

Framework Data & IA Industrial
Marco de referencia
de datos e inteligencia
artificial para el sector
industrial español

INDESIA



01

INTRODUCCIÓN

¿Qué es la IA?

¿Por qué es importante la IA en la industria?

¿Por qué es necesario un marco para Data & IA específico para las empresas industriales?

02

ÁMBITO DEL VALOR PARA NEGOCIO

Reto 1.

Identificación del valor y los casos de uso

Reto 2.

Medición del Impacto

03

ÁMBITO DE LA CIENCIA DE DATOS

Reto 3.

Desarrollo de modelos analíticos más complejos

Reto: 4.

Puesta en producción de los modelos analíticos

04

ÁMBITO DE LOS DATOS

Reto 5.

Accesibilidad a los datos

Reto 6.

Calidad y seguridad de los datos

05

ÁMBITO DE LA ORGANIZACIÓN

Reto 7.

Modelo organizativo

Reto: 8.

Modelo de Sourcing y Financiación

06

ÁMBITO DE PERSONAS Y CULTURA

Reto 9.

Conocimiento y capacidades

Reto 10.

Adopción y gestión del cambio

07

ÁMBITO DE LA SOSTENIBILIDAD

Reto 11.

Ética de la Inteligencia Artificial

Reto 12.

Algoritmos verdes

08

CONCLUSIONES



01

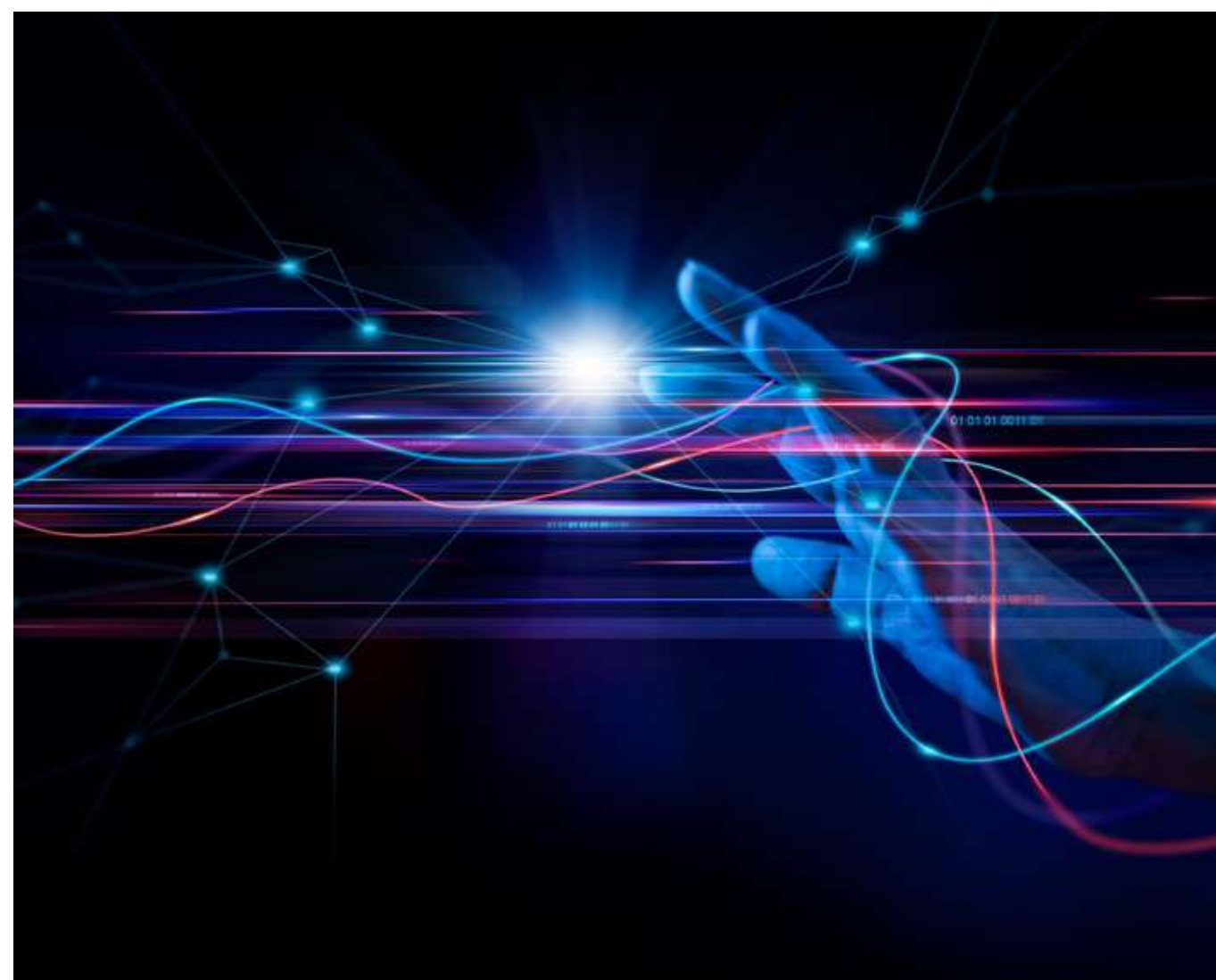
INTRODUCCIÓN

1.1. ¿Qué es la IA?

La IA son las siglas de Inteligencia artificial, que no es otra cosa que la ciencia de hacer máquinas inteligentes. Esto lo definió así, y acuñó este término, John McCarthy en 1956. Más tarde fue Marvin Minsky el que lo definió como “la ciencia de hacer que las máquinas hagan cosas que requerirían inteligencia si las hubiera hecho un humano”.

Esta segunda definición es la que originó un debate, que a día de hoy sigue abierto, en el que se pretende averiguar cómo reconocer si una máquina es inteligente o no. Para esto existen dos aproximaciones: evaluar el resultado o evaluar el proceso. En este sentido, la prueba de Turing (Turing test) ha tenido un papel importante en la historia de la IA: permite decidir si una máquina puede pensar o no, basado en una evaluación del resultado. Un interlocutor hace preguntas y, o una máquina, o una persona le responde. El interlocutor no puede ver quién responde. La máquina pasa la prueba de Turing en el momento en que el interlocutor no es capaz de distinguir si las respuestas vienen de la máquina o de la otra persona.

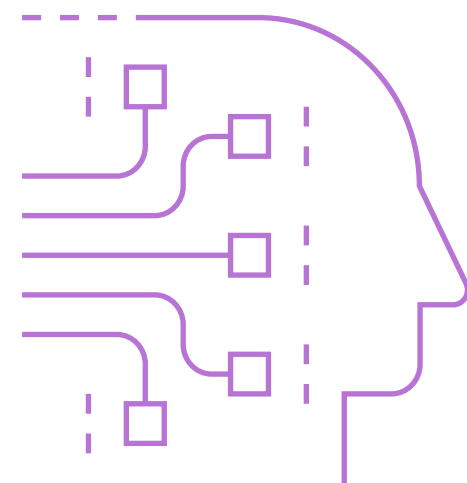
Lo que está claro es que la IA está demostrando ser una tecnología transformadora con impacto importante en la economía, la sociedad y la vida. Cada vez existen más aplicaciones para la inteligencia artificial (IA). Desde el diagnóstico médico hasta los vehículos autónomos pasando por la traducción automática y el reconocimiento de voz. O, por ejemplo, la optimización de procesos, la eficiencia energética, el mantenimiento predictivo, la creación de nuevos productos o la mejora de la relación con los clientes.



Pero, ¿cómo es un software de IA?

Se trata de un *software*, un programa informático, que procesa datos y genera decisiones basándose en los datos analizados. Lo que más se está usando en estos momentos para su desarrollo es el llamado *machine learning*, aprendizaje automático. En este caso, al contrario del cerebro humano, la Inteligencia Artificial necesita grandes cantidades de datos para realizar el análisis (el aprendizaje) que le lleva a sus conclusiones.

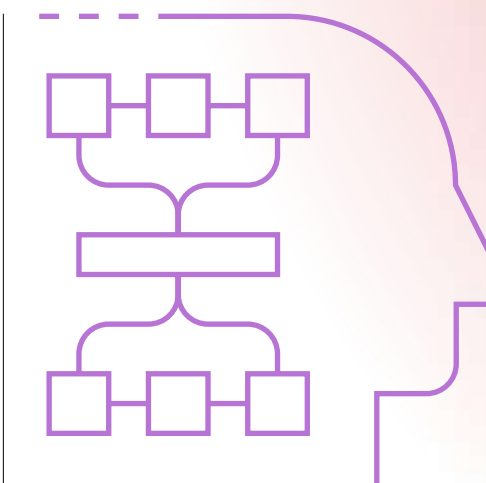
Existen tres tipos principales de aprendizaje automático: supervisado, no supervisado y el llamado *reinforcement learning*. Vamos a bucear, a través de un ejemplo, en cada uno de estos conceptos para quedarnos con la idea general necesaria para entender el proceso.



Aprendizaje automático supervisado:

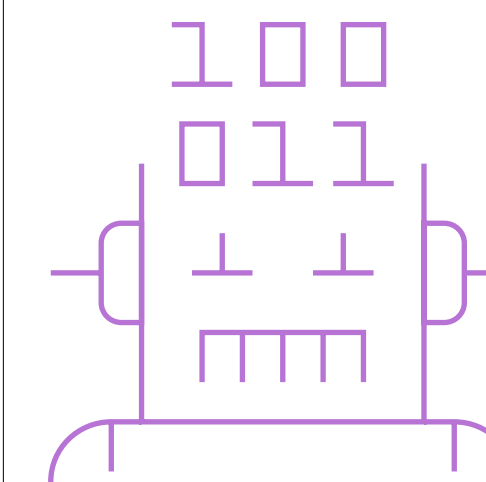
Imaginemos que queremos entrenar un sistema de IA para reconocer gatos en fotos. En este caso, le enseñamos miles de fotos con gatos y otras que no los muestran. De esta manera, el sistema aprende qué características de las fotos corresponden a gatos y cuáles no. Así, cuando hay una nueva foto, el sistema es capaz de decir si aparece un gato o no. Estos sistemas no siempre suelen acertar. Cometan errores que se conocen como “falsos positivos” al reconocer gatos donde no los hay, y “falsos negativos” al no reconocer un gato donde sí lo hay.

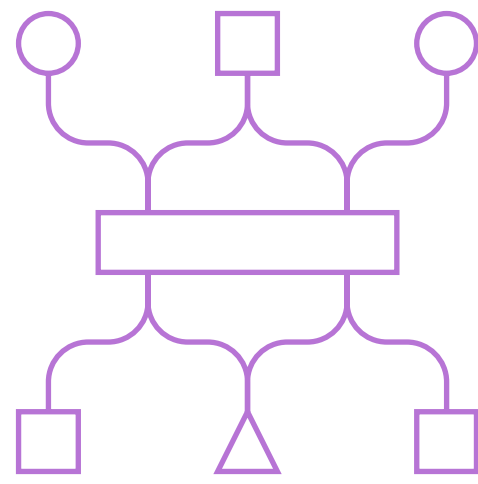
Aprendizaje automático no supervisado: Pongamos, por ejemplo, que se presenta al sistema de Inteligencia Artificial la base de datos de clientes de una empresa. El algoritmo, por sí mismo, se encarga de agrupar los datos en grupos que maximizan las características comunes y minimizan las características distintas. De esta manera, se puede segmentar automáticamente una base de datos con muchos clientes en grupos interesantes.



Reinforcement learning:

El uso de algoritmos de IA para jugar al ajedrez o al juego del Go es uno de los mejores ejemplos. En este caso, el sistema aprende por prueba y error. Recibe una recompensa o penalización por cada movimiento que propone según se acerca o se aleja del objetivo de ganar. Aprende a base de jugar muchas partidas para quedarse con los mejores movimientos.





1.2. ¿Por qué es importante la IA en la industria?

El concepto de Industria 4.0 representa en sí mismo la transformación digital aplicada al ámbito industrial. Si bien debemos considerarlo como un fenómeno de evolución, son las personas el principal agente del cambio.

Los expertos, que aportan su experiencia en los casos de uso, se apoyan en nuevas herramientas y tecnologías que son la principal palanca de cambio en este proceso. Según los expertos, tecnologías como *big data*, *Cloud Computing* o el *Internet of Things* son habilitadores de nuevos enfoques y plataformas a la hora de obtener respuestas novedosas a todo tipo de preguntas, tanto las viejas como las nuevas.

Pero, sin duda, la Inteligencia Artificial, presente desde hace décadas tanto en investigación como en algunos proyectos, es uno de los principales agentes disruptivos en la completa transformación de la industria.

La aplicación de la IA a casos industriales cambia totalmente el enfoque de los posibles retornos o resultados a la hora de usar las nuevas plataformas. Esto se debe a que los operarios de las industrias pueden obtener retorno donde el proceso automatizado hace una predicción sobre un KPI en clave industrial o una recomendación en cuanto a ciertos ajustes del proceso.

Para entender mejor este punto, planteamos algunos ejemplos recopilados del conocimiento y experiencia de grandes multinacionales españolas en los que la Inteligencia Artificial es el eje central de su estrategia digital:

- Optimización de procesos logísticos en todas las fases de la cadena de producción.
- Recomendador en los flujos de mantenimiento, preventivo y prescriptivo.
- Sistemas de visión artificial para procesos de inspección de calidad automáticos.
- Proceso de planificación automática industrial.
- Sistemas de monitorización de procesos de producción basados en Inteligencia Artificial para la detección de anomalías y toma de acciones.
- Sistemas de IA para control de seguridad del empleado en el puesto de trabajo.
- Asistentes virtuales en diversos casos de uso del empleado.
- Sistema de eficiencia energética con Inteligencia Artificial para la optimización de consumos.

1.3 ¿Por qué es necesario un marco para Data & IA específico para las empresas industriales?

Pese a la promesa de valor que supone la Inteligencia Artificial para las empresas industriales, el grado de adopción de estas tecnologías es todavía, a día de hoy, relativamente bajo.

La realidad es que, pese a que las empresas industriales disponen de una gran cantidad de datos producidos por sus sistemas de producción y cuentan con perfiles ingenieros habituados a trabajar con datos, los sectores que han sido pioneros en el desarrollo de soluciones de IA han sido los más ligados al sector servicios, como la banca, el *retail*, las telecomunicaciones o los grandes *players* digitales. Estas industrias han adoptado el uso de Inteligencia Artificial muy centradas en mejorar la relación con sus clientes explotando todas las posibilidades que el *big*

data y la IA brindan para conocer mejor a los clientes, personalizar su experiencia y maximizar su impacto comercial.

Y ¿por qué las empresas industriales no lo usan? Porque no han popularizado de la misma manera el uso de algoritmos para la optimización de sus procesos. Y eso que el impacto en la cuenta de resultados puede ser aún más directo y de mayor volumen que el obtenido de la aplicación de la IA a la sistemática comercial en las empresas de servicios.

¿Por qué se da esta circunstancia?

Hemos analizado las principales barreras que se encuentran las empresas industriales para escalar el uso de la Inteligencia Artificial, y los hemos agrupado en diferentes ámbitos y retos. Para cada reto se ha descrito cuál es el principal problema que declaran tener los ejecutivos que trabajan en empresas industriales para ser capaces de escalar la IA. Se han descrito

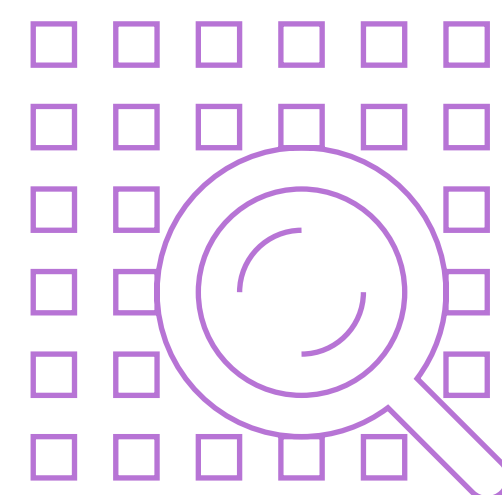


las diferentes soluciones que pueden ayudar a resolver cada reto apoyándonos en las mejores prácticas que hemos observado en empresas del sector.

También, se ha querido definir un modelo de madurez que facilite a cada empresa saber en qué nivel está dentro de la solución de cada uno de los retos y cuál es el siguiente paso que puede dar para seguir avanzando en el escalado de la Inteligencia Artificial en su organización.

Todo lo anterior se puede resumir en el siguiente Marco de referencia de datos e inteligencia artificial para el sector industrial español.

A partir de ahora abordaremos los distintos ámbitos en los que se puede adoptar la Inteligencia Artificial en la industria. Son un total de 6 los principales a destacar, con 12 retos que se señalan para ser capaces de escalar el uso de la IA en las organizaciones industriales.



2.1. Reto 1. Identificación del valor y los casos de uso

2.1.1. Descripción del problema

Uno de los grandes problemas que existen es el grado de desconocimiento de los beneficios que la Inteligencia Artificial puede aportar, y más al sector industrial. Es cierto que esto varía según el estamento de la empresa de la que hablemos.

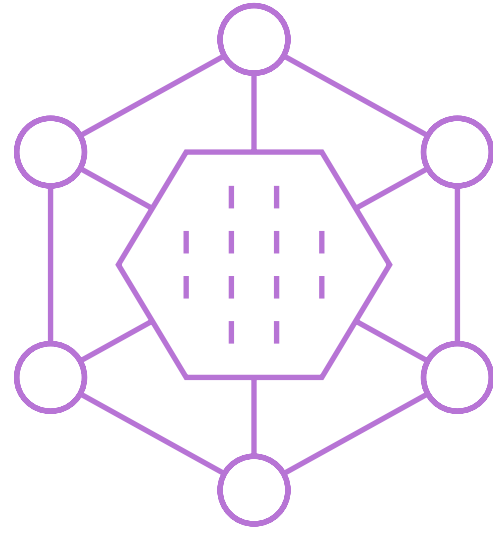
El 42% de los ejecutivos piensa que “comprender los beneficios de la IA para el negocio y los posibles casos de uso” es el principal desafío de la adopción de Inteligencia Artificial. Aquí ya se denota que no solo es conocer qué hace la IA, también es importante ver la aplicación y el uso que se puede hacer de ella. Y es que la mayoría de empleados de las empresas industriales y sus managers desconocen con un grado profundo qué es realmente la Inteligencia Artificial y cómo puede ayudarles en su día a día.

Algo muy habitual es que se confunda el término

IA con conceptos como la robótica. Por lo que se olvida que también se puede asociar a la ciencia de los datos o el *big data*.

Por tanto, además del desconocimiento de los posibles casos de uso que pueden implementarse hay una percepción que el uso de estas tecnologías es demasiado complejo y costoso en términos de recursos y de tiempo, como para que merezca la pena explorar sus posibilidades.

Aquellas empresas suficientemente valientes para plantearse el reto de descubrir estas tecnologías, carecen, con frecuencia, de un método estructurado para ayudarles en la búsqueda de situaciones que les generen valor y en su correcta priorización. En demasiadas ocasiones se acaban realizando casos de uso que resultan muy vistosos o que se pueden contar bien a la prensa pero que apenas generan impacto económico, perdiendo la oportunidad de demostrar el valor de estas tecnologías.



2.1.2. Soluciones

2.1.2.1. Sesiones de ideación

Es fundamental desarrollar de manera estructurada sesiones en las que participen personas de distintos niveles de la compañía para identificar un número suficiente de casos de uso de la IA. De esta manera se puede dar prioridad a unos u otros en función de su viabilidad y su impacto.

Lo ideal es que estas sesiones comiencen con una explicación detallada de qué es la Inteligencia Artificial y sus diferentes dominios, aplicaciones y capacidades. Después se debería realizar una sesión de análisis en la que se compartan los casos de uso que están implementando otras empresas del mismo sector o sectores adyacentes. Finalmente, deben identificarse, por parte de los participantes, los problemas reales de su negocio que si se resolvieran tendrían un impacto económico significativo; fijándose en aquellos que puedan resolverse analizando datos disponibles, y así imaginar qué tendría que darle un modelo de IA para resolverlo.



2.1.2.2. Definición de los casos de uso de Inteligencia Artificial aplicables al negocio

Un caso de uso es la descripción de una actividad de negocio en los cuales puede aplicarse una tecnología para optimizar el proceso u obtener un mejor resultado. Es decir, es como realizar una hipótesis del uso de la IA con un ejemplo concreto.

En la descripción de los casos de uso de Inteligencia Artificial es importante detallar cuál es el valor que se va a obtener con el modelo analítico, qué datos se van a necesitar para desarrollarlo, las variables que pueden predecirse, qué tipo de algoritmos van a ser necesarios utilizar, el resultado que esperemos que produzcan y cómo van a ser utilizados esos resultados para optimizar el proceso. Lo más importante de todo es saber cómo se va a evaluar el impacto en el negocio.

Todo esto que se acaba de explicar se puede resumir, de una manera muy visual, en un canvas como el propuesto por Louis Dourard en su publicación *The Machine Learning Canvas*.

Expliquemos un poco este gráfico que ilustra tan bien cómo utilizar los casos de uso de Inteligencia Artificial.

El Machine Learning Canvas comienza con un bloque central dedicado a la propuesta de valor del sistema donde se va a utilizar; es decir, qué es lo que vamos a hacer, por qué es importante, cómo lo haremos, por qué, y quién será el encargado de usar el sistema o se verá afectado por él.

Luego está el cómo, que se puede dividir en dos partes: aprender hacer predicciones. La parte del lado izquierdo está dedicada a las predicciones, basadas en los modelos que aprenderemos de los datos. Hacer predicciones permite aprender de los resultados.

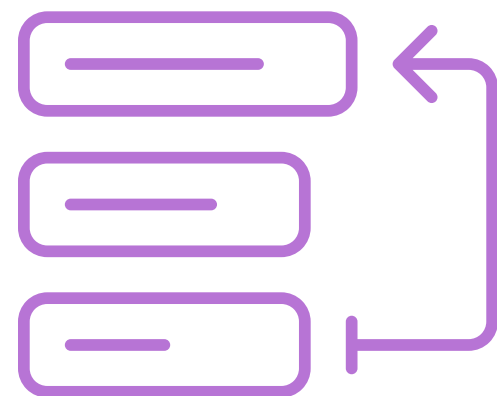
Composición del *Machine Learning Canvas*

Resumiendo, la definición de un caso de uso estaría compuesto por los siguientes bloques de información, teniendo en cuenta que en el lado izquierdo están:

- **Tarea de *Machine Learning Canvas*.** ¿Qué tipo (por ejemplo, clasificación, regresión...), cuál es la entrada y cuál es la salida para predecir (junto con los valores posibles)?
- **Decisiones.** ¿Cómo se utilizan las predicciones para tomar decisiones que proporcionen el valor propuesto al usuario final?

- **Predicciones.** ¿Cuándo hacemos predicciones sobre nuevas entradas y cuánto tiempo tenemos para eso?
- **Evaluación.** ¿Qué métodos y métricas podemos usar para evaluar la forma en que se realizarán y utilizarán las predicciones antes de la implementación?

