determinar cuáles serían las características sobre big data que les resultaban más interesantes a las empresas.

El trabajo abarcó a más de 1144 negocios y profesionales de las Tecnologías de la Información y su extensión alcanzó los 95 países. A continuación, en la Figura 1, se expone el resultado de dicha investigación.

---

<sup>2</sup> <http://searchdatacenter.techtarget.com/es/definicion/big-data>

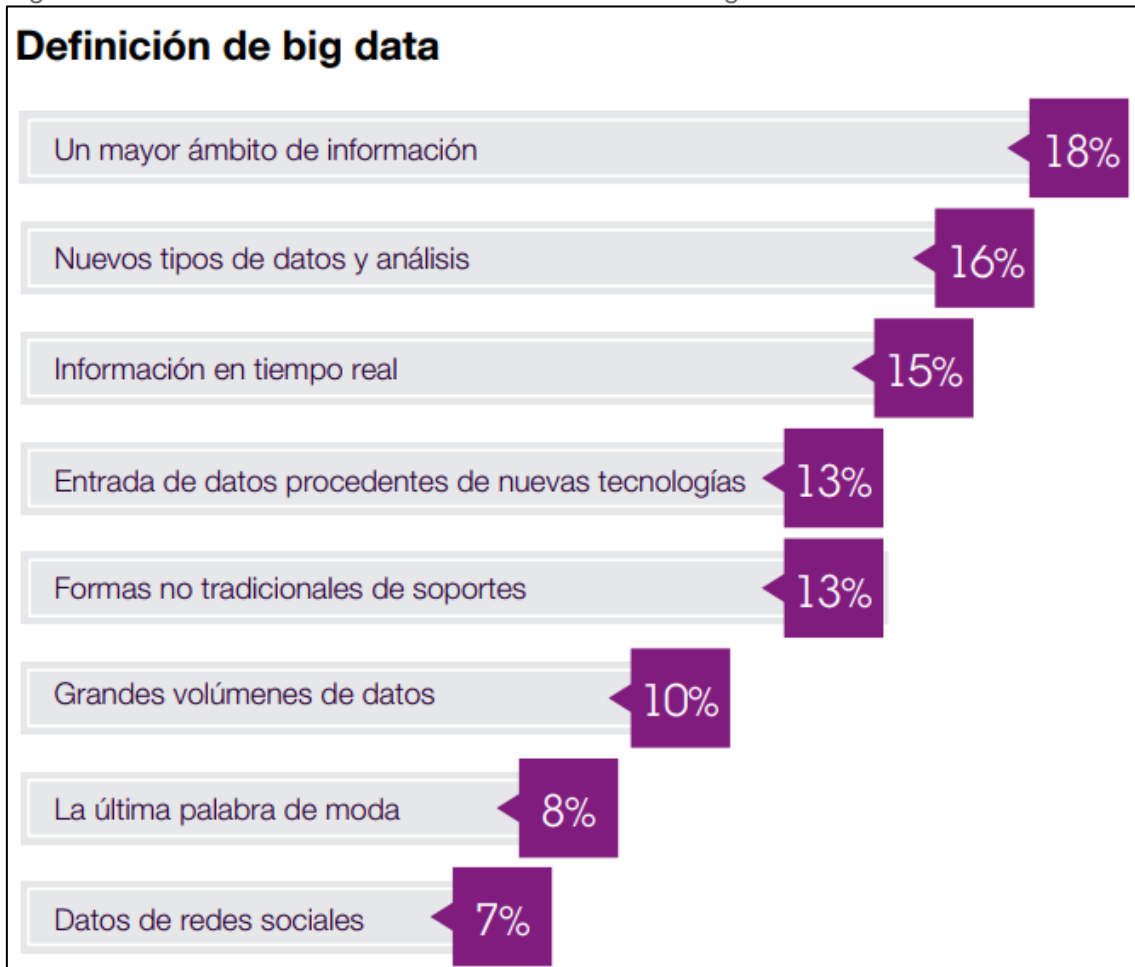
<sup>3</sup> <https://research.gartner.com/definition-what-is-big-data?resId=3002918&srcId=1-8163325102>.

<sup>4</sup> <http://www.eleconomista.es/tecnologia/noticias/5578707/02/14/La-moda-del-big-Data-En-que-consiste-en-realidad.html>

<sup>5</sup> <https://go.forrester.com/blogs/12-12-05-the-pragmatic-definition-of-big-data/>.



Figura 1: Resultados de la encuesta de Definición de big data.



Fuente: IBM Institute for Business Value.

## 1.2. Nacimiento y evolución del big data.

La idea de sobrecarga de información apareció por primera vez en la década de 1930. El crecimiento exponencial de la población, con la consecuente multiplicación de registro de datos, desembocó en la creación de un nuevo sistema registral más exhaustivo.

Esto supuso un problema para las bibliotecas, ya que se quedaban sin espacio para almacenar la información. Por esto, y durante años posteriores, las empresas y organizaciones comenzaron a diseñar e implantar sistemas computacionales que les facilitaban automatizar sus inventarios y, al mismo tiempo, mejorar su estrategia empresarial.

Una vez que nacen las bases de datos relacionales surge la Inteligencia Empresarial (BI) y es con el nacimiento de Internet y la web 2.0., cuando surge el concepto y desarrollo del big data. A continuación se muestra, en la Figura 2, un breve cronograma que resume el nacimiento del big data:





Figura 2: Cuadro cronológico de nacimiento del big data.

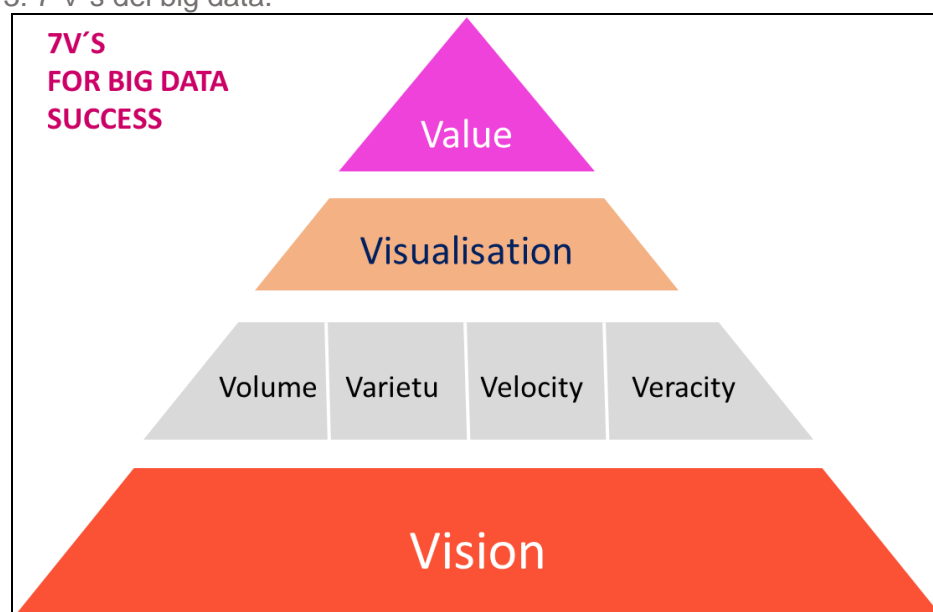
Año	Hito histórico
1970	Nacimiento de la base de datos relacional
1976	Sistemas de planificación de necesidades de material (MRP).
1980	La Ley de datos de Parkinson
1985	Sistemas de planificación de recursos de fabricación (MRP II).
1989	Inteligencia Empresarial (BI) de Dresner
1992	Crystal Reports emite el primer informe de base de datos.
1996	Inteligencia empresarial (BI) 2.0.
1997	Surge el término "big data"
1999	Internet de las cosas (IOT)
2001	Software como servicio (SAAS)
2002	Sistemas ERP extendidos o ampliados
2005	Nacimiento de la Web 2.0.
2006	Creación de Hadoop
2009	Linked Data para la interconexión
2010	Aparición del ERP en la nube
2011	IBM lanza el tweet #IBMbigdata
2012	Lanzamiento del nuevo sistema IPV6
2013	Tecnología de base de datos SAP HANA
2014	Adopción de ERP en la nube
2015	Nacen las Smart Cities

Fuente: elaboración propia.

### 1.3. Principales características


Una vez definido qué es el big data se profundiza en el concepto a través de las 7 v's que lo caracterizan: Volumen, Velocidad, Variedad de los datos, Veracidad de los datos, Viabilidad, Visualización de los datos y Valor de los datos. IBM fue el encargado de definir tres de ellas, para posteriormente ser ampliadas hasta siete, dependiendo de la acotación documental de las fuentes, como se puede observar en la Figura 3:

Figura 3: 7 V's del big data.



Fuente: <https://www.linkedin.com/pulse/7vs-successful-big-data-project-vit-soupal>




- 
1. Volumen de información: con la llegada de las nuevas tecnologías y la conexión al mundo 2.0. se genera gran cantidad de información de datos. Éstos se crean masivamente cada minuto, almacenando cantidades desorbitadas para procesar dicha información y transformando los datos en acciones.
  2. Velocidad de los datos: en el mundo online la inmediatez es la clave del éxito. Los datos están en constante movimiento debido a las interconexiones que se realizan, permitiendo crearlos, almacenarlos y procesarlos en tiempo real.
  3. Variedad de los datos: son de numerosa procedencia, formas y tipos (documentos de texto, correos electrónicos, datos de sensores, audios, vídeos o imágenes). Pueden ser estructurados, como las bases de datos clásicas, o no estructurados, siendo necesario para éstos últimos poder ser almacenados y procesados sin depender de un pre proceso que los estructure e indexe su contenido.
  4. Veracidad de los datos: en base al grado de fiabilidad de la información recibida. Una de las premisas del big data es utilizar la tecnología para reducir la incertidumbre y conseguir datos de calidad.
  5. Viabilidad: el big data permite conocer la viabilidad de un proyecto de manera inteligente. Profundiza en las relaciones entre las variables ocultas e insta a diseñar estrategias eficaces. Así se pueden presentar interesantes hojas de ruta y planes de negocio fructíferos.
  6. Visualización de los datos: es necesario procesarlos para pasar de tablas y hojas de cálculo a imágenes que sean fácilmente visualizables. Es preciso hacerlos más legibles, accesibles y ágiles y para ello existen herramientas que ayudan a comprenderlos gráfica y contextualmente.
  7. Valor de los datos: radica en la capacidad de transformar datos en información y ésta, a su vez, en decisiones. El objetivo es que los responsables de las empresas puedan tomar la mejor resolución en base a estos datos recopilados.

## 1.4. Mapa de tecnologías

### 1.4.1. Tipos de información

La información relativa al big data procede de varias fuentes, como exponen desde el Grupo TRC -empresa referente en el mercado de las telecomunicaciones-:

- **Personas:** a través de mensajes de texto, vídeos, notas de voz en sus dispositivos móviles, ordenadores...
- **Transacción de big data:** registros de facturación y de llamadas telefónicas.
- **Redes sociales y web:** correos electrónicos, redes sociales, blogs y contenidos de las páginas web.
- **Machine to machine (M2M):** son las tecnologías que comparten datos a través de dispositivos (medidores de temperatura, de altura, presión, química, etc. que transforman información en valores).
- **Biométrica:** los datos biométricos usados en el mundo de la seguridad e inteligencia.



IBM, gracias a su estudio sobre big data, refleja el porcentaje de procedencia de las fuentes de datos en la Figura 12:

Figura 12: procedencia de las fuentes de big data.

Transacciones	88%
Datos de registro	73%
Eventos	59%
Correos electrónicos	57%
Redes sociales	43%
Sensores	42%
Canales externos	42%
Escaneos RFID o datos de TPV	41%
Texto sin información	41%
Geoespacial	40%
Sonido	38%
Imágenes fijas/videos	34%

Fuente: IBM Institute for Business Value.

Al mismo tiempo se clasifican los tipos de datos según su naturaleza en:

- **Datos estructurados**<sup>6</sup>: es información ya procesada, filtrada y con un formato estructurado. Es el tipo de datos más usado en la actualidad.
  - o Creados: datos generados por los sistemas de una manera predefinida (registros en tablas, ficheros XML asociados a un esquema).
  - o Provocados: datos creados de manera indirecta a partir de una acción previa (valoraciones de restaurantes, películas, empresas (Yelp, TripAdvisor, ...)).
  - o Dirigido por transacciones: datos que resultan al finalizar una acción previa de manera correcta (facturas autogeneradas al realizar una compra, recibo de un cajero automático al realizar una retirada de efectivo, ...)
  - o Compilados: resúmenes de datos de empresa, servicios públicos de interés grupal. Entre ellos se encuentra el censo electoral, vehículos matriculados, viviendas públicas, ...).
  - o Experimentales: datos generados como parte de pruebas o simulaciones que permitirán validar si existe una oportunidad de negocio.
- **Datos semi-estructurados o híbridos**<sup>7</sup>: es información procesada y con un formato definido, pero no estructurado que puede variar.
  - o Datos de mercados emergentes.
  - o E-commerce.
  - o Datos meteorológicos.

<sup>6</sup>: <http://www.bit.es/knowledge-center/tipos-de-datos-en-big-data/>.

<sup>7</sup> <http://www.bit.es/knowledge-center/tipos-de-datos-en-big-data/> .

- **Datos no estructurados**<sup>8</sup>: es información sin procesar y que puede adquirir cualquier estructura. Aquí se engloba cualquier formato como los textos, las imágenes, los vídeos, los códigos, etc.
  - o Capturados: datos creados a partir del comportamiento de un usuario (información biométrica de pulseras de movimiento, aplicaciones de seguimiento de actividades (carrera, ciclismo, natación, ...), posición GPS).
  - o Generados por usuarios: datos que especifica un usuario (publicaciones en redes sociales, vídeos reproducidos en Youtube, búsquedas en Google,...).

La gran novedad del big data es esa capacidad de trabajar con datos no estructurados, impidiendo así una pérdida de información con respecto a los modelos antiguos o la BI. Mientras antes era necesario pasar por un proceso de filtrado y transformación de los datos no estructurados -con el coste y la pérdida que lleva aparejada-, ahora se aumenta la variabilidad y la variedad de los datos y, como consecuencia, su volumen y valor.

#### 1.4.2. Conceptos tecnológicos relacionados

Para entender el funcionamiento y potencial del big data es fundamental hacer referencia a otros conceptos tecnológicos que están estrechamente ligados al mismo. Es el caso de: Data Warehouse, Data Mining, Cloud Computing y Business Intelligence.

- **Data Warehouse**: es una base de datos corporativa que se caracteriza por integrar y depurar información de una o más fuentes distintas, para luego procesarla permitiendo su análisis desde infinidad de perspectivas y con grandes velocidades de respuesta. La creación de un data warehouse representa en la mayoría de las ocasiones el primer paso, desde el punto de vista técnico, para implantar una solución completa y fiable de Business Intelligence<sup>9</sup>.
- **Data Mining**: es el conjunto de técnicas y tecnologías que permiten explorar grandes bases de datos, de manera automática o semiautomática, con el objetivo de encontrar patrones repetitivos, tendencias o reglas que expliquen el comportamiento de los datos en un determinado contexto. Básicamente, el Data Mining surge para intentar ayudar a comprender el contenido de un repositorio de datos. Con este fin, hace uso de prácticas estadísticas y, en algunos casos, de algoritmos de búsqueda próximos a la Inteligencia Artificial y a las redes neuronales<sup>10</sup>.
- **Cloud Computing**: es el conjunto de programas y servicios alojados en un servidor conectado a la Red, accesible desde cualquier ordenador (sea cual sea el sistema operativo que éste ejecute) con conexión a Internet sin necesidad de instalar aplicaciones ejecutables en su disco duro y donde también se almacena la información generada por estas mismas aplicaciones o servicios.<sup>11</sup>
- **Business Intelligence**: es la habilidad para transformar los datos en información, y la información en conocimiento, de forma que se pueda optimizar

<sup>8</sup> <http://www.bit.es/knowledge-center/tipos-de-datos-en-big-data/>

<sup>9</sup> [http://www.sinnexus.com/business\\_intelligence/datawarehouse.aspx](http://www.sinnexus.com/business_intelligence/datawarehouse.aspx).

<sup>10</sup> [http://www.sinnexus.com/business\\_intelligence/datamining.aspx](http://www.sinnexus.com/business_intelligence/datamining.aspx) (

<sup>11</sup> <http://www.eleconomista.es/gestion-empresarial/noticias/4087167/07/12/Cloud-Computing-que-es-para-que-sirve-y-cuales-son-sus-aplicaciones-.html>

el proceso de toma de decisiones en los negocios. Desde un punto de vista más pragmático, y asociándolo directamente con las tecnologías de la información, se puede definir Business Intelligence como el conjunto de metodologías, aplicaciones y tecnologías que permiten reunir, depurar y transformar datos de los sistemas transaccionales e información desestructurada (interna y externa a la compañía) en información estructurada, para su explotación directa (reporting, análisis OLTP / OLAP, alertas...) o para su análisis y conversión en conocimiento, dando así soporte a la toma de decisiones sobre el negocio<sup>12</sup>.

En la siguiente figura se recogen las principales diferencias entre Business Intelligence (BI) y big data:

Figura 4: diferencias entre BI y big data.




Fuente: <http://www.tuataratech.com/2016/06/el-futuro-se-escribe-con-b-de-big-data.html>

### 1.4.3. Ecosistema del big data (landscape 2017)

Matt Turck, Director General de First Mark, junto a Jim Hao, ha elaborado el nuevo Landscape 2017 para contextualizar el ecosistema del big data (tecnologías y herramientas existentes para hacer big data en una empresa). La publicación aborda un análisis profundo de los fabricantes y nuevas tecnologías disponibles en el mercado de la Tecnología Informática (IT). La información resulta de mucha utilidad, tanto para los desarrolladores como para los responsables de infraestructuras, ya que ofrece una visión bastante exhaustiva de la oferta actual en el sector. Este ecosistema se divide en diferentes nodos:

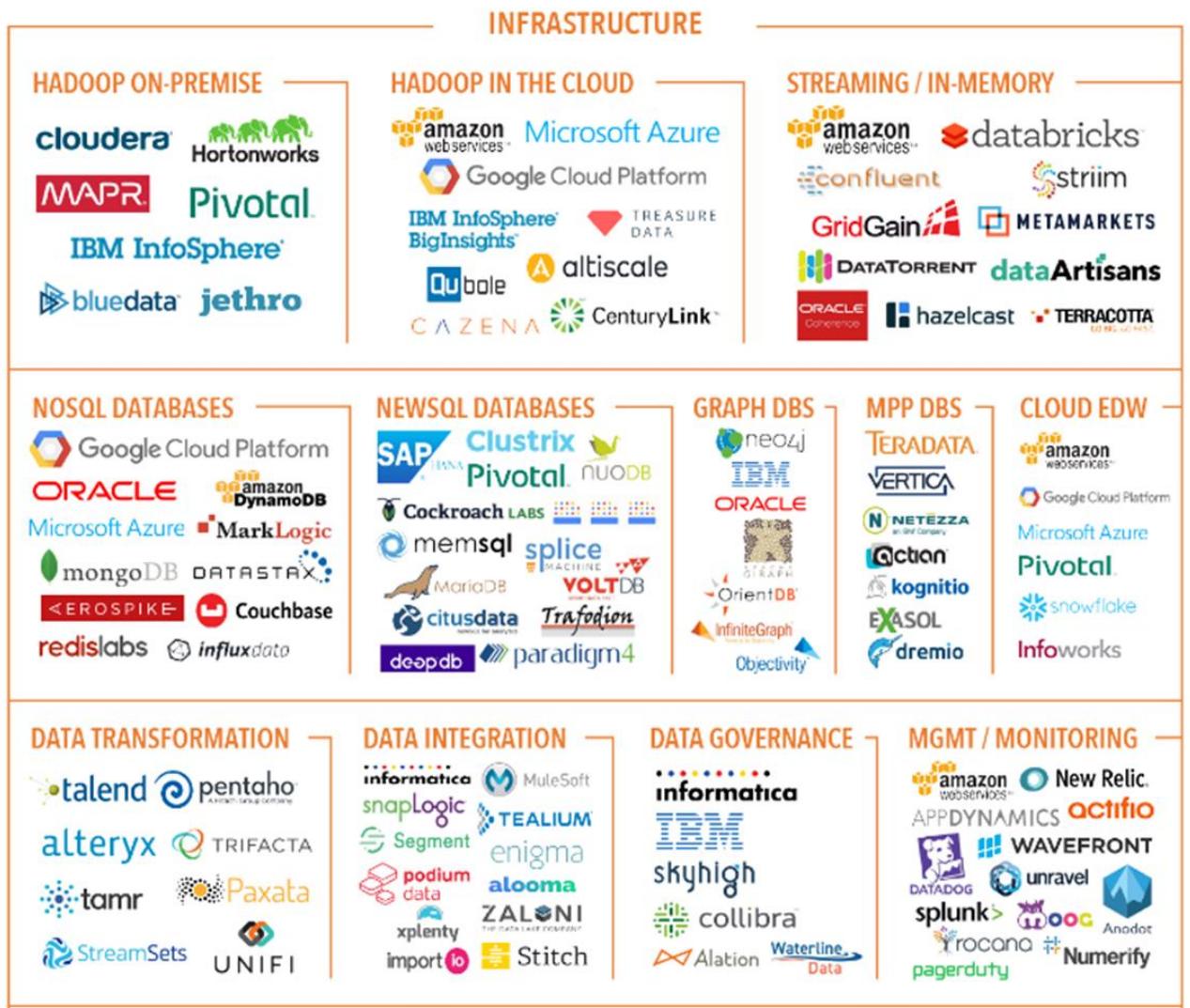
- Infraestructura, donde se incluyen las empresas y las tecnologías que ofrecen servicios de infraestructuras de big data, es decir aquellas empresas que venden los soportes que permiten la recogida de datos y su almacenamiento.

<sup>12</sup>: [http://www.sinnexus.com/business\\_intelligence/](http://www.sinnexus.com/business_intelligence/).

- 
- Analítica, en este cuadro se incluyen las empresas y tecnologías disponibles para analizar datos, se trata de software que permiten trabajar con la gran cantidad de datos que se han generado.
  - Aplicaciones: se trata de aquellos software diseñados para generar servicios concretos a los diferentes sectores o a determinadas aplicaciones de gestión. Dentro de este apartado encontramos: aplicaciones industriales, infraestructuras analíticas y sistemas open source.
  - Bases de datos e interfaz y recursos de datos. Se incluyen las empresas y aplicaciones que trabajan en el entorno de las bases de datos.

A continuación, se presenta el esquema fragmentado para exponer de manera individual cada uno de estos nodos:

Figura 5: esquema de la infraestructura.



Fuente: <http://mattturck.com/bigdata2017/>

Figura 6: esquema de infraestructuras analíticas.



Fuente: <http://mattturck.com/bigdata2017/>



Figura 7: esquema de analíticas.



Fuente: <http://mattturck.com/bigdata2017/>

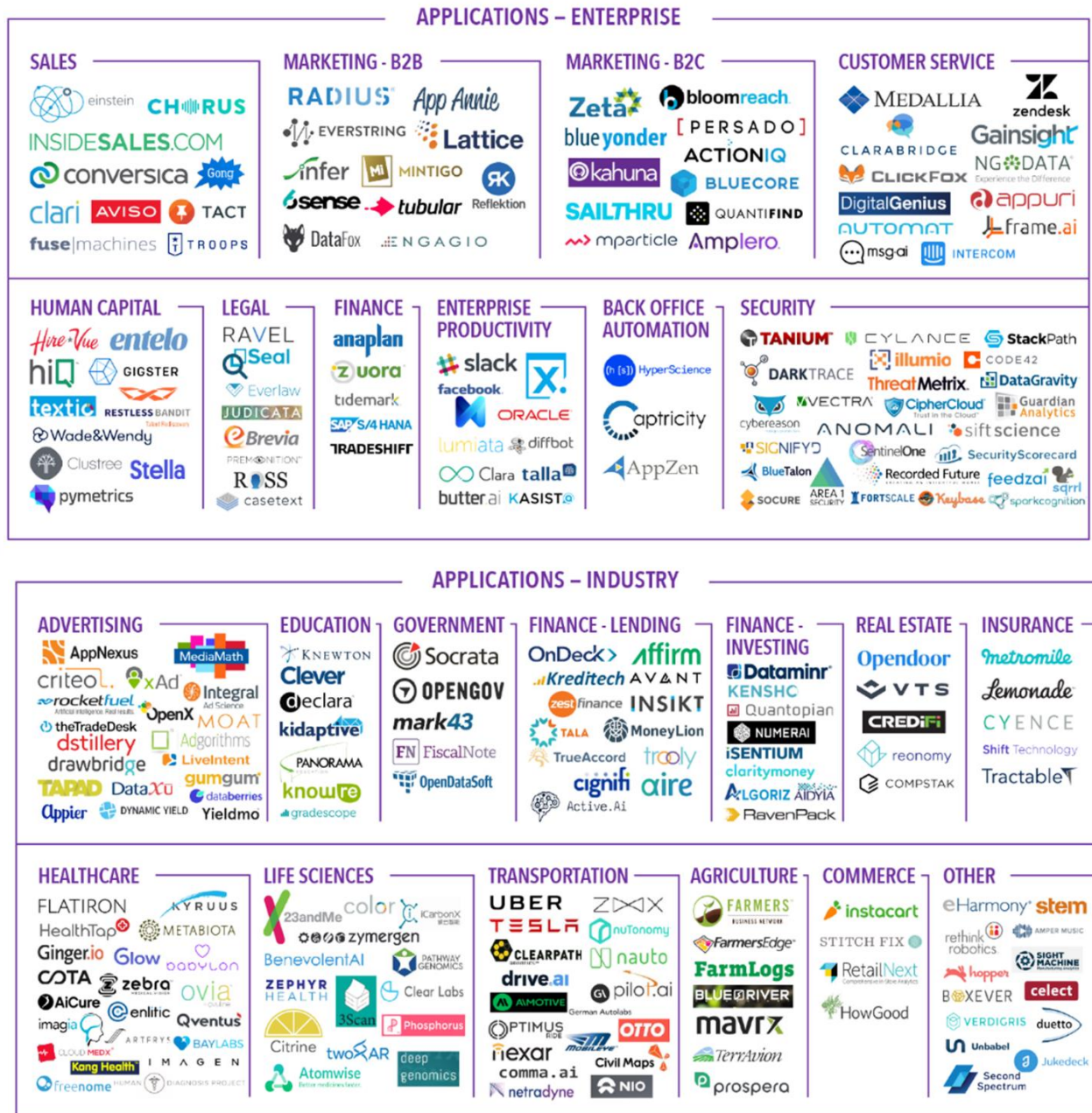
Figura 8: esquema de arquitectura de código abierto.



Fuente: <http://mattturck.com/bigdata2017/>



Figura 9: esquema de aplicaciones empresariales e industriales.



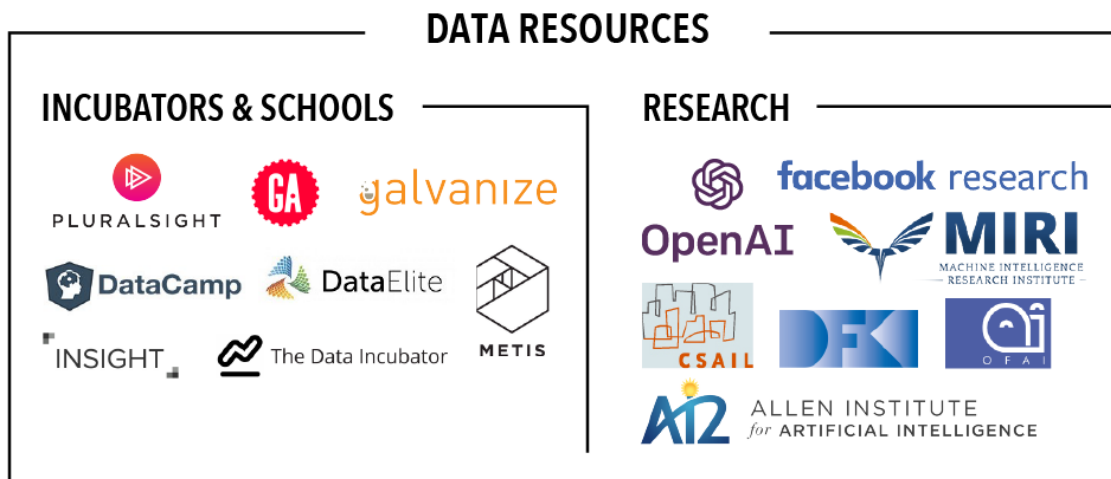
Fuente: <http://mattturck.com/bigdata2017/>

Figura 10: esquema de bases de datos y APIS.



Fuente: <http://mattturck.com/bigdata2017/>

Figura 11: esquema de recursos de datos.



Fuente: <http://mattturck.com/bigdata2017/>

#### 1.4.4. Arquitectura Lambda.

Nathan Marz publicó en su libro *big data: Principles and best practices of scalable realtime data systems* de 2015 un modelo que explica la arquitectura del big data. En concreto, la denominó Arquitectura Lambda y la definió como un conjunto de principios para arquitecturar sistemas big data en Tiempo Real.

