

# Introducción: la Economía de los Datos

## 1.1

**Definición, evolución y situación actual desde una perspectiva económica**



La progresiva incorporación de los avances acumulados en las últimas décadas en materia de tecnología y redes de comunicación —articuladas en torno a Internet— a la producción de bienes y servicios ha contribuido decisivamente a reconfigurar las relaciones económicas y sociales a escala global. De la época analógica a la era digital, se han sucedido notables transformaciones en todos los ámbitos: desde la manera de consumir bienes y servicios —por ejemplo, a través del *eCommerce*— a la manera de producirlos —los servicios financieros o los medios de comunicación son buenos exponentes— o regularlos.

No hay que entender este proceso como un cambio estructural repentino —pocos en la historia pueden caracterizarse así—, sino que es producto de una evolución paulatina, que alterna periodos de mayor intensidad transformadora con otros de menor profusión. Sin embargo, a lo largo de toda esa línea evolutiva, el elemento nuclear que permanece y que posibilita estas transformaciones es la información. En otras épocas fue la aparición de fuentes de energía o el descubrimiento de innovaciones en el transporte. No es de extrañar, por consiguiente, que se aluda a este episodio de la historia en el que nos encontramos como la «era de la información»; desde los años setenta del siglo xx se viene gestando una revolución en la que los datos pasan a ocupar un lugar central para explicar la realidad económica y social y sus transformaciones.

En el ámbito de la economía, la información adquiere una relevancia creciente a la hora de explicar el producto nacional de los países, dado el protagonismo adquirido por las industrias intensivas en datos y la importancia que reviste como *input* fundamental de producción. En este sentido, los datos se convierten en esenciales para el desarrollo de la Economía Digital, en la medida en que habilitan la innovación, la eficiencia de los procesos y la sofisticación de los bienes y servicios que se producen. Y es que una de las características de la época digital es la producción y el almacenamiento de datos e información a bajo coste, lo que retroalimenta el círculo virtuoso de la innovación.

El proceso evolutivo de transformación económica y social antes aludido se encuentra en un momento de pleno apogeo. La tecnología avanza a pasos agigantados —la inteligencia artificial y el Internet de las Cosas son ya realidades que comienzan a penetrar en la sociedad— y, con ella, surgen nuevos retos, como la preocupación sobre la utilidad que se le brinde a la utilización de la ingente cantidad de datos generados de forma automática, a pesar de ser este el punto donde reside la oportunidad de la transformación digital.

Los datos generados por ciudadanos, empresas, Administraciones públicas y, más recientemente, objetos (Internet de las Cosas, IoT) han aumentado exponencialmente. En muchos casos, se trata de información asociada al usuario, esto es, datos personales que permitirían identificar a una persona concreta, por lo que requieren —los datos y las personas que los generan y/o comunican— una protección especial en ámbitos relacionados con la transparencia, el consentimiento, la finalidad, la protección, la seguridad, la conservación, la calidad y la exactitud, aspectos que se irán analizando en los diversos capítulos de esta publicación.

La seguridad digital, por motivos análogos, se convierte en un aspecto clave en la nueva Economía de los Datos.



En el ámbito europeo, el Mercado Único Digital es considerado uno de los ámbitos de progreso más prometedores y de los que más retos plantean. Los avances que conlleva dicho mercado requieren un marco normativo que facilite el desarrollo de la computación en nube, la conectividad de datos móviles sin fronteras y un acceso simplificado a la información y al contenido, preservando la privacidad, los datos personales o la seguridad cibernética, entre otros.

