

**21 CLAVES PARA CREAR
VALOR A TRAVÉS DE LOS DATOS Y
LA INTELIGENCIA ARTIFICIAL**

A DATA-DRIVEN COMPANY

RICHARD BENJAMINS



ÍNDICE

Agradecimientos	13
Prólogo de Enrique Mota	15
Introducción	17
PARTE I: ORGANIZACIÓN	
1. Dónde debe situarse el <i>Chief Data Officer</i> o CDO en el organigrama	35
1. ¿Qué relación de proximidad debe tener el CDO con el CEO?	36
2. ¿Dónde situar el área de CDO?	37
3. Conclusión	39
2. Datos y TI, ¿cómo conviven?	47
1. Ventajas y desventajas de las alternativas para organizar los departamentos de datos y TI.....	48
2. Conclusión	52
3. Cómo relacionar la IA con el área de datos	55
1. Cómo organizar las áreas de datos, analítica e IA....	56
2. La relación con la madurez de los datos.....	57
3. ¿Científicos o ingenieros de datos?	58
4. Conclusión	59
4. Cómo medir la madurez de los datos	63
1. Dimensiones de la madurez de los datos	64
2. Cómo evaluar la madurez de los datos	69
3. Conclusión	69

5. Cómo monetizar los datos de forma externa.....	73
1. Una nueva propuesta de valor.....	74
2. Desafíos	75
3. Ejecutar una estrategia de monetización externa.....	76
4. Conclusión	83

PARTE II: NEGOCIOS Y FINANZAS

6. Cómo seleccionar casos de uso de IA y big data.....	93
1. Matriz de oportunidades.....	94
2. Encontrar casos de uso	96
7. Cómo medir el impacto económico	103
1. Reducción de costes de la infraestructura de TI de <i>big data</i>	104
2. Casos de uso de la analítica para optimizar el negocio	105
3. Generar nuevas fuentes de ingresos mediante la monetización de datos externos.....	108
4. Conclusión	109
8. Cómo financiar el proceso de transformación	113
1. Modelos alternativos para financiar las inversiones en datos	114
2. Invertir en datos e IA	118
3. Conclusión	119
9. Cómo utilizar los datos abiertos de forma efectiva	123
1. ¿Qué son los datos abiertos?	124
2. Ventajas del uso de datos abiertos	125
3. Desafíos del uso de datos abiertos	126
4. Posibles soluciones a los retos de los datos abiertos	127
5. Algunas pautas comenzar a usar los datos abiertos...	128
6. Publicación de datos abiertos.....	129
10. Ocho claves para crear valor en las pymes a través de los datos y la inteligencia artificial	133
1. IA y <i>big data</i> en pymes.....	133
2. ¿Cuáles son las claves para crear valor a través de los datos y la inteligencia artificial en las pymes?	135
3. Un breve manual para crear valor con inteligencia artificial y datos en las pyme.....	143

PARTE III: TECNOLOGÍA

11. Almacenamiento en la nube o local	149
1. Ventajas de la nube	150
2. Ventajas del <i>on-premise</i>	153
3. Diferentes estrategias en la nube	154
4. Conclusión	156
12. ¿Almacenamiento local o global? ¿Modelo de datos unificado o no?	161
1. Conceptos relevantes	161
2. Diferentes escenarios prácticos para el almacenamiento y el modelo de datos	162
3. Conclusión	171
13. ¿Dónde ejecutar la analítica de datos?	175
1. Conceptos relevantes	176
2. Decisiones importantes	177
3. El enfoque del centro de excelencia	177
4. El enfoque local	180
5. Escenario típico o ideal relacionado con la madurez de los datos	182
6. Conclusión	183
14. Estrategia de recogida de datos	187
1. Abundancia de datos	187
2. De las fuentes de datos tradicionales a las digitales	188
3. Por qué la recogida de datos no es tan fácil como se cree	190
4. ¿Qué debe abarcar una estrategia de recogida de datos?	197
15. Trabajar con proveedores y socios externos	199
1. Por qué las organizaciones trabajan con terceros para los datos, la analítica y la IA	199
2. Modos de colaboración con terceros	202
3. Experiencias concretas y posibles escollos	203
4. Conclusión	204

PARTE IV: PERSONAS

16. Convencer a los escépticos	213
1. ¿Qué motiva a los escépticos?	213
2. Estrategias para trabajar con éxito con escépticos	215
3. Conclusión	216

17. Democratización de los datos	219
1. Mayores equipos de datos para cubrir una mayor demanda de <i>insights</i> y analítica	220
2. Democratización de los datos.....	221
3. Herramientas de autoservicio de datos y analítica...	221
4. Conclusión	226
18. Cómo crear una dinámica positiva con los datos	231
1. El problema	232
2. Presentación de proyectos y resultados: cómo impactar y generar un enfoque positivo en la organización.....	233
3. Conclusión	235
PARTE V: RESPONSABILIDAD	
19. Los desafíos sociales y éticos de la inteligencia artificial y el <i>big data</i>	245
1. Desafíos éticos y sociales del uso de la inteligencia artificial y el <i>big data</i>	247
2. Cómo afrontar el impacto social y ético de la inteligencia artificial y el <i>big data</i>	251
3. Conclusión	252
20. De los principios de la IA al uso responsable de la misma	255
1. Selección de los principios adecuados para tu organización	259
2. Aplicación de los principios en la organización.....	261
21. Los datos como herramienta para el bien	269
1. Los Objetivos de Desarrollo Sostenible (ODS).....	269
2. <i>Big data</i> móvil contra la COVID-19	273
3. Consideraciones de las empresas para unirse al movimiento <i>Data for Good</i>	276
4. Cómo y por dónde empezar.....	278
5. Estado actual de los datos para el bien social y cómo avanzar.....	283
Notas	289
Bibliografía	293
Sobre el autor	303

PRÓLOGO

Yuval Noah Harari en su magnífico libro *21 lecciones para el siglo XXI* nos alertaba del gran impacto que la infotecnología y en especial la inteligencia artificial (IA) podrían tener en la política, la economía, las empresas, etc., y nos descubría el camino hacia una sociedad distópica que resulta casi inexorable.