A lo largo de este apartado se presenta el proceso que debe llevar a cabo cualquier empresa que quiera emplear big data para mejorar su rendimiento. La arquitectura Lambda es, por tanto, "el proceso", que engloba desde la recolección de datos (información proveniente de numerosas fuentes), pasando por el pre-procesamiento (para desechar lo innecesario y transformar lo interesante) y por la capa de análisis big data (eje central de la aplicación), para desembocar en la capa de aplicación (sectores en los que se aplicarán las soluciones big data).

La arquitectura Lambda, consta, por ende, de cuatro grandes capas como se define a continuación.

1. **Capa de recolección de datos.** Engloba las tecnologías y dispositivos que obtienen los datos en bruto como la analítica web, el Open Data, las apps, las bases de datos, los medios sociales o los datos monitorizados por sensores.
2. **Capa de pre-procesamiento.** Aquí se llevan a cabo servicios de preparación, limpieza, transformación, evaluación e integración de datos con el objetivo de preparar de forma óptima los datos para la siguiente capa.
3. **Capa de análisis big data.** En este nivel se engloban varias subcapas en función del tipo de procesado que se lleva a cabo en cada una de ellas:
  - Capa de velocidad: se encarga de realizar el cómputo en tiempo real.
  - Capa batch (por lotes): es la responsable de almacenar los datos monitorizados y pre-procesados, permitiendo realizar los cálculos necesarios para dar acceso a vistas concretas de los datos almacenados o a la información arbitraria de los mismos.

- Capa de servicios: se encarga de indexar y exponer las vistas de la capa anterior para que puedan ser buscadas a través de queries.

4. **Capa de aplicación.** En esta última capa se desarrollan los servicios concretos asociados en torno a múltiples sectores como el financiero, salud, administraciones públicas, medio ambiente, etc.

### 1.4.5. Tecnologías y herramientas

Actualmente son muchas las tecnologías basadas en big data, pues es un sector que está en continua expansión. Desde Paradigma Digital<sup>13</sup>, empresa por excelencia de big data en España, que organiza el evento internacional big data Spain, se resume en siete categorías el mapa tecnológico de las herramientas big data.

Figura 13 Mapa tecnológico de las herramientas big data

Tecnología	Concepto	Herramientas
<b>Ingesta</b>	Recolección de los datos desde su origen	- Flume - Sqoop
<b>Almacenamiento</b>	Guardar y gestionar grandes volúmenes de datos	- Cassandra - Hadoop HDFS - MongoDB - Elastic
<b>Gestión de recursos</b>	Planificación y asignación de recursos: procesamiento de datos	- Yarn: - Mesos:.
<b>Motores de procesamiento</b>	Núcleo central de la tecnología big data, se realiza el cómputo de manera distribuida entre varios nodos de computación	- Hadoop MapReduce - Spark - Storm - Flink
<b>Mensajería</b>	permiten el intercambio de datos entre los diferentes componentes software	- Kafka - RabbitMQ
<b>Bibliotecas de consulta</b>	orientadas a simplificar el acceso a los datos	- Hive - Pig - Spark SQ
<b>Bibliotecas de Machine Learning</b>	se encargan de implementar algoritmos para clasificar, predecir o perfilar estrategias	- Spark MLlib - FlinkML

Fuente propia basada estudio de paradigma Digital

<sup>13</sup>: <https://www.paradigmadigital.com/lineas-servicio/big-data/> .

Figura 14: los 7 campos de la tecnología big data.



Fuente: <https://www.paradigmadigital.com/ecosistema-big-data/>

Es imposible abordar el mundo del big data sin hablar de la veterana Apache Hadoop. Esta herramienta open source se considera el framework estándar para el almacenamiento de grandes volúmenes de datos; se usa también para analizar y procesar, y es utilizado por grandes multinacionales como Facebook y Yahoo!.

El proyecto Apache™ Hadoop® desarrolla software de código abierto para computación distribuida, confiable y escalable. Está diseñado para escalar de servidores individuales a miles de máquinas, cada una ofreciendo computación y almacenamiento local. En lugar de confiar en hardware para ofrecer alta disponibilidad, la propia biblioteca está diseñada para detectar y gestionar fallos en la capa de aplicación, por lo que ofrece un servicio altamente disponible encima de un grupo de equipos, cada uno de los cuales puede ser propenso a fallos.

El proyecto incluye estos módulos<sup>14</sup>:

- Hadoop Common: las utilidades comunes que soportan los otros módulos de Hadoop.
- Hadoop Distributed File System (HDFS™): sistema de archivos distribuido que proporciona acceso de alto rendimiento a los datos de la aplicación.
- Hadoop YARN: un marco para la programación de tareas y gestión de recursos de clúster.

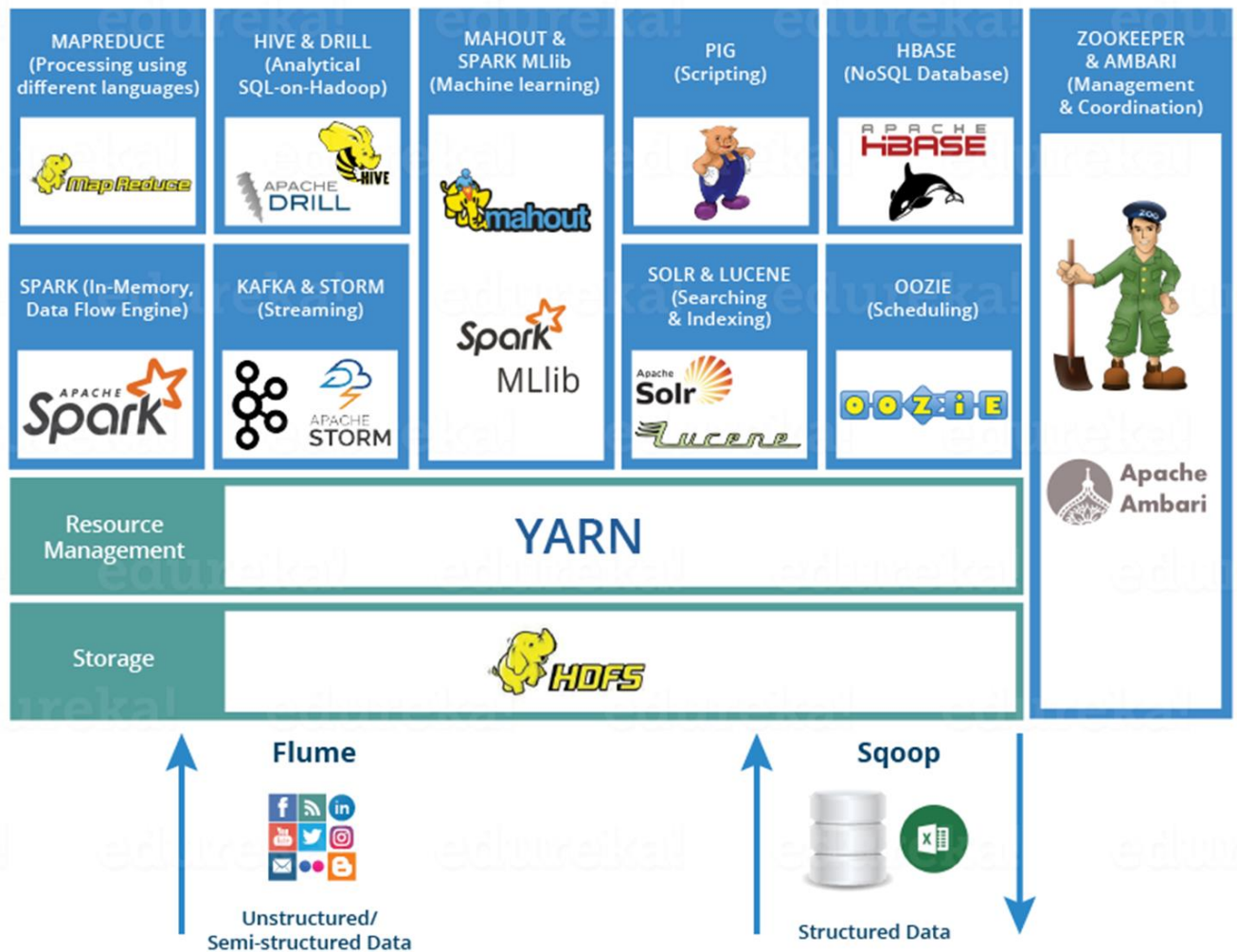
<sup>14</sup>: <http://hadoop.apache.org/> .



- Hadoop MapReduce: un sistema basado en YARN para el procesamiento paralelo de grandes conjuntos de datos.

El ecosistema de Hadoop es demasiado complejo para ser explicado programa a programa, aunque es interesante mostrarlo a través de la Figura 15 para tener un breve acercamiento de su situación en la actualidad (en pleno proceso de expansión):


Figura 15: ecosistema de Hadoop en la actualidad.



Fuente: <https://www.edureka.co/blog/hadoop-ecosystem>

### 1.5. Mapa de servicios: Ventajas del big data para las empresas

Las empresas e instituciones están en pleno descubrimiento de la tecnología big data para mejorar su estrategia y posicionarse en el mercado. Gracias a esta nueva forma de explotar la información y los datos, las empresas ofertan mejores productos, desarrollan una mejor relación con sus clientes y, por ende, se transforman en más ágiles y competitivas.



A continuación, se exponen las ventajas a las que optarán las empresas que usen big data para la explotación de sus servicios<sup>15</sup>:

- Nuevas oportunidades de negocio.
- Mejor segmentación, basada en la probabilidad de que el cliente contrate servicios o productos.
- Mejora la estrategia: se puede predecir con mayor exactitud qué productos son más apropiados para cada cliente. También forja la unión cliente-agente.
- Negocio más visible gracias a informes más detallados.
- Analiza las Redes Sociales y dentro de ellas a las personas: determina los círculos sociales de los clientes, identificando qué papel desempeñan en sus círculos y su grado de influencia.
- Mejora el conocimiento de clientes y mercados a través del Marketing Viral.
- La información en tiempo real: siempre está disponible, sin tener que esperar una actualización de datos.
- Anticipación al problema: con el análisis predictivo y el cruce de datos es posible anticiparse a los problemas futuros.
- Reducción de costes: gracias a la simplificación de los procesos actuales y control de negocio, y a la analítica predictiva frente a fraudes y transacciones y operaciones sospechosas.
- Ayuda en la toma de decisiones: gracias a algoritmos automáticos se puede descubrir información que se mantenía oculta o a la que no se le daba importancia.
- Reducción del tiempo.
- Filtros inteligentes de seguridad en el negocio electrónico.

Para extraer todo el potencial de esta nueva tecnología es necesario entender el negocio y los datos. Por ello, es preciso un análisis detallado de las herramientas que la empresa maneja y que condiciona sus procesos. En segundo lugar, se deben localizar los problemas del negocio y cómo los datos pueden ayudar a combatirlos. A continuación, se deben establecer expectativas razonables y metas alcanzables, para probar si con las nuevas soluciones prestadas por el big data, se revierte la situación.

Una premisa a tener en cuenta según los expertos es que al implementar un proyecto de big data se debe ser flexible con la metodología y las herramientas, pues la reconfiguración de la estrategia empresarial lleva tiempo y ensayos prueba-error. Por último, una recomendación es trabajar en paralelo con el sistema que hoy está funcionando en cada organización o empresa.

---

<sup>15</sup>: <https://repositorio.unican.es/xmlui/bitstream/handle/10902/4528/TFM%20-%20David%20L%C3%B3pez%20Garc%C3%ADaS.pdf?sequence=1> .

## BLOQUE II: ANALISIS Y DIAGNÓSTICO DE LOS NEGOCIOS Y EXPERIENCIAS EMPRESARIALES EN BIG DATA.

### 2.1. Mapa de sectores y actividades para sacar partido al big data

Multitud de empresas e instituciones a nivel mundial precisan almacenar y gestionar la infinidad de datos e información que generan y reciben. Ahí es donde entra la funcionalidad del big data para sacar lo mejor de la empresa.

La tecnología big data se aplica indistintamente en numerosos campos y sectores, dentro del ámbito de la empresa privada y en organizaciones y centros científicos, incluso en terreno gubernamental, ayudando así a optimizar los resultados y mejorar la toma de decisiones.

Algunos de estos sectores son: Turismo, Salud, Agricultura, Medio ambiente, Educación, Gobierno e instituciones, Banca, Deportes y Sector Retail, entre otros.

Figura 16: utilidades del big data por sectores.



Fuente: big data España 2015. IBM.

A continuación, se desarrollan los principales servicios que actualmente ofrece el big data en cada uno de los sectores:





<b>Gobiernos, Instituciones y Organizaciones:</b>	<b>Ciencia y Salud:</b>
<ul style="list-style-type: none"><li>• Gestión del conocimiento: proceso por el cual una organización facilita la transmisión de informaciones y habilidades a sus empleados, de una manera sistemática y eficiente.</li><li>• Mejoras en rendimiento y servicios.</li><li>• Soporte a aplicaciones de misión crítica: geoespacial, investigación científica.</li><li>• Detección automática de vulnerabilidades, fraude y malversación de fondos públicos.</li><li>• Ciberseguridad: detección de filtración de datos, detección de actividades criminales y patrones.</li><li>• Rendimiento de infraestructuras.</li><li>• Análisis de datos para anticipar quejas y preocupaciones de los ciudadanos.</li><li>• Optimización del tráfico: mejorar atascos en las ciudades.</li></ul>	<ul style="list-style-type: none"><li>• Predicción meteorológica.</li><li>• Estudios sobre el genoma.</li><li>• Análisis de composición de la materia.</li><li>• Detección de posibles brotes víricos o epidemias.</li><li>• Registros de pacientes más eficientes.</li><li>• Planes de salud.</li><li>• Diagnósticos y tratamientos inmediatos.</li><li>• Permitir a los equipos de sistemas, negocio e informática descubrir mejoras y errores analizando las comunicaciones entre aplicaciones.</li><li>• Seguridad en sistemas de salud.</li><li>• Auditorías y detección de anomalías. Prevención de fraude y abuso.</li><li>• Monitorización y alertas en tiempo real.</li><li>• Análisis predictivo. Riesgos de intervención clínica.</li><li>• Gestión y análisis de salud en grupos de población de riesgo para emitir recomendaciones.</li></ul>
<b>Política</b>	<b>Banca, Finanzas y Seguros:</b>
<ul style="list-style-type: none"><li>• Conocer las preferencias y necesidades de los ciudadanos.</li><li>• Desarrollar programas y políticas públicas acordes a las demandas.</li><li>• Discursos más cercanos y directos.</li><li>• Predecir el comportamiento electoral al conocer mejor al electorado.</li><li>• Se potencia la comunicación y la interacción digital porque hay una mejora en la escucha-participación del ciudadano.</li><li>• Gestión de crisis.</li></ul>	<ul style="list-style-type: none"><li>• Mitigar riesgos y anticipar oportunidades.</li><li>• Detectar fraude.</li><li>• Gestión de aplicaciones de forma inteligente.</li><li>• Visibilidad Avanzada Persistente.</li><li>• Correlacionar ingresos con transacciones.</li><li>• Optimizar balanceos de carga: para que las peticiones de Internet sean distribuidas sobre una fila de servidores y gestionar más fácilmente los usuarios que realizan peticiones online.</li><li>• Nueva generación de monitorización de rendimiento de redes.</li></ul>
<b>Logística y Manufacturas:</b>	<b>Cultura:</b>
<ul style="list-style-type: none"><li>• Optimización de inventarios y volumen de aprovisionamiento constante.</li><li>• Análisis de transacciones de pedidos.</li><li>• Backups de almacenamiento.</li><li>• Solución de incidencias RFID: chips que informan sobre su ubicación exacta y su estado físico en cualquier momento durante el proceso de producción o envío.</li><li>• Optimización de rutas.</li><li>• Datos de telemetría a través de sensores (patrones de uso, tasas fracaso y montaje).</li></ul>	<ul style="list-style-type: none"><li>• Conocimiento de los gustos, tendencias y apetencias del público.</li><li>• Diseño de estrategias cercanas a las peticiones.</li><li>• Nuevos modelos de planificación y desarrollo cultural: mediante el uso de la estadística y de algoritmos.</li><li>• Conservación del patrimonio mediante sensores.</li><li>• Localización de nuevos nichos en el ámbito de la cultura (música, teatro, espectáculos).</li><li>• Planear exposiciones o reubicar colecciones en galerías de arte y museos</li></ul>





Sector Industrial	Deportes
<ul style="list-style-type: none"><li>• Mantenimiento efectivo de la maquinaria: sensores para registrar datos sobre el estado de sus componentes.</li><li>• Optimización de procesos.</li><li>• Detección de nuevas oportunidades de eficiencia.</li><li>• Identificar, corregir y reducir los riesgos involucrados en la cadena de suministro.</li><li>• Hiperconectividad.</li></ul>	<ul style="list-style-type: none"><li>• Análisis de la competencia mediante sensores.</li><li>• Diseño de la estrategia de entrenamiento y competición más conveniente.</li><li>• Estudio sobre posibles riesgos de lesiones.</li><li>• Análisis y puesta en funcionamiento de un plan nutricional y de sueño personalizado.</li><li>• Mejora de los movimientos de los jugadores a través de sensores.</li></ul>
Retail	Turismo
<ul style="list-style-type: none"><li>• Adoptar enfoque orientado a datos para ofrecer mejores experiencias a los clientes.</li><li>• Ciberseguridad: detección de filtración de datos.</li><li>• Monitorización de clientes en tiempo real. Segmentación y perfiles.</li><li>• Programas de fidelización de clientes.</li><li>• Predecir tendencias.</li><li>• Recomendación de nuevos productos.</li><li>• Análisis de cestas de compra.</li><li>• Gestión de inventarios.</li><li>• Análisis de localización de locales comerciales con potencial de negocio.</li><li>• Ayuda en la creación de marca.</li><li>• Entender el sentimiento del cliente y la interacción con el mismo.</li><li>• Política de precios</li></ul>	<ul style="list-style-type: none"><li>• Aplicar análisis y anticiparse a problemas en los alojamientos, en restauración, en transportes, etc.</li><li>• Información en tiempo real sobre los usuarios, sus movimientos y sus preferencias.</li><li>• Crecimiento de Smart Cities: destinos inteligentes que cuentan con sensores conectados a la red (smartphones, puntos de venta, vehículos, etc.). Capturan los datos generados por los turistas, antes, durante y después de la estancia.</li><li>• Conocer al visitante: cuántos visitantes se reciben, de dónde proceden y tiempo medio por nacionalidades.</li><li>• Niveles de gasto: en global y en la categoría de alojamiento.</li><li>• Detección de influencers en redes sociales para que promocionen nuestro destino o empresa.</li><li>• Marketing digital.</li></ul>
Agricultura	Medio Ambiente
<ul style="list-style-type: none"><li>• Vigilancia del rebaño.</li><li>• Mejora la eficacia de las inspecciones pesqueras.</li><li>• Lectura óptica de la maduración y composición de producto mediante sensores desde tres puntos: suelo, planta y clima</li><li>• Economía del agua.</li><li>• Predicción de producción o de necesidades existentes.</li><li>• Control de plagas.</li><li>• Control de precios</li><li>• Control de sistemas de producción agroindustrial</li></ul>	<ul style="list-style-type: none"><li>• Control de la contaminación y de la dirección del viento en tiempo real.</li><li>• Predicción y control de incendios.</li><li>• Simulación de propagación del fuego en incendios forestales.</li><li>• Depuración de aguas residuales con menos energía y menor coste.</li><li>• Mejora de los servicios climatológicos.</li><li>• Mantenimiento de la biodiversidad y protección de especies.</li></ul>





---

<b>Publicidad, Medios y Telecomunicaciones</b>	<b>Educación</b>
<ul style="list-style-type: none"><li>• Mantener la infraestructura del backend -área del administrador del sitio- siempre alerta a la temática que pueda existir en la web sobre su servicio para actuar automáticamente.</li><li>• Detectar incidencias en tiempo real y darles respuestas.</li><li>• Obtener predicciones relevantes: picos/valles de actividad.</li><li>• Oportunidad de dirigirse a los consumidores cuando estén cerca de un comercio: gracias a los teléfonos y dispositivos GPS.</li><li>• Analizar y crear tendencias</li></ul>	<ul style="list-style-type: none"><li>• Análisis crítico del discurso: es un tipo de investigación analítica sobre el discurso que estudia primariamente el modo en que el abuso del poder social, el dominio y la desigualdad son practicados, reproducidos, y ocasionalmente combatidos, por los textos y el habla en el contexto social y político.</li><li>• Permite filtrar y evaluar recursos de aprendizaje.</li><li>• Permite valorar la motivación de los alumnos, su relación con otros compañeros o la propia capacidad individual para el aprendizaje.</li><li>• Tipo de participación del alumno (tiempo de conexión, lugar, etc.).</li><li>• Diseñar clases efectivas y personalizadas.</li></ul>

---

## 2.2. Estado y evolución del big data

### 2.2.1. Contexto internacional

El impacto actual del big data es tan amplio como su transversalidad. Se considera una palanca de valor para el PIB de países, empresas, ámbitos económicos e incluso para la calidad de vida de los ciudadanos. Un informe de la compañía de investigación de mercados IDC apunta a que los ingresos alrededor de los datos y su análisis crecerán desde 122 mil millones de dólares en 2015 hasta más de 187 mil millones de dólares en 2019. Es decir, aumentará más del 50% en menos de cinco años

Un estudio de la Omidyar Network, una red de inversión filantrópica, ya establecía en 2014 que las políticas de datos abiertos podían aumentar los ingresos de los países del G20 entre 700 y 950 mil millones de dólares al año. Algunos de los beneficios provenían de la reducción de la corrupción gracias a la transparencia, de las mejores condiciones en el lugar de trabajo, de la mayor eficiencia energética y de un mejor comercio exterior. Otro beneficio de la gran disponibilidad de datos fue la reducción de las barreras para que emprendedores y pequeñas empresas entren en el mercado

Ya de hecho durante el pasado 2016 se produjo una gran explosión en el mundo del big data. Esta tecnología, que en un principio fue utilizada por grandes multinacionales y en sectores muy precisos, hoy ha llegado a las startups y pymes para implantarse en su modelo de negocio.

Una mayor cantidad de organizaciones ha comenzado a almacenar y procesar datos de todo tipo de formatos y tamaños, además de extraer valor de ellos, como ya lo hacían a través de la BI. A lo largo del 2018 y siguientes años continuará la expansión de los sistemas que admiten grandes volúmenes de datos, tanto estructurados como no estructurados.

Inciendo en los beneficios del big data para las empresas, se puede afirmar que el aprovechamiento de los datos tiene la capacidad de transformar cualquier industria. Erik Brynjolfsson, director de Initiative on the Digital Economy del prestigioso Instituto Tecnológico de Massachusetts (MIT), sostiene que las empresas que adoptan decisiones basadas en datos logran entre 5% y 6% más productividad que aquellas que no lo hacen. Un informe de McKinsey Global Institute estima que big data podría generar un valor adicional anual superior a tres mil millones de dólares en siete industrias

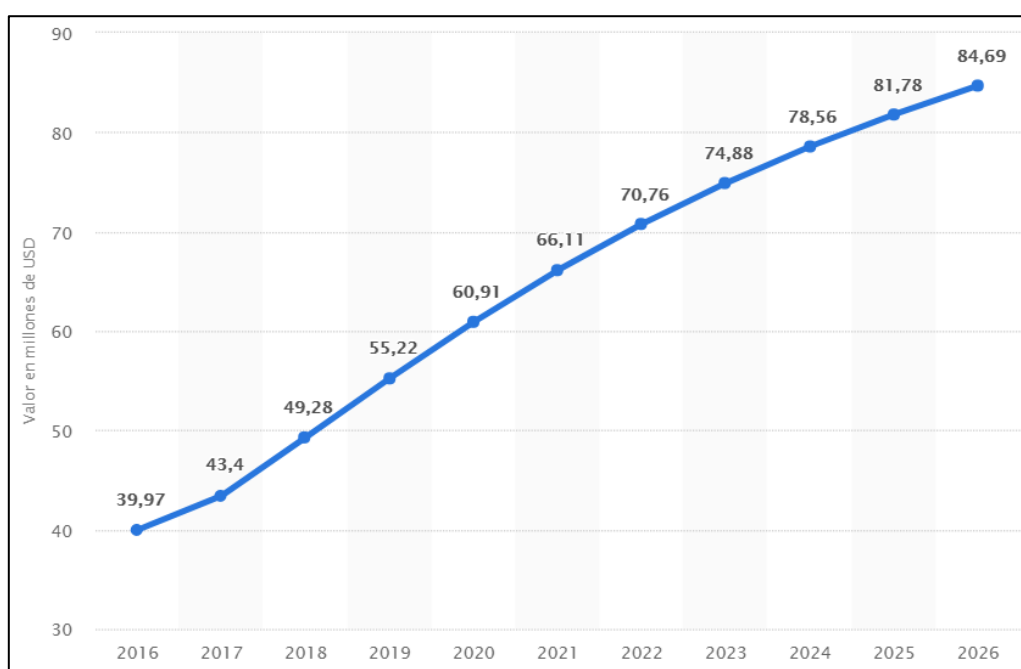


analizadas a nivel mundial. El informe estima que alrededor de la mitad repercutirá de forma directa en los ciudadanos por mejora de servicios públicos y privados.

El volumen de negocio del sector del big data en Europa se prevé que supere los 5.500 millones de euros en 2018, muy por encima de los 2.000 millones de euros que facturaba en 2013, según el Colegio Universitario de Estudios Financieros (CUNEF)<sup>16.</sup>, lo que representa el 1,87% del PIB de los países miembros y podía ascender al 4,7% en 2020.

Juan Manuel López Zafra y Ricardo Queralt, codirectores del Máster de Data Science para Finanzas de CUNEF, han apuntado que “el big data ha provocado un cambio de paradigma en diferentes sectores productivos, con la aparición de nuevos actores, que se configuran como los competidores de las compañías tradicionales, en sectores tan dispares como el de las telecomunicaciones, el turismo o el financiero”.

Figura 16: previsión del valor de mercado del big data en el mundo entre 2016 y 2026.




Fuente: Statista 2017.

Por otra parte, y según un estudio de Accenture -consultora de referencia en las TIC- las grandes empresas con ingresos anuales por encima de 10.000 millones de dólares (7.000 millones de euros), enfocan el big data de un modo distinto a las pequeñas empresas (con menos de 500 millones de dólares de ingresos anuales, 350 millones de euros):

“El 67% de los ejecutivos de grandes empresas consideran que el big data es extremadamente importante, frente a solo un 43% de los encuestados en empresas más pequeñas. Los ejecutivos de grandes empresas tienen una percepción del big data más amplia que la de sus colegas en pequeñas empresas y utilizan más fuentes de datos en sus iniciativas de big data, como datos de redes sociales (54% frente a 29%), datos de visualización (50% frente a 29%) o datos sin estructura (49% frente a 36%). El 62% de los ejecutivos de

<sup>16</sup>: <http://www.europapress.es/economia/noticia-sector-big-data-facturara-mas-5500-millones-euros-2018-europa-cunef-20170329173524.html>.



grandes empresas afirman que los altos directivos comprenden y apoyan las iniciativas de big data, frente al 42% de los encuestados en pequeñas empresas”.<sup>17</sup>

En los siguientes subapartados se retratarán algunos casos de éxito que se han producido a nivel internacional, nacional y andaluz gracias a la aplicación de big data en los negocios, en la investigación y en los organismos oficiales”.

### 2.2.2. Contexto nacional

En la actualidad no existen demasiados informes internos que dimensionen el mercado y la evolución de los datos en España. Un estudio de 2014 —elaborado, entre otras, por la Asociación Española de Estudios de Mercado, Marketing y Opinión— presentó una cifra neta de negocio de 441,41 millones de euros para el mercado big data en España.

El informe también cuantificaba que la industria del big data española empleaba en 2014 a más de 10.000 personas, de las que el 57% eran altamente cualificadas. Otros análisis, como el de Carmen Artigas, cofundadora de Synergic Partners (Grupo Telefónica), muestran la demanda actual de capacidades de datos en España: “el mercado de big data crece un 30% cada año en España, siete veces más que la inversión en tecnologías de la información tradicionales. Es una apuesta sólida para aumentar las fuentes de ingresos de las compañías, consolidar sus estrategias de personalización y fidelización de los clientes e impulsar su transformación digital”.