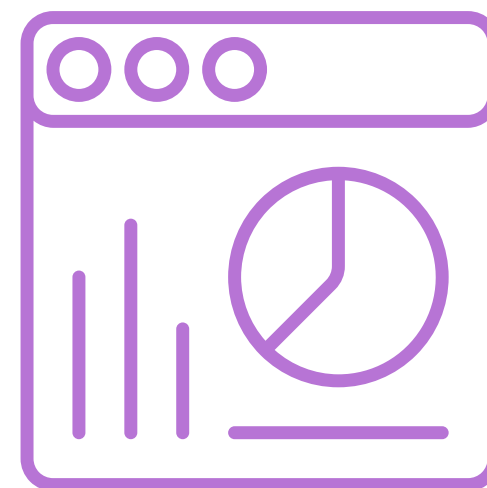




2.1.2.3. Priorización de los casos de uso

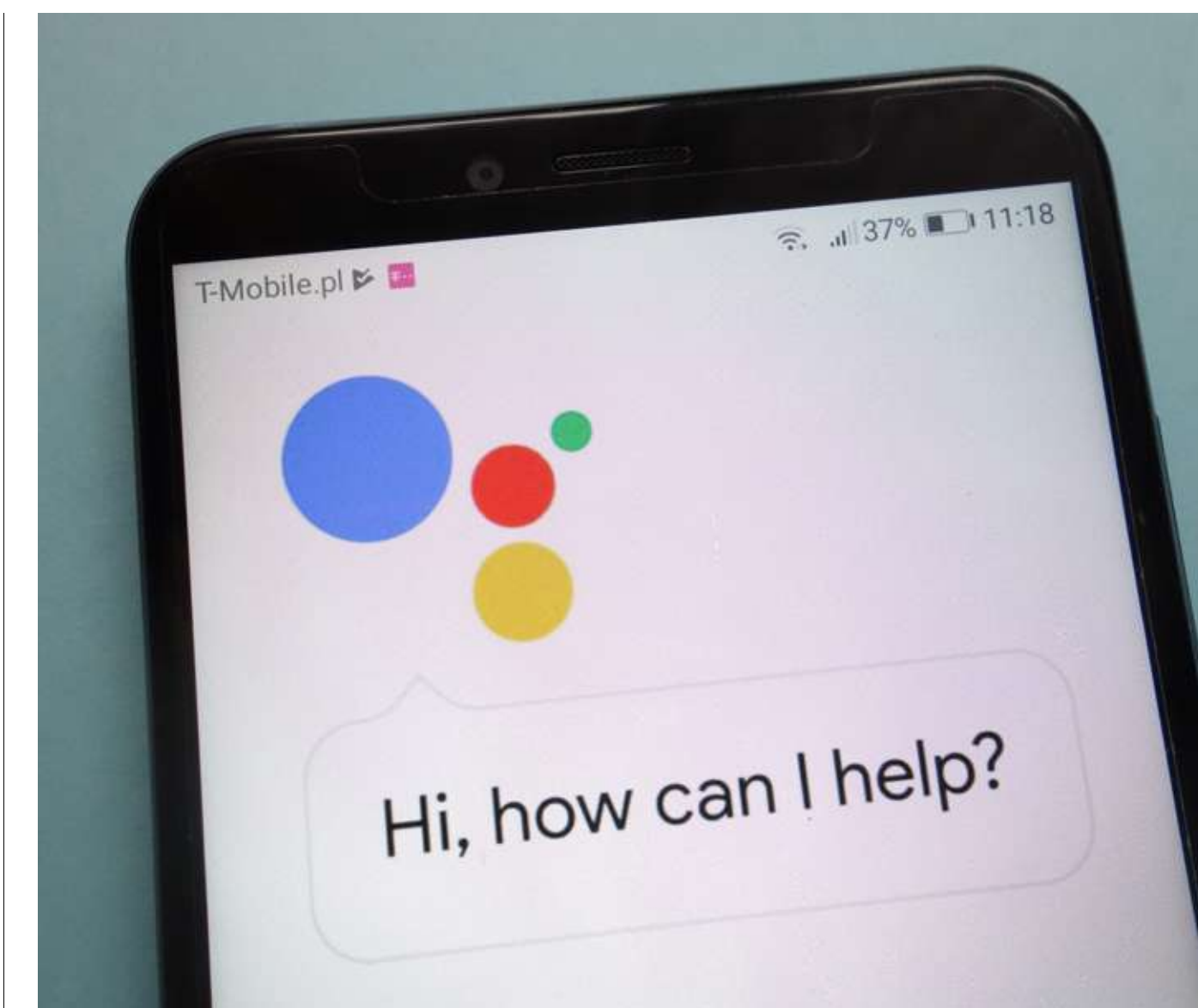
Es importante crear un sistema que nos ayude a priorizar los casos de uso con el objetivo de lograr el mayor impacto posible en el negocio, lo antes posible, para demostrar a la organización el valor que pueden llegar a aportar estas tecnologías y conseguir cierta “tranquilidad” organizativa.

Debe conseguirse un portafolio de casos equilibrado entre la necesidad de obtener resultados en el corto plazo y lograr el mayor impacto. Los casos identificados como *bottom up* (casos principales de uso) tienden a ser excesivamente operativos, y aunque dan resultados en el corto plazo y pueden considerarse *low hanging fruit*. Se corre el riesgo de dedicar excesivos esfuerzos a iniciativas de bajo impacto, por el contrario, los casos identificados *top down* tienden a ser estratégicos y de alto impacto, pero tienen un alto riesgo de no ser factibles o necesitar bastante tiempo para su implementación. Por eso es conveniente definir una agenda equilibrada con casos dando resultados en el corto y en el medio plazo desde lo más táctico a lo más estratégico.



2.1.2.4. Establecimiento de un programa estructurado de IA

Una vez desarrolladas estas sesiones de ideación y priorización de los posibles casos de uso a desarrollar, la organización debe estructurar el lanzamiento de estas iniciativas en un programa único que disponga de una agenda visible por toda la organización que permita hacer seguimientos de los avances conseguidos, y de los impactos obtenidos por cada una de las iniciativas priorizadas. Este programa deberá estar respaldado por toda la organización, incluidos los altos cargos, por el carácter transformador que va a tener en la organización y el seguimiento de sus avances deberá estar en la agenda del Comité Ejecutivo de la organización para reforzar la importancia de estas iniciativas.



2.1.3. Niveles de madurez

2.1.3.1. Nivel 1. Fundacional. Experimentos desestructurados

Se busca entender cómo la Inteligencia Artificial puede crear valor de negocio mediante pruebas de casos sueltos desarrollados por *early adopters*. Es decir, cuando las pruebas de concepto no han sido pensadas en un escenario real productivo sino como para ser prueba de la tecnología.

Lo importante de esta fase es que hay mucho interés por saber qué hacen los demás en la industria para buscar retorno rápido mediante casos sueltos no relacionados entre sí y no siempre asociados a un valor de negocio claro.

2.1.3.2. Nivel 2. Acercamiento. Pilotos con foco en comunicación (Innovación de rueda de prensa)

En este nivel se identifican algunos ejemplos de casos de uso que ayudan a comprender cómo la IA puede generar valor “en mi organización”. Los casos iniciales generalmente se enfocan en la reducción de costes, automatización de procesos o personalización de recomendaciones, donde ya hay asistentes habilitados para voz o escenarios similares donde la Inteligencia Artificial ha demostrado valor y retornos en otras empresas o sectores.

El principal valor de retorno de estos casos suele ser a nivel de marketing y concienciación dentro de la organización.

2.1.3.3. Nivel 3. Aspiracional. Agenda de IA Bottom UP (Quick wins-casos operativos)

El valor de negocio comienza a provenir no solo de un enfoque de reducción de costes, sino también de la identificación de nuevas formas de crear y capturar valor, flujos de ingresos y modelos comerciales.

Existe una agenda de IA a largo plazo que se define por los negocios más innovadores dentro de la compañía y que se empieza a priorizar según el retorno de las iniciativas.

2.1.3.4. Nivel 4. Maduro. Agenda de Transformación del Negocio IA Top Down (Big Bets -casos disruptivos-)

En este nivel, la organización define y ofrece modelos de negocio nuevos o disruptivos que aprovechan la IA a nivel interdepartamental y con apoyo explícito desde los principales ejecutivos de la compañía. La Inteligencia Artificial está totalmente incorporada en la estrategia de la organización y el retorno de los casos de IA forma parte la PNL de cada negocio como algo intrínseco a la estrategia y no como casos o escenarios aislados.

2.1.4.

CASOS DE ÉXITO



Programa Digital con más de 150 iniciativas de Data & Analytics lanzadas en 3 años

En el anterior Plan Estratégico de Repsol (2016-2020), así como en el actual (2021-2025), Repsol identificó la digitalización como una de las principales palancas para alcanzar los objetivos marcados. Con este motivo, en 2017 se lanzó el Programa de Digitalización de Repsol, para acelerar la transformación digital de la compañía, lanzando cientos de iniciativas digitales en todos sus negocios, que son los verdaderos protagonistas de todo ello.

El Plan Estratégico recoge, dos objetivos transversales para el Programa digital:

- La contribución positiva en materia de sostenibilidad, ayudando a Repsol a conseguir su objetivo de ser Net Zero en emisiones para 2050.
- Promover la transformación de formas de trabajo de Repsol y el desarrollo de sus empleados en capacidades digitales.

El Programa también ha promovido las capacidades tecnológicas en los empleados de Repsol, extendiéndose a toda la organización e involucrando a más de 1200 personas en 4 continentes y más de 20 países.

Además, el Programa cuenta actualmente con más de 150 iniciativas con impacto directo en reducción de emisiones de CO2 o mejora de la eficiencia energética.

Dentro del Plan Estratégico 2021-2025, todos los negocios de Repsol han identificado retos y necesidades tecnológicas y de digitalización. Dichos retos y necesidades se han condensado dentro de nueve tendencias tecnológicas en las que se van a centrar en los próximos años:

- Data & Analytics.
- Experiencia de los usuarios y personalización.
- Comunicaciones abiertas.
- Automatización de operaciones.
- Multinube.
- Arquitecturas de última generación.
- Ciberseguridad mejorada.
- Blockchain.
- HPC y Quantum.

CASOS DE ÉXITO

Para promover la adopción de estas tecnologías y asegurar que desde la empresa se posee el talento que les permita mantenerse a la vanguardia en cada una de ellas, han creado 9 HUBs tecnológicos integrados a lo largo de toda la compañía:

- Analítica de Datos.
- Experiencia de Cliente (UX) y diseño.
- Blockchain.
- Omnicanal.
- Metodología Ágil.
- Robótica de software.
- Centro de Competencias Cloud (CCC).
- Ciberseguridad.
- Robótica del hardware.

Los HUBs son grupos de expertos internos y externos especializados en una tecnología determinada, que, dentro

del Programa Digital, proporcionan recursos y apoyo en el desarrollo de iniciativas digitales y desarrollan plataformas transversales para promover el uso de dichas tecnologías. Adicionalmente, los HUBs trabajan mano a mano con las distintas áreas de negocio de Repsol para identificar nuevas oportunidades de digitalización y así capturar las últimas tendencias en digitalización del sector.

Un ejemplo de plataforma creada por un HUB es ARiA. Es la plataforma del dato de Repsol que centraliza la información necesaria para el funcionamiento de los casos digitales y la generada por estos, favorece su accesibilidad, permitiendo aprovechar el valor de los datos de forma transversal en toda la organización. Además, permite la industrialización del desarrollo de casos de analítica de datos e inteligencia artificial.

CASOS DE ÉXITO

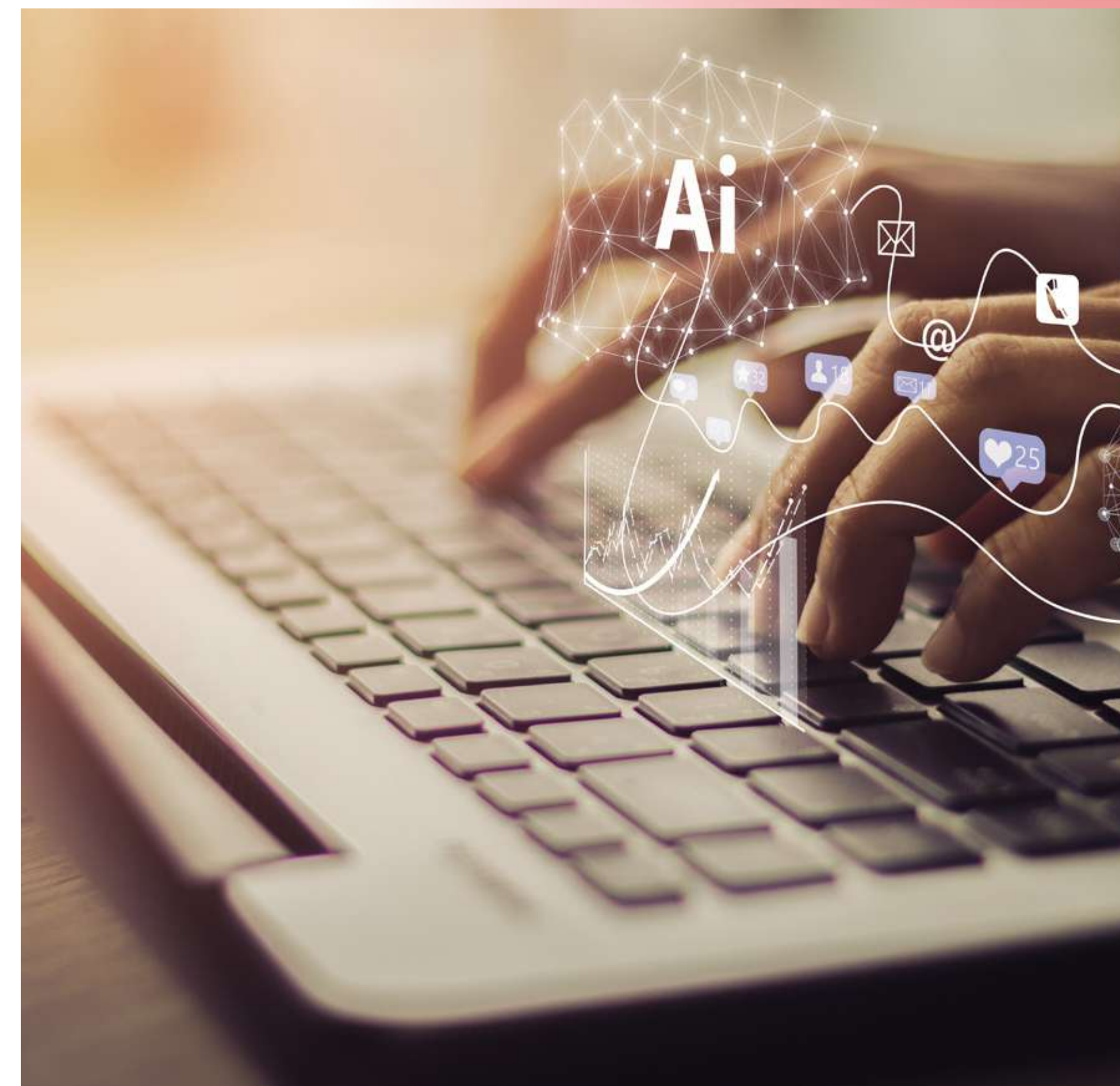


Matriz de oportunidades: impacto / viabilidad. Cuantificar la oportunidad

Muchas organizaciones que empiezan a trabajar con *big data* e Inteligencia Artificial se hacen la pregunta: ¿por dónde empezar?

En general, desde el punto de vista de Telefónica, hay dos formas:

- Comenzar a desarrollar las capacidades necesarias para usar IA y *big data* (infraestructura, datos, habilidades, etc.).
- Comenzar con casos de uso que muestren el valor potencial para la organización.



La mayoría de las organizaciones eligen la segunda opción ya que es más fácil invertir en capacidades una vez que se tiene una comprensión más clara del valor que se puede generar.

Pero, ¿cómo elegir el mejor caso de uso para empezar?

Bajo la experiencia de Telefónica se vio que una buena manera de atacar este problema es mediante la creación de una matriz de oportunidades (también llamada Matriz de Ansoff) con las oportunidades específicas de datos e IA para la organización.

2.1.4.

CASOS DE ÉXITO

Por lo general, los dos ejes principales representan, por un lado, el valor o el impacto comercial, y, por otro lado, la viabilidad. El valor es importante porque demostrar un caso de uso de importancia lateral para el negocio no convence a la organización para invertir. La viabilidad es importante porque los resultados no deben llegar en años, sino en meses: la paciencia de las empresas por los resultados de cosas nuevas es limitada.

A veces es difícil estimar el valor comercial de un caso de uso antes de ejecutarlo. Una buena manera de estimar el valor comercial de un caso de uso es multiplicar el volumen de negocio o coste por el porcentaje estimado de optimización. Por ejemplo, si la tasa de cancelación de una empresa es del 1% (mensual) y hay unos 10 millones de clientes, con unos

ingresos medios mensuales de 10€, entonces el volumen de negocio asciende a 1M€ al mes o 12M€ al año. Si la inteligencia artificial pudiera reducir la tasa de cancelación en un 25%, es decir, del 1% al 0,75%, entonces el valor estimado sería de 250.000 € al mes.

La otra dimensión importante a estimar es la viabilidad de los casos de uso. Esta es una estimación más cualitativa que puede ser diferente según la organización. Básicamente, estima qué tan fácil o difícil es ejecutar el caso de uso, incluyendo factores como la disponibilidad de los datos (ubicación, propiedad, coste), la calidad de los datos, la colaboración con el área de negocios (algunos son campeones, otros son defensivas), el riesgo de privacidad, etc.

2.1.4.

CASOS DE ÉXITO

ferroviario

Identificación de oportunidades. Matriz palancas de valor vs viabilidad

Desde Ferrovial se está trabajando también en la implantación de la Inteligencia Artificial para poder mejorar procesos y la experiencia de usuario.

A la hora de identificar casos de uso donde la Inteligencia Artificial puede aportar un valor diferencial al negocio, las sesiones de ideación o *design thinking* pueden ser una gran herramienta porque permiten hacer un descubrimiento *top down* de distintos casos de uso, partiendo de los objetivos de negocio. Para que estas sesiones tengan éxito es importante tener en cuenta varios factores clave:



- Deben ser sesiones guiadas para facilitar la participación de todas las partes.
- Se debe elegir un conjunto de usuarios para las sesiones que tenga un pensamiento estratégico o que sepa identificar claramente cuáles son los objetivos que se persiguen desde el punto de vista de negocio.
- Es necesario que parte del grupo tenga conocimientos de datos e IA para poder complementar la visión más pura de negocio con una visión que permita proponer soluciones que puedan contribuir a la consecución de esos objetivos estratégicos.

2.1.4.

CASOS DE ÉXITO

Pongamos un ejemplo para ilustrarlo mejor:

- **Objetivo de negocio:** nos gustaría ahorrar energía en la producción de agua en plantas.
- **Posible oportunidad:** crear un algoritmo de IA que pueda predecir la demanda real de agua para ajustar así los parámetros de producción al máximo.
- **Posible impacto:** disminuir el excedente de producción en un 2%, implicaría disminuir el consumo energético en x% y por lo tanto se produciría un ahorro en costes estimado de Y€/año.

Si las sesiones de ideación son fructíferas, al final se obtiene de las mismas un conjunto de oportunidades que pueden contribuir parcial o totalmente a la consecución de ciertos objetivos. A partir de aquí será necesario estandarizar

las palancas de valor para poder comparar diferentes oportunidades y explorar más en profundidad cada una de ellas para valorar su viabilidad, típicamente en términos de esfuerzo-coste antes de lanzarnos a su implementación.

A la hora de estandarizar, cada organización puede definir sus propias palancas de valor, pero deben ser lo suficientemente generales como para poder servir de marco común para la comparativa. Como no todas las palancas se pueden traducir en impacto financiero, habrá que definir también otro tipo de palancas más cualitativas e incluirlas en la valoración.

El análisis posterior de impacto para el negocio vs viabilidad, es lo que permitirá priorizar todas estas iniciativas y descartar aquellas que no tenga sentido implementar en el corto y/o medio plazo bien porque aportan poco valor frente a otras, bien porque el esfuerzo de llevarlas a cabo sea excesivo.

CASOS DE ÉXITO

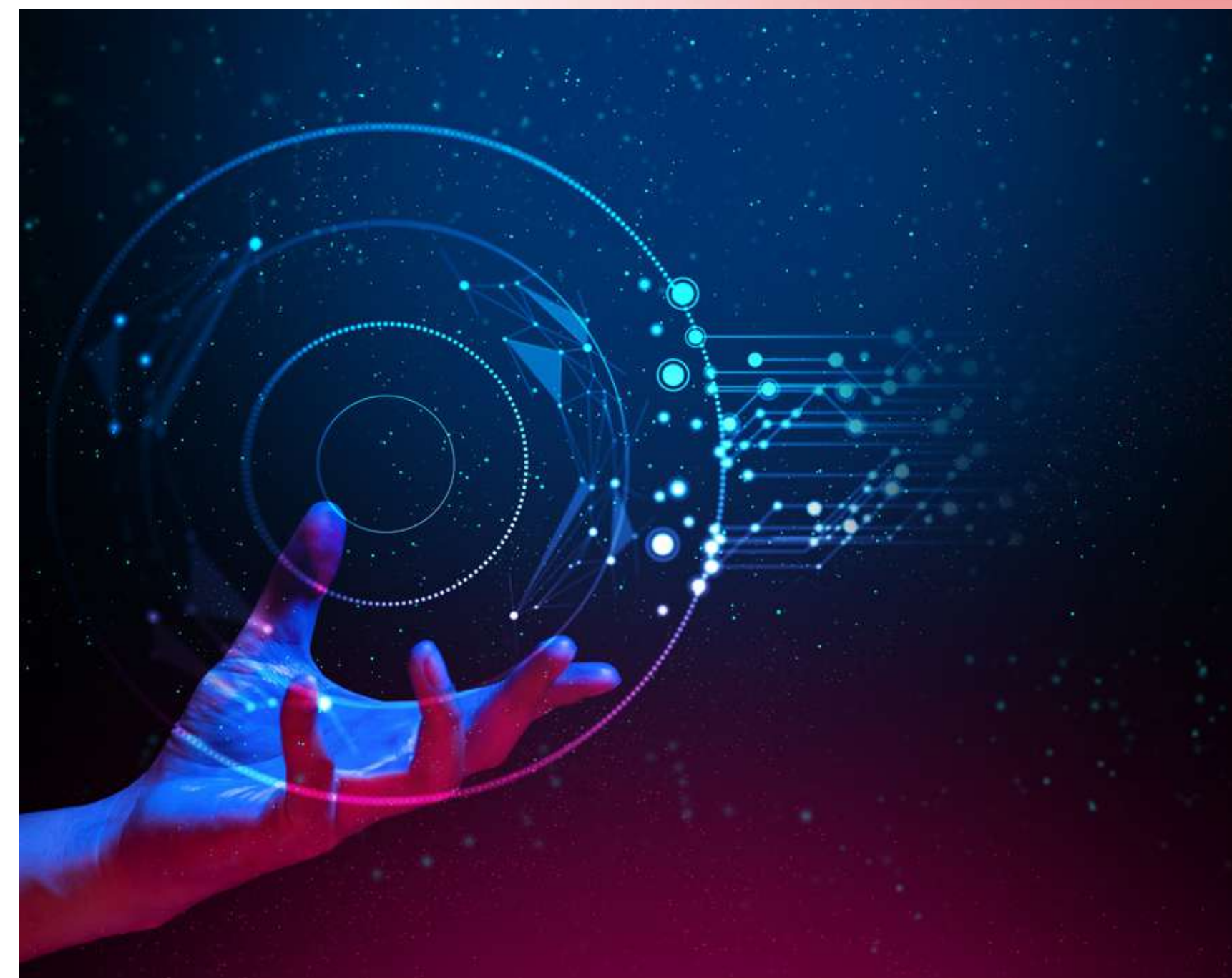


TECNICAS REUNIDAS

Programa de talleres de empatía e inspiración para el levantamiento de casos de uso. Portafolio de iniciativas digitales.

La estrategia corporativa de transformación digital de Técnicas Reunidas se basa en un cambio cultural de la compañía apalancada en las nuevas tecnologías. Es un cambio que necesita de la implicación de todas las personas, y solo se puede producir si se comprende el nuevo paradigma digital y los beneficios que puede aportarnos.