En la semana en la que escribo este prólogo una de las noticias más comentadas en los foros tecnológicos ha sido la polémica sobre si la IA de Google llamada laMDA podría tener conciencia y, por tanto, capacidad de reflexionar sobre sí misma, pero lo inquietante de esta noticia no es si esta IA conversacional tiene o no conciencia hoy; el verdadero desafío consiste en que noticias como esta surgirán en los próximos meses y años hasta que realmente se convierta en una profecía autocumplida de la singularidad tecnológica, lo que pondrá sobre nuestras cabezas como humanos la responsabilidad de lidiar con las consecuencias de una distopía largamente baruntada por los expertos que dibuja una sociedad cada vez menos ficticia en la que los humanos multimillonarios, las corporaciones o los gobiernos podrían hacer un abusivo uso de ella y convertirla en un instrumento de alienación en vez de en un instrumento de progreso y de cambio positivo.

Richard Benjamins no solo nos acerca de una forma positiva al mundo de la IA, sino que además nos ofrece una nueva perspectiva

clarificadora sobre el futuro de los negocios basados en datos haciendo hincapié en el valor que pueden aportar a las empresas modernas y cómo podemos aprovechar todo su potencial y aplicarlo a mejorar los procesos y la toma de decisiones al tiempo que se contribuye a la transformación de los negocios y de las sociedades de forma más amable y justa.

Con un estilo claro y directo, *A Data-Driven Company* nos revela los beneficios y también los desafíos a los que tendremos que hacer frente con el desarrollo de la IA y cómo esta puede ser una palanca para crear nuevos servicios o negocios basados en datos.

Estoy seguro de que este libro contribuirá de forma significativa a inspirar a toda una generación de CXO responsables de liderar el devenir de las organizaciones *data-driven* que eviten una distopía que parece inevitable.



Enrique Mota
Chief Technology and Innovation Officer en Scalian



INTRODUCCIÓN

POR QUÉ ESCRIBÍ ESTE LIBRO

A lo largo de mi carrera, he dado cientos de charlas en eventos de distintos sectores. En los últimos cinco años, cada vez que he dado una presentación mostrando las 20 decisiones que las organizaciones deben afrontar en su camino a orientarse hacia los datos y la IA, he notado un fuerte reconocimiento por parte de la audiencia. Cuando hablo de «datos», pienso en la relación entre las áreas de datos y TI, a cuál función debe reportar el Departamento de Gestión de Datos (CDO), en las formas de medir el impacto económico, etc. Se ha escrito mucho sobre qué es el *big data*, sobre cómo funciona, sobre cómo aplicarlo a diferentes sectores, sobre los retos que conlleva, sobre qué significa para las empresas, etc. Sin embargo, casi no existe bibliografía que recoja la experiencia práctica de una transformación empresarial con base en los datos y la IA, con consejos concretos sobre las distintas opciones que permiten a las organizaciones aprovechar el potencial de estas herramientas, minimizando decisiones inadecuadas. Eso me hizo decidirme a escribir este libro.

Como científico cognitivo de formación, con especialización en inteligencia artificial y habiendo trabajado siempre en el campo de los sistemas de información inteligentes, el *big data* y la IA, he visto cómo muchas organizaciones han pasado de considerar los datos

como algo residual a verlos como un activo estratégico. He tenido el honor de formar parte de esta transformación en Telefónica durante más de diez años y también he sido testigo de estos cambios en AXA. En el camino, he participado en más de un centenar de eventos, talleres y conferencias sobre *big data*, analítica de datos e IA, lo que me ha servido para confirmar la necesidad de una orientación clara y práctica. He aprendido que alrededor del 70 % de la transformación hacia los datos y la IA es similar en todos los sectores de la industria, mientras que el 30 % es específico de cada sector. En este libro, trato de captar el 70 % que es común a cualquier organización que quiera orientarse hacia el dato y crear valor a partir de la IA.

¿POR QUÉ LAS ORGANIZACIONES QUIEREN CONVERTIRSE EN EMPRESAS DATA-DRIVEN IMPULSADAS POR LA IA?

Hace una década, un informe de McKinsey —«*Big data: La próxima frontera para la innovación, la competencia y la productividad*» (Manyika et al., 2011)— llamó la atención de los consejos de administración de las empresas sobre el *big data*. Desde entonces, numerosas organizaciones han iniciado su viaje de *big data*, como parte de una transformación digital más amplia, para convertirse en empresas *data-driven*: el *big data* ayudaría a las organizaciones a optimizar su negocio principal, a lanzar nuevos productos y servicios basados en los datos e incluso a crear negocios completamente nuevos. Hoy en día, prácticamente todos los procesos empresariales se ejecutan en un sistema informático y, por tanto, generan datos que pueden analizarse y utilizarse para optimizar estos procesos, que pueden combinarse con otras fuentes de datos para permitir una mayor optimización en toda la empresa, o pueden combinarse con fuentes de datos externas para ayudar a generar aún más oportunidades. Más recientemente, ocurrió algo similar con la IA, que ha eclipsado al *big data* en términos de valor potencial y de atención. De hecho, se considera que la IA es una de las tecnologías más prometedoras de la actualidad, y se la ha comparado con cambios revolucionarios como la electricidad e Internet. De hecho, muchos gobiernos han puesto

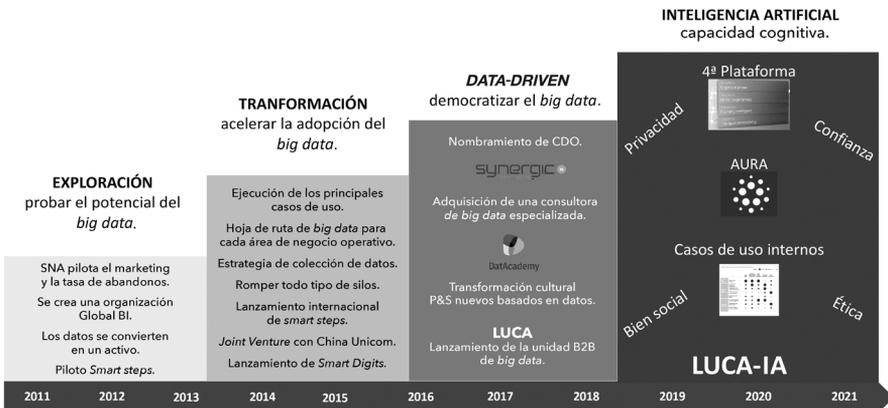
en marcha estrategias para maximizar sus beneficios, al tiempo que se aseguran de que los efectos secundarios potencialmente negativos sean comprendidos y tratados.