## 1.1

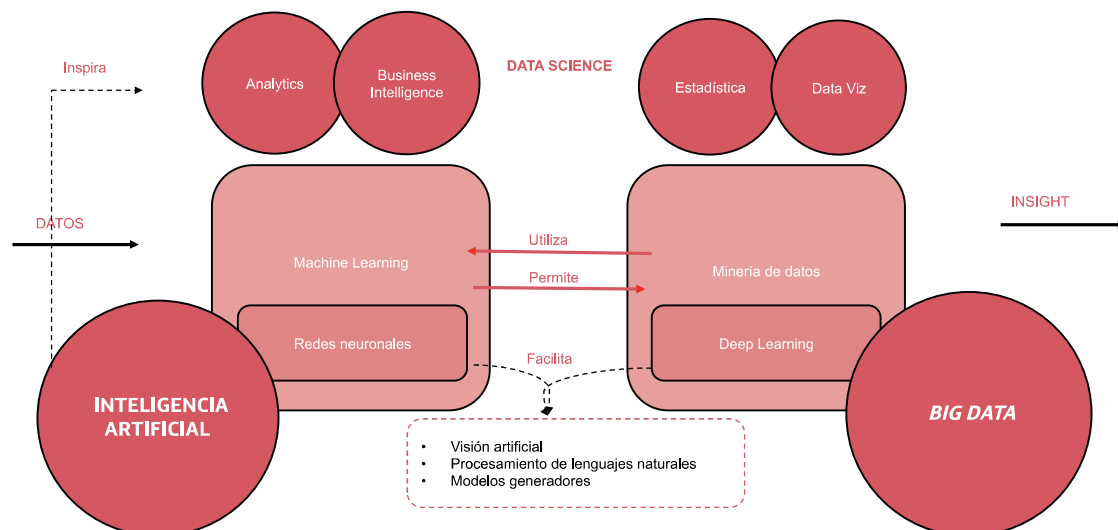
### Definición, evolución y situación actual desde una perspectiva económica

#### 1.1.1 Definición: conceptualización y perímetro

En los últimos años, especialmente desde la popularización del acceso y el uso de Internet, los entornos sociales, económicos, culturales y productivos enfrentan cambios que afectan desde las tareas más sencillas —como la comunicación interpersonal o las decisiones de compra cotidiana— a otras más complejas, como el funcionamiento de la cadena de producción de la industria o el *marketing* de una empresa, que están suponiendo una auténtica transformación de carácter digital de corte transversal en todos los ámbitos de nuestra vida. Muchos de estos cambios vienen explicados por la innovación tecnológica, que en el siglo *xxi* ha alcanzado velocidades vertiginosas, habilitada en buena medida por la creciente capacidad para obtener, procesar y analizar datos.

Fig. 1.1

La relación entre *big data*, *data science*, *machine learning*, *deep learning* e inteligencia artificial.



Fuente: whatsthebigdata.com.

La Economía de los Datos puede definirse como el conjunto de iniciativas, actividades y/o proyectos cuyo modelo de negocio se basa en la exploración y explotación de las estructuras de bases de datos existentes (tradicionales y procedentes de nuevas fuentes) para identificar oportunidades de generación de productos y servicios.

El nuevo contexto de relaciones en torno al dato trae consigo nuevas reglas de juego para los agentes —ciudadanos, Gobiernos y empresas— de carácter transversal —no discrimina entre sectores— y supone un cambio en las formas de producción y consumo, de planificación y gestión, de distribución, movilidad, etc., derivadas de los cambios tecnológicos relacionados con la información, las comunicaciones y la globalización.

Los datos son el centro de la nueva economía. Una definición práctica del término «datos» es la que recoge la Norma ISO/IEC 2382-1, que señala que un dato es la «representación reinterpretable de información de una manera formalizada, adecuada para la comunicación, la interpretación o el procesamiento». Dicho de otro modo, los datos son valores que reflejan situaciones o hechos que se producen y que permiten recrear o conocer una realidad. Se trata de una representación simbólica (numérica, alfabética, algorítmica, espacial, etc.) de un atributo o variable cuantitativa o cualitativa que describe hechos empíricos, sucesos y entidades. Un dato puede surgir de distintas fuentes y presentar diversas formas: por ejemplo, una imagen de una cámara de tráfico, una fotografía, un mensaje en las redes sociales, una voz en una conversación, la temperatura de una habitación, el grado de humedad de un cultivo, el número de unidades vendidas por una tienda o la hora exacta en la que se entregó un pedido.

Las imágenes, por ejemplo, están dotadas de metadatos, o información adicional sobre los datos, de contexto, tales como el tipo de extensión del archivo, el tamaño

y la fecha de creación, el dispositivo utilizado, la versión del *software* correspondiente, la ubicación (georreferenciación), la fecha y la hora, etc. Los metadatos son datos que describen otros datos, esto es, ayudan a ubicar datos.