En el plan estratégico se ha definido un programa de talleres de innovación digital para enriquecer la visión de nuestros directivos. El programa nace en 2019 con el objetivo de formar en las tendencias digitales que están teniendo mayor impacto, e inspirar con referencias de casos de éxito de otras empresas.



Lograr permeabilizar en la capa directiva con la cultura digital fue un elemento clave para desarrollar un portafolio de iniciativas de datos e inteligencia artificial. Las sesiones de innovación están inspiradas en la metodología *design thinking*, y organizadas a medida para cada una de las áreas de la compañía. Generalmente se definen de la siguiente manera:

- **Sesiones inspiracionales** sobre tendencias para comprender el impacto de las tecnologías en los procesos, los nuevos modelos de negocio y las implicaciones de las tendencias actuales, tanto a nivel macro como sector industrial. Es una inmersión tecnológica que permite a la dirección identificar posibles escenarios futuros que puedan influir en el negocio de la compañía.

CASOS DE ÉXITO

- **Talleres de empatía y descubrimiento de retos.** Tras la inmersión tecnológica tenemos una sesión de innovación para reflexionar sobre los puntos de dolor, y transformarlos en oportunidades descritas en forma de retos.
- **Sesión de ideación.** Una vez hemos determinado los puntos de dolor y focalizado los tramos, hay que esbozar soluciones basadas en nuevas tecnologías. En las sesiones anteriores, se centraron los esfuerzos en comprender y concretar la información obtenida en áreas de acción. Después de esto hay que valorar posibles soluciones innovadoras que podrían resolver el desafío planteado.

Este tipo de dinámicas se han convertido en una práctica habitual en la compañía, 3-4 sesiones anuales, y han ayudado a definir el portafolio con más de 60 iniciativas identificadas por la dirección y el negocio.

Posteriormente, cada una de estas iniciativas se valoran considerando múltiples aspectos, como: esfuerzo, recursos, dificultad técnica, tipo de innovación, horizonte temporal, riesgos, valor o retorno previsto, entre otros. El portafolio de iniciativas se completa con una matriz esfuerzo-impacto, y el ranking de priorización, lo que permite comparar y seleccionar las iniciativas que se desarrollarán en la compañía. Las seleccionadas se desarrollan internamente de acuerdo con la metodología propia de Técnicas Reunidas "LeaDIn" que define modelo de ejecución hacia la industrialización e implantación.



Plan de Transformación Digital - Identificando Oportunidades.

El Plan de Transformación Digital de Navantia se convierte en uno de los pilares de su Plan Estratégico (PEN), aprobado a finales de 2018, pudiéndose extender hasta el año 2025.

En el marco del PEN, la transformación digital se une con la modernización de algunos centros de producción donde se llevarán a cabo inversiones en tecnología, innovación y formación, consolidando el astillero 4.0. y continuando la evolución al astillero 5.0, que suma a la digitalización el poner en el centro a las personas y la transformación digital.

Este Plan de Transformación Digital de Navantia busca dar respuesta a los retos a los que se enfrenta la compañía en un mercado cada vez más exigente y tecnológico con el objetivo



de incrementar su competitividad, reforzando su propuesta de valor en productos y servicios nacionales e internacionales, rejuveneciendo la plantilla y modernizando los procesos, productos e instalaciones.

Se fundamenta, por tanto, en desarrollar nuevos modelos de negocio y orientar las operaciones hacia la eficiencia transformando los actuales astilleros y fábricas de Navantia en verdaderos astilleros y fábricas 5.0 capaces de afrontar con éxito el futuro bajo parámetros de colaboración, adaptabilidad, transparencia e integración con el objetivo de mejorar la competitividad a través de la reducción de plazos, de costes y de la mejora de la calidad, en los procesos productivos y corporativos, así como con la definición de nuevos productos y servicios que exploten estas tecnologías y sitúen a Navantia en la vanguardia del sector a nivel mundial.

2.1.4.

CASOS DE ÉXITO

Para esta transformación digital, Navantia ha acometido este reto basándose en tres ejes:

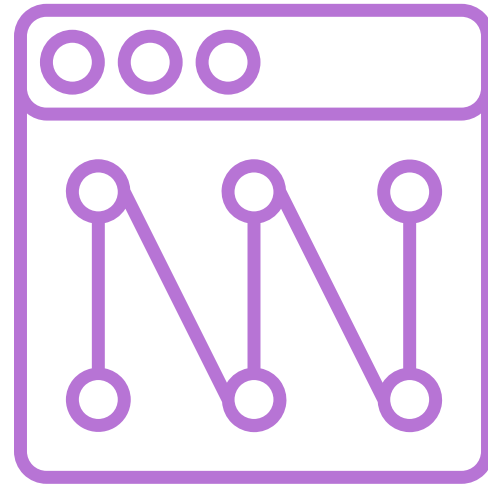
- **El primer eje** está enfocado a los productos y servicios inteligentes, con el objetivo de aumentar las capacidades de los buques mejorando su posición comercial e incrementando la cartera de pedidos.
- **El segundo eje** tiene que ver con desarrollar la fábrica inteligente e integración de la cadena de suministro; incorpora además mejoras en las partes más corporativas de soporte al negocio con objeto de mejorar la eficiencia de los procesos.
- **El tercer eje** está orientado a las personas, la cultura digital y las nuevas herramientas que permitan conectar lo que hacemos desde la ingeniería, las compras, la producción y los servicios.

Estos ejes se despliegan en seis programas o líneas que engloban al conjunto de iniciativas diseñadas para promover y desarrollar la digitalización y la transformación tecnológica de la empresa en un marco ágil que facilite el trabajo con visión, flexibilidad y confianza para dar una rápida respuesta a los desafíos y cambios del entorno, y con el objetivo de crear valor para la empresa.

Dentro de las tecnologías habilitadoras para este plan de Transformación Digital, el uso de la Inteligencia Artificial tiene un papel fundamental.

Para lograr identificar las oportunidades y casos de uso de Inteligencia Artificial dentro de Navantia, se combina con la realización de Sesiones de Descubrimiento, mediante la aplicación de metodología *Garage*, basada en prácticas de estrategias probadas como *Agile* y *Design Thinking* con la investigación del estado del arte de la tecnología.

Con estas fuentes, se realiza una programación mediante un *Funnel* de Innovación, en el que se categorizan y miden los proyectos para verificar su aplicabilidad y posibilidad de desarrollo, iniciando los mismos mediante POC o MVP y escalándolos a toda la organización si resultan satisfactorios. Asimismo, se tienen en consideración las lecciones aprendidas en la implementación de nuevas tecnologías en los programas en vigor, así como la previsión de nuevas capacidades que serán requeridas en futuros proyectos, las cuales emanan de las actividades de vigilancia tecnológica desarrollada por Navantia.



2.2. Reto 2. Medición del Impacto

2.2.1. Descripción del Problema

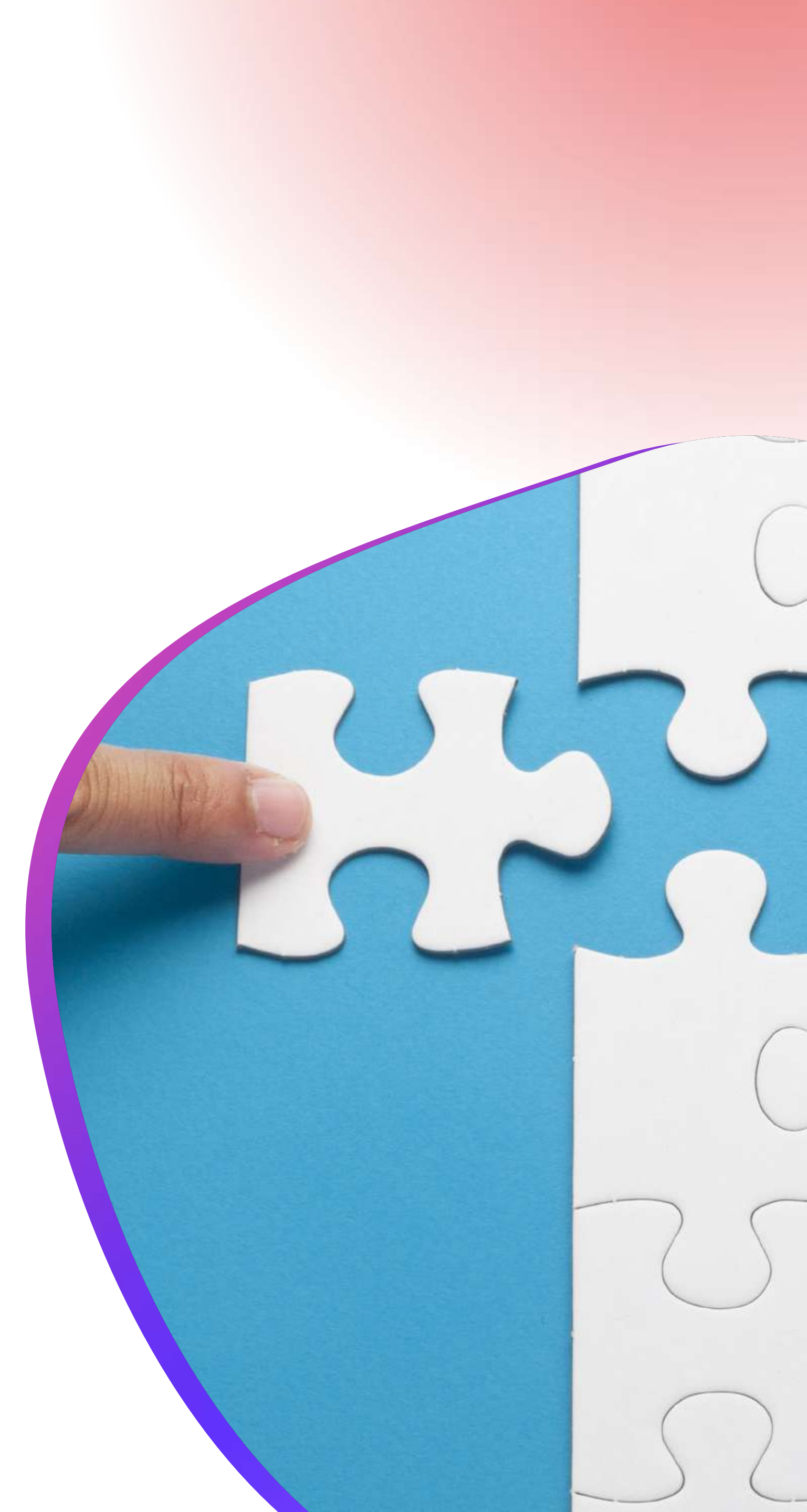
Desarrollar iniciativas basadas en modelos de Inteligencia Artificial es algo costoso, porque lleva bastante tiempo y es más complicado medir el impacto real que tienen los casos de uso en la cuenta de resultados de las organizaciones.

Como en cualquier proyecto tecnológico, cuando se implementa una iniciativa de IA, hay que cuantificar el beneficio que aporta a la organización vs el esfuerzo que conlleva ejecutarla.

En la parte de beneficio o impacto para el negocio, se busca, a priori, obtener un ROI que justifique la inversión necesaria para desarrollar el proyecto. Sin embargo, los

proyectos de Inteligencia Artificial tienen ciertas peculiaridades que hacen que no siempre se pueda medir su impacto en términos financieros. Incluso cuando puede hacerse, éste puede ir cambiando a lo largo del tiempo a medida que cambian los factores que determinan el resultado del algoritmo de IA. Por eso, a la hora de medir el potencial impacto para negocio, se deben combinar palancas cuantitativas (preferiblemente llevadas a términos financieros) y cualitativas (gestión del riesgo, ventaja competitiva, nuevas oportunidades de negocio, etc.).

Esto se verá mejor a través de algunos ejemplos que ilustran estas afirmaciones:



INTRODUCCIÓN

ÁMBITO DEL VALOR PARA NEGOCIO

ÁMBITO DE LA CIENCIA DE DATOS

ÁMBITO DE LOS DATOS

ÁMBITO DE LA ORGANIZACIÓN

ÁMBITO DE PERSONAS Y CULTURA

ÁMBITO DE LA SOSTENIBILIDAD

CONCLUSIONES



Ejemplo 1:

“Se quiere predecir la demanda de coches en una determinada zona para ajustar el precio y maximizar el beneficio por ventas”.

En este caso, es posible que desarrollemos un algoritmo de Inteligencia Artificial sencillo, que tenga en cuenta únicamente el histórico de ventas de la planta en los últimos 10 años y, con estos datos, determine el precio objetivo. Esto permitiría tener un beneficio estimado (frente a no desarrollar el proyecto), que incrementaría las ventas en un X% gracias a la elasticidad en el precio.

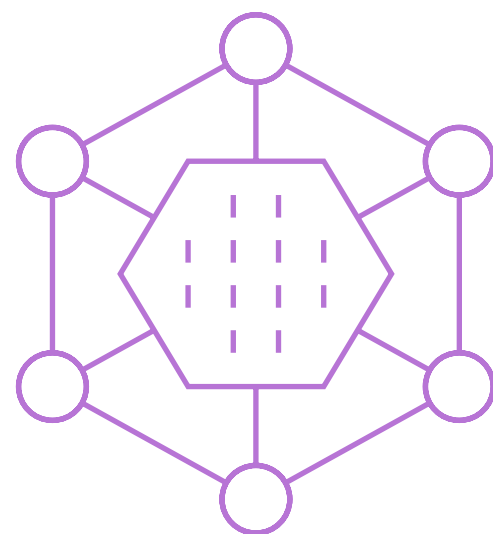
Sin embargo, la demanda no es un parámetro estático. ¿Qué ocurre si se establece un nuevo competidor en la zona o si cierra una de las empresas de flota a las que se abastece, y eran variables que no estaban en el algoritmo de IA? En ese caso, la precisión del algoritmo variará y no permitirá obtener el beneficio previsto. El ROI habrá cambiado. Será necesario por tanto incorporar nuevas variables para ajustar el modelo y mejorar su precisión a futuro. Con cada nueva iteración, el modelo será más preciso y tendrá en cuenta más factores, pero esto requiere una inversión recurrente para su mantenimiento y evolución que debe ser tenida en cuenta desde el inicio en el proyecto. En este ejemplo, el ROI inicial varía a lo largo del tiempo y el esfuerzo necesario, también. Por eso se debe medir el impacto cada cierto tiempo.

Ejemplo 2:

“Queremos mejorar la percepción de calidad de nuestro servicio”.

En esta ocasión se va a usar Inteligencia Artificial, en concreto el procesamiento del lenguaje natural, para interpretar automáticamente los comentarios de los usuarios en redes sociales y poder identificar y clasificar los principales puntos de mejora. Así, lo que se mejora, si se ponen en marcha acciones concretas, es la percepción del servicio y eso se traducirá en mayores ventas. Qué porcentaje de beneficio viene determinado por el IA y cuánto por la acción posterior es algo difícil de determinar. Ahí puede ser útil establecer el coste de oportunidad - no hacer nada (pérdida de clientes frente a competencia) vs el coste de implementar estas acciones.

En términos de esfuerzo, siempre será bueno evaluar el desarrollo de un nuevo algoritmo vs usar uno ya existente en el mercado e incluso se pueden combinar ambas opciones: usar un algoritmo ya existente y, en el caso de querer mejorar la precisión o diferenciarse de la competencia en un proceso concreto, entonces evaluar el coste de desarrollarlo internamente e ir evolucionando con el tiempo.



2.2.2. Soluciones

2.2.2.1. Business Case Previo.

Hipótesis de partida y escenarios alternativos. En muchas organizaciones, los recursos necesarios para poner en marcha una iniciativa de Inteligencia Artificial dependen del potencial beneficio que ésta pueda aportar. Y la manera más “sencilla” de demostrar ese beneficio es a partir de la hipótesis de partida.

Porejemplo: “Si consiguiera mejorar la predicción de la demanda de este producto en un X%... eso me permitiría ahorrar un Y% en costes de almacenamiento.”.

Las hipótesis de partida, deben ser verosímiles y realistas. Debe existir un consenso respecto a las probabilidades de que ocurran para dar por válido el business case.

Otra clave de las hipótesis de partida, es que deben ser reflejadas en términos cuantitativos y siempre que sea posible, en el ámbito financiero.

Y para ello, es bueno tener una serie de palancas de valor para analizar el beneficio del caso de uso:

- ¿Este caso de uso me va a permitir generar más ingresos? ¿Por qué? ¿Cuántos?
- ¿Me va a permitir ahorrar costes (eficiencia)? ¿Por qué? ¿Cuánto?
- ¿Me va a permitir disminuir los riesgos que asumo? ¿Cómo? ¿En qué porcentaje?
- ¿Este algoritmo me va a dar una ventaja competitiva? ¿En qué se traduce eso? ¿Mayores ventas? ¿Cuántas?

Si no existen muchas certezas sobre cómo se puede cuantificar una hipótesis, se puede jugar con escenarios alternativos: optimista, medio y pesimista. En cuyo caso, la organización debe elegir cuál considera el más adecuado para tomar una decisión sobre si merece la pena o no implantar un caso de uso.

El escenario que se elija, con las hipótesis asociadas, nos estará marcando el umbral de aceptación de nuestro caso de uso.



2.2.2.2. Medición de resultados post-implantación

Con independencia de si se ha hecho un *business case* previo o no, una vez desarrollado un caso de uso de la Inteligencia Artificial, es conveniente medir su impacto y éste dependerá del tipo de caso de uso y contexto de negocio en el que se aplique. Con ejemplos se podrá ver con mayor claridad:

ejemplo 1: Si lo que se quiere es predecir el volumen de ventas de coches con 1 mes de antelación mediante un algoritmo de IA, habrá que comparar el resultado de la predicción con las ventas reales y ver cuál es el grado de precisión y si éste supera o no, al que se esperaba.

Ejemplo 2: Si lo que se quiere es automatizar la clasificación de miles de documentos mediante un algoritmo de IA, lo que hay que hacer es medir los tiempos actuales, compararlos con los pasados y a partir de ahí, cuantificar ahorros económicos y/o de eficiencia en el proceso.

Ejemplo 3: Si lo que hay que detectar es un defecto sobre un objeto a partir de un algoritmo de *vision computing*, entonces la métrica será más compleja y deberá tener en cuenta una ponderación entre el número de aciertos del nuevo proceso automático vs el número de aciertos en manual, el tiempo de detección del algoritmo vs manual y el coste de uno frente a otro.

Es decir, para cada caso de uso hay que haber identificado qué se quiere conseguir y a partir de ahí determinar qué métricas o conjunto de ellas, ponderadas, van a dar el beneficio esperado.

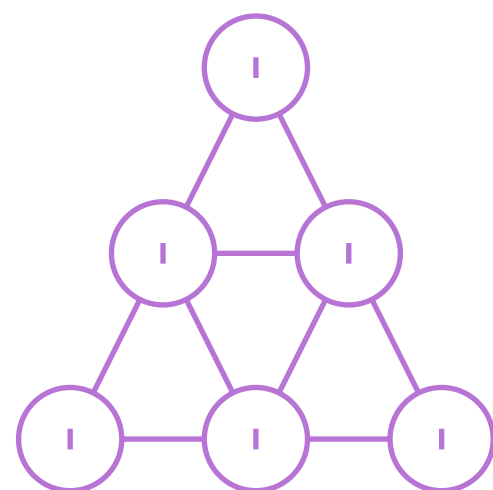
2.2.2.3. Medición recurrente de resultados

Hay momentos en los que el contexto puede variar de manera imprevista y puede afectar al resultado del caso de uso. En estas ocasiones, que se aplican además a procesos de negocio críticos, interesa especialmente hacer mediciones recurrentes para asegurar que los beneficios no se deterioren con el paso del tiempo o las condiciones cambiantes del contexto.

Pongamos un ejemplo. Si queremos predecir la inflación mediante un algoritmo de IA, interesará comprobar de forma automática y recurrente, que el modelo no va perdiendo precisión a lo largo del tiempo porque el contexto haya variado.

Medir de manera recurrente, permite no solo detectar desviaciones respecto a lo previsto, sino también mejorar los modelos e identificar variables que deben ser tenidas en cuenta para su mejora continua.





2.2.3. Niveles de madurez

2.2.3.1. Nivel 1. Fundacional. No hay medición

Puede existir cierta conciencia sobre la necesidad de medir el impacto, pero probablemente no existen las capacidades, herramientas o enfoques adecuados para medirlo. Hay una sensación optimista de que añadiendo Inteligencia Artificial a cualquier proceso implica reducciones de costes sin saber exactamente qué hay que medir o qué parte del proceso potencialmente se puede mejorar. Se considera más fácil medir una vez el caso esté funcionado o en producción y muchas veces por la propia falta de medición el propio caso se desecha o no pasa de una mera prueba de concepto.

2.2.3.2. Nivel 2. Acercamiento. Business Case Previo

La necesidad de medir el impacto para el negocio es evidente y se trabaja en una justificación inicial del caso antes de tomar ninguna decisión de implementación. Se definen unas hipótesis y sobre ellas se calculan los potenciales beneficios de negocio. Este caso de negocio normalmente no se revisa conforme avanza la implantación del caso de uso y suele quedar relegado a una presentación inicial que no se actualiza o comprueba una vez llegan los primeros resultados si el caso llega realmente a producción.

2.2.3.3. Nivel 3. Aspiracional. Medición recurrente

La medición recurrente del impacto de cada caso forma parte intrínseca de la definición y puesta en marcha de la iniciativa.

Desde el punto de vista técnico y de implementación el cómo se va medir y con qué frecuencia se tiene en cuenta desde el inicio y se definen límites o valores objetivos para saber en todo momento si el rendimiento está por encima de lo esperado o, por el contrario, el retorno no cumple con la hipótesis inicial.

2.2.3.4. Nivel 4. Maduro. Auditado por el CFO (Cuantificando Market Value Added)

Existe una fuerte cultura de medición y cuantificación del impacto de cada caso consolidada en todos los niveles de la organización. Business as Usual es la idea de que el equipo financiero sea el responsable de auditar el retorno de los casos desde el punto de vista del negocio como cualquier otra iniciativa estratégica de la compañía. Una mentalidad de toma de decisiones basada en impacto es parte del ADN de la organización.

CASOS DE ÉXITO



Metodología de medición y auditoría del Impacto económico.

El programa tiene un objetivo claro de alcanzar 800 millones de euros de impacto económico acumulado en 2022.

El criterio para medición de dicho objetivo es muy claro, medir el impacto económico positivo antes de los impuestos en flujo de caja operativo y ahorros de CapEx. Dicho impacto puede proceder de:

- aumento de ingresos o margen bruto (ej. Aumento de volumen de ventas por nuevos clientes o mayor precio medio de venta).
- Reducción de OpEx (ej. eficiencias operativas).
- Reducción de CapEx (ej. mejora de vida útil o eficacia de activos).

Se realiza un amplio seguimiento de las iniciativas digitales y se recopila información tanto de las iniciativas que cumplen las características 3 + 1 (descrito en un punto anterior) como de otras iniciativas destacadas fuera de las carteras digitales

(por ejemplo, adquisiciones de *startups* relacionadas con el programa). Hay que supervisar de manera integral el programa, incluyendo:

- KPIs económicos: inversión, gasto, estimaciones de retornos previstos y obtenidos.
- KPIs cualitativos: situación cualitativa según el resultado/ impacto esperado (verde, amarillo o rojo) y situación según el calendario (en plazo, retraso medio o retraso grave).

Para asegurar la homogeneidad de las cifras, se establecen y difunden criterios comunes para supervisar el avance del Programa Digital. De forma continua, las funciones de Planificación y Control de las Unidades de Negocio ayudan a los equipos de caso a medir y realizar un seguimiento de las cifras económicas reales y estimadas. Trimestralmente, la función de Planificación Corporativa Integrada del CFO lidera un proceso de consolidación de cifras económicas que involucra a las Unidades de Negocio y sus funciones de Planificación y Control.