Sin embargo, convertirse en una empresa *data/IA-driven* no es una tarea fácil. Representa un largo proceso paralelo a la transformación digital por la que están pasando todas las grandes organizaciones. La figura 0.1 muestra el viaje de Telefónica hacia los datos y la IA, que ha sido largo y complejo, con muchos altibajos. De él se desprenden dos lecciones principales:

1. El viaje debe afrontarse por fases; es casi imposible saltar directamente al estado final deseado sin pasar por las fases anteriores. Sin embargo, esto no significa que las fases no puedan desarrollarse parcialmente en paralelo.
2. Para poder implementar la IA en todo su potencial, primero hay que convertirse en una empresa impulsada por los datos.

Según un artículo de la *Harvard Business Review* (Bean y Davenport, 2019), «el 77 % de los ejecutivos informan de que la adopción empresarial de iniciativas de *big data* e IA es un reto importante», pero menos del 10 % citan la tecnología como un desafío.

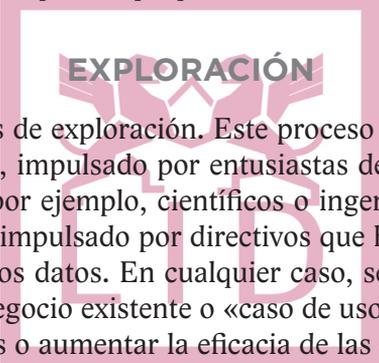
Figura 0.1 El viaje de los datos y la IA de Telefónica: exploración, transformación, *data-driven* y la IA.



EL VIAJE HACIA EL *BIG DATA* Y LA IA

El viaje hacia los datos y la IA desempeña un papel central en este libro porque la toma de decisiones importantes depende de la fase del viaje en que se encuentre la organización, es decir, de su madurez en datos e IA. El viaje pasa de explorar las oportunidades a consolidar la organización y los beneficios de los datos. A medida que una organización avanza, más datos pueden transformarse en valor. Y, en las últimas fases, se pueden aprovechar al máximo las técnicas de aprendizaje automático y la inteligencia artificial.

Para ilustrar este punto, hablaré del viaje de la transformación de los datos en Telefónica, tal y como se representa en la figura 0.1. Aunque el viaje de cada organización es único, las principales fases, los retos y las actividades por las que pasan son más o menos similares.



EXPLORACIÓN

La primera fase es de exploración. Este proceso puede iniciarse de abajo hacia arriba, impulsado por entusiastas de los datos con capacidad técnica (por ejemplo, científicos o ingenieros de datos), o de arriba a abajo, impulsado por directivos que han oído hablar de los beneficios de los datos. En cualquier caso, se suele seleccionar un problema de negocio existente o «caso de uso», como reducir la pérdida de clientes o aumentar la eficacia de las campañas de marketing. Entonces, se recogen los datos y se aplica la analítica para resolver el problema.

En Telefónica, esta fase de exploración comenzó en 2011 cuando se experimentó con el marketing viral. El objetivo principal del caso de uso era aumentar la eficacia de las campañas de marketing para la televisión de pago. Las campañas tradicionales se basaban en perfiles de clientes individuales, pero el equipo quería explorar el potencial de las campañas virales. Entonces, comenzamos por utilizar los registros de llamadas de las líneas fijas y extrajimos grafos sociales que reflejaban quién se comunicaba con quién. Después, identificamos los grupos que se caracterizaban por tener más comunicación y, por tanto, mayores vínculos (por lo general, grupos pequeños de entre cuatro y seis hogares). A continuación, nos fijamos en los grupos en los que al

menos dos hogares ya habían contratado el servicio de televisión de pago para aplicar la hipótesis de que estos hogares tendrían un efecto viral sobre el resto y, por tanto, aumentaría la posibilidad de compra del servicio en comparación con el cliente medio. Lo más innovador fue que hasta entonces considerábamos a los clientes de forma aislada; con esta nueva tecnología, tuvimos en cuenta sus relaciones sociales.

Los resultados de esta investigación fueron sorprendentes. Si se consideran todos los clientes incluidos en la campaña, no hubo realmente ningún efecto, ni positivo ni negativo. Sin embargo, al analizar los nuevos clientes, resultó que muchos pertenecían a un grupo que hasta entonces no se había considerado como segmento: los hogares con personas mayores (de más de 55 años). Dentro de este grupo hubo un claro efecto viral, lo que permitió diseñar una campaña específica que impulsara un alto índice de conversión.

Tras estos resultados tan prometedores que unían datos, analítica y negocio, Telefónica decidió explorar otras formas de aprovechar los datos a un nivel más global, en todas sus operaciones. Para ello, se creó una unidad global de inteligencia de negocio (BI) que permitiera obtener una visión consolidada de todas las operaciones y adaptar los procesos para implementar mejores prácticas y las lecciones aprendidas. Tuve el honor de dirigir la primera unidad global de BI de Telefónica en 2012.



TRANSFORMACIÓN

El objetivo de la siguiente fase, la transformación, es preparar a la organización para tratar los datos como un activo estratégico y crear valor de forma sistemática. Las actividades que se iniciaron en esta fase incluyeron la selección e implementación de casos de uso estratégicos que tenían un impacto significativo en el negocio. La selección de casos de uso se hizo a nivel global, utilizando una especie de «menú» con opciones, pero la implementación se decidió localmente, en cada unidad de negocio.

En esta fase, también iniciamos una hoja de ruta global de *big data* para transformar la práctica tradicional de BI (basada en *data-warehouses* de proveedores) en una arquitectura de *big data* más

abierta. Telefónica seleccionó la distribución abierta Hadoop de Hortonworks como arquitectura de referencia.

También se desarrolló una estrategia de colección de datos. El sector de las telecomunicaciones dispone de una gran cantidad de datos (CDR, datos de red, datos web, datos de aplicaciones, datos de llamadas a centros de atención al cliente, etc.), pero recopilarlos y almacenarlos no es una cuestión trivial. A menudo, estos datos se encuentran en sistemas de proveedores y no siempre está claro si los operadores de telecomunicaciones tienen acceso a ellos y en qué condiciones. Los proveedores son cada vez más conscientes del valor de los datos que almacenan y no siempre estaban dispuestos a facilitarnos el acceso a ellos. En este sentido, una lección importante que aprendimos fue añadir siempre una cláusula sobre acceso a los datos en los contratos de proveedores. Con estas limitaciones en mente, Telefónica empezó a construir una hoja de ruta de recopilación de datos, añadiendo progresivamente más cantidad y priorizándolos en función de los casos de uso.