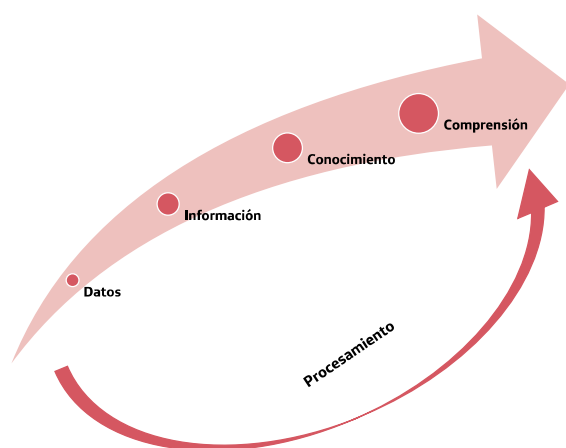
Los datos son el nuevo recurso considerado por muchos «la nueva gasolina» (Kroes, 2011; Rotella, 2012; Arthur, 2013), quienes reflejan su relevancia y la cada vez mayor dependencia de este recurso. También han sido definidos como un recurso infraestructural (OECD, 2014), ya que pueden ser utilizados por un número ilimitado de agentes para un infinito número de aplicaciones para producir bienes y servicios, considerados el equipamiento y estructura básica necesaria para que un país, una región o una organización funcione de forma adecuada (Merriam-Webster, 2004).

Sin embargo, un dato por sí mismo no proporciona valor, igual que la imagen de una cámara de tráfico, sin más contexto o tratamiento, presenta una utilidad muy limitada. Para obtener información de los datos y, por tanto, utilidad de los mismos, es necesario su procesamiento. Por esta razón, una de las metáforas más utilizadas es que los datos, más que gasolina, son el nuevo petróleo (Michael Palmer, Association of National Advertisers), por la necesidad de transformarlos (refinar) para ser convertidos en insumos de valor y utilidad.

Una vez que a los datos brutos (por ejemplo, un sonido) se les agrega información de contexto básica (metadatos), estos se convierten en información (por ejemplo, el sonido pertenece a una máquina) que, por sí misma, no proporciona ningún valor: no se conoce si el sonido es indicativo de una alarma, si se trata de un sonido normal emitido de forma constante por la máquina o si tiene algún otro significado. Por tanto, es necesario considerar un contexto mayor para que el dato pueda ser de alguna utilidad, para que la información se convierta en conocimiento. Por ejemplo, el volumen a partir del cual el sonido que produce la máquina indica un incorrecto funcionamiento de la misma y puede ser un síntoma de fallo de ciertas piezas.

**Fig. 1.2**

El ciclo del dato



Fuente: Afi.

En esta fase, el conocimiento puede ser útil por sí mismo, pues puede aportar información que previamente no era conocida. Sin embargo, en la mayor parte de los casos, es necesario tener una pregunta que responder. Generalmente, los análisis de datos se realizan para dar respuesta a determinadas preguntas: ¿cuál es la ruta óptima para llegar de un punto a otro?, ¿qué probabilidades hay de que llueva mañana?, ¿qué modificaciones en la cadena de producción pueden hacer los productos más rentables?, ¿qué clientes tienen más posibilidades de abandonar la compañía? En el ejemplo anterior, la pregunta podría ser la siguiente: ¿cuál es el sonido que indica que es necesaria una revisión antes de que se rompa la máquina?

El análisis de los datos con las distintas técnicas aplicables que detallaremos en los siguientes capítulos ayuda a dar respuesta a las distintas preguntas planteadas, lo que en sí mismo supone una aportación directa de valor al revelar la respuesta a preguntas que antes, por falta de información o de capacidad tecnológica, quedaban sin responder.

La nueva realidad en torno a los datos determina que mientras que un solo dato (un único sonido) puede ser observado, analizado, comprendido y utilizado, estas cualidades dejan de ser abordables cuando se analizan cientos, miles o millones de datos, a menos que se utilicen nuevas disciplinas y herramientas, entre las que cabe distinguir, fundamentalmente, estas dos:

- *Business intelligence* o inteligencia de negocio: permite explicar el pasado, crear cuadros de mando con visualizaciones de la información o informes de errores.
- *Data science* o ciencia de datos: se basa en las técnicas que, además de realizar un análisis descriptivo, posibilitan anticipar o predecir sucesos futuros (análisis predictivo) e incluso prescriptivos.