CASOS DE ÉXITO



Medir impacto económico (infraestructura open source vs propriety, iniciativas nuevas y mejorar procesos existentes midiendo el uplift)

Todas las organizaciones ya saben que la inteligencia artificial tiene un valor económico enorme.

PWC estimó en 2017 que en 2030 el valor total iba a ser 15,7 billones. Sin embargo, estos números altísimos no les dicen a las empresas e instituciones individuales cuánto valor hay para ellas y cómo medir el valor que generan con sus iniciativas de Inteligencia Artificial. De hecho, muchas organizaciones luchan por poner un valor económico a sus inversiones en IA, que es una de las principales razones por las que tantas iniciativas no alcanzan los ambiciosos objetivos que alguna vez establecieron.

Hay varias fuentes de valor económico:

- Reducción de costes de infraestructura. El valor que se puede crear con el cambio de infraestructura de propietario



a abierto puede ayudar al principio para demostrar valor en poco tiempo y retorno rápido de la inversión, pero no es escalable porque es un "one off". El valor se puede medir con el TCO (total cost of ownership). Aparte de los menores costes, la ventaja es evitar un "vendedor lock-in". La desventaja es que para poder aprovechar las soluciones de código abierto la organización debe tener disponible el conjunto de habilidades y la experiencia adecuada, que no siempre es el caso al principio.

- **Optimización de negocio.** Hay dos formas de optimizar el negocio con Inteligencia Artificial:
 - Generar ingresos adicionales significa hacer más con lo mismo.
 - Reducir costes significa hacer lo mismo con menos.

CASOS DE ÉXITO

El problema con esto es que no es fácil averiguar cómo medir este “más”. La IA es casi nunca la única razón para una mejora. Otras áreas de negocio estarán involucradas y entonces se vuelve difícil decidir cuánto valor asignar a IA versus a por ejemplo el departamento de marketing o retención.

Las soluciones para este problema son:

- Hacer cosas nuevas, que antes nunca se habían hecho. Con esto es poco discutible que el valor se pueda asignar a la inteligencia artificial.
- Trabajar con grupos de control para poder medir la diferencia entre el uso de la inteligencia artificial frente a su no uso.
- **Monetización de datos externos.** Aquí, el valor económico

no se genera a partir de la optimización del negocio, sino que se genera a partir de nuevos negocios centrados en datos. Esto es solo para organizaciones que han alcanzado un cierto nivel de madurez en IA. Estas organizaciones avanzadas pueden comenzar a buscar nuevos negocios en torno a los datos, ya sea creando nuevas propuestas de valor de datos, es decir, nuevos productos donde los datos están en el centro, o creando “insights” para ayudar a otras organizaciones a optimizar su negocio. En este caso, medir el valor económico del *big data* no es diferente a lanzar nuevos productos al mercado y gestionar su P&L (Profits and Loss, ganancias y pérdidas).

La experiencia ha enseñado que la mayor parte del valor de IA en los siguientes años provendrá de la optimización del negocio, es decir, al convertir a las empresas en organizaciones que toman sus decisiones basadas en datos.

ferroviario

Medición de impacto (cualitativo y cuantitativo)

A la hora de identificar los casos de uso, ya se hace un trabajo previo de estimación del valor o impacto de una iniciativa.

Normalmente, la estimación suele ser a alto nivel ya que el objetivo en el momento de identificación inicial de un reto es contar con un rango que permita priorizar y comparar diferentes oportunidades antes de profundizar en su análisis y lanzamiento.

Una vez evaluado a alto nivel el valor potencial de un caso de uso, cuando se profundiza sobre la oportunidad, hay que partir de una serie de hipótesis que permiten cuantificar en términos económicos el impacto esperado (cómo afecta al flujo de caja del negocio). A veces este impacto económico no es posible y hay que hablar de otro tipo de palancas de valor, más cualitativas (ventaja competitiva, reducción de riesgos, otros).

Una vez que se ha implementado un caso de uso, hay que medir resultados contra lo esperado. Y éste es quizás el reto mayor



en Inteligencia Artificial porque la precisión de un algoritmo puede ir variando con el tiempo (si cambia el contexto en el que actúa). Y por extensión, su impacto, también. Por lo tanto, se debe hacer una medición de impacto cada cierto tiempo.

En el caso de algunos algoritmos (por ejemplo, predictivos), la medición puede hacerse esperar porque pueden estar diseñados para predecir algo en el largo plazo (a 1 año vista) aunque ayuden a la toma de decisiones en el corto plazo. En estos casos, hasta que eso no ocurra, no se podrá de manera factible medir resultados.

En otros casos, la medición del impacto se debe basar en hipótesis. ¿Cómo hay que medir el valor de contar con un algoritmo que da una ventaja competitiva en el mercado por ser único dentro del sector? ¿Permite ganar más concursos, licitaciones, ofertas? Pero, ¿cuántas? ¿Cómo se mide? Cada organización debe desarrollar su business case y tratar de contrastarlo contra mercado.

CASOS DE ÉXITO

AIRBUS

Digital Transformation Program

Es fundamental, para que las aplicaciones de IA tengan éxito, que se pueda demostrar su beneficio de cara a la cuenta de resultados, es decir el negocio.

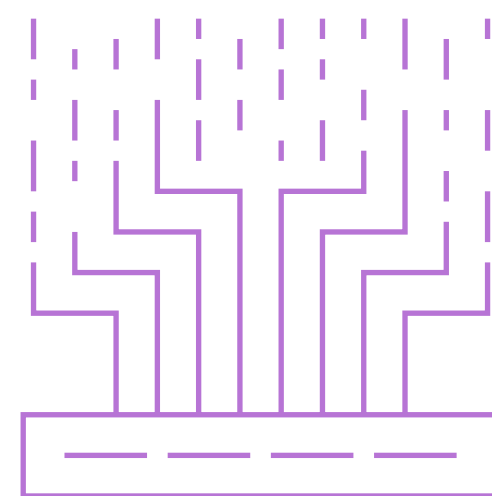
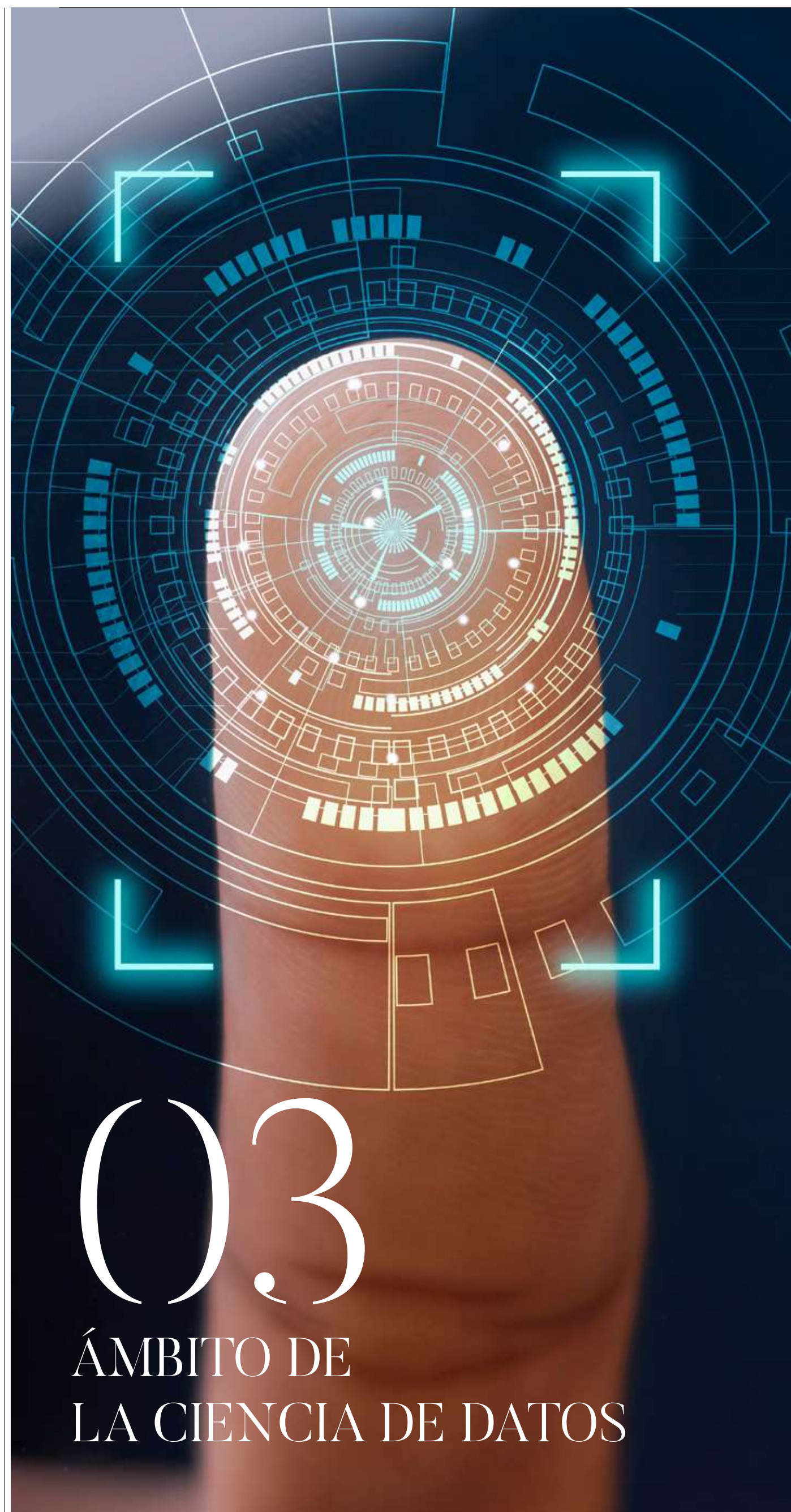
Por ello, hay que definir procesos claros que cuantifiquen el impacto de los proyectos de IA. La recomendación es seguir un proceso colaborativo y que ayude a la toma de decisiones de forma rápida.

El primer paso es realizar un primer *value assessment* a alto nivel de las ideas que evalúen los expertos técnicos que pueden tener potencial. Una vez ese *value assessment* es positivo e interesante para el negocio se procede a una primera fase de estudio de viabilidad donde se empieza a trabajar en preparar un *business case*. De momento será cualitativo ya que no se tendrán todos los datos.



Tras este paso, se recopila información detallada para madurar el business case junto con la unidad de negocio a la que debería revertir los beneficios del proyecto de IA. Para ello, habrá que recopilar datos financieros relevantes y desarrollar la propuesta de concepto que dará paso a una propuesta madura de MVP (minimum viable product). Estos dos pasos irán en paralelo y se retroalimentan, es decir la maduración del business case y los resultados de la prueba de concepto contribuyen a definir el MVP. En este punto, en función del resultado de la prueba de concepto y del business case, es el negocio el que tiene que decidir si lanza el MVP.

Una vez finalizado el MVP, es necesario volver a revisar los resultados con el negocio, y decidir si hay que ajustar alguno de los parámetros del business case. Es en este momento donde el negocio tiene que decidir si se procede a la industrialización de la solución.



3.1. Reto 3.

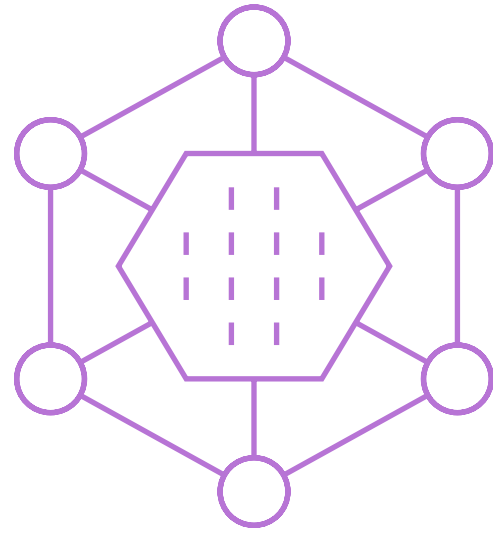
Desarrollo de modelos analíticos más complejos

3.1.1. Descripción del problema: No estamos preparados para desarrollar modelos demasiado complejos y confiar en sus resultados para tomar decisiones

Los modelos analíticos complejos transforman una serie de datos de entrada en uno o varios datos de salida, pero tienden a ser considerados "cajas negras" al ser muy difícil comprender qué proceso aplica el modelo al dato de entrada para obtener la salida. De hecho, existe una compensación entre la exactitud de los resultados y su interpretabilidad, de modo que los modelos más sencillos son más fácilmente interpretables, mientras que los más complejos son prácticamente imposibles de interpretar para el ser humano.

Del mismo modo, es necesario saber si los modelos utilizados están sesgados a la hora de realizar predicciones o cuándo un modelo se desvía del criterio previsto en determinadas decisiones. Si los datos que utilizamos tienen sesgos, ocasionados mayormente por las personas al introducir la información, los modelos analíticos también aprenderán conforme a esos sesgos.

Debido a la criticidad de ciertas decisiones, el proceso de toma de decisión no puede recaer sobre un modelo que no sea explicable, interpretable y fiable, o sobre un modelo que contenga sesgos que condicionan sus resultados, a no ser que se utilicen técnicas y herramientas, cada vez más populares, que permitan eliminar o reducir esta problemática.



3.1.2. Soluciones

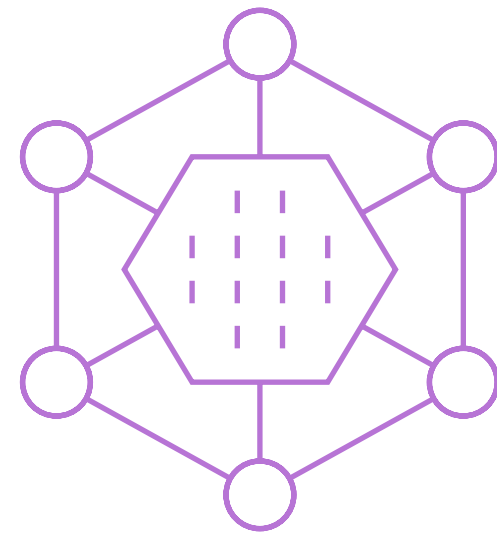
3.1.2.1. Identificar los niveles de madurez

Siguiendo el paradigma de que no se puede gestionar lo que no se puede medir, es necesario desarrollar un sistema para medir la madurez analítica de cada solución o caso de uso que se quiera implementar y garantizar que la organización vaya avanzando gradualmente. Un ritmo demasiado acelerado significa que los usuarios de negocio seguramente no tengan conocimiento ni confianza en las soluciones que se están implementando y por otro lado un ritmo demasiado pausado significa que se están perdiendo oportunidades de aumentar el impacto de las soluciones analíticas que se están implementando.

3.1.2.2. Involucrar personal propio para conocer los modelos

3.1.2.3. Definir y gestionar el ciclo de vida de la ciencia de datos





3.1.3. Niveles de madurez

En el camino de cada organización para convertirse en una compañía verdaderamente *Data Driven*, que tome todas sus decisiones basándose en los datos, se puede definir distintos niveles de madurez en el uso de la analítica en función de las preguntas que se están tratando de responder, la complejidad de los algoritmos utilizados y qué peso tienen las máquinas frente a los humanos en la toma de decisiones.

3.1.3.1. Foco en analítica descriptiva

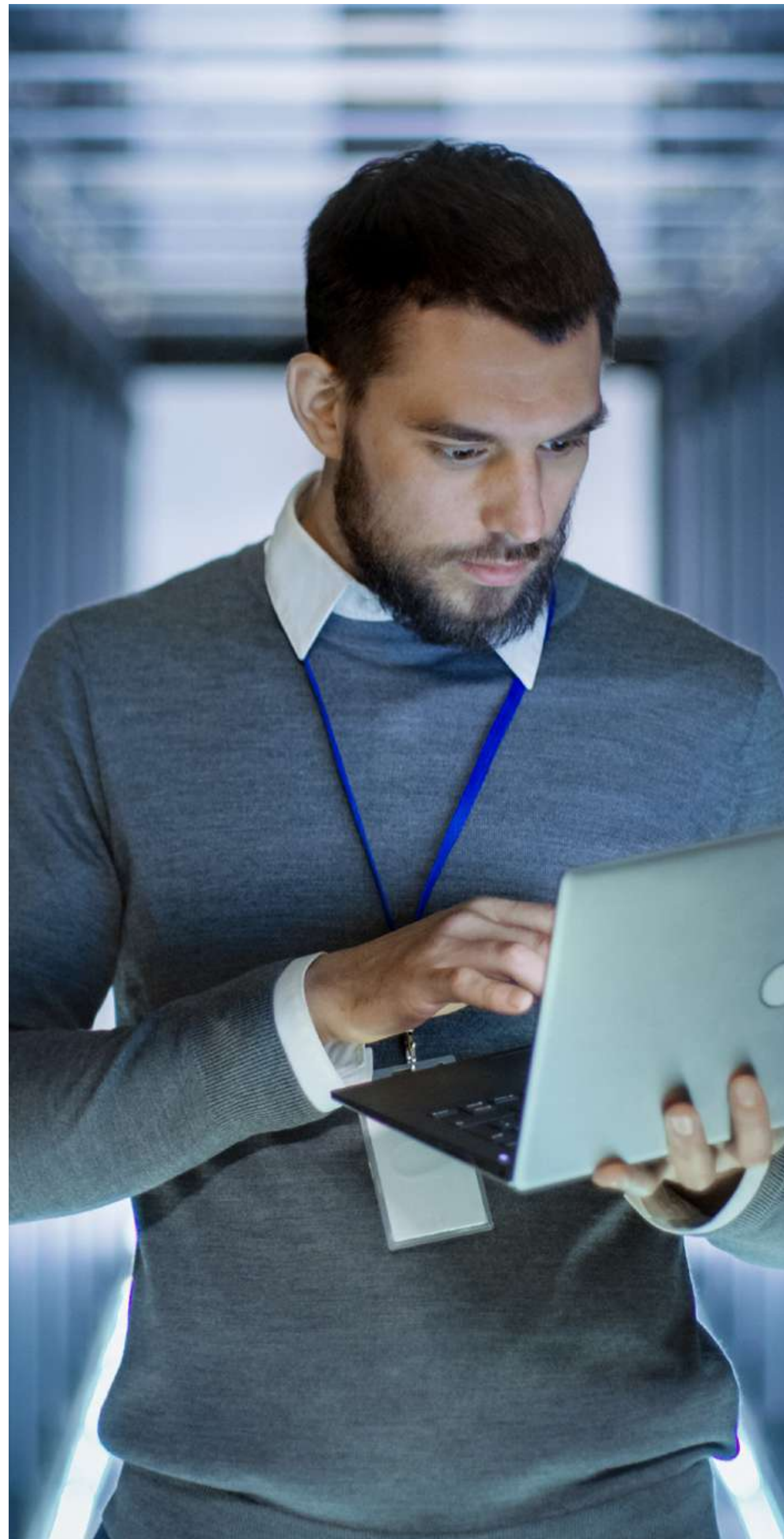
Es el nivel más básico de la analítica y suele estar asociado a preguntas del tipo «¿qué ha pasado?, ¿qué está pasando? o incluso ¿por qué está pasando?»

Para responder estas preguntas se utilizan datos de diversas fuentes que son procesados de tal manera que su visualización por sí sola, o combinada con algunas estadísticas muy básicas, permite responder a estas preguntas. Aquí se habla, fundamentalmente, de los famosos

dashboards que usan, por ejemplo, los analistas de datos trabajando con un equipo de marketing de comercio electrónico para revisar los datos de ventas para identificar las tendencias y patrones de ventas. Se verá un aumento o disminución en las ventas del año pasado, específicamente en qué región y en qué porcentaje.

Aunque es el nivel más básico es fundamental porque con el uso de *boxplots*, correlaciones, *dashboards*, distribuciones de los datos, histogramas o detectando *outliers*, se consigue también tener un conocimiento de la calidad de la información que se podría utilizar en otras fases. Con una adecuada aplicación de estas técnicas, conseguimos eliminar el conocido «*garbage in*» en las siguientes fases.





3.1.3.2. Foco en analítica predictiva

Aquellas organizaciones y usuarios que ya tienen cierto dominio de la analítica descriptiva en su toma de decisiones empiezan a hacerse preguntas más relacionadas con anticipar el futuro para tomar mejores decisiones: ¿qué va a pasar?

Para resolver estas preguntas hay que aplicar algoritmos estadísticos y técnicas de aprendizaje automático que permiten identificar la probabilidad de resultados futuros basados en datos históricos del pasado. Es común utilizar técnicas de regresión o el análisis de series temporales para ser capaz por ejemplo de predecir la demanda de un producto con el objetivo de adecuar los niveles de stock y garantizar que no hay roturas de servicio ni sobrecostes. En este nivel de madurez la máquina proporciona información que nosotros como seres humanos incorporamos a nuestro análisis para la toma de decisión.



3.1.3.3. Foco en analítica prescriptiva

Según las organizaciones empiezan a acostumbrarse al uso de algoritmos como parte del proceso de decisión. Es habitual plantearse preguntas más avanzadas, como: ¿cuál es la mejor opción en esta situación de negocio? Para ello se utilizan avanzadas técnicas matemáticas de optimización que permite encontrar el mínimo o máximo de una situación de negocio dado ciertas restricciones. Las organizaciones comienzan a apoyar sus procesos de planificación, scheduling o logísticos en este tipo de técnicas para minimizar los costes o maximizar los beneficios. En este nivel de madurez las personas no solo confían en la máquina para que le un input importante en la resolución del problema si no que le pregunta la máquina que haría ella en esta situación, la máquina da una recomendación o prescripción que el humano incorpora a su proceso de decisión final.

3.1.3.4. Foco en el desarrollo de sistemas autónomos

Los sistemas autónomos operan en entornos complejos y abiertos con altos niveles de independencia y autodeterminación. Perciben, aprenden, razonan y actúan con conciencia de sí mismos y responden de manera inteligente a los cambios imprevistos en el entorno.

Para su desarrollo se utilizan técnicas complejas de aprendizaje profundo combinadas con aprendizaje de refuerzo y entornos de simulación digitales. Se utilizan fundamentalmente para automatizar un proceso productivo, en el que la máquina ha sido entrenada por los seres humanos para ejecutar con niveles óptimos de perfección. Como por ejemplo un vehículo autónomo capaz de conducir por sí solo o un brazo robótico capaz de ejecutar tareas en una línea de producción son algunos de los ejemplos más habituales de cómo utilizar estas técnicas.

Significan el nivel más alto de madurez analítica en tanto en cuanto el humano deja de participar en el proceso de decisión relevando su papel al entrenamiento inicial y la supervisión de las acciones y decisiones que va tomando la máquina.



3.1.4.

CASOS
DE ÉXITO3.1.4.1 Equipo de *Data Science*, modelo de Peer Review y madurez Analítica

Dentro del HuB de Data & Analytics se estableció en Repsol un equipo de *Data Science* que tiene fundamentalmente los siguientes objetivos:

- **Desarrollar casos de manera interna.** El equipo de *Data Science* de Repsol está formado por profesionales senior del sector con amplia experiencia en el desarrollo de soluciones analíticas. Los perfiles están seleccionados de tal manera que cubren todo el abanico de dominios de la inteligencia artificial con expertos en las siguientes especialidades: *knowledge graph*, análisis de series temporales, algoritmos de clasificación y *scoring* de clientes, modelos de optimización, aprendizaje profundo, *computer vision*, procesos naturales del lenguaje y aprendizaje de refuerzo. Esta diversidad permite al equipo interno tanto supervisar y desarrollar autónomamente cualquier iniciativa que por considerarse estratégica para Repsol no quiera desarrollarse con proveedores externos o bien retomar iniciativas que un proveedor externo no ha conseguido desarrollar con la bondad estadística mínimamente exigida por el negocio.