Romper los silos es uno de los retos menos técnicos que conlleva el proceso de transformación, pero es crucial en este momento, puesto que las fuentes de datos siempre están asociadas a determinados procesos empresariales, y el propietario de cada función empresarial tradicionalmente tenía el control de los datos generados y podía decidir cómo extraer los indicadores clave de rendimiento (KPI) y con quién compartir subconjuntos de datos. Cuando quedó claro que los datos tenían valor, se convirtieron en una fuente de poder, por lo que compartirlos con otras áreas de la empresa llegó a percibirse como una pérdida de control y de poder. Con el tiempo, sin embargo, esta resistencia se ha debilitado y ahora todos ven el valor de compartir datos en beneficio de la organización.

Uno de los puntos de inflexión en la fase de transformación y en la orientación hacia el dato fue la exigencia de transparencia del presupuesto para *big data* en el plan estratégico anual de todas las unidades de negocio. Esto hizo visible algo que hasta entonces estaba «oculto» en los presupuestos de TI, entre otros. Por primera vez, quedó claro para todos —desde los empleados de base hasta los mandos intermedios y superiores— cuánto invertía cada unidad en sus datos y cuánto valor esperaba a cambio. De hecho, hubo muchas excepciones a la expectativa de que las unidades más grandes

invirtieran y obtuvieran más beneficios de los datos. Esta medida tan sencilla de transparencia facilitó el cambio de cultura necesario. Hasta entonces, los profesionales de los datos tenían que convencer a las unidades de negocio de que los utilizaran; y ahora estos pedían a los profesionales de los datos que los ayudaran. Cuando Telefónica comenzó su proceso, el Área de Gestión de Datos estaba a unos seis pasos de reporte del CEO. A finales de 2019, el director de datos o *chief data officer* (CDO) reportaba directamente al CEO. Con la creciente madurez de los datos, su importancia estratégica ha ido progresando, aunque lentamente, en la jerarquía corporativa.

DATA-DRIVEN

La fase del *data-driven journey* implica que muchas de las decisiones importantes de la empresa se tomen con base en los datos, es decir, que a la sabiduría convencional, la experiencia y la intuición se suma el conocimiento analítico con base en los datos de la empresa. Sin embargo, todavía hay importantes retos que superar.

Todas las empresas que se encuentran en esta fase cuentan con un CDO, o similar, que dirige un equipo de profesionales de datos, pero el punto clave es alcanzar la democratización de esta capacidad para que no solo se generen beneficios a partir de un grupo selecto de profesionales de datos, sino de todos los empleados de la empresa. Aumentar la creación de valor es más una cuestión cultural que tecnológica.

Este momento también sirve para comenzar a pensar en nuevas formas de obtener valor de los datos. Hasta esta fase, el valor se ha creado sobre todo internamente, para mejorar el negocio. Sin embargo, en algunos sectores, los *insights* revelados por los datos pueden implicar un valor significativo para otros sectores: en el caso de Telefónica, se trata de los datos que fluyen por la red móvil. Hoy en día, la mayoría de las personas tienen uno o dos teléfonos móviles, además de otros dispositivos, conectados a Internet y generando actividad en las antenas, lo que permite analizar una cantidad ingente de datos anonimizados y agregados que se obtienen de cada antena. La actividad en la red móvil genera datos de afluencia y de movilidad, de gran valor para sectores como el transporte, el turismo, la administración pública, el comercio minorista y las finanzas.

Los datos de la red móvil no solo tienen valor comercial, sino también social. En 2016, Telefónica creó un departamento llamado *Big Data for Social Good* (BD4SG) que utiliza los datos de dicha red móvil —combinados con datos abiertos y otros datos— para ayudar a alcanzar los Objetivos de Desarrollo Sostenible (ODS) de las Naciones Unidas. Este esfuerzo implica una estrecha colaboración con organizaciones humanitarias y otras ONG que trabajan en asuntos como la migración forzada, el coste humano de los desastres naturales, la pobreza infantil, el impacto de la movilidad en el cambio climático y las enfermedades contagiosas.

INTELIGENCIA ARTIFICIAL

Es en esta fase final donde se explota todo el valor de los datos, a través de la analítica, del aprendizaje automático y otras tecnologías de IA. En este momento, las empresas pueden reconsiderar su estrategia de datos original, tras años de experiencia y de aprendizaje, con el objetivo de alcanzar su máximo potencial. Esto puede estar relacionado con la revisión de decisiones tecnológicas (como el almacenamiento local, global o en la nube), decisiones organizativas o la dependencia de un equipo de datos centralizado frente a un equipo distribuido. Incluso puede extenderse a la creación de nuevas unidades de negocio, al tiempo que se reconsideran las responsabilidades locales frente a las globales. En Telefónica, esta fase comenzó con la aplicación de la IA para cambiar la forma de interactuar con los clientes, utilizando la llamada «computación cognitiva», que implicaba el uso de la tecnología de procesamiento del lenguaje natural (PLN) para entender automáticamente las intenciones de los clientes y luego conectarlas con registros individuales para responder a sus solicitudes.

En esta fase, Telefónica pasó a ser una empresa compuesta por cuatro capas o plataformas. La primera plataforma corresponde a la infraestructura física (la red, las antenas, la fibra, las tiendas, etc.). La segunda incluye los sistemas informáticos para operar el negocio (llamados OSS y BSS¹ en la industria de las telecomunicaciones). Y la tercera comprende los servicios digitales que se ejecutan sobre las otras dos anteriores. El punto clave es que las tres primeras plataformas generan enormes cantidades de datos que tradicionalmente se almacenaban (o no) localmente. Por esta razón costaba tanto

esfuerzo conseguir los datos apropiados para los casos de uso: estaban dispersos por toda la empresa, técnicamente en diferentes sistemas físicos, con diferentes formatos y proveedores, y eran propiedad de diferentes departamentos. Finalmente, la cuarta plataforma (ahora llamada Telefónica Kernel) recoge todos los datos de las otras plataformas en un formato interoperable y con una semántica clara. Es la base de todas las iniciativas de datos e IA de la empresa.

Sacar el máximo partido del *big data* y la IA también conlleva nuevos riesgos, especialmente aquellos relacionados con la privacidad y otras consecuencias no deseadas. Colocar los datos y la IA en el núcleo de una organización exige prestar más atención a la seguridad de los datos de los clientes, por motivos de cumplimiento legal, pero también para establecer y mantener su confianza. Muchos de los datos que utilizan las empresas se generan a partir de la actividad de los clientes y de su interacción con los servicios de la empresa, por lo que requieren de muchas obligaciones legales establecidas en el Reglamento General de Protección de Datos Europeo (GDPR), pero también se corre el riesgo de reducir la confianza de los clientes si estos datos no se utilizan de manera transparente. Además, el uso de la IA en toda la organización implica que las empresas se aseguren de que esta tecnología se utiliza siempre de forma responsable, es decir, evitando la discriminación, colocando a los humanos en el centro y abriendo los algoritmos *black box* o de caja negra cuando sea necesario.

