

Maria PÉREZ RÀFALES

---

*BIG DATA: LA ERA DEL MARKETING DEL DATO*

Trabajo Fin de Grado  
dirigido por  
Josep Ramon ROBINAT RIVADULLA

Universitat Abat Oliba CEU  
Facultad de Ciencias Sociales  
Grado en Marketing y Dirección Comercial

---

2016



*El marketing se está convirtiendo en una batalla basada más en la información que en el poder de las ventas*

PHILIP KOTLER



## **Resumen**

El presente Trabajo de Fin de Grado en Marketing y Dirección Comercial expone la relación entre el ámbito del marketing y una tendencia que hoy en día está en pleno auge, *Big Data* (del inglés, los datos masivos). En primer lugar, se ha desarrollado un marco teórico donde se identifican los conceptos tecnológicos clave, necesarios para el entendimiento de las prácticas de *Big Data*. En segundo lugar, se indaga en el fuerte vínculo que hay entre el marketing y los datos masivos, así como en el importante factor de competitividad que aportan estos datos para cualquier organización. Asimismo, se ha adjuntado un caso práctico de una empresa donde se demuestra que, la correcta recolección de datos es una verdadera fuente de eficiencia y de conocimiento sobre el cliente.

## **Resum**

*L'actual Treball de Fi de Grau en Màrqueting i Direcció Comercial exposa la relació entre l'àmbit del màrqueting i una tendència que avui en dia està en ple auge, Big Data (del anglès, les dades massives). En primer lloc, s'ha desenvolupat un marc teòric on s'identifiquen els conceptes tecnològics clau, necessaris per entendre les pràctiques de Big Data. En segon lloc, s'indaga en el fort vincle que hi ha entre el màrqueting i les dades massives, així com en l'important factor de competitivitat que aporten aquestes dades per a qualsevol organització. Així mateix, s'ha adjuntat un cas pràctic d'una empresa on es demostra que, la correcta recol·lecció de dades és una verdadera font d'eficiència i de coneixement sobre el client.*

## **Abstract**

*This Final Project Degree in Marketing and Sales Management describes the relationship between the field of marketing and a tendency that today is booming, Big Data. On the one hand, we have argued a theoretical framework where we can identify the key technology concepts which are necessary for understanding the practices of Big Data. On the other hand, we explored the strong link between marketing and mass data as well as the important factor of competitiveness that provides these data to any organization. We also attached a case study of a company which shows that the correct data collection is a real source of efficiency and customer knowledge.*

## **Palabras claves / *Keywords***

Algoritmo – Análisis de datos en tiempo real – Big Data – Competitividad – Dato –  
Eficiencia – Fidelización – Marketing – Satisfacción del usuario – Valor añadido

## Sumario

Introducción .....	9
I. MARCO TEÓRICO .....	11
1. Introducción a las TIC y al <i>Big Data</i> .....	11
1.1. Business Intelligence o Inteligencia de negocio .....	12
1.2. Origen de los datos masivos .....	14
2. Definición del concepto <i>Big Data</i> .....	16
2.1. Las cinco V's de los Big Data .....	17
2.2. Fuentes de los Big Data .....	19
3. Infraestructura y Arquitectura de un sistema <i>Big Data</i> .....	20
3.1. Sistema de ficheros distribuidos tipo Hadoop .....	22
3.2. Base de datos NoSQL .....	22
3.3. Motor de trabajo MapReduce .....	23
4. Otras herramientas que complementan la estrategia de los <i>Big Data</i> .....	24
5. Aplicaciones de los <i>Big Data</i> .....	25
II. <i>BIG DATA</i> DENTRO DEL MARKETING ESTRATÉGICO .....	27
1. Cómo <i>Big Data</i> afecta a las 4 P's del Marketing .....	28
1.1. Producto .....	28
1.2. Precio .....	29
1.3. Distribución .....	31
1.4. Comunicación .....	32
2. La gestión estratégica de la actividad de marketing .....	35
3. Repercusión de los <i>Big Data</i> en los modelos de creación de valor de Porter .....	39
3.1. Análisis Porter de las cinco fuerzas competitivas .....	39
3.2. Análisis Porter de la cadena de valor .....	40
3.3. Ejemplo real de un proceso de creación de valor con los Big Data .....	41
4. <i>Dirty Data</i> y la dificultad de identificar datos relevantes para la correcta estrategia de marketing orientada al dato .....	43
5. Nueva profesión entorno a los <i>Big Data</i> : El científico de datos .....	45
6. Limitaciones legales entorno a los <i>Big Data</i> .....	46
III. CASO PRÁCTICO .....	49
1. Breve introducción del caso práctico de Affinity .....	49
2. Diagnóstico de la situación inicial y Análisis DAFO .....	50
2.1. Estrategia principal a seguir .....	52
3. Objetivos generales planteados por Affinity en el proyecto .....	52
4. Visión estratégica de Affinity desde 2014 hasta 2016 .....	53

5. Tácticas de Affinity empleadas en el proyecto.....	55
5.1. Canales de marketing .....	56
5.2. Embudo de leads .....	57
5.3. Contenido .....	57
5.4. Clientes.....	58
5.5. Movilidad .....	58
5.6. Redes Sociales .....	60
5.7. Alertas inteligentes.....	62
5.8. Reporting.....	62
6. Indicadores que recibe Affinity tras la implantación del proyecto.....	63
7. Valoración final del proyecto de <i>Data Business Intelligence</i> de Affinity .....	63
Conclusiones .....	65
Bibliografía.....	67
Anexo I: Glosario.....	73

## Introducción

El presente Trabajo de Fin de Grado en Marketing y Dirección Comercial tiene por objeto demostrar que el apropiado análisis científico de los datos masivos que las organizaciones recolectan sobre los consumidores, son la mejor solución para abordar una estrategia de marketing que esté, realmente, orientada a las personas, a su satisfacción y a conseguir la fidelización de marca. Además, este proyecto pretende abordar el poder de la estrategia del dato como ventaja competitiva para una organización que, además, aporta beneficios a sus consumidores.

En nuestro trabajo hablamos, constantemente, de datos masivos pero no estaremos haciendo referencia a simples datos, sino a grandes cantidades de información que se generan de forma exponencial, diariamente, en todos los ámbitos de nuestra sociedad, a través de múltiples dispositivos como los *smartphones* (del inglés “teléfonos inteligentes”), las tabletas, los sensores, la domótica, etc.

Siguiendo la metodología del razonamiento, a partir de los principios y objetivos del marketing y viendo los efectos de los diferentes ejemplos reales en empresas actuales, hemos intentado explicar la importancia del correcto uso del *Big Data* en este sector, entendiendo que es la mejor solución, que existe hoy en día, para alcanzar sus metas.

El cuerpo del proyecto se divide, básicamente, en tres bloques que pasamos a resumir. En el primero, se presenta el marco teórico necesario para comprender el alcance tecnológico de la estrategia del dato masivo. Se define el concepto, las fuentes que generan estos datos, se hace repaso de una técnica anterior y colindante como el *Business Intelligence*, seguimos con la arquitectura e infraestructura de un sistema *Big Data*, imprescindible para entender como poder manejar grandes cantidades de información, así como sus principales algoritmos de análisis como el *MapReduce*, claves para la correcta interpretación.

Asimismo, en este primer capítulo, comentaremos dos estrategias aliadas: el *CRM* y la nube. Por último, ejemplificaremos algunos sectores que ya obtienen beneficios de la aplicación de los *Big Data*.

En la segunda parte del proyecto, nos focalizaremos en el sector del marketing. Desarrollaremos como el uso de los datos soporta, con gran éxito, los procesos estratégicos del marketing y se hacen imprescindibles para crear una oferta personalizada para el consumidor, una experiencia positiva y un retorno para la corporación. A partir de repasar las cuatro herramientas estratégicas de P. Kotler, veremos que la estrategia empresarial orientada al dato pone el énfasis en un nuevo

concepto: el proceso. También, a partir de la cadena de valor de M. Porter, ejemplificaremos la oportunidad que los datos masivos nos dan para la creación de valor de la empresa. Por último, en este segundo capítulo, abordaremos la dificultad para conseguir trabajar con datos relevantes y sobre los aspectos legales que limitan la aplicación de este nuevo paradigma de la era digital.

El tercer bloque de este estudio, se basa en un caso práctico sobre el análisis de información de los clientes, realizado por Affinity, una empresa multinacional del ámbito de la venta de productos para animales domésticos, y las ventajas y beneficios que se obtuvieron.

Finalmente, agradecer la colaboración que para este proyecto me ha dispensado mi tutor, el señor Josep Ramon Robinat Rivadulla, por la guía realizada en todo el proceso de elaboración, al profesor Josep Lluís Del Olmo Arriaga por sus comentarios que han servido de mucha ayuda, a mis padres por el apoyo prestado y a la empresa que ha tenido la bondad de facilitarme el caso práctico.

# I. MARCO TEÓRICO

## 1. Introducción a las TIC y al *Big Data*

Hoy en día, cualquier empresa, por pequeña o grande que sea, es imprescindible que haga uso de las nuevas tecnologías para tener un mayor grado de conocimiento del cliente y poder actuar en beneficio de la satisfacción del mismo.

Para poner el foco en el cliente, el auge de las nuevas tecnologías de la información y comunicación (en adelante TIC) ha sido clave. Las TIC engloban las actividades relacionadas con plataformas informáticas hardware y software. Es mediante éstas herramientas que las organizaciones pueden segmentar a los clientes y determinar sus preferencias, así como conocer el grado de rentabilidad (Barcelona Activa, 2013, p. 5).

Asimismo, el uso de las TIC también ha sido esencial en los procesos de internacionalización de las corporaciones, ya que han permitido una conectividad permanente y, efectiva, con el cliente, en cualquier lugar del mundo. Frente a un mundo globalizado, las empresas toman conciencia de ello y rentabilizan sus estrategias de expansión. Además, las TIC les ayudan a optimizar recursos; aspecto crítico para el éxito.

A diario, la tecnología genera miles de datos sobre los usuarios que se conectan desde los teléfonos inteligentes y otros dispositivos, hecho que proporciona muchísima información sobre las personas y, por tanto, incrementa la posibilidad de aumentar el grado de conocimiento de éstas. En la época actual, todo lo que hacemos está, continuamente, dejando un rastro digital que se puede utilizar y tratar; los avances en tecnología, junto a la expansión de Internet y la capacidad de almacenamiento en la nube, han provocado que crezca, exponencialmente, la cantidad de datos que podemos obtener de los consumidores.

La creación de una ingente cantidad de datos es permanente y da lugar a los llamados datos masivos, también conocidos como los *Big Data*, multitud de datos que requieren de estrategias de análisis e interpretación para encontrar las soluciones más óptimas.

En consecuencia, gracias al uso intensivo de las facilidades de las TIC en todos los ámbitos de la sociedad, las empresas pueden generar valor añadido para la oferta de negocio, tras una mejor adaptabilidad de esta oferta, una vez conocida la demanda, y por otro lado, los consumidores también se benefician, en tanto que están más satisfechos al proporcionarnos conocimiento sobre sus comportamientos, gustos y preferencias.

Por tanto, los *Big Data* son producto de las TIC y, se refieren, a miles y miles de datos que, a máxima velocidad y, a tiempo real con las herramientas adecuadas, son capaces de crear información muy válida para las compañías, y crear una fuerte ventaja competitiva.

Sin embargo, para que éste uso de los datos se lleve a cabo, es necesario que las compañías tengan una base de inteligencia de negocio, llamada inteligencia empresarial o IE (en inglés "*Business Intelligence* o *BI*") que será el fundamento previo necesario para la aplicación de los *Big Data* (Puyol, 2014, p. 471).

### **1.1. *Business Intelligence* o *Inteligencia de negocio***

Actualmente, hay una mayor difusión del término anglosajón *Business Intelligence* que de la traducción habitual "inteligencia de negocio". Por ello, en este estudio nos vamos a referir a este concepto en su abreviación inglesa, en adelante *BI*.

El principal objetivo de la *Business Intelligence* es facilitar información y conocimiento en la toma de decisiones de una organización, para que la misma pueda ser más competitiva en el sector que actúa.

Según el glosario de Gartner Group, consultoría líder en las tecnologías de la información, encontramos la siguiente definición en términos de Cano (2007, p. 23) sobre *Business Intelligence*: "*BI* es un proceso interactivo para explorar y analizar información estructurada sobre un área (normalmente almacenada en un *data warehouse*), para descubrir tendencias o patrones, a partir de los cuales derivar ideas y extraer conclusiones".

El beneficio básico que nos proporciona la *BI* es ayudar en la toma de decisiones de una organización a través de los datos proporcionados por un proceso tecnológico. Mediante un sistema de *BI*, los datos pueden llegar a convertirse en conocimiento y además, son muchas más las ventajas que podemos extraer como: reducir costes; generar más ingresos; reducir los tiempos de actividad empresarial; generar una mayor fuente de ventaja competitiva o definir mejor a qué clientes y a qué mercados nos dirigimos.

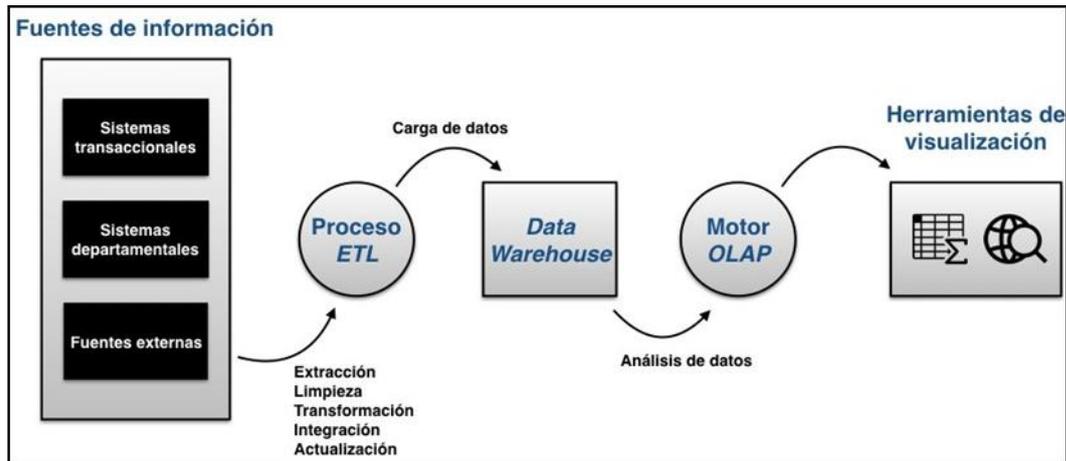
Dentro de la arquitectura de un sistema de *BI* vemos claramente que hay una serie de componentes esenciales que aplican a todos los proyectos de *data business intelligence* (Cano, 2007, pp. 93-141):

- a) Fuentes de información, de las cuales partimos para alimentar un almacén de datos. Las fuentes de información a las que se pueden acceder son: sistemas transaccionales que incluyen aplicaciones como *ERP* (acrónimo del

inglés *Enterprise Resource Planning*, es decir, un sistema de planificación de recursos empresariales), *CRM (Customer Relationship Management*, es decir, gestión de relaciones con los clientes) o *SCM (Supply Chain Management*, es decir, gestión de la cadena de suministro); sistemas de información departamentales como previsiones, presupuestos o hojas de cálculo y fuentes de información externa, como estudios de mercado contratados a terceros. Toda ésta información, podemos decir que es estructurada, ya que proviene de bases de datos relacionales, y debe cargarse en un almacén de datos (del inglés *data warehouse*). Para realizarlo correctamente, se debe asegurar que la calidad de los datos que se cargan en el depósito de datos sean los correctos.

- b) Proceso de extracción, transformación y carga de los datos en el *data warehouse* (en inglés *ETL*, acrónimo de *Extract, Transform and Load*). El mismo, trata de recoger los datos de las fuentes de información y alimentar el almacén de datos. Éste trabajo, delicado y lento, consume entre el 60% y el 80% del tiempo de un proyecto de *BI*. Esta fase se divide en cinco etapas: la extracción de fuentes de información; la limpieza de los datos en bruto para comprobar la calidad de los mismos; la transformación, para estructurar los datos de calidad; la integración para validar que los datos incorporados en el *data warehouse* sean consistentes y la actualización para añadir nuevos datos en el depósito.
- c) El propio almacén de datos o *data warehouse*, donde se permite analizar la información conceptualmente y relacionarla dentro de la organización. Se debe asegurar una correcta gestión del depósito de datos que permita: alta disponibilidad; rendimiento; copias de seguridad (en inglés *backup*) y recuperación de la información (en inglés *recovery*).
- d) El motor de procesamiento analítico en línea (en inglés *OLAP*, acrónimo de *On-Line Analytical Processing*) es la tecnología más extendida que permite tratar y visualizar la información que reside en un *data warehouse*. El motor *OLAP* debe proveer capacidad de cálculo, consultas, funciones de planteamiento, pronóstico y análisis de escenarios en grandes volúmenes de datos.
- e) Herramientas de visualización que permiten el análisis y la navegación de la información. Para visualizar la información extraída de un proyecto de *BI* podemos contar con hojas de cálculo como Excel, herramientas específicas como *Tableau* o simples navegadores (ver figura 1).

Figura 1. Componentes de un proyecto de *Business Intelligence*.



Fuente: Elaboración propia.

A modo genérico, hemos explicado cómo funciona un proyecto de *BI* que puede extrapolarse a un proyecto de *Big Data* siempre que analicemos unas cantidades de datos mucho mayores, incorporemos no solo datos estructurados, sino también no estructurados y *semi-estructurados* y el proceso se haga en tiempo real. Es decir, en una aplicación del *Big Data* se requiere grandes volúmenes de datos con mucha variedad de tipologías y, además, se precisa trabajar a mucha velocidad para aportar valor en tiempo real en la toma de decisiones empresarial. Por tanto, podríamos resumir que las limitaciones que posee un proyecto de *BI* que se solventan con la aplicación de un sistema de los datos masivos, son:

- La cantidad de datos que se maneja en sistemas de *BI* es pequeña en comparación con los que se tratan hoy en día.
- Los datos que se incorporan en la *BI* son estructurados, por tanto, son cuidadosamente añadidos y tratados previamente para que sean coherentes.
- El tiempo necesario para procesar los datos es demasiado elevado, por lo que solo se pueden implantar estrategias a largo y medio plazo.

## 1.2. Origen de los datos masivos

A diario, se produce una enorme cantidad de datos por todo tipo de dispositivos, máquinas, sensores e incluso objetos que, cualquier empresa es capaz de recolectar y extraer valor, si posee las herramientas pertinentes.

Más en detalle, cuando hablamos de datos nos referimos a *bytes* de información. Entendemos por *byte* según la RAE (2016): “unidades de información compuestas generalmente de ocho bits”. La escala de unidades en *bytes* es la siguiente:

1 *byte* (B) – *bits* (b)

1 *kilobyte* (Kb) – 1024 *bytes*

1 *megabyte* (Mb) - 10242 *bytes*

1 *gigabyte* (Gb) - 10243 *bytes*

1 *terabyte* (Tb) - 10244 *bytes*

1 *petabyte* (Pb) - 10245 *bytes*

1 *exabyte* (Eb) - 10246 *bytes*

1 *zettabyte* (Zb) - 10247 *bytes*

1 *yottabyte* (Yb) - 10248 *bytes*

Cabe mencionar que, un proyecto de datos masivos está considerado, tal y como comentamos posteriormente, a partir de *petabytes* y *exabytes* de información.

Para hacernos una idea de lo fácil que resulta generar datos actualmente, vamos a mostrar una serie de ejemplos reales.

En primer lugar, Twitter, red social que permite un servicio de *microblogging*, gestiona más de 500 millones de *tuits* (es decir, publicaciones) al día. Esto representa un total de 45 *terabytes* de datos cada día (Quincoces, 2016).

Otro ejemplo es Wal-Mart, la cadena minorista más grande del mundo, gestiona un millón de transacciones de clientes por hora, hecho que alimenta una base de datos de unos tres *petabytes*.

También, en el mundo científico podemos plasmar la facilidad de producir datos. Tenemos el caso del colisionador de partículas del CERN, que puede llegar a generar 40 *terabytes* de datos por segundo durante los experimentos.