- **Peer Review de los modelos desarrollados por partners externos.** La función fundamental del equipo de *Data Science* es supervisar la calidad del desarrollo realizado por proveedores externos y asegurar la transferencia de conocimiento con el objetivo de poder operar esos modelos internamente una vez puestos en producción. Esto significa conocer su funcionamiento para poder monitorizar su performance, mantenerlos y aplicar una mejora continua.

Uno de los primeros pasos que se dio como compañía fue la realización de un Peer Review para las empresas colaboradoras de Repsol que, a modo de "sello", se concede, desde el Hub de Data & Analytics y mediante un procedimiento específico, a los colaboradores que desarrollan Inteligencia Artificial con Repsol.

El objetivo de este documento es proporcionar una plantilla, que será de obligado cumplimiento para que el equipo de *Data Science* perteneciente al HUB de Data Management and Analytics, dé por válidos los trabajos realizados por los distintos proveedores durante la fase de MVP desde el punto de vista analítico.

CASOS DE ÉXITO

El *Peer Review* establece un formato obligatorio para la documentación de los distintos modelos analíticos. En las secciones de esta documentación, se demanda información sobre:

- **Los metadatos.** Más en detalle, para cada caso será de obligado cumplimiento los siguientes metadatos relativos al caso digital, los cuales serán compatibles con la herramienta de AirBnB KnowledgeRepo.
- **Resumen ejecutivo.** Incluirá una descripción del caso desde un punto de vista de negocio, incluyendo el contexto del caso, una imagen que resuma los resultados obtenidos, y las conclusiones o recomendaciones.
- **Informe científico.** Un informe científico con la descripción de los datos, su preprocesamiento, el procedimiento de entrenamiento de los modelos, el procedimiento de

medición del éxito del modelo, el impacto en el negocio, así como conclusiones y puntos de mejora.

- **Operación de todos los modelos desarrollados.** Una de las tareas claves es la operación de los modelos que consiguen ponerse en producción. Los modelos analíticos tienen la característica que con el tiempo pierden eficacia por múltiples motivos y por tanto es necesario una monitorización y reentrenamiento frecuente. Esta tarea solo puede realizarse si se tiene un conocimiento suficiente del modelo, de ahí la importancia del *Peer Review* anterior.

Cuando los modelos tienen un impacto económico importante para la organización estas soluciones como es deseable se convierten en un activo crítico para la compañía y por tanto su operación y correcto funcionamiento se convierte en algo estratégico sobre lo que deberíamos tener absoluto control.

Análisis de madurez Analítica

Con el objetivo de ayudar a todas las áreas de la organización a ir avanzando progresivamente en la complejidad de las técnicas analíticas utilizadas definimos un modelo que consta de 9 niveles y en el que se clasifican todos los casos analíticos que desarrollamos.

- 1. Disponibilización de datos.** Para aquellos casos cuya única misión es servir de habilitador para garantizar el acceso a ciertos dominios de información.
- 2. Analítica descriptiva.** Iniciativas basadas en visualización de datos mediante *dashboards* y soluciones de *business intelligence*, centradas en entender qué ha ocurrido y por qué.
- 3. Analítica predictiva básica.** Permite realizar predicciones que van a ser útiles en el futuro. En este nivel, asociado a la pregunta ¿qué va a pasar? Se utilizan un gran número de algoritmos para construir los conocidos como modelos, a modo de ejemplo, se aplican regresiones no programáticas con herramientas ofimáticas.
- 4. Predictiva media.** Subnivel en el que englobamos árboles de decisión¹, aproximaciones bayesianas², algoritmos de «*boosting*» o «*bagging*»³, redes neuronales⁴, modelos ARIMA, ARMA⁵, ... Aunque algunas técnicas pudieran ser similares a las anteriores, el nivel de intervención humana requerido es inferior.

- 5. Predictiva avanzada.** Se hace uso de técnicas de computer vision⁶, data augmentation⁷, deep learning⁸, feature engineering⁹ avanzado, procesamiento del lenguaje natural¹⁰, combinación de simuladores con técnicas de predicción.
- 6. Analítica prescriptiva básica.** Este nivel de analítica permite determinar la predicción óptima para un problema concreto. En este nivel, se utilizan los denominados modelos de optimización basados en reglas¹¹ o las técnicas de optimización lineal¹².
- 7. Prescriptiva media.** Nivel en el que agrupamos las conocidas técnicas de programación lineal entera mixta (también conocida como MILP¹³) o la programación no lineal entera mixta (conocida como MINLP¹⁴).
- 8. Prescriptiva avanzada.** Último subnivel en el que agrupamos los muy aplicados simuladores¹⁵ de la empresa, las denominadas metaheurísticas¹⁶, optimizaciones estocásticas¹⁷, robustas o multiobjetivo¹⁸.
- 9. Sistemas autónomos de decisión.** Nivel máximo de madurez en el que las técnicas de Inteligencia Artificial toman decisiones de alto valor en el mundo real. Permite entrenar aplicaciones en las que estas técnicas aprendan probando y recibiendo una recompensa en función de su respuesta¹⁹. En este nivel aplicamos necesariamente Deep Reinforcement Learning²⁰ o Reinforcement Learning²¹.

Notas

1 <https://www.sciencedirect.com/topics/computer-science/decision-trees>

2 <https://www.stata.com/features/overview/bayesian-analysis/>

3 <https://mitpress.mit.edu/books/boosting>

4 <https://towardsdatascience.com/deep-learning-101-neural-networks-explained-9fee25e8ccd3>

5 <https://towardsdatascience.com/advanced-time-series-analysis-with-arma-and-arima-a7d9b589ed6d>

6 <https://www.sciencedirect.com/topics/computer-science/computer-vision>

7 <https://towardsdatascience.com/data-augmentation-for-deep-learning-4fe21d1a4eb9>

8 https://www.tutorialspoint.com/keras/keras_overview_of_deep_learning.htm

9 <https://github.com/MaartenGr/feature-engineering>

10 <http://language.worldofcomputing.net/nlp-overview/natural-language-processing-overview.html#:~:text=%20Natural%20Language%20Processing%3A%20Overview%20%20>

11 <https://www.sciencedirect.com/topics/computer-science/linear-optimization>

12 <https://www.sciencedirect.com/topics/computer-science/linear-optimization>

13 <https://www.sciencedirect.com/topics/engineering/mixed-integer-linear-programming>

14 https://www.cambridge.org/core/services/aop-cambridge-core/content/view/2D0CE8CDA53363A31ADE8689565517BD/S0962492913000032a.pdf/mixedinteger_nonlinear_optimization.pdf

15 <https://www.springer.com/gp/book/9783319689128>

16 https://www.researchgate.net/publication/326723086_5_Metaheurísticas_de_investigación_en_complejidad

17 <https://www.sciencedirect.com/topics/computer-science/stochastic-optimization>

18 <https://gray.mgh.harvard.edu/research/optimization/39-multicriteria-optimization#:~:text=%20Multicriteria%20Optimization%20%20%20Overview.%20For%20several,available%20in%20radiotherapy%20is%20growing%20at...%20More%20>

19 Amazon.com: Designing Autonomous AI: A Guide for Machine Teaching: 9781098110758: Anderson, Kence: Libros

20 <https://ieeexplore.ieee.org/document/8585411#:~:text=An%20Introduction%20to%20Deep%20Reinforcement%20Learning%20Abstract%3A%20Deep,were%20previously%20out%20of%20reach%20for%20a%20machine.>

21 <https://web.stanford.edu/class/psych209/Readings/SuttonBartoPRLBook2ndEd.pdf>

3.1.4.2. Team Data Science process

El proceso de ciencia de datos en equipo (TDSP) es una metodología de ciencia de datos ágil e interactiva para proporcionar soluciones de análisis predictivo y aplicaciones inteligentes de manera eficiente. TDSP ayuda a mejorar la colaboración y el aprendizaje en equipo al sugerir cómo los roles de equipo funcionan mejor juntos. TDSP incluye procedimientos recomendados y estructuras de Microsoft y otros líderes del sector para ayudar a implementar correctamente iniciativas de ciencia de datos. El objetivo es ayudar a las empresas a que se den cuenta de las ventajas de su programa de análisis.

TDSP tiene los siguientes componentes principales:

- Una definición de ciclo de vida de la ciencia de datos.
- Una estructura de proyecto estandarizada.
- Infraestructura y recursos recomendados para proyectos de ciencia de datos.
- Herramientas y utilidades recomendadas para la ejecución de proyectos.

El proceso de ciencia de datos en equipo (TDSP) proporciona un ciclo de vida para estructurar el desarrollo de los proyectos de ciencia de datos. En el ciclo de vida se describen todos los pasos que siguen los proyectos correctos. Aunque esté usando otro ciclo de vida de ciencia de datos, como CRISP-DM, KDD o el proceso personalizado de su organización, puede usar también el TDSP basado en tareas en el contexto de esos ciclos de vida de desarrollo. A alto nivel, estas distintas metodologías tienen mucho en común.

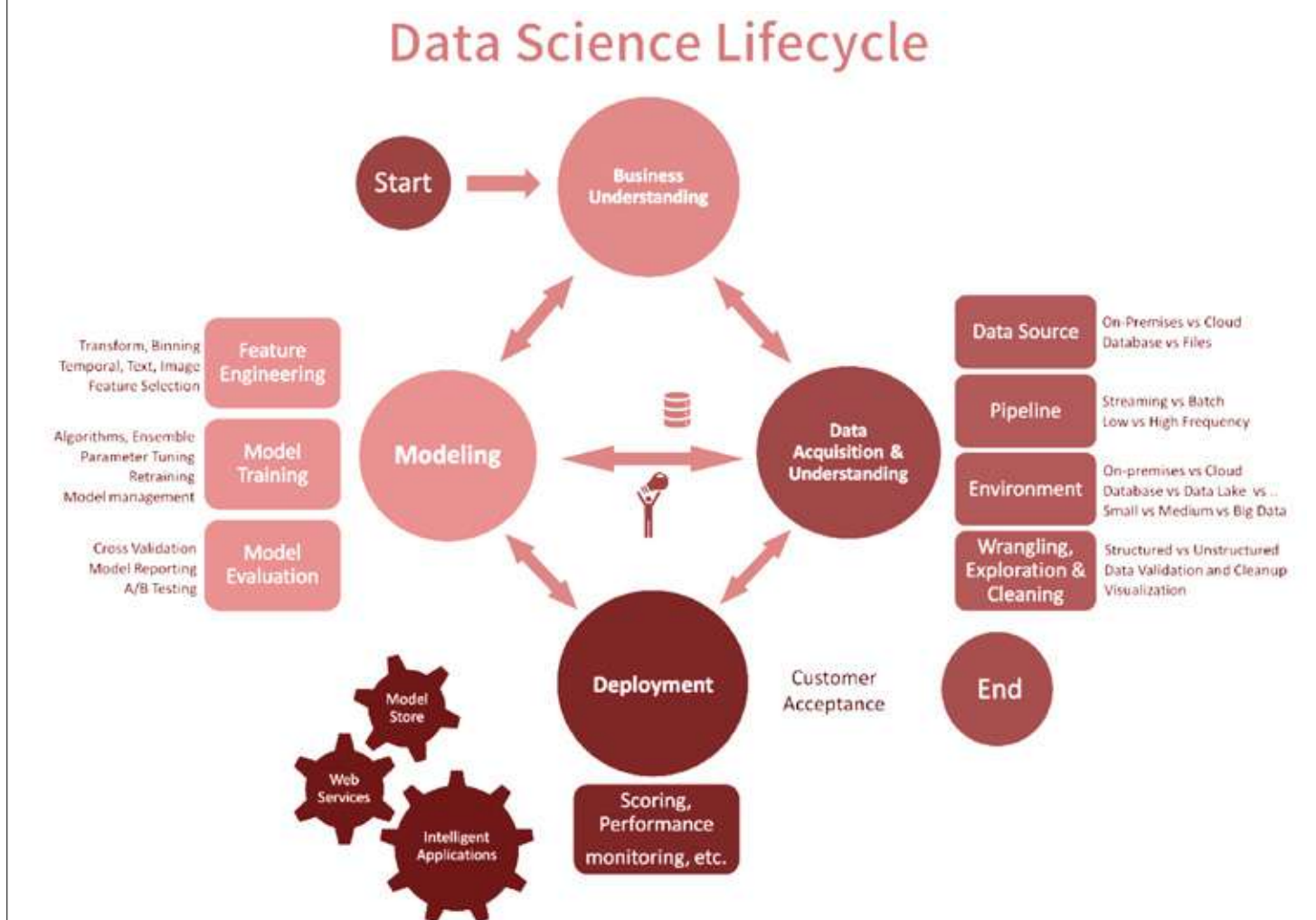
Este ciclo de vida se ha diseñado para proyectos de ciencia de datos que se enviarán como parte de aplicaciones inteligentes. Estas aplicaciones implementan modelos de aprendizaje o inteligencia artificial de máquina para realizar un análisis predictivo. Los proyectos de ciencia de datos exploratorios o proyectos de análisis improvisados también se pueden beneficiar del uso de este proceso. Pero, en estos casos, puede que algunos de los pasos descritos no sean necesarios.

El ciclo de vida describe las fases principales por las que pasan normalmente los proyectos, a menudo de forma iterativa:

- Conocimiento del negocio.
- Adquisición y comprensión de los datos.
- Modelado.
- Implementación.

- Modelado.
- Implementación.

Esta es una representación visual del ciclo de vida del proceso de ciencia de datos en equipo.



3.1.4.3. Equipo de Data Scientists propio, escenarios de simulación y validación de resultados a pasado.

Cuando una organización se enfrenta a un reto de IA, es posible que existan modelos en el mercado disponibles para cubrir un caso de uso concreto y para ser consumidos por la organización, siempre y cuando se supere la barrera inicial de confianza en el modelo.

Para ello y dado que muchos modelos funcionan como “cajas negras”, lo primero que hay que hacer es establecer un plan de pruebas sobre escenarios conocidos que permitan evaluar la fiabilidad de ese modelo y su nivel de precisión para un caso de uso concreto.

Pero, ¿qué ocurre cuando el modelo no alcanza los umbrales deseados o no existe un modelo que cubra esas necesidades? En ese caso, puede interesar desarrollarlo internamente para que se adapte a unas necesidades concretas o para controlar todas las variables que intervienen en el modelo, de manera que se pueda explicar y variar su comportamiento.

En estos casos se necesita:

- Contar con un equipo de data scientists propio o externo, capaz de diseñar y programar el modelo que se necesita.
- Complementar el equipo con personas de negocio, capaces de identificar variables relevantes para el diseño del modelo.
- Establecer un plan de pruebas exhaustivo que permita testar de manera recurrente y continua la fiabilidad del modelo a lo largo del tiempo.

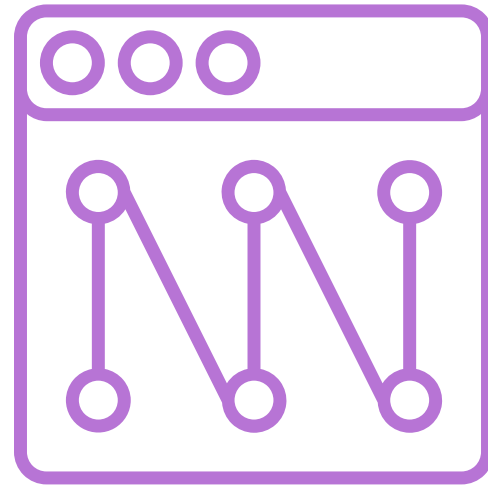
Los casos de uso que requieren el desarrollo de modelos predictivos suelen requerir elaborar numerosos tests que deben validarse con datos del pasado que no han formado parte de los datos de entrenamiento del modelo. Estos tests permiten identificar los puntos de desviación sobre el valor deseado y encontrar nuevas variables o factores de corrección al modelo, hasta encontrar el punto óptimo para considerarlo un modelo estable que pudiera pasar a producción. Para poder realizar tests fiables, se necesita contar

con un historial lo suficientemente amplio como para probar en diferentes ventanas temporales.

Ciertos casos de uso, como los relacionados con imágenes, requieren un histórico de imágenes también elevado. A veces no se cuenta con un banco de imágenes lo suficientemente grande como para cubrir distintas casuísticas

(escenarios con más luz, con menos luz, imágenes borrosas, desenfocadas, defectuosas, etc.) y puede ser necesario generarlas de manera artificial mediante herramientas de simulación y etiquetarlas a posteriori para que sean útiles durante las fases de entrenamiento del modelo. Este tipo de escenarios son costosos en términos de esfuerzo y debe valorarse bien el valor obtenido.

En resumen, el nivel de confianza en el modelo dependerá de cada organización y del nº de tests mínimos que considere adecuados y que sea capaz de lanzar para cada caso de uso. Cuanto más complejo es el caso de uso, mayor esfuerzo requerirá su desarrollo y testeo.



3.2. Reto 4. Puesta en producción de los modelos analíticos

3.2.1. Descripción del Problema

El 87% de los modelos de *Machine Learning* nunca llegan a ponerse en producción.

Este es uno de los grandes retos de los últimos años en materia de Inteligencia Artificial: pasar de las bien conocidas pruebas o pilotos a modelos que trabajen en producción. Es decir, que se usen en casos reales y con datos reales para, como se comenta en otras secciones, medir el impacto de estos modelos. El paso de piloto a producción da vértigo en muchas organizaciones y solo se pueden conseguir si se considera como un proceso continuo y se tiene en cuenta desde el inicio de la planificación del caso.

El reto de incluir a todos los roles necesarios desde la planificación implica resolver la desconexión que históricamente han tenido los principales actores necesarios para la puesta en producción de un modelo: IT, ingeniero de datos, científicos de datos y usuarios de negocio.

En muchas compañías estos roles residen en distintas partes de la organización o trabajan de manera secuencial compartiendo el producto final, una vez supuestamente acabado, como entregable de cada fase. La naturaleza de los experimentos de los científicos de datos implica romper cosas, probar enfoques y probar hasta que se encuentra la solución adecuada, o suficientemente buena, para la hipótesis inicial. Este modelo de trabajo requiere un cambio sustancial en los procesos de aprobación o traspaso entre los implicados en la puesta en producción de modelos.

En el desarrollo de software este problema se ha resuelto hace tiempo con la adopción de prácticas de *DevOps*. En Inteligencia artificial la solución pasa por aplicar ese mismo paradigma que tan buenos resultados ha demostrado en las organizaciones que lo han sabido adoptar. Esta adopción de *DevOps* para AI es lo que se llama *MLOps*.

INTRODUCCIÓN

ÁMBITO DEL
VALOR PARA NEGOCIO

ÁMBITO DE
LA CIENCIA DE DATOS

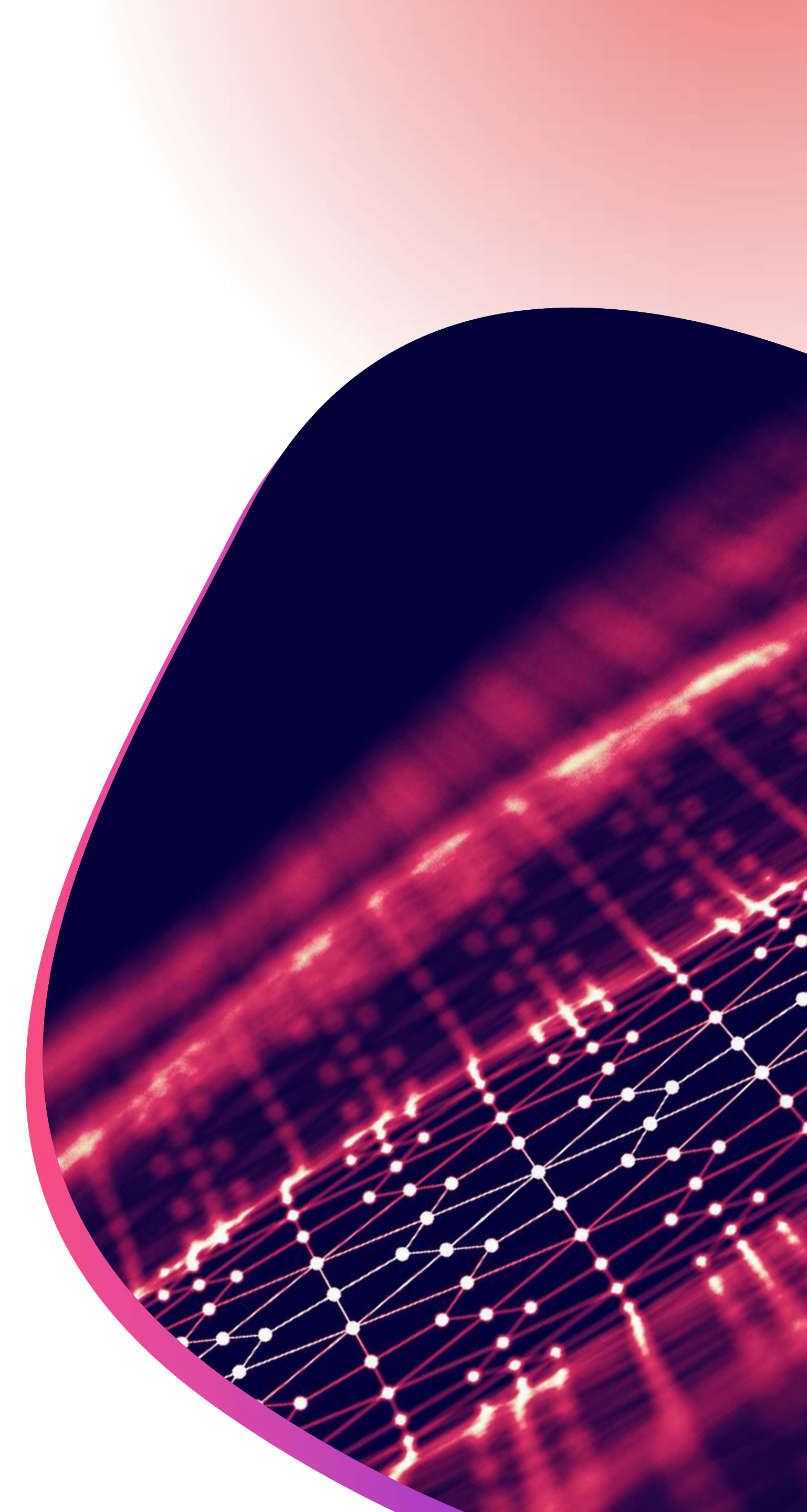
ÁMBITO DE
LOS DATOS

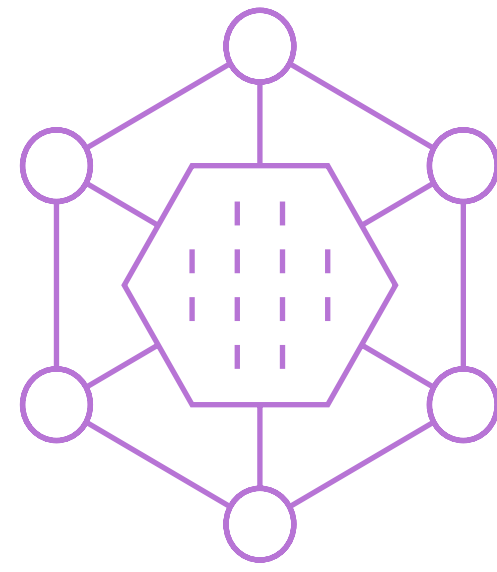
ÁMBITO DE
LA ORGANIZACIÓN

ÁMBITO DE
PERSONAS Y CULTURA

ÁMBITO DE
LA SOSTENIBILIDAD

CONCLUSIONES





INTRODUCCIÓN

ÁMBITO DEL VALOR PARA NEGOCIO

ÁMBITO DE LA CIENCIA DE DATOS

ÁMBITO DE LOS DATOS

ÁMBITO DE LA ORGANIZACIÓN

ÁMBITO DE PERSONAS Y CULTURA

ÁMBITO DE LA SOSTENIBILIDAD

CONCLUSIONES

3.2.2. Soluciones

MLOps o Machine Learning Operations, es una aplicación de la metodología DevOps que incluye los procesos de inteligencia artificial, aprendizaje automático y ciencia de datos en la cadena de desarrollo y operaciones para hacer que el desarrollo de los algoritmos sea predecible, carente de errores y, siempre que se pueda, automático.