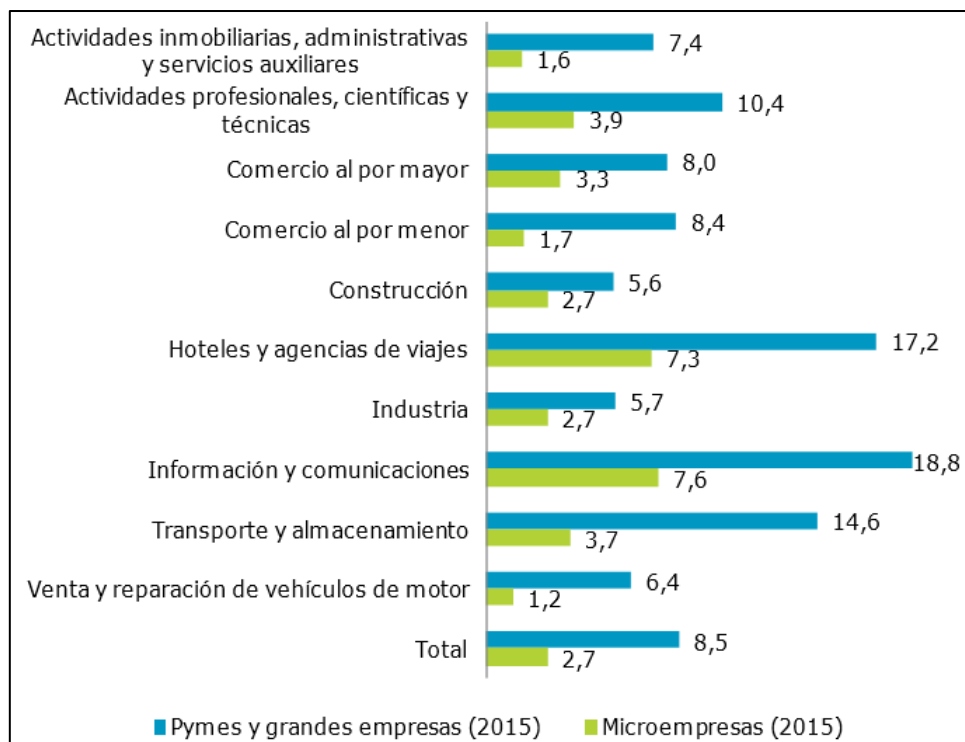
Estas cifras ilustran el impacto real que el big data y la economía del dato van a tener en el crecimiento económico de España durante los próximos cinco años. Esta oportunidad económica trae consigo una mayor: la oportunidad de crear nuevos perfiles profesionales cualificados. La creación de puestos de trabajo cualificados y destinados a formar una industria que puede convertirse en un motor de la economía nacional, como ya está sucediendo en otros países.

En España el big data se está instalando en diversos sectores, sacando el máximo rendimiento a la empresa. Mediante la Figura 17 se observa cuáles son los mercados que más lo usan, haciendo una comparativa entre Pymes, grandes empresas, por un lado, y microempresas por otro.

Figura 17: sectores que utilizan big data en España.

---

<sup>17</sup>: <https://www.accenture.com/es-es/company-big-data-fundamental-transformacion-digital>.



Fuente: ONTSI a partir de datos del INE 2016.

El sector de la información y las comunicaciones, junto a los hoteles y agencias de viajes son los que más se sirven del big data para potenciar sus negocios con un 18,8% y un 17,2% respectivamente en las pymes y grandes empresas. Le siguen el transporte y almacenamiento con un 14,6% y las actividades profesionales, científicas y técnicas con un 10,4%.

Cuando se habla de microempresas, el sector de la información y las comunicaciones sigue siendo el más potente con un 7,6%, seguido de los hoteles y agencias de viaje con un 7,3%, de la misma forma que ocurría con las pymes y grandes empresas.

Si se visualiza el compendio total se puede destacar que existe una importante brecha en la utilización de big data entre las grandes y medianas empresas (8,5%) y las microempresas, que solamente utilizan el potencial de los datos masivos en un 2,7%.

La gestión inteligente de los datos ha crecido un 30% en España el último año, según la consultora informática Synergic Partners. El último estudio del Instituto Nacional de Estadística (INE) refleja que 8 de cada 100 empresas españolas manejan el big data de manera habitual para dirigir sus negocios, mientras que en Europa su uso está más asentado.

Por otra parte, los resultados del Informe ePyme 2016 del Observatorio Nacional de las Telecomunicaciones y Sociedad de la Información (ONTSI), pone de manifiesto que tan solo 10 de cada 100 empresas grandes, medianas y pequeñas usan big data.

Las empresas son cada vez más conscientes del beneficio que supone usar big data en su estrategia empresarial y así lo demuestra un estudio realizado por EMC -consultora tecnológica- entrevistando a 510 profesionales de TI en España (Figura 18).

Figura 18: Beneficios del big data. Encuesta española.

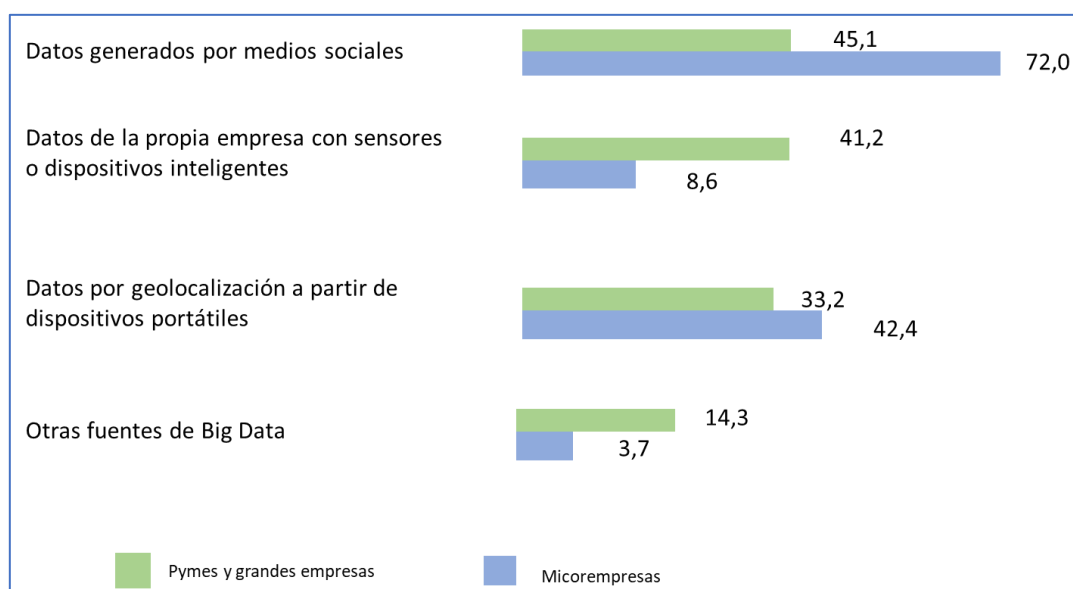




Fuente: [http://www.silicon.es/las-empresas-espanolas-empiezan-comprender-las-ventajas-del-big-data-47366?inf\\_by=59c7ba63671db839478b459a](http://www.silicon.es/las-empresas-espanolas-empiezan-comprender-las-ventajas-del-big-data-47366?inf_by=59c7ba63671db839478b459a)

En cuanto a la procedencia de los datos, las empresas españolas se nutren en su mayoría de los datos generados por los medios sociales, seguido de los datos por geolocalización a partir de dispositivos móviles y los datos extraídos de la propia empresa a través de sensores o dispositivos inteligentes, como se observa en la Figura 19.

Figura 19: procedencia de los datos de las empresas españolas que usaron big data.




Fuente: ONTSI a partir de datos INE 2016.

En definitiva, según Forbes, el big data como servicio generará 30 mil millones de dólares en 2021, lo que representa un 34% del total del negocio. Las previsiones son muy positivas, pues se prevén aumentos anuales de volumen de negocio superiores al 30% en la próxima década.

### 2.2.3. Contexto andaluz





Miguel Cardoso, economista jefe para España de BBVA Research, sitúa a España y Andalucía en un nivel de adaptación a las nuevas tecnologías digitales por debajo de la media europea:

“Solo el 72% de las empresas andaluzas dispone de página web, lo que supone 5 puntos menos que la media española. Además, únicamente un 21% de ellas permite hacer reservas online y, aunque la cifra es ligeramente superior a la de media estatal, todavía es reducida”.<sup>18</sup>

En el sector público, el indicador sintético de BBVA Research asegura que Andalucía se encuentra “en línea con la situación en el resto del Estado, a un nivel similar al de la Unión Europea, aunque todavía lejos de los niveles que alcanzan algunos países del norte de Europa”.

De hecho existen importantes iniciativas públicas y privadas en nuestra Comunidad que trabajan con tecnología big data. La Consejería de Turismo y Deporte y la Consejería de Agricultura, Pesca y Desarrollo Rural de la Junta de Andalucía son un referente en ello, con importantes programas para implementar el uso del big data dentro de estos sectores.

Concretamente resalta la propuesta de la Consejería de Turismo y Deporte para poner en marcha el primer Smart Data de carácter regional y público destinado al turismo, como una herramienta tecnológica que permitirá a destinos y empresas perfeccionar la experiencia del cliente, conocer mejor el mercado al que se dirigen, optimizar las acciones promocionales y de planificación en función del interés y el grado de satisfacción del viajero y programar las campañas turísticas.

En esta misma línea es necesario señalar la apuesta que la Junta de Andalucía ha realizado dentro del sector TIC a través de la aprobación de la Estrategia de Impulso del Sector TIC hasta el año 2020, con el objetivo de generar empleo en el sector, incrementar la dimensión de las empresas para hacerlas más competitivas e incentivar su internacionalización.

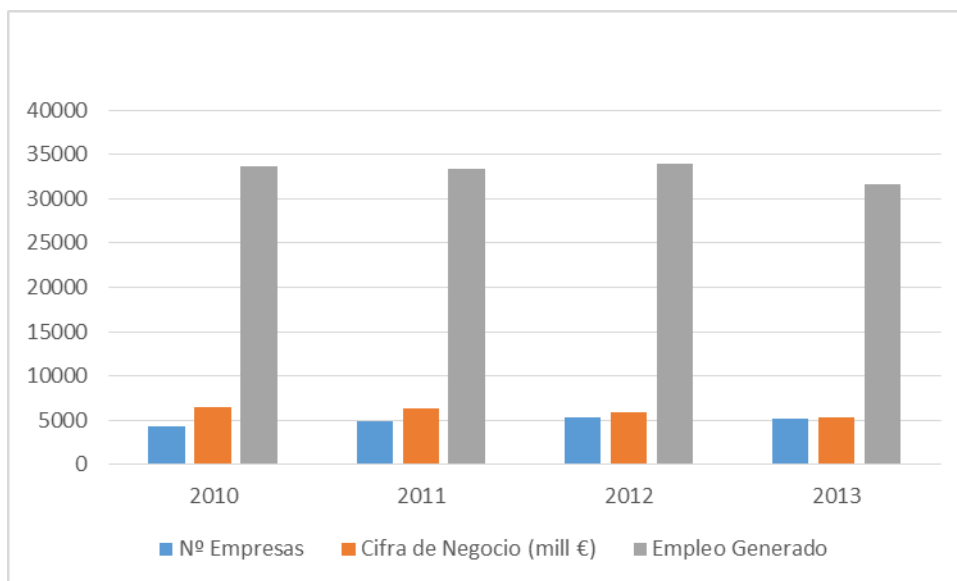
A día de hoy, la actividad de las TIC en Andalucía factura 5.300 millones de euros al año, genera 31.200 empleos y cuenta con 5.100 empresas innovadoras ubicadas principalmente en Málaga y Sevilla, las cuales se dedican en su mayoría a la programación y a la consultoría.

En la siguiente figura se contempla la evolución del número de empresas, cifra de facturación en millones de euros y el empleo generado por el sector TIC Andaluz durante los últimos años de los que se disponen cifras, y que se reflejan en la Estrategia Andaluz de Impulso del Sector TIC.

Figura 20: Evolución del Sector TIC en Andalucía

---

<sup>18</sup>: [http://www.lavozdigital.es/cadiz/provincia/lvdi-digitalizacion-y-tecnologias-data-claves-para-supervivencia-empresas-201706100922\\_noticia.html](http://www.lavozdigital.es/cadiz/provincia/lvdi-digitalizacion-y-tecnologias-data-claves-para-supervivencia-empresas-201706100922_noticia.html)



Fuente: Estrategia del Sector TIC 2020 Junta de Andalucía

Con este plan se pretenden crear 5.000 nuevos empleos hasta 2020, hacer crecer en un 30% el número de empresas innovadoras y aumentar un 25% su aportación al PIB regional hasta situarse en 2,14%. Además también tiene el objetivo de favorecer la internacionalización, ya que sólo una de cada cinco empresas TIC se ha expandido mundialmente.

La estrategia establece 23 programas englobados en nueve ejes con especial atención a las pequeñas y medianas empresas y los autónomos. Entre ellos destacan por su novedad el de especialización tecnológica, dirigido a posicionar al sector TIC andaluz en áreas como el 'big data', la ciberseguridad o 'el Internet de las cosas', y otro específico de igualdad de género.

## 2.3. Identificación de experiencias empresariales con big data

### 2.3.1. Contexto internacional

#### 1. Atención al cliente:

- Xerox (EE.UU.): utiliza big data para mejorar la tasa de desgaste en sus centros de llamadas. Gracias a esto logró una reducción del 20%, ya que detectó qué es lo que estaba influyendo en su volumen de negocio y determinó formas para mejorar la participación de sus empleados.
- Unilever (Reino Unido): utiliza el big data para conocer si las conversaciones de los potenciales consumidores en medios sociales difieren mucho de su comportamiento real de compra, y también analiza el comportamiento de los usuarios en las búsquedas que éstos realizan en internet.

## 2. Deporte:

- Oakland Athletics (EE.UU.): el exjugador de baseball y ejecutivo Billy Beane utilizó una serie de métodos estadísticos propios de los mercados financieros para determinar la valía de sus jugadores y de otros potenciales.

## 3. Inmobiliaria:

- Twiddy (EE.UU.): esta inmobiliaria ha logrado un descenso del 15% en los costes de mantenimiento comparados con el historial y la media de los otros 1.200 suministradores. Lo han logrado identificando y eliminando errores de procesamiento de facturas, y mediante la automatización de la planificación de servicios.

## 4. Investigación:

- PSG College of Technology (India): esta universidad de la India analiza múltiples secuencias de proteínas para determinar los enlaces evolutivos y predecir estructuras moleculares. El uso de big data a través de Hadoop mejora la velocidad y exactitud de estas secuencias.
- Universidad de Maryland (EE.UU.): colabora en la iniciativa académica de cómputo en la nube de IBM/Google. Sus investigaciones incluyen proyectos en la lingüística computacional, modelado del lenguaje, bioinformática, análisis de correo electrónico y procesamiento de imágenes.<sup>19</sup>
- CERN (Suiza): el laboratorio suizo de física nuclear hace big data desde hace más de 20 años. Cada año produce 30 petabytes de nueva información, acumulando ya unos 250 Pb de datos en sus centros de datos, sobre los que se realizan unos dos millones de tareas cada día.
- Language, Interaction and Computation Laboratory (CLIC) (Italia): junto a la Universidad de Trento en Italia estudian la comunicación verbal y no verbal tanto con métodos computacionales como cognitivos.
- Universidad Distrital Francisco José de Caldas: utiliza big data para investigar el sistema de inteligencia territorial de la ciudad de Bogotá.

## 5. Logística:


- Grupo Vaasan (Finlandia): mediante big data ha mejorado su entrega de pedidos en un 98,5% y ha reducido el riesgo de negocio alineando materias primas, recursos humanos y calendario de producción a la demanda de los clientes.

## 6. Medio ambiente:

- The Weather Company (EE.UU.): junto a IBM están trabajando en mejorar el impacto del clima sobre el rendimiento de un negocio, que se prevé que supone

---

<sup>19</sup>[https://www.researchgate.net/profile/Renato\\_Gonzalez6/publication/311950470\\_Inteligencia\\_Computacional\\_Aprendizaje\\_de\\_Maquinas/links/58645b5208ae8fce490b74f3/Inteligencia-Computacional-Aprendizaje-de-Maquinas.pdf](https://www.researchgate.net/profile/Renato_Gonzalez6/publication/311950470_Inteligencia_Computacional_Aprendizaje_de_Maquinas/links/58645b5208ae8fce490b74f3/Inteligencia-Computacional-Aprendizaje-de-Maquinas.pdf).



medio billón de dólares al año sólo en los EEUU. Los datos sobre climatología están siendo recogidos por más de 100.000 sensores y aviones meteorológicos, así como millones de teléfonos inteligentes, edificios y vehículos en movimiento. Esos datos se combinan con los de otras fuentes para lograr 2,2 millones de puntos de pronóstico únicos, y un promedio de más de 10 mil millones de pronósticos del tiempo en un día activo<sup>20</sup>.

## 7. Política:

- Campaña Barack Obama (EE.UU.): usó big data para su reelección en 2012. La estrategia se enfocó en tres aspectos: registro, persuasión y voto del electorado. La clave fue trabajar de manera conjunta el equipo de campo, el digital y el de comunicación. La tecnología utilizada fue la plataforma inteligente HP Vertica, que permitía recoger datos a pie de campo y realizar un feedback muy rápido vía notificaciones email por parte del equipo online o detectar los nichos en los que funcionaría mejor la publicidad en TV cruzando datos de los votantes con otros demográficos, audiencias, precios de publicidad, programas, etc.<sup>21</sup>

## 8. Salud:

- Bristol-Myers Squibb (EE.UU.): la farmacéutica estadounidense redujo en un 98% el tiempo que se tarda en ejecutar simulaciones de ensayos clínicos mediante la implementación de un entorno de red alojado internamente en los sistemas cloud de AWS (Amazon Web Services). También optimizó los niveles de dosificación de los medicamentos, hizo que éstos sean más seguros y que los ensayos clínicos precisaran un menor número de muestras de sangre de los pacientes.
- Universidad del Estado de Nueva York (SUNY) (EE.UU.): emplean big data para estudiar la esclerosis múltiple y determinar un mejor diagnóstico, tratamiento y posible cura de la enfermedad.
- Instituto de Tecnología de la Universidad de Ontario (UOIT) (Canadá): gracias al big data monitorea bebés prematuros en neonatología para determinar cambios en la presión arterial, temperatura, alteraciones en los registros del electrocardiograma y electroencefalograma, etc.
- Lineberger Comprehensive Cancer Center - Bioinformatics Group (EE.UU.): utiliza Hadoop y HBase para analizar los datos producidos por los investigadores de The Cancer Genome Atlas (TCGA) para trabajar en las investigaciones relacionadas con el cáncer.

## 9. Sector retail:

- Nike (EE.UU.): gracias a la creación de wearables -dispositivo que permite aplicar inteligencia artificial a objetos como relojes o pulsómetros- la compañía está localizando nuevos nichos de mercado.

---

<sup>20</sup>: <https://www-03.ibm.com/press/es/es/pressrelease/46477.wss>.

<sup>21</sup> <http://www.elmundo.es/cronica/2016/07/03/57779fcoca4741301d8b4609.html>).

- Macy's (EE.UU.): estos grandes almacenes estadounidenses utilizan la tecnología del SAS Institute para mejorar sus ingresos y la experiencia del usuario, además de reducir en 500.000 dólares el gasto anual de analítica. Mediante la utilización de un algoritmo y al control de la demanda e inventario, lanzan ofertas cruzadas, ajustan precios y hacen rebajas casi en tiempo real.<sup>22</sup>
- Kroger (EE.UU.): esta empresa de supermercados estadounidenses consiguió mediante big data implantar en sus programas de fidelización de clientes una campaña de correo directo que logró una tasa de uso de cupones de más del 70% en un plazo de seis semanas.
- Target (EE.UU.): esta cadena de distribución fue capaz de predecir cuándo sus clientes esperaban un bebé, asignando un número de identificación único a cada cliente, asociado a su tarjeta de crédito para analizar el historial de compra y después generar cupones de descuento.
- L'Oreal (Francia): desde que tiene implementado el big data en su estrategia empresarial han conseguido una mayor calidad de datos, tener nuevas iniciativas y acumular una visión de 360 grados de toda la información sobre productos.
- Alibaba (China): gracias al big data analiza el comportamiento de sus compradores, pudiendo adelantarse a sus deseos y planificar jornadas como el Día del Soltero, que le permitió facturar hasta 16.300 millones de euros en 24 horas.
- Amazon (EE.UU.): para conocer los gustos de los lectores, la compañía se basa, no sólo en los datos del usuario, sino también en las tendencias que es capaz de generar a partir de los amplios catálogos de datos que recopila. Una de las estrategias para conseguirlo es a través de las recomendaciones que ofrece con cada producto, basadas en las pautas de comportamiento de otros compradores similares al propio usuario.

#### **10. Telecomunicaciones:**

- Netflix (EE.UU.): la reina de las plataformas de series y películas online supo utilizar el big data para conocer los gustos de sus usuarios. Su éxito se debe a las recomendaciones que hace a sus consumidores en base a los patrones de consumo de contenido. Tanto es así, que la serie House of Cards fue creada a partir de los patrones de consumo obtenidos de más de 40 millones de consumidores, después de detectar que lo que atraía al público en ese momento eran contenidos que incluyeran drama, política, sensualidad y poder.
- T-Mobile (Alemania): después de implantar big data, esta empresa de telecomunicaciones consiguió reducir el número de portabilidades hasta un 50% analizando los datos de las quejas y conversaciones que los clientes dejaban en redes sociales.

#### **11. Transporte:**

- Pratt & Whitney (EE.UU.): esta empresa dedicada a la aviación, junto a su socio IBM, está intentando reducir los servicios de mantenimiento no planificado en sus aparatos implantando big data.

<sup>22</sup>: <http://laeradigital11.blogspot.com.es/>.

- Avis Budget (EE.UU.): esta empresa dedicada al alquiler de vehículos ha implementado una estrategia destinada a aumentar la cuota de mercado que le ha proporcionado cientos de millones de dólares de ingresos adicionales. Para ello, su socio tecnológico CSC aplicó un modelo que calculaba el valor de la base de datos de clientes de Avis y, a continuación, lo validaba mediante una campaña de marketing multicanal con análisis simultáneo.<sup>23</sup>

### 2.3.2. Contexto nacional

A continuación se destacan algunas de las iniciativas empresariales divididas sectorialmente sobre big data en España. En este apartado cabría destacar el trabajo llevado a cabo por la empresa Delfos Research<sup>24</sup> que ha elaborado la primera guía en España de empresas que trabajan en el ámbito del big data, con un directorio de 300 empresas que ofertan productos y servicios dirigidos a ayudar a las organizaciones que deseen desarrollar una estrategia de datos.

#### 1. Administración Pública:

- Consult@web: aplicación que lanzó la Consejería de Sanidad de la Comunidad de Madrid para facilitar el acceso a la información de pacientes a médicos de familia, pediatras y enfermeras.
- Análisis de movilidad con información big data en Aragón<sup>25</sup>: el objetivo era conocer el número de desplazamientos medios diarios (en fin de semana y en día laborable) entre las diferentes comarcas aragonesas.
- Smart-Santander: ha comenzado a gestionar su tráfico basándose en big data. Esta aplicación ayuda a los conductores a reducir la contaminación y el tráfico, ofreciendo la mejor ruta para los usuarios.
- Servicio 012: el Centro de Excelencia (CoE) en big data de Barcelona -de Eurecat-, la Generalitat de Catalunya y la empresa Atento pusieron en marcha un proyecto que utilizaba técnicas de analítica avanzada y big data en el ámbito de la atención ciudadana para mejorar la calidad y eficiencia del servicio de atención telefónica del 012.<sup>26</sup>

#### 2. Banca y finanzas:

- La Caixa: este banco ha comenzado un macroproyecto con el apoyo de Oracle para crear una nueva infraestructura big data. El reto, extensible a toda Europa, se basa en mejorar los servicios que ofrecen a sus clientes. En 2012 lanzó una aplicación capaz de detectar las ofertas cercanas al usuario.
- BBVA: es uno de los “reyes” del big data en España. Mediante esta tecnología extrae información de sus clientes (cómo utilizan sus cuentas, qué tipo de movimientos hacen, qué tipo de vida tienen, etc.). Así, consiguen ofrecer productos y servicios más adaptados a sus necesidades.


---

<sup>23</sup>: <http://www.ticbeat.com/empresa-b2b/casos-exito-aplicacion-big-data/7/>

<sup>24</sup> <http://www.computerworld.es/tendencias/ya-disponible-la-guia-de-estrategia-de-datos-2017>

<sup>25</sup> <http://datos.gob.es/es/catalogo/a02002834-analisis-de-movilidad-con-informacion-big-data>

<sup>26</sup>: <http://www.elmundo.es/cataluna/2017/09/19/59cof61746163fab248b4614.html>.



También están trabajando en un proyecto piloto en el que utilizan los datos de pago del pequeño y mediano comerciante, con el objetivo de estimar mejor el riesgo que tiene dicho empresario a la hora de concederles un crédito. Marco Bressan, Chairman & CEO BBVA Data & Analytics, explica que de esta manera la entidad está otorgando aproximadamente un 4% más de créditos.<sup>27</sup>

Asimismo, han creado una aplicación para el sector inmobiliario en la que los usuarios podrán seleccionar el barrio en el que desean vivir para conocer el precio y su evolución, así como la superficie media de los pisos de la zona y el porcentaje de inmuebles disponibles. Además, podrán adquirir información acerca de los colegios, farmacias y supermercados que tienen en su entorno. Otro de los estudios que BBVA elaboró partiendo del big data fue en 2011, cuando analizó el uso de las tarjetas de crédito en España durante la Semana Santa en cuatro sectores: mercados y alimentos, bares y restaurantes, moda y gasolineras.

### **3. Cultura:**

- Taburete: este grupo musical ha alcanzado la fama gracias al big data. Sus componentes contrataron a The Good Click, una empresa de estrategia online, para diseñar su estrategia y analizar tendencias musicales nacidas de bases de datos en plataformas digitales como Spotify, YouTube y otras redes sociales. Estas bases informativas, que se pueden comprar, permiten ilustrar hábitos de escucha.<sup>28</sup>

### **4. Deportes:**

- Real Madrid: fue el primer club en utilizar big data. Entre algunos de los recursos que obtienen al usar esta tecnología destaca el de calcular la longitud y tipos de pases de los jugadores. Esta información se une a la obtenida de un sujetador deportivo que integra un contador de pulsaciones, un acelerómetro en la espalda y un GPS y de ahí se sacan mejores estrategias para los partidos.
- FC Barcelona: ha implantado de la mano de Telefónica un sistema que le permite conocer más en detalle los patrones de juego de los equipos rivales y mejorar los movimientos de sus jugadores.

### **5. Medio ambiente:**

- Aqualia: gracias al big data implanta una tecnología pionera en Baleares de recogida y análisis de caudal y presión de los 278 kilómetros de canalizaciones. A través de los sensores instalados detecta fugas de agua.

### **6. Retail:**


- Inditex: utiliza big data y cloud computing para conocer los gustos que tiene el cliente y diseñarlo y ponerlo a su disposición en un tiempo récord, “monitorizando todo el sistema tecnológico en tiendas, logística y comercio electrónico, lo que

---

<sup>27</sup> <http://datosconciencia.com/big-data-banca/>.

<sup>28</sup>: [https://elpais.com/cultura/2017/03/15/actualidad/1489607603\\_172367.html](https://elpais.com/cultura/2017/03/15/actualidad/1489607603_172367.html).





hace que se conozcan en tiempo real las tendencias, la situación de los pedidos, el volumen de ventas, etc”<sup>29</sup>.

- El Corte Inglés: desde 2012 lleva recopilando los datos de su web, y cuenta con unos consumidores muy recurrentes, lo que le ha permitido disponer de un histórico de datos de gran valor. Han construido un modelo que predice cómo actúan los clientes en determinados momentos y campañas.
- Mercadona: utilizan la plataforma SAP S/4 HANA para proporcionarse información inmediata con análisis en tiempo real, así como para “simplificar y estandarizar de forma masiva procesos y tecnologías mediante un sistema que elimina replicaciones de datos, conciliaciones o redundancias. Todo ello permitirá a Mercadona obtener en todo momento un dato único, veraz y en tiempo real de cualquier punto de sus procesos”.<sup>30</sup>

## 7. Salud:

- ERESA-CETIR: este grupo biomédico de expertos en diagnóstico y tratamiento ha puesto en marcha el proyecto Oncobytes. Este sistema permitirá obtener toda la información acerca de los tumores, pudiendo detectar células cancerígenas en sus fases iniciales, haciendo uso del big data. Gracias al análisis de multitud de datos en función del tipo de tumor se puede llegar a realizar tratamientos más tempranos y personalizados.
- Proyecto Visc+: este programa desarrollado en España utiliza el big data para ofrecer información de pacientes en Cataluña. El objetivo de este proyecto es “relacionar la información de salud que se genera en Cataluña de una manera totalmente anonimizada y segura para impulsar y facilitar la investigación y la innovación en medicina y las ciencias de la salud”.<sup>31</sup>
- Polen control: este proyecto se basa en la combinación de apps para smartphones y el potencial del big data para prevenir o evitar los efectos de las patologías derivadas de la profusión de polen en la atmósfera. En esta aplicación han trabajado la Sociedad Española de Alergología e Inmunología Clínica (SEAIC) y los Laboratorios Almirall.<sup>32</sup>
- SMUFIN (Somatic Mutations Finder): nuevo método basado en big data, desarrollado por un equipo de investigadores españoles y publicado en Nature Biotechnology que hace posible la detección rápida y precisa de los cambios genómicos causantes de la aparición y progresión de tumores, logrando analizar el genoma completo de un tumor y detectando sus mutaciones en pocas horas al mismo tiempo que se localizan alteraciones hasta ahora ocultas<sup>33</sup>.