Con los ejemplos mencionados, podemos ver que ya son muchos, por no decir todos, los sectores que hoy en día generan enormes volúmenes de datos, hecho que comporta un crecimiento exponencial de la información disponible. Sin embargo, la facilidad de crear datos no implica que sea fácil procesarlos y transformarlos en información de valor. Todo ello nos lleva a la gran tendencia que domina la industria de las TIC hoy en día, los *Big Data* (Torres i Viñals, 2012, p. 24).

## 2. Definición del concepto *Big Data*

De manera genérica, podemos definir el concepto de *Big Data* (en español, “grandes datos”) como una marea de datos masivos generados a diario por los distintos *stakeholders* de una organización. Pero, entrando más en detalle, no todos los datos comparten la misma tipología. La multitud de datos distingue entre datos estructurados, como el género y la edad, datos *semi-estructurados*, como las páginas web, y datos no estructurados, como el contenido publicado en redes sociales.

Más en profundidad, los datos estructurados son aquellos que contienen una longitud y un formato definido, y se encuentran almacenados en bases de datos relacionales. Entre los datos estructurados encontramos: censos de población, datos de transacciones bancarias, compras en tiendas online, entre otros. (Lucas, 2015).

Los datos *semi-estructurados* son los que no pertenecen a bases de datos relacionales, ya que no se limitan a campos determinados aunque poseen un orden interno o marcadores que facilitan el tratamiento de sus elementos. Por ejemplo: documentos *HTML* en páginas web (Marketingdirecto.com, 2005).

Por último, los datos no estructurados son lo que carecen de un formato determinado y no pueden ser almacenados en una tabla o base de datos relacional. Pueden ser de tipo texto como informes generados por los usuarios en foros, redes sociales, documentos de Word o correos electrónicos; y tipo no-texto como ficheros de imagen, audio o video (Lucas, 2015).

Más aún, cuando nos referimos a la palabra *Big Data*, no solo identificamos grandes volúmenes de datos con diferentes tipologías, sino que también toma importancia la velocidad de tratamiento de información en tiempo real, el valor de transformar los datos en conocimiento válido y la veracidad de los mismos.

De forma precisa, podemos hallar varias definiciones de *Big Data* descritas por muchos profesionales del sector TI, pero vamos a centrarnos en una destacada, la de Albert Solana y Genís Roca (2015, p.11): “se entiende por *Big Data* la toma de decisiones o la prestación de servicios basada en el uso de flujos de datos digitales y la capacidad de procesarlos a tiempo real”.

En otras palabras, es el proceso de recolección de grandes cantidades de datos con la finalidad de obtener información y conocimiento sobre los usuarios en el momento adecuado. De esta forma, se consigue generar valor añadido a una organización por un lado, y satisfacer mejor las necesidades de los consumidores, por otro lado. Asimismo, el uso de los *Big Data* refuerza la estrategia de marketing *one-to-one* de

cualquier organización porque permite establecer pautas y perfiles de consumidores más detallados con sus gustos y preferencias. Es decir, permite personalizar más en detalle la oferta de productos a clientes y mejorar las relaciones con el mismo.

Además, el análisis de los datos masivos con técnicas algorítmicas conlleva a modelos de reconocimiento de pautas que posibilitan predecir comportamientos, deducir patrones de consumo, realizar agrupaciones y segmentos con grupos homogéneos de consumidores e incluso analizar textos para extraer información útil en tiempo real.

Otra definición que cabe señalar de *Big Data* es la que desarrolla la Agencia Gartner (2016, párr.3): “*Big Data* se refiere al volumen, variedad y velocidad de datos estructurados y no estructurados que se vierten a través de redes en los procesadores y dispositivos de almacenamiento, así como la conversión de dichos datos para el asesoramiento empresarial”.

A modo teórico, teniendo en cuenta las anteriores declaraciones de lo que se entiende por *Big Data*, debemos puntear una serie de características que conforman los datos masivos: volumen, velocidad, variedad, veracidad y valor; también conocidas como las cinco “V’s” de los *Big Data*.

### **2.1. Las cinco V’s de los Big Data**

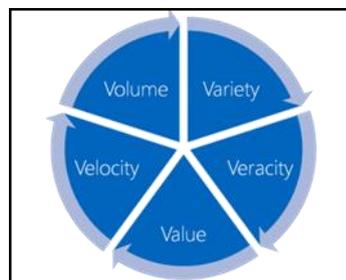
Se puede considerar que los datos masivos comparten una serie de características que agrupadas se conocen como las cinco V’s de los *Big Data*. Son las que detallamos a continuación (Lucas, 2015; Hilbert, 2015; Schroeck, Shockley, Smart, Romero-Morales y Tufano, 2012, p.6).

- 1) Volumen: captar y organizar absolutamente toda la información que nos llega es esencial para tener registros completos y no sesgados, y en consecuencia, que las conclusiones que obtengamos sirvan de forma eficaz a la hora de emprender estrategias. Según un informe de *IBM Institute for Business Value* y la Escuela de Negocios Saïd en la Universidad de Oxford, donde las personas encuestadas representaban una mezcla de disciplinas (entre las que se incluían profesionales de negocios en un 54% y profesionales de TI en un 46% de la muestra), algo más de la mitad de los encuestados consideraba por grandes volúmenes de datos entre un *terabyte* y un *petabyte*.
- 2) Velocidad: el proceso de análisis de datos masivos pide agilidad a la hora de obtener información relevante frente a un mundo cada vez más competitivo.

El uso de los *Big Data* permite reaccionar de forma inmediata para extraer oportunidades de negocio en tiempo real.

- 3) Variedad: el análisis de los datos masivos tiene la peculiaridad de poder exponer información homogénea y consistente dentro de una marea heterogénea de datos masivos. Entre los distintos tipos de datos que pueden ser manejados encontramos datos procedentes de correos electrónicos, registros de datos, registros de tarjetas de crédito, datos de estudios científicos, registros de salud, datos financieros y multimedia como fotos, presentaciones, gráficos, audio o video. Es decir, los datos que se generan poseen innumerables formatos entre los que se incluyen texto, datos web, tuits, datos de sensores, audio, video, archivos de registro, entre otros.
- 4) Veracidad: la información presentada por herramientas *Big Data* hace posible que las conclusiones sean veraces y propongan líneas de actuación dentro de una organización que optimicen los resultados. Por ello, las tecnologías creadas al servicio de los *Big Data* son imprescindibles y eficientes para afrontar retos.
- 5) Valor: todo el proceso de análisis de datos beneficia tanto a empresas como a consumidores. Por un lado, aporta valor a las organizaciones, ya que mejora las estrategias de marketing personalizado con clientes y, por otro lado, proporciona valor a los consumidores, porque sus necesidades son satisfechas con mayor precisión y a tiempo real. Por tanto, *Big Data* favorece la calidad de vida de las personas y ayuda a las compañías a ser más competitivas en un entorno cada vez más exigente (ver figura 2).

Figura 2. Características de un sistema *Big Data*.



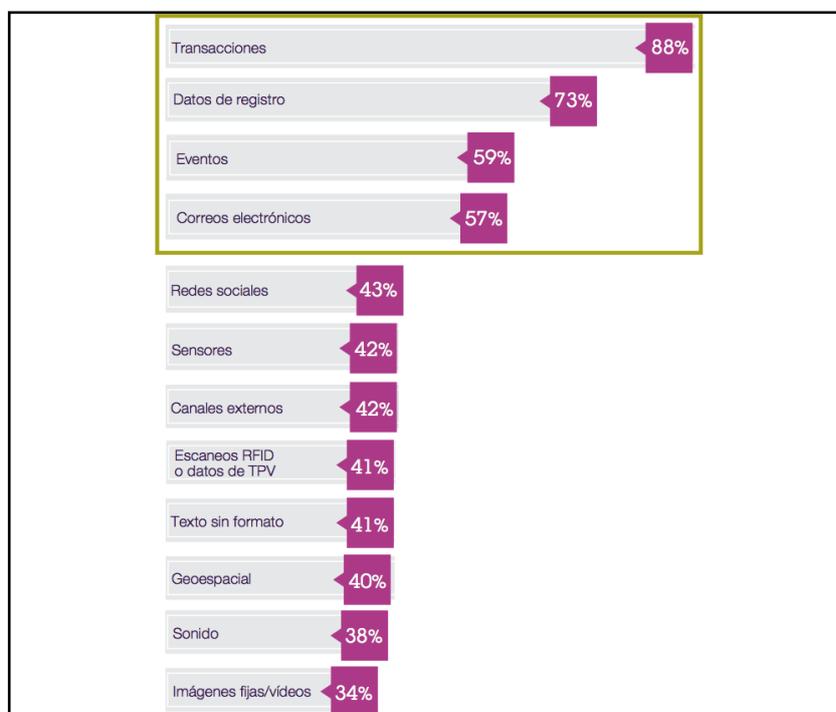
Fuente: Bannister (2015).

## 2.2. Fuentes de los Big Data

El uso de *Big Data* sugiere nuevos resultados centrados en el cliente y proporciona soluciones a retos empresariales (Schroeck, Shockley, Smart, Romero-Morales y Tufano, 2012, p.1). Son muchas las empresas que, desde hace años, manejan grandes cantidades de datos, como por ejemplo, una empresa de extracción de petróleo que analiza *terabytes* de información geológica para determinar la rentabilidad de explotar o no los yacimientos de petróleo. Aun así, hoy en día, dos tendencias importantes se encaran a la nueva era del *Big Data*: la digitalización de prácticamente todo y las tecnologías avanzadas de análisis de hoy en día que permiten extraer conocimiento de prácticamente cualquier dato (Schroeck, Shockley, Smart, Romero-Morales y Tufano, 2012, p.3).

La información que recibimos proviene de dos tipos de fuentes. Por un lado, encontramos las fuentes internas de información que provienen del núcleo de una organización. Tras los resultados obtenidos de la encuesta mencionada del capítulo anterior, observamos que la mayor parte de los participantes consideran que las fuentes de información en procesos de *Big Data* provienen de: transacciones, datos de registro, eventos y correos electrónicos (ver figura 3) (Schroeck, Shockley, Smart, Romero-Morales y Tufano, 2012, p. 11).

Figura 3. Resultado de la encuesta de IBM *Institute for Business Value* y la Escuela de Negocios Saïd sobre qué fuentes de datos recaban y analizan los encuestados.



Fuente: Schroeck, Shockley, Smart, Romero-Morales y Tufano (2012, p.11).

Por otro lado, encontramos las fuentes externas de las que las empresas extraen información con la ayuda de herramientas de recolección de datos, las cuales pueden clasificarse en dos grupos:

- En primer lugar encontramos las “*batch*” o por lotes, las cuales se conectan de manera periódica a la fuente originaria de datos en busca de nuevas informaciones. Normalmente se usa para conectarse a sistemas de ficheros o bases de datos en busca de cambios desde la última vez que se actualizaron.
- En segundo lugar, encontramos la herramienta “*streaming*” o transmisión en tiempo real, la cual permite una conexión continua a la fuente de datos. Esto implica una descarga de información cuando la misma es transmitida. Éste método de recolección de datos se suele usar para monitorizar sistemas como sensores o para conectarse a redes sociales y obtener información en tiempo real (Serrat, 2013, p. 17).

### **3. Infraestructura y Arquitectura de un sistema *Big Data***

La tecnología de los *Big Data* tiene el potencial de reforzar la inversión en almacenamiento de datos y en *Business Intelligence* con nuevas capacidades y arquitecturas posibles. *Big Data* ofrece las siguientes oportunidades (Schmarzo, 2014, pp. 211- 212):

- En almacenamiento, *Big Data* permite el acceso y análisis de un enorme volumen de datos (miles de *terabytes* y *petabytes*), entre ellos datos estructurados como por ejemplo de transacciones de terminales de puntos de venta (en adelante usaremos el acrónimo TPV); de tarjetas de crédito, de pedidos, entre otros. Además, *Big Data* capacita la integración de datos *semi-estructurados* como registros de páginas web, sensores o GPS y datos sin estructurar como vídeos, correos electrónicos o comentarios en foros de consumidores.
- Respuestas de datos en tiempo real. Es decir, captura, análisis, identificación y actualización de información producida por datos que van llegando a una organización.
- Analítica predictiva. Una arquitectura *Big Data* puede prever y detectar posibles acontecimientos y dar recomendaciones que se pueden integrar en los sistemas de la compañía.

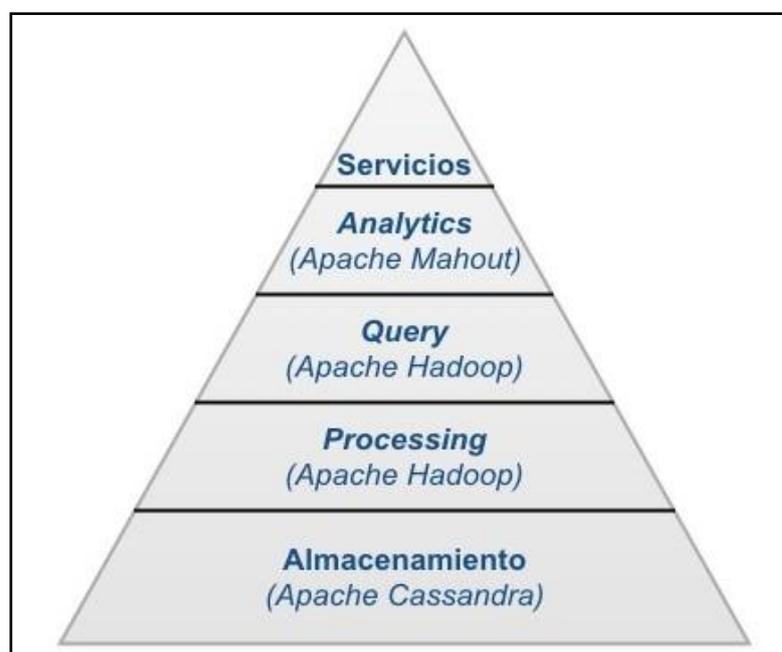
En cuanto a la interfaz que requiere el uso de *Big Data*, existe una comunidad de desarrolladores y usuarios organizados llamada “*Apache Software Foundation*” que tiene como objetivo coordinar un conjunto de productos de código abierto (en inglés, “*open source*”, es decir, que pueden descargarse de manera libre o gratuita) y promocionar el desarrollo de los mismos (The Apache Software Foundation, 2016).

*Apache* es el marco que proporciona el procesamiento de información de forma efectiva y económica. Algunas de las principales tecnologías de *Apache* en *Big Data* son: *Apache Hadoop*, *Hadoop MapReduce*, *Apache Hive*, *Apache Pig* (Schmarzo, 2014, pp. 212-214).

Dependiendo de criterios como el sector de actividad, los objetivos estratégicos de la organización o el nivel de consultas que se requiere responder, una compañía opta por una arquitectura de *Big Data* u otra.

En la siguiente figura se presentan algunas de las tecnologías de *Big Data* de código abierto de *Apache Software Foundation* (ver figura 4) (Torres i Viñals, 2012, p. 32).

Figura 4. Herramientas de *open source* de *Big Data*.



Fuente: Adaptado de Torres i Viñals (2012, p. 32).

Como hemos visto hasta ahora, para implantar un sistema de *Big Data*, es necesario disponer de la infraestructura tecnológica. A grandes rasgos, podemos

resumir que, la arquitectura tecnológica necesaria se compone de: un sistema de ficheros distribuidos tipo *Hadoop*; una base de datos *NoSQL* y un motor de trabajo de procesamiento de datos que, en la mayoría de los casos hablamos de *Hadoop MapReduce* . Más adelante detallaremos cada una de éstas partes tecnológicas (ver figura 5) (Ayora, 2015, p. 19).

Figura 5. Arquitectura básica de un sistema Big Data



Fuente: Elaboración propia.

### **3.1. Sistema de ficheros distribuidos tipo Hadoop**

La premisa básica de un sistema de archivos distribuido es que permite manejar grandes cantidades de datos en un mismo tiempo real.

Actualmente, se considera *Hadoop* como el principal sistema de archivos distribuido. *Hadoop* es un *software* de *open source* bajo licencia de *Apache*, creado por Doug Cutting (Terdoslavich, 2016).

La esencia de ser un sistema de archivos distribuido reside en que permite almacenar grandes cantidades de datos en diferentes servidores que pueden estar en diferentes emplazamientos físicos. Estos servidores distribuidos son llamadas “nodos”.

Es importante en este sistema que los datos se puedan procesar en paralelo y horizontalmente, ya que, por consiguiente, se evita la existencia de errores en diferentes nodos como pueden ser duplicados o diferentes formas de procesamiento de datos.

### **3.2. Base de datos NoSQL**

Cualquier proyecto de inteligencia empresarial se enmarca en bases de datos relacionales, también denominadas *SQL* (acrónimo del término anglosajón *Structured Query Language*, es decir, lenguaje de consulta estructurado).

Hoy en día, el mercado posee muchas opciones de bases de datos SQL como *DB2*, *Oracle*, *MySQL*, *Informix*, *Microsoft SQL*, etc. Éstas bases de datos siguen las reglas *ACID* (acrónimo del inglés *Atomicity*, *Consistency*, *Isolation*, *Durability*; es decir, atomicidad, consistencia, aislamiento y durabilidad) (Torres i Viñals, 2012, p. 25).

Las bases de datos relacionales o SQL pierden interés cuando nos referimos a la arquitectura de un sistema *Big Data*, porque las claves de estos, como se ha comentado anteriormente, reside en almacenar todo tipo de datos no solo los estructurados sino también *semi-estructurados* o sin estructurar.

La infraestructura *Big Data* implica manejar un tamaño elevado de datos, de formatos diferentes de los mismos y una velocidad de entrega de información. Es entonces cuando hablamos de bases de datos *NoSQL* (acrónimo inglés de *Not only SQL*, es decir, no solo bases de datos relacionales o SQL).

Las bases de datos llamadas *NoSQL* no suponen que no se haya de usar las SQL sino que aglutinan diferentes sistemas de almacenamiento para gestionar datos que pueden ser flexibles, es decir, cambiantes en el tiempo y de diferentes tipologías de información (Torres i Viñals, 2012, p. 26).

### **3.3. Motor de trabajo MapReduce**

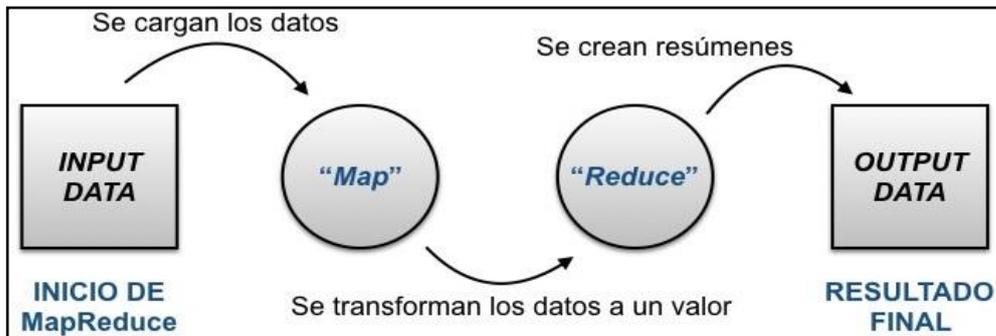
*MapReduce* es un modelo de programación introducido por *Google*, pero en su desarrollo, han participado decenas de colaboradores y, por consiguiente, aparecen múltiples implementaciones. De entre todas ellas, la más destacada y común es *Hadoop MapReduce* (Serrat, 2013, p. 21). *Hadoop MapReduce* fue creada por *Yahoo*, y se usa actualmente como *software* de *open source* bajo licencia de *Apache* (Torres i Viñals, 2012, p. 29).

“*MapReduce* es un modelo de programación para procesar grandes conjuntos de datos en grupos más pequeños mediante un algoritmo distribuido y en paralelo” (Schmarzo, 2014, p. 213). Es decir, el objetivo es procesar al mismo tiempo grandes volúmenes de datos de servidores en lugares diferentes y tratar ficheros de gran tamaño (de orden de *gigabytes* y *terabytes*) para extraer conclusiones de valor y precisas.

Los programas de *MapReduce*, como bien indica su nombre, se componen de dos fases. La primera es la fase “Map”, y la otra es la “Reduce”. El procedimiento “Map” filtra y ordena datos que son procesados uno a uno y transformados en un conjunto para asignar un valor. Después, la fase “Reduce” crea resúmenes, o sea, muestra el resultado final a través de la combinación de valores. Podemos afirmar que el

sistema *MapReduce* controla los servidores distribuidos, ejecuta tareas en paralelo, gestiona las comunicaciones y transferencias de datos y se ocupa de la redundancia y los fallos. Es decir, administra todo el proceso de procesamiento de datos (ver figura 6) (Schmarzo, 2014, p. 213).