La ciencia de datos permite responder preguntas que contribuyen a conseguir mejoras de eficiencia (tiempos de transporte más reducidos, campañas de publicidad individualizadas, diseño de productos optimizado), ahorros de costes (reparaciones de maquinaria en el momento preciso, ofertas exclusivas para aquellos clientes con más riesgo de abandonar la compañía, ahorro de costes energéticos...) y otros muchos tipos de ganancias. Entre estos usos, los más destacados son la capacidad de análisis pormenorizado del cliente, la prevención de fraude, los mecanismos de recomendación, el análisis de redes sociales y el Internet de las Cosas.

Estas nuevas herramientas están siendo utilizadas por distintos agentes, desde empresas tradicionales que se transforman en *data-driven companies* (DDC, por sus siglas en inglés) a *startups* surgidas de forma nativa en la Economía Digital, hasta Administraciones públicas que mejoran sus servicios a partir del aprovechamiento del dato (para diseñar o mejorar sus planes de movilidad o para favorecer la detección del fraude, por poner algunos ejemplos). También existen usuarios particulares que se valen de información pública para crear y/o usar aplicaciones que

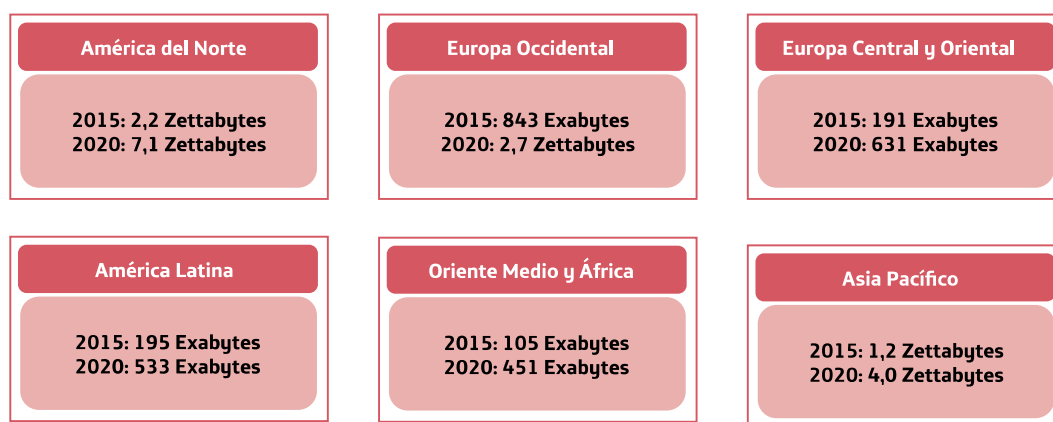
mejoran su día a día (localización de plazas de aparcamiento, análisis de tiempo de llegada de autobuses, servicios *peer-to-peer*, P2P, etc.). No obstante, la adopción de estas nuevas herramientas sustentadas en el dato no es homogénea ni entre sectores ni entre distintos territorios, de manera que sus efectos en algunos casos han quedado más patentes, mientras que en otros aún no han llegado a su máxima capacidad.

### 1.1.2 Evolución y situación actual desde una perspectiva económica

Una de las claves más destacadas que ha llevado al surgimiento de la Economía de los Datos ha sido precisamente el crecimiento en el volumen de este recurso. Según la Software & Information Industry Association (SIIA, 2013), en 2013 la información almacenada en el mundo alcanzó los 161 exabytes (EB =  $10^{18}$  bytes) por año —5 EB, en 2003—. En 2020 se prevé que el tráfico de datos desde y hacia los centros de datos supere los 15 zettabytes (ZB =  $10^{21}$  bytes).

**Fig. 1.3**

Tráfico global en centros de datos, por región (2015-2020)



Fuente: CISCO Global Cloud Index 2015-2020.