## 8. Sector agroalimentario:

- Dinsa: usa desde 2015 tecnología big data desarrollada por la startup Cubenube llamada solución bynse. Este software ofrece a gestores y productores agrícolas

---


<sup>29</sup>: [http://www.fundacionorange.es/wp-content/uploads/2016/07/eE\\_La\\_transformacion\\_digital\\_del\\_sector\\_retail.pdf](http://www.fundacionorange.es/wp-content/uploads/2016/07/eE_La_transformacion_digital_del_sector_retail.pdf)

<sup>30</sup>: [https://cincodias.elpais.com/cincodias/2017/07/21/companias/1500626918\\_610182.html](https://cincodias.elpais.com/cincodias/2017/07/21/companias/1500626918_610182.html).

<sup>31</sup> <http://aguas.gencat.cat/es/projectes/visc/index.html>.

<sup>32</sup><http://www.ontsi.red.es/ontsi/sites/ontsi/files/Informe%20big%20Data%20en%20Salud%20Digital.pdf>.

<sup>33</sup>[http://www.immedicohospitalario.es/uploads/2017/07/nuevas\\_maneras\\_hacer\\_12031\\_20170731014939.pdf](http://www.immedicohospitalario.es/uploads/2017/07/nuevas_maneras_hacer_12031_20170731014939.pdf)



la posibilidad de controlar, desde cualquier dispositivo con conexión a Internet, el estado actual y las necesidades futuras de sus cultivos desde tres puntos: suelo, planta y clima.<sup>34</sup>

## 9. Sector energético:

- Endesa: usa estratégicamente big data, entre otros objetivos, para reducir el fraude. También ha puesto en marcha un sistema de 'Machine Learning', después de dos años de trabajo junto a empresas líderes del sector en Estados Unidos y España, que “permite analizar los datos de consumo y contrato de los últimos diez años de cada instalación conectada a la red de Endesa, enriquecidos con fuentes externas y centenares de variables propias de los procesos de 'machine learning’”.<sup>35</sup> Además, la empresa ha incorporado nuevas herramientas tecnológicas para localizar los puntos de la instalación eléctrica donde se ha producido manipulación (a través de un videoscopio y un detector de dobles), que permite localizar derivaciones irregulares, aunque estén enterradas bajo tubos o empotrados en la pared.

## 10. Sector jurídico:

- Chávarri Abogados: gracias al big data predicen cómo va a evolucionar un expediente, ayudando al abogado a prevenir posibles incidencias a corto, medio y largo plazo.

## 11. Sector TIC

- big ML: Empresa del sector TIC centrada en desarrollo de modelos predictivos para desarrolladores utilizando Machine learning. Tienen sede en Valencia y Oregon, USA.
- Inteliderm: Proyecto que nace en el Programa ejecutivo de big data y Business Analytics de EOI (Madrid). Utiliza modelos predictivos de big data aplicado a la cosmética personalizada.
- New Solutions: Empresa del sector TIC creada en 2007, especializada en la utilización de big data para la monitorización y control en el sector del ferrocarril y eólico.

## 12. Telecomunicaciones:


- Telefónica: creó en 2016 LUCA, la unidad especializada en big data que impulsará la transformación digital de los clientes “al permitir a las organizaciones acelerar su viaje hacia el big data a través de la ingeniería de datos y los business insights”<sup>36</sup>. A través de los servicios del IoT -Internet de las

---

<sup>34</sup>: [http://almeria360.com/el-mundo-en-360/21092017\\_la-big-data-una-revolucion-la-agricultura\\_159942.html](http://almeria360.com/el-mundo-en-360/21092017_la-big-data-una-revolucion-la-agricultura_159942.html)

<sup>35</sup>: <http://www.europapress.es/economia/energia-00341/noticia-endesa-destinara-13-millones-2019-combatir-fraude-electrico-causa-agujero-150-millones-20161204120437.html>

<sup>36</sup> [https://www.telefonica.com/es/web/sala-de-prensa/detalle-actos-y-eventos/-/asset\\_publisher/xEr5SqPC9DVI/content/telefonica-aporta-en-espana-el-1-8-del-pib-y-paga-en-impuestos-22-euros-de-cada-100](https://www.telefonica.com/es/web/sala-de-prensa/detalle-actos-y-eventos/-/asset_publisher/xEr5SqPC9DVI/content/telefonica-aporta-en-espana-el-1-8-del-pib-y-paga-en-impuestos-22-euros-de-cada-100).



cosas- y del big data, Telefónica también ha ayudado a empresas y ciudades a evitar emisiones y a mejorar la gestión de los residuos (Smart Waste) y de recursos como la energía y el agua.

Otro de los servicios que ofrece la compañía es poner a disposición de las grandes empresas los servicios de su plataforma de big data Hortonworks de Hadoop: “La plataforma de datos Hortonworks de Hadoop con Telefónica tiene dos funciones esenciales: por un lado, permite a las organizaciones recolectar y almacenar en un repositorio todo tipo de información; por otro, lanzar procesos de transformación de datos, analítica predictiva y prescriptiva sobre esta misma información”.<sup>37</sup> También usan big data para saber dónde poner antenas, para tratar mejor al cliente y prestarles un servicio más efectivo.

- Phone House: están implementando sistemas de sincronización de datos para conseguir que todo el personal pueda acceder a los datos de los clientes, independientemente del canal. El objetivo final es conocer las necesidades del consumidor.

### 13. Turismo:

- Beonprice: especializada en el sector hotelero, usa big data para gestionar variables como los precios de la competencia, demanda y disponibilidad de hoteles para conseguir el precio más económico.
- Segittur: ha llevado a cabo un proyecto centrado en Badajoz para analizar algunos de sus problemas de ocupación en determinadas fechas o la posibilidad de que los comercios abrieran los domingos.
- Patrimonio Nacional: Telefónica ha trabajado con Patrimonio Nacional para analizar la afluencia de turistas en la zona del Palacio Real, las horas a las que transitaban o la nacionalidad de los visitantes, todo ello mediante big data.
- Protección del Patrimonio en Ávila: mediante la disposición de 226 sensores -el 60% ya instalado-, en 29 puntos de la ciudad (la muralla, la catedral, edificios singulares y calles), se está generando información sobre humedades, temperatura, luminosidad, grietas, xilófagos o la vibración por el tráfico en las puertas de acceso a la muralla, contribuyendo así a una mejor conservación del patrimonio.<sup>38</sup>
- BBVA y Ayuntamiento de Madrid: bajo el título “Dinámicas del Turismo en la Ciudad de Madrid” se analizó el comportamiento de los turistas en el terreno comercial durante 2012. El estudio sirvió, entre otras cosas, para cuantificar el impacto económico del Orgullo Gay en diversas zonas de la ciudad. También aportó otros datos interesantes como los turistas que más gastaron, en qué lo gastaron, por dónde se movieron, etc.
- Barcelona: la GSMA -asociación de operadores móviles-, el Mobile World Capital, el Ayuntamiento de Barcelona, Orange y el big data Center of Excellence de Eurecat -grupo especializado en análisis de datos y big data- han elaborado un análisis para detectar el perfil de los turistas en la ciudad condal (de dónde vienen, qué edad tienen, cuál es su poder adquisitivo, en qué gastan su dinero...).

---

<sup>37</sup> <https://www.telefonica.com/es/web/sala-de-prensa/-/telefonica-ofrece-a-las-empresas-espanolas-los-servicios-de-su-plataforma-big-data-con-tecnologia-de-hewlett-packard-enterprise> .

<sup>38</sup> [https://elpais.com/cultura/2017/06/21/actualidad/1498047062\\_122018.html](https://elpais.com/cultura/2017/06/21/actualidad/1498047062_122018.html).

### 2.3.3. Contexto andaluz

#### 1. Cultura:

- Lantia: el negocio principal de Lantia son los servicios de autoedición, en digital y papel, para editoriales, aunque también tiene su propio sello, Samarcanda. Ha firmado recientemente un acuerdo con Telefónica para implantar el big data y el Internet de las cosas en su estrategia empresarial para la creación de una multiplataforma de gestión inteligente del mercado turístico, entre otros.

#### 2. Derecho:

- Legal Innovation: esta startup acelerada en el centro de crowdworking El Cubo, ha lanzado un nuevo producto, Legal Data. Esta innovadora herramienta basada en big data e Inteligencia Artificial es “capaz de predecir los resultados de litigios en función de las búsquedas y parámetros que introduzca el usuario. Gracias a esta tecnología, Legal Innovation pretende reducir la litigiosidad de la justicia española y contribuir a la transparencia y democratización de los procesos judiciales”.<sup>39</sup>

#### 3. Medio ambiente:


- CycleSMARTcity: esta empresa que sale de la Tecnoincubadora El Cubo se encarga de la medición ambiental en tiempo real de las ciudades para los ciudadanos. Su objetivo principal es conectarse a través de la instalación de sensores en las infraestructuras móviles existentes en la ciudad (bici, bus, vehículos de limpieza, etc.) para la medición ambiental de las calles y barrios y mejorar la calidad del aire.
- Irsolav: es una empresa acelerada en el Cable que ha desarrollado una herramienta de big data para facilitar a los productores la información y la previsión de recursos solares con la suficiente antelación y fiabilidad para tomar decisiones de forma más eficiente. Esta herramienta sigue dos pasos, “por un lado, la evaluación detallada de los recursos solares y la producción eléctrica orientada a la inversión y gestión de plantas solares. Por otro lado, la predicción de la producción eléctrica para la optimización de la operación y el mantenimiento de las plantas solares”.
- Geovisor CambiaA: el Grupo de Climatología del Departamento de Geografía Física de la Universidad de Sevilla y la Consejería de Medio y Ordenación del Territorio (CMAOT) de la Junta de Andalucía han creado un geovisor basado en la generación de servicios interoperables OGC (Open Geospatial Consortium) de visualización. “El conocimiento generado en términos de técnicas de geovisualización de big data espaciales ha sido transferido a la Red de Información Ambiental de Andalucía (REDIAM), con el objeto de poder ser aplicado a otros ámbitos de la información climatológica y ambiental”.<sup>40</sup>

#### 4. Salud:

---

<sup>39</sup>: <https://andalucia.openfuture.org/legal-innovation-queremos-abanderar-cambio-tecnologico-sector-legal-espana/> .

<sup>40</sup>[https://idus.us.es/xmlui/bitstream/handle/11441/48791/geovisualizacion\\_de\\_escenarios\\_AEC\\_2016.pdf?sequence=1&isAllowed=y](https://idus.us.es/xmlui/bitstream/handle/11441/48791/geovisualizacion_de_escenarios_AEC_2016.pdf?sequence=1&isAllowed=y) .

- 
- ehCOS SmartICU: esta nueva plataforma tecnológica ha sido creada por Everis y NTT DATA junto al Servicio Andaluz de Salud y el Hospital Universitario Virgen del Rocío. El objetivo es la creación de una nueva solución tecnológica que mejorará el proceso de atención de pacientes en estado crítico en las unidades de cuidados intensivos (UCIs) y ayudará a reducir la morbilidad, mortalidad y tasas de eventos adversos evitables en esta área de atención hospitalaria.<sup>41</sup>

## 5. Sector agroalimentario:

- Agrowing Data: este equipo multidisciplinar ha sabido aprovechar la ventaja competitiva de usar big data en el sector agrícola. Gracias a sus conocimientos sobre agricultura e informática han sido capaces de construir modelos predictivos para conocer mejor la información del campo con business intelligence. Actualmente está en proceso de internacionalización y colaboran con la Universidad de Madrid, con el Ceigram y con expertos en Smart Cities.
- S3 Agrifood de la Comisión Europea: Andalucía lidera la subplataforma temática 'Trazabilidad y big data en la cadena de valor agroalimentaria' dentro de la plataforma de especialización inteligente en agroalimentación S3 Agrifood de la Comisión Europea. El objetivo principal de este proyecto radica en la posibilidad de conocer las experiencias de los consumidores en tiempo real para adaptar los patrones de producción, siendo más efectivos; así como “la de medir, predecir y valorar la sostenibilidad ambiental de la actividad agroalimentaria a través de un sistema ligado a big data para determinar el comportamiento ambiental de los productos, la huella de carbono”.<sup>42</sup>
- Ec2ce: es una empresa especializada en el big data, aplicada al sector agrícola, que tiene proyectos de éxito como el de predecir la plaga de la mosca del olivo. El big data, unido a la Inteligencia Artificial crea “modelos predictivos que ayudan a la industria agroalimentaria a gestionar riesgos, reducir la incertidumbre y tomar decisiones, tanto en el campo como en los mercados”. A través de esta iniciativa han conseguido predecir la plaga con cuatro semanas de antelación y con una efectividad del 100% la primera semana y de hasta el 95% en la semana tercera. Gracias a esta previsión el agricultor reduce sus pérdidas y, además, “el aceite resultante es de mayor calidad, ya que la aceituna picada produce un aceite de mayor acidez que muchas veces tampoco se puede utilizar”.<sup>43</sup>

## 6. Turismo:

- Andalucía Lab y la Fundación CINNTA: han realizado un estudio sobre la reputación hotelera online usando big data. Esta investigación analiza 793 hoteles y 400.000 comentarios obtenidos durante mayo de 2014. Se ha diferenciado la valoración de los hoteles por canales y localización, incluyendo el estudio de 7 atributos específicos: comida, habitación, instalaciones, limpieza, calidad/precio, servicio y ubicación.<sup>44</sup>
- Consejería de Turismo y Deporte de Andalucía: utiliza big data para obtener información acerca del tipo de turista que visita la Comunidad Autónoma.


---

<sup>41</sup>: <https://www.everis.com/spain/es/news/newsroom/everis-y-ntt-data-anuncian-una-nueva-solucion-de-analisis-predictivo-que-mejorara-la>.

<sup>42</sup>: <http://www.lavanguardia.com/vida/20161206/412438207795/andalucia-liderara-una-subplataforma-europea-sobre-trazabilidad-y-big-data-para-la-cadena-de-valor-agroalimentaria.html>.

<sup>43</sup>: <https://www.techfoodmag.com/ec2ce-inteligencia-artificial-y-big-data-agtech/>.

<sup>44</sup>: <https://www.andalucialab.org/blog/big-data-sector-turistico/>

- 
- La Red Andaluza de Posicionamiento: esta red activa de estaciones GNSS - sistema global de navegación por satélite-, gestionada por el Instituto de Estadística y Cartografía de Andalucía, ofrece un servicio público para el posicionamiento preciso en tiempo real. La información, generada como consecuencia de la utilización de la red para fines topográficos, es susceptible de análisis bajo el marco de big data<sup>45</sup>

---

<sup>45</sup>: <http://www.juntadeandalucia.es/institutodeestadisticaycartografia/blog/tag/big-data/> .



## BLOQUE III: INNOVAR Y EMPRENDER CON BIG DATA

### 3.1. Análisis DAFO

Para la elaboración de este estudio, se han realizado 20 encuestas a profesionales del sector (ANEXO 3), así como cinco entrevistas personalizadas (ANEXO 4) a expertos en el campo del big data.

En esta misma línea, se organizó un Focus Group con expertos en la materia de diversos campos: ámbito universitario e instituciones de formación, administraciones públicas y empresas privadas del sector que están en la actualidad a la vanguardia de la aplicación de las tecnologías big data en Andalucía (ANEXO 5).

Estas tres iniciativas se han desarrollado con el objetivo de analizar la situación del sector y profundizar en las actuaciones que permitirían el desarrollo del mismo en Andalucía a través de la puesta en marcha de iniciativas públicas y privadas, con 4 objetivos definidos:

- Creación de nuevas empresas proveedoras de servicios y soluciones de big data.
- Impulsar el uso del big data como instrumento para mejorar la competitividad del sector productivo andaluz.
- Impulsar el uso de soluciones de big data por parte del sector TIC andaluz para la generación de valor añadido dentro de su cartera de servicios.
- Implementación de soluciones basadas en big data en el sector público a fin de mejorar su eficiencia y eficacia en la disposición de recursos públicos.

Teniendo en cuenta los resultados obtenidos durante la fase de exploración, se lleva a cabo un análisis DAFO con las principales fortalezas, oportunidades, debilidades y amenazas detectadas para favorecer el emprendimiento en el ámbito del big data en Andalucía.



## Emprendimiento en el ámbito del big data: DAFO

### FORTALEZAS

- Experiencias positivas de empresas andaluzas, casos de éxito extrapolables.
- Fortalezas del sector turístico, energético y agro con un alto grado de aplicabilidad.
- Existencia de incubadoras de empresas de alto valor tecnológico (CUBO, CABLE).
- Compromiso de grandes compañías multinacionales con el desarrollo de nuevas oportunidades de negocio basados en nuevas tecnologías en Andalucía (Telefónica, Vodafone, IBM, etc).
- Desarrollo de iniciativas públicas en materia de formación (Master EOI, Cátedra de UPO, Estrategia de Impulso del Sector TIC de Andalucía 2020, Andalucía Tech-Universidad de Málaga - Sevilla).
- La capacidad de predecir del big data.
- En el sector agro el big data ayudará a optimizar recursos (tanto humanos como inputs materiales, económicos y ambientales) lo cual además de aumentar la productividad y competitividad de las explotaciones las hace más eficientes en materia de sostenibilidad.
- Importante base de datos del emprendimiento andaluz, tras más de quince años de fomento del emprendimiento de Andalucía Emprende

### OPORTUNIDADES

- Aplicabilidad, capacidad de mejorar la competitividad de las empresas, mejorando procesos y desarrollando modelos predictivos.
- Estado del arte en continua evolución.
- Mayor concienciación de las aplicaciones de big data.
- Empuje de grandes multinacionales proveedoras de servicios.
- Existencia de nuevos nichos de mercado.
- Empuje de gobiernos a través de políticas activas para impulsar la mejora de la competitividad a través del uso del big data.
- Necesidad de dar soluciones en base a algoritmos. Los datos por sí mismos no ofrecen nada si no se les aplica algoritmos que sean capaces de dar respuestas a preguntas concretas.
- Apertura de un campo inmenso para dar soluciones concretas a problemas de los emprendedores y/o empresarios. Lo que se hace actualmente es ofrecer tecnología no soluciones.
- Ante la dificultad de encontrar perfiles con capacidad de interpretar/modificar/adaptar los algoritmos matemáticos que extraen el valor de los datos, se detecta una oportunidad si se ponen los medios para formar a este tipo de profesionales.



---

## DEBILIDADES

- Falta de conocimiento por parte de las pymes.
- Falta de personal especializado en big data.
- Escasa dimensión de las compañías que prestan sus servicios en Andalucía.
- Escaso desarrollo de iniciativas empresariales en el sector de big data.
- Falta de conexión entre técnicos y gestores de pymes, entre posibles soluciones y problemas detectados.
- Alto desconocimiento del concepto y en algunos sectores con mucho potencial en big data, como el Agro, la falta de cultura digital es una debilidad notable.
- No existe una definición clara aceptada por todos respecto a lo que significa big data, es necesario acotar el concepto.
- La falta del volumen de Datos necesario en algunos sectores para poder desarrollar big data.
- Bajo grado de implementación de infraestructuras digitales y de baja calidad. Necesidad de mejorar la calidad de las telecomunicaciones, sobre todo en espacios ligados al sector agro.
- Necesario avanzar en patrones comunes de datos.
- Muy concretas del sector agro:
  - o La estructura de las explotaciones agrarias: pequeñas dimensiones, edad avanzada, etc.
  - o Escasa cultura digital.
  - o Escasa especialización en el Agro de las empresas TICs.

---

## AMENAZAS

- Desconocimiento de las tecnologías por parte de las pymes productoras.
- Dificultad cognitiva del concepto y su aplicabilidad y para identificar casos de negocios atractivos.
- Disponer del volumen necesario de datos y de la calidad de los mismos.
- Falta de conexión entre soluciones y necesidades.
- Falta de personal con habilidades analíticas y técnicas necesarias.
- Desconfianza del empresario respecto a la utilidad.
- Andalucía se está quedando atrás en el desarrollo de esta tecnología, existe mucha competencia y hay zonas geográficas de Europa y resto del mundo que están apostando con mucha más fuerza.
- Limitación normativa de la UE respecto a la utilización de los datos, ello limita la capacidad de competir con otras zonas geográficas donde la normativa es más laxa y la utilización del big data mucho más competitiva. Ello tendrá un doble efecto; las empresas europeas no van a poder competir con las empresas de estos territorios en mejorar su competitividad, y por otra parte se van a generar mayores iniciativas empresariales para dar solución a los problemas concretos de las empresas, ya que los entornos van a ser más competitivos que los andaluces.
- La utilización del big data va a provocar la integración horizontal de sectores. Si Andalucía no está preparada para este reto tecnológico vendrán empresas de fuera a ocupar este terreno, lo que preocupa especialmente en el sector agro.
- Amortizaciones de puestos de trabajo, especialmente de baja cualificación.
- Andalucía debe estar preparada para absorber las nuevas posiciones de personal cualificado que se debería generar con las oportunidades de comercio sobre big data, o serán posiciones para personal de fuera en detrimento del desempleo andaluz.

### 3.2. Factores para emprender con el uso de big data

La transformación digital de empresas es un hecho y una exigencia de futuro. Sin embargo, Europa se encuentra rezagada y mostrando unos indicadores que un informe de McKinsey Global Institute (PDF) considera “alarmantes”. Europa funciona sólo al 12% de su potencial digital, por detrás del 18% de Estados Unidos. Este índice mide la intensidad digital a través de docenas de indicadores, reflejando el estado de transformación digital de diferentes sectores y países.

Un aumento del esfuerzo en la digitalización podría añadir 2,5 billones de euros al PIB en 2025, lo que permitiría añadir un punto porcentual al año durante la próxima década. El mercado único digital podría añadir hasta 415.000 millones de euros al año y proporcionar una plataforma común para que las empresas se beneficien de las economías de escala.

Europa debe redoblar sus esfuerzos en “intensidad digital”, según el informe de McKinsey, que explica este concepto como “el grado en que la digitalización impulsa sectores y empresas”.

Con este escenario, es necesario señalar que uno de los principales factores a tener en cuenta a la hora de desarrollar cualquier tipo de negocio basado en big data es el factor tecnológico, ya no solo a la hora de emprender, que también, si no a la hora de plantearse cualquier tipo de actuación basado en la utilización de un gran volumen de datos.

En primer lugar están todos los aspectos relacionados con la captura de esos datos. El primer obstáculo que existe en este sentido es el acceso a los mismos, ya que no todas los datos disponibles son de libre acceso.

Como ya se ha mencionado, existen dos tipos de datos en referencia a big data:

- **Datos estructurados:** información ya procesada, filtrada y con un formato estructurado. Es el tipo de datos más usado en la actualidad, y en muchos casos no son de libre disposición. Son, fundamentalmente, las instituciones públicas y las grandes corporaciones empresariales las que disponen de este tipo de información, que barajan multitud de información de usuarios, datos estadísticos, informes, recogida de pruebas y análisis y un largo etcétera en multitud de sectores.

No obstante no siempre esta información está disponible para el conjunto de la sociedad, encontrando barreras legales, que impiden que los mismos puedan ser objeto de análisis por parte de las empresas.

Un claro ejemplo se encuentra en Andalucía con la información disponible por parte de la Consejería de Salud o la Consejería de Agricultura, Pesca y Desarrollo Rural, los cuales disponen de bases de datos donde recogen información de todo tipo y que no pueden ser utilizadas por cualquier empresa.

Por otra parte se encuentran con los datos que disponen las grandes compañías suministradoras de servicios (telefonía, energética, telecomunicaciones, retail, entre otros) las cuales han iniciado ya, en la gran mayoría de los caso proyectos, proyectos basados en tecnología big data para mejorar la competitividad de sus

negocios, pero que en ningún caso ofrecen estos datos de manera pública y gratuita.

- **Datos no estructurados:** información sin procesar y que puede adquirir cualquier estructura. Aquí se engloba cualquier formato como textos, imágenes, vídeos, códigos, etc, que se encuentran en redes sociales y en el mundo on-line de cualquier aplicación.

En este punto, por un lado está la barrera derivada de la privacidad de las redes, ya que en multitud de ellas los códigos son cerrados, caso de Facebook o Instagram, por lo que no son de libre acceso a todos los usuarios, y por otro lado las barreras tecnológicas para la captura de los mismos.

No obstante, las barreras tecnológicas son cada vez más salvables a la hora de desarrollar cualquier proyecto de big data. Existen multitud de empresas que ofrecen servicios en la nube que posibilitan el almacenamiento de los datos a bajo coste. Un claro ejemplo de éxito en este campo es la empresa andaluza ubicada en Granada Grupo Trevenque, quien posee uno de los data center más importantes de Europa.

En cualquier caso el uso de los datos no es una preocupación nueva. Ya en el año 2013 las conclusiones del Consejo Europeo se centraban en la economía digital, la innovación y la calidad de los servicios como motores del crecimiento y el empleo. En esta línea se fijaba la necesidad de la intervención de la UE para establecer las condiciones marco adecuadas de un mercado único de los grandes volúmenes de datos (big data).

En Marzo de 2016, el Parlamento Europeo terminó de dar forma a estas directrices y publicó una resolución denominada “Hacia una economía dirigida por los datos” presentada por la Comisión de Industria, Investigación y Energía.<sup>46</sup>

<b>Observatorio EU big data</b>
<b>2017</b>
Tasa de crecimiento mercado TIC x 6 = tasa de crecimiento mercado big data
Movimiento de mercado: 50.000M€ y generación de 3,75 M de empleo
<b>2020</b>
Disponibles 16 billones de Gb
Crecimiento anual de un 236€ en la generación de datos
Crecimiento del 1,9% en el PIB
<b>Necesidades</b>
Mayor control de riesgos en datos de carácter personal
Salvaguarda y protección de los derechos de los consumidores
Equilibrio sectorial en beneficio de explotación big data

*Fuente: Observatorio UE para big data (Fuente: Resolución del Parlamento Europeo sobre la iniciativa «Hacia una economía de los datos próspera» (2015/2612(RSP))*

<sup>46</sup> Resolución del Parlamento Europeo sobre la iniciativa «Hacia una economía de los datos próspera» <http://www.europarl.europa.eu/sides/getDoc.do?pubRef=-//EP//TEXT+MOTION+B8-2016-0308+0+DOC+XML+V0//ES>



Ante esta realidad a nivel europeo, la Junta de Andalucía debe apostar por perseguir la correcta alineación de esta estrategia europea a nivel regional, planteando sobre ello un análisis de fortalezas y necesidades en las entidades del territorio.

De esta manera, y teniendo en consideración que los datos personales han adquirido una importancia económica de gran relevancia en el mercado actual, el objetivo al unificar las normas europeas sobre protección de datos, es fomentar las oportunidades de negocio y la innovación. De entre todas las medidas que se han propuesto por la UE se pueden destacar las siguientes:

- El reglamento establecerá un único conjunto de reglas que simplificarán y abaratarán para las empresas hacer negocios en la UE.
- Las empresas sólo tendrán que hacer frente a una autoridad única de supervisión.
- Las empresas con sede fuera de Europa tendrán que aplicar las mismas reglas que las empresas europeas cuando ofrezcan servicios en la UE.
- Enfoque de las normas basado en el riesgo de cada organización.
- La regulación garantizará que las garantías de protección de datos estarán integradas en productos y servicios desde las primeras etapas de su desarrollo.
- Así mismo, en el ámbito empresarial de la pequeña y mediana empresa, el objetivo desde el Reglamento es la reducción de costes y burocracia llegando a eliminar la exigencia de realizar notificaciones a las autoridades de control, favoreciendo el cobro de cuotas si se hacen solicitudes de datos excesivas, o incluso librándose de la obligación de designar a un “Data Protection Officer” si el procesamiento de datos no es su actividad principal o de la realización de evaluaciones de impacto a menos que exista un alto riesgo.

En este contexto, se detecta que existen cuatro factores diferenciadores que marcarán el futuro del big data y son los más relevantes a tener en consideración de cara a favorecer el nacimiento de iniciativas empresariales en el campo del big data:

- Factores tecnológicos
- Formación y capacitación
- Información y conocimiento
- Ciberseguridad

### **3.2.1. Factores tecnológicos**

Si se analizan las predicciones de fuentes prestigiosas como la revista FORBES, Forrester, Gartner, IDC, el Instituto Internacional de Análisis o algunos informes como el realizado por Planetic, se concluye que las tendencias tecnológicas del sector big data en los próximos años serán las siguientes:


- Los volúmenes de datos continuarán creciendo, hecho absolutamente incuestionable.