Figura 6. Procedimiento del motor de trabajo *MapReduce*.



Fuente: Elaboración propia.

#### 4. Otras herramientas que complementan la estrategia de los *Big Data*

Actualmente, las empresas están expuestas a grandes cantidades de datos almacenados, o no, en una organización. Dada la complejidad de recogida de información y extracción de conclusiones, hay una serie de herramientas que ayudan y complementan las prácticas de *Big Data* como las estrategias de *CRM* y el *Cloud* (la nube).

*CRM* se puede definir como el desarrollo de la actividad comercial y de marketing, centrada en el mercado, constituido por los clientes de la empresa, con la finalidad de mantener su fidelidad e incrementar su valor a través de la gama de productos y servicios que puede aportarles la empresa para cubrir sus necesidades (Robinat, 2015, p. 27).

La finalidad de la estrategia de *CRM* es conseguir relaciones a largo plazo con los clientes de una empresa y aumentar el grado de satisfacción de los mismos mediante la recogida de información de estos. La herramienta permite realizar un marketing relacional haciendo foco en el cliente. Cabe señalar que, hoy en día la orientación al cliente es clave para generar ventaja competitiva.

Sin embargo, una organización por el hecho de realizar estrategias de *CRM* no está predestinada al éxito. Muchas veces, son tantas las cantidades de datos que se

requieren almacenar sobre clientes que se imposibilita el proceso. Es entonces cuando las herramientas de *Big Data* toman su mayor importancia, ya que permiten procesar, filtrar y almacenar datos relevantes para extraer información útil y aumentar el grado de conocimiento de los clientes.

En segundo lugar, vamos a referirnos a otra estrategia que ayuda a la realización de prácticas de *Big Data*: la tecnología *Cloud* (también llamada “la nube”).

Por un lado, el término *Cloud* se refiere a contratar como servicio externo un entorno de almacenamiento o analítica bajo demanda, reduciendo costes de inversión y operación y además solo se desembolsa una suscripción por el servicio necesario. Por otro lado destacar que, la infraestructura de la nube es esencial porque muchos de los datos e información que una organización debe extraer se hallan en la nube, como puede ser el contenido de publicaciones de usuarios en redes sociales (Navarro, 2015, pp. 30-41). En este sentido, el *Cloud* es el aliado perfecto para afrontar los nuevos retos que desafían las prácticas de *Big Data*.

Finalmente, cabe destacar que ambas tecnologías descritas, *CRM* y *Cloud*, hoy en día continúan evolucionando y están obligas a hacerlo de la mano de los sistemas tecnológicos de *Big Data*.

## 5. Aplicaciones de los *Big Data*

Respecto la infinidad de aplicaciones que puede tener *Big Data*, vamos a señalar una serie de sectores dónde se visualiza claramente su funcionalidad y práctica (Lucas, 2015; Think Up-LKS, 2016a; Think Up-LKS, 2016b):

- a) Dentro del sector *retail*, una empresa que comercializa sus productos vía *online*, puede optimizar el inventario en sus almacenes a través de la información recogida, como puede ser lo que más busca la gente en su página web, las tendencias en redes sociales y foros, lo que más se vende, lo que menos, etc. También, la compañía puede fijar precios dinámicos en sus productos de venta gracias a la recopilación de datos como de acciones que realizan los clientes, preferencias de venta de los proveedores, precios de los competidores, etc.
- b) El ámbito de las telecomunicaciones es la industria privilegiada en *Big Data*, debido a las redes que ofrece y a la proliferación actual de dispositivos móviles, tabletas y *smartphones*. La ventaja más clara de los *Big Data* es conseguir información de la experiencia del usuario gracias al tráfico de

datos, y de esta manera, ofrecer altas en contratos personalizados, optimizar la batalla por la competencia e incluso crear nuevas fuentes de ingresos.

- c) Mediante *Big Data*, la banca está afrontando su transformación digital mejorando la experiencia de compra de sus clientes gracias a su gestión *omnicanal* (es decir, desde cualquier dispositivo) y a ganar competitividad frente a las empresas “*fintech*”, las cuales se están convirtiendo en un voraz competidor (ver anexo 1).
- d) Otro sector que se ha beneficiado de la tecnología de *Big Data* es el deportivo. Un ejemplo es la Federación Alemana de Fútbol quien analiza grandes cantidades de datos para la mejora del rendimiento de sus jugadores así como la preparación física. Hoy en día, existen “camisetas inteligentes” que recogen pulsaciones, calorías consumidas, etc. La recogida de estos datos ayudan a clubes y federaciones a realizar ejercicios cardiovasculares de forma más saludable. Cabe añadir que en cuanto al calzado se están realizando pruebas para que sea inteligente, de manera que éste advierte al usuario si pisa correctamente o si su manera de colocar el pie puede ocasionar problemas de rodilla o de espalda.
- e) El entorno de la criminología y seguridad también se ve alterado por las prácticas de *Big Data*, ya que las fuerzas de seguridad de la mayoría de países analizan datos obtenidos en tiempo real con el fin de proteger a sus ciudadanos. Tratan de prevenir delitos en general como robos o atracos así como acciones terroristas.
- f) En el ámbito sanitario, la tecnología *Big Data* también ha impactado la actividad. A través del cruce de información de historiales clínicos, antecedentes familiares, el clima o el entorno así como los hábitos de consumo de las personas, se ha llegado a predecir modelos para cada paciente con el fin de detectar de forma precoz enfermedades y así pues, establecer pautas y estrategias efectivas para combatirlas.
- g) Por último, ya son varias las ciudades, como Barcelona, que usan el análisis de datos para convertirse en “*Smart Cities*” es decir, ciudades inteligentes. Las mismas, pueden ofrecer opciones más eficientes en diferentes intereses ciudadanos como por ejemplo en movilidad gracias a las plataformas de servicios de las ciudades en colaboración con diversas empresas. Alguna de las aplicaciones de las ciudades inteligentes es minimizar los tiempos de espera en transportes públicos.

## II. **BIG DATA DENTRO DEL MARKETING ESTRATÉGICO**

En el siglo XX, Philip Kotler y Gary Armstrong exponen unas bases y fundamentos del ámbito de la mercadotecnia. Kotler y Armstrong (2013, p. 5), definen el marketing como “el proceso mediante el cual las empresas crean valor para sus clientes y generan fuertes relaciones con ellos para, en reciprocidad, captar valor de los clientes”.

Es, a partir de entonces, que entendemos como marketing la ciencia y el arte de explorar, crear y entregar valor para satisfacer necesidades a un *target* o público objetivo. El marketing identifica necesidades y deseos insatisfechos y busca la satisfacción de los mismos. La actividad del marketing se encarga de definir y cuantificar un mercado objetivo, segmentarlo, crear una oferta adecuada para cada grupo, y satisfacer las necesidades de los consumidores, promoviendo los productos y servicios que ofrece una compañía para así, obtener a su vez, un retorno de la inversión a nivel corporativo. Para que un proceso de marketing sea exitoso debemos detectar necesidades y con ello identificar oportunidades de negocio. Seguidamente, tenemos que desarrollar nuevos productos que se adapten a los consumidores potenciales. Después, debemos atraer la atención de éstos potenciales consumidores. Luego, una vez ya han comprado el producto nuevo, fidelizarlos y retenerlos como consumidores leales de la compañía. Por último, seguir cumpliendo las expectativas que requieren los clientes (Kotler, 2001). Hoy en día, teniendo en cuenta que las acciones de cualquier compañía están condicionadas por una competencia feroz, un error en alguna de las etapas del proceso de marketing conllevará estar más cerca de desaparecer en el mercado.

Asimismo, podemos destacar que existe una relación entre el proceso de ventas de Kotler y *Big Data*. Actualmente, es de primera necesidad por parte de las empresas (debido al alto grado de competitividad), identificar oportunidades de negocio en tiempo real en base a los datos masivos que generan a diario los consumidores. Posteriormente, es necesario desarrollar ofertas adaptadas rápidamente según los nuevos productos o experiencias que se requieren. Además, se debe saber cómo se puede atraer la atención de los consumidores para que compren a una compañía y, no a la competencia de la misma. Y por consiguiente, se debe saber qué elementos valoran los consumidores de la compañía y cómo se puede seguir creando una relación de valor mutuo entre comprador y empresa.

[...] la ciudadanía exige ser atendida las 24 horas del día todos los días del año, sea donde sea [...]. Esto requiere un profundo rediseño de procesos, sistemas, perfiles profesionales e

incluso de cultura de la compañía y nos lleva de nuevo al territorio de las bases de datos segmentados, la personalización y la experiencia de usuario [...] (Roca, 2015, párr.4).

Para identificar toda esta información de los consumidores, requerida para obtener importantes ventajas competitivas frente a la competencia voraz, cualquier empresa debe poner énfasis en la correcta estrategia de recolección del dato.

La obtención de datos masivos, tanto de clientes como de potenciales por parte de las compañías, es la esencia para identificar realmente qué necesidades se pueden satisfacer de un determinado *target*, y qué elementos son los más y los menos valorados de una marca por parte de los consumidores. La obtención de datos permite a cualquier empresa conocer al detalle, qué beneficio aporta con precisión a la sociedad y, es en este punto, donde la empresa debe focalizar su estrategia corporativa.

Hay que encontrar entonces productos y servicios relevantes y de valor. Buscar solucionar los problemas de nuestros clientes y la belleza sigue siendo algo por lo que todos estamos dispuestos a pagar (Cabezas, 2016, párr.8).

## **1. Cómo *Big Data* afecta a las 4 P's del Marketing**

Kotler define cuatro herramientas básicas e interrelacionadas para alcanzar los objetivos de la actividad del marketing. Las mismas, son conocidas como estrategias de Marketing Mix o las cuatro P's del marketing, las cuales son: *product* (producto); *price* (precio); *placement* (distribución); y *promotion* (comunicación) (Kotler y Armstrong, 2013, p. 52).

Centrándonos en las cuatro P's, citadas anteriormente, podemos explorar qué oportunidades ofrece *Big Data* en el ámbito del marketing y en qué manera pueden ser utilizadas para optimizar cualquier negocio.

### **1.1. Producto**

Desde una perspectiva de marketing, se entiende por producto “todo aquello que se puede ofrecer en un mercado para su atención, adquisición o consumo, y que satisface un deseo o una necesidad” (Kotler y Armstrong, 2013, p. 51). Es decir, un producto básico no es más que un bien sensible que satisface necesidades de consumidores.

El uso de las técnicas de *Big Data* es una oportunidad para generar productos más personalizados para cada consumidor. Es decir, los datos masivos que desprenden

los usuarios ayudan a generar nuevos productos haciéndolos más adecuados para cada uno.

Un ejemplo de un nuevo producto generado con el uso de los *Big Data* es el que ha creado la compañía Netflix de vídeos bajo demanda. Ésta empresa está completamente regida por los datos que recoge a gran velocidad de sus cincuenta millones de suscriptores. Netflix recoge datos masivos para entender cómo sus usuarios se comportan y qué preferencias tienen. La empresa almacena las trazas de lo que los suscriptores ven, cuándo lo ven, dónde lo ven, qué dispositivos utilizan para verlo, qué búsquedas realizan, cuándo paran o retroceden escenas de las series que están viendo, etc. Esto permite a la compañía proveer un servicio único, personalizado, para cada individuo suscrito, adaptando específicamente la oferta a sus gustos y deseos. La compañía entrega una perspectiva interesante de todos los datos personalizados de cada suscriptor para mejorar la experiencia del usuario en la plataforma de vídeo (Morris, 2016).

El éxito de Netflix en la recolección de datos masivos conlleva a la compañía a producir una de las series más exitosas de los últimos tiempos, “*House of Cards*” asegurándose que el desarrollo del argumento a partir de la segunda temporada esté alineado con las preferencias de los consumidores. La primera temporada ya estaba grabada por completo, pero las temporadas siguientes de la serie están generadas a través de algoritmos que, en base a otros guiones de producciones que han funcionado en taquilla, deciden cuál es el mejor contenido para incorporar a la serie con la seguridad que fidelizará a sus usuarios (Marketingdirecto.com, 2014).

Además, la compañía de video bajo demanda no solo utiliza las prácticas de *Big Data* para determinar qué contenido lanzar en las series, sino que también realiza varias versiones dirigidas a diferentes públicos sobre un mismo *trailer* de una serie como por ejemplo de “*House of Cards*”. Netflix también es capaz de analizar datos obtenidos de imágenes concretas para obtener conclusiones sobre gustos en la intensidad de imagen, el volumen o las paletas de color preferidas por los suscriptores (Hedge, 2014).

## **1.2. Precio**

Otra estrategia de marketing mix necesaria para el ámbito de la mercadotecnia es el precio. Cualquier producto tangible o intangible creado por una organización o persona lleva aparejado un valor. Este valor se llama precio (Barrabino, 2014, p. 15).

Los precios impactan en la percepción e imagen de los consumidores sobre un producto o servicio. Por ello, podemos decir que el precio posiciona. Los factores que influyen en la fijación de precios son: los costes del producto, la normativa legal sobre el precio, el mercado y la competencia, los objetivos de la empresa, la percepción de los consumidores, las reacciones de los intermediarios, los precios de los proveedores, la demanda del mercado y el ciclo de vida del producto (Del Olmo, 2016, p. 29). Cabe señalar que, en la estrategia de fijación de precios, el uso del *Big Data* adquiere importancia porque, a partir de las herramientas de *business intelligence* que nos permiten conocer los precios que marcan los competidores, un posterior análisis de datos masivos permite a cualquier empresa fijar un precio mejor para sus clientes y por consiguiente, aportar un valor añadido en términos de precio.

Una compañía que ejemplifica y comunica, a diario, con un proyecto de *business intelligence* que lleva a cabo en sus supermercados es Caprabo. La empresa muestra en los tickets de compra de sus productos qué precio se gasta un consumidor y cuánto se ahorra teniendo en cuenta los precios de la competencia. Esto no es más que la recolección de datos de la competencia, concretamente de precios, y depositarlos en un *data warehouse* para que, con esta información se pueda dar respuesta a la pregunta “cuánto me ahorro comprando en Caprabo respecto a otros supermercados” introduciendo los artículos comprados.

Si este proyecto de Caprabo fuese multiplicado por miles y miles de datos generados a diario por usuarios permitiría usar las prácticas de *Big Data*. Es decir, no solo tendría en cuenta el factor precio estrictamente, sino también otros datos que influyen en la vida de un consumidor.

Asimismo, otro ámbito de aplicación, por ejemplo, se sitúa en las aseguradoras de coches. Progressive, una compañía aseguradora de Reino Unido, es capaz de fijar los precios de las pólizas de seguro de los conductores de coches a través de la incorporación de sensores, llamados “*Snapshot*”, en los automóviles para monitorizar toda la actividad del coche y, a partir de aquí, con tecnología *Big Data*, que almacena también información de sus ámbitos de actividad, como sería su vida social, les permite extraer conclusiones que acaban definiendo un precio personalizado para cada asegurado. Entre los datos que recoge el sensor y la tecnología encontramos cuántas horas al día el conductor utiliza el coche, cuál es la velocidad media a la que suele ir, qué tipo de conducción presenta (es decir, si es agresiva o no), cuántas veces se ha siniestrado, en qué momento del día utiliza el coche, qué comentarios publica en redes sociales que incrementan el riesgo de un conductor, entre otros. Tras toda esta recolección de datos masivos, la compañía de seguros Progressive es capaz de fijar un precio personalizado (Meek, 2014).

Por último, mencionar que ya son muchas las compañías de seguros que se suman a esta tendencia como Tesco Bank en Reino Unido y Generali Group en Italia (Gittleson, 2013).

Finalmente, otro ejemplo muy significativo fuera del sector seguros es la compañía UBER de reservas de taxis particulares que impone los precios según la demanda que presenta y según condicionantes como el tiempo meteorológico (si llueve, los precios aumentan) o los eventos programados (como conciertos) (O'Reilly, 2015).

### **1.3. Distribución**

En tercer lugar, la estrategia de marketing mix es la distribución. Ésta política se encarga de situar un producto en el lugar y en el momento adecuado. Las claves de la distribución son la selección de canales, el transporte, los almacenes y las rutas. El canal de distribución es el camino que sigue un producto desde la producción hasta el consumo. En la cadena de distribución intervienen distintos sujetos (el fabricante, el mayorista, el detallista y el consumidor) cambiantes según los niveles del canal (Barrabino, 2014, p. 17).

Para una empresa es determinante definir qué tipo de distribución se va a implantar en el proceso de negocio. Los tipos pueden ser: distribución exclusiva, intensiva o selectiva.

La distribución exclusiva concede a un número limitado de intermediarios el derecho exclusivo de distribuir los productos de la empresa en sus territorios. La distribución selectiva permite el uso de más de un intermediario, pero menos de la totalidad, que están dispuestos a trabajar los productos de la empresa. Para terminar, la distribución intensiva implica tener en existencia el producto en tantos expendios como sea posible.

FedEx es un claro ejemplo de empresa internacional líder en materia de distribución que ofrece entregas a multitud de países. FedEx actualmente se beneficia de las herramientas de *Big Data* en su estrategia empresarial, pudiendo ser capaz de usar los datos proporcionados por los GPS de sus vehículos, los datos meteorológicos, o la información sobre el tráfico como por ejemplo las posibilidades de encontrar una retención o un accidente en la ruta a seguir.

El éxito de FedEx es por el uso de los *Big Data*. Los consumidores de la compañía son capaces de ver el "*tracking*" (rastreo) de sus entregas en tiempo real y recibir información sobre la localización y condiciones del paquete (Capron, 2013). Además, la compañía también se beneficia, ya que disminuye los costes de manera radical y reduce el tiempo de entrega de los paquetes gracias al conocimiento de las

posibilidades para cada ruta en tiempo real. Para este cálculo, la corporación posee un algoritmo capaz de analizar todos los datos en el momento que transcurren (Carter, 2015).

Otro ejemplo real sobre la política de distribución y *Big Data* es la compañía Inditex. Concretamente cuando hablamos de su logística, Inditex es capaz de renovar su inventario en función de lo que más o menos se vende en cada tienda de la cadena consiguiendo así una mayor rapidez, una reducción de los costes y un mayor nivel competitivo (Marketingdirecto.com, 2016).

#### **1.4. Comunicación**

Por último pero no menos importante, la herramienta controlable e interrelacionable con las demás que conforma el marketing mix es la comunicación.

La misma posee una serie de técnicas para alcanzar sus objetivos de marketing, las cuales son llamadas el mix de la comunicación: la publicidad, la promoción de ventas, las relaciones públicas, los patrocinios (*esponsorizaciones* y *mecenazgos*), el marketing directo y la venta personal (Del Olmo, 2015, p.4). Todas estas acciones son susceptibles de usar los *Big Data* pero, la promoción de ventas y las actividades de marketing directo son las que más saben aprovechar el análisis de los datos masivos.

Kotler y Armstrong (2013, p. 409) definen las promociones de venta como “incentivos a corto plazo que fomentan la compra o venta de un producto o servicio”. Entre ellas encontramos descuentos, cupones, regalos, etc. Las acciones de promoción de ventas se han generalizado en los últimos años gracias a la flexibilidad y eficiencia en costes que se consigue con el uso de las nuevas tecnologías. Las promociones de venta online son llamadas “*e-promotions*”.

Por otro lado, entendemos por marketing directo el conjunto de actividades de comunicación de marketing que realiza una compañía con la finalidad de generar en el público objetivo ciertas reacciones que tienen que plasmarse en una respuesta directa e inmediata para la empresa. Algunas de las posibles acciones de marketing directo son: llamar a la compañía al centro de atención telefónica, enviar un formulario para difundir datos personales, comprar un producto anunciado en el sitio web de la compañía, etc. En resumen, podemos decir que el marketing directo considera la realización de una serie de tareas interactivas de comunicación entre empresa y consumidor con el fin de explotar la relación directa entre ambos mediante la personalización (Rodríguez, 2007).