MLOps se basa en los mismos principios y prácticas de *DevOps* aplicadas a la creación de modelos, buscando no reinventar la rueda y aplicar las mejores prácticas que vienen demostrando desde hace años que mejoran la robustez y velocidad de puesta en producción de tareas relacionadas con el desarrollo del software. Usar *MLOps*, en definitiva, aumenta la eficacia en el proceso de creación de pipelines. En tareas como, por ejemplo, integración continua, puesta en producción o el propio proceso de desarrollo e implementación. Los objetivos principales de *MLOps* se pueden resumir como:

1. Conseguir una experimentación y desarrollo de modelos más rápidos.
2. Conseguir una implementación más rápida de los modelos en producción, no dependientes de los sistemas o de las responsabilidades divididas entre departamentos y roles.
3. Control de calidad y seguimiento de linaje de un extremo a otro que incluye desarrollar, construir, testear, desplegar y monitorizar los modelos *end to end*.



Las mejores prácticas de MLops se pueden resumir como:

- **Observar de manera automática los datos y su desviación.** Los datos de las compañías están constantemente en cambio y evolución. Cuando se entrenan modelos de *Machine Learning* estos están basados en datos. Una variación en la tendencia de los datos provoca que los modelos que una vez fueron válidos dejen de serlo con el tiempo. Nuestra principal preocupación debe ser monitorizar las variables estadísticas relevantes asociadas a los datos empleados en el entrenamiento de modelos. Con mecanismos de alertas y una gestión de eventos adecuados, así se pueden monitorizar dichos cambios y reentrenar modelos, tanto manual como automáticamente para mantener su rendimiento y precisión a lo largo del tiempo. De esta manera se alimenta el modelo con datos de manera continua y se mantiene el entrenamiento actualizado.
- **Desplegar y monitorizar el rendimiento de manera automática.** Así se pueden sustituir o reentrenar los modelos cuando su rendimiento esté bajo el umbral esperado. Hay que ser

capaces de cruzar los datos inferidos por los modelos en el tiempo con los datos reales posteriormente recopilados y verificar que los modelos que se emplean están realizando predicciones correctas. Así se verifica que la compañía está tomando decisiones correctas en todo momento. Si se detecta una desviación en el funcionamiento de dichos modelos se puede corregir a tiempo, antes de incurrir en errores provocados por el incorrecto funcionamiento de dichos modelos

- **Crear modelos con pipelines de ML reusables.** Se harán utilizando como base las herramientas de control de código. Es importante integrar los *pipelines* dentro del proceso automático de control y despliegue de código. Los *Data Scientist* no son ingenieros de software, y, la mayoría, está especializado en la construcción de modelos, pero no son expertos en desarrollar aplicaciones que cumplan los estándares de seguridad, buenas prácticas de desarrollo, ciclo de vida de software entre entornos (*Continuous Integration/Deployment*), monitorización de la infraestructura y los modelos desplegados en la misma. Hay que apostar por la estandarización del desarrollo de software desde los entornos

previos, con componentes desarrollados dentro de la compañía que favorecen y permiten implementar de manera sencilla y estándar el entrenamiento e inferencia de modelos. Dichos componentes incluyen *out of the box* la integración de forma segura con los datos de la compañía y permiten implementar de manera sencilla los procesos de *Continuous Integration/Development* y el despliegue de los modelos entre entornos automatizado de manera robusta, asegurando la calidad y pruebas de los componentes de software desplegados. Además, se aplica una monitorización *end to end* tanto a los datos que permiten entrenar los modelos como al rendimiento y precisión real de los mismos.

- **Automatizar los despliegues.** Utilizando *DevOps* para el versionado y enriquecimiento de los metadatos de cada despliegue.
- **Crear automáticamente los datos de auditoría** para todos los artefactos que se generen en un pipeline, asegurando que cumple todas las políticas y requerimientos de calidad establecidos por cada empresa.

INDESIA

Framework Data & IA
Industrial Marco de
referencia de datos e
inteligencia artificial para el
sector industrial español

INTRODUCCIÓN

ÁMBITO DEL VALOR PARA NEGOCIO

ÁMBITO DE LA CIENCIA DE DATOS

ÁMBITO DE LOS DATOS

ÁMBITO DE LA ORGANIZACIÓN

ÁMBITO DE PERSONAS Y CULTURA

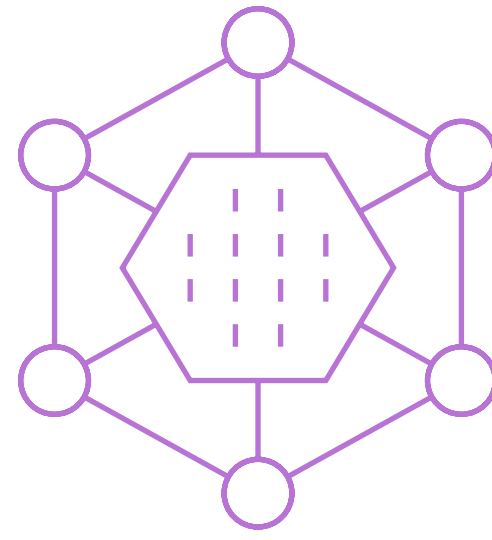
ÁMBITO DE LA SOSTENIBILIDAD

CONCLUSIONES

Se están utilizando estos modelos de operación y desarrollo altamente entrelazado, y de esta manera se aprende de lo que ha costado adoptar DevOps en las organizaciones y se asegura que las buenas prácticas de *MLops* se aplican desde el inicio del planteamiento del caso, los beneficios son:

- **Antes de plantear un caso.** Reusar conocimiento, herramientas y procesos y por tanto hacer más fácil que nuevos miembros del equipo comiencen a ser productivos. Con la escasez de buenos perfiles de ciencia de datos, la adopción de *MLops* de manera corporativa puede acelerar drásticamente los tiempos de aprendizaje para cada nuevo científico de datos que se una al equipo o empiece un caso nuevo
- **Durante la implementación del caso.** Simplificar la implementación de soluciones complejas, reduciendo la posibilidad de errores manuales y sobre todo asegurando un menor tiempo de implementación y por tanto aceleración en el proceso de crear valor
- **Después de la puesta en producción de caso.** Monitorización automática que asegura que los modelos ya siguen siendo útiles y lo suficientemente precisos para seguir aportando valor





3.2.3. Niveles de madurez

3.2.3.1. Nivel 1. Fundacional. Puesta en producción manual

El proceso de puesta en producción de los casos es manual y depende de los conocimientos y habilidades de los Data Scientist. Las organizaciones en esta etapa comienzan a definir e implementar prácticas, generalmente en forma de documentación, wikis o guidelines que los encargados de la puesta en producción deben seguir, pero no se comprueba de manera automática que así sea. Los errores son comunes y por supuesto la revisión y reentrenamiento de los modelos suele ser costosa y a menudo no se tiene en cuenta.

3.2.3.2. Nivel 2. Acercamiento. Puesta en producción sistemática y automatizada

La organización conoce y práctica MLOps, por lo que automatiza la puesta en producción y sistematiza pruebas y despliegues. No obstante, en este nivel de madurez, los MLOps se utilizan exclusivamente durante el ciclo de desarrollo del modelo de Inteligencia Artificial, es decir, durante el

entrenamiento, el empaquetado y la validación hasta el punto de implementación. Y tampoco de manera sistemática en todos los casos

3.2.3.3. Nivel 3. Aspiracional. Sistemas de monitorización

La práctica de MLOps que se integra completamente el ciclo de vida de Machine Learning de extremo a extremo, desde el desarrollo hasta la implementación, el monitoreo y el reentrenamiento (si es necesario) de manera fluida y eficiente. La mayoría de los casos implementados usan estos sistemas de monitorización y detección de pérdida de precisión o necesidad de reentrenamiento.

3.2.3.4. Nivel 4. Maduro. Desarrollo de soluciones propias

La solución de MLOps es un estándar de la compañía que forma parte de la plataforma de Data&AI y el 100% de los casos en producción cumplen con las políticas y de los sistemas de puesta en producción y monitorización de IA.

CASOS DE ÉXITO



3.2.4.1. Kernel

Los casos de uso que se desarrollan sobre Telefónica Kernel utilizan *big data*, Machine Learning y otras técnicas de Inteligencia Artificial, sobre los datos normalizados y utilizan el resto de capacidades para implementarse en producción. Con estos casos de uso se pretende maximizar la creación de valor, enfocándose en la optimización, gestión del valor de los clientes y marketing inteligente.

Los problemas más frecuentes que se presentan para la puesta en producción de modelos analíticos incluyen las diferentes configuraciones software de los entornos de producción, las configuraciones de privacidad y de acceso a los datos, la configuración de los nodos necesarios para el despliegue y ejecución de los modelos, etc.



Telefónica Kernel da solución a los anteriores problemas ofreciendo un entorno único de ejecución de modelos con una configuración software normalizada, un software development kit para el acceso a los datos, gestión centralizada de la privacidad, aislando de esta complejidad a los modelos y un entorno de ejecución auto escalable en función las necesidades de computación de los modelos.

3.2.4.

CASOS DE ÉXITO



3.2.4.2. Framework de industrialización MLOps

En Repsol el programa de transformación digital ha impulsado el desarrollo de cientos de modelos analíticos avanzados en los últimos tres años. Para poder hacer frente al reto de desplegar y operar este volumen han desarrollado un framework de MLOPS soportado sobre su propia Plataforma ARIA que permite desarrollar, desplegar y operar todos los modelos de la compañía de manera centralizada.

El objetivo de las prácticas de MLOps es la estandarización y la eficiencia de la gestión del ciclo de vida de los procesos de Machine Learning.

Todo el proceso de MLOPS está soportado por el módulo MLLab de la plataforma ARIA. Para la industrialización del despliegue y la monitorización de modelos, ARiA ofrece un conjunto de componentes que permiten:

- Dar cobertura del ciclo de vida del modelo (MLOps):



entrenamiento, registro y disponibilización de modelos analíticos avanzados y de IA.

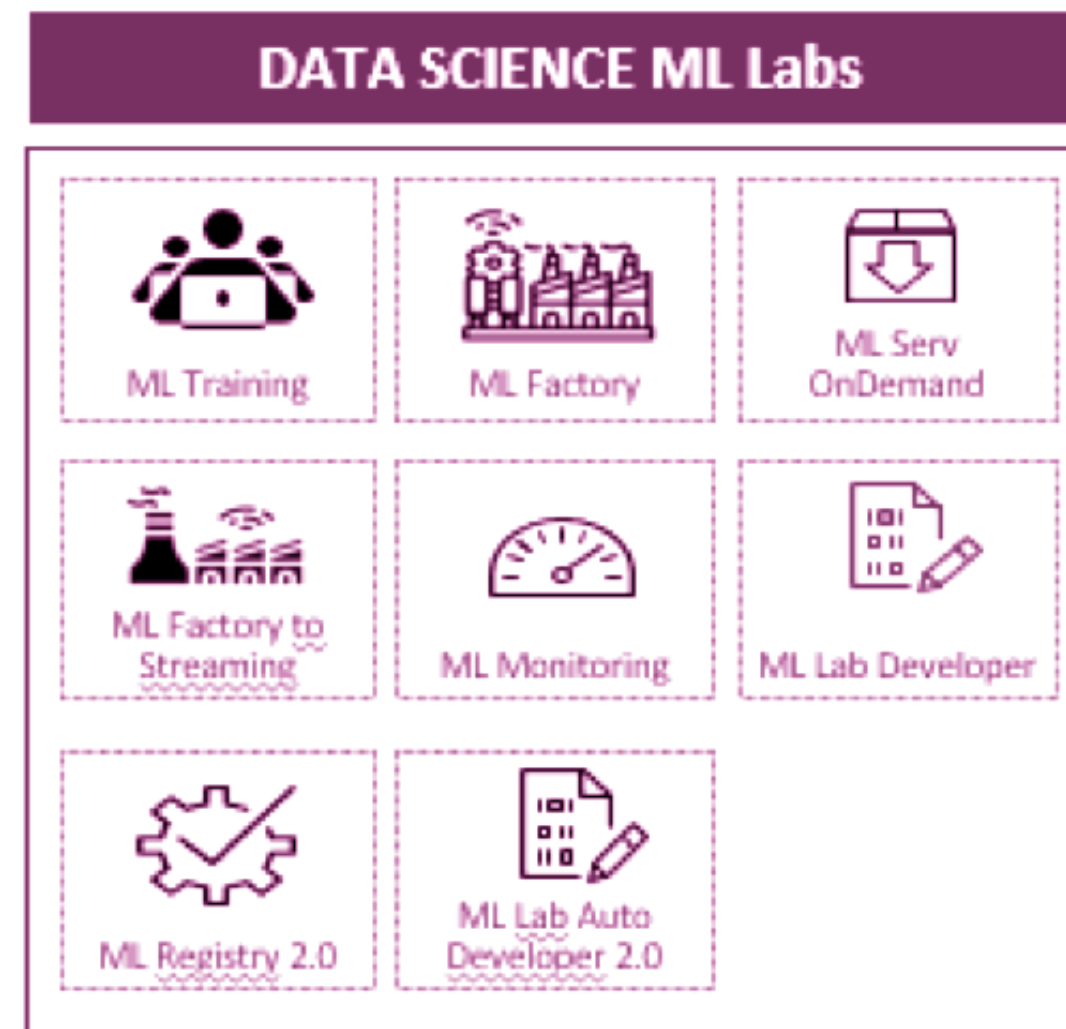
- Aprovisionar Sandboxes basados en Azure ML Studio con interfaz gráfico para usuarios ocasionales.
- Facilitar el Metadato de los modelos procesos analíticos y prevista la monitorización de los mismos.
- Aplantillar los modelos para despliegue rápido y homogeneización de la operativa de los mismos.
- Visualizar los modelos disponibles, así como sus dependencias.

3.2.4.

CASOS DE ÉXITO

Para la industrialización de la ideación, desarrollo, despliegue y la monitorización de modelos, ARiA ofrece un conjunto de componentes y buenas prácticas:

- **Know How (WiKi).** Las fases de desarrollo de un modelo analítico en ARiA están procedimentadas en su WIKI, determinando todos los pasos desde la ideación del caso analítico hasta su puesta en producción, pasando por las fases de exploración, alta, desarrollo, publicación, entrenamiento, ejecución o monitorización, y tiene las siguientes características:
 - Todo el conocimiento adquirido por Repsol evoluciona en su WiKi alojada en Azure DevOps.
 - Punto de consulta y sincronización común de los equipos.
 - Modo de transmitir el conocimiento y las buenas prácticas a la hora de crear casos.
 - Método de hacer más productiva la analítica de extremo a extremo.

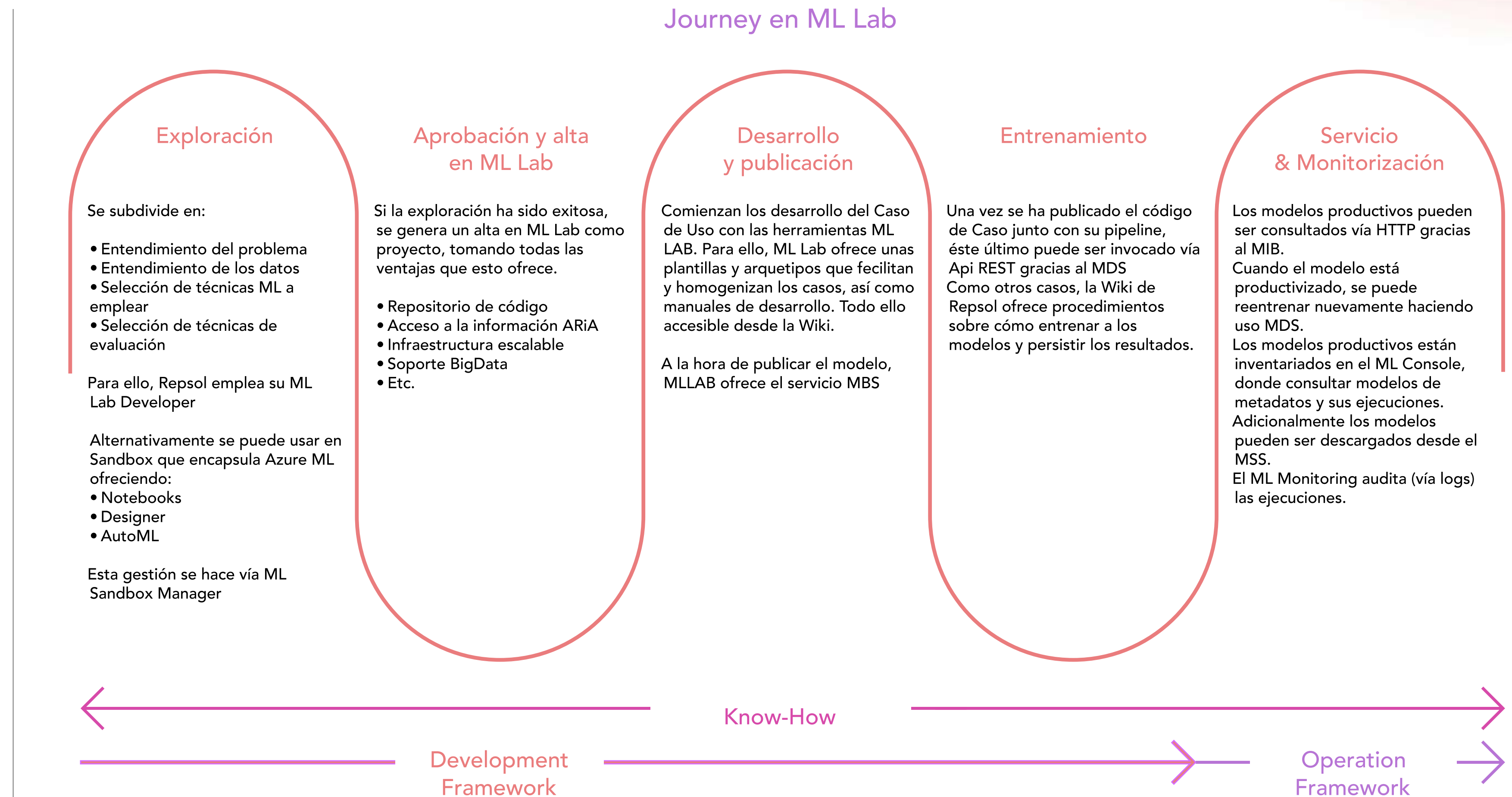


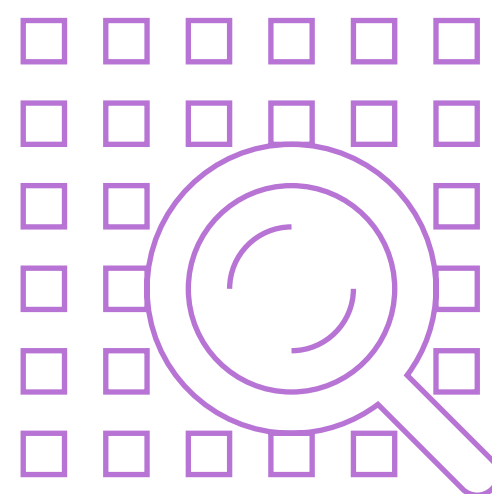
- **Development Framework.** Conjunto de herramientas y servicios que materializan la creación y desarrollo de los casos analíticos, basados en herramientas de Azure y desarrollos ad-hoc:
 - ML Lab Developer. Herramienta basada en Jupyterhub que permite a los Data Scientists crear sus Notebooks con los modelos analíticos que den respuesta a los casos de uso.
 - Model Build Service (MBS). Es un microservicio desarrollado con I, el cual está desplegado en el AKS de ML Lab. Registra los pipelines de entrenamiento de los modelos desarrollados por un caso de uso en una instancia de AML (Azure Machine Learning).
 - Model Deploy Service (MDS). Es otro microservicio con una orientación más operativa para el despliegue en el AKS de los modelos, ejecución de los pipelines de training de MBS, recuperación del estado de los entrenamientos, de las métricas de los trainings, de los logs de los trainings, despliegue de modelos entrenados y endpoints de AML, así como su monitorización o borrado, creación, eliminación y listado de triggers de DataFactory para lanzamiento de pipelines MBS

• Operations Framework

Conjunto de herramientas que permiten monitorizar y operar los modelos analíticos, así como su uso productivo.

Todas estas soluciones soportan el Journey de desarrollo y operación de modelos analíticos avanzados que se ha definido en Repsol y que se detalla en el siguiente gráfico:





4.1. Reto 5. Accesibilidad a los datos

4.1.1. Descripción del problema

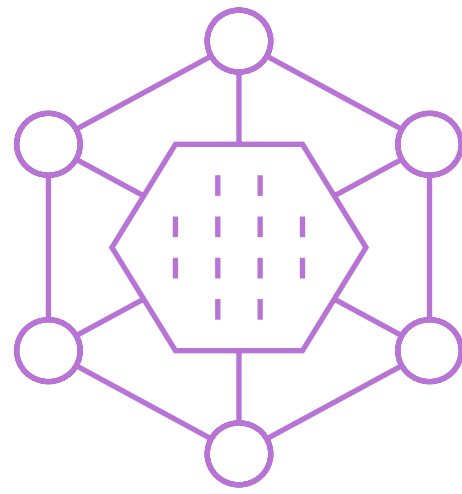
Para las organizaciones involucradas en el desarrollo de proyectos de Inteligencia Artificial la disponibilidad de los datos es uno de los principales problemas a resolver.

Como bien es sabido, sin datos no es posible hacer Inteligencia Artificial. En la mayor parte de las empresas los datos aparecen aislados o la calidad de los mismos no es la debida. Es fundamental, desde un principio, contar con una estrategia clara de recolección, almacenamiento y gestión de la demanda de los datos, de forma que a la hora de aplicar algoritmos de IA estos datos puedan ser extraídos de manera consistente, ordenada y segura. Así mismo, se ha de considerar el valor del uso de un espacio de datos compartido.

El reto de la accesibilidad a los datos conlleva una serie de preguntas que hay que responder para poder resolverlo de manera correcta:

- ¿Qué datos tengo?
- ¿Qué datos necesito? ¿cómo los voy a obtener?
- ¿Qué tipo de datos son (estructurados, no estructurados, nivel de seguridad...)?
- ¿Quién o quiénes son los dueños de los datos?
- ¿Dispongo de los perfiles necesarios para gestionar estos datos?
- ¿Qué tipo de infraestructura necesitaré? ¿Cloud, On-prem, híbrida?

En los puntos siguientes se desgranarán las soluciones existentes a día de hoy, los niveles de madurez de las tecnologías y algunos casos de éxito de algunos integrantes de IndesIA.



4.1.2. Soluciones

Las principales soluciones de *big data* para disponibilizar los datos están basadas en:

4.1.2.1. Data Lake

Una manera de clasificar y almacenar gran cantidad de datos de una forma organizada que facilita su análisis. Sus principales características son:

- Los datos se almacenan en un estado no transformado o casi no transformado.
- Todos los datos se cargan desde los sistemas fuente. No se rechazan los datos.
- Los datos se transforman y el esquema se aplica para satisfacer las necesidades del análisis.