Cada organización debe definir o experimentar su propio proceso para transformarse en una empresa *data/IA-driven*. Como he mencionado, nunca habrá dos procesos iguales, pero sí muchos temas comunes. Este libro recoge estos aspectos comunes para que cada organización pueda tomar decisiones informadas sobre los puntos clave a lo largo de su viaje.

¿DE QUÉ TRATA ESTE LIBRO?

Este libro dará a los lectores la capacidad de reconocer los retos a los que se enfrentarán durante los procesos de transformación digital y de orientación al dato y a la IA. Incluye claves para abordar estos retos, a través de decisiones con sus pros y sus contras. Del

mismo modo, analiza 21 decisiones clave a las que se enfrentará cualquier gran organización a lo largo de su viaje para convertirse en una empresa orientada a los datos y la IA. Las grandes organizaciones no siempre toman estas decisiones explícitamente, pero sí forman parte de sus procesos. Las lecciones aquí incluidas se basan en mi experiencia adquirida a lo largo de diez intensos años de actividad en dos grandes organizaciones y en conversaciones con muchas otras en más de cien eventos de diferentes sectores. Resulta sorprendente lo similares que son los retos en distintos sectores industriales, y estos puntos comunes son precisamente lo que recoge este libro.

Los capítulos incluyen al final una perspectiva externa de un profesional con experiencia en datos en diversas empresas y sectores. El objetivo ha sido enriquecer las lecciones aprendidas con experiencias del mundo real y conclusiones concretas.

Este libro no trata los aspectos técnicos de la ciencia de datos o la arquitectura de *big data*, ni tampoco de explicar la ciencia de datos o la IA a los empresarios. Más bien, es una guía para el viaje de la transformación digital, de la que los datos y la IA son ingredientes clave. Del mismo modo, esta no es una lista exhaustiva de lo que hay que tener en cuenta, sino un resumen de las lecciones aprendidas que destaca los puntos clave a los que se enfrentan las empresas.

¿QUIÉN DEBERÍA LEER ESTE LIBRO?

Este libro está dirigido a los profesionales de los datos que estén interesados en orientar su organización al dato y a la IA, y que quieran aprovechar al máximo las posibilidades que ofrecen estas potentes tecnologías. Aportará valor tanto a las organizaciones que desean empezar como a las que ya están en el proceso pero quieren pasar a la siguiente fase. Aquellos profesionales que pueden beneficiarse de este libro son:

- Directores de datos (CDO), de inteligencia artificial o de analítica de datos con más o menos experiencia, responsables de guiar a su organización en la transformación hacia los datos y la IA.

- Científicos de datos, ingenieros de IA y otros profesionales de datos con interés en progresar hacia un rol de CDO.
- Directores ejecutivos interesados en saber cómo orientar su organización al dato y a la IA (aspectos empresariales, inversiones, retos...).
- CFO que necesitan saber cómo financiar este proceso y qué beneficios esperar.
- Cualquier persona con interés en entender cómo las empresas avanzan (lentamente) en su transformación digital orientada a los datos y la IA.

¿QUÉ VAS A APRENDER?

Al final de este libro, entenderás:

- Las diferentes etapas del *data-driven journey* y sus características.
- Las decisiones organizativas, tecnológicas, de negocio, de personas y éticas que debe plantearse cualquier organización que quiera guiarse por los datos y por la IA.
- Las diferentes opciones disponibles a la hora de tomar una decisión para avanzar en esta transformación.
- Los pros y los contras de las distintas opciones disponibles y las mejores alternativas en determinadas circunstancias.
- Ejemplos prácticos de opciones disponibles y decisiones tomadas.
- La experiencia de veinte profesionales de datos que están impulsando la transformación de los datos en sus organizaciones.

¿CUÁL ES LA ESTRUCTURA DEL LIBRO?

El libro está estructurado en diferentes partes, cada una de las cuales agrupa las decisiones relevantes relacionadas con un aspecto específico del proceso examinándolo desde diferentes perspectivas.

La primera parte, «Organización», trata de las decisiones relacionadas con los aspectos organizativos de la orientación a datos y a IA. Esto incluye decisiones como:

- Lugar del CDO en el organigrama.
- Convivencia de las áreas de datos y TI.
- Posición de la IA en el área de datos o en otro lugar.
- Análisis de la madurez de los datos.
- Monetización externa de los datos.

La segunda parte, «Negocios y finanzas», analiza las principales decisiones de negocio relacionadas con los datos y la IA, como:

- Selección de casos de uso de IA y *big data*.
- Medida del impacto económico.
- Financiación del proceso.
- El papel de los datos abiertos.
- Implementación de IA y de *big data* en pequeñas y medianas empresas.

La tercera parte, «Tecnología», trata las decisiones tecnológicas a las que se enfrentarán las organizaciones, incluyendo cuestiones relacionadas con:

- Alojamiento en la nube frente a alojamiento local.
- Almacenamiento local o global de los datos y adopción de un modelo de datos unificado.
- Ejecución global o local de la analítica de datos.
- Estrategia de recogida de datos.
- Trabajo con proveedores y socios externos.

La cuarta parte, «Personas», aborda consideraciones relacionadas con las personas, como:

- Escepticismos.
- Democratización de los datos y ampliación de conocimientos en toda la empresa.
- Comunicación de los resultados y creación de impulso.

La quinta parte, «Responsabilidad», engloba las implicaciones sociales y éticas del uso de la IA y el *big data*, como:

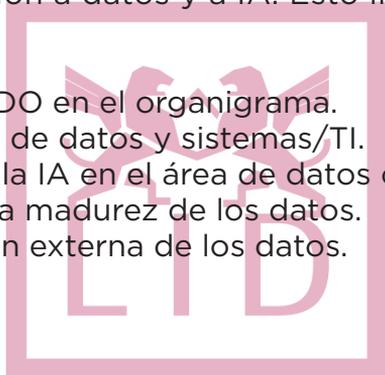
- Retos sociales y éticos de la IA y el *big data*.
- Importancia de los principios de la IA y del uso responsable de la misma.
- Uso del *big data* y la IA para el «bien común».