En términos de *Big Data*, un ejemplo donde puede plasmarse la política de comunicación es el caso de la compañía Eroski, concretamente, cuando hablamos de la tarjeta de fidelización “*Eroski Club*” que se aplica en todos los puntos de venta que lleven la enseña Eroski. La empresa realiza promociones de venta a través de propuestas personalizadas para cada cliente que posea la tarjeta mencionada. Para ello, realiza técnicas de *Big Data*, ya que posee el sistema informático pertinente y la ayuda de analistas de datos. Eroski incorpora en su negocio una serie de herramientas con las que puede segmentar y ofrecer un valor más personalizado a cada cliente. La estrategia de negocio de la compañía coloca al cliente en el centro y, estructura, una oferta comercial atractiva según sus intereses. Con toda esta actividad de comunicación, Eroski permite a sus consumidores ahorrar en un año hasta 110 millones de euros mediante propuestas comerciales *ad hoc*. Además, la compañía ha llegado a personalizar los folletos y los dípticos que llegan a los hogares con el nombre y los apellidos de cada cliente para incrementar aún más su fidelización (Larrakoetxea, 2015).

En segundo lugar, respecto a las técnicas de marketing directo y qué relación tienen con los datos masivos, vemos el ejemplo de uno de los comercios minoristas más importantes de Estados Unidos, Macy's. Actualmente, Macy's es capaz de conocer el impacto que generan los envíos de su *newsletter* periódica y de las notificaciones. Además, conoce mejor la satisfacción de los usuarios sobre ello. Todos los datos que Macy's recoge permiten a la compañía segmentar al máximo sus envíos, de manera que, envía menos emails, pero con mucho más impacto. Así, como mínimo, consigue reducir las desuscripciones a la *newsletter* hasta un 20%. Gracias a la utilización de un algoritmo y el control de la demanda e inventario, los comercios minoristas de Macy's pueden lanzar ofertas cruzadas, ajustar precios y hacer rebajas casi en tiempo real para sus 73 millones de artículos de venta. Macy's, gracias a la tecnología incorporada en la compañía, ha conseguido, por un lado, optimizar sus ingresos y costes reduciendo en 500.000 dólares el gasto anual de analítica y por otro lado, mejorar la experiencia del usuario (BBVA *Innovation Center*, 2015).

Dentro del ámbito de la comunicación, podemos hacer mención a la publicidad, sector que hoy en día está explotando enormemente el uso de los datos masivos. Es realmente a través de esta aplicación que, los *Big Data* permiten impactar a los segmentos correctos de clientes con acciones a medida y en el momento preciso. Desde hace años, ya se analizan variables de comportamiento, como por ejemplo si el consumidor, previamente, ha visitado una web, se le puede impactar de nuevo con técnicas de *remarketing*.

El *remarketing* es una técnica de comunicación que permite aprovechar los datos de navegación de los visitantes de una web que se interesan por un determinado producto pero no lo llegan a comprar, con la finalidad de impactarlos a través de publicidad, sobre ese mismo artículo buscado mientras están visitando otras plataformas o páginas web. Así, provocan la conversión de compra haciendo que el usuario vuelva a la página web donde había buscado anteriormente el producto de interés. Facebook es uno de los soportes que permite atraer la atención del consumidor con información de artículos buscados en otra *site*, también llamada en español sitio web.

Una compañía que quiera realizar *remarketing* deberá disponer de una cuenta en *Google Adwords* y hacer una campaña de publicidad con la misma. Entre las principales ventajas encontramos: atraer a clientes potenciales interesados en productos de una web, focalizar los anuncios a segmentos de consumidores, lograr una mayor exposición de la marca de la compañía y aumentar el ratio de conversión (Moll, 2014). Son muchas las empresas que realizan técnicas de *remarketing*, entre ellas encontramos a Amazon que es capaz de recordar a los usuarios los productos que han visitado e introducido en su carrito virtual de compra, pero finalmente no han comprado. Por consiguiente, Amazon consigue un gran porcentaje de visitantes que acaban adquiriendo el producto y aumenta así los beneficios (Activa internet, 2014).

Sin embargo, cuando hablamos de publicidad y *Big Data* no solo debemos tener en cuenta el *remarketing*, sino que hay muchas más variables en los procesos de comunicación y que intervienen en la conversión de compra. Por ello, tenemos que ser capaces de recolectar más información sobre el comportamiento y las interacciones que los consumidores tienen con cualquier medio de comunicación, es decir, debemos aprovechar la fuerza de los datos masivos. Algunas de estas variables a recopilar pueden ser: la intención de compra de un usuario que muestra interés (ya que, ha consultado varias veces un mismo artículo y ha consultado opiniones en foros, blogs y otras redes sociales sobre ese mismo artículo, y además lo ha realizado desde diferentes dispositivos) o el grado de receptividad que el estado mental puede demostrar ante un mensaje, mediante el conocimiento del tiempo de exposición del mismo que observa un usuario (San Agustín, 2016; Cordero, 2016).

La clave para crear campañas de comunicación más precisas está en la correcta recolección y análisis del *data*. Hasta ahora, recabar datos de medios digitales se ha predefinido por segmentos amplios de consumidores predeterminados, donde se asocia información a cada uno de ellos. Sin embargo, la tecnología hoy en día

permite recolectar datos de manera flexible, proceso denominado “Taxonomía abierta”, el cual predefine categorías de consumidores y asocia comportamientos con cientos, incluso miles, de características. Esto permite en cualquier momento generar grupos más pequeños de consumidores que se adaptan, en tiempo real, a las propiedades de las campañas de publicidad lanzadas y eleva, aún más, la capacidad de segmentación empresarial.

Cabe mencionar, que todo el proceso de recolección de los datos masivos, puede generar mensajes positivos para el usuario, pero también hay una posible derivación negativa, debido a que el conocimiento que tienen las empresas sobre los consumidores, puede llegar a vulnerar el espacio de privacidad del individuo. Por ello, se debe encontrar el equilibrio entre los intereses de los consumidores y entre la compañía, para conseguir apaciguar el riesgo del uso de los *Big Data* (San Agustín, 2016).

## **2. La gestión estratégica de la actividad de marketing**

El concepto *Big Data* se vincula con cualquier disciplina capaz de mejorar las relaciones entre personas a través de la recolección de grandes cantidades de información. Hemos visto que es una práctica aplicable al área de marketing y permite, generar relaciones, más eficientes, entre empresa y consumidores. La recolección de datos es tanto cuantitativa como cualitativa, pero es realmente la parte cualitativa la que puede hacer evolucionar y mejorar, sustancialmente, la relación entre las marcas y sus usuarios.

Cuando una organización está preparada para la aplicación de *Big Data-Marketing*, es decir, tiene las infraestructuras pertinentes y quiere optar por conseguir una oportunidad de negocio mediante los datos masivos, debe guiarse por las siguientes etapas (Arthur, 2013, pp. 71-143; Roca, 2015):

- 1) *Ser inteligente, ser estratégico.* Actualmente y de cara a los próximos años, la revolución digital está y seguirá estando en auge, hecho que exige a los negocios adaptarse a los cambios de comportamiento tanto de sus organizaciones como de sus consumidores. Este cambio en la forma en que la empresa se relaciona con los clientes explotado en todo el negocio a toda la potencia del canal digital se conoce como transformación digital. La digitalización está cambiando profundamente la sociedad y el tejido empresarial a un ritmo veloz y de manera integral. Esto obliga a las empresas a repensarse los procesos. La transformación digital aporta

experiencias más relevantes para los consumidores y aumenta el grado de retorno de las marcas. Para desarrollar la estrategia de marketing en el ámbito de la transformación digital, es necesario tener una visión clara de negocio porque afectará a todos los procesos. La visión permite conectar con los puntos de los proyectos que se llevan a cabo y ayuda a ilustrar cómo los pequeños trabajos aportan un valor global a toda la compañía. Una vez se establece una visión clara, es necesario realizar un plan de acción considerando en él unos aspectos fundamentales: la estrategia de interacción con el consumidor, la estrategia de análisis, la estrategia del dato, la estrategia organizacional y la estrategia tecnológica. Respecto a la estrategia de interacción con el consumidor, es necesario que la compañía se involucre en el papel de un comprador desde el primer momento de contacto con la marca hasta el último de servicio post-compra. La estrategia de análisis es también relevante para la empresa. Hay tres categorías básicas: analítica de negocio (es decir, los indicadores básicos que miden la ejecución del negocio, también conocidos como *KPI's*, acrónimo del inglés "*Key Performance Indicators*"); analítica predictiva (o sea, modelos estadísticos que predican qué pasará debido a eventos pasados); y analítica prescriptiva (la cual son percepciones que identifican qué opción proponer mejor a los clientes). La estrategia del dato también es crítica para el buen desarrollo de la actividad de marketing, en tanto que debe prestar colaboración con los factores clave de negocio así como con las herramientas tecnológicas pertinentes. Cabe mencionar que, la estrategia del dato tiene que estar relacionada con la estrategia global de la compañía, es decir debe compartir la visión empresarial. La estrategia organizacional, no menos importante que las anteriores, parte de la premisa que la colaboración es clave. Trabajar conjuntamente con todos los departamentos de una compañía hace que la información no se fragmente y se produzca una experiencia única para los consumidores sobre el valor que transmite la organización. Por último, la estrategia tecnológica permite al responsable de marketing y al responsable de sistemas de información trabajar conjuntamente. Ambos departamentos deben evaluar el beneficio de integrar una plataforma de marketing capaz de gestionar los aspectos clave de un proyecto como el presupuesto, las campañas, la segmentación, etc.

- 2) *Eliminar los "silos"*. Unificar el marketing con la compañía es crucial y para ello, los esfuerzos de la empresa deben centrarse en la colaboración, la organización y la consolidación. La colaboración permite mejorar la ejecución

global del negocio y generar compromiso (en inglés *engagement*) con los clientes de forma más particular, mediante interacciones en tiempo real que hacen posible la comunicación. La colaboración es esencial para no generar “silos”, entendidos como partes de un proyecto que separan la actividad de marketing de la tecnológica, las ventas, y las actividades legales. La colaboración necesita de la comunicación, para adjudicar el rol a cada individuo dentro de una organización y clarifica los niveles de liderazgo que hacen posible juntar los silos de información. Por otro lado, la organización es la que hace posible la colaboración dentro de una compañía. Para acabar, la consolidación parte de alinear la visión y los objetivos de la compañía con los de marketing, de integrar la actividad de marketing hacia una estrategia guiada por los datos, y de implementar los sistemas y procesos que permiten interactuar con los usuarios.

- 3) *Desenredar el rizo de datos*. La expresión metafórica “desenredar el rizo de datos” no es más que un juego de palabras para definir el desorden que conlleva acumular multitud de datos y poder extraer una información útil de ellos. Los responsables de marketing son los primeros en recibir toda la información digital que desprende un consumidor, las veinticuatro horas del día, los siete días de la semana. Las personas que se dedican al marketing son quienes reconocen la enorme complejidad que conlleva la recolección de datos. Tras toda la acumulación de información, es importante que una organización esté conectada, es decir, que no pierda datos entre departamentos, hecho que si sucediere no se podrían ver todas las interacciones que realiza un consumidor a través de los diferentes canales que puede presentar una empresa para relacionarse con el mercado. Cabe mencionar que los sistemas de información son la herramienta para unificar todos los datos que proceden de diferentes departamentos, como por ejemplo en el caso del servicio de atención al cliente se recoge el grado de satisfacción de un consumidor, o en el departamento financiero se filtra la cantidad de compras de un consumidor. Dividir todos los datos y generar información relevante no es un proceso rápido, conlleva mucho tiempo de análisis. Sin embargo, cada paso que se realice, mejora la estrategia del dato sobre la que se fundamenta la actividad de marketing que esté orientada al incremento de valor, tanto corporativo, como a nivel de usuario. Cabe destacar que, es importante no unificar en un solo proyecto corporativo el deseo de obtener información útil tras recolectar miles de *terabytes* y *petabytes*, sino que es más óptimo dividir en pequeños grupos que

información necesitamos y qué datos debemos recabar para conseguir un conocimiento adaptado a cada situación. Para ello, debemos poner el foco en observar el rendimiento de los *KPI's*, comentados anteriormente, que nos indican cuál es la información que se debe mejorar y, en base a ella, deberemos abordar proyectos que encabecen la estrategia del dato.

- 4) *Hacer de las métricas el “mantra”*. En este punto de trabajo, lo que una organización debe plantearse es cómo medir en términos tangibles si se están realizando bien los proyectos y, por consiguiente, si se está consiguiendo una buena estrategia de marketing y del dato. Solo mediante indicadores de negocio podremos medir qué progreso hemos realizado y ver, de forma cuantitativa, qué ganamos y qué valor estamos desarrollando. Entre los principales indicadores de rendimiento de proyectos de marketing encontramos el coste por *lead*, es decir, de un cliente potencial que recibimos una respuesta, y el *ROMI* (acrónimo del inglés “Return On Marketing Investment”), es decir, medir el retorno de la inversión que realiza el departamento de marketing.
- 5) *Convertir el proceso en “el nuevo negro”*. La expresión anglosajona “to be the new black” (es decir, ser el nuevo negro) significa estar de moda, ser tendencia. En el contexto del *Big Data*, “el nuevo negro” pasa a ser el proceso. Por proceso nos referimos a actividades de marketing construidas con un propósito y llevadas a cabo gracias a los avances tecnológicos. Hoy en día, los procesos son sofisticados y transparentes. Hemos visto que Kotler, describe la estrategia de marketing orientada al producto en base a cuatro *P's*. Sin embargo, la sociedad ya no está orientada solo al producto sino que, cada vez más está dirigida a las experiencias y, por tanto, están entrando en escena nuevas características que no son meramente consecuencias de un bien tangible. Entre ellas encontramos el proceso, las personas que interactúan en el servicio, la ejecución del mismo y el beneficio que se reporta. Es decir, podemos añadir nuevas *P's* a las ya conocidas: proceso (*process*), personas (*people*), ejecución (*performance*) y beneficio (*profit*). Alinear a las personas para que compartan conocimientos y colaboren entre ellas es tan importante como realizar las correctas prácticas de marketing. Los procesos de marketing permiten a las organizaciones colaborar entre departamentos y fomentan la cesión de responsabilidades entre ellos. Una vez definidos los procesos, los equipos de marketing son capaces de orientar mejor el foco y el valor que se quiere adoptar, reducir costes y recursos no deseados, y agilizar la ejecución de los procesos en

tiempo real (los cuales están afectados por las cambiantes condiciones del mercado).

### **3. Repercusión de los *Big Data* en los modelos de creación de valor de Porter**

Michael Porter (1982), reconocido autor en estrategia competitiva, divide los procesos de creación de valor en: el análisis de las cinco fuerzas competitivas y el análisis de la cadena de valor (Porter, 2009, p. 16).

Los datos masivos pueden repercutir en los procesos de creación de valor de una compañía. Para comprobarlo, explicamos primero el marco teórico y después ejemplos de aplicación del *Big Data*.

#### **3.1. Análisis Porter de las cinco fuerzas competitivas**

El análisis de las cinco fuerzas de Porter de 1982 establece un escenario para analizar el nivel de competencia dentro de una industria. Schmarzo, realiza un estudio de las fuerzas competitivas de Porter. El mismo, define que las cinco fuerzas estratégicas que proporcionan una perspectiva global del sector son las siguientes (Schmarzo, 2014, p. 125) :

- La rivalidad entre los competidores: competencia actual del sector.
- El poder de negociación de los proveedores.
- El poder de negociación de los clientes.
- El desarrollo de productos y tecnología: amenaza de productos sustitutivos.
- Competidores potenciales: amenaza de nuevos entrantes en el mercado.

El comportamiento de estas cinco fuerzas varía según el sector que se analice o incluso el segmento específico de actividad, dado que la intensidad de las distintas fuerzas no opera de la misma manera en todos los sectores.

En primer lugar, la rivalidad entre los competidores incluye el número y tamaño de empresas que compiten en el sector, el tamaño global del sector, las principales tendencias del sector, los criterios de costes fijos frente a variables del sector, los rangos de productos y los servicios ofrecidos, y las estrategias para conseguir una diferenciación competitiva.

En segundo lugar, el poder de negociación de los proveedores incluye la reputación de la marca del proveedor, su cobertura geográfica, la calidad de productos y servicios, la profundidad de las relaciones con los principales clientes y la capacidad de pujar por un rango de productos y servicios más amplio.

En tercer lugar, el poder de negociación de los compradores incluye las preferencias de los consumidores, el número de compradores, la frecuencia de cambio y los costes de cambio relacionados, la importancia del producto o servicio en cuanto a valor y diferenciación, los descuentos por cantidad, la planificación a tiempo y la disponibilidad de productos.

En cuarto lugar, el desarrollo de productos y tecnología incluye el precio y la calidad de productos y servicios alternativos, la vulnerabilidad a la distribución del mercado, las tendencias de moda, el impacto de las acciones legales o gubernamentales y los costes de realización.

Por último, los factores de los nuevos entrantes del mercado incluyen las barreras de entrada en el sector, los factores geográficos y culturales, el posicionamiento, las posibilidades financieras y estrategias para los nuevos competidores y la dificultad para establecer una presencia sostenible en el mercado.

### **3.2. Análisis Porter de la cadena de valor**

El análisis de la cadena de valor de Porter de 1982, en términos de Schmarzo, cubre dos categorías de actividades como pilares estratégicos necesarios para el correcto desarrollo empresarial de una organización que aporte valor añadido al negocio. Los dos eslabones esenciales se dividen en primarios y secundarios.

Primeramente, las actividades primarias son aquellas necesarias para llevar un producto o servicio desde su materia prima hasta el cliente final. Dentro de las actividades primarias encontramos: la logística interna (o sea, el abastecimiento y gestión de la materia prima); las operaciones (es decir, la fabricación del producto o servicio final); la logística externa (incluye la distribución del producto o servicio final); el marketing y las ventas (lo cuales son necesarios para llevar el producto terminado al cliente final); y el servicio (es decir, la instalación, el soporte al cliente, la resolución de quejas, reparaciones, y todo el apoyo necesario para el cliente después de haberle entregado el producto o servicio final).

En segundo lugar, las actividades secundarias o también llamadas de soporte se encargan de añadir valor a las actividades primarias. Las actividades secundarias implican: el abastecimiento o compras de materiales y servicios en los que se

apoyan las operaciones, mantenimiento y reparaciones; el desarrollo de la tecnología (es decir, el diseño de producto, y tecnología de apoyo así como las tendencias de mercado); los recursos humanos (o sea, la contratación, motivación y despido de personal); y la infraestructura física de la empresa como los edificios y almacenes (ver figura 7) (Schmarzo, 2014, p. 127).

Figura 7. Cadena de valor de Porter (1982).



Fuente: Elaboración propia.

### 3.3. Ejemplo real de un proceso de creación de valor con los Big Data

Vamos a centrarnos en un ejemplo real de *merchandising* de la empresa Foot Locker, la cual es minorista líder en el sector del calzado y en la vestimenta deportiva para hombre y mujer y, posee, tanto tienda física como *online* para comercializar sus productos. El proceso de creación de valor de la compañía tiene como fin ser el principal vendedor de calzado y vestimenta deportiva a través de una clara visión de las marcas que oferta. Además, entre otras prioridades estratégicas, Foot Locker quiere disponer de un buen surtido de vestimenta deportiva, convertir las tiendas y sitios web en lugares atractivos para ir de compras, y aumentar la productividad de todas las áreas de negocio (Schmarzo, 2014, p.128).

En primer lugar, vamos a ejemplificar el uso de los *Big Data* la cadena de valor de Foot Locker. Respecto al área de la logística interna, se puede optimizar la efectividad del *merchandising* con la información que proviene de los TPV's en tiempo real. Así, podremos informar a los proveedores de casos potenciales de desabastecimiento o exceso de inventario, antes de que se convierta en una realidad y implique un factor crítico (Schmarzo, 2014, p. 131).

En el ámbito de operaciones, también se puede mejorar la efectividad del *merchandising* mediante la integración en tiempo real de datos de los TPV's y de los

stocks para predecir la potencial demanda de productos. De esta forma, se pueden gestionar los tiempos de rebajar e identificar los productos que se venden más o menos.

En cuanto a la logística externa, se puede usar la analítica de los *Big Data* utilizando datos de medios sociales y móviles para conocer opiniones sobre el *merchandising* de la tienda y, con ello, saber qué repercusiones tiene en los niveles de inventario.

En el área de marketing y ventas, mediante el análisis de los datos masivos, se permite optimizar la efectividad del *merchandising* de Foot Locker con la recolección de datos móviles y de medios sociales sobre búsquedas y visualizaciones que han llegado a ser conversiones de compra. Esto ayuda a gestionar cuáles deben ser las palabras clave de campañas de publicidad para posicionarse correctamente, así como conocer las ubicaciones de anuncios en las *sites* idóneas. Además, se impacta de forma más precisa en los consumidores.