- Se mejorarán las vías para analizar datos. Mientras que SQL sigue siendo el estándar en gestión de datos, tecnologías como Spark están emergiendo como herramienta de análisis complementaria, y se prevé un mayor crecimiento.
- Se construirán un mayor número de herramientas prescriptivas de análisis de datos. La mitad del software analítico de negocio, incluirá toda la inteligencia necesaria para antes de 2020.
- El streaming de datos en tiempo real y su ingesta marcará las diferencias en los próximos años. Los usuarios querrán tener disponible datos actualizados que ayuden a tomar decisiones en tiempo real, y a ser posible codificado bajo estándares de fuentes abiertas.
- Otra de las tendencias será el Machine Learning: la explosión actual de fuentes de información y la complejidad de dicha información convierte la clasificación manual y el análisis de datos en algo totalmente inviable, lo que hace necesario un sistema que aprenda de manera autónoma como percibir el entorno.
- big data afrontará retos relacionados con la privacidad, entre otras cosas por la nueva regulación de privacidad establecida por la Unión Europea.
- El modelo de negocio de datos como servicio está en el horizonte, y cada vez más empresas intentarán monetizar sus datos.
- Los mercados de algoritmos también irán creciendo, las empresas preferirán comprar algoritmos que programarlos, y agregar sus propios datos.
- La tecnología cognitiva será la nueva moda. Para muchas empresas, el vínculo entre computación cognitiva y analítica se convertirá en sinónimo, de la misma manera que las empresas ahora ven similitudes entre analítica y big data.
- Para 2020, las empresas que extraen valor de sus datos verán \$ 430 mil millones en beneficios de productividad sobre la competencia que no hace uso inteligente de ellos.
- "Datos rápidos" y "datos procesables" reemplazarán el término big data, según algunos expertos. El argumento es que la parte "grande" (big) no es necesariamente mejor cuando se trata de datos, y que las empresas no utilizan una fracción de los datos a los que tienen acceso. En su lugar, la idea sugiere que las empresas deben centrarse en hacer las preguntas correctas y hacer uso de los datos que tienen, ya sea la colección de estos grandes o no.

Tras este análisis, se plantean las siguientes cuestiones:

- ❖ ¿Es fácil y económico el acceso a las tecnologías basadas en big data?
- ❖ ¿Dentro del sector TIC andaluz existe un alto grado de conocimiento de estas tecnologías?
- ❖ ¿Se promueve por parte de las instituciones públicas españolas y andaluzas la implementación de soluciones basadas en big data?
- ❖ ¿Se promueve por parte del conjunto de administraciones de una manera ágil y flexible el uso de datos para la generación de valor añadido?






Para resolver las mismas, los expertos entrevistados en el Focus Group que han formado parte de la investigación, han llegado a las siguientes conclusiones.

- ✓ El acceso a la Tecnología big data es adecuado, el coste de infraestructuras y de algoritmos es asequible.
- ✓ Lo que falta es market pull o detectar necesidades del mercado. Se detecta en el sector una falta de generadores de negocio. Es necesario que los conocedores de negocio trabajen conjuntamente con el científico de datos, para generar soluciones.
- ✓ No existe dentro del sector TIC Andaluz un alto grado de conocimiento de estas tecnologías.
- ✓ Se recomienda que las instituciones públicas andaluzas y españolas proporcionen más ayudas a universitarios emprendedores para que inicien su empresa en big data.
- ✓ Respecto a los datos, es necesario que las Administraciones Públicas anonimicen los datos que disponen, sobre todo del sector Salud, para ponerlos a disposición de las empresas.
- ✓ El uso de datos o la accesibilidad de datos por parte de las empresas tiene varias limitaciones a día de hoy. Por un lado es necesario un gran esfuerzo económico y de recursos para conseguir poner en abierto los datos públicos disponibles actualmente, y por otro la confidencialidad /propiedad de los datos privados en la normativa europea. Se trata de un problema legal que requiere trabajar en posibles soluciones.
- ✓ Existen importantes iniciativas en el ámbito autonómico para fomentar el desarrollo del big data en Andalucía, llevadas a cabo por la Consejería de Empleo, Empresa y Comercio, así como la Consejería de Agricultura, Pesca y Desarrollo Rural de la Junta de Andalucía, detalladas en el punto 4.

### **3.2.2. Capacitación y formación**

Según Harvar Bussines Review el experto en big data ejerce en la actualidad la profesión más atractiva del siglo XXI. La consultora MC Kinsey asegura que en los próximos cinco años habrá un déficit de especialistas en esta área de 1,5 millones de profesionales, mientras que la consultora Gartner, asegura que el 73% de las compañías invertirán en tecnologías de interpretación y análisis de datos durante los próximos 24 meses.

Ante este panorama, es fácil concluir que uno de los factores esenciales a tener en consideración a la hora de valorar el papel que jugará el big data en cualquier economía avanzada, va a depender en gran medida de la capacidad de formar y capacitar a jóvenes que atiendan las necesidades del mercado.



Es por ello que la falta de capacitación y el conocimiento es uno de los factores que más nos pueden limitar el desarrollo de aplicaciones basadas en big data, y con ello la generación de riqueza en Andalucía.<sup>47</sup>

Tal y como se detalla en el Estudio de mercado sobre talento, capacitación y demanda en big data (designed by big data 4 Success <sup>48</sup>) el talento en big data se encuentra en las universidades, en los centros superiores de formación y en el sector Business Intelligence.

No obstante el big data va a requerir no solo analistas de datos especializados en determinadas aplicaciones, sino multidisciplinares, con un alto índice de especialización transversal a todas las áreas de una compañía, ya que de su análisis y predicción se tomarán decisiones y soluciones que afectarán a todos los departamentos de una empresa u organización.

Estos nuevos profesionales deben ser capaces de entender la manera tradicional de analizar datos y ser capaces de atender nuevos tipos de datos de manera diferentes mediante el descubrimiento de correlaciones que den lugar a nuevas oportunidades de negocio.

En este sentido, las tendencias para los próximos años dentro de esta área para el campo del big data serán las siguientes:

- La mayoría de las compañías necesitarán la figura de un jefe de datos. Este rol será visto cada vez con mayor respeto e importancia.
- Los agentes y cosas autónomas, seguirán siendo tendencia, esto incluye a robots, vehículos autónomos, asistentes personales virtuales y consejeros inteligentes.
- La gran escasez de personal especializado en big data será tratada realizando contratos a analistas y científicos expertos en Datos.
- La crisis de grandes talentos especializados en big data puede aliviarse a medida que las empresas empleen nuevas tácticas a través del reclutamiento y capacitación interna para resolver sus problemas de personal.


En la actualidad el gran hándicap formativo es que la demanda de capacitación ha crecido exponencialmente, pero de momento es incapaz de satisfacer las necesidades del sector, por lo que los profesionales acaban en la mayoría de los casos trabajando para las grandes compañías que a nivel nacional están avanzando más en el campo del desarrollo de iniciativas de big data.

Ante esta realidad uno de los principales objetivos a la hora de plantearse cualquier política de fomento del emprendimiento en el entorno del big data es doble: Por un lado duplicar los esfuerzos en formación y capacitación, y por otro lado la necesidad de retener el talento. En Andalucía existen importantes iniciativas en materia de formación como el Máster de la UPO en big data, dirigido fundamentalmente a estudiantes con

---

<sup>47</sup> Manual sobre utilidades de big data para Bienes Públicos (Instituto Universitario de Investigación Ortega y Gaset)

<sup>48</sup> [http://blog.soydata.net/wp-content/uploads/resumenEntrevistasFondo\\_v31.pdf](http://blog.soydata.net/wp-content/uploads/resumenEntrevistasFondo_v31.pdf)



poca o ninguna experiencia profesional, y el Máster en Business Intelligence y big data de EOI, dirigido a profesionales que requieren una mayor especialización o que quieren redirigir su vida laboral hacia el big data. Ambos casos de éxito pueden servir de guía y estímulo para el lanzamiento de nuevas iniciativas en este campo.

En este contexto surgen las siguientes cuestiones:

- ❖ ¿Existe una oferta de formación en big data estructurada y sólida en Andalucía, tanto a nivel académico como en programas especializados?
- ❖ ¿Son coherentes con las necesidades del mercado los programas y cursos de formación existentes en Andalucía?
- ❖ ¿Las empresas andaluzas están demandando perfiles especializados en big data para su incorporación a proyectos concretos?

Para resolver estos interrogantes, los expertos entrevistados en el Focus Group que han formado parte de la investigación han arrojado las siguientes conclusiones:

- ✓ Se recomienda a las instituciones públicas andaluzas y españolas que realicen formación en big data desde etapas tempranas como Bachillerato, para no quedarse solo en formación de postgrado.
- ✓ Se destaca la importancia de una formación adecuada y adaptada, siguiendo la línea de “Life Long Learning” (aprendizaje permanente)
- ✓ Falta formación en la parte de analítica de datos gestionada con algoritmos.
- ✓ El perfil más demandado es el científico de datos.
- ✓ Las empresas que demandan perfiles especializados en big data no encuentran candidatos por falta de formación o por un desajuste entre nivel salarial con el nivel de estudio/experiencia que requiere este tipo de perfiles cualificados.


### 3.2.3. Información y conocimiento acerca del big data

Según los investigadores del MIT, McAfee y Brynjolfsson<sup>49</sup>, las organizaciones que desean obtener un beneficio de la implantación de soluciones de big data deben gestionar el cambio de forma efectiva en los siguientes cinco aspectos críticos:

- **El liderazgo:** Las organizaciones que destaquen por el uso de soluciones big data no lo harán simplemente porque tengan más o mejores datos, sino porque hayan sido capaces de crear equipos bien liderados que se plantearán las preguntas correctas, fijarán objetivos y definirán métricas.
- **El talento:** La aparición de las nuevas soluciones big data lleva aparejada no solo la creación de nuevas tecnologías de gestión y de analítica de datos, sino también el nacimiento de nuevas profesiones y roles dentro de las organizaciones. Un informe de la consultora McKinsey alerta sobre la escasez de profesionales con estos perfiles y de la necesidad creciente de su reclutamiento en las organizaciones.

---

<sup>49</sup> <https://www.technologyreview.es/s/3615/de-como-la-tecnologia-esta-destruyendo-el-empleo>



Es este rol, el del científico de datos, el que va a resultar clave para se generen iniciativas empresariales en este ámbito.


- **La tecnología:** Las herramientas adquiridas para desarrollar una estrategia basada en la explotación de grandes volúmenes de datos deben cumplir con los requerimientos que cada caso precise. Existen multitud de tecnologías en la actualidad capaces de satisfacer cada necesidad, por lo que en cualquier caso es necesario conocer nuestras necesidades para poder conectar con la tecnología que nos puede ayudar a resolverla.
- **La toma de decisiones:** El primer paso para tomar decisiones acertadas es hacerse las preguntas adecuadas, es decir, definir la pregunta a la que se quiere dar respuesta con precisión. El siguiente paso es identificar las fuentes de información que pueden contribuir a responder esa pregunta, sean internas o externas. Por último, definir los modelos de análisis de esos volúmenes de datos que generan respuestas sencillas a la pregunta inicial.
- **La cultura:** La implantación de soluciones big data en organizaciones precisa de un cambio cultural profundo. Debe imponerse una cultura analítica, basada en la utilización de indicadores, orientada al resultado y a la mejora continua de los procesos. Esta cultura de organización es el sustrato óptimo para poder afrontar innovaciones basadas en la utilización de soluciones de big data.

Llegados a este punto se plantean las siguientes cuestiones:

- ❖ ¿Existe en Andalucía un conocimiento real de las posibilidades que ofrece el big data como instrumento para mejorar la competitividad y generar valor añadido en las empresas?
- ❖ ¿Cuál es el grado de capacitación del empresariado andaluz para liderar los cambios que ofrece el big data?
- ❖ ¿Existe una oferta de servicios basados en big data adecuada a las necesidades del empresariado andaluz?
- ❖ ¿Está cambiando la cultura de las empresas en Andalucía respecto a las nuevas realidades tecnológicas?

Los expertos participantes en el Focus Group que formo parte de la investigación, llegaron a las siguientes conclusiones

- ✓ Se recomienda a las instituciones públicas andaluzas y españolas que realicen más concienciación acerca de las aplicaciones y existencia del big data.
- ✓ Es necesario que el empresariado andaluz se forme y sea consciente de las posibilidades del big data.
- ✓ Es fundamental detectar en primer lugar las necesidades y retos y en segundo lugar buscar la solución TIC. Este orden es importante para que sea efectivo y se demuestre al conjunto del sector los beneficios de la aplicación de estas tecnologías.

- 
- ✓ Necesidad de establecer mecanismos de sensibilización: exposición de proyectos pilotos concretos, divulgación/demostración sobre casos de éxito/ tecnologías implantadas (LABs), charlas empresarios “pioneros” que hagan de atractivo a otros. Así como necesidad de información sobre fuentes de financiación.
  - ✓ Relativo a la divulgación, habría que hacer especial énfasis en demostrar los beneficios (económicos, ambientales, optimización tiempos, etc.) que el big data podría aportar.
  - ✓ Es importante que se extienda y consolide la figura del asesor tecnológico.

#### **3.2.4. Ciberseguridad.**

El mundo del big data está estrechamente relacionado con la ciberseguridad, pues todos los datos a los que se opta parten de la red y necesitan estar protegidos de posibles ataques. Por tanto, el vínculo entre ambos es imprescindible para mantener la información de manera segura en el ámbito digital.

El big data se utiliza como herramienta de análisis y pronóstico, y es dentro de esta actividad desde donde se puede gestionar y analizar toda la información nueva que se recabe, detectando así posibles ataques en tiempo real. Es decir, tanto el big data se sirve de la ciberseguridad, como a la inversa. La protección y la conservación de la integridad de esas fuentes masivas de datos con las que trabaja el big data precisa de expertos especializados en ciberseguridad, y las empresas son conscientes de ello.

Los ciberataques aumentan y las vías de ataque se dispersan, haciendo más complicado mantener la seguridad online. El big data se utiliza también en este campo para detectar potenciales problemas que amenacen la seguridad digital como herramienta de pronóstico. Además, gracias al big data se pueden monitorizar y rastrear sistemas, normalmente contenidos dentro de la nube, para captar irregularidades y posibles brechas.

Otro potencial de la ciberseguridad con big data es usarlo de manera ofensiva. Las empresas no solo quieren big data para tener a buen recaudo sus datos, sino también para identificar a quienes están tratando de entrar a su red o ya están allí, y así atraparlos. Esto se realiza mediante predicciones exactas mezclando datos estructurados con análisis de sentimientos, ubicación y otros datos: “esto requiere de avanzadas tecnologías de machine learning (aprendizaje automático o aprendizaje de máquinas) en conjunto con cloud intelligence, que ya están empezando a surgir”.<sup>50</sup>


### **3.3. Nuevos nichos de mercado: el enfoque sectorial**

big data es una disciplina transversal, que se puede aplicar en todos los sectores, siempre que en ellos se disponga de los siguientes elementos:

1. Gran volumen de datos.
2. Talento y capacitación de las personas para que trabajen con los datos.
3. La existencia de infraestructuras y tecnología para trabajar con datos.
4. La disposición dentro de la organización de una estructura que aproveche el activo de big data.

---

<sup>50</sup> <http://ilifebelt.com/4-maneras-big-data-la-ciberseguridad-se-complementan/2017/05/>.



La mayoría de las organizaciones se acercan al big data con un objetivo centrado en el cliente como la máxima prioridad. Las organizaciones están comprometidas con la mejora de la experiencia del cliente y una mejor comprensión de sus preferencias y comportamientos. big data proporciona la capacidad para comprender y predecir mejor los comportamientos de los clientes y poder mejorar su experiencia.

La analítica de los clientes a través de transacciones, interacciones multicanal, redes sociales, datos sindicados a través de fuentes como tarjetas de fidelidad, entre otros, han servido para que las empresas tengan una imagen clara de las preferencias y demandas de los clientes. A través de esta comprensión las organizaciones encuentran nuevas formas de interactuar con sus clientes.

Otro de los objetivos que pretenden alcanzar las organizaciones con el uso del big data, es la optimización de los procesos, que conlleve la mejora operativa e incremento de beneficios en todos los sentidos.

Igualmente, otra aplicación de big data, es la gestión financiera, la gestión de riesgos, la prevención del fraude, la colaboración de los empleados y la habilitación de nuevos modelos de negocio.

Pero no todas las organizaciones necesitarán gestionar todo el espectro de funcionalidades de big data. Sin embargo, en todos los sectores existe en cierta medida la posibilidad de utilizar nuevos datos, tecnologías y analíticas. Las empresas generan valor al analizar el volumen, la velocidad y la variedad de datos nuevos y ya existentes y al aplicar las habilidades y las herramientas adecuadas para comprender mejor sus operaciones, clientes y el mercado en su conjunto.

En un reciente estudio realizado por IBM en colaboración con el Saïd Business School de la Universidad de Oxford, sobre el uso del big data en el mundo real<sup>51</sup>, en el que realizaron encuestas a 1.144 negocios y profesionales TI de 95 países diferentes, se concluye que existen cuatro fases en la adopción y progresión del big data en una organización:

**1. Educación:** crear una base de conocimiento

En esta primera etapa, se realizan sobre todo acciones de concienciación y desarrollo del conocimiento. En ella los empleados son los encargados de recabar la información, a diferencia de grupos de trabajo formales, y sus conocimientos aún no están siendo usados por la organización. Como consecuencia de esta falta de conocimiento, los directivos aún no han comprendido totalmente el potencial de big data.

**2. Exploración:** definir el caso de negocio y la hoja de ruta


En esta etapa los objetivos son elaborar la hoja de ruta de la organización para el desarrollo de big data y crear un caso de negocio cuantificable y un proyecto de big data.

**3. Organización:** adoptar big data.

---

<sup>51</sup> [https://www-5.ibm.com/services/es/gbs/consulting/pdf/El\\_uso\\_de\\_big\\_Data\\_en\\_el\\_mundo\\_real.pdf](https://www-5.ibm.com/services/es/gbs/consulting/pdf/El_uso_de_big_Data_en_el_mundo_real.pdf)





En esta fase, la organización comienza a comprobar el valor de negocio de big data y valora sus tecnologías y habilidades en este campo.

#### **4. Ejecución:** implementar big data a escala.

En la fase de ejecución, el nivel de operatividad e implementación de las funciones analíticas y de big data es mayor dentro de la organización.

Cuando una organización se enfrenta a la necesidad de incorporar soluciones de big data para mejorar su actividad, la mayor dificultad la encuentra en determinar cómo y dónde abordar el big data.

Para ello IBM identifica cinco casos de uso que pueden servir de guía sobre cómo abordar big data en una organización:

- 1) Exploración de big data: búsqueda, visualización y comprensión de big data para mejorar la toma de decisiones.
- 2) Perspectiva integral mejorada del cliente: ampliación de la perspectiva actual de los clientes con la incorporación de fuentes de información internas y externas. Consiste en el análisis de clientes para poder “comprender las necesidades de los clientes y anticiparse a futuros comportamientos”.

Además de estos análisis que proporcionan perspectivas estratégicas sobre el comportamiento del cliente, la importancia del enfoque integral se aplica a los empleados de primera línea. Las organizaciones más innovadoras reconocen la necesidad de proporcionar a estos profesionales información que ayude a comprometer a los clientes, consolidar las relaciones y obtener resultados positivos como, por ejemplo, solucionar problemas de clientes o fomentar las ventas cruzadas y complementarias de productos. Para ello, estos empleados deben contar con medios para analizar rápidamente grandes volúmenes de información e identificar las necesidades específicas de los clientes.

- 3) Mejora de la inteligencia/seguridad: reducción de riesgos y fraudes, y supervisión de la ciberseguridad en tiempo real.

El enfoque de mejora de la seguridad/inteligencia permite a las organizaciones:

- Filtrar los datos (tanto internos como externos) para detectar relaciones ocultas y patrones, así como para evitar amenazas de seguridad.
  - Detectar casos de fraude al correlacionar el historial de actividad de las cuentas con los datos en tiempo real y, de este modo, localizar transacciones o comportamientos sospechosos.
  - Examinar nuevas fuentes de datos para obtener pruebas de actividades delictivas como, por ejemplo, Internet, los dispositivos móviles, las transacciones, el correo electrónico o los medios sociales.
- 4) Análisis de operaciones: análisis de datos informáticos para obtener mejores resultados de negocio y aumentar la eficacia.

La abundancia y el aumento de los datos de sistemas (generados por equipos y dispositivos de red, sensores, medidores y dispositivos GPS) es otro factor clave para las soluciones de big data.



La combinación de los datos de sistemas con los datos empresariales existentes mediante análisis de operaciones permite a las organizaciones:

- Obtener visibilidad en tiempo real de las operaciones, la experiencia del cliente, las transacciones y los comportamientos.
- Realizar una planificación proactiva para incrementar la eficacia operativa.
- Identificar e investigar las anomalías.
- Supervisar la infraestructura para evitar de forma proactiva las incidencias o la degradación de los servicios.

5) Modernización de los Data Warehouse: integración de big data con las capacidades de almacenamiento para obtener perspectivas de negocio y optimizar la infraestructura existente.

Finalmente IBM recomienda las siguientes claves para que las organizaciones puedan avanzar en sus iniciativas de big data y obtener el máximo valor de negocio:

- Dedicar los esfuerzos iniciales a obtener resultados centrados en el cliente.
- Desarrollar un plan de big data para toda la empresa.
- Comenzar con datos ya existentes para lograr resultados a corto plazo.
- Desarrollar capacidades analíticas sobre la base de prioridades de negocio.
- Crear un caso de negocio sobre la base de resultados cuantificables.

Por otra parte, el informe big data, “El poder de los datos”, realizado por la Fundación Innovación Bankinter<sup>52</sup>, revisa las claves y retos del big data, que sin duda está transformando y revolucionará en los próximos años distintos ámbitos de la sociedad, desde la economía a la Administración Pública, la sanidad o la educación. Hay aún mucho recorrido por hacer y todavía las distintas organizaciones no le están sacando al big data todo el rendimiento y beneficios que ofrece.

Según el citado informe, esta tecnología ofrece un mundo de oportunidades a la hora de incrementar la rentabilidad y eficiencia operativa de las organizaciones o elevar sus ingresos a través de la oferta personalizada de servicios adaptados a las características y circunstancias de cada consumidor, de la retención de los mejores clientes, de la definición de nuevos productos, de la identificación de nuevas oportunidades en el mercado, e incluso de la conversión de los clientes en agentes proactivos de venta. Asimismo permite la reducción de los costes, a través de la optimización de los canales de aprovisionamiento, la limitación de la comunicación a aquella realmente relevante, de la detección del fraude y el dimensionamiento de las plataformas comerciales.

---

<sup>52</sup><https://www.fundacionbankinter.org/documents/20183/42758/Publicaci%C3%B3n+big+data/cc4bd4e9-8c9b-4052-8814-ccb48324147>

Igualmente, el análisis efectivo de los datos permitirá identificar nuevas oportunidades de mercado, acelerar el lanzamiento de nuevos productos, modificar precios, valorar con exactitud el retorno de las inversiones, prever la disposición del cliente a adquirir otro producto o elegir el momento más adecuado para hacer una oferta. “Permitirá generar patrones de comportamiento y predecir hábitos, no sólo para su uso comercial, sino también para combatir la delincuencia, incrementar la inteligencia de las ciudades hasta el punto de poder autogestionarse mediante algoritmos o mejorar aspectos de la vida humana como la salud, el rendimiento personal, deportivo, laboral, etc.”

Aunque la aplicación de big data es enormemente transversal, algunas instituciones y analistas consideran que existen sectores que están mejor posicionados para obtener ganancias a corto plazo mediante big data.

De las encuestas y entrevistas realizadas en la investigación de este trabajo, así como del estudio realizado por PLANETIC, Análisis de la Estrategia big data en España, se exponen a continuación los sectores y nichos de mercado existentes para desarrollar big data<sup>53</sup>:

SECTOR	NICHO DE MERCADO
<b>Tecnologías de la Información</b>	<p><b>Smart Grids</b></p> <ul style="list-style-type: none"> <li>• Mantenimiento predictivo.</li> <li>• Atención personalizada.</li> <li>• Nuevas capacidades de Almacenamiento.</li> <li>• Mejora de la operación de la red.</li> </ul> <p><b>Sector Digital</b></p> <ul style="list-style-type: none"> <li>• Marketing micro personalizado y contextual</li> <li>• Personalización del uso de clientes</li> <li>• Integración de servicios</li> <li>• Analítica para redes</li> <li>• Gestión avanzada de call centers: identificación de problemas en tiempo crítico, maximización de retorno de los clientes y eficiencia en la conservación de clientes.</li> <li>• Robotización.</li> </ul>
<b>Teleco</b>	<ul style="list-style-type: none"> <li>• Eficiencia Energética</li> <li>• Protección Digital</li> <li>• Ciberseguridad</li> </ul>
<b>Turismo</b>	<ul style="list-style-type: none"> <li>• Optimización de procesos.</li> <li>• Conocimiento del cliente.</li> <li>• Análisis de la demanda y del mercado.</li> <li>• Pricing predictivo.</li> <li>• Modelado de clientes.</li> <li>• Valoración y creación.</li> <li>• Atracción y fidelización de clientes mediante técnicas de data mining, machine learning y procesamiento del lenguaje natural, para satisfacer, atraer y retener a los clientes.</li> </ul>

SECTOR	NICHO DE MERCADO
--------	------------------

<sup>53</sup> <http://planetic.es/iniciativa-big-data>

<b>Agroalimentario</b>	<ul style="list-style-type: none"> <li>• Optimización de procesos</li> <li>• Lectura óptica y continua de la maduración y composición química del producto mediante sensores.</li> <li>• Economía del agua.</li> <li>• Predicción de producción o de necesidades existentes.</li> <li>• Control de plagas.</li> <li>• Control de precios.</li> <li>• Open data.</li> <li>• Modelos predictivos.</li> </ul>
<b>Industrial</b>	<ul style="list-style-type: none"> <li>• Protección digital.</li> <li>• Ciberseguridad.</li> <li>• Fabricación personalizada en serie, flexibilidad en la fabricación.</li> <li>• Optimización logística y de procesos de almacenaje.</li> <li>• Mantenimiento predictivo de maquinaria.</li> <li>• Planificación y asignación óptima de recursos de producción.</li> <li>• Predicción de fallos en cadenas de producción.</li> <li>• Optimización logística de suministro y distribución.</li> <li>• Pronóstico de la demanda.</li> <li>• Eficiencia energética en producción.</li> <li>• Despliegue óptimo de bienes y cadenas de producción (layouts).</li> <li>• Seguridad en la planta.</li> <li>• Análisis de riesgos y predicción de fallos.</li> </ul>
<b>Administración Pública</b>	<ul style="list-style-type: none"> <li>• Ahorro de costes por las mejoras operativas de sus procedimientos.</li> <li>• Reducción del fraude y los errores.</li> <li>• Mejorar la recaudación de impuestos.</li> </ul>
<b>Financiero y Seguros</b>	<ul style="list-style-type: none"> <li>• Prevención del fraude interno y externo, detección de incidencias con clientes.</li> </ul>
<b>Salud</b>	<ul style="list-style-type: none"> <li>• Prevención y diagnóstico de enfermedades; mejorar los resultados de los pacientes y reducción del gasto.</li> <li>• Mejorar la eficacia y calidad del servicio prestado</li> <li>• Análisis comparativo de tratamientos</li> <li>• Medicina personalizada y de precisión</li> <li>• Sistemas de soporte a las decisiones clínicas</li> <li>• Automatización de procesos repetitivos</li> <li>• Desarrollo de nuevos medicamentos</li> <li>• Mejora en la calidad de los datos de salud</li> <li>• Integración de información desde los wearables y apps de salud para la monitorización de datos biológicos y enfermedades</li> <li>• Posibilitar la integración y análisis de fuentes de información de salud muy diversas y/o aisladas</li> <li>• Emerge como fenómeno el autocuidado, con nuevas aplicaciones y dispositivos que proporcionan valiosa información sobre salud y hábitos de vida</li> </ul>

**SECTOR**

**NICHO DE MERCADO**



<b>Energético</b>	<p style="text-align: center;"><b>Energías renovables:</b></p> <ul style="list-style-type: none"><li>• Reducción de gases de efecto invernadero.</li><li>• Descentralización y fragmentación de la generación.</li><li>• Sistemas de gestión avanzada.</li><li>• Reducción del coste de la energía.</li><li>• Predicción de producción.</li><li>• Despliegue óptimo de infraestructuras para la recolección de energías renovables</li><li>• Planificación de despacho y almacenamiento de energía renovable producida, basada en criterios económicos (ejemplo: placas solares en casas particulares)</li><li>• Diseño estructural de los elementos que componen las infraestructuras de captación de la energía renovable.</li><li>• Monitorización avanzada de infraestructuras.</li></ul> <p style="text-align: center;"><b>Eficiencia Energética en Edificios:</b></p> <ul style="list-style-type: none"><li>• Mantenimiento predictivo de sistemas de climatización.</li><li>• Predicción de fallos en sistemas de climatización.</li><li>• Normativa en eficiencia energética.</li><li>• Consumo eficiente y responsable.</li><li>• Una gestión más delegada hacia el usuario.</li><li>• Predicción de la Demanda.</li><li>• Detección de fraude eléctrico.</li></ul> <p style="text-align: center;"><b>Gas</b></p> <ul style="list-style-type: none"><li>• Optimización de la Red de infraestructuras.</li><li>• Detección de fraude.</li></ul>
<b>Movilidad y Logística</b>	<ul style="list-style-type: none"><li>• Predicción y trazado de rutas óptimas, basado en información de incidencias, meteorológica y criterios de polución ambiental</li><li>• Transporte multimodal</li><li>• Agregación de flujos logísticos urbanos para análisis agregados de cara a la planificación de la logística de la ciudad.</li><li>• Despliegue óptimo de infraestructuras de coche eléctrico</li><li>• Aparcamiento disuasivo</li><li>• Diseño de planes de emergencia ante eventualidades en la red vial</li><li>• Reclutar todas las calles ante cortes locales en la red vial</li><li>• Detección de puntos de interés según rutas del usuario.</li><li>• Mantenimiento predictivo de infraestructuras viales</li><li>• Gestión de carga de vehículos eléctricos y roaming de gestores de carga</li><li>• Aplicaciones de transporte compartido</li><li>• Optimización de planes de distribución logística</li><li>• Trazado de isócronas geolocalizadas y dinámicas</li><li>• Transporte Urbano eléctrico</li><li>• Cambio de intervinientes: operadores logísticos y zonas Gestión de almacenes inteligente y automáticos</li><li>• Gestión de carga inteligente.</li><li>• Creación de plataformas digitales abiertas que integren de forma óptima los datos de la logística interconectada en gobernanza colaborativa</li><li>• Capacidad de readaptar procesos logísticos en tiempo real.</li></ul>



SECTOR	NICHO DE MERCADO
<b>Retail</b>	<ul style="list-style-type: none"> <li>• Detectar necesidades de los clientes para satisfacer la demanda e incrementar así el margen operativo.</li> </ul>
<b>Media</b>	<ul style="list-style-type: none"> <li>• Optimización de procesos.</li> <li>• Conocimiento del cliente.</li> </ul>
<b>Sector de Envase y Embalaje.</b>	<ul style="list-style-type: none"> <li>• La aplicación práctica de criterios de big data en el sector del packaging mediante la conectividad a través de sensores inteligentes proporciona sin duda mayores cuotas de calidad, eficiencia, seguridad, amén de contribuir a la optimización de costes de todo el proceso.</li> <li>• La interacción de dispositivos tecnológicos con envases activos e inteligentes proporciona gran información acerca del producto: frescura, valor nutricional, etc.</li> <li>• En el caso del packaging farmacéutico, el envase pueden además aportar datos sobre composición, posología, contraindicaciones, modo de conservación, etc.</li> <li>• A través de diversas tecnologías (RFID, NFC, Bluetooth, etiquetado inteligente, etc.) contenidas en el envase se puede obtener información más depurada en cuanto a la trazabilidad, lo que permite a las empresas tomar mejores decisiones en cuanto al "time to value".</li> <li>• De la misma manera la inserción de sensores inteligentes bien en máquinas de envasado, bien en líneas completas de envasado permitirá perfeccionar la automatización del proceso.</li> </ul>

### 3.4. Experiencias internacionales para fomentar el uso del big data

A continuación se señalan una serie de experiencias internacionales para fomentar el uso del big data a nivel de la Unión Europea:


La **BDVA, BIG DATA VALUE ASSOCIATION**<sup>54</sup>, es una asociación sin ánimo de lucro creada bajo la ley belga y parte privada del contrato denominado big data Value PPP (Public private Partnership), cuya parte pública está representada por la Comisión Europea. El objetivo de la PPP es la creación de un mercado de datos funcional y económico en la Unión, tratando de otorgar a la misma un papel destacado en la gestión de grandes volúmenes de datos en el mercado global.