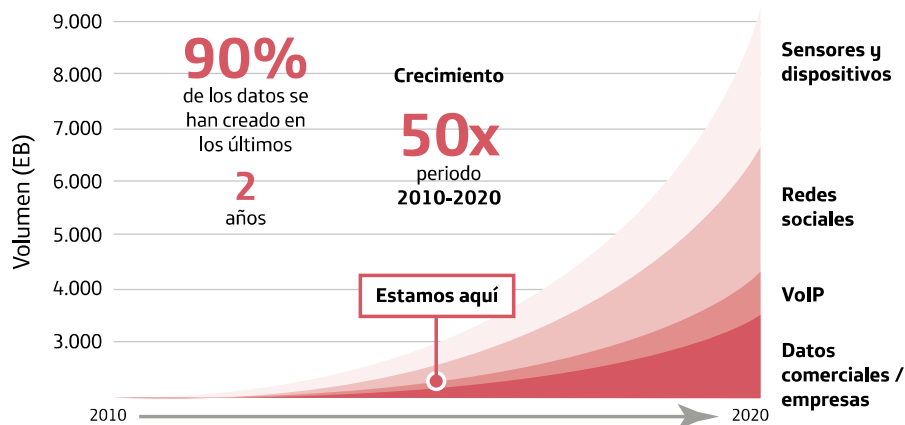
Nota: 1 ZB =  $10^3$  EB =  $10^6$  PB =  $10^9$  TB =  $10^{12}$  GB =  $10^{15}$  MB =  $10^{18}$  kB =  $10^{21}$  bytes.

Según José María Álvarez-Pallete,<sup>1</sup> presidente ejecutivo de Telefónica, «los recursos más valiosos en el futuro serán los datos»; hoy la mitad del tráfico en Internet es de origen no humano. Además, según Telefónica, el volumen de datos en los próximos años se multiplicará por once, la velocidad de transferencia de información será cien veces la disponible actualmente y la capacidad se multiplicará por cien.

1. [https://cincodias.elpais.com/cincodias/2017/06/19/companias/1497861452\\_912952.html](https://cincodias.elpais.com/cincodias/2017/06/19/companias/1497861452_912952.html)



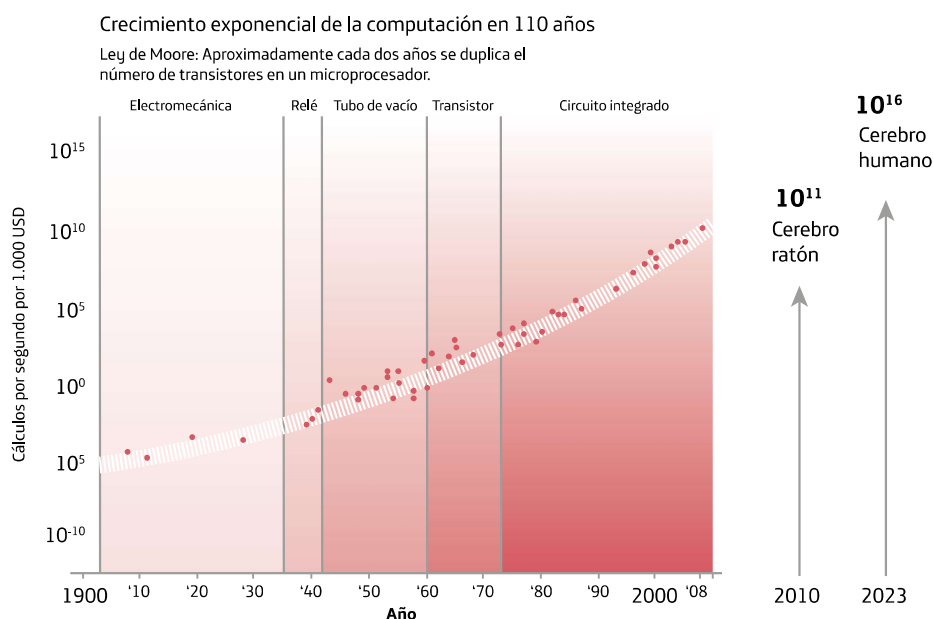
Fig. 1.4

Volumen de *big data*Fuente: (Familian 2016) *Working with Big Data*.

A pesar de ello, el incremento en el volumen de información no es el único factor relevante que ha permitido esta aceleración. La gran variedad de tipos de datos capturados y analizados, y la propia capacidad de procesarlos para obtener valor, son parte consustancial del elemento revolucionario que ha provocado este cambio global. El núcleo principal de estos avances ha estado en el abaratamiento de la computación, exponencial en los últimos años, con el que la expansión de la Economía de los Datos se ha reforzado.

Fig. 1.5

Evolución de la velocidad de cálculo



Fuente: Kurzweil 2005.