PARTE I
ORGANIZACIÓN

La primera parte, «Organización», trata de las decisiones relacionadas con los aspectos organizativos de la orientación a datos y a IA. Esto incluye decisiones como:

- Lugar del CDO en el organigrama.
- Convivencia de datos y sistemas/TI.
- Posición de la IA en el área de datos o en otro lugar.
- Análisis de la madurez de los datos.
- Monetización externa de los datos.



1

DÓNDE DEBE SITUARSE EL *CHIEF DATA OFFICER* O CDO EN EL ORGANIGRAMA

Ahora que el *big data* y la IA se han hecho un hueco en el mundo de los negocios, no es de extrañar que el *chief data officer* (CDO) se haya ganado un puesto en la mesa de la junta directiva. El primer CDO fue nombrado en 2002 por Capital One (Davenport y Randy, 2020). En los últimos años, un número cada vez mayor de organizaciones —sobre todo en el sector privado, pero también empresas de propiedad pública— consideran los datos como un activo estratégico e incorporaron un CDO. En 2021, una encuesta reveló que el 65 % de las empresas habían nombrado también un CDO en 2021, frente a solo el 55.6 % en 2018 (New Vantage Partners, 2021).

Hasta la aparición de esta nueva función, las iniciativas de inteligencia empresarial (BI) y de *big data* solían estar dispersas en las organizaciones, en departamentos aislados, aunque supuestamente

hubiera un equipo central de *Business Intelligence* (BI) que controlara la estrategia general de la empresa. Con el aumento de la importancia estratégica de los datos, muchas organizaciones han decidido nombrar a un CDO como responsable central de su estrategia y gestión de datos. Las preguntas típicas a las que se enfrentan las organizaciones a la hora de nombrar un CDO incluyen:

- ¿A qué distancia debe estar el CDO del CEO? ¿CEO-1¹, CEO-2 o CEO-n?
- Si es el CEO-1, es decir, si reporta directamente al CEO, ¿cómo se relaciona el CDO con los demás directivos de la empresa, en particular el CIO y el CTO?
- Si es CEO-n, ¿a qué ejecutivo debe reportar el CDO? ¿Debe ser al director de información (CIO), al director de operaciones (COO), al director de marketing (CMO), al director financiero (CFO), al director de transformación, al director de tecnología o al director digital?

1. ¿QUÉ RELACIÓN DE PROXIMIDAD DEBE TENER EL CDO CON EL CEO?

Para aprovechar todo el potencial de los datos, lo mejor es que el CDO se sitúe en un área cuya misión sea transversal y que represente una parte significativa del negocio. De este modo, la creación de valor no se limita a un área específica (por ejemplo, marketing) y se extiende a todo el negocio. Hacer lo contrario limitaría el valor a una sola área o a un área que no está directamente enfocada al negocio.

Por lo tanto, muchos sostienen que el mejor lugar para situar al CDO es reportando directamente al CEO (CEO-1) o bajo el COO (director de operaciones, CEO-2), lo que daría al ejecutivo una responsabilidad transversal. El hecho de que el CDO dependa directamente del CEO le otorga un puesto en el Comité Ejecutivo, lo que transmite un fuerte mensaje tanto interna como externamente. Hay dos directivos alternativos que garantizan la aplicación y la relevancia interorganizativa de la orientación al dato: el director de transformación y el director digital. Si bien estos dos cargos son de carácter

temporal (aunque a veces activos durante muchos años), trabajan de forma transversal y tienen la misión de adaptar el negocio al mundo digital, del que los datos son una parte fundamental.

Por supuesto, que el CDO dependa directamente del CEO no es necesariamente adecuado para todas las organizaciones. Requiere de un nivel alto de madurez de datos, y es probable que se reserve para organizaciones con más visión de futuro que aceptan la necesidad de adaptarse al mundo digital impulsados por los datos.

Entonces, ¿por qué algunas organizaciones se resisten a que el CDO reporte al CEO?

- Las empresas con menos experiencia en el uso de datos podrían querer situar al CDO bajo el CIO, dentro del área de TI. Esto ayuda a garantizar un conocimiento tecnológico antes de utilizar los datos para orientar el negocio.
- Algunas organizaciones tienen una idea muy clara de por dónde empezar a aplicar sus datos, por lo que sitúan al CDO bajo el departamento correspondiente. Por ejemplo, empresas de sectores como el de bienes de consumo rápido (FMCG, por sus siglas en inglés) —con un gran interés en mejorar su estrategia de marketing— podrían situar al CDO bajo el CMO. Otras más innovadoras podrían incluso situar al CDO bajo el CTO (I+D), mientras que las organizaciones que quieran ahorrar dinero podrían situarlo bajo el responsable de Recursos Globales.

2. ¿DÓNDE SITUAR EL ÁREA DE CDO?

En general, si el CDO se sitúa dentro de un área específica, esto implica que el ejecutivo herede algunos de los objetivos de esa área: si se encuentra en el área de Marketing, los objetivos estarán probablemente orientados a las ventas o a los ingresos; si se encuentra en Recursos Globales, estarán probablemente relacionados con el ahorro... Ayudar al resto del negocio es entonces un «esfuerzo extra», más que una responsabilidad central, y dependerá del ancho de banda del CDO y de su grado de colaboración. Es un verdadero reto ver este tipo de cooperación cuando excede el día a día impuesto por los KPI.

Por lo tanto, si una organización decide colocar al CDO bajo uno de los directivos que reporta al CEO, sin responsabilidad transversal, se crea una limitación innecesaria a la creación de valor a partir de los datos. Pero, entonces, ¿por qué la mayoría de los CDO no dependen directamente (ni están cerca) del CEO? El cuadro 1.1 enumera los pros y los contras que influyen en que una organización lo haga de esta manera:

Figura 1.1 Ventajas e inconvenientes de la posición de un CDO en el organigrama.

ÁREA	VENTAJAS	DESVENTAJAS
CMO	Marketing y Ventas proporcionan casos de uso con impacto directo.	Normalmente está centrado en B2C y no tanto en B2B, sin crear o capturar valor de otras áreas.
CFO	El libro de contabilidad requiere datos de alta calidad	Menos foco en el negocio. La gestión financiera no necesita <i>big data</i> .
CIO	Uso de la tecnología según los estándares de la empresa.	Menos foco en el negocio, más foco en la tecnología.
CTO	Acceso a las nuevas innovaciones tecnológicas.	Menos foco en el negocio, más foco en tecnología nueva.
CSO (seguridad)	Ventajoso para la seguridad y la privacidad de los datos.	Menos foco en el negocio.
CRO (recursos)	El ahorro de costes va directamente a la cuenta de resultados	Menos foco en el crecimiento del negocio y más foco en la eficiencia.