Europa no está jugando el papel que le corresponde en el mercado mundial. Sólo 2 de las 20 empresas punteras que están a la vanguardia tecnológica a nivel usuario y haciendo dinero gracias al procesamiento eficiente de grandes volúmenes de datos, son europeas. Para que esta situación se invierta, Europa tiene que reforzar todas las partes de la "cadena de valor de datos" de modo que se pueda presentar un ecosistema eficiente de grandes volúmenes de datos y modelos de negocio de datos innovadores.

Esto supone que personas y organizaciones estén implicadas en la gestión de datos cualquiera que sea su función, ya sea producción, análisis, o creación de valor y conocimiento a partir de los mismos.

<sup>54</sup> <http://www.bdva.eu/>





La creación de BDVA surge de la alineación de la Comisión con la industria europea (tanto grandes empresas como pymes), los Centros de Investigación y las Universidades, en un Partenariado Público Privado que coopere en la Investigación sobre datos, construir una Comunidad alrededor de los mismos y fijar las bases de una economía próspera basada en los datos.

El acuerdo contractual se firmó el 13/10/2014 y actualmente se está aplicando en el período 2016-2020 a través de convocatorias europeas de Liderazgo Industrial como parte del programa de trabajo de H2020 a través de cuatro instrumentos:

- a. Demostradores a gran escala o Lighthouse en sectores industriales que mejor se pueden beneficiar de big data (transporte, agricultura, bioeconomía, etc.)
- b. Acciones de experimentación e integración de datos o Innovation Spaces: entornos donde convergen proveedores de tecnología y usuarios finales para identificar servicios, habilidades y modelos de negocio.
- c. Proyectos técnicos facilitadores y que supongan un avance en la adquisición de conocimiento y en el uso de metodologías.
- d. Redes, comunidades y soporte a la definición de políticas, de forma que se cree una unión activa de actores principales trabajando con un objetivo común social y organizacional.

Incluidas en el portfolio “big data and Open Data”, las convocatorias de financiación actuales creadas por la Unión Europea en el marco de esta iniciativa son:


- ICT 14 2017: Integración de datos y experimentación intersectorial y en varios idiomas.
- ICT 15 2017: Grandes pilotos en sectores beneficiados por la innovación orientada a datos. Generar soluciones a gran escala que sirvan como estándar y puedan ser replicados con facilidad.
- ICT 16 2017: Investigación dirigida a los retos tecnológicos de la economía de datos, donde se incentiva a los socios industriales con grandes Datasets.
- ICT 17 2017: Apoyo, expertise de la industria, benchmarking y evaluación. CSA de soporte a la construcción de la comunidad y gobernanza del PPP de big data

**SMART DATA FORUM**<sup>55</sup> : fundada por el Ministerio Federal Alemán de asuntos Económicos y Energía, tiene como principales objetivos: (1) generar un ecosistema de innovación que integre distintos actores, incluyendo Industria, Centros de Investigación y políticos, para promover la investigación, innovación y educación en torno a los datos; (2) alinear las diferentes iniciativas relacionadas con big data en el país, enlazarlas con actividades internacionales y apoyar la transferencia de conocimiento y buenas prácticas; y (3) servir como punto de referencia para casos de éxito en distintos sectores, así como punto de contacto para empresas emergentes en el mercado de los datos.

Smart Data Forum juega un rol importante en poner en contacto a los distintos actores e iniciativas centradas en el mercado de datos en Alemania con programas de financiación pública para potenciar proyectos de interés en el área, pero también incluye fondos europeos como el programa H2020.

---

<sup>55</sup> <http://smartdataforum.de/en/>



**RED EUROPEA DE CENTROS DE EXCELENCIA EN BIG DATA<sup>56</sup>** : esta iniciativa está promovida por Know Center, centro de investigación líder en Austria en big data analytics y negocios basados en datos. Aunque impulsada desde Austria con apoyo del Ministerio que posee las competencias en datos, se trata de una iniciativa conjunta con el gobierno Alemán, y por tanto cuenta con el apoyo político de Smart data Forum y del Ministerio correspondiente.

Sus principales objetivos son entender cuáles son los centros de excelencia de big data en Europa y conectarlos de acuerdo a distintos temas de investigación, promocionar la transferencia de conocimiento de la Universidad y centros de Investigación a la Industria, y fortalecer la posición europea en big data. Se trata por tanto de una red más académica, muy orientada a promover cooperación entre centros punteros en big data y establecer un puente que permita que los resultados de investigación de los equipos conjuntos lleguen a aplicaciones industriales. A nivel temático, la red incluye áreas como data analytics, visualización e interacción hombre-máquina, HW, sistemas y HPC, modelos de negocio o educación.

El equipo central de la red está integrado por las siguientes organizaciones: Know-Center, BBDC, INSIGHT, SICS, ADS, UniResearch, aunque cuenta con el apoyo directo, entre otros, del proyecto BDVe y tiene ya más de 60 centros de excelencia de 17 países diferentes.

**BIG DATA FORUM10 (FINLANDIA)<sup>57</sup>**: se trata de una plataforma similar a BDVA, pero cuyas actividades se circunscriben al país de referencia. La inclusión en la misma está sujeta al pago de una tasa cuyo importe depende del tipo de organización. Sus objetivos principales se resumen como sigue: generar conocimiento y soluciones en big data entre Universidad, Centros de Investigación e Industria privada; promover nuevos negocios y mercados donde los datos sean pieza central, esencialmente en la intersección entre big data y el llamado Internet Industrial; promover la colaboración entre los distintos actores para generar un sistema de innovación, y actuar de vehículo hacia otras iniciativas Europeas en big data.

**RISE DATA SCIENCE<sup>58</sup>**: se trata de la iniciativa nacional Sueca para la Innovación centrada en Datos. Al igual que las anteriores actúa como un hub para integrar y combinar capacidades y conocimientos de las diferentes entidades punteras del país y fomentar proyectos de innovación con un impacto mayor en los distintos sectores económicos del país. Sin embargo, en este caso particular hablamos de una entidad única que aglutina a las principales organizaciones de investigación y tecnología del país, como SP Technical Research Institute of Sweden, Swedish ICT e Innventia.

Tiene en su haber 2200 empleados y posee centros de demostración y experimentación abiertos a otras instituciones y clientes. En este sentido cabe destacar SICS ICE Data Center, que es la oferta sueca en cuanto a Data Centers para ejecutar experimentos y demostraciones. Debido a sus capacidades y excelencia ha recibido recientemente la etiqueta de Innovation Space por parte de la BDVA.

**COMUNIDAD BIG DATA DE LA REGIÓN EMILIA ROMAGNA (ITALIA)<sup>59</sup>**: destacada como iniciativa regional comenzó como resultado de una encuesta sobre proveedores de infraestructura que mostró que un 70% de los recursos de computación en Italia estaban concentrados en la región de Emilia Romagna.


---

<sup>56</sup> <https://www.bigdatabcn.com/es/>

<sup>57</sup> <http://www.bigdatasolutions.fi/>

<sup>58</sup> <https://www.sics.se/media/news/rise-data-science-ett-nationellt-initiativ-for-datadriven-innovation>

<sup>59</sup> <http://formazionelavoro.regione.emilia-romagna.it/notizie/big-data-dati-grandi-come-una-regione>



Aunque académica inicialmente, el 7 de febrero de 2016 se abrió también a la industria. La integración de los distintos actores se ve como un elemento necesario para crear una plataforma de ciencia abierta que mejore el conocimiento científico y su transferencia a las 5 o 6 industrias con mayor impacto en la competitividad de las empresas y el crecimiento de la región. La iniciativa involucra ya casi 2.000 investigadores e incluye una larga lista de dominios de trabajo donde se ha hecho un análisis exhaustivo de qué actores están involucrados, qué proyectos se han hecho o están bajo ejecución en las distintas áreas, la infraestructura asociada y los datasets disponibles.

La lista incluye, entre otros, capítulos como: big data en ICT y Contenido Digital, big data en Ciencias de la Vida, big data en Agrifood y Bio-industria, big data en Transporte, big data en materiales, big data en mecánica y procesamiento industrial, big data en cambio climático, big data en medio ambiente y energía, entre otros.

La región de Emilia Romagna tiene previsto invertir 7M Euro de financiación regional precisamente en infraestructuras de Investigación en big data, materiales avanzados y genoma.

Como puede verse, existe una gran diversidad de iniciativas centradas en Datos y particularmente en los retos asociados a big data en Europa. En unos casos se trata de iniciativas más académicas y en otros de plataformas más industriales, regionales o nacionales, pero en todos los casos con la ambición de crear un ecosistema de innovación abierto que permita la generación de conocimiento en big data y su transferencia a casos de uso industrial en uno o varios sectores. Algunas involucran actividades educativas; otras no. Algunas tienen fondos públicos directamente asociados a sus proyectos; otras no.

**S3 AGRIFOOD PLATFORM<sup>60</sup>**. Es una iniciativa que se está llevando a cabo en Andalucía, por parte de la Agencia IDEA y la Consejería de Agricultura, Pesca y Desarrollo Rural para favorecer el uso de las nuevas tecnologías en la cadena alimentaria. Se trata de un proyecto de la UE que lidera esta Consejería dentro de la iniciativa europea para la creación de la plataforma especializada en agricultura y alimentación, “S3 Agrifood Platform”, liderando el lanzamiento de la subplataforma “SMARTFOOD” cuyo objetivo general es apoyar los procesos de innovación en la producción y distribución de alimentos cada vez más sanos y seguros, de mayor calidad y perfectamente trazables, con óptimos ciclos de vida, menor producción de residuos y más eficientes energéticamente.

La S3-Platform es una iniciativa impulsada por la CE y la Red de Regiones Europeas para la Innovación en la Agricultura, Alimentación y Silvicultura (ERIAFF por sus siglas en inglés), para impulsar las actividades que las regiones establecieron en sus estrategias de especialización inteligente (RIS3).

Esto se conseguirá teniendo en cuenta a toda la cadena alimentaria, desde el campo hasta la mesa, buscando la modernización de sus procesos y enfoques a través de una herramienta esencial y común a todas ellas, como el BIG DATA, que permitirá mejorar aspectos clave relacionados con los siguientes conceptos:

- a. Experiencia del consumidor: atendiendo a sus demandas, no solo adaptando la oferta de alimentos a sus exigencias sino también mejorando su experiencia en

---

<sup>60</sup> <http://s3platform.jrc.ec.europa.eu/agri-food>



fases como la decisión de compra, la trazabilidad del producto o la información que desea percibir.

- b. El valor compartido: generando valor añadido en todas las etapas de la cadena y permitiendo que éste se repercuta al resto de fases.
- c. La vigilancia tecnológica: principalmente vigilancia inteligente de la cadena (smartfarming, smartland, smartindustries...).
- d. Cooperación territorial: como base para la transferencia de tecnología y resultados de la investigación, de experiencias, de personal investigador y entre empresas, en las regiones que compartan interés por esta misma temática.
- e. Open Data: la transparencia de los datos públicos y en algunos casos también los privados, para mejorar la toma de decisiones, mejorar las oportunidades de negocio y la mejora del servicio a la ciudadanía por parte de las distintas instituciones.


El lanzamiento y desarrollo de la subplataforma SMARTFOOD influirá de forma muy positiva en aspectos como el incremento de la competitividad del sector agroalimentario, reduciendo su impacto medioambiental, incrementando su eficiencia energética, optimizando los ciclos de vida de los productos agroalimentarios, reduciendo la producción de residuos, mejorando los sistemas de trazabilidad, diversificando la información disponible para los consumidores y creando nuevos productos y servicios.

La Consejería de Agricultura Pesca y Desarrollo Rural está liderando el lanzamiento de esta iniciativa, para lo cual ya ha puesto en marcha un grupo de trabajo que integra a todos los implicados en este ámbito, donde participan, además el Campus de Excelencia Internacional Agroalimentario (ceiA3), la Agencia IDEA y el Instituto de Formación Agraria y Pesquera de Andalucía (IFAPA).

Gracias al trabajo realizado hasta el momento, se ha conseguido el apoyo tanto de la CE como de la ERIAFF (Red de Regiones Europeas para la Innovación en la Agricultura, Alimentación y Silvicultura) para lanzar este proyecto, mediante la tutorización y puesta a disposición del grupo de trabajo de una persona experta internacional, que ayudará en la búsqueda de socios y definición de conceptos y objetivos concretos de la subplataforma.

Por otra parte, y en línea con los proyectos comentados, se recogen otros proyectos destacados en el ámbito internacional y nacional que han sido señalados por los expertos participantes en la investigación, como casos de éxito.

- ✓ Se señala el éxito cosechado en Canadá a través de la creación de un Hub y un centro de inteligencia artificial en Toronto y Montreal, que ha permitido situar a estas dos ciudades a la vanguardia del big data a nivel internacional, permitiendo la atracción de las principales empresas del sector que se dedican a la investigación y comercialización de estas tecnologías. Todo ello mediante la generación de un marco de competitividad promovida por los gobiernos locales y regionales para favorecer la implantación de empresas de alto valor tecnológico dentro de este sector.
- ✓ Otra de las iniciativas a destacar son los Hackathon, puntos de encuentro entre oferta y demanda del sector tecnológico y otros sectores que buscan soluciones a sus necesidades. Por ejemplo los desarrollados por: Endesa Datathon (<https://www.enel.com/es/medios/news/d/2016/02/endsa-datathon-la-innovacin-en-los-datos>); o FIRMART el Agrohackathon organizado por la Diputación Provincial de Córdoba.

- 
- ✓ Destacan los encuentros organizados por la Consejería de Salud de la Junta de Andalucía, sobre Tecnología en general, llamados “Innovando en Jueves”, en los que se muestran casos de éxito del Sistema sanitario Público Andaluz en el área de innovación tecnológica.
  - ✓ Por último señalar que el Ministerio de Agricultura y Pesca, Alimentación y Medio Ambiente, ha creado un Focus Group de Digitalización y big data en el Sector Agroalimentario, Forestal y Medio Rural, del que forma parte la Consejería de Agricultura. Pesca y Desarrollo Rural de la Junta de Andalucía.



## BLOQUE IV: ANÁLISIS DEL BIG DATA EN ANDALUCÍA

### 4.1. Experiencias empresariales en el uso de big data: casos de éxito en Andalucía

Además de las experiencias empresariales en Andalucía detalladas en el apartado 2.3.3. a continuación se exponen algunos casos de éxito de emprendimiento basados en big data en Andalucía, con el objetivo de servir de ejemplo para emprendedores que quieran desarrollar nuevas líneas de negocio en este campo.

**EC2CE:** Empresa creada en 2014 especializada en Inteligencia Artificial aplicada al sector agrícola, cuenta con proyectos de éxito como el de predecir la plaga de la mosca del olivo. El big data, unido a la Inteligencia Artificial crea “modelos predictivos que ayudan a la industria agroalimentaria a gestionar riesgos, reducir la incertidumbre y tomar decisiones, tanto en el campo como en los mercados”. A través de esta iniciativa han conseguido predecir la plaga con cuatro semanas de antelación y con una efectividad del 100% la primera semana y de hasta el 95% en la semana tercera.

**DINAMIC AREA:** Empresa del sector TIC que trabaja con minería de datos y ha desarrollado un producto, opileak, basado en big data que analiza en redes sociales y web en tiempo real, lo que se esté opinando sobre la temática/empresas que se desee analizar.

**GEOGRAPHICA:** Empresa del sector TIC que utiliza tecnologías en Geo big data y Machine Learning para desarrollar modelos predictivos.

**GRUPO TREVENQUE:** Empresa del sector TIC con más de 25 años de experiencia, utiliza big data para los sectores salud y editorial.

**TECHIN CROP AGRORESEARCH:** Empresa de reciente creación, del sector TIC, centrada en el sector Agro. Han desarrollado una solución de mercado, UNIFIT, que permite la monitorización de cultivos.

**AXESOR:** Empresa consultora del sector financiero, que están trabajando con big data para prevención del fraude.

**GENERA GAMES:** Empresa del sector del video juegos está utilizando big data para modelos predictivos de comportamiento de los jugadores, que permitan adaptar la oferta a la demanda.

**EMERGYA:** Empresa del sector TIC que está utilizando big data en el sector salud.

**DATA CO TECHNOLOGIES SOLUTIONS:** Empresa de reciente creación, partner de Oracle, ofrece soluciones en big data.

**MORASCHI DATA STUDIO:** Microempresa del sector TIC que está utilizando big data para análisis predictivo con Azure Machine Learning Studio.

**CV LABS:** Han desarrollado un producto para realizar y analizar el CV propio con Inteligencia Artificial.





**BO TRUE ACTIVITIES:** Empresa del sector TIC que está utilizando big data para análisis predictivo sector Agro.

**AGROPLANNING:** Empresa incubada por Andalucía Open Future, en el CUBO, especializada en el sector Agro, utiliza big data para ofrecer servicios de monitorización de suelos y cultivos y para el control de flotas agrícolas.

**PREOBAR:** Empresa incubada por Andalucía Open Future, en el CUBO, especializada en el sector TIC con soluciones para el sector salud, utiliza big data para ofrecer servicios de monitorización y preparación para la cirugía de pacientes bariátricos mediante sistemas inteligentes de recomendación y soporte a la decisión tanto a pacientes como para equipo médico.

**CICLOGREEN:** Empresa creada en 2013, en el Startup Weekend, especializada en la utilización de big data para Smart Cities. El objetivo de su solución es incentivar la movilidad sostenible, para mejorar las ciudades y hacerlas más habitables, reduciendo las emisiones de CO2 y el uso del vehículo privado.

**ISOTROL:** Empresa andaluza del sector TIC, con 34 años de experiencia, y amplia extensión internacional, está utilizando big data para monitorización y control en el sector de las Energías renovables y Eléctrico, en el que está muy bien posicionada.

**AYESA:** Empresa andaluza del sector Ingeniería y TIC, con más de 50 años de experiencia y amplia expansión internacional, aplica big data al sector eléctrico, en concreto a los contadores inteligentes de electricidad.


**ABENGOA:** Empresa andaluza del sector Ingeniería y TIC centrada en los sectores Energía y Medio Ambiente, con más de 75 años de experiencia y amplia expansión internacional, que cotiza en bolsa, aplica big data en los sectores Energía y Medio Ambiente.

#### **4.2. Nuevos nichos de mercado para el uso del big data en Andalucía: energías renovables, salud, sector agro y turismo.**

A la hora de analizar los nichos de mercado en Andalucía lo primero a resaltar es el hecho de que el big data está de moda. Durante el estudio se han podido resaltar multitud de aplicaciones que tienen el uso de esta tecnología y un alto grado de aplicabilidad. En segundo lugar hay que destacar el hecho de que nos encontramos con una tecnología transversal, de aplicación a todos los sectores y a todos los niveles del proceso de fabricación o prestación de servicios, tanto desde una perspectiva empresarial o desde una perspectiva de servicios públicos.

No obstante, y dada su complejidad y alto grado de aplicabilidad, hay que aclarar que al margen de poner en marcha políticas activas que permitan el desarrollo de esta tecnología de forma transversal, es necesario establecer un campo de acción en el que fomentar el usos de big data en Andalucía desde una perspectiva sectorial, con un doble enfoque.

En primer lugar para que sirva de estímulo y tirón para el resto de sectores, y en segundo lugar para seguir mejorando la competitividad en aquellos sectores en los que Andalucía presenta ventajas competitivas con respecto al resto del sistema productivo.



Bajo estas premisas, y tras la fase de exploración, donde se ha podido analizar los nuevos nichos de mercado con mayor viabilidad en Andalucía, los sectores detectados con mayor potencial serían: energías renovables, salud, sector agro y turismo.

Los motivos se pueden encontrar en la existencia de un tejido empresarial potente y competitivo en la región, en la existencia de abundantes datos en dichos sectores, y por último por la capacidad de mejorar la competitividad en dichos sectores a través de la utilización de análisis basados en big data.

Como ya se ha mencionado, Andalucía cuenta con una de las mejores bases de datos del mundo en sectores como salud, agro y turismo; bases de datos fundamentalmente en manos de la Junta de Andalucía, la cual puede aprovechar estos activos para mejorar la competitividad de dichos sectores a través de técnicas big data.

#### 4.2.1. Sector Agro

Andalucía ha jugado un papel determinante durante los últimos años en la generación de valor añadido dentro del sector, habiéndose convertido en la despensa de Europa liderando las exportaciones de España y la apertura de nuevos mercados dentro del sector agrícola, y a nadie escapa que esto no hubiera sido posible sin haber llevado a cabo una importante transformación y modernización de las explotaciones.

No obstante, nos enfrentamos a nuevos retos, las últimas estimaciones de la FAO arrojan los siguientes datos: en 2050 seremos 9.000 millones de personas en el mundo, las cuales deberán comer un mínimo de tres veces al día. Para esto, la agricultura deberá aumentar los rendimientos hasta un 70% en algunas regiones del mundo, y este aumento de producción, se tiene que realizar de forma sostenible, consumiendo menos agua, menos fertilizantes, y utilizando la misma cantidad del suelo, con el fin de preservar el medio ambiente.


Con este escenario, el big data se posiciona como uno de los principales ejes en el que deberá girar la nueva revolución tecnológica dentro del sector. La capacidad computacional moderna ha permitido aumentar la capacidad de recolectar, intercambiar, procesar y sintetizar datos de una forma tal que está impactando en todo el ámbito agrícola: maquinaria agrícola, optimización de semillas, fertilizantes e insumos, riego y gestión predial.

No debe olvidarse que la producción agrícola es muy compleja ya que interactúan la biología, el clima y las acciones humanas, no obstante con la disponibilidad de GPS y de otras tecnologías, los productores pueden hacer seguimiento a los rendimientos, guiar y controlar las máquinas, monitorear las condiciones del campo y gestionar los insumos a niveles muy precisos dentro de los campos, aumentando de esta manera la productividad y la rentabilidad.

En este sentido las empresas de *big data* pueden evaluar muchísimas variedades genéticas, agroinsumos y condiciones de campos, suelos y climas, así como conectarse con las necesidades de los mercados.

Las principales acciones que se pueden desarrollar en esta línea se resumen en los siguientes puntos:

- ✓ Reducir el gasto de agua hasta en un 40%. Gracias a las sondas instaladas en el suelo de la finca se puede medir la humedad del suelo en todo momento, con lo que se puede ajustar los parámetros de riego para no desperdiciar ni una gota de agua. Además, con las previsiones meteorológicas actualizadas, se puede



prever cuando lloverá, y por tanto retrasar riegos o ajustarlos a dichas condiciones meteorológicas, con lo que el ahorro hídrico es mayor.

- ✓ Reducir un 30% el uso de fitosanitarios. Se ha comprobado que el big data permite predecir cuándo puede atacar una enfermedad o una plaga, gracias a la combinación de datos meteorológicos y el estado del cultivo. Gracias a esta predicción, se puede ajustar la aplicación de fitosanitarios, reduciendo la cantidad necesaria para aplicar en la parcela.
- ✓ Aumenta el rendimiento hasta un 20%. Otro de los aspectos interesantes del big data es que permite ajustar las diferentes labores agrícolas en nuestra parcela, que van desde la poda hasta la recolección, lo que se traduce en poder realizar calendarios de trabajo en el campo mucho más optimizados, permitiéndonos por ejemplo, realizar las podas en el momento adecuado. Recoger el producto en su estado óptimo de maduración, y ajustar la cosecha a los mejores precios en el mercado. Todo esto se traduce no solo en un aumento de rendimiento del cultivo, si no de la rentabilidad del mismo.
- ✓ Mejora la trazabilidad de los alimentos desde el campo hasta la mesa. De forma adicional, el big data, permite registrar todas las fases que recorre nuestro producto agrícola, desde que es plantado, hasta que llega al consumidor, aportando valor añadido a los cultivos; permitiendo incrementar la rentabilidad de las explotaciones agrícolas.

#### **4.2.2. Sector de energías renovables**

España deberá realizar un sprint final si quiere cumplir los objetivos en materia de energías renovables que fija la Unión Europea. Estos compromisos, marcados para todos los países de la UE, imponen el objetivo de que al menos un 20% de la energía consumida en cada estado miembro proceda de fuentes renovables. España está aún a 3,8 puntos de ese objetivo. Puede llegar, pero hay que hacerlo bien. Y la digitalización del sector es un instrumento clave.


Andalucía ha desarrollado un papel fundamental en esta carrera hacia la generación de energías a través de fuentes renovables, desarrollando un tejido empresarial muy competitivo que debe seguir a la vanguardia con la incorporación de este tipo de tecnologías basas en big data.

Los principales campos donde el uso del big data tiene mayor poder de generación de beneficios dentro del sector se resumen en los siguientes puntos:

- ✓ Predicción de producción de energía de origen renovable. Se pueden hacer predicciones a partir de datos de predicción meteorológica y de producción de energía de origen renovable (eólica, fotovoltaica y termosolar). Pueden aplicarse tanto en un parque individual concreto como en áreas más amplias.

Gracias a esta tecnología las empresas que producen o gestionan la energía renovable pueden realizar ofertas precisas de generación al operador del mercado correspondiente y planificar mantenimientos.

- ✓ Predicción de la demanda de energía. En la actualidad se pueden ofrecer servicios de predicción horaria de demanda de energía eléctrica dirigidos principalmente a compañías comercializadoras.

- 
- ✓ Detección de fraude o incidencias en la red de distribución. Haciendo uso de la información estructural de los clientes y del histórico de su consumo, se pueden determinar patrones fraudulentos en los datos de consumo.
  - ✓ Predicción y simulación asociada al vehículo eléctrico: La incorporación de vehículos eléctricos al sistema eléctrico supone un reto considerable, por lo que para afrontarlo de forma eficiente es muy útil contar, por un lado, con escenarios de simulación a largo plazo y, por otro, a corto plazo, con predicción de la carga que supone el parque de vehículos eléctricos.
  - ✓ Mantenimiento predictivo. A través de la instalación de sensores y del análisis de producción de energía en tiempo real, se pueden predecir averías y planificar el mantenimiento de las instalaciones.

#### 4.2.3. Sector salud

Los centros de salud y los pacientes acumulan grandes cantidades de datos en distintos formatos, ya sea en papel o en versión electrónica, que por su dispersión resultan imposibles de utilizar. big data ofrece la posibilidad de organizar información de forma efectiva, por lo que se podrían integrar a los datos estructurados ya existentes hoy en día (fichas personales de los pacientes, etc) y aquellos que permanecen ocultos al sistema actual de almacenamiento y sólo existen de forma analógica en poder de los pacientes (recetas de papel, registros médicos, notas manuscritas de los doctores o resultados de pruebas médicas).