Este entorno favorecedor se ha visto además enriquecido por la convergencia de otros avances como el crecimiento de los objetos conectados (IoT) y en movilidad, las redes sociales, la computación en la nube (*cloud computing*) y el despliegue de herramientas analíticas. El futuro que se avecina con las capacidades esperadas de la computación cuántica supondrá con seguridad una revolución en el procesamiento de los datos y, por ende, en la economía con ellos generada.

El *cloud computing* o procesamiento en la nube permite utilizar capacidad de computación remota, masiva y escalable. Está basada en la servificación, es decir, la conversión de la capacidad de procesamiento computacional de un producto a un servicio contratable por una tarifa que posibilita hacer uso de unos recursos mientras están siendo utilizados. Este uso bajo demanda de la computación evita la necesidad de realizar una gran inversión en infraestructuras, plataformas y *software* y se encuentra a disposición del usuario de forma más directa. Gracias a este servicio, el análisis de datos deja de estar únicamente a disposición de grandes empresas con una gran capacidad de inversión, y pasa a estar al alcance de todos; se nivela el terreno de juego para empresas y Administraciones públicas de cualquier tamaño, al eliminar parte de las barreras de entrada.

Otro de los elementos clave ha sido la proliferación de los dispositivos conectados a Internet, entre los que se encuentran los teléfonos móviles y las tabletas como elementos principales, pero a los que se suman sensores, relojes inteligentes, coches o incluso hogares conectados. Todos estos dispositivos ofrecen nuevas oportunidades gracias a la conectividad prácticamente ininterrumpida, el acceso a la computación en la nube desde casi cualquier localización y la capacidad de acceso a cualquier fuente de información online o la conexión a cualquier dispositivo en cualquier momento. El elemento más relevante, si cabe, al que dan acceso estos dispositivos conectados es la creación de nuevos servicios personalizados basados en los datos recabados de forma autónoma, tales como recomendaciones de salud según las constantes vitales, alertas de retraso en medios de transporte según localización, ofertas de restaurantes o tiendas basadas en los gustos, o alertas de posible fraude ante la detección de situaciones anómalas. Servicios, en definitiva, que permiten ajustarse al máximo a las preferencias y los estilos de vida de los clientes o usuarios.

En relación con la ubicuidad de Internet, el tercer elemento que favorece el gran desarrollo actual es la interconectividad social. Las redes sociales son un fenómeno que ha visto una expansión global en los últimos años, no solo en número sino también en capacidad, incorporando componentes de otros programas, nuevos servicios y diversas plataformas, consiguiendo una conjunción de funciones que alcanzan desde la capacidad de compartir historias, imágenes o vídeos hasta capacidades relacionadas con el mundo laboral o la formación. El análisis de la información recabada en las redes sociales ha permitido un incremento en la comprensión de los individuos y una mejora en los servicios ofertados por distintos proveedores de forma personalizada, tanto dentro de la comunicación, como del mercado de trabajo y el entretenimiento, entre otros muchos.

Al mismo tiempo, las capacidades tecnológicas surgidas en relación con estos puntos han permitido la aparición de nuevos modelos de negocio basados en el entorno colaborativo, ya sea por la posibilidad del libre acceso y reutilización de datos (*open data*), como por las numerosas plataformas y aplicaciones habilitantes de nuevos modelos de negocio (*apificación*) sustentados en la compartición o intercambio de activos P2P (vivienda, vehículos, objetos, tiempo, etc.), la información a usuarios sobre el estado de la movilidad en las ciudades y carreteras (atascos, accidentes o retrasos en medios de transporte) o la puesta en común de ideas para el desarrollo de soluciones y/o iniciativas emprendedoras.

Por último, los avances en capacidad de computación, almacenamiento, sensores y tecnologías de la comunicación han sido el factor determinante en el avance del análisis de datos. Las empresas tienen ahora a su alcance programas estadísticos, de análisis predictivo, prescriptivo, de visualización y de aplicación para la toma de decisiones, independientemente del lugar donde se encuentren los datos, de su volumen, heterogeneidad, etc. Estas herramientas de análisis permiten a los agentes analizar y utilizar su información y la generada por terceros, ya sean datos estructurados o no estructurados, en tiempo real o estáticos, para innovar y mejorar los procesos, productos y servicios ofrecidos.