Las organizaciones que han decidido convertirse en *data-driven* como parte importante de su transformación digital suelen nombrar un CDO a nivel de CEO-2. En este punto, es importante reflexionar sobre el área de CDO y tomar una decisión basada en los consejos de este capítulo. Cuando la madurez de los datos de la organización aumente, el CDO podría convertirse en CEO-1, pero eso no suele ocurrir.

Por supuesto, que un CDO tenga éxito no solo depende de su lugar en el organigrama, pero este sigue siendo un factor importante. El consultor de gestión de datos estratégica Jay Zaidi (Zaidi, 2015) formuló otras consideraciones relevantes que pueden explorarse, como el patrocinio empresarial o la falta de claridad en el rol.

En Telefónica, la función de CDO se introdujo en el Comité Ejecutivo a finales de 2015 y desde entonces ha sido desempeñada por tres personas diferentes. Hace diez años, las áreas de datos estaban dispersas por la empresa, y los ejecutivos responsables se situaban entre el CEO-5 y el CEO-6. Hace siete años, pasó a ser -4, y hace seis años ascendió a CEO-2. En 2015, por fin se elevó a CEO-1 en el organigrama, mostrando lo fundamental que se habían vuelto los datos en la estrategia de transformación digital de Telefónica.

El debate sobre dónde situar al CDO y a qué nivel es relevante para aquellas organizaciones que están en camino de orientarse a los datos. Sin embargo, en aquellas empresas «nativas en datos», cuyo negocio son los datos, el CDO tiene requisitos muy diferentes. Un informe de Gartner sobre los cuatro tipos de organizaciones CDO (Gartner, 2016) señaló que el CDO es aún más crítico en empresas que negocian con datos, pues podría considerarse que, en estos casos, el CDO es realmente el CEO. Puede que no sepamos qué depara el futuro a las grandes empresas, pero sí sabemos que estará impulsado por los datos.

3. CONCLUSIÓN

El CDO es el director del viaje impulsado por los datos y la IA. Como tal, su función —similar a la de un director de transformación o director digital— es de naturaleza temporal. Cuando el proceso se completa, los datos tienen el estatus de un activo

estratégico y se gestionan en consecuencia. Mientras que una de las principales responsabilidades del CDO al inicio de su tarea es servir como agente de cambio, en las empresas impulsadas por los datos, es habitual la gestión de datos y la creación de valor a partir de ellos. Esto puede resultar en que el CDO tenga un perfil algo más bajo, pero no menos importante, en organizaciones maduras en términos de datos.

Aunque esto podría interpretarse con cierto dramatismo como «el ascenso y la caída de los *chief data officers*», en realidad es la forma lógica de afrontar cambios organizativos disruptivos. Los CDO no desaparecerán, pero la naturaleza del cargo es intrínsecamente evolutiva. Al principio, pueden ser responsables de todo lo relacionado con los datos y la correspondiente creación de valor, mientras que, en empresas desarrolladas e impulsadas por sus datos, la función del CEO se centrará en la gestión de los datos para mantener su estatus de activo estratégico. La creación de valor a partir de ese activo se asigna ahora a los departamentos apropiados, lo que permite extraer todo su potencial.



ERRORES COMUNES EN EL CAMINO PARA CONVERTIRSE EN *DATA-DRIVEN*

En los últimos diez años he tenido la suerte de participar en la transformación digital de una empresa líder en tecnología en una compañía *data-driven*. Durante este tiempo hemos asumido proyectos e iniciativas de todo tipo, la mayoría exitosos, que sin duda nos han permitido acercarnos al objetivo de la conversión de los datos como principal fuente de toma de decisiones y generación de valor. Sin embargo, en el camino se han cometido inevitablemente errores que han ralentizado un poco este proceso. Me gustaría destacar los que considero más importantes, que pueden servir de aprendizaje para otras iniciativas similares en el futuro:

1. **Estar cegado por la tecnología.** Las nuevas tecnologías, como *big data* e IA, son muy atractivas, sobre todo para los más tecnológicos. Cuando se inicia un proyecto basado en datos, inevitablemente ellos se sienten atraídos por este y lo primero en lo que piensan es en la plataforma de *big data* o en el algoritmo de IA que hay que utilizar: cuanto más sofisticado, mejor. Sin embargo, esto es como construir una casa empezando por el tejado. Para comenzar, resulta fundamental tener claros los problemas del negocio que hay que resolver, identificar los datos relevantes a los que hay que recurrir para buscar la solución y asegurarse de que son de suficiente calidad.

Una vez que los datos están en orden, se puede empezar a analizarlos para resolver un problema empresarial. Ten en cuenta que la última tecnología no siempre resulta esencial para lograrlo; a veces basta una simple hoja de cálculo Excel o una base de datos SQL para empezar. Más adelante, cuando la cantidad de datos es mucho mayor y los problemas más complejos, las tecnologías de *big data* e IA ayudarán y serán de vital importancia a la hora de sacar el máximo partido a los datos.

2. **No proteger tus datos en los servicios prestados por plataformas externas.** En la actualidad cada vez son más los servicios digitales prestados por plataformas externas alojadas en la nube (herramientas de comercio electrónico, servicios financieros, logística, etc.). Un error recurrente a la hora de ejecutar acuerdos comerciales con estos proveedores es olvidar incluir cláusulas que aseguren que los datos generados por los servicios pertenecen a la empresa cliente y que el proveedor los proporcionará a través de la API o de mecanismos similares para que puedan usarse internamente. Sin este acuerdo contractual, puede ser muy difícil acceder posteriormente a los datos y tener permiso para emplearlos. Es importante no dejar abiertas a la interpretación las cuestiones relativas a la propiedad de los datos.

- 3. Concentrar la generación de conocimiento en un solo departamento.** También es un error común designar a los centros de experiencia (p. ej., en funciones de BI o *big data*) como los únicos responsables de analizar los datos de la empresa y obtener *insights* de negocio. Estas áreas suelen tener expertos en analítica de datos, pero sin un contexto empresarial resulta imposible obtener conocimientos relevantes. La generación de *insights* tiene que ser un proceso totalmente transversal a la compañía, es decir, debe cruzarse y entrelazarse con las distintas líneas de negocio.