Por último, los servicios ayudan a combinar los datos de los medios sociales con los de fidelidad del cliente para crear estimaciones con mayor frecuencia y grado de fiabilidad sobre asuntos como: la retención, el fraude, la venta cruzada y la promoción en la red (Schmarzo, 2014, p.133).

Por otro lado, teniendo en cuenta el estudio de Porter (1982), vamos a utilizar el mismo ejemplo de *merchandising* de Foot Locker para demostrar cómo el uso de los *Big Data* ayuda en las cinco fuerzas competitivas.

En primer término, en el área de rivalidad entre competidores del mismo sector, se puede optimizar el *merchandising* de la tienda *online* comparando la gestión que realizan los competidores en sus pertinentes plataformas web (siempre, teniendo en cuenta la opinión de los clientes de Foot Locker que realizan conversiones de compra en otros rivales de la marca). Esto también permite mejorar la eficacia de la política de comunicación y precios en cuanto a la ubicación *online* de productos.

Respecto el poder de negociación de los compradores, mediante el uso de los datos masivos se puede descubrir conocimientos únicos sobre el mercado, el producto y el cliente. Por ejemplo, se permite analizar opiniones de redes sociales para identificar tendencias en términos de producto. Además, se puede detallar éste aprendizaje en grupos de consumidores y así mejorar la segmentación y la información sobre los clientes. También se garantiza aprovechar los datos de ventas de los clientes para evaluar el compromiso de los mismos con la compañía mediante un sistema de *CRM*. Y todo, en tiempo real, para realizar una optimización productiva de las campañas de publicidad (Schmarzo, 2014, p.134).

En tercer lugar, el ámbito del poder de negociación de los proveedores, se puede optimizar con la gestión del *merchandising* de tienda física con los datos provinientes de los TPV's. Esto sirve para cancelar o devolver productos que se venden, poco o nada, antes que los competidores. También permite identificar los productos más destacados antes que los rivales y obtener de los proveedores condiciones mas favorables.

En cuarto lugar, encontramos el ámbito de tecnología e innovación de productos que, con los *Big Data*, se puede llegar a realizar analítica predictiva de los proveedores de Foot Locker para aprovechar sus conocimientos de *merchandising* y profundizar en el ahorro de costes de abastecimiento, inventario y distribución.

Finalmente, en el área de nuevos competidores en el mercado, el uso de los datos masivos posibilita la identificación y anticipación de las acciones de los rivales. Hecho crítico ante un entorno de competencia voraz.

#### **4. Dirty Data y la dificultad de identificar datos relevantes para la correcta estrategia de marketing orientada al dato**

*Dirty data* (en español, datos sucios) es la expresión utilizada actualmente para definir los datos que enmarcan información errónea.

Los datos correctos se han convertido en el elemento esencial para lograr la información que permita entender el comportamiento del consumidor (qué busca, de qué forma es más atractiva ofrecérselo, a qué segmento pertenece, entre otras variables fundamentales para la estrategia de marketing de una compañía). Por tanto, el primer paso es tratar y analizar con herramientas estadísticas los datos que proporcionan los consumidores antes de volcarlos en cualquier base de datos, ya que los clientes no suelen proporcionar toda la información, ni ésta suele ser totalmente real (Páramo, 2016).

Para una organización, el hecho de eliminar todos los datos falsos que generan los consumidores es prácticamente imposible; por tanto, se debe contemplar la posibilidad de no obtener información totalmente exacta y a veces obtener resultados distorsionados. Sin embargo, las prácticas de *Big Data* aseguran la veracidad de extracción de conocimiento, ya que el volumen de la muestra, implica a una población realmente representativa.

Los siguientes datos pueden ser considerados como *dirty data*: datos engañosos, datos incorrectos, datos imprecisos o datos incorrectamente deletreados.

Como habíamos mencionado en otro capítulo, *Big Data* trabaja con bases de datos *NoSQL*; por lo que resulta común que existan duplicados o datos contradictorios que se agregan de distintas bases de datos no estructuradas. En términos de Moore (2007), consultor de Gartner Group: “el 25% de los datos esenciales para el negocio de las primeras 1000 empresas en la lista *Forbes* son erróneos”.

Cabe mencionar, que muchos de los consumidores mienten multitud de veces conscientemente, al proporcionar datos cuando les son solicitados, para comprar productos en Internet o para acceder a determinados contenidos. Los motivos más habituales de mentir son: proteger la privacidad, evitar la publicidad online, y miedo de cómo serán utilizados los datos personales. Los datos que suelen falsearse más son: la dirección; la edad y el correo electrónico (Chahal, 2015; Puromarketing.com, 2015).

Como hemos destacado, la estrategia orientada al dato de las compañías, sugiere una difícil extracción de valor debido a la cantidad ingente de información que generan a diario los consumidores. Es decir, la expresión “desenredar el rizo” toma en este punto un papel relevante. Es importante destacar que, para depurar los datos sucios una compañía puede optar por dos vías.

La primera vía se ejecuta cuando disponemos de pocos datos a analizar y los podemos trabajar manualmente. En este caso, las etapas a tener en cuenta son: definir una visión de negocio clara; esquematizar las cuestiones que se requieren responder; asignar el equipo y el liderazgo adecuado; identificar los requisitos del dato (es decir, entender qué tipo de dato se necesita para satisfacer las necesidades de los consumidores); encontrar la fuente de información pertinente; identificar la verdad del dato (es decir, trabajar conjuntamente departamento de TI y de marketing para determinar qué datos son veraces y cuáles no); consolidar e integrar los datos; y probar, expandir y evolucionar (es decir, verificar que se respondan las cuestiones de negocio que previamente se plantean responder y después, si el objetivo es satisfactorio, seguir progresando frente a nuevos retos) (Arthur, 2013, pp. 105-107).

Por otro lado, en situaciones donde se hallan grandes volúmenes de *bytes* de información, donde depurar bases de datos manualmente resulta una tarea imposible, debemos trabajar con algoritmos. Los mismos, son específicamente creados para solucionar los problemas más comunes de recolección de datos. Puede que no solucionen a la perfección todos los fallos e inexactitudes de los datos sucios, pero limitan mucho más las probabilidades de error en todos ellos (Buckley, 2015). Hay muchas grandes corporaciones que ofrecen servicios y *softwares* para limpiar datos, entre ellas IBM y SAP (Kopytoff, 2014).

Por último, mencionar que no todos los datos sucios conllevan desventajas. En este caso tenemos el ejemplo de Google, compañía que reconoce los errores de ortografía de sus usuarios y recomienda palabras de búsqueda similares. Esto, solo es posible gracias a que Google ha recogido previamente miles y miles de consultas con errores ortográficos y en lugar de desecharlos, como datos sucios, ha sabido reconocer una oportunidad en su plataforma web de búsquedas alternativas para los usuarios y ha propuesto de nuevas (Kopytoff,2014).

## **5. Nueva profesión entorno a los *Big Data*: El científico de datos**

*Big Data* además de ofrecer ventajas a corporaciones y generar nuevas oportunidades de negocio, también beneficia al ámbito laboral con nuevos lugares de trabajo.

Actualmente, la inteligencia de negocio y los datos masivos son el mayor destino de la inversión y la principal fuente de empleo cualificado en el ámbito de los sistemas de información. Tal y como apunta Gartner en términos de Tilves (2016, párr. 1): “A lo largo de 2016 se llegará a los 16.900 millones de dólares por analistas y *Business Intelligence*, de acuerdo con cálculos de Gartner”. Por ello, cabe destacar que los expertos dedicados a analizar e interpretar grandes bases de datos, se han duplicado en los últimos cuatro años, según menciona un blog de José Ramon Rodríguez, director del máster de *Business Intelligence* de la *Universitat Oberta de Catalunya* (ProfesionalesHoy, 2016).

La profesión que hoy en día está en pleno auge es la del científico de datos. Éstos son profesionales que comparten una parte analítica y otra creativa. Su rol es determinar la expresión cualitativa de cuantiosas ideas. En otras palabras, un científico de datos es un apasionado de los datos que cuidadosamente traduce la masa de datos en respuestas clave para el negocio y para la estrategia de marketing. Éstos profesionales están preparados en materias como la estadística y la analítica predictiva avanzada y tienen experiencia trabajando con grandes conjuntos de datos. Por tanto, se puede decir, que para los apasionados de los números se abre un interesante campo laboral. Los científicos de datos son capaces de ayudar a los mercados a combinar el arte de la creatividad con la ciencia de los números para dirigir exitosas oportunidades de negocio. Ellos entienden que las operaciones de marketing son estratégicas y su trabajo es determinar la expresión cualitativa de los datos masivos. Es decir, detectan modelos útiles de relaciones

entre individuos y productos o servicios con el objetivo de aportar beneficio a las organizaciones (Arthur, 2013, p.58).

Los profesionales de los datos entienden que, en el momento actual, el conocimiento dirige los ingresos de una compañía y ellos están formados para ser líderes en detectar nuevas oportunidades de negocio latentes en el mercado. Estos analistas son capaces de trabajar en equipo con toda la organización a fin de obtener un beneficio común (Arthur, 2013, p.98).

La entrada en escena de esta profesión se debe a que las organizaciones se han dado cuenta que mediante el análisis de datos, se cubre la posibilidad de incrementar el valor de los negocios y proyectos, pero también a que la mayoría de la información que se genera en las compañías no puede ser analizada de forma efectiva con las herramientas de uso tradicional (Muñoz, 2016).

*Byte TI*, una de las revistas de actualidad más importantes del sector tecnológico, indica que la profesión de moda de los próximos años y la escasez de profesionales con este perfil hace que sea uno de los trabajos mejor pagados en el sector TI. (Navarro, 2016). Entre las salidas profesionales de un científico de datos encontramos:

Consultores y gestores de proyectos de *business intelligence* (*BI* o en español, inteligencia de negocio), arquitectos de datos y sistemas de *BI*, analistas digitales que trabajan con departamentos de marketing, directores de *BI*, y una nueva posición directiva, el *chief data officer* (*CDO*) responsable de los datos empresariales de la empresa y de la estrategia de la información (ProfesionalesHoy, 2016, párr.4).

## **6. Limitaciones legales entorno a los *Big Data***

Las prácticas de los *Big Data* plantean muchas veces controversia sobre limitaciones legales. Siempre hay que respetar la privacidad de los usuarios y para precisar y concretar hasta donde llegan los derechos y obligaciones tanto de empresarios como de consumidores, vamos a focalizar este apartado en conocer la Ley 15/19999 de protección de datos conocida como LOPD, y en el ámbito de internet, la Ley 34/2002 de servicios de la sociedad de la información y de comercio electrónico. Ambas leyes condicionan el uso de los datos.

La protección de datos personales es un derecho que otorga a las personas la facultad de control sobre los propios datos. Es un derecho fundamental, pero no es absoluto y sólo se refiere a personas físicas, por tanto, no protege a personas jurídicas. La ley de protección de datos concede derechos a las personas ante la

exposición de sus datos y establece obligaciones para los responsables que traten los datos. En España la autoridad de control de la normativa es la Agencia Española de Protección de Datos.

La Ley 34/2002, es una ley muy vinculada a la protección de datos porque regula todo lo relativo a páginas web, de los proveedores de servicios, contenidos, contratos online y todo lo concerniente a dispositivos o información que un usuario pueda reclamar a través de la red.

Es interesante clarificar una serie de conceptos clave sobre la LOPD. En primer lugar, destacamos qué entiende la ley por dato de carácter personal. Éste es cualquier tipo de información referente a una persona física identificada o identificable y en cualquier formato, ya sea por grabaciones de voz, imágenes, etc.

En segundo lugar, define qué es un fichero. Es un conjunto organizado de datos, los cuales pueden estar en un archivo documental, bases de datos, datos gravados en un sistema de video-vigilancia, etc.

Por último, es conveniente destacar qué se entiende por tratamiento de datos. Esta palabra es la que se ha determinado para denominar cualquier actividad que hagamos con los datos como puede ser acceder a ellos, eliminarlos, bloquearlos, enviarlos, manipularlos, etc.

Una vez definidos los conceptos clave, vamos a explicar los principios que regulan la ley. Primero, el principio de calidad: establece que los datos que recogemos han de ser adecuados, pertinentes y no excesivos, de acuerdo con la finalidad para los que se recaban. Además, es obligatorio conservar los datos actualizados y exactos. Sin embargo, la ley establece una conservación limitada, o sea, se deben definir periodos de duración concretos para guardar los datos.

Por otro lado, encontramos el principio de deber de información: es necesario facilitar información previa antes de recoger cualquier dato. Se tiene que informar a la persona que va a dar sus datos, quién es el responsable de sus datos, para qué se van a utilizar y si quiere ejercer sus derechos de revocación a quién se debe dirigir.

En este principio, cabe destacar el caso de *Google Search*. Las autoridades pidieron a la organización que la página principal de Google incluyera un aviso de privacidad. A nivel europeo, se ha abierto un expediente contra Google porque la compañía realizó una unificación de sus políticas de privacidad y adujo que era para facilitar el entendimiento y así simplificarlas. Sin embargo, la realidad es que el usuario pierde el control de sus datos cuando utiliza este servicio, los periodos de retención no

quedan claros ni queda de forma clara qué se hace con esos datos y a quién se comunican. Es, en este ámbito de aplicación, donde entra en juego las prácticas de *Big Data*, dado que se pueden hallar multitud de datos en la red sobre usuarios, los cuales no son conscientes que los están generando ni a quien se dirigen.

Otro principio que regula la ley de protección de datos es el de legitimidad y consentimiento. Esto quiere decir que para tratar datos o para cederlos a terceros, ha de ser una finalidad legítima y se necesita en muchos casos el consentimiento, salvo que exista una relación laboral o administrativa con la persona en cuestión. Tampoco cuando haya un interés vital. En relación a las cesiones se debe saber que publicar datos en internet se considera cesión. Entonces, si publicamos datos de terceros en internet necesitamos su consentimiento. Esto aplica siempre, excepto si estamos hablando del ámbito doméstico y particular, como familia o amigos.

Por último, el deber de secreto y seguridad también es partícipe de la ley en cuestión. La ley obliga a guardar secreto de los datos que se facilitan y además obliga a adoptar una serie de medidas de seguridad que son más o menos estrictas en función del nivel de sensibilidad de los datos. Las medidas de seguridad se establecen en el Real Decreto de desarrollo 1720/2007 y hay unas medidas u otras en función del tipo de dato.

Actualmente, hay mucho debate sobre las sanciones generadas a empresas que gestionan grandes multitudes de datos. Las sanciones más graves están entre los 100 millones de euros y el 5% del volumen de negocio anual. Cabe mencionar que hoy en día muchas empresas están optimizando sus relaciones en ámbitos legales para proveer mejor visibilidad de sus proyectos de marketing (Arthur, 2013, p. 93).

### III. CASO PRÁCTICO

Cabe mencionar, que resulta imposible abarcar un caso entero de datos masivos ya que, como hemos visto en todo el trabajo, recoge un volumen de datos demasiado elevado, donde analizarlos estaría realmente fuera de nuestro alcance. Sin embargo, podemos centrarnos en un proyecto de igual dinámica pero de menor magnitud; es decir, un caso de inteligencia empresarial que como explicamos en el marco teórico se circunscribe a un entorno más limitado y estructurado. El caso práctico del trabajo en curso refleja la relación e implicación con el área de marketing.

#### 1. Breve introducción del caso práctico de Affinity

El caso en cuestión versa sobre la compañía Affinity, empresa internacional experta en nutrición animal de perros y gatos. La empresa comercializa, con tienda física y *online*, sus productos. Está presente en más de 20 países y los pilares de su estrategia de crecimiento empresarial son: innovación, investigación, calidad, internacionalización, y construcción de marcas fuertes. Cabe destacar que, la empresa pone al consumidor en el centro de su visión empresarial, lo escucha, lo conoce y muestra un aprendizaje continuo. Affinity posee más de 50 años de experiencia en el sector *retail* de mascotas. Además, es líder en alimentación para perros y gatos en España, y ocupa el sexto lugar a nivel mundial. Asimismo, la organización se divide en tres marcas dentro de su portafolio de productos: *Advance*, *Ultima* y *Brekkies* (Affinity Petcare, 2016).

Podemos entender el objeto de negocio de la compañía en tanto que numerosos estudios demuestran los fuertes vínculos emocionales entre animales y dueños. Son conexiones hormonales que se acentúan cuando hablamos de mascotas domésticas. Cualquiera que tenga una en casa, siempre busca lo mejor para ella, y también, en términos de alimentación. Por ello, Affinity identifica esta oportunidad de mercado y proporciona una óptima estrategia de negocio afín a la misma. Para explotarla, Affinity utiliza el siguiente proyecto de *data business intelligence*.

El beneficio principal que este proyecto persigue es incrementar la satisfacción de los clientes, situándolos en el foco de la estrategia empresarial, a través de la plataforma *online* de venta que dispone la compañía. La estrategia de la *site* se centra en analizar datos con las herramientas de inteligencia de negocio. Por consiguiente, aporta un valor añadido a los clientes y a la empresa.

La red *online* que se dispone a mejorar, pretende facilitar la elección de nutrición animal a sus clientes de forma más rápida y concreta. Es decir, busca eliminar los problemas de ineficacia debidos al exceso de oferta que aparece en la página web. Por ejemplo, es esencial que un cliente que posee un “*caniche toy* en edad cachorro” como mascota, no se le impacte con información para otra raza, o edad, como podría ser la de un “*ladrador* adulto”. Dicho de otra manera, la nueva plataforma se adecua a las necesidades de cada cliente y de su mascota, con una oferta más personalizada en base a cada necesidad y gusto, y todo ello, lo hace mediante el análisis de datos que desprenden los mismos usuarios de manera inconsciente.

Finalmente, destacar que la acción en cuestión, también constituye un rediseño de la página web de Affinity (la cual persigue adaptarse a los nuevos cambios del entorno empresarial) y una mejora de las campañas de publicidad y marketing de la marca.

## **2. Diagnóstico de la situación inicial y Análisis DAFO**

En España, el mercado de mascotas acapara un volumen de negocio de unos 900 millones de euros, de los cuales un 51% corresponde a la nutrición para perros, un 24% a la alimentación para gatos y el resto engloba un 25% de cuidado de mascotas (Marketingnews.es, 2010).

Dentro del mercado español, Affinity es líder como vendedor de alimentación para perros y gatos, en 2015 su facturación fue de 458 millones de euros (Europa Press, 2016). Le sigue Mars España con una facturación de 300 millones de euros y después Nestlé (Ribé, 2015).

Cabe señalar que en la mitad de los hogares españoles se hallan mascotas, un total que representa 20 millones de animales de compañía, entre los que dominan perros en un 5,4 millones y gatos en un 3,8 millones (Blázquez, 2015).

Con toda esta información, partimos a desarrollar un análisis DAFO que se presenta como diagnóstico frente al proyecto de análisis de datos planteado (ver cuadro 1).

Cuadro 1. Análisis DAFO.