4.1.2.2. Data Warehouse

Son estructuras destinadas a almacenar gran cantidad de datos de una forma integrada y organizada, permitiendo su análisis. Sus principales características son:

- Define un uso específico para los datos antes de su carga.

- Representa una imagen resumida del negocio organizado por área temática.
- Está altamente transformado y estructurado.

Las principales diferencias entre ambos son:

- El *Data Lake* conserva todos los datos y el *Data Warehouse* decide si almacena o no el dato.
- El *Data Lake* almacena información que no está preparada y lista para el consumo, sino que se recoge en estado natural.
- El *Data Lake* admite todos los tipos de datos, independientemente de su tipo, formato o procedencia.
- El *Data Lake* puede dar servicio a todos los usuarios de la organización.
- El *Data Lake* se adapta fácilmente a los cambios y es más ágil, mientras el *Data Warehouse* tiene un diseño complejo que dificulta la adaptación.

En función de los casos de uso y tanto el volumen como las necesidades de datos se podrá elegir la solución adecuada.

4.1.3. Niveles de madurez

Las principales soluciones de *big data* para disponibilizar los datos están basadas en:

4.1.3.1. Data Warehousing y soluciones de BI departamentales

Las empresas utilizan tecnologías SQL tradicionales normalmente on premise combinadas con herramientas de business intelligence para solucionar el reporte y hacer un análisis estático de la información.

4.1.3.2. Arquitecturas *big data* independientes por caso de uso

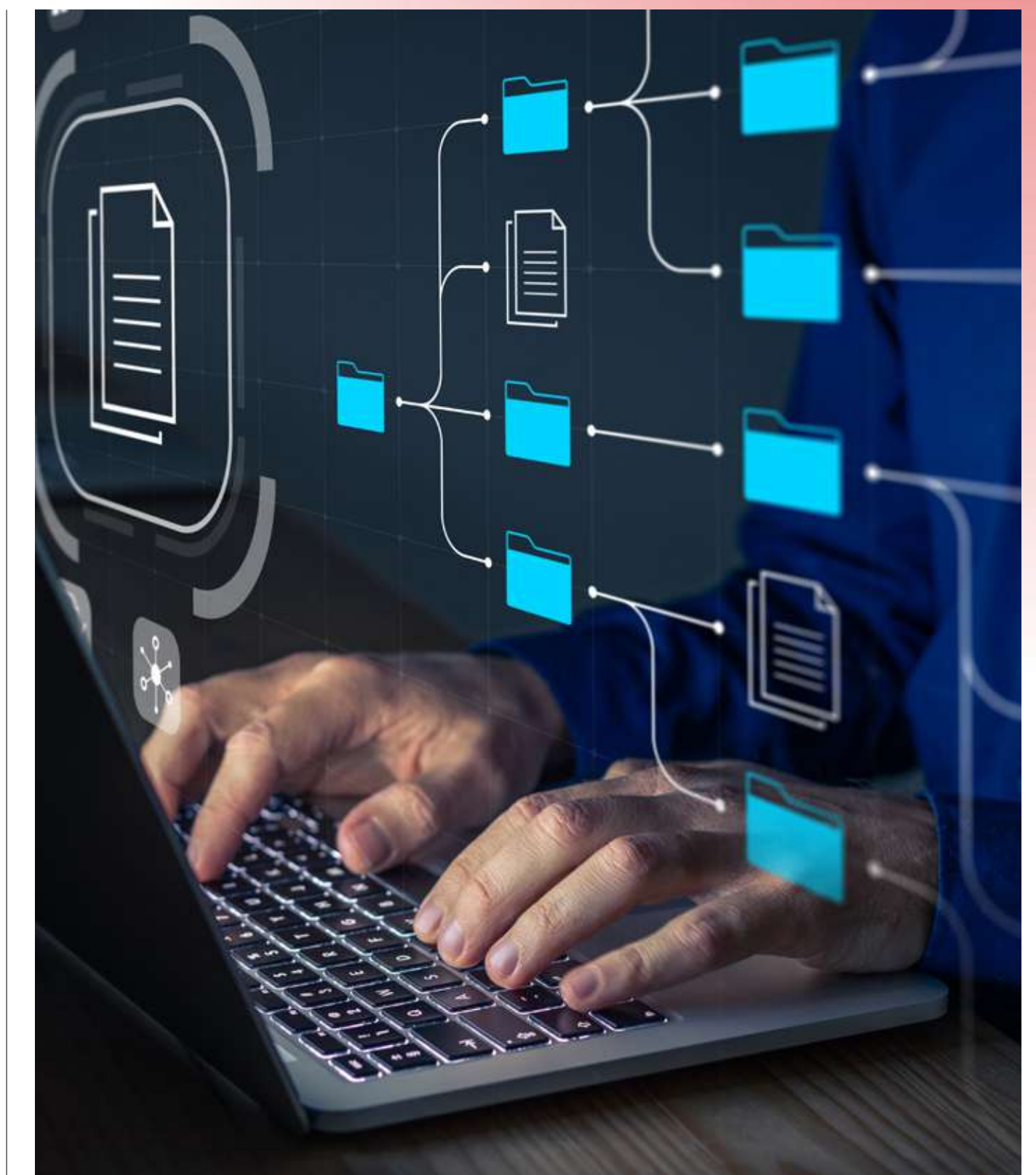
Según las compañías avanzan en su transformación digital van aumentando su necesidad de explotar el *big data* aumentando el volumen de datos, el tipo de fuentes no estructuradas como audio, video, logos o texto y analizando la información cada vez más rápido hasta llegar al tiempo real. Al principio lo realizan sobre casos de uso individuales en que cada caso tiene su propia arquitectura de *big data* con sus capas de extracción, almacenamiento y explotación analítica individuales. Estas nuevas infraestructuras suelen desplegarse en la nube.

4.1.3.3. Plataformas *big data* Corporativas

Según el volumen de casos de uso digitales aumenta las empresas que se ven en la necesidad de capturar sinergias para poder escalar un gran volumen de casos digitales analíticos con lo que comienzan a construir una plataforma única corporativa, transversal entre unidades de negocio y que soporte todos los casos de uso de manera que se reutilizan los componentes desarrollados y las ingestas realizadas con el objetivo de acelerar el time to value, disminuir el coste de los desarrollos y evitar la formación de nuevos silos.

4.1.3.4. Lakehouse corporativos

La democratización del dato que suponen las plataformas *big data* despierta la necesidad por parte de los usuarios de tener acceso a toda esa nueva información que está almacenada en los Data Lakes sin tener dependencias de las áreas técnicas para analizarlas. En esta fase los tradicionales DWH on premise se sustituyen por soluciones DWH en cloud que conviven con los Data Lakes y que mediante la combinación de catálogos de datos y herramientas de autoservicio analítico dotan de independencia a todos los empleados de la compañía para poder tomar decisiones en base a datos.



4.1.3.5. Espacio de datos compartidos

En su estadio más avanzado las organizaciones comienzan a fomentar el uso y el compartir datos con terceros para ampliar sus posibilidades y mejorar la calidad de su analítica con un mayor volumen de información. Lo hacen sin querer perder el gobierno y la soberanía del dato de una manera federada, fomentando el desarrollo de soluciones y casos transversales a múltiples organizaciones y que generan una nueva economía del dato.

4.1.4.

CASOS DE ÉXITO

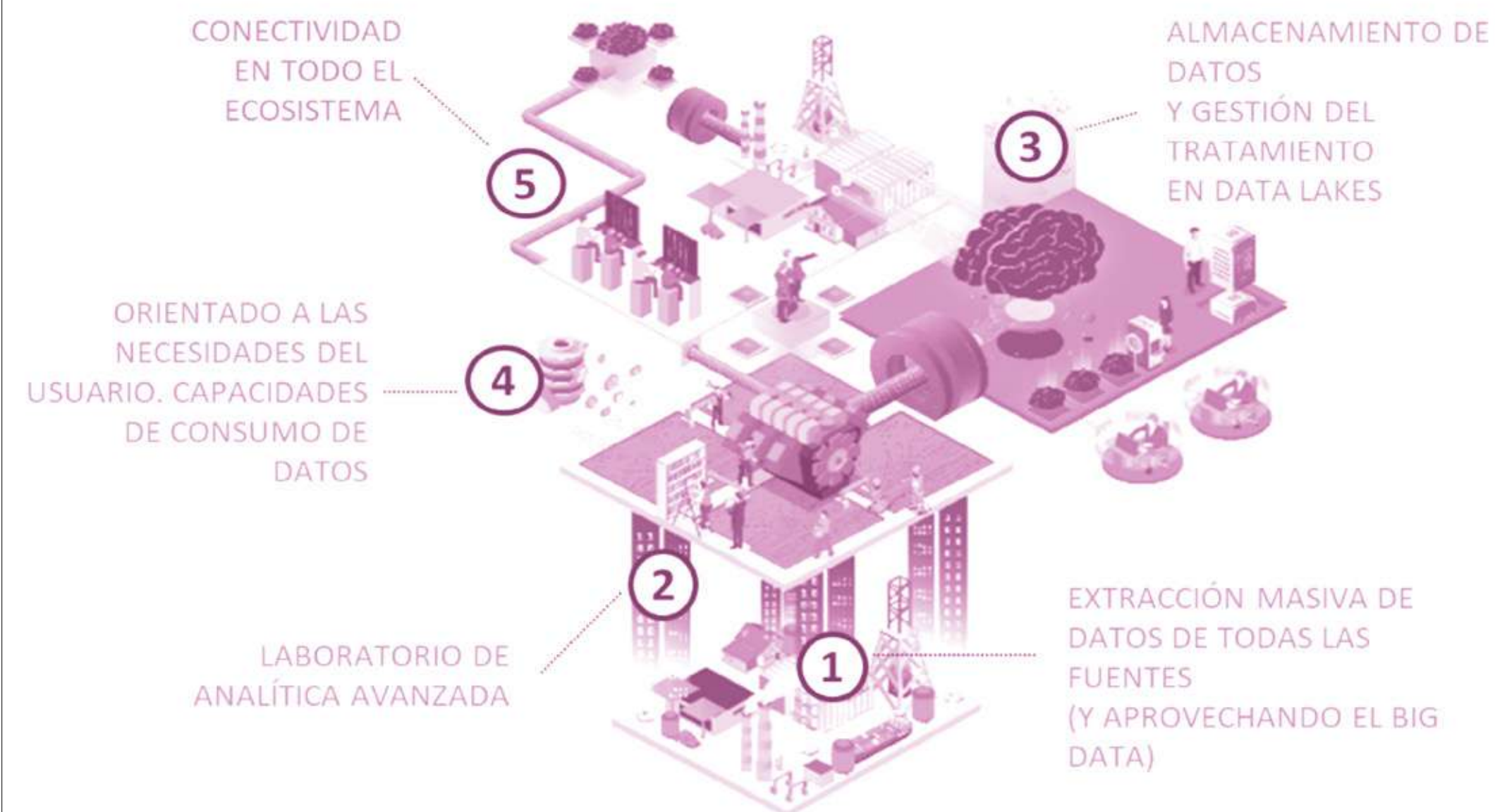


4.1.4.1. ARIA

Con el objetivo de ser capaz de acelerar el desarrollo de soluciones de Inteligencia Artificial y de escalar el uso de los datos y la analítica a la mayoría de los procesos en todas sus unidades de negocio Repsol decidió invertir en el desarrollo de su propia plataforma *big data* Analytics en Cloud: ARiA.

El objetivo de ARIA es ser el cerebro digital de Repsol. La plataforma consiste en un conjunto de componentes de servicios Cloud donde se integran los elementos para: extraer, procesar, custodiar, explotar la información con diferentes técnicas analíticas, de Machine Learning y de exploración de la información con consumos ágiles y flexibles.

ARiA está gobernado por un catálogo único que da consistencia a todo el ciclo y es el que permite realmente romper los silos de información. Este es el esquema de ARIA:



4.1.4.

CASOS DE ÉXITO

ARIA tiene una estructura muy cuidada. Los componentes técnicos desarrollados en la nube consisten en piezas de funcionalidad reaprovechable e independiente entre sí (se pueden cambiar sin que afecte al resto de componentes), y entiende las necesidades a resolver de cada caso de uso.

ARIA tiene más de 80 de componentes técnicos, con más de 70 casos analíticos, 100 fuentes de datos gobernadas, más de 80 TB de información en producción con más de 100 Data Steward y +50M operaciones en datos por semana en la nube que permiten democratizar el acceso al dato y la Industrialización del desarrollo de casos de uso de Datos & Analítica e Inteligencia Artificial.

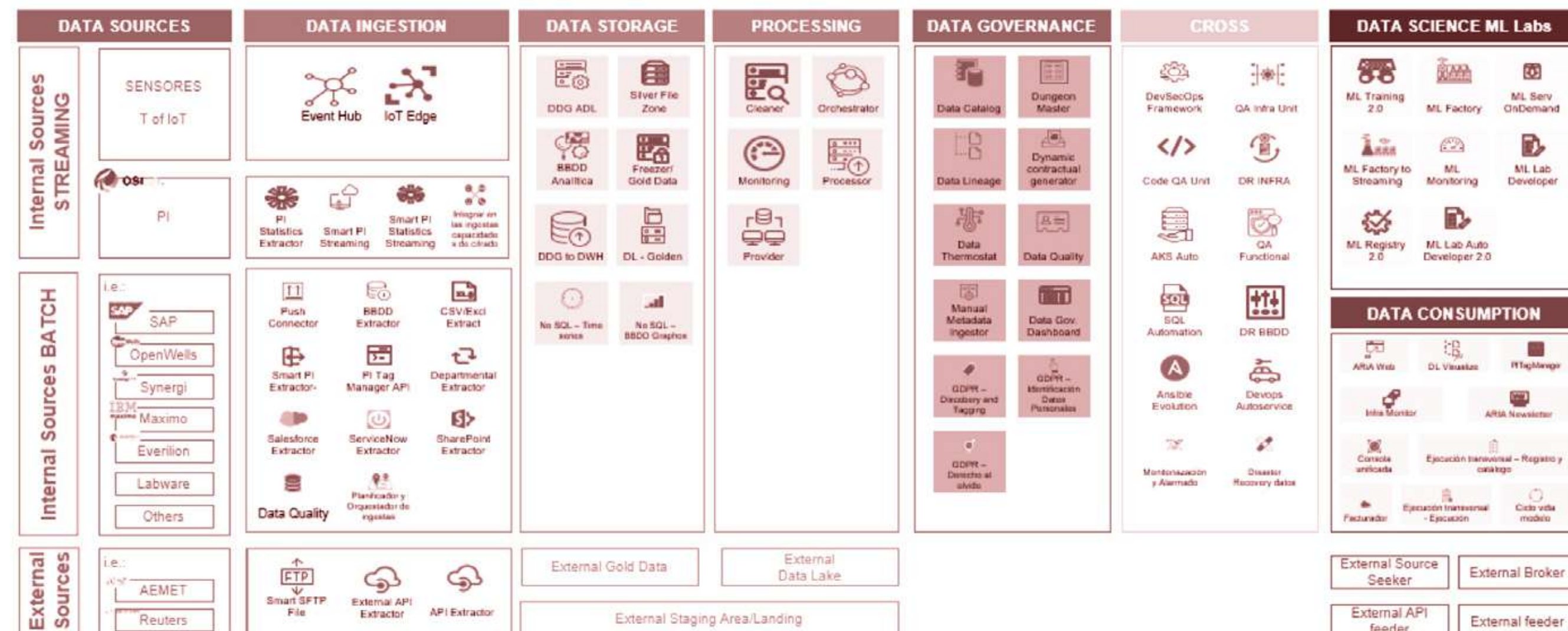
ARIA ha introducido una manera de trabajar en Repsol donde se aprecian, claramente, los distintos procesos que sigue. Son los siguientes:

- **Industrialización de ingestas.** Se habilita un modelo de trabajo tipo "factoría" para industrializar las ingestas,

respondiendo a los casos de uso y automatizando los datos que más se repiten en la compañía con tal de ayudar al time to value. Se crean y reutilizan conexiones de flujos de información "tuberías" conectadas a diferentes fuentes origen de datos.

- **Almacenamiento y procesamiento.** Los datos recopilados, se almacenan en Azure Data Lakes(1), que proporcionan una arquitectura de datos seguros y escalables basado en *big data*. Cada tipo de dato requiere almacenamientos diferentes que optimizan el uso de los recursos de computación en la construcción de las transformaciones de los casos de uso.

Hay un alto porcentaje de reaprovechamiento y optimización de Data Lakes, con interconexión entre sí para evitar silos de información en las consultas, consiguiendo una mayor celeridad en la obtención de los datos.



4.1.4.

CASOS DE ÉXITO

- **Gobierno del dato.** El diccionario de datos se pone a disposición mediante un catálogo en un modelo *Data Marketplace* de autoservicio con una gobernanza *end-to-end* y seguridad desde su recolección hasta su entrega para consumo del dato.

Se garantiza que los datos puedan usarse de la manera más flexible y efectiva posible para lograr su máximo valor con el adecuado acceso y en conformidad con la regulación o relación con terceras partes de la información que contiene.

Se catalogan y clasifican todos los datos para facilitar su consumo de acuerdo al propósito, para acelerar la identificación, ingesta y disponibilidad de los datos y permitir su reutilización.

De esta manera, se garantiza “que todo el mundo tenga acceso al dato adecuado, en el momento oportuno, con la calidad idónea y las garantías necesarias” como palanca de democratización del dato.

- **Cumplimiento en los objetivos del programa digital.** Consecución de los objetivos del programa digital de Repsol que acumula retornos de hasta 300M€ acumulados en 2020 de los casos digitales desde el inicio del programa digital, de los cuales 150M€ son específicos de los casos de uso desarrollados en base al uso de datos, analítica e Inteligencia Artificial.

Sobre ARiA se construye la mayor parte de casos digitales analíticos de Repsol:

- **Exploración y producción:** mantenimiento predictivo y modelos prescriptivos para la optimización de operaciones facilitando la asistencia remota y monitorización de operaciones, etc.
- **Industrial:** Predice condiciones inseguras en operaciones a través de IA, desarrollo de plantas autónomas apalancado en el ML, mejora la fiabilidad de los activos para optimizar el tiempo de actividad, etc.
- **Comercial:** permite una visión 360° que permite la ultra personalización, Pricing dinámicos y mejorar fidelización, ofrece una experiencia diferencial a través de modelos de Inteligencia Artificial mejorados, crea nuevos productos y servicios basados en datos, etc.

4.1.4.

CASOS DE ÉXITO

ARiA, además, se está utilizando en uno de los mayores retos que tiene en estos momentos Repsol. Y es la descarbonización. Así es como está contribuyendo a ello:

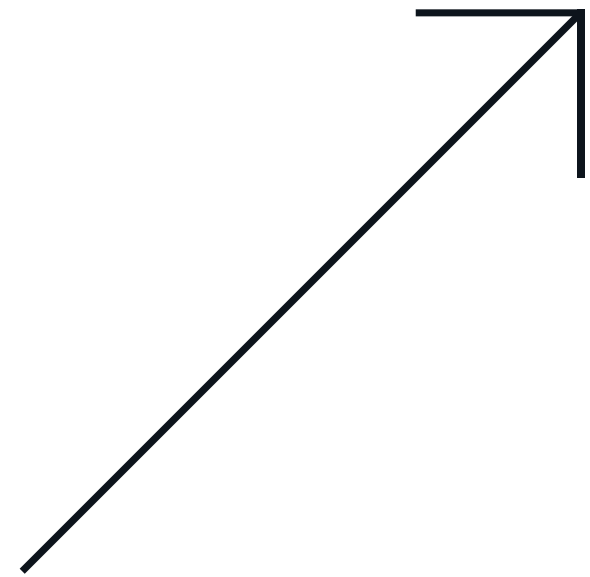
- A través del desarrollo de herramientas de simulación y escenarios *what if* que permitan el diseño de hojas de ruta hacia un futuro bajo de emisiones.
- Facilitando la medición y detección de emisiones contaminantes como el metano mediante el uso de modelos de visión artificial aplicados a imágenes satelitales de los pozos de petróleo.
- Mejorando la eficiencia energética de los activos mediante modelos de *Machine Learning* que permiten identificar desviaciones en los consumos energéticos de forma predictiva. Con esto, Repsol ha sido capaz de dejar de emitir 28.000 toneladas de CO2 durante este año.
- La aplicación ofrecerá recomendaciones a sus usuarios sobre las fluctuaciones que sufre el coste de la tarifa eléctrica, el histórico de consumos en el hogar y los propios hábitos personales de los consumidores.

Resumiendo. El objetivo de ARIA es ser el cerebro digital de Repsol, la plataforma analítica que permita facilitar el acceso a los datos, extraer conocimiento y desarrollar modelos y algoritmos que doten de la máxima inteligencia los procesos corporativos, industriales y comerciales de Repsol. Y, además, hacerlo de manera ágil, eficiente y a gran escala.



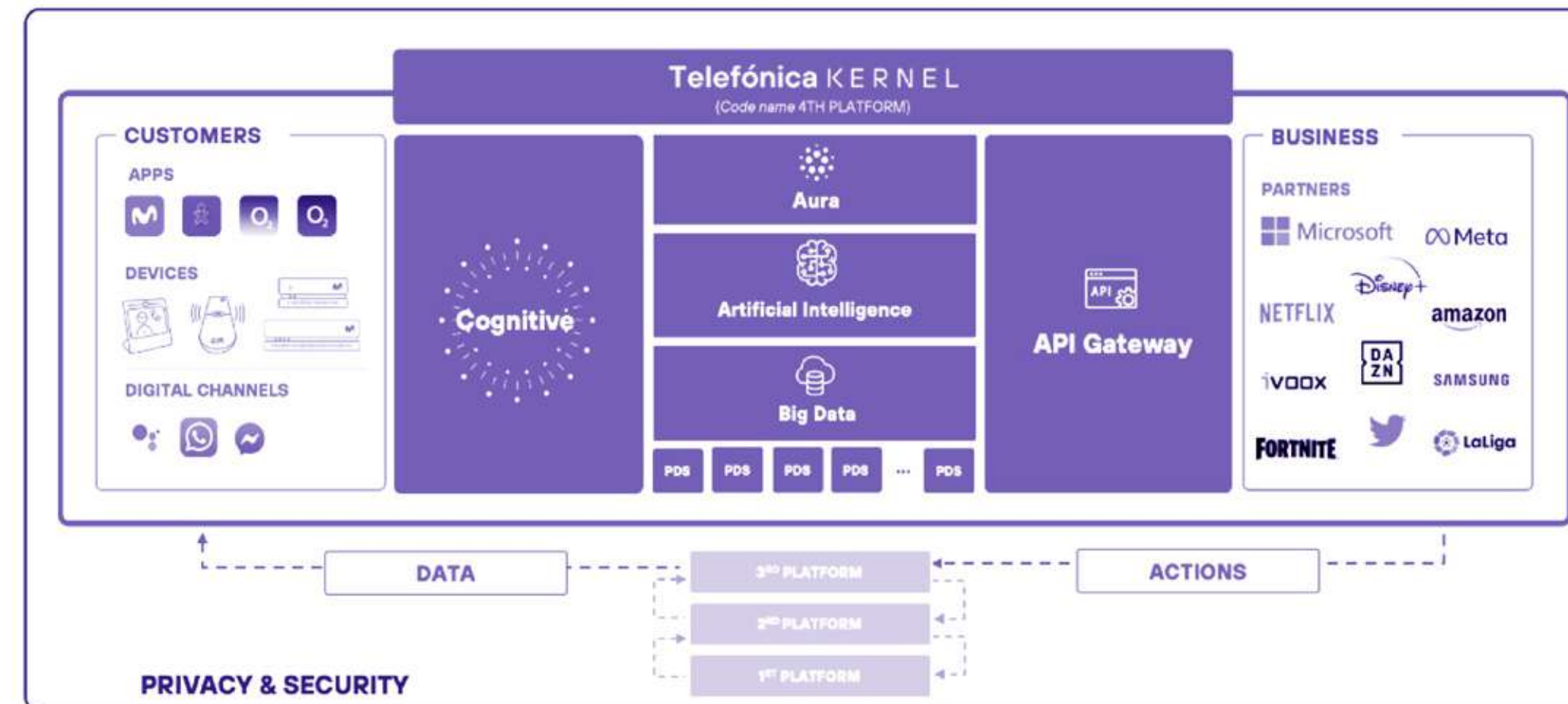
4.1.4.