A todos estos datos hay que añadirles los que provienen de las redes sociales y de otros dispositivos que permiten monitorizar al paciente y se irán incorporando a la recopilación de información gracias al internet de las cosas: dispositivos, sensores, instrumentos médicos, aparatos de fitness... Las aplicaciones relacionadas con la sanidad permitirán mejoras en el área médica, en la síntesis de datos de las historias médicas y análisis clínicos, la gestión de dispensarios y hospitales, la administración hospitalaria, la distribución de material sanitario y medicamentos (incluso según necesidades puntuales por epidemias eventuales), la detección y prevención de posibles efectos secundarios de medicamentos y tratamientos, o la generación, almacenamiento y explotación de la documentación científica.

Según un estudio Kinsey Global Institute se calcula que las aplicaciones de big data en el sector sanitario podrían representar unos beneficios de hasta 250.000 millones de euros en los sistemas de salud públicos en Europa y de hasta 300.000 millones de dólares en Estados Unidos. Un ahorro considerable si se tiene en cuenta que la mayoría de sistemas sanitarios de la sociedad occidental presentan unas pérdidas mayores cada año.

Estos ahorros de costes estarían protagonizados por las siguientes actuaciones:

- ✓ Optimización sociosanitaria: La optimización de los servicios sociosanitarios de teleasistencia conlleva segmentar a los pacientes o usuarios en situación de dependencia, predecir su posible evolución, adelantarse a sus necesidades, cubrirlas mejor y mejorar la calidad del servicio a menor coste.

El servicio de optimización sociosanitaria mejora el rendimiento de los servicios sociosanitarios al desarrollar modelos predictivos capaces de predecir patrones de comportamiento y métodos algorítmicos propios de análisis de la información que permiten optimizar la atención que se les presta a los pacientes en situación de dependencia.

- 
- ✓ Segmentación de pacientes crónicos. Predicción de necesidades. Consiste en desarrollar sistemas de estratificación de la población en función de niveles de riesgo para detectar precozmente a aquellos pacientes a los que no sería necesario derivar a servicios de atención especializada, urgencias u hospitalizaciones si recibieran un control proactivo desde la atención primaria.

Tener la población clasificada permite al gestor sanitario incrementar la toma de decisiones proactivas. Igualmente, la predicción de reingresos, del gasto sanitario (individual o por patologías crónicas), etc., supone un impacto económico significativo para las entidades sanitarias

- ✓ Análisis de hiperfrecuentación en atención primaria. La Sociedad Española de Médicos de Atención Primaria define como 'hiperfrecuentador' al paciente que acude a consulta 12 veces al año o más. Esta definición, aunque precisa, es deficiente. Se podría obtener patrones de comportamiento de los pacientes y determinar aquellos perfiles que corresponden realmente a pacientes hiperfrecuentadores.
- ✓ Sistemas de alertas inteligentes, permiten a través de técnicas big data proporcionar datos útiles de apoyo a la práctica de los profesionales sanitarios en el establecimiento de protocolos, el diagnóstico temprano de enfermedades, el pronóstico de la evolución de enfermedades y la planificación del tratamiento de sus pacientes.
- ✓ Optimización de recursos en servicios asistenciales: permite segmentar la población en grupos con características sociosanitarias similares y niveles de riesgo determinados para asignar los servicios y recursos más adecuados en función de las necesidades reales de los pacientes.

#### **4.2.4. Sector Turístico**

El turismo es el principal sector de la economía andaluza, llegando a alcanzar unas cifras del 13% del PIB. Al margen de las ventajas competitivas que tiene Andalucía dentro del sector turístico, es un hecho contrastado los esfuerzos que vienen realizando las empresas e instituciones en promoción del sector, por situarse a la vanguardia de esta industria, tanto a nivel tecnológico, como a nivel de servicios.

Los profesionales del turismo saben que las nuevas tecnologías pueden traer consigo claras ventajas competitivas para la mejora de los productos y servicios ofrecidos, lo que repercute directamente en los beneficios, tanto para el cliente como para la empresa.

Lo tradicional en Turismo venía siendo hacer estudios basados en encuestas, para acercarse a los gustos y preferencias del viajero, e intentar adaptarse así a la demanda del cliente. Pero este tipo de estudios suponían dos inconvenientes principales: el tamaño de la muestra y los sesgos de deseabilidad social.

En la actualidad, cada vez más empresas confían en el uso de análisis big data, con una muestra de datos exponencialmente mayor, y que les aportan además una mayor fiabilidad en los resultados, dado que se nutren ya no solo de las opiniones que vierten los usuarios, sino de la huella digital que dejan en la web, que describe lo que el viajero o futuro viajero efectivamente hace, no lo que cree o desearía hacer.



Los principales beneficios que aportan las tecnologías big data dentro del sector se pueden resumir en los siguientes puntos:

- ✓ Atraer y fidelizar al viajero con big data: Un mejor conocimiento del cliente puede dar como resultado dos claros beneficios: un producto o servicio más adaptado a él y, por tanto, un cliente más satisfecho. Cada vez más el consumidor se está convirtiendo en prosumidor: el viajero, además de consumir un producto o servicio, produce contenido en la web.
- ✓ El análisis de las conversaciones de los usuarios en las redes sociales en relación a una marca o temas de interés unido al estudio de la huella digital (rastreo y análisis de los datos generados por las compras de los viajeros en internet, duración de su estancia, zonas en las que se alojan...) puede ofrecer información de valor para facilitar la captación y fidelización del cliente.
- ✓ A través de una segmentación del público más ajustada, es posible medir el momento y manera óptimos de lanzar un producto o servicio según las preferencias y acciones observadas en los usuarios, prediciendo sus gustos y su propensión a compra.

Por ejemplo, si un estudio aporta información sobre los porcentajes por nacionalidad de los turistas que visitan una ciudad, se puede personalizar el producto para saber satisfacer las necesidades específicas de ese turista: traducir los menús de restaurantes para ese tanto por ciento elevado de turistas que no entran en un restaurante porque no comprenden la carta, ofrecer precios más económicos o productos de alta gama (hoteles de más o menos estrellas, restaurantes familiares o gourmet, etc.) en función de las preferencias detectadas según la nacionalidad del turista, entre otros.

En el caso concreto del turismo, se puede analizar el sentimiento de los comentarios realizados por los turistas en las redes sociales a la hora de visitar una ciudad. Esto es de mucha utilidad a las instituciones ya que es un termómetro de reputación de los servicios que ofrecen a los visitantes.

- ✓ Anticiparse a las necesidades del cliente turístico: Saber lo que va a pasar puede ayudar a anticiparse a los comportamientos tanto del cliente como del mercado. Como se ha comentado, el análisis predictivo puede ser utilizado para conocer mejor al viajero.


Un análisis de los datos puede ofrecer las estancias medias de los turistas: si la estancia media del grupo de turistas en el que se quiere enfocar un tour operador es de dos días, es posible anticiparse a las necesidades del viajero, ofreciéndole paquetes con actividades que se ajusten preferentemente a ese periodo de tiempo y adaptadas a las preferencias turísticas detectadas según las nacionalidades.

- ✓ big data en turismo para mejorar la toma de decisiones: El análisis big data puede ayudar a las empresas del sector turismo a realizar una gestión de su negocio más eficiente, menos costosa; tomando decisiones más inteligentes y fundamentadas en los datos.

El big data facilita saber qué está ocurriendo en cada momento, lo que permite a las empresas del sector turístico reaccionar de una forma más eficiente e inmediata, que no sería posible sin los análisis de datos. Se puede aprender de







los resultados anteriores de una campaña de marketing y modificar en función de los resultados, prediciendo cómo van a responder los usuarios.

El conocimiento de todos estos datos facilita, pues, la toma de decisiones dentro de la agencia turística, hotel, tour operador, etc., empresa, orientando la estrategia comercial, las acciones de marketing y las políticas de calidad.

Un hotel puede llevar a cabo una estrategia de captación de clientes que se base en un canal de comercialización o en otro dependiendo del gasto medio por persona y día que se haya analizado: así, para un segmento de viajeros que tengan un gasto medio, se podría recomendar el uso de intermediarios más que la venta directa (agencias de viajes online y offline, tour operadores, entre otros).

- ✓ Mejorar el impacto de las acciones de marketing. El departamento de Marketing de empresas del sector turístico, como una cadena hotelera o una agencia de viajes, es uno de los más beneficiados por el análisis de los datos, pues permite hacer un seguimiento de las acciones de marketing, y mantenerlas o modificarlas en función de los resultados obtenidos. Con big data se puede segmentar el público objetivo de las campañas, analizar la intención de compra, descubrir el momento adecuado para lanzar un producto o servicio y afinar precios, entre otros.

Una hotelera puede llevar a cabo campañas de reservas online más efectivas basándose en los datos obtenidos de analizar las últimas visitas de sus huéspedes, su duración, sus horarios, etc., para personalizar la estancia: camas, asientos, música, temperatura, baño, iluminación, entre otros, todo para acercarse cada vez con más precisión a los gustos del viajero.

- ✓ Encontrar nuevas oportunidades en turismo. Las empresas que comiencen a invertir en soluciones big data y a comprender el valor de integrar sus fuentes de datos internas con fuentes de datos externas (opiniones vertidas en las redes sociales, información del clima, geolocalización, o incluso balizas inteligentes) podrán disponer de información muy valiosa que puede ser utilizada o bien para mejorar productos y servicios ya existentes (habitaciones de hotel, alquiler de coches, vuelos, entradas a museos...) o bien para proporcionar nuevos, descubriendo nuevas necesidades o detectando, con lo que se conseguirá aumentar los beneficios de una forma exponencial y poner a disposición del viajero 2.0 una experiencia de viaje más inteligente y personalizada.

El sector turístico está viviendo una evolución de mano del mundo digital y los dispositivos móviles, que permiten conocer en tiempo real las opiniones, reacciones y la actividad en la red de los usuarios. Toda esta información a nuestro alcance debe tratarse de forma ágil y eficaz con el fin de generar valor a partir de ella para un mayor conocimiento de los clientes y sus preferencias, identificación de aspectos a mejorar, optimización de los procesos, costes y servicios, prevención de comportamientos perjudiciales, detección de tendencias, etc.

### 4.3. Propuestas de actuaciones a desarrollar

Antes de analizar las posibles acciones a desarrollar, sería interesante conocer las políticas activas aprobadas en la actualidad por la Administración nacional y regional para favorecer el uso del big data, con el objetivo de conocer el marco general de actuación en el que moverse:

En primer lugar y a nivel nacional esta la Agenda Digital para España<sup>61</sup>, la cual establece el desarrollo de un Plan de innovación del sector TIC que busca focalizar los recursos en las iniciativas más relevantes del ámbito digital como el cloud computing, las smart Cities o el big data, para situar a España como un destino atractivo para realizar inversiones en industrias de futuro.

En lo que se refiere a convocatorias específicas para financiación de proyectos en torno a big data a nivel nacional, su adopción se está realizando de una forma moderada. Los proyectos financiados en la línea de I+D desde el Centro para el Desarrollo Tecnológico Industrial (CDTI) van en una línea genérica que no establece temáticas concretas como big data, esto es, no existe una línea específica aunque un proyecto de esta índole sí que puede ser financiado si obtiene valoración positiva por el Centro.

Por lo que respecta a Andalucía las políticas activas en vigor son competencias de la Consejería de Empleo, Empresa y Comercio, la cual dentro del marco de la Estrategia TIC 2020, está fomentando el desarrollo del big data en la región a través de las siguientes iniciativas:


- Realización de un mapa de capacidades Andaluzas en big data.
- Realización de un estado del arte en big data.
- Servicios de consultoría personalizada para empresas TIC con la finalidad de ayudarlas a diseñar nuevos modelos de negocio sustentados en tecnologías big data. El objetivo es ayudar a las empresas a evolucionar de vendedores de tecnología a vendedores de soluciones (previsto 2019).
- Formación en big data:  
<http://www.tic2020.andaluciaesdigital.es/campuscorporativo>  
(Adicionalmente esta previsto lanzar dos cursos de big data para jóvenes desarrolladores).
- Servicios de información y asesoramiento a las empresas del Sector TIC andaluz sobre los instrumentos de ayuda y financiación pública disponible: <http://www.tic2020.andaluciaesdigital.es/oficinadeayudas> (la herramienta permite filtrar ayudas para big data).

Por parte de la Consejería de Agricultura Pesca y Desarrollo Rural de la Junta de Andalucía, se están realizando acciones de concienciación, de fomento de la cultura digital, generar un ecosistema de big data en el sector Agro, y a través de un efecto demostrador con proyectos para dar a conocer al sector Agro la tecnología big data y su potencial.

Es importante destacar que Andalucía ostenta el liderazgo del Partenariado Temático S3P Agrifood sobre Trazabilidad & big data en la cadena de valor agroalimentaria (PT S3AGRIFOOD T&BD).

---

<sup>61</sup> <http://www.agendadigital.gob.es/Paginas/index.aspx>



Este Partenariado reúne a las entidades públicas y privadas clave en trazabilidad y big data de distintas regiones de la UE, tanto Administraciones Públicas, Universidades y centros tecnológicos, como empresas privadas del sector agroalimentario y del sector de nuevas tecnologías de la información y la comunicación, y organizaciones de la sociedad civil (asociaciones de consumidores y fundaciones, entre otras). En total y hasta la fecha lo conforman 20 regiones, 4 clusters público-privados y 3 centros de investigación.

Con este partenariado se pretende ayudar a construir un ecosistema y promover el compromiso interregional que respalda la innovación digital y el fomento de las acciones de mejora a través de la digitalización en toda la cadena del sector agroalimentario, desde el campo hasta el consumidor final, en todas las etapas de la cadena de valor.

También es importante reseñar la creación del “Andalusia Agrifood Digital Innovation HUB” que se trata de un Ecosistema de referencia para toda Europa, a través del cual se conectan Administraciones, centros de competencia, empresas TIC, empresas agro, consumidores y otros agentes para, entre otros objetivos, crear conexiones, impulsar nuevos negocios y acceso a mercados, dar respuesta a retos del sector agroalimentario, etc.

Con este escenario y fruto del trabajo desarrollado en este trabajo de investigación, a continuación se señalan una serie de propuestas de actuación para su desarrollo por parte del conjunto de Administraciones cuyo objetivo es promover el desarrollo de nuevas oportunidades de negocio en el campo del big data. Todas y cada una de las actuaciones propuestas han surgido durante la fase de entrevistas y reuniones mantenidas con los agentes del sector, lo que por una parte valida su conveniencia y por otra resalta su eficacia como instrumentos para el fomento del uso del big data en Andalucía, habiendo sido contrastadas todas ellas como referentes de acciones implementadas por los organismos nacionales e internacionales de más prestigio dentro del sector.

Como punto de partida es necesario resaltar el hecho del doble enfoque que presenta esta tecnología: el enfoque de usuario (empresas y/o emprendedores que comienzan a utilizar tecnología big data para mejorar la competitividad de sus negocios o sus proyectos) y el enfoque de proveedor (empresas y/o emprendedores que ofrecen servicios de big data a sus clientes).

Con este marco se quiere señalar que todas y cada una de las líneas planteadas parten de la idea de convertir a Andalucía en una región usuaria y proveedora de servicios de big data.

#### **4.3.1. Acciones de comunicación, divulgación y formación**

##### **❖ Organización de jornadas y seminarios para dar a conocer los beneficios del big data.**

Se trataría de organizar por parte del conjunto de Administraciones competentes en la materia una serie de jornadas y seminarios con un doble objetivo: dar a conocer las bondades que ofrece el big data como instrumento para mejorar la competitividad de las empresas y por otro ofrecer a los emprendedores y empresas del sector TIC información sobre las tecnologías existentes para aplicar soluciones de big data a sectores concretos, a fin de que puedan incorporar en su cartera de servicios este tipo de tecnologías.



- ❖ **Organización de congresos especializados en big data de carácter sectorial y/o técnico.**

Con el mismo objetivo que el punto anterior se trataría de organizar con cierta periodicidad (anualmente) un congreso en el que se puedan analizar y debatir las últimas tecnologías aplicables y la evolución del sector en cuanto a soluciones, que sirva de punto de encuentro entre tecnólogos, empresas que ofrecen los servicios de big data, usuarios y emprendedores que se inicien en la materia.

- ❖ **Organización de ferias de carácter sectorial, punto de encuentro de oferta y demanda.**

Consistiría en la organización de un encuentro de varios días de duración que serviría de punto de encuentro entre empresas que ofrecen servicios de big data (tanto de infraestructuras como de soluciones) y empresas que buscan soluciones en big data para sus negocios, situando a Andalucía en un referente internacional dentro del sector, para lo que sin duda sería necesario la colaboración de empresas privadas que lideren y patrocinen el evento.

- ❖ **Patrocinio de estudios y tesis sobre la materia.**

Poner a disposición de centros de investigación de las Universidades los fondos necesario para desarrollar soluciones en big data de aplicación al mundo empresarial.

- ❖ **Desarrollo de una mayor oferta formativa por parte de las Universidades andaluzas.**

Coordinación de la oferta formativa, de tal manera que se puedan abordar mayores temáticas relacionadas con el big data, abordando tanto los aspectos técnicos como los relacionados con el negocio.


- ❖ **Apoyo al desarrollo de máster privados, a través de becas.**

Apoyar, a través de becas, la participación en máster de aquellos alumnos más brillantes que quieran acceder a una formación más cualificada, tanto de universidades públicas como de entidades privadas.

#### **4.3.2. Financiación**

- ❖ **Creación de fondos de iniciativa público/privada para el apoyo de proyectos empresariales nuevos.**

Creación de fondos públicos –privados con la finalidad de promover el desarrollo de proyectos empresariales y el emprendimiento, y dotar de más líneas de apoyo financiero a la I+D+i, tanto a nivel europeo como a nivel nacional y regional, con el objetivo de potenciar la colaboración entre los diferentes agentes (empresa, universidad y centros tecnológicos) y generar un entorno favorable para la inversión empresarial en el campo del big data

- 
- ❖ **Compra pública innovadora, política pública de fomento de la contratación pública con PYMES, emprendedores y empresas que realicen una apuesta tecnológica.**

La compra pública de innovación (CPI); se subraya muy especialmente como una de las mejores actuaciones a poner en marcha por parte de la administración para el fomento de la innovación, orientada a potenciar el desarrollo de nuevos mercados innovadores desde el lado de la demanda, a través del instrumento de la contratación pública

- ❖ **Establecimiento de líneas de subvenciones concretas para la implementación de soluciones basadas en big data.**

Tal y como ocurre en la actualidad con otras tecnologías se propone dedicar un apartado específico dentro de las líneas de subvención existentes al big data como instrumento para la mejora de la competitividad de las empresas. Existen casos de éxito que ofrecen subvenciones directamente a las empresas proveedoras de los servicios para que implementen las soluciones en los beneficiarios finales.

- ❖ **Puesta en marcha de acciones de cooperación público-privada para la participación en proyectos europeos y/o internacionales, como los desarrollados por la UE.**

Fortalecimiento de la cooperación público-privada en materia de I+D+i, de manera estable, especialmente en el sector productivo y los agentes generadores de conocimiento y la contribución a la articulación del Sistema Andaluz de Tecnología, apoyando a las empresas andaluzas su participación en proyectos nacionales o internacionales.

#### **4.3.3. Desarrollo sectorial**

- ❖ **Creación de un centro de excelencia en big data de carácter público – privado cuyo objetivo sea generar un ecosistema de innovación que integre a todos los actores: universidades, centros de investigación, empresas e instituciones públicas.**

Se trataría de crear un centro similar al existente en la Región Emilia Romana como una plataforma abierta donde se integren los distintos actores involucrados en esta tecnología: investigadores, empresas tecnológicas, industria, sectores relevantes, universidades, con el objetivo de mejorar el conocimiento científico y su transferencia a las empresas

- ❖ **Creación de centros demostradores de carácter sectorial con la participación de agentes del sector.**

Iniciativa similar al RISE DATA SCIENCE<sup>62</sup>: se trataría de la creación de un hub para integrar y combinar capacidades y conocimientos de las diferentes entidades punteras del país y fomentar proyectos de innovación con un impacto mayor en los distintos sectores económicos de Andalucía, donde se aglutinen a las principales organizaciones de investigación y tecnología del país, para ejecutar experimentos y demostraciones, desarrollando acciones de experimentación e integración de datos o Innovation Spaces: entornos donde convergen proveedores de tecnología y usuarios finales para identificar servicios, habilidades y modelos de negocio.

---

<sup>62</sup> <https://www.sics.se/media/news/rise-data-science-ett-nationellt-initiativ-for-datadriven-innovation>

#### ❖ Puesta a disposición de los agentes del sector de los datos públicos susceptibles de aprovechamiento.

Andalucía cuenta con una de las mayores bases de datos a nivel mundial en temas agrícolas o sanitarios, que en la actualidad no están disponibles para su uso por investigadores y empresas del sector. La idea es generar plataformas de datos abiertos que permitan a las empresas y emprendedores su utilización bajo determinados parámetros de seguridad para el desarrollo de soluciones basadas en tecnologías big data para mejorar la competitividad de los sectores económicos.

#### 4.3.4. Conclusiones:

- ✓ Necesidad de crear un marco legal y administrativo que genere iniciativas empresariales en el ámbito del big data, destacando como ejemplo la puesta en marcha de Digital Innovation Hub y centros demostradores, como iniciativas de éxito a través de la transferencia del conocimiento.
- ✓ Los Hub cubren una de las principales carencias detectadas en este ámbito que es la necesidad de integrar: necesidades, conocimientos, servicios, oferta de tecnología, etc., entre emprendedores TICs y los diferentes sectores.
- ✓ Necesidad de la puesta en marcha de iniciativas que conecten oferta y demanda, necesidades con soluciones de manera sectorial.
- ✓ Se destaca el potencial que tienen las Administraciones para poner a disposición de las empresas y/o emprendedores sus datos como herramienta para buscar soluciones y ponerlas en el mercado, y mejorar la competitividad de los negocios.
- ✓ Se descarta la idea de que sean las Administraciones Públicas las que a través de licitaciones impulsen de manera artificial el sector. Se propone impulsar la compra pública innovadora como instrumento para desarrollar iniciativas en este campo.
- ✓ Se pone de ejemplo la actuación de la Administración Estadounidense durante el mandato del Presidente Obama, con el impulso dado al Open Data, a la normalización de los datos, y la incorporación en los centros de decisión públicos de empresarios de éxito para conocer de primera mano su experiencia y enriquecer la toma de decisiones públicas.
- ✓ Necesidad de suministrar información sobre productos/información relativa a big data a los nuevos emprendedores.
- ✓ Necesidad de recopilar las investigaciones que se están desarrollando en Andalucía y ponerlas a disposición de los emprendedores.
- ✓ Necesidad de aumento de concienciación en los líderes políticos sobre la importancia de la inclusión del fomento de big data en las estrategias políticas futuras.
- ✓ Necesidad de crear un sistema de vigilancia tecnológica que permita conocer las nuevas tecnologías y nuevas soluciones de big data existentes a nivel internacional.
- ✓ Necesidad de medir la eficacia de los proyectos que se están desarrollando hasta el momento.



## ANEXO 1: RESULTADOS DE ENCUESTAS, ENTREVISTAS.

Como ya se ha mencionado en el documento, durante la fase de elaboración del mismo se han realizado 20 encuestas y 5 entrevistas a diferentes profesionales del sector relacionados con el big data.

La relación de personas entrevistadas:

Nombre	Empresa	Ciudad
1. Francisco Huidobro	Orange Spagne S.L.U.	Madrid
2. Pablo Alonso Aguilar	Irradia Energía	Sevilla
3. Victoriano Madroñal Santos	Bolaboo	Sevilla
4. Angel Quiles Morcillo	CEA	Sevilla
5. Ana Cirera	Considera S.L.	Sevilla
6. Juan Reig Redondo	Free Lance Consultant	Sevilla
7. Fabián Varas Sánchez	CTA	Sevilla
8. Luis Reina Juliá	IBM	Madrid
9. Isaac Ferrer Garcia	Instituto de Estudios Cajasol	Sevilla
10. Pablo Adanero González	Dinamic Área Software	Sevilla
11. Pedro Luis Navarro Gómez	Fundación Ssg	Sevilla
12. José Baena De La Iglesia	Everis	Sevilla
13. Fernando Moreno-Torres Camy	Mtc Soft, S.L.	Granada
14. Tomás Cruz Barrios	Endesa	Sevilla
15. Juan Manuel Crespo Pérez	Productos Vegetales Del Mar S.L.	Cádiz
16. Javier Viñuales Gutiérrez	Open Sistemas	Sevilla
17. Gonzalo Fernandez Candau	Expedia	Londres
18. Diego Hueltes	Ravenpack	Málaga
19. José M <sup>a</sup> De La HIGUERA González	SAS-Consejería Salud	Sevilla
20. Antonio Jesús Salvador Rodríguez	Taller Emarketing	Sevilla

Relación de personas encuestadas:

Nombre	Empresa	Institución
1. José García Franquelo	ISOTROL	Empresa
2. Juan Ignacio de Arcos	EOI	Formación
3. Ricardo Arjona	EC2CE	Empresa
4. Norberto Díaz	UPO	Formación
5. Loreto del Valle	Consejería de Innovación	Administración

Relación de participantes en el focus group

Nombre	Empresa	Institución
José García Franquelo	ISOTROL	Empresa
Juan Ignacio de Arcos	EOI	Formación
Ricardo Arjona	ec2ce	Empresa
Ovidio González	Sandetel	Administración
Norberto Díaz	UPO	Formación
Esperanza Perea Acosta	Consejería de Agricultura	Administración
Francisco Alonso de Castro	CEA	Institución
Andoni Recabarren	Ciclogreen	Empresa

## Conclusiones de ENCUESTAS y ENTREVISTAS

- ✓ Nivel de conocimiento del Big Data en el ámbito internacional, nacional y andaluz (siendo 0 el mínimo y 5 el máximo)

Internacional	Nacional	Andalucía
4,15	3,05	2,10

- ✓ Cuáles son las principales implicaciones que destacarías del uso de big data

Actuación	Puntuación
Mejora la toma de decisiones	94
Predecir el comportamiento de tus clientes	93
Optimiza y permite segmentar tu oferta de productos y servicios	89
Análisis predictivo	86
Marketing digital	83
Optimización de procesos productivos	77
Evita incidencias, anticipa problemas y evita el fraude	75
Ciberseguridad	68
Optimización de inventarios y volumen de aprovisionamiento	67
Ahorro de costes de producción	65

- ✓ Cuáles son las principales barreras existentes actualmente para utilizar big data

Barreras	Puntuación
Falta de RRHH con conocimiento especializado	87
Falta de información	81
Tamaño de las empresas	60
Elevado coste de los servicios y las tecnologías	47
Otros	33

- ✓ Qué tipo de actuaciones deben llevar a cabo las instituciones públicas para fomentar el uso del big data

Actuación	Puntuación
Formación	91
Programas de consultoría de apoyo a pymes	85
Iniciativas de Networking	84
Comunicación y difusión	82
Ayudas y subvenciones	74
Premios y distinciones	56
<i>Otros</i>	25

- ✓ Cuáles son los sectores más relevantes actualmente en el uso del big data y cuales consideras que serán los más relevantes en un futuro

Sector	Puntuación a nivel Internacional	Puntuación a nivel Andalucía
Turismo	77	88
Agro	66	78
Energético	80	61
Telecomunicaciones	86	61
Salud	75	60
Retail	77	60
Público	62	59
Financiero y seguros	94	56
Industrial	67	55
<i>Otros</i>	9	5

## ANEXO 2: BIBLIOGRAFÍA

### Libros y publicaciones:

- BBVA (2013): big data. Es hora de generar valor de negocio con los datos. Innovation Edge. Disponible en Internet: <http://www.centrodeinnovacionbbva.com/infografia/big-data-es-hora-de-generar-valor-de-negocio-con-los-datos>
- CABALLERO, R. y MARTÍN, E. (2015): Las bases de big data. Madrid: Los Libros De La Catarata.
- CAMARGO-VEGA, J.J. et al. (2015): Conociendo big data. Fac. Ing. vol.24 no.38 Tunja Jan./Apr. 2015. Disponible en Internet: [http://www.scielo.org.co/scielo.php?pid=S0121-11292015000100006&script=sci\\_arttext&tlng=es](http://www.scielo.org.co/scielo.php?pid=S0121-11292015000100006&script=sci_arttext&tlng=es)
- CARRASCO, F. (2013): Los 6 pasos que su organización debe seguir para confiar en big data. América Latina. CIO, América Latina. Disponible en Internet: <http://www.cioal.com/2013/07/31/los-6-pasos-que-su-organizacion-debe-seguir-para-confiar-en-big-data/>
- CASTELLS, M. (2000): La Sociedad Red. Madrid: Alianza. - \_\_\_\_\_ (2001): La galaxia Internet. Reflexiones sobre Internet, empresa y sociedad. Barcelona: Areté.
- CENATIC. Open Smart Cities II: big data de Código Abierto. Badajoz: Cenatic. Disponible en Internet: [http://observatorio.cenatic.es/index.php?option=com\\_content&view=article&id=804:open-smart-cities-ii-open-big-data-&catid=94:tecnologia&Itemid=137#ftn3](http://observatorio.cenatic.es/index.php?option=com_content&view=article&id=804:open-smart-cities-ii-open-big-data-&catid=94:tecnologia&Itemid=137#ftn3)
- DANS, E. (2011): big data, una pequeña introducción. Disponible en Internet: <https://www.enriquedans.com/2011/10/big-data-una-pequena-introduccion.html>
- DHIRAJ, A., CHAMBERS, M. y MINELLI, M. (2013): big data, big Analytics: Emerging Business Intelligence and Analytic Trends for Today's Businesses. Nueva Jersey: John Wiley & Sons.
- FORRESTER. The pragmatic definition of big data. Cambridge: Mike Gualtieri. Disponible en Internet: [http://blogs.forrester.com/mike\\_gualtieri/12-12-05-the-pragmatic-definition-of-big-data](http://blogs.forrester.com/mike_gualtieri/12-12-05-the-pragmatic-definition-of-big-data)
- GARTNER. big data. Connecticut: Gartner. Disponible en internet: <http://www.gartner.com/it-glossary/big-data/>
- JOYANES AGUILAR, L. (2013): big data. Análisis De Grandes Volúmenes De Datos En Organizaciones. Barcelona: Marcombo.
- LAYNEZ QUIRÓS, M. (2017): Estudio del big data. Sevilla: Universidad de Sevilla.
- LOSHIN, D. (2013): big data Analytics. Massachusetts: MorganKaufmann.