Fortalezas	Debilidades
<p>Affinity posee un fuerte compromiso y pasión por las mascotas y sus dueños. Es decir, pone al cliente en el centro de su estrategia empresarial.</p> <p>Affinity tiene más de 50 años de experiencia en nutrición animal.</p> <p>Affinity consta de un departamento de I+D+i consolidado con una amplia red de profesionales.</p> <p>Affinity tiene una gran capacidad logística debido a las múltiples marcas que comercializa.</p> <p>La compañía posee tres marcas fuertes que le permiten abarcar mucha cuota de mercado, sobretodo en España donde es líder.</p> <p>Affinity contrata a profesionales externos para realizar estudios de mercado y así se actualiza según los cambios del entorno.</p>	<p>La principal debilidad a cubrir es el exceso de oferta en la red <i>online</i> de Affinity que implica una dificultad de búsqueda de referencias para los clientes.</p> <p>Canal <i>online</i> como medio de comercialización desconocido, ya que la compra de nutrición animal suele ser más en tienda física.</p> <p>Exceso de variedad de marcas que hace dudar al comprador.</p> <p>Falta de personal especializado en el canal digital para clarificar las dudas a los clientes.</p>
Oportunidades	Amenazas
<p>Empresa líder del mercado alimentario de perros y gatos en España por lo que puede explotar su capacidad productiva a nivel internacional donde ocupa la sexta posición.</p> <p>Tendencia al crecimiento de compañías que poseen capacidad para invertir en I+D+i y adoptar nuevas tecnologías.</p> <p>Mercado de mascotas en crecimiento a nivel internacional.</p> <p>Tendencia de mercado de compromiso con los animales, el medio ambiente y la sostenibilidad.</p> <p>Tendencia al crecimiento del canal <i>online</i> en el <i>retail</i> para la comercialización de productos.</p> <p>Tendencia al auge de las redes sociales, blogs y páginas web para potenciar la imagen de marca.</p> <p>Tendencia al <i>omnichannel</i> (multicanalidad, para transmitir una sola imagen de marca)</p> <p>Tendencia al envejecimiento de la población en España, cambio demográfico que hace que los animales se perciben como un componente más de la familia y reciben más cuidados.</p> <p>Affinity a través de sus tres marcas puede segmentar mejor su oferta afín a cada cliente.</p>	<p>Mercado maduro en España donde Affinity posee una alta cuota de mercado y pocos competidores.</p> <p>Situación de crisis económica y reducción del poder adquisitivo de los consumidores en España.</p> <p>Auge de las marcas de distribución en el ámbito de la alimentación.</p> <p>Globalización que implica una fuerte competitividad en el sector de nutrición animal.</p> <p>Consumidores más informados, son proactivos y por tanto, más exigentes cuando compran.</p> <p>Necesidad de adaptación al entorno cambiante: hay nuevas demandas de los consumidores, nuevas leyes y regulaciones medioambientales, sanitarias y de seguridad alimentaria.</p>

Fuente: Elaboración propia.

### **2.1. Estrategia principal a seguir**

A partir del análisis DAFO y, una vez determinada la situación inicial donde se enmarca la compañía, detectamos una serie de oportunidades que se pueden alcanzar gracias a las fortalezas corporativas.

Como consecuencia de que la empresa pone el foco de la estrategia empresarial en los clientes, Affinity entiende la importancia que aplica el canal *online* para ellos y fija una estrategia orientada a éste medio de comunicación. El sitio web es, por tanto, el canal escogido para desarrollar el proyecto. Cabe mencionar, que la implicación en I+D+i (es decir, “Investigación, Desarrollo e innovación”) de la empresa y el auge de las nuevas tecnologías en el mercado permiten llegar al camino propuesto. Para terminar, destacar que el nuevo objetivo de la estrategia corporativa requiere de un proceso de seguimiento y de indicadores de evaluación ante los resultados obtenidos.

## **3. Objetivos generales planteados por Affinity en el proyecto**

En el presente proyecto, Affinity plantea unos objetivos a conseguir. El primero a desarrollar, es implementar y configurar la plataforma de *Universal Analytics* de *Google Tag Manager* en la web corporativa para asegurar que se obtienen datos fiables. Tal y como menciona Pérez (2016, párr.1): “*Google Tag Manager* es una herramienta que simplifica en gran medida la gestión de etiquetas o fragmentos de código para poder hacer el seguimiento y medición de cualquier hito que acontece en un sitio web”.

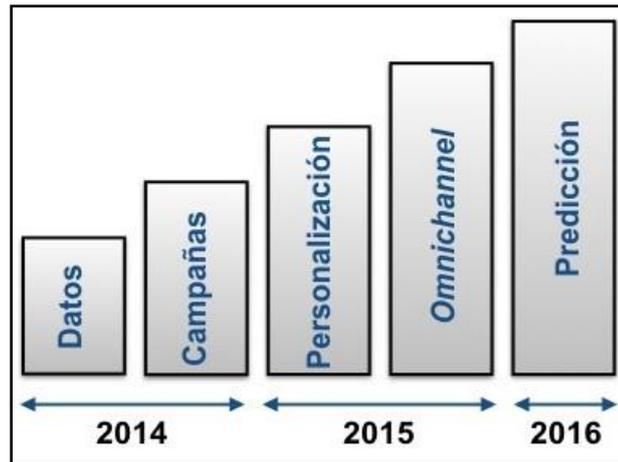
Por otro lado, se busca medir el éxito de todas las acciones de publicidad y marketing que realiza la compañía mediante un cuadro de mando personalizado.

Por último, se plantea utilizar los datos cuantitativos y cualitativos de comportamiento de los usuarios dentro de la web corporativa, para mejorar los resultados de las campañas de publicidad y marketing, así como optimizar los resultados de la propia página web de Affinity.

## 4. Visión estratégica de Affinity desde 2014 hasta 2016

El planteamiento estratégico del proyecto analiza los objetivos empresariales desde 2014 hasta 2016. Se estructura en una serie de fases donde la marca obtiene ciertos beneficios que explicamos más adelante (ver figura 8).

Figura 8. Visión estratégica de Affinity de 2014 a 2016



Fuente: Elaboración propia.

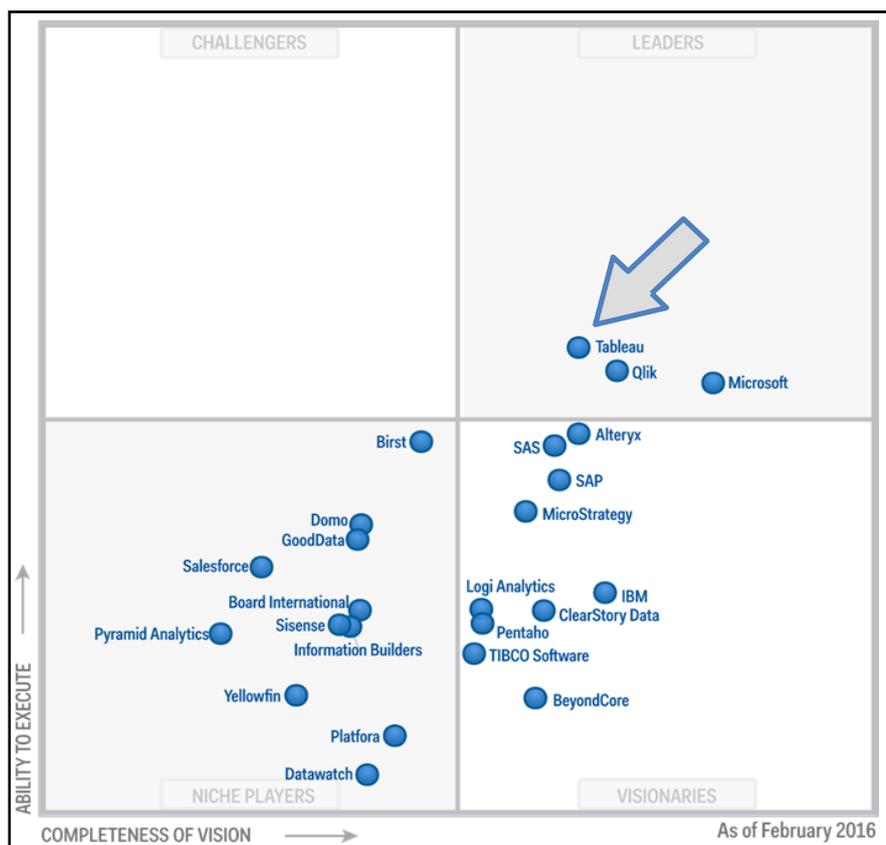
La primera etapa, en 2014, persigue el objetivo de mejorar el negocio y el marketing de la compañía con la implementación y uso adecuado de herramientas de análisis de datos. Para ello, pone el foco, durante este año, en los datos y en las campañas publicitarias y de marketing. Los beneficios que se obtuvieron son: disponer de datos correctos, de manera ágil y fácil, para tomar las mejores decisiones de negocio, producto, clientes y marketing; y entender mejor el comportamiento de los usuarios de la empresa; y mejorar los resultados mediante cambios en la web y en las campañas.

Para conseguir los beneficios mencionados, la compañía, desde el punto de vista de los datos, configura *Google Analytics* para conocer las palabras esenciales y claves de búsqueda que utilizan sus usuarios y para analizar las campañas de publicidad. Además, establece un cuadro de mando con *Tableau*: explicamos a continuación las dos herramientas.

*Google Analytics* es una herramienta de analítica web de la compañía Google, hábil para cualquier empresa, que permite monitorizar acciones de marketing y por consiguiente, conocer la audiencia, el comportamiento y las conversiones de compra de los usuarios en un sitio web.

Por otro lado, *Tableau* es un programa informático americano que proporciona la visualización de datos en materia de inteligencia empresarial. *Tableau* es líder en herramientas de inteligencia empresarial y plataformas analíticas. Otros competidores de *Tableau* que también son líderes en el mercado actualmente dentro del mismo ámbito, son: *Qlik* y *Microsoft* (ver figura 9) (Gartner Group, 2016).

Figura 9. Cuadrante de Gartner Group sobre inteligencia empresarial y plataformas analíticas.



Fuente: Gartner Group (2016)

La principal ventaja que ofrece *Tableau* frente a otro *software* de inteligencia empresarial es que éste posee una clara capacidad de visualización y exploración interactiva de diferentes fuentes de datos. La plataforma permite a los usuarios, ver y entender los datos procedentes de diferentes fuentes de información y, en base a estos, dar respuesta a las preguntas que se plantean éstos. Asimismo, la interfaz no solo está habilitada para expertos en materia tecnológica, sino que está abierta a todo tipo de público, ya que posee una dinámica visual e intuitiva. También *Tableau* ofrece una amplia gama de opciones de aprendizaje (incluye en su programa

tutoriales, seminarios, vídeos, entre otros) para educar y ayudar. (Parenteau, et al., 2016).

Volviendo a la visión estratégica de Affinity, en el año 2015 y 2016, es decir, en la segunda y tercera fase, se pretende mejorar el negocio y el marketing de la compañía mediante la personalización de la comunicación con sus clientes. Para conseguirlo, en 2015 la compañía centra la atención en la personalización de los mensajes y la estrategia *omnichannel* (en español, omnicanal). A través de la misma se comunica con cada cliente mediante mensajes hechos a medida para cada canal. Así, también contribuye a mejorar la imagen, ya que proporciona un mejor servicio. Asimismo, Affinity es capaz de fidelizar a sus clientes y aumentar las tasas de conversión de compra en su *site*.

Destacar que, entendemos por *omnichannel* la experiencia exclusiva que recibe un consumidor sobre todos los canales comerciales que intervienen en la compra de un usuario, ya sea en la propia tienda o a través del ordenador, tableta, correo o *smartphone*. Dicho camino es capaz de unir la imagen del entorno *online* y *offline* de la marca, y permite crear una visión única de marca a los consumidores. (Trazada Internet Marketing, 2016)

Los objetivos planteados en 2015 son llevados a cabo con una serie de acciones: se personalizan las campañas de *email* marketing, se personaliza la web corporativa, se personaliza el “*search*” (es decir, las búsquedas en la *home* de Affinity) y se personaliza el “*display*” (o sea, se proponen datos de utilidad según cada usuario).

Por otro lado, en la tercera fase, que se ejecuta en el 2016, la empresa propone llevar a cabo un análisis predictivo. El mismo, permite desarrollar algoritmos de predicción para la compañía con el fin de diseñar escenarios de inversión en medios de comunicación y también en precios, además de predecir el impacto en *leads*.

## **5. Tácticas de Affinity empleadas en el proyecto**

La principal táctica que se emprende en el presente caso es la adopción del *software Universal Analytics* de Google. Esta herramienta es un completo sistema de métricas de marketing, clientes y negocio, totalmente adaptable a las necesidades de una compañía. En particular, el caso de *data business intelligence* de Affinity deja la herramienta en uso para trabajar los siguientes aspectos y poder conseguir una serie de beneficios para cada uno de ellos.

- *Canales de marketing*: análisis y mejora del retorno de la inversión de los canales de marketing.

- *Embudo de leads*: mejora de la tasa de conversión a *lead*.
- *Contenido*: análisis y mejora de la estrategia de marketing de contenidos.
- *Clientes*: análisis de las características y preferencias.
- *Movilidad*: análisis del comportamiento del cliente a través de móviles, tabletas y escritorios (en inglés *desktops*).
- *Redes Sociales*: análisis y mejora del retorno de la inversión en redes sociales.
- *Alertas inteligentes*: la herramienta de Google avisa de cambios que afectan al negocio..
- *Reporting*: cuadro de mando e informes personalizados para gestionar el día a día.

### **5.1. Canales de marketing**

Son varios los canales de marketing donde una empresa puede fomentar una relación con su público. Entre ellos encontramos: búsquedas orgánicas (es decir, sin coste para la empresa), búsquedas de pago, campañas de *display* de Google, y otras herramientas de marketing directo.

La estrategia de canales de marketing de Affinity parte, como habíamos mencionado, de la configuración de *Google Analytics*, pero también de *Artemis*.

*Artemis* es un producto de evaluación de campañas *online* guiado por los datos. La plataforma *Artemis* es capaz de conjugar información arrojada por diferentes fuentes de datos y con ello, analizar campañas en conjunto y no por separado. Es decir, la herramienta no solo atribuye valor al medio donde se realiza la conversión, sino que tiene en cuenta, todos los medios digitales por los cuales un usuario pasa antes de generar la acción de compra. *Artemis* funciona en base a escenarios de *causa-efecto*, por lo que dispone de modelos matemáticos de simulación donde, a través de la incorporación de datos externos del mundo real y la experiencia, se miden y se predicen patrones de comportamiento. Por consiguiente, se permite disminuir la creciente complejidad de los medios digitales especialmente cuando se trata de analizar campañas multicanal y su impacto en los resultados de negocio (Marketingdirecto.com, 2012).

Ambas plataformas, tanto *Artemis* como *Google Analytics*, conllevan los siguientes beneficios a Affinity: saber qué canales ayudan a la empresa a encontrar nuevos clientes; conocer qué canales ayudan a la empresa a fidelizar los clientes actuales;

definir qué presupuesto se debe atribuir a cada canal; conocer qué canales se solapan con otros, y por tanto, obviarlos, evitando así costes innecesarios; y precisar mejor en qué canales se debe apostar más y en cuáles disminuir la inversión.

En la siguiente ilustración vemos un ejemplo puesto en práctica del informe de canales basado en conversiones y coste por adquisición (CPA). Así podemos ver qué canales atraen más clientes a un precio más económico. Por lo que se observa, el canal directo es el más eficiente, ya que no implica costes y obtiene altas conversiones de los usuarios (ver figura 10).

Figura 10. Informe de canales basado en conversiones y coste por adquisición.

MCF Channel Grouping	Spend (for selected time range)	Conversions & CPA				% change in Conversions (from First Interaction)
		First Interaction		Last Interaction		
		Conversions	CPA	Conversions	CPA	
1. Organic Search	—	55,727.00	—	53,119.00	—	-4.68%
2. Direct	—	20,671.00	—	23,986.00	—	16.04%
3. Social Network	—	13,662.00	—	13,286.00	—	-2.75%
4. Referral	—	3,454.00	—	3,183.00	—	-7.85%
5. Display	\$2,671.51	2,098.00	\$1.27	2,078.00	\$1.29	-0.95%
6. Paid Search	\$1,241.54	866.00	\$1.43	830.00	\$1.50	-4.16%
7. Other Advertising	—	81.00	—	77.00	—	-4.94%
8. Email	—	3.00	—	3.00	—	0.00%
9. (not set)	—	2.00	—	2.00	—	0.00%

*clasificaremos los canales según la nomenclatura y estructura*     
 *podrá utilizar diferentes modelos de atribución y comparar el rendimiento de sus canales de marketing*     
 *importaremos los costes de campaña para que pueda analizar en base a métricas de negocio*

Fuente: Elaboración propia a partir de Google Analytics.

## 5.2. Embudo de leads

En este proceso, como se ha mencionado, también se parte de Google Analytics para realizar un seguimiento de toda la transacción para conseguir un lead, desde que se inicia en la web hasta la visita y cierre del mismo. Las ventajas que podemos adquirir son: saber en qué punto del proceso se pierden a más clientes potenciales; cuántos leads se pierden a causa de tener que cumplimentar el formulario, previo, de datos; y qué contenidos tienen mejor o peor conversión a lead.

## 5.3. Contenido

Se configura Google Analytics para medir el rendimiento de la estrategia de marketing de contenido del sitio web. Esto, les permite responder a preguntas como:

¿qué tipo de contenidos en la *home* inducen al riesgo o al cumplimiento de los objetivos de la web?; ¿qué tipo de contenidos posiciona mejor en buscadores?; ¿qué tipo de contenido es el más importante sobre los productos?; ¿cuáles son los productos que más visitan consiguen?; ¿cuántos usuarios utilizan el buscador de la *home* y qué tipo de búsquedas realizan?; ¿qué marcas tienen un mayor retorno?; ¿qué contenidos fidelizan al usuario para que retorne al sitio web?; y ¿qué contenidos provocan que el usuario abandone el sitio web?. Con todas las respuestas podrán mejorar la estrategia.

#### **5.4. Clientes**

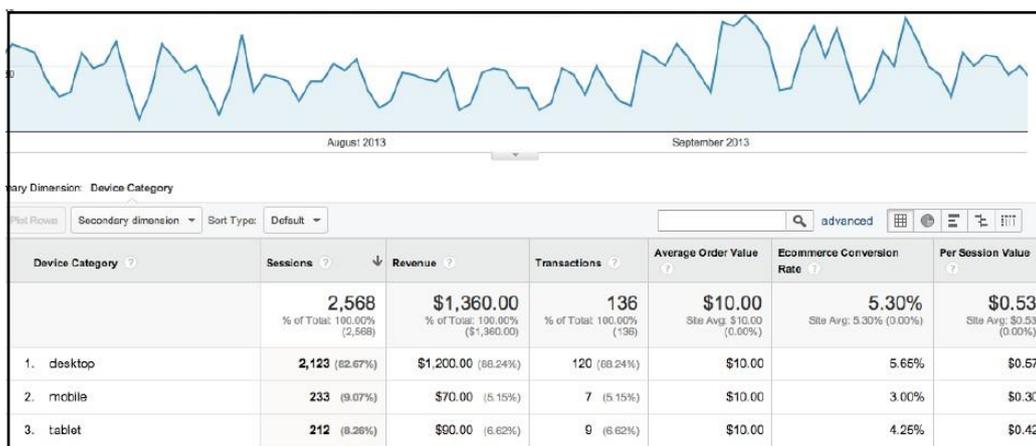
La configuración de *Google Analytics* permite analizar la tipología de clientes, la fidelización y los comportamientos más habituales de los mismos. Además, detecta qué perfil socio-demográfico prima en los clientes, con cuánta frecuencia visitan la web, antes de realizar una compra, y en qué horas y días consultan los contenidos.

#### **5.5. Movilidad**

La movilidad es un factor crítico en la adaptación a las nuevas tecnologías para todas las compañías. En particular, el caso de Affinity presenta también una estrategia de movilidad. La configuración de *Google Analytics* para realizar un seguimiento al detalle de todos los dispositivos de acceso a la web (ordenadores o escritorios “*desktops*”, tabletas, televisiones inteligentes “*Smart TV*”, móviles) les permitió obtener beneficios como: conocer la diferente efectividad entre tabletas y escritorio; saber qué diferencias de contenido demanda el usuario de móvil frente a un usuario de escritorio; entender qué problemas puede encontrar un usuario desde un móvil; conocer qué canales de marketing funcionan mejor para móviles y cuáles para tabletas; y saber cuáles son los móviles más populares entre los clientes de la compañía para adaptarse mejor.

Como mostramos en la figura siguiente, Affinity presenta un problema de movilidad, ya que la cantidad de *leads* que recibe son mayormente desde *desktop* y quedan, muy en segundo plano, las entradas procedentes de móviles y tabletas (ver figura 11).

Figura 11. Informe de la probabilidad de lead desde diferentes dispositivos.



Fuente: Elaboración propia a partir de Google Analytics.

Por otro lado, la próxima figura que mostramos a continuación, enseña el seguimiento de los móviles que tienen los usuarios que acceden a la web de la compañía. Por consiguiente, la empresa puede conocer qué dispositivos móviles son los más populares entre sus clientes y detectar problemas frente a modelos concretos. Como se muestra en la imagen, *iOs* y *Android* son los más usados y por tanto la empresa debe orientar sus esfuerzos hacia la mejora de estos en términos de adaptabilidad y movilidad (ver figura 12).

Figura 12. Informe de la probabilidad de lead desde diferentes modelos de móvil.