CASOS DE ÉXITO



4.1.4.2. Kernel

Telefónica Kernel es el ecosistema digital sobre el que se apoya la transformación digital de Telefónica. Permite crear capacidades para desarrollar nuevos productos digitales poniendo al cliente en el centro y ofreciendo un control total sobre su experiencia digital.



Para conseguir todo eso, Kernel ofrece distintas capacidades de gestión de identidad, "apificación" (APIs) y de abstracción de datos que implementa bajo la máxima de privacidad-por-diseño. Con este ecosistema de capacidades que ofrece, se pueden desarrollar aplicaciones, productos y servicios digitales más rápido, acelerando así el time to market y habilitar la creación de productos que incorporan inteligencia artificial para enriquecer la experiencia del cliente en Telefónica.

Para la gestión de la identidad, Telefónica Kernel permite el acceso a datasets y a APIs, de manera controlada para los diferentes servicios tanto internos como externos. Kernel provee la habilidad de gestionar diferentes credenciales de los usuarios independientemente del canal que utilicen para acceder a esos datos y APIs.

4.1.4.

CASOS DE ÉXITO

Además, provee de **APIs normalizadas**, que facilitan de forma homogénea el acceso a los datos de los sistemas operacionales. Y orquesta la información que proveen los diferentes sistemas que se encuentran alojados en nuestras otras 3 plataformas (red, sistemas y servicios) usando un modelo unificado y estándar de datos **URM (Unified Reference Model)**. Así, los distintos productos, algoritmos, consumen estos datos normalizados cuando se despliegan en producción.

A partir de los datos normalizados, en Telefónica Kernel se genera el **CBD, Consolidated Business Data**, donde a través de un proceso de orquestación de dichos datos, se genera una capa de negocio, con todos los datos disponibles, generando la visión 360 del cliente. Es decir, convergen en esta plataforma, de manera digital, más de 400 *insights* que permiten a los algoritmos de IA no tener que preocuparse por la coherencia de los datos de negocio del cliente sino de la explotación y consumo específico, centralizando y democratizando el conocimiento del cliente de manera homogénea en todas las operadoras de Telefónica.

Así mismo, Kernel implementa la **privacidad por diseño** apoyándose en la utilización de propósitos estandarizados, información y catálogos de bases legitimadoras para el tratamiento de datos personales. De esta manera, cada vez que un producto o servicio hace uso de una API o un *dataset*, lo hace de acuerdo con el propósito y bases legitimadora aplicable de manera automática. Kernel gestiona la privacidad de dichos datos haciendo el filtrado, hashado o anonimizado de los datos según las bases legitimadoras acordes con GDPR que tenga cada producto en base al propósito de explotación de dichos datos. De esta manera se garantiza la privacidad de los clientes desde el diseño y por defecto en todos los P&S a través de sus sistemas y plataformas, definiendo los requisitos técnicos, estandarizando y automatizando el proceso de toma de decisiones de

Privacidad. Y se garantiza que una vez integrado en Kernel, funciona igual en todos los países y los productos no tienen que preocuparse de integrarse todas las veces en cada uno de ellos, ni de la privacidad de los datos, lo cual permite una gran escalabilidad y sostenibilidad de la productivización de productos basados en datos.

El resultado final hace posible la reutilización a gran escala de casos de uso en diferentes países. Eso permite una gran capacidad de desarrollo y despliegue, ya que se desarrolla una única vez y se puede desplegar en cualquier país con un coste mínimo de integración, y mejora del *time to market* exponencial. Por lo que se pueden operar todos los productos desplegados de una manera totalmente automatizada con total control garantizando la privacidad de los datos. En resumen, Telefónica Kernel nos permite crear capacidades digitales a partir de sistemas complejos para crear nuevas aplicaciones digitales que colocan al cliente en el centro, brindándoles un control total sobre su experiencia digital en todos los canales y dispositivos. Y reduce el tiempo necesario para lanzar nuevos productos y servicios digitales aprovechando capacidades estandarizadas en gestión de identidad, APIs, datos, privacidad y seguridad.

4.1.4.

CASOS DE ÉXITO



TECNICAS REUNIDAS

4.1.4.3. Proyecto TRDat@ arquitectura de plataforma corporativa *big data*

Técnicas Reunidas lanzó el proyecto TRDat@ en 2020 que articula el objetivo estratégico para llegar a ser una compañía Data Driven. La compañía es consciente del potencial valor que se puede extraer de los datos, pero cuenta con un gran número de sistemas de información que funcionan en su mayoría como silos de datos y que resultan complicados de gestionar de forma eficiente, entre los que se puede destacar:

- Sistemas transaccionales.
- Sistemas de ingeniería.
- Sistemas BIM y gemelos digitales.
- Sistemas de gestión con clientes, CRM.
- Fuentes de datos IIOT de la construcción.

Manejar las fuentes de información es complicado y actualmente requiere de un esfuerzo elevado para alimentar los informes operativos y estratégicos que extraen los datos cargados en cada uno de los sistemas.

En definitiva, mejorar la toma de decisiones de Técnicas Reunidas pasa por extraer el valor de los datos mediante un sistema centralizado de plataforma *big data* que unifica los silos de datos actuales, para que proporcione a los diferentes usuarios de negocio el poder contar con una plataforma

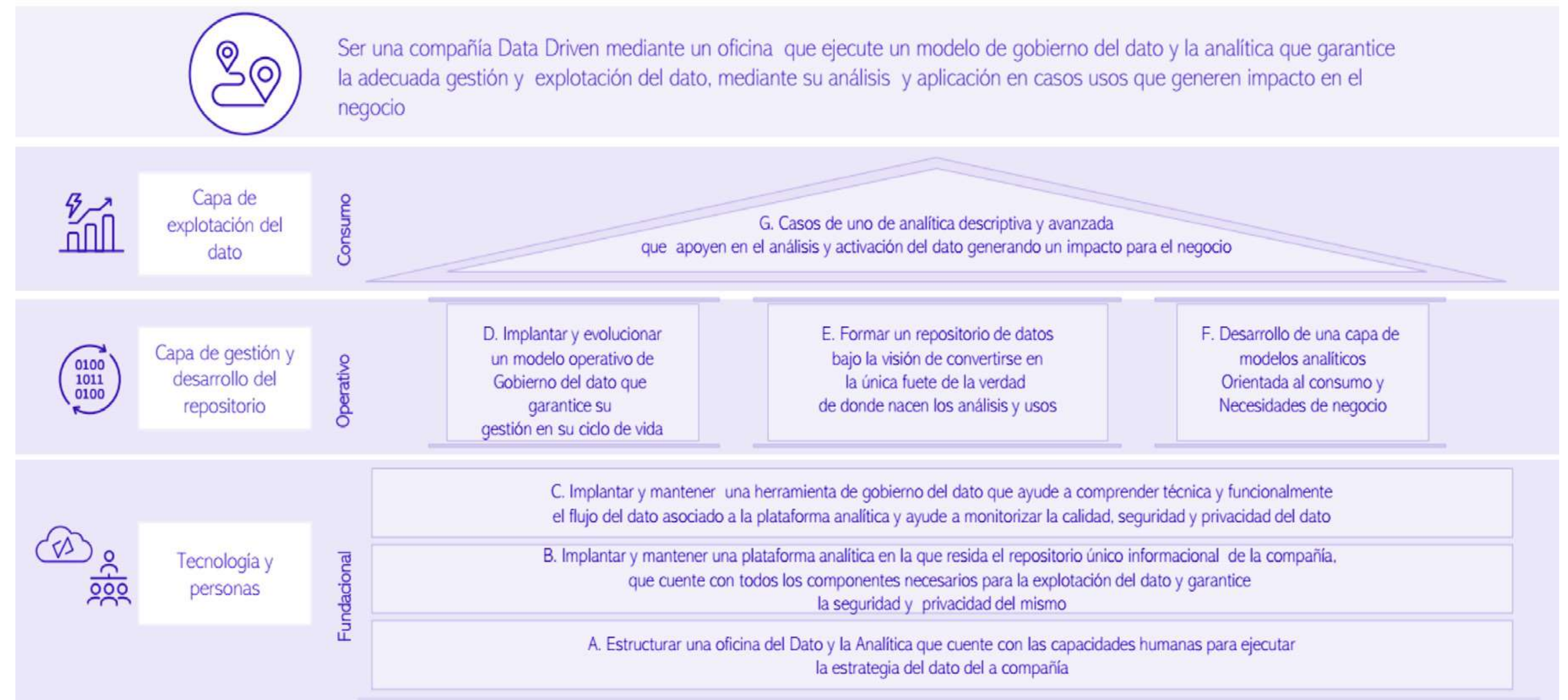
de datos de diferentes fuentes para el análisis y toma de decisiones ágil y con el correspondiente gobierno del dato. TRDat@ define el plan para el gobierno del dato y la arquitectura de la plataforma de datos corporativa que asegura a Técnicas Reunidas un acceso seguro y controlado de los datos, un común entendimiento de fuentes, dimensiones e indicadores y establecer los responsables de las definiciones y mantenimiento de los mismos, objetivos:

- Crear un repositorio único de información tipo Data Warehouse y/o Data Lake, con un esquema de datos único.
- Reutilizar/reaprovechar en la medida de lo posible los cuadros e informes ya desarrollados
- Mejora de la calidad de los datos y por lo tanto confianza en la información. Visión única del dato (estructura alrededor del dato).
- Unificación de mecanismos de tratamiento de la información ETL, gestión de datos maestros, transformación y explotación...
- Mayor grado de automatización de flujos de datos.
- Proactividad en el control sobre los datos para conseguir con ello un mayor control de calidad y trazabilidad de la información. Potenciación de políticas proactivas y habilitación de procedimientos preventivos y/o correctivos.
- Gestión de roles y acceso único.
- Analítica avanzada y predictiva (Machine Learning, IOT...).

4.1.4.

CASOS DE ÉXITO

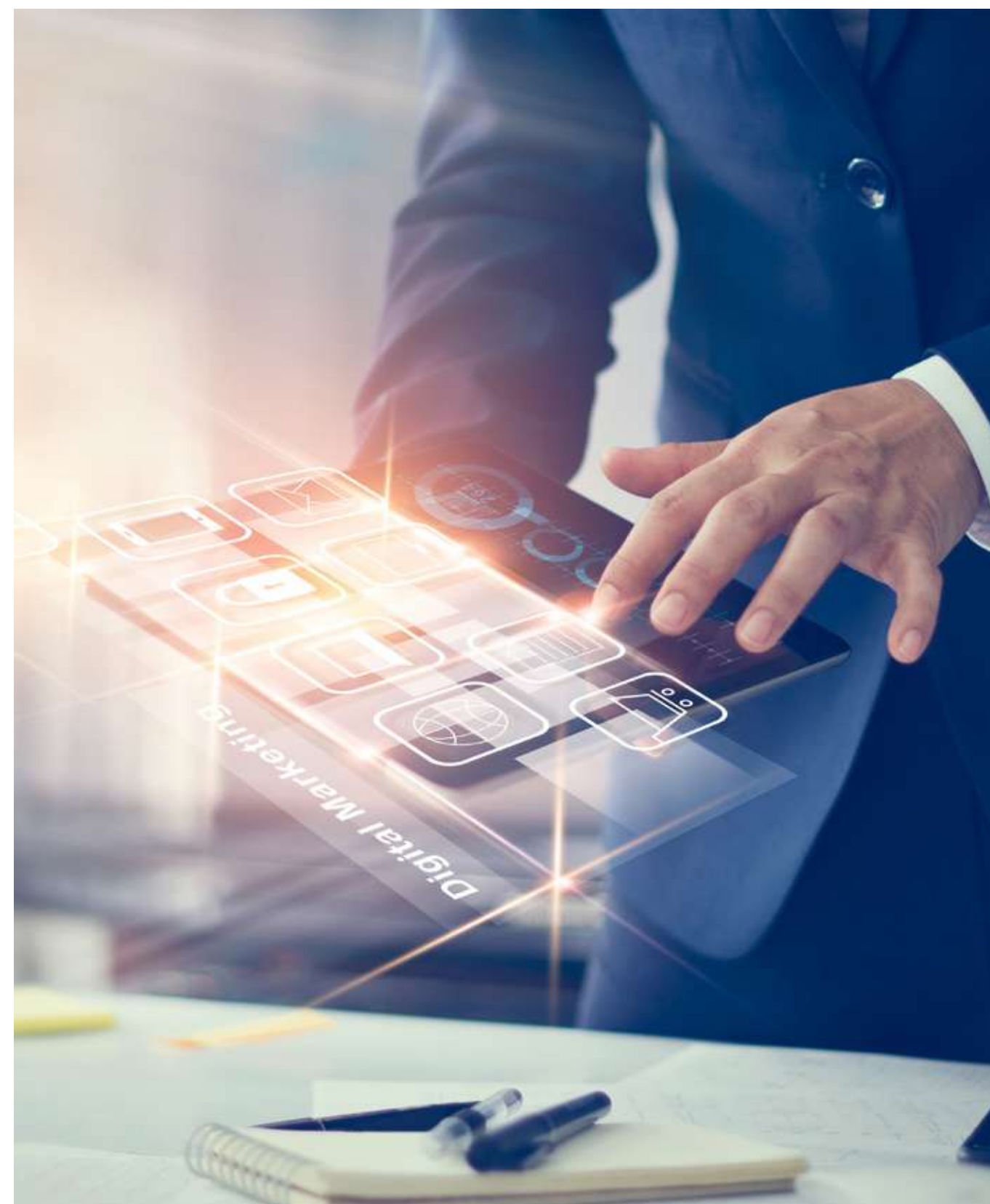
Estrategia de datos. Definición de líneas de trabajo y hoja de ruta para cada una de ellas: oficina del dato, plataforma, modelo gobierno, modelo operativo, ingesta, y analítica.



4.1.4.

CASOS DE ÉXITO

Así mismo, se ha definido la arquitectura física necesaria para la construcción y evolución de la plataforma analítica de la compañía, y que se está desplegando de forma progresiva para cubrir los casos de uso de datos. La plataforma y todos los componentes que son necesarios se desplegará durante los próximos años. A grandes rasgos la arquitectura global cubre los siguientes aspectos:



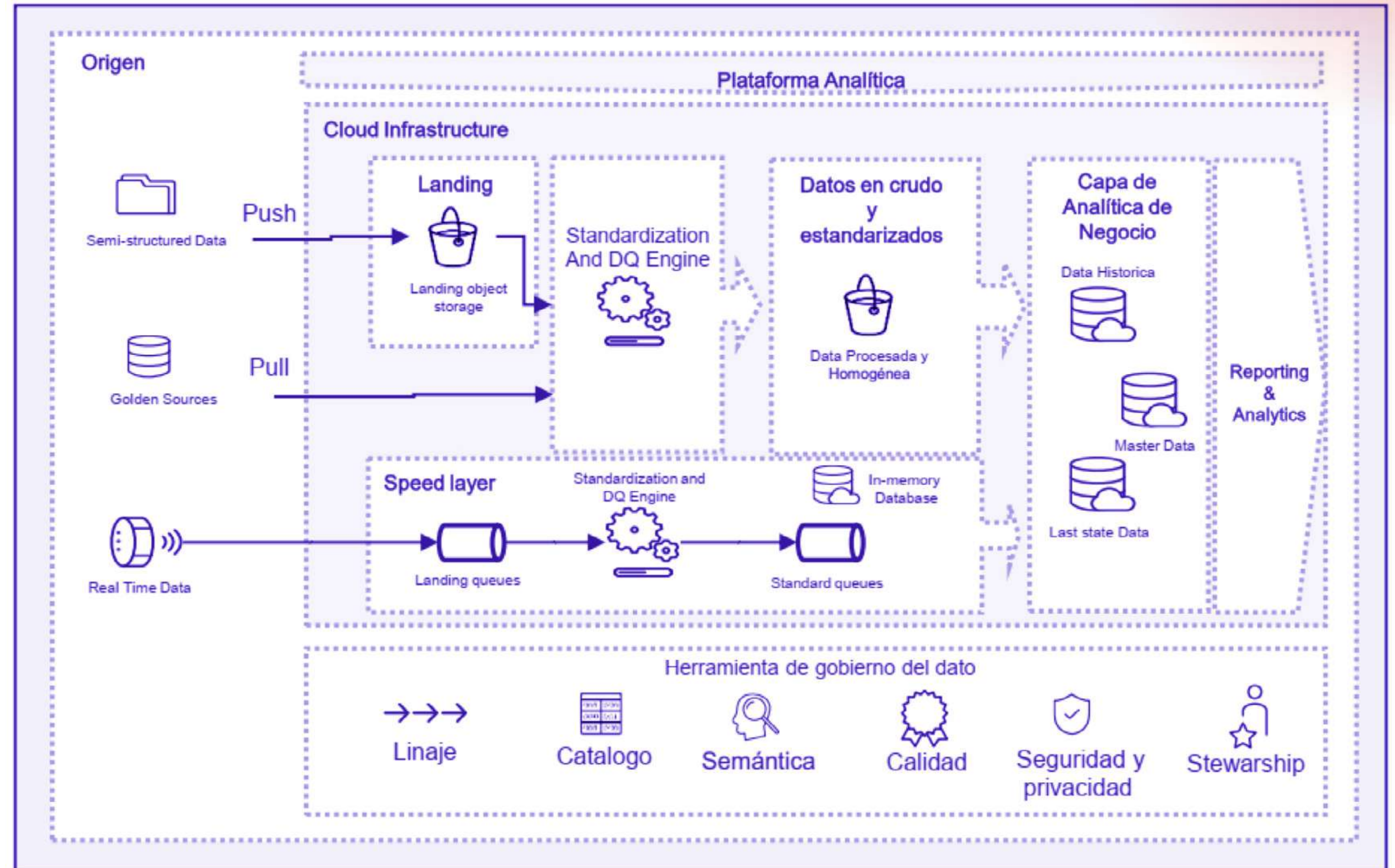
- **Gobierno del dato**
 - Origen y consumo de la información.
 - Seguridad y control de acceso.
 - Gestión de metadatos.
 - Gestión de datos maestros.
 - Semántica del negocio.
 - Gestión de calidad de los datos.
 - Mayor control sobre los indicadores de la actividad, de lo que se reporta y de lo que se comparte en la organización.
 - Poner en marcha una estrategia escalable de gobierno del dato.
 - Gestión de metadatos.
 - Gestión de datos maestros.
 - Semántica del negocio.
 - Gestión de calidad de los datos.
- **Transformación y almacenamiento de datos**
 - Limpieza de datos en origen: reglas de transformación, procesamiento de datos y cálculos avanzados.
 - Extracción del dato en plataformas comunes.
 - *Data Lake*: almacenamiento de datos flexible y escalable y/o.
 - *Data Warehouse*: agregados y procesamiento de *queries* en tiempo real.
 - Homogeneización del dato.
 - Catálogo de datos: metadatos y *KPIs*.
- **Capacidad analítica e inteligencia artificial**
 - Explotación de información mediante distintas plataformas de *reporting*: PowerBi, Qlikview, BO ANALYSIS...
 - Analítica avanzada.
 - *Machine Learning*.
 - Visión artificial.
 - Procesamiento del lenguaje natural.
 - *Deep Learning*.

4.1.4.

Plataforma de datos.

Modelo funcional y arquitectura de la plataforma de datos corporativa

Aproximación agnóstica a las funcionalidades de la plataforma

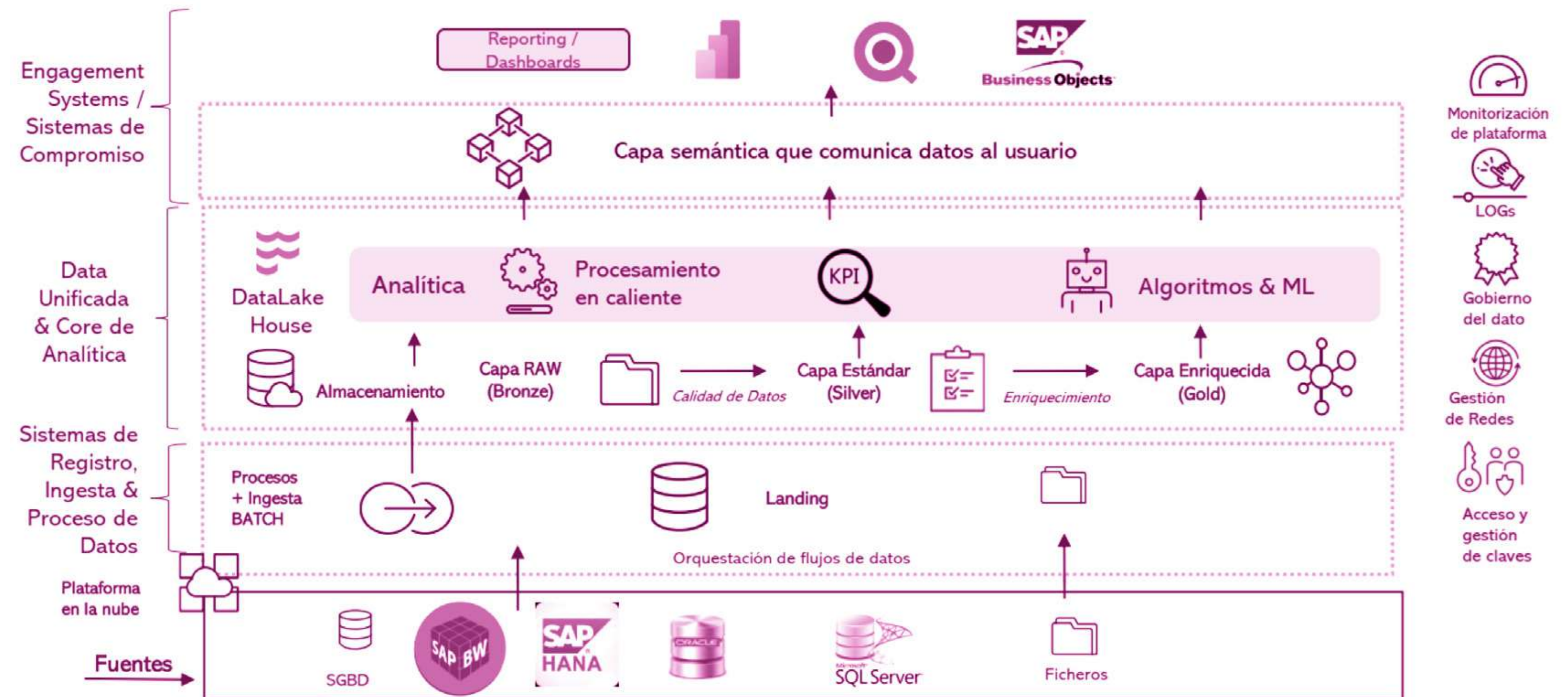


4.1.4.

CASOS DE ÉXITO

Accionar el valor de datos a través de la plataforma corporativa es un proyecto ambicioso y un reto para la compañía. Durante los próximos años se seguirá desarrollando el proyecto TRDat@ en todas sus dimensiones de acuerdo con la hoja de ruta definida para alcanzar el objetivo de compañía *Data Driven*.

Diagrama de componentes básicos



4.1.4.

CASOS DE ÉXITO