- LUDLOFF, M.E. y CRAIG, T. (2011): Privacy and big data. Massachusetts: O'Reilly Media, Inc.
- MCKINSEY (2011): big data: The next frontier for innovation, competition, and productivity. McKinsey Global Institute. Junio 2011. Disponible en Internet: [http://www.mckinsey.com/insights/business\\_technology/big\\_data\\_the\\_next\\_frontier\\_for\\_innovation](http://www.mckinsey.com/insights/business_technology/big_data_the_next_frontier_for_innovation)
- MCKINSEY (2013): The big-data revolution in US health care: Accelerating value and innovation. Center for US Health System Reform Business Technology Office. Enero 2013. Disponible en Internet: [http://www.mckinsey.com/insights/health\\_systems\\_and\\_services/the\\_bigdata\\_revolution\\_in\\_us\\_health\\_care](http://www.mckinsey.com/insights/health_systems_and_services/the_bigdata_revolution_in_us_health_care)
- MARZ, N. y WARREN, J. (2015): big data. Principles and best practices of scalable realtime data systems. Disponible en Internet: <https://www.manning.com/books/big-data>
- MIRÓN, F. *et al.* (2017): Análisis de la estrategia big data en España. Planetic.
- PANIAGUA, E. (2015): big data, el poder de los datos. Fundación Innovación Bankinter.
- RHEINGOLD, H. (2004): Multitudes inteligentes: la próxima revolución social. Madrid: Gedisa.
- SALVADOR, F. (2014): big data. ¿La ruta o el destino? Fundación Advanced Series. N°3.
- SERRANO-COBOS, J. (2014): big data y analítica web. Estudiar las corrientes y pescar en un océano de datos. El profesional de la información, v. 23, n°6, noviembre-diciembre, pp. 561-565. Disponible en Internet: <http://www.elprofesionaldelainformacion.com/contenidos/2014/nov/01.html>
- SERRAT MORROS, R. (2013): big data. Análisis de herramientas y soluciones. Barcelona: Everis, Facultat d'Informàtica de Barcelona y UPC.
- SOARES, S. (2012): Not Your Type? big data Matchmaker On Five Data Types You Need To Explore Today. Disponible en Internet: <http://www.dataversity.net/not-your-type-big-data-matchmaker-on-five-data-types-you-need-to-explore-today/>
- TASCÓN RUÍZ, A.M. y COULLAUT SANTURTÚN, A. (2016): big data. Madrid: Los Libros De La Catarata.
- TASCÓN, M. (2013): big data. Revista TELOS (dossier). N°95, junio-septiembre. Disponible en Internet: [https://telos.fundaciontelefonica.com/docs/2013/11/11/11400001\\_4\\_4\\_0.pdf](https://telos.fundaciontelefonica.com/docs/2013/11/11/11400001_4_4_0.pdf)
- WALLACE, P. (2001): La Psicología de Internet. Barcelona: Paidós.

## Recursos digitales:

- <http://www.iic.uam.es/innovacion/big-data-caracteristicas-mas-importantes-7-v/>
- <http://www-01.ibm.com/software/data/bigdata>
- <http://www.ibm.com/developerworks/ssa/local/im/que-es-big-data/index.htm>
- <https://www.tableau.com/es-es/resource/top-10-big-data-trends-2017>
- <https://www.winshuttle.es/big-data-historia-cronologica/>
- <http://hadoop.apache.org/>
- [http://www.trc.es/pdf/descargas/big\\_data.pdf](http://www.trc.es/pdf/descargas/big_data.pdf)
- <http://www.cloudera.com/content/cloudera/en/why-cloudera/hadoop-and-big-data.htm>
- <http://www.puromarketing.com/12/23633/big-data-como-grandes-empresas-utilizan-estrategia-producto.html>
- <http://www.iic.uam.es/innovacion/herramientas-big-data-para-empresa/>
- <http://www.datamation.com.ar/dominando-big-data-con-herramientas-open-source-6357>
- <http://www.ticbeat.com/empresa-b2b/casos-exito-aplicacion-big-data/4/>
- <http://blog.es.mailify.com/envios-newsletter/5-casos-de-exito-en-big-data-marketing/>
- <https://www.bbva.com/es/ejemplos-reales-uso-big-data/>
- <https://www.baoss.es/big-data-para-pymes-caso-de-exito-de-una-inmobiliaria/>
- <http://www.intelectium.com/big-data-la-primer-inversion-35-las-grandes-empresas-28-las-medianas-20-las-pequenas/>
- <http://www.customercentric.info/is-bdaas-next-years-big-thing/>
- <http://web.mit.edu/smadnick/www/wp/2014-06.pdf>
- <http://www.zdnet.com/big-data-as-a-service-is-here-but-is-anybody-ready-7000013257/>
- <http://cmsreport.com/articles/big-data-as-a-service-has-arrived-5805>
- <http://upcommons.upc.edu/bitstream/handle/2099.1/19855/90807.pdf?sequence=1>
- <https://www.paradigmadigital.com/ecosistema-big-data/>
- <https://repositorio.unican.es/xmlui/bitstream/handle/10902/4528/TFM%20-%20David%20L%C3%B3pez%20Garc%C3%ADaDaS.pdf?sequence=1>





- <https://www.ibm.com/developerworks/ssa/local/im/que-es-big-data/>
- <http://blog.soydata.net/entradas/el-big-data-para-combatir-el-cancer/>
- [http://www.silicon.es/las-empresas-espanolas-empiezan-comprender-las-ventajas-del-big-data-47366?inf\\_by=59c7ba63671db839478b459a](http://www.silicon.es/las-empresas-espanolas-empiezan-comprender-las-ventajas-del-big-data-47366?inf_by=59c7ba63671db839478b459a)
- [http://www.ey.com/es/es/home/ey-informe-sobre-big-data-y-analytics-en-el-sector-financiero-espanol#.WdTN\\_2i0PIU](http://www.ey.com/es/es/home/ey-informe-sobre-big-data-y-analytics-en-el-sector-financiero-espanol#.WdTN_2i0PIU)
- [http://www.habber.com/es/files/2014/10/Casos-de-uso-de-big-Data\\_Wolfram.Rozas\\_1.pdf](http://www.habber.com/es/files/2014/10/Casos-de-uso-de-big-Data_Wolfram.Rozas_1.pdf)
- <http://www.centrodeinnovacionbbva.com/blogs/emprendedores/post/casos-de-innovacion-traves-de-big-data>
- [http://www.fundacionorange.es/wp-content/uploads/2016/07/eE\\_La\\_transformacion\\_digital\\_del\\_sector\\_retail.pdf](http://www.fundacionorange.es/wp-content/uploads/2016/07/eE_La_transformacion_digital_del_sector_retail.pdf)
- <https://andalucia.openfuture.org/agrowingdata-el-nexo-entre-el-sector-agricola-y-el-big-data/>
- <https://www.winshuttle.es/big-data-historia-cronologica/>
- [http://cdn2.hubspot.net/hub/239039/file-1955764138-pdf/docs/%5BIC%5D\\_OFFER\\_-\\_EBOOK\\_-\\_Las\\_10\\_mejores\\_estrategias\\_de\\_big\\_data\\_en\\_retail.pdf](http://cdn2.hubspot.net/hub/239039/file-1955764138-pdf/docs/%5BIC%5D_OFFER_-_EBOOK_-_Las_10_mejores_estrategias_de_big_data_en_retail.pdf)
- <https://www.cronista.com/especiales/big-Data-mejor-informacion-mas-ventas-20170922-0004.html>
- <http://www.elmundofinanciero.com/noticia/70425/empresas/big-data-para-evitar-fugas-de-clientes-en-el-sector-bancario.html>
- <http://www.diariodeibiza.es/pitiuses-balears/2017/09/19/big-data-prevenir-fugas-agua/941119.html>
- <https://www.tableau.com/es-es/resource/top-10-big-data-trends-2017>
- <http://www.europapress.es/economia/noticia-sector-big-data-facturara-mas-5500-millones-euros-2018-europa-cunef-20170329173524.html>
- <https://blog.ferrovial.com/es/2017/04/big-data-control-de-contaminacion-ciudad/>
- [https://retina.elpais.com/retina/2015/12/06/tendencias/1449379780\\_144937.html](https://retina.elpais.com/retina/2015/12/06/tendencias/1449379780_144937.html)
- <https://aunclidelastic.blogthinkbig.com/big-data-en-el-deporte-dopaje-tecnologico-o-mejor-toma-de-decisiones/>
- <http://www.elmundo.es/andalucia/2017/10/06/59d7b7ad268e3e81388b45ae.html>






## ANEXO 3: GLOSARIO

- **ANALÍTICAS EN TIEMPO REAL:** es la capacidad de las nuevas herramientas tecnológicas de big data para analizar grandes conjuntos de datos en el mismo momento en el que se están generando dichos datos.
- **ANALÍTICAS PREDICTIVAS:** se generan como resultado de la extracción y análisis del conocimiento histórico que se encuentra en los datos masivos y señalan posibles patrones, resultados o tendencias futuras.
- **APACHE COUCHDB.** es una base de datos NoSQL, que carece de un esquema o de estructuras de datos predefinidas como las tablas en las bases de datos relacionales, la información almacenada son documentos JSON. Y la estructura de los datos o documentos puede acomodarse a las necesidades de cambio o a la evolución del software.
- **API:** la Interfaz de Programación de Aplicaciones es el punto de contacto a través del cual un informático puede acceder a una aplicación sin necesidad de conocer su funcionamiento, usando el lenguaje informático.
- **APRENDIZAJE COMPUTACIONAL AUTOMÁTICO:** modelo de autodidáctica computacional basado en inteligencia artificial y que se realiza a partir de análisis comparativos de datos estructurados, semiestructurados y no estructurados. A menudo requieren interacción humana por parte del analista y se nutren de ella.
- **APRENDIZAJE MULTITAREA:** es un modelo aprendizaje computacional automático que analiza de forma conjunta varios problemas o tareas relacionadas para extraer los puntos comunes y más útiles para así mejorar la situación de partida.
- **AVRO:** es un proyecto de Apache que provee servicios de serialización. Cuando se guardan datos en un archivo, el esquema que define ese archivo es guardado dentro del mismo; de este modo es más sencillo para cualquier aplicación leerlo posteriormente puesto que el esquema está definido dentro del archivo.
- **BATCH:** es una ejecución de una serie de trabajos o procesos informáticos que no requieren de la intervención manual para iniciarse.
- **BASE DE DATOS:** una base de datos o banco de datos (en inglés: database) es un conjunto de datos pertenecientes a un mismo contexto y almacenados sistemáticamente para su posterior uso.
- **BASES DE DATOS NOSQL:** son estructuras que nos permiten almacenar información en aquellas situaciones en las que las bases de datos relacionales generan ciertos problemas debido principalmente a problemas de escalabilidad y rendimiento de las bases de datos relacionales donde se dan cita miles de usuarios concurrentes y con millones de consultas diarias.
- **BACK-OFFICE:** en términos informáticos el back-office es el software que procesa información de la empresa como bases de datos, registros, operaciones, inventarios, etc. y que se encarga del mantenimiento de esta información de trastienda.

- **BIG TRANSACTION DATA:** es un tipo de dato que se utiliza en registros de facturación, en telecomunicaciones registros detallados de las llamadas (CDR), etc. Estos datos transaccionales están disponibles en formatos tanto semiestructurados como no estructurados.
- **BIOMETRICS:** es un tipo de dato en la que se incluye huellas digitales, escaneo de la retina, reconocimiento facial, genética, etc. En el área de seguridad e inteligencia, los datos biométricos han sido información importante para las agencias de investigación.
- **BIT:** unidad básica de información en la informática y las comunicaciones digitales.
- **BSON:** es una serialización binaria con codificación similar de documentos JSON. BSON apoya la incorporación de documentos y matrices, también contiene extensiones que permiten la representación de los tipos de datos que no son parte de la especificación JSON.
- **B-TREE:** estructuras de datos de árbol que se encuentran comúnmente en las implementaciones de bases de datos y sistemas de archivos.
- **BYTE:** unidad de información digital formada por una secuencia de ocho bits
- **CASSANDRA:** es una base de datos no relacional distribuida y basada en un modelo de almacenamiento de <clave-valor>, desarrollada en Java. Permite grandes volúmenes de datos en forma distribuida.
- **CHUKWA:** diseñado para la colección y análisis a gran escala de "logs". Incluye un toolkit para desplegar los resultados del análisis y monitoreo.
- **CLUSTER:** según Strauch, es otro enfoque para la partición de datos que se esfuerza por la transparencia hacia los clientes que deberían manejar un grupo de servidores de bases de datos en lugar de un único servidor.
- **COUCHDB:** es una base de datos que abarca completamente la red, utiliza documentos en JSON para guardar los datos, permite acceder a los datos desde un navegador web a través del protocolo http y permite realizar operaciones utilizando JavaScript.
- **CONSOLA DE COMANDOS:** también conocida como línea de comandos, es una herramienta que permite a los usuarios de una aplicación darle instrucciones a través de comandos de texto sencillos.
- **CROWDSOURCING:** técnica que consiste en implicar a las masas en la búsqueda de soluciones a problemas complejos. En el ámbito corporativo requiere abrir los datos necesarios para el trabajo de las personas ajenas a la organización.
- **CUADRO DE MANDO:** pantalla que presenta todas las métricas clave del negocio -los KPI más importantes- en un solo lugar. Esboza una imagen del panorama general con respecto a esos indicadores y facilita la detección de tendencias.
- **DASHBOARD:** es una interfaz desde la que el usuario puede controlar y administrar una aplicación.

- DATO: representación de hechos referentes a una persona, DATAFLOW: diseño de ejecución de una serie de procesos o autómatas que se comunican enviándose información a través de distintos canales.
- cosa o transacción. Incluyen atributos, variables cuantitativas y cualitativas
- DATOS BASURA O DIRTY DATA: son grandes volúmenes de datos inútiles que no aportan información de valor en un contexto de análisis big data.
- DATOS LIMPIOS O CLEAN DATA: son aquellos datos de calidad que determinarán los resultados del análisis big data.
- DIRECTIVA DE GRUPO (GPO): sistema de configuración de la infraestructura informática que sigue cada organización y que determina lo que los usuarios pueden hacer en su equipo dentro del entorno computacional
- ERLANG: lenguaje de programación utilizado para construir sistemas escalables en tiempo real, con alta disponibilidad.
- ESCALABILIDAD: es una propiedad de un sistema que indica su capacidad de reacción y de adaptación a los cambios de envergadura, ya sean al crecer o disminuir.
- ETL: proceso de transformación de datos realizado para extraer datos de una fuente y almacenarlos en una base de datos o data warehouse.
- FRAMEWORK: conjunto de conceptos y tecnologías que sirven de base para el desarrollo de aplicaciones. Acostumbra a incluir bibliotecas de software, lenguajes, soportes a aplicaciones de desarrollo...
- FLUME: tal como su nombre lo indica, su tarea principal es dirigir los datos de una fuente hacia alguna otra localidad, en este caso hacia el ambiente de Hadoop. Existen tres entidades principales: sources, decorators y sinks.
- GAMIFICACIÓN: se trata de la aplicación de técnicas de la teoría de juegos a la interacción con diferentes públicos (internos y externos) de la organización y que sirve para atraer y para motivar a dichas personas.
- GRAFO: conjunto, no vacío, de objetos llamados vértices o nodos y una selección de pares de vértices llamados aristas.
- HACKATHON: se trata de una especie de maratón intensiva de programación que se lleva a cabo durante un elevado número de horas seguidas, con el objetivo de aportar soluciones -en forma de prototipo de API- a uno o varios retos propuestos.
- HADOOP: es un marco de desarrollo de código abierto que permite el procesamiento de grandes conjuntos de datos, de manera distribuida a través de un grupo o clúster de computadoras, usando un modelo de programación sencillo.
- HADOOP COMMON: Hadoop Common Components son un conjunto de librerías que soportan varios subproyectos de Hadoop.


- HANDSHAKING: es un protocolo de comunicación que sirve para definir un proceso de comunicación e identificación entre dos entidades a través de un canal.
- HARDWARE: unidades de almacenamiento secundario, principalmente discos duros, discos compactos, cintas magnéticas etc.
- HBASE: es el sistema de almacenamiento no relacional para Hadoop. Es una base de datos de código abierto, distribuido y escalable para el almacenamiento de big data. Está escrita en Java e implementa el concepto de bigtable desarrollado por Google.
- HDFS: es el sistema de almacenamiento, es un sistema de ficheros distribuido. Fue creado a partir del Google File System (GFS). HDFS se encuentra optimizado para grandes flujos y trabajar con ficheros grandes en sus lecturas y escrituras. Su diseño reduce la E/S en la red. La escalabilidad y disponibilidad son otras de sus claves, gracias a la replicación de los datos y tolerancia a los fallos.
- HIVE: es una infraestructura de data warehouse que facilita administrar grandes conjuntos de datos que se encuentran almacenados en un ambiente distribuido. Hive tiene definido un lenguaje similar a SQL llamado Hive Query Language(HQL).
- HOLOCRAZIA: es un sistema de organización que distribuye la autoridad en grupos autónomos o círculos que identifican, gestionan y resuelven los objetivos y retos de forma independiente, aunque de forma integrada con el resto de la organización.
- HTTP: de Hypertext Transfer Protocol. Es un protocolo de comunicación sin estado -no guarda información sobre las conexiones realizadas con anterioridad- para el intercambio de peticiones y documentos. Es el protocolo usado en la World Wide Web.
- HUMAN GENERATED: es un tipo de dato que las personas generan, como la información que guarda un call center al establecer una llamada telefónica, notas de voz, correos electrónicos, documentos electrónicos, estudios médicos, etc.
- IDE: significa Integrated Development Environment y consiste en una aplicación que agrupa un conjunto de herramientas -como editor de texto, gestor de proyectos, compilador o debugger- para facilitar el desarrollo de aplicaciones.
- INDICADORES CLAVE DE RENDIMIENTO (KPI): son parámetros medibles que fijan los objetivos de la organización y que sirven para medir y controlar el desempeño a nivel estratégico y operativo de una empresa al comparar su evolución con dichos parámetros. De esta manera se identifican deficiencias y se visualiza el potencial de mejora de procesos de negocio.
- INNOVACIÓN ABIERTA: es un tipo de innovación que se basa en un proceso de integración de personas y comunidades ajenas a la organización en su cadena de innovación. Por tanto, aprovecha los recursos externos para encontrar de forma más rápida y ágil nuevas y mejores soluciones a los problemas y retos corporativos.

- 
- **INTELIGENCIA DE NEGOCIOS O BUSINESS INTELLIGENCE (BI):** es el conjunto de métodos para mejorar la toma de decisiones de negocio mediante el uso de sistemas de apoyo basados en hechos. Aplicado al big data se puede definir como la técnica a seguir para la toma de decisiones basadas en datos aplicada a los negocios.
  - **INTRAEMPRENDEDOR:** término acuñado recientemente para definir a aquellas personas emprendedoras dentro de la organización, que no solo son proactivas y aportan ideas, sino que las llevan a cabo: desde la concepción hasta su materialización.
  - **IOT:** el Internet of Things (Internet de las cosas) es un concepto que se basa en la interconexión de cualquier producto con cualquier otro de su alrededor. Desde un libro hasta el frigorífico de tu propia casa. El objetivo es hacer que todos estos dispositivos se comuniquen entre sí y, por consiguiente, sean más inteligentes e independientes.
  - **JAQL:** fue donado por IBM a la comunidad de software libre. Query Language for Javascript Object Notation (JSON) es un lenguaje funcional y declarativo que permite la explotación de datos en formato JSON diseñado para procesar grandes volúmenes de información. Para explotar el paralelismo, Jaql reescribe los queries de alto nivel (cuando es necesario) en queries de "bajo nivel" para distribuirlos como procesos MapReduce.
  - **JAVA:** lenguaje de programación orientado a objetos creado por Sun Microsystems que permite crear programas que funcionan en cualquier tipo de ordenador y sistema operativo.
  - **JAVA SCRIPT:** lenguaje de programación que permite a los desarrolladores crear acciones en sus páginas web.
  - **JSON:** es un formato ligero de intercambio de datos, basado en un subconjunto del lenguaje de programación JavaScript.
  - **LENGUAJE SQL:** es el idioma para bases de datos relacionales. Bases de datos relacionales de código abierto, como MySQL y PostgreSQL han aumentado la comprensión de SQL entre los desarrolladores de software.
  - **LOG:** es un fichero que archiva un conjunto de entradas que informan de los distintos eventos -cambios en los estados de los procesos, comunicaciones, errores...- que se producen en un ordenador.
  - **LUCENE:** es un proyecto de Apache para realizar búsquedas sobre textos. Lucene provee de librerías para indexación y búsqueda de texto. Ha sido principalmente utilizado en la implementación de motores de búsqueda.
  - **MACHINE-TO-MACHINE (M2M):** se refiere a las tecnologías que permiten conectarse a otros dispositivos. M2M utiliza dispositivos como sensores o medidores que capturan algún evento en particular (velocidad, temperatura, presión, variables meteorológicas, variables químicas como la salinidad, etc.) los cuales transmiten a través de redes alámbricas, inalámbricas o híbridas a otras aplicaciones que traducen estos eventos en información significativa.
  - **MALWARE:** es un tipo de software malicioso diseñado para causar daños informáticos.



- MANEJO RESPONSABLE DE DATOS: es una forma de usar los datos que respeta la privacidad de las personas de las que se han obtenido.
- MAPREDUCE: es el corazón de hadoop, es el paradigma de programación que permite escalabilidad a través de cientos y miles de servidores en un clúster hadoop. El término MapReduce se refiere actualmente a dos tareas distintas que los programas en hadoop ejecutan: “map job” y “reduce job”.
- MASTER-SLAVE: es un modelo para una comunicación de protocolo en el que un dispositivo o proceso (conocido como el maestro) controla uno o más de otros dispositivos o procesos (conocida como esclavos). Una vez establecida la relación maestro /esclavo, la dirección de control es siempre desde el maestro al esclavo.
- MINERÍA DE DATOS: es la ciencia que se dedica a encontrar las agujas en el pajar del big data, que extrae lo valioso de entre la ingente cantidad de información existente.
- MODELO DE DATOS ENTIDAD/RELACIÓN: es la representación en escala de la realidad de la base de datos, además refleja la estructura de negocio de la organización, por medio de datos y relaciones.
- MONGODB: es un sistema de base de datos multiplataforma orientado a documentos, de esquema libre. Esto significa que cada entrada o registro puede tener un esquema de datos diferentes, con atributos o “columnas” que no tienen por qué repetirse de un registro a otro
- MOTOR DE BASE DE DATOS: software dedicado a servir de interfaz entre la base de datos, el usuario y las aplicaciones que la utilizan.
- MULTI-MASTER REPLICATION: en una configuración multi-master, los datos se actualizan en varios patrones. Cada maestro mantiene un registro de cambios, y los cambios realizados en cada master se replican en los demás servidores. Cada maestro tiene el papel de proveedor y el consumidor.
- NOSQL: es un término que describe las bases de datos no relacionales de alto desempeño. Las bases de datos NoSQL utilizan varios modelos de datos, incluidos los de documentos, gráficos, claves-valores y columnas. Las bases de datos NoSQL son famosas por la facilidad de desarrollo, el desempeño escalable, la alta disponibilidad y la resiliencia.
- OOZIE: es un proyecto de código abierto que simplifica los flujos de trabajo y la coordinación entre cada uno de los procesos. Permite que el usuario pueda definir acciones y las dependencias entre dichas acciones.
- OPEN-SOURCE: hace referencia al código distribuido y desarrollado libremente, de manera abierta a todo el mundo, dando acceso al código fuente del proyecto.
- OUTSOURCING: es el proceso empresarial de destinar recursos a contratar una empresa externa para realizar ciertas tareas determinadas.
- PHISHING: ataque informático que consiste en suplantar la identidad de una persona o una entidad para engañar al atacado.

- PHP: lenguaje de programación gratuita y multiplataforma, se ejecuta en el servidor web justo antes de enviar la página web a través de internet al cliente.
- PIG: inicialmente desarrollado por Yahoo para permitir a los usuarios de Hadoop enfocarse más en analizar todos los conjuntos de datos y dedicar menos tiempo en construir los programas MapReduce.
- REDES PROFUNDAS: sistema de desarrollo computacional que trata de imitar el funcionamiento de las redes neuronales profundas para lograr que una máquina pueda detectar, reconocer, recordar y responder como la mente humana. Su manera de aprender se basa en representaciones de datos y en modelos que aprenden de estas representaciones.
- REPOSITORIO: servidor centralizado donde se almacena y mantiene información digital -como bases de datos, aplicaciones o código- para poder ser accedidos remotamente.
- RFID: sistema de almacenamiento y recuperación de pequeñas cantidades de datos que se utiliza mediante etiquetas RFID y que generalmente se usan para la identificación.
- SCRIPT: archivo de texto plano que contiene una serie simple de comandos -o códigos sencillos en lenguajes interpretados- cuyo objetivo es generalmente el de realizar tareas de orquestación de procesos o monitorización.
- SENSOR: dispositivo capaz de detectar eventos físicos -cambios de luz, movimiento, temperatura- y convertirlo en información analógica o digital.
- SERIALIZACIÓN/DESERIALIZACIÓN DE DATOS: codificación o decodificación de un objeto en un medio con el fin de transmitirlo en forma de series de bytes.
- SHARDING: medios para dividir los datos de tal manera que los datos solicitados y actualizados estén en el mismo nodo y que el volumen de carga y almacenamiento se distribuya entre los servidores. Los fragmentos de datos también pueden ser replicados por razones de fiabilidad y de equilibrio de carga.
- SISTEMA GESTOR DE BASES DE DATOS (SGBD): conjunto de programas que permiten el almacenamiento, modificación y extracción de la información en una base de datos. También incluyen métodos de administración y monitorización.
- SISTEMA OPERATIVO: conjunto de programas de software que controlan el funcionamiento general del computador y que administran los recursos del mismo.
- SOCKET: concepto abstracto mediante el que dos aplicaciones pueden conectarse e intercambiar datos. Generalmente se usa una API para protocolos TCP/IP para realizar la comunicación mediante sockets.
- SOFTWARE: programas usados para dirigir las funciones de un sistema de computación o un hardware.
- SQL: lenguaje de consultas estructuradas para bases de datos relacionales.

- 
- **STREAMING:** distribución de datos de manera constante en forma de flujo continuo -sin interrupción-, usada por ejemplo en la transmisión de contenido multimedia a través de internet.
  - **TCP:** el Protocolo de Transmisión de Control es un sistema que utiliza protocolos comunes para que diferentes sistemas informáticos puedan intercambiar información a través de internet.
  - **TCP/IP:** conjunto de protocolos de red definidos con el fin de hacer una definición de la transferencia de datos entre computadoras.
  - **URL:** es un sistema de localización de recursos uniforme que usa una secuencia de caracteres con un formato modélico y estándar para Internet.
  - **VIRTUALIZACIÓN:** emular un entorno físico mediante software dentro de otro entorno físico real.
  - **WEB AND SOCIAL MEDIA:** es un tipo de dato que incluye contenido web e información que es obtenida de las redes sociales como Facebook, Twitter, LinkedIn, etc.
  - **WORKFLOW:** flujo de trabajo que define como tienen que ejecutarse y comunicarse entre ellas diversas aplicaciones o ejecuciones.
  - **XQUERY:** es un lenguaje de consulta para XML.
  - **ZOOKEEPER:** es otro proyecto de código abierto de Apache que provee de una infraestructura centralizada y de servicios que pueden ser utilizados por aplicaciones para asegurarse de que los procesos a través de un cluster sean serializados o sincronizados.



[www.andaluciaemprende.es](http://www.andaluciaemprende.es)