Mobile Device Info	Sessions	Revenue	Transactions	Average Order Value	Ecommerce Conversion Rate	Per Session Value
	445 % of Total: 17.33% (2,568)	\$160.00 % of Total: 11.76% (\$1,360.00)	16 % of Total: 11.76% (136)	\$10.00 Site Avg: \$10.00 (0.00%)	3.60% Site Avg: 5.30% (-32.11%)	\$0.36 Site Avg: \$0.53 (-32.11%)
1. Apple iPad				\$10.00	4.89%	\$0.47
2. Apple iPhone				\$10.00	5.00%	\$0.50
3. (not set)				\$0.00	0.00%	\$0.00
4. Samsung GT-I9300 Galaxy SIII				\$0.00	0.00%	\$0.00
5. Samsung GT-I9100 Galaxy S II				\$0.00	0.00%	\$0.00
6. Google Nexus 4	5 (1.12%)	\$0.00 (0.00%)	0 (0.00%)	\$0.00	0.00%	\$0.00
7. Samsung GT-I8190L Galaxy S III Mini	4 (0.90%)	\$0.00 (0.00%)	0 (0.00%)	\$0.00	0.00%	\$0.00
8. Samsung GT-I9001	4 (0.90%)	\$0.00 (0.00%)	0 (0.00%)	\$0.00	0.00%	\$0.00
9. Samsung GT-S5830 Galaxy Ace	4 (0.90%)	\$0.00 (0.00%)	0 (0.00%)	\$0.00	0.00%	\$0.00
10. SonyEricsson ST25i Xperia U	4 (0.90%)	\$0.00 (0.00%)	0 (0.00%)	\$0.00	0.00%	\$0.00

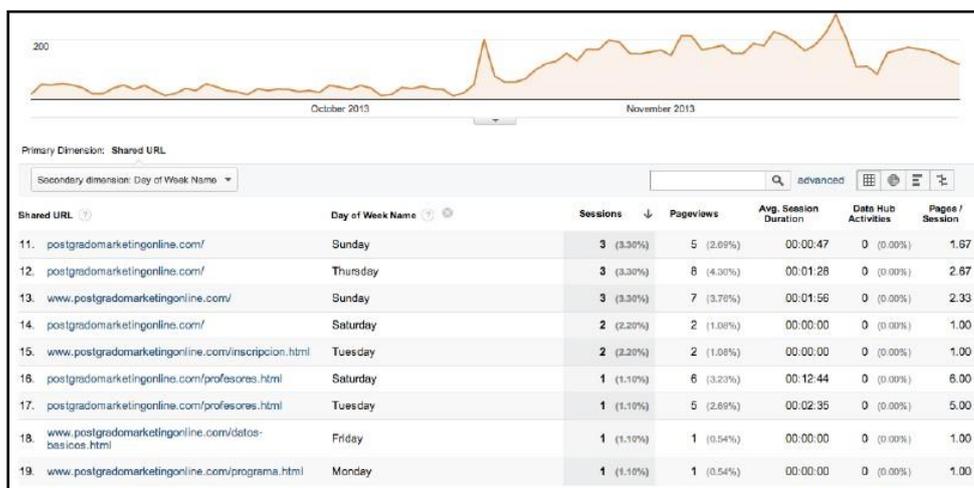
Fuente: Elaboración propia a partir de Google Analytics.

## 5.6. Redes Sociales

Respecto a las acciones de redes sociales, también se hace un informe de seguimiento con *Google Analytics*. Se analizan los contenidos web que se publican en redes sociales, el tráfico, la fidelización y los *leads* procedentes de las mismas. Con todo ello, se hallan una serie de respuestas cómo: los contenidos más populares o virales en redes sociales; el tráfico que las mismas aportan; la influencia que tienen en cuanto a conversiones; el grado de fidelización hacia la marca a causa de la actividad en redes sociales; entre otras.

La figura siguiente muestra el contenido más compartido en redes sociales por días de la semana. Esto permite conocer qué tipo de contenido se debe publicar en redes sociales cada día de la semana para maximizar la viralización (ver figura 13).

Figura 13. Informe del contenido más compartido en redes sociales por día de la semana.



Fuente: Elaboración propia a partir de *Google Analytics*.

Seguidamente, la imagen posterior permite a Affinity analizar qué sitios web o redes sociales están enlazando contenido de la compañía. Así, se pueden planificar acciones específicas con los sitios web y redes que aportan más valor (ver figura 14).

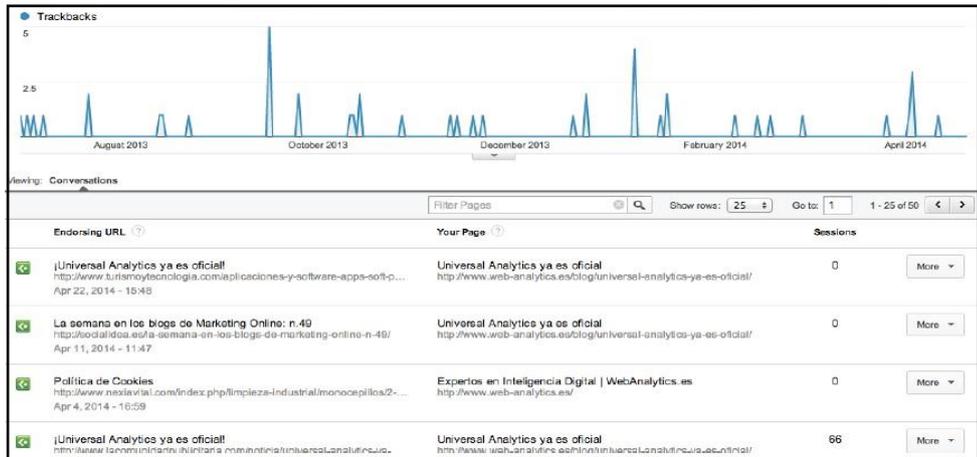
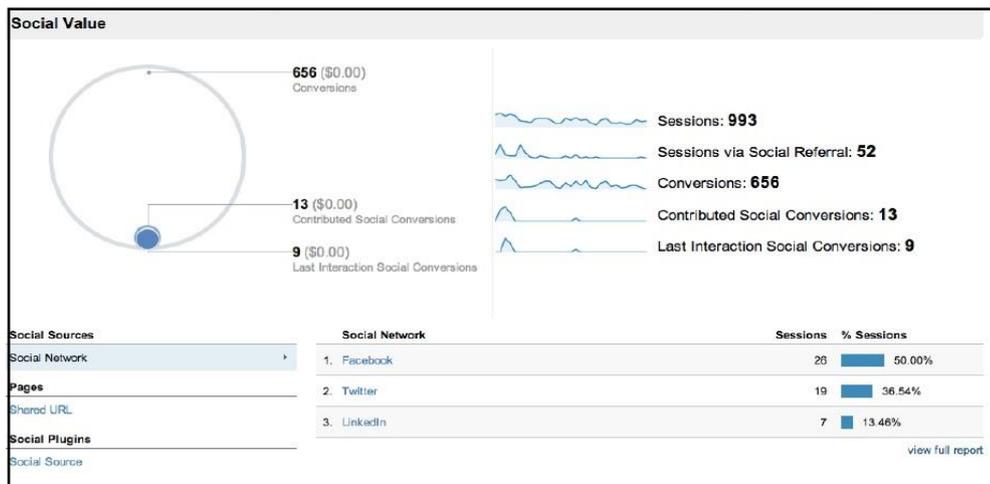


Figura 14. Informe del contenido más compartido en redes sociales por día de la semana.

Fuente: Elaboración propia a partir de Google Analytics.

Por último, otra utilidad en cuanto a redes sociales tiene que ver con la fidelización y el tráfico. Es, en la siguiente ilustración, donde vemos un informe comparativo de las diferentes redes sociales de Affinity y un análisis de qué redes aportan más tráfico, cuáles tienen usuarios más afines a la compañía, y qué tasa de conversión a lead representan (ver figura 15).

Figura 15. Informe comparativo de las diferentes redes sociales.



Fuente: Elaboración propia a partir de Google Analytics.

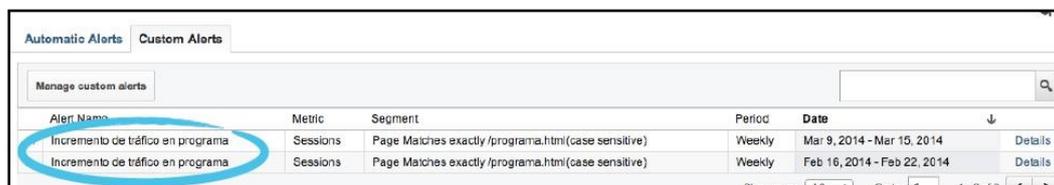
## 5.7. Alertas inteligentes

Destacar que *Google Analytics* permite también la fijación de alertas inteligentes. Son mensajes que detectan automáticamente las variaciones significativas en los diferentes *KPI's*, que hemos explicados en el marco teórico. Las alertas son complementarias a las tareas de análisis previas y aseguran que no surjan variaciones en la nueva estrategia de valor que se dispone a conseguir la empresa. Las alertas son necesarias porque los datos que se generan por las actuaciones de los consumidores son cambiantes, constantemente, y esto puede implicar una actualización de comportamiento de *Affinity* en términos de estrategia.

Entre las distintas ventajas que disponemos cuando incorporamos mecanismos de alertas inteligentes encontramos: conocer qué *KPI's* han sufrido mayores cambios significativos y cuáles son las causas; determinar cómo ha impactado una campaña de marketing o nuevo contenido en los *KPI's*; y saber qué productos han aumentado o disminuido en popularidad.

La siguiente figura adjunta una informe donde se alerta de una variación del tráfico de usuarios de una *site*, donde se había predefinido un número que ha sido excedido (ver figura 16).

Figura 16. Variaciones significativas para *KPI's* y segmentos previamente definidos.



Alert Name	Metric	Segment	Period	Date	
Incremento de tráfico en programa	Sessions	Page Matches exactly /programa.html(case sensitive)	Weekly	Mar 9, 2014 - Mar 15, 2014	Details
Incremento de tráfico en programa	Sessions	Page Matches exactly /programa.html(case sensitive)	Weekly	Feb 16, 2014 - Feb 22, 2014	Details

Fuente: Elaboración propia a partir de *Google Analytics*.

## 5.8. Reporting

Finalmente, la última táctica implementada en la empresa en cuestión, es el "*reporting*" o informe personalizado. Consta de crear cuadros de mando para reportar los datos que cada cliente quiere conocer. En consecuencia, se permite personalizar informes con las necesidades específicas de cada cliente con la finalidad de que *Affinity* pueda analizarlas a diario y de forma individual.

A la práctica, se configuran informes personalizados como: un cuadro de mando de campañas; cuadro de mando de buscadores; cuadro de mando de redes sociales; cuadro de mando de contenidos; cuadro de mando de *leads*; cuadro de mando de

rendimiento técnico; cuadro de mando de movilidad; entre otros. Cabe destacar que, en este caso, Affinity realiza más de 200 informes personalizados de modo que, al final consigue que cada usuario obtenga un acceso personalizado afín a sus necesidades y preferencias.

## **6. Indicadores que recibe Affinity tras la implantación del proyecto**

Una vez finalizado el proyecto, teniendo presente el objetivo principal a abarcar y las tácticas o estrategias mencionadas, como resultado la compañía dispone de un completo sistema de indicadores y cuadros de mando que le permiten analizar los resultados conseguidos y actuar en caso de desviación.

En concreto, Affinity presenta los siguientes servicios e indicadores tras finalizar la implantación del proyecto: configuración a medida de *Universal Analytics* de Google; cuadros de mando específicos con *Tableau*; formación sobre *Google Analytics* y *Tableau* a los empleados de la compañía; informe mensual de resultados y soporte técnico; y una consultoría en análisis así como en optimización de campañas.

## **7. Valoración final del proyecto de *Data Business Intelligence* de Affinity**

En primer lugar, hoy en día, destacamos la utilización del canal *online* como medio imprescindible de las organizaciones para llegar al público objetivo de forma más eficiente. Las ventas *online* son muy importantes para conocer mejor a los clientes en términos de qué marcas y productos prefieren, qué precios compran, qué días de la semana compran, cuánto compran, etc. En este ámbito, Affinity sabe detectar esta oportunidad y consigue optimizar la experiencia de sus clientes, segmentar mejor su estrategia empresarial y orientar más eficientemente su objetivo de negocio.

Asimismo, el mercado en auge de las tecnologías de la información y comunicación da una oportunidad, a todas las empresas ya implantadas en tienda *offline*, de enmarcarse dentro de una estrategia *omnichannel* donde se pretende dar una imagen única a través de todos los canales de comunicación de una compañía. Una vez más, Affinity sabe aprovechar esta oportunidad y se presenta al mercado con una visión única y una estrategia *omnichannel*.

También mencionar que Affinity reconoce la ventaja de la plataforma *Universal Analytics*, la cual predispone una serie de acciones muy productivas en varios ámbitos de actividad como en el marketing y la publicidad. Entre los beneficios que Affinity obtiene vemos: obtener nuevas fuentes de información; analizar contenidos en redes sociales y en la *site* de Affinity; determinar la movilidad de los clientes; o detectar la cantidad de *leads* y conversiones, entre otras. Todas ellas monitorizadas y seguidas con *KPI's* de negocio.

Por último, destacar la utilización de la herramienta *Tableau* para la visualización de todo el proyecto de datos de inteligencia empresarial, capaz de poner el foco en el cliente, pilar fundamental en los valores de Affinity.

## Conclusiones

Affinity, la empresa de nuestro caso, ha logrado colocarse en el número uno de su mercado gracias a implantar una estrategia de marketing continuada basada en la tecnología del dato. Ha sabido recolectar, mediante las herramientas explicadas, toda la información precisa de sus clientes para personalizar y presentarles una oferta capaz de cubrir todas las necesidades y deseos insatisfechos, tanto para sus consumidores reales como aquellos potenciales, y lo ha hecho, además, avanzándose a sus movimientos, es decir, ha utilizado la funcionalidad algorítmica de predecir los resultados antes de que sucedan. Con ello, Affinity ha podido adaptar, de forma dinámica, sus *KPI's* a las respuestas reales y posibles que ha ido obteniendo de los usuarios. Por tanto, vemos como esta empresa multinacional, se ha convertido en líder de su sector basando su modelo de negocio en obtener del análisis de los datos una ventaja competitiva constante. Por consiguiente, este caso es el ejemplo donde el *Big Data* y el marketing funcionan como el binomio perfecto para el éxito, pudiendo afirmar que esta empresa se ha adentrado ya en la nueva era de la transformación digital.

A diario, aparecen artículos sobre este nuevo concepto y el *Big Data* como eje fundamental, o como el nuevo impulsor, de la economía de nuestro tiempo. Vista la potencia de estas herramientas concluimos que realmente resulta esencial, para la vida de las empresas, entrar en el mercado de la tecnología del dato y el marketing para poder posicionarse y satisfacer a los usuarios.

Comprobamos, que a las cuatro *P's* del marketing que comentábamos en el cuerpo del trabajo, ha llegado otra para quedarse, la que hemos llamado “proceso”, y que no es otra cualquiera, sino la que esconde la piedra filosofal de este nuevo paradigma, el algoritmo matemático, que nos permite hacer llegar a nuestro cliente una experiencia de compra única; con ello obtendremos el premio de su fidelidad y un canal de constante aprendizaje de valor incalculable para seguir adaptando nuevas ofertas de marketing a sus gustos, costumbres o necesidades.

Cuando hemos analizado las fuerzas competitivas y, la influencia en ellas del *Big Data*, podemos afirmar que nuestra “*P*” de proceso cobra especial relevancia. Los competidores ya no están llegando por los canales tradicionales del mercado, son nuevos competidores, lejos del negocio tradicional y expertos en hacer llegar al cliente sus soluciones a través de Internet, por ejemplo, la mayor empresa de creación de contenidos, Facebook, no tiene ni un contenido propio, la mayor empresa de gestión de reservas de taxis, UBER, no tiene ni un taxi propio y la mayor

empresa de venta de libros, Amazon, no vende ni un libro propio, y así podríamos enumerar muchas empresas hasta hace poco desconocidas.

Se abren nuevas oportunidades, donde el valor del negocio se ha desplazado a quién está más cerca del cliente, mejora su experiencia hábilmente y se aplica en cómo llegar digitalmente a él. Nuestra empresa del caso práctico lo sabe muy bien.

Destacar que hemos aprendido, como diríamos, en lenguaje coloquial, que no es oro todo lo que reluce y que, también en esta disciplina, encontramos límites como la protección de los datos personales de los clientes que siempre tendremos que salvaguardar y respetar o barreras potenciales como los datos sucios, datos inexactos, distorsionados o erróneos que precisan de análisis pero que continúan siendo datos de valor, ya que lo que para algunos puede ser información desechable, para otros, como Google o IBM, hemos visto que pueden ser oportunidades de negocio.

Por otro lado, y no menos importante, al redactar este proyecto hemos visto las nuevas posibilidades de trabajo que se están y se seguirán generando, en el mercado del *Big Data* orientado al marketing. Hemos hablado del “científico de datos” que se encarga de la recolección de los mismos desde distintos silos, ya sean estructurados o no estructurados, pero, además, todos estos datos deberán ser solicitados por una nueva figura que podríamos llamar nuestro “científico del marketing”, quien, constantemente, tiene que trabajar los resultados para aumentar el grado de conocimiento de nuestros clientes y optimizar nuestras estrategias de mercadotecnia. Solo así, daremos con los gustos, necesidades y preferencias adecuadas de nuestros consumidores.

Por último, concluir que, desde un punto de vista empresarial, hemos visto como ya no se demandan organizaciones seguras y fiables, sino ágiles y flexibles en las que nuestra “P” de los procesos debe ser capaz de adaptarse a la “P” de las personas, a las necesidades de los clientes que son y serán, cada vez, más cambiantes. En este sentido, vemos que un buen consejo para las empresas es invertir en el dato para poder alcanzar el conocimiento de sus clientes.

Apoyarnos en las tecnología de *Big Data* para orientar correctamente nuestra estrategia de marketing, nos permitirá observar a nuestros clientes y disponer de un aprendizaje constante de ellos, pilares estratégicos para el éxito que debemos tener siempre en consideración.