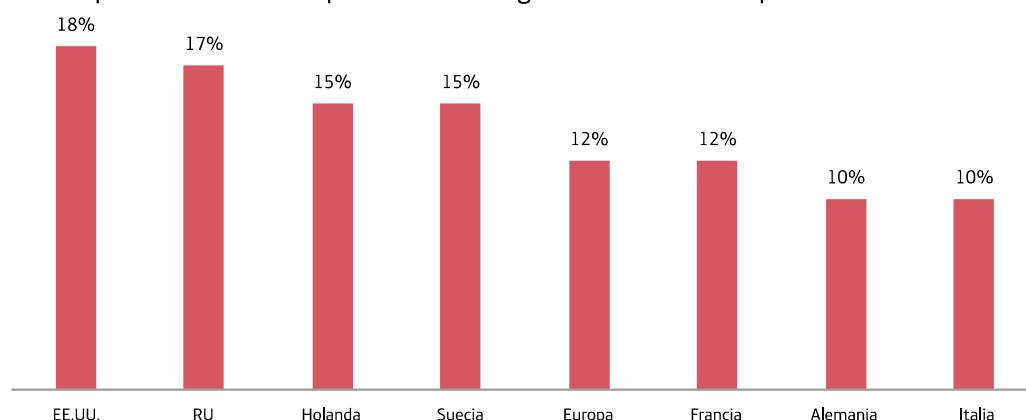
Aunque las tecnologías mencionadas son las más relevantes, hay muchas otras que se engloban en esta corriente y que han evolucionado a lo largo de los años. Prueba de esta vertiginosa evolución tecnológica es la diferencia existente entre las sucesivas curvas de expectativas tecnológicas elaboradas por Gartner en su *Hype Cycle Report* desde 2005.

Es aún difícil cuantificar el impacto económico de la Economía de los Datos, debido a su amplia diversificación, su gran alcance y su alta capacidad de transformación. Según la Comisión Europea, se estima que la Economía de los Datos en la Unión Europea ascendía a un valor de 272.000 millones de euros en 2015. McKinsey (Henke, et al. 2016), por su parte, estima que el PIB de la economía europea podría incrementarse en 2,5 billones en 2025 si alcanzara un mayor potencial, especialmente en las empresas más retrasadas en términos de digitalización. En el caso de Estados Unidos, McKinsey estima que el Internet de las Cosas, la analítica de datos y las plataformas de talento online podrían añadir 2,2 billones de dólares al PIB de la economía americana en 2025.

Sin embargo, a pesar de que el 90% de los datos digitales ha sido generado en los últimos dos años, solamente un 1% de esa información ha sido analizada. Se estima que Europa ha alcanzado apenas un 12% de su potencial, mientras que Estados Unidos se encuentra en el 18%.

**Fig. 1.6**

Grado de aprovechamiento del potencial de la digitalización en varios países



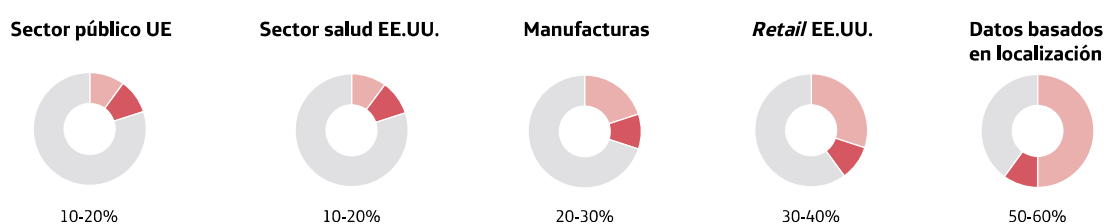
Fuente: McKinsey Global Institute Analysis.

Nota: Europa representa el promedio de los seis países mostrados, que suponen el 60% de la población, el 72% del PIB de la UE-28.

La mayor parte de las compañías captan aún hoy una mínima parte del valor potencial de sus datos. De entre los dominios existentes, el mayor progreso se ha producido en el área de servicios basados en localización y las ventas al por menor, gracias a los negocios nativos digitales que han surgido en ambos ámbitos. En el sector manufacturero, las Administraciones públicas y la salud, se ha conseguido capturar menos del 30% del valor potencial previsto hace cinco años.

**Fig. 1.7**

% del valor capturado (2016) respecto al valor previsto en 2011

Fuente: (Henke, et al. 2016) <http://www.mckinsey.com/business-functions/mckinsey-analytics/our-insights/the-age-of-analytics-competing-in-a-data-driven-world>.