Cualquier área (de negocio, RR. HH., marketing de producto, etc.) debe tener las herramientas y el conocimiento para acceder a los datos, analizarlos y llegar a conclusiones que permitan tomar las mejores decisiones en cada momento. Obviamente, los expertos en BI y *big data* deben apoyar la iniciativa con su mayor conocimiento de la tecnología de analítica de datos y algoritmos, pero debe ser un esfuerzo conjunto.

Para ello, tiene sentido posicionar a los expertos en *big data* e IA dentro de las áreas de negocio de una empresa para facilitar este trabajo colaborativo y ayudar a cambiar la cultura organizativa.

- 4. Usar excesivamente algoritmos *black box*.** Otro error frecuente es confiar ciegamente en los resultados de los algoritmos *black box* o de caja negra. sistemas tan complejos que su funcionamiento interno no puede ser observado ni interpretado fácilmente, ni siquiera por los científicos de datos que los diseñan. Gracias a las matemáticas, ofrecen una forma rápida de detectar a los clientes que potencialmente van a abandonar la marca (*churners*) o a los que tienen propensión a comprar un producto o servicio. Sin embargo, en muchas ocasiones es fundamental no solo conocer el resultado del algoritmo, sino también su funcionamiento. En el caso de los

churners, hay que saber no solamente quiénes son los más propensos a llevarse su negocio a otra parte, sino también los motivos. Con un mayor *insight* de ese razonamiento, la organización adquiere más confianza a la hora de tomar decisiones basadas en los datos y puede tomar decisiones más concretas sobre cómo actuar según el resultado del algoritmo. En el caso del cliente con probabilidad de irse a otra marca, se puede saber qué hacer para retenerlo. Por ejemplo, en el ámbito de las telecomunicaciones, es importante saber si se debe a que tienen un móvil anticuado, un plan de servicio que no se ajusta a sus necesidades o **problemas técnicos con la línea**.

Convertir un negocio en una operación *data-driven* es un viaje largo y complejo, especialmente para las grandes empresas. Resulta inevitable cometer errores, pero lo importante es aprender de ellos. Por supuesto, es mucho mejor y más eficaz conocer los errores cometidos por otros antes que cometerlos uno mismo.

PEDRO A. ALONSO BAIGORRI

Global Product Lead en TelefónicaTech
IoT & BigData

EL ROL DEL CDO EN LA ORGANIZACIÓN

Uno de los grandes retos para cualquier compañía que quiera verdaderamente convertirse en una empresa *data-driven* es entender cómo debe adaptar la organización para facilitar esta nueva transformación. Y es que, como habitualmente se dice, nuevas ideas requieren siempre nuevas organizaciones.

Desde mi punto de vista, el primer paso que debe dar una compañía es crear una figura que lidere esta transformación en la organización; una figura cuya misión sea facilitar y acelerar la creación de una cultura *data-driven* en la que los datos generan valor para el negocio y se convierten en una de las principales ventajas competitivas: el *Chief Data Officer*.

La visión de este profesional como garante de la cultura *data-driven* y con foco en la transformación y la obtención de impacto económico a través de los datos y la inteligencia artificial dista mucho de las primeras aproximaciones. Los primeros *chief data officers* se centraron principalmente en los aspectos más regulatorios de gobierno y uso *compliance* de los datos con una visión muy limitada de sus funciones y, en general, con poca aportación de valor para el negocio.

Sin embargo, desde mi punto de vista el CDO debe desarrollar una serie muy amplia de funciones por lo que necesita también desarrollar una organización específica con diferentes equipos:

1. **Dar soporte a la ideación y *delivery* de casos analíticos.** El equipo de *data translators* ayuda a los negocios a identificar oportunidades en las que los datos puedan generar valor para la compañía. Este equipo debe facilitar no solo la ideación si no también el desarrollo de los casos de uso analíticos, ayudando a los *product owners* (que deben ser del negocio) en las decisiones relacionadas con la gestión de los datos y el desarrollo de modelos analíticos avanzados, garantizando que las soluciones desarrolladas sean viables y escalables para el negocio
2. **Facilitar el acceso a los datos.** El equipo de *data engineers* y *data architects* deben permitir extraer los datos de los sistemas transaccionales y almacenarlos en *datalakes* y *lakehouses* en la nube que

faciliten su disponibilidad a toda la organización a través de herramientas de autoservicio. Es por tanto su responsabilidad la creación de una plataforma corporativa que garantice la reutilización de las arquitecturas de datos para realmente poder escalar de manera eficiente a toda la organización y a todos los empleados.

3. **Orquestar el gobierno de los datos.** El equipo de *data managers* debe desarrollar las herramientas y políticas que faciliten ejercer el gobierno de los datos a toda la organización y orquestar este proceso. Para convertirse en una compañía *data-driven* es clave contar con un dato de calidad y garantizar el uso correcto de los datos desde el punto de vista regulatorio interno y externo. En los comienzos se atribuía al *Chief Data Officer* la responsabilidad de gobernar los datos de la compañía, hoy en día sin embargo existe una visión de que el gobierno de los datos debe ocurrir de una manera federada en toda la organización, y los negocios tienen que contar con un rol protagonista en ejercer esta actividad.
4. **Desarrollo, evolución y gobierno de los modelos de ciencia de datos.** El equipo de ciencia de datos, independientemente de la decisión de desarrollar interna o externamente estos modelos, deben ser capaces de hacer *ownership* de los algoritmos, velando porque todos los desarrollos tengan la calidad adecuada y conociendo su funcionamiento para poder monitorizar su *performance*, mantenerlos y aplicar una mejora continua. Además debe velar por el uso ético y sostenible de la inteligencia artificial, facilitando la definición de principios y la aplicación de herramientas desarrolladas para tal propósito.

5. **Impulsar una cultura *data-driven* en toda la organización.** El equipo de cultura debe ser responsable de crear una comunidad del dato que facilite la adopción de estas tecnologías en toda la organización reduciendo las barreras del cambio. Para ello debe fomentar la realización de actividades de comunicación interna y externa, y definir e implementar planes formativos que garanticen la capacitación analítica de toda la organización en base a sus necesidades.

La mayor o menor centralización de estas funciones dependerá de la madurez de la compañía, que podrá evolucionar de un modelo más centralizado a un modelo de *hub & spoke* o incluso descentralizado.

Pero independientemente de este último punto para mí es clave que la organización identifique al *Chief Data Officer* como la figura que debe guiar el viaje de transformación hacia una compañía en la que los datos y la inteligencia artificial sean clave del éxito futuro de la empresa.

JUAN JOSÉ CASADO

Chief Data Officer en Repsol y Academic Director
& Professor Master Big Data & Business Analytics en IE