## Bibliografía

- Activa internet. (24 de febrero de 2014). Ejemplos exitosos con campañas de remarketing [Mensaje en un blog]. Recuperado el 15 de abril de 2016 de <http://www.activainternet.es/ejemplos-exitosos-campanas-remarketing/>
- Affinity Petcare. (2016). Recuperado el 5 de mayo de 2016 de <http://www.affinity-petcare.com/es>
- Arthur, L. (2013). *Big Data Marketing: Engage your customer more effectively and drive value*. New Jersey: Wiley.
- Ayora (2015). *Mentiras a lo grande: una introducción al Big Data en el marketing y el efecto de los dirty data*. (Trabajo de fin de grado). Universitat Oberta de Catalunya, Barcelona.
- Barcelona Activa. (2013). *Telecomunicacions i TIC. (Informe sectorial 2013)*. Barcelona: Ajuntament de Barcelona.
- Barrabino, A. (2014). *Nivel operativo de actuación del marketing: políticas de marketing*. Manuscrito inédito, Departamento de Marketing y Dirección Comercial, Universitat Abat Oliba CEU, Barcelona, España.
- BBVA *Innovation Center*. (30 de enero de 2015). Ejemplos reales del uso de Big Data: Macy's y sus precios en tiempo real [Comentario en un grupo de noticias]. Recuperado el 12 de abril de 2016 de <http://www.centrodeinnovacionbbva.com/noticias/ejemplos-reales-del-uso-de-big-data>
- Blázquez, S. (27 de junio de 2015). Animales que generan millones. *El País*. Recuperado el día 5 de mayo de 2016 de [http://economia.elpais.com/economia/2015/06/26/actualidad/1435315494\\_296503.html](http://economia.elpais.com/economia/2015/06/26/actualidad/1435315494_296503.html)
- Buckley, J. (27 de agosto de 2015). Causes of Dirty Data and how to combat them [Mensaje en un blog]. Recuperado el 17 de abril de 2016 de [https://www.qubole.com/blog/big-data/causes-of-dirty-data-and-how-to-combat-them/?nabe=5695374637924352:1&utm\\_referrer=https://www.google.es/](https://www.qubole.com/blog/big-data/causes-of-dirty-data-and-how-to-combat-them/?nabe=5695374637924352:1&utm_referrer=https://www.google.es/)
- Cabezas, M. (28 de abril de 2016). Valor, valor y más valor [Mensaje en un blog]. Recuperado el 8 de mayo de 2016 de <http://www.transformaciondigital.es/blog/valor-valor-y-mas-valor/>
- Cano, J. Ll. (2007). *Business Intelligence: Competir con información* [versión PDF]. Recuperado el 21 de abril de 2016 de [http://itemsweb.esade.edu/biblioteca/archivo/Business\\_Intelligence\\_competir\\_con\\_informacion.pdf](http://itemsweb.esade.edu/biblioteca/archivo/Business_Intelligence_competir_con_informacion.pdf)
- Capron, E. (2013). How Big Data has delivered for FedEx for 25 years [Versión Electrónica]. *Digitalist Magazine*, (1 de noviembre de 2013). Recuperado el 8 de abril de 2016 de <http://www.digitalistmag.com/future-of-work/2013/11/01/big-data-delivered-fedex-25-years-0887095>

- Chahal, M. (2015). Consumers are 'dirtying' databases with false details. *Marketing Week*, (8 de julio de 2015). Recuperado el 17 de abril de 2016 de <http://www.marketingweek.com/2015/07/08/consumers-are-dirtying-databases-with-false-details/>
- Cordero, M. (2016). La influencia del Big Data en la publicidad. *El Publicista*, (28 de marzo de 2016). Recuperado el 15 de abril de 2016 de [http://www.elpublicista.es/frontend/elpublicista/noticia2.php?id\\_noticia=23685&id\\_seccion=33](http://www.elpublicista.es/frontend/elpublicista/noticia2.php?id_noticia=23685&id_seccion=33)
- Del Olmo, J. Ll. (2016). *El marketing mix de la moda: factores que influyen en la fijación de precios*. Manuscrito inédito, Departamento de Marketing y Dirección Comercial, Universitat Abat Oliba CEU, Barcelona, España.
- Del Olmo, J. Ll. (2015). *Introducción a la promoción de ventas: la promoción como parte integrante del Marketing Mix*. Manuscrito inédito, Departamento de Marketing y Dirección Comercial, Universitat Abat Oliba CEU, Barcelona, España.
- Europa Press. (12 de mayo de 2016). Affinity facturó 458 millones en 2015 y espera superar los 500 por la compra de Nova Foods. *La Vanguardia*. Recuperado el 16 de mayo de 2016 de <http://www.lavanguardia.com/vida/20160512/401748178484/catalunya-affinity-facturo-458-millones-en-2015-y-espera-superar-los-500-por-la-compra-de-nova-foods.html>
- Gartner. (2016). *IT Glossary: Big Data*. Recuperado el 25 de febrero de 2016 de <http://www.gartner.com/it-glossary/big-data/>
- Gittleson, K. (15 de noviembre de 2013). How big data is changing the cost of insurance. *BBC News*. Recuperado el 6 de abril de 2016 de <http://www.bbc.com/news/business-24941415>
- Hedge, H. (2014). *Netflix and its revolutionary use of Big Data*. Recuperado el 8 de abril de 2016 de <http://dataconomy.com/netflix-and-its-revolutionary-use-of-big-data/>
- High, P. (2015). CIO Rob Carter transforms FedEx into a cloud centric enterprise. *Forbes*, (20 de abril de 2015). Recuperado el 12 de abril de 2016 de <http://www.forbes.com/sites/peterhigh/2015/04/20/cio-rob-carter-transforms-fedex-into-a-cloud-centric-enterprise/#6e86c5644f3a>
- Hilbert, M. (Productor). (2015). *What is Big Data?* [Video]. Estados Unidos: Digital Technology & Social Change University of California course.
- Kopytoff, V. (2014). Big Data's dirty problem. *Fortune*, (30 de junio de 2014). Recuperado el 17 de abril de 2016 de <http://fortune.com/2014/06/30/big-data-dirty-problem/>
- Kotler, P. (2001). *Dr. Philip Kotler answers your questions on marketing*. Recuperado el 7 abril de 2016 de [http://www.kotlermarketing.com/phil\\_questions.shtml#answer5](http://www.kotlermarketing.com/phil_questions.shtml#answer5)
- Kotler, P. y Armstrong, G. (2013). *Fundamentos de marketing*. (11ª ed. rev.). México: Pearson Educación.
- Larraoetxea, C. (14 de diciembre de 2015). El 'big data' permite a Eroski lanzar 30.000 promociones por quincena, casi todas diferentes. *El Economista*. Recuperado el 17 de abril de 2016 de [http://www.economista.es/pais\\_vasco/noticias/7217840/12/15/EI-big-data-permite-a-Eroski-lanzar-30000-promociones-por-quincena-casi-todas-diferentes.html](http://www.economista.es/pais_vasco/noticias/7217840/12/15/EI-big-data-permite-a-Eroski-lanzar-30000-promociones-por-quincena-casi-todas-diferentes.html)

- Ley orgánica 34/2002, de 11 de julio, de servicios de la información y de comercio electrónico. (BOE, núm. 166, 12-07-2002, pp. 25388-25403).
- Ley orgánica 15/1999, de 13 de diciembre, de Protección de Datos de Carácter Personal. (BOE, núm. 281, 14-12-1999, pp. 43088-43099).
- Lucas, M. A. (23 abril 2015). Big Data, significado y su utilidad en la sociedad [Mensaje en un blog]. Recuperado el 25 de febrero de 2016 de <http://mibloguel.com/big-data-significado-y-su-utilidad-en-la-sociedad/>
- Marketingdirecto.com. (2016). *Oportunidades y retos del Big Data (más allá del social media) #Santander SW*. Recuperado el 12 de abril de 2016 de <http://www.marketingdirecto.com/marketing-general/eventos-y-formacion/oportunidades-y-retos-big-data-mas-alla-social-media-santandersw/>
- Marketingdirecto.com. (2014). *House of Cards, la primogénita del “matrimonio” formado por la televisión y el Big Data*. Recuperado el 8 de abril de 2016 de <http://www.marketingdirecto.com/anunciantes-general/medios/house-of-cards-la-primogenita-del-matrimonio-formado-por-la-television-y-el-big-data/>
- Marketingdirecto.com. (2012). *Havas presenta Artemis, una herramienta de medición y predicción para optimizar campañas de publicidad online*. Recuperado el 7 de mayo de 2016 de <http://www.marketingdirecto.com/digital-general/digital/havas-presenta-artemis-una-herramienta-de-medicion-y-prediccion-para-optimizar-campanas-online/>
- Marketingdirecto.com. (2005). *Datos simples semiestructurados*. Recuperado el 25 de febrero de 2016 de <http://www.marketingdirecto.com/actualidad/marketing/datos-simples-semiestructurados/>
- Marketingnews.es. (2010). *Affinity Petcare estrena web y hace su entrada en las redes sociales*. Recuperado el 5 de mayo de 2016 de <http://www.marketingnews.es/gran-consumo/noticia/1046457028005/affinity-petcare-estrena-web-hace.1.html>
- Meek, T. (2014). In-car sensors put insurers in the driver's seat. *Forbes*, (27 de junio de 2014). Recuperado el 8 de abril de 2016 de <http://www.forbes.com/sites/ptc/2014/06/27/in-car-sensors-put-insurers-in-the-drivers-seat/#dbb3d7a39786>
- Moll, E. (29 de julio de 2014). ¿Qué es el remarketing, cómo funciona y para qué sirve? [Mensaje en un blog]. Recuperado 15 de abril de 2016 de <https://masquesocial.com/2014/07/29/remarketing-definicion-funcionamiento-beneficios/>
- Moore, S. (2 de marzo de 2007). 'Dirty Data' is a Business Problem, not an IT problem, says Gartner [Comentario en un grupo de noticias]. Recuperado el 17 de abril de 2016 de <http://www.gartner.com/newsroom/id/501733>
- Morris, D. (2016). Netflix says Geography, Age, and Gender are “garbage” for predicting taste. *Fortune*, (27 de marzo de 2016). Recuperado el 8 de abril de 2016 de [http://fortune.com/2016/03/27/netflix-predicts-taste/?xid=for\\_fb\\_sh](http://fortune.com/2016/03/27/netflix-predicts-taste/?xid=for_fb_sh)

- Muñoz, A. (2016). Por qué ser científico de datos es la profesión del año. *TIC Beat*, (27 de enero de 2016). Recuperado el 28 de abril de 2016 de <http://www.ticbeat.com/bigdata/por-que-ser-cientifico-de-datos-es-la-profesion-del-ano/>
- Navarro, M. (2016). Los científicos de datos, una profesión al alza. *Byte TI*, (11 de febrero de 2016). Recuperado el 21 marzo de 2016 de <http://www.revistabyte.es/actualidad-byte/62971cientificos-de-datos-una-profesion-al-alza/>
- Navarro, M. (2015). Big Data, el análisis correcto del dato [Versión Electrónica]. *Byte TI, época III* (231), 30-41. Recuperado el 11 marzo de 2016 de <http://www.revistabyte.es/tema-de-portada-byte-ti/big-data-el-analisis-correcto-del-dato/>
- O'Reilly, T. (15 de septiembre de 2015). Improving Uber's surge pricing: harnessing the power of markets and multi-factor algorithms [Mensaje en un blog]. Recuperado 8 de abril de 2016 de <http://medium.stfi.re/the-wtf-economy/improving-uber-s-surge-pricing-3fd2fe108bd6?sf=jxoyko#.xix2kvvcs>
- Páramo, A. (10 de febrero de 2016). El peligro del "dirty data" cuando no todos los datos valen [Mensaje en un blog]. Recuperado el 17 de abril de 2016 de <http://www.adalidinternet.com/el-peligro-del-dirty-data-cuando-no-todos-los-datos-valen-49532.htm>
- Parenteau, J. et al. (2016) *Gartner: Magic Quadrant for Business Intelligence Analytics Platforms*. Recuperado el 1 mayo 2016 de <https://www.gartner.com/doc/reprints?id=1-2XXET8P&ct=160204&st=sb>
- Pérez, B. (20 de enero de 2016). Qué es Google Tag Manager [Comentario en un blog]. Recuperado el 1 mayo 2016 de <http://www.idento.es/blog/marketing-online/que-es-google-tag-manager-y-para-que-sirve/>
- Porter, M. E. (2009). *Estrategia competitiva: técnicas para el análisis de la empresa y sus competidores*. Madrid: Pirámide.
- ProfesionalesHoy. (15 de febrero de 2016). Los científicos de datos. Una profesión en alza [Comentario en un grupo de noticias]. Recuperado el 28 de abril de 2016 de <http://profesioneshoy.es/tics/2016/02/15/los-cientificos-de-datos-una-profesion-al-alza/19270>
- Puromarketing.com. (2015). *Los retos del 'Dirty Data' cuando los consumidores engañan con sus datos falsos a las empresas*. Recuperado el 17 de abril de 2016 de <http://www.puromarketing.com/12/25111/retos-dirty-data-cuando-consumidores-enganan-datos-falsos-empresas.html>
- Puyol, J. (2014). *Una aproximación al Big Data*. Revista de Derecho UNED, (14), 471- 505.
- Quincoces, A. (17 de marzo de 2016). España un país muy tuitero. *Diario Información*. Recuperado el 22 de mayo de 2016 de <http://www.diarioinformacion.com/vida-y-estilo/tecnologia/2016/03/17/espana-pais-tuitero/1740107.html>
- REAL ACADEMIA ESPAÑOLA. (2016). Byte. En: *Diccionario de la lengua española*. [Versión electrónica] Recuperado el 1 de mayo de 2016 de <http://dle.rae.es/?id=6MofNvJ>

- Real Decreto 1720/2007, de 21 de diciembre, por el que se aprueba el Reglamento de desarrollo de la Ley Orgánica 15/1999, de 13 de diciembre, de protección de datos de carácter personal. (BOE, núm. 17, 19-01-2008, pp. 4103-4136).
- Ribé, N. (2015). Grupo Mars factura más de 300 millones en España y prevé un incremento del 7% en 2015. *Revista Aral*, (29 de junio de 2015). Recuperado el día 5 de mayo de 2016 de <http://www.revistaaral.com/es/notices/2015/06/grupo-mars-factura-mas-de-300-millones-en-espana-y-preve-una-incremento-de-un-7-75018.php>
- Robinat, J. R. (2015). *CRM gestión de la cartera de clientes P1*. Manuscrito inédito, Departamento de Marketing y Dirección Comercial, Universitat Abat Oliba CEU, Barcelona, España.
- Roca, G. (11 de diciembre de 2015). Los seis pilares de una transformación digital. *Expansión*. Recuperado el 7 de abril de 2016 de <http://www.expansion.com/economia-digital/2015/12/11/566abf4aca4741577e8b45a9.html>
- Rodríguez, I. (2007). *Estrategias y técnicas de comunicación: una visión integrada en el marketing*. Manuscrito inédito, Departamento de Marketing, Universitat Oberta de Catalunya, Barcelona, España
- San Agustín, A. (2016). Big Data: el tótum revolútum de la publicidad que quiere dejar de serlo. *El Publicista*, (17 de marzo de 2016). Recuperado el 15 de abril de 2016 de [http://www.elpublicista.es/frontend/elpublicista/noticia2.php?id\\_noticia=23684&id\\_seccion=33](http://www.elpublicista.es/frontend/elpublicista/noticia2.php?id_noticia=23684&id_seccion=33)
- Schmarzo, B. (2014). *Big Data. El poder de los datos*. Madrid: Anaya Multimedia.
- Schroeck, M., Shockley, R., Smart, J., Romero-Morales, D. y Tufano, P. (2012). *Analytics: el uso de big data en mundo real. Cómo las empresas más innovadoras extraen valor de datos inciertos. (Informe ejecutivo)* [Versión PDF]. Recuperado el 20 de marzo de 2016 de [http://www-05.ibm.com/services/es/gbs/consulting/pdf/El\\_uso\\_de\\_Big\\_Data\\_en\\_el\\_mundo\\_real.pdf](http://www-05.ibm.com/services/es/gbs/consulting/pdf/El_uso_de_Big_Data_en_el_mundo_real.pdf)
- Serrat, R. (2013). *Big Data: análisis de herramientas y soluciones*. (Trabajo de fin de carrera). Facultat d'Informàtica de Barcelona - UPC, Barcelona.
- Solana, A y Roca, G. (2015). *Big Data para directivos: guía rápida y ejemplos prácticos*. Barcelona: Empresa Activa.
- Terdoslavich, W. (2016) Hadoop at 10: Doug Cutting on making Big Data work. *Information Week*, (2 de enero de 2016). Recuperado el 21 de abril de 2016 de <http://www.informationweek.com/big-data/software-platforms/hadoop-at-10-doug-cutting-on-making-big-data-work/d/d-id/1324138>
- The Apache Software Foundation. (2014). *Welcome to Apache™ Hadoop®! What is Apache Hadoop?*. Recuperado el 21 de abril de 2016 de <http://hadoop.apache.org>
- Think Up-LKS. (2016a). *7 Ejemplos de cómo el Big Data ha cambiado nuestra vida y nuestros negocios*. Recuperado el 27 de abril de 2016 de <http://www.thinkuplks.com/7-ejemplos-de-como-el-big-data-ha-cambiado-nuestra-vida-y-nuestros-negocios/>

- Think Up-LKS. (2016b). *Aumenta la inversión en servicios de movilidad inteligente para las Smart cities*. Recuperado el 27 de abril de 2016 de <http://www.thinkuplks.com/aumenta-la-inversion-en-servicios-de-movilidad-inteligente-para-las-smart-cities/>
- Tilves, M. (2016). *Los ingresos por Business Intelligence están a punto de rozar los 17.000 millones de dólares*. Recuperado el 28 de abril de 2016 de <http://www.silicon.es/los-ingresos-business-intelligence-estan-punto-rozar-los-17-000-millones-dolares-94444>
- Torres i Viñals, J. (2012). *Del cloud computing al big data: visión introductoria para jóvenes emprendedores*. Manuscrito inédito, Departamento de Tecnología, Universitat Oberta de Catalunya, Barcelona, España.
- Trazada Marketing. (8 de agosto de 2013). *Omni-channel, el nuevo paso del retail marketing* [Comentario en un blog]. Recuperado el 1 mayo 2016 de <http://trazada.com/omni-channel-el-nuevo-paso-del-retail-marketing/>

## Anexo I: Glosario

*Algoritmo*: es un conjunto ordenado de operaciones sistemáticas que permite hacer un cálculo y extraer una solución a un tipo de problema.

*Almacén de datos*: En el contexto de la informática, un almacén de datos (del inglés *data warehouse*) es un repositorio de recolección de datos que aplica a una determinada organización. El almacén de datos ayuda a la toma de decisiones de la entidad que lo utiliza.

*Base de datos relacional*: es un tipo de base de datos que permite establecer relaciones entre datos, almacenados en tablas. El lenguaje que maneja la base de datos relacional es denominado *SQL* (acrónimo del inglés *Structured Query Language*), es decir, solo engloba datos estructurados.

*Empresas fintech*: son empresas especializadas en ofrecer servicios financieros vía tecnología. Servicios financieros como por ejemplo descuentos de letras, poner en común empresas para que entre ellas se presten créditos sin pasar por un banco como intermediario, entre otros.

*ERP*: acrónimo del inglés *Enterprise Resource Planning*, es decir, planificación de los recursos empresariales. El *ERP* trata de unir todos los flujos de una organización para que no existan pérdidas de información a nivel interno.

*ETL*: acrónimo del inglés *Extract, Transform and Load*. Es el proceso que extrae, limpia, transforma, integra y actualiza datos que provienen de diferentes fuentes de información para nutrir un almacén de datos.

*Hardware*: Es el conjunto de infraestructura física que constituye el sistema informático.

*KPI's*: acrónimo del inglés *Key Performance Indicators*, es decir, indicadores clave para la ejecución del negocio. Son mediciones cuantificables que reflejan factores críticos de negocio.

*Motor OLAP*: *OLAP* es el acrónimo en inglés de procesamiento analítico en línea (es decir, *On-Line Analytical Processing*). Es una solución utilizada en el campo de la llamada Inteligencia de negocios (o *Business Intelligence*) cuyo objetivo es agilizar la consulta de grandes cantidades de datos.

*Nodo*: Se entiende por nodos los diferentes servidores que almacenan grandes cantidades de datos y, además que son procesados en paralelo y horizontalmente para evitar errores.

*NoSQL*: En informática, *NoSQL* (a veces llamado "no sólo *SQL*") es una amplia clase de sistemas de gestión de bases de datos que difieren del modelo de base de datos relacional. En este caso, los datos almacenados no requieren de una estructura.

*Nube*: En informática, se entiende la nube como un espacio de almacenamiento de datos y archivos ubicado en internet, al que puede acceder el usuario desde cualquier dispositivo.

*Procesamiento por lotes o batch:* es un tipo de programa que permite la ejecución del mismo sin ningún tipo de control o supervisión directa de un usuario. Es decir, este tipo de programas no precisan de interacción con el usuario y permiten ejecutar tareas de grandes volúmenes de datos sin probabilidad de error.

*Remarketing:* Es una técnica de marketing que consiste en recoger datos de navegación de productos que no han llegado a ser conversiones de compra de usuarios, y permite impactarlos de nuevo con publicidad en otras plataformas web con esos mismos productos.

*SCM:* acrónimo del inglés *Supply Chain Management*, es decir, gestión de la cadena de suministro. Son las herramientas y métodos cuyo propósito es mejorar y automatizar el suministro a través de la reducción de las existencias y los plazos de entrega.

*Servidor:* Es una unidad informática que proporciona diversos servicios a través de una red de conexión a diferentes sistemas informáticos.

*Sistema transaccional:* Es un tipo de sistema de información diseñado para recolectar, almacenar, modificar y recuperar cualquier tipo de información generada por transacciones dentro de una organización.

*Software:* es un conjunto de programas que permite a cualquier sistema informático realizar determinadas tareas.

*Stakeholder:* Son las distintas personas que pueden ser afectadas por la actividad de una organización.

*Tecnologías de la información (TI)* o término también conocido como IT, del inglés *Information Technology*: Son todos los equipos de telecomunicaciones y otras aplicaciones de ordenadores. Todo ello para almacenar, recuperar, transmitir y manipular datos. Con frecuencia, TI es utilizado en el contexto de los negocios u otras empresas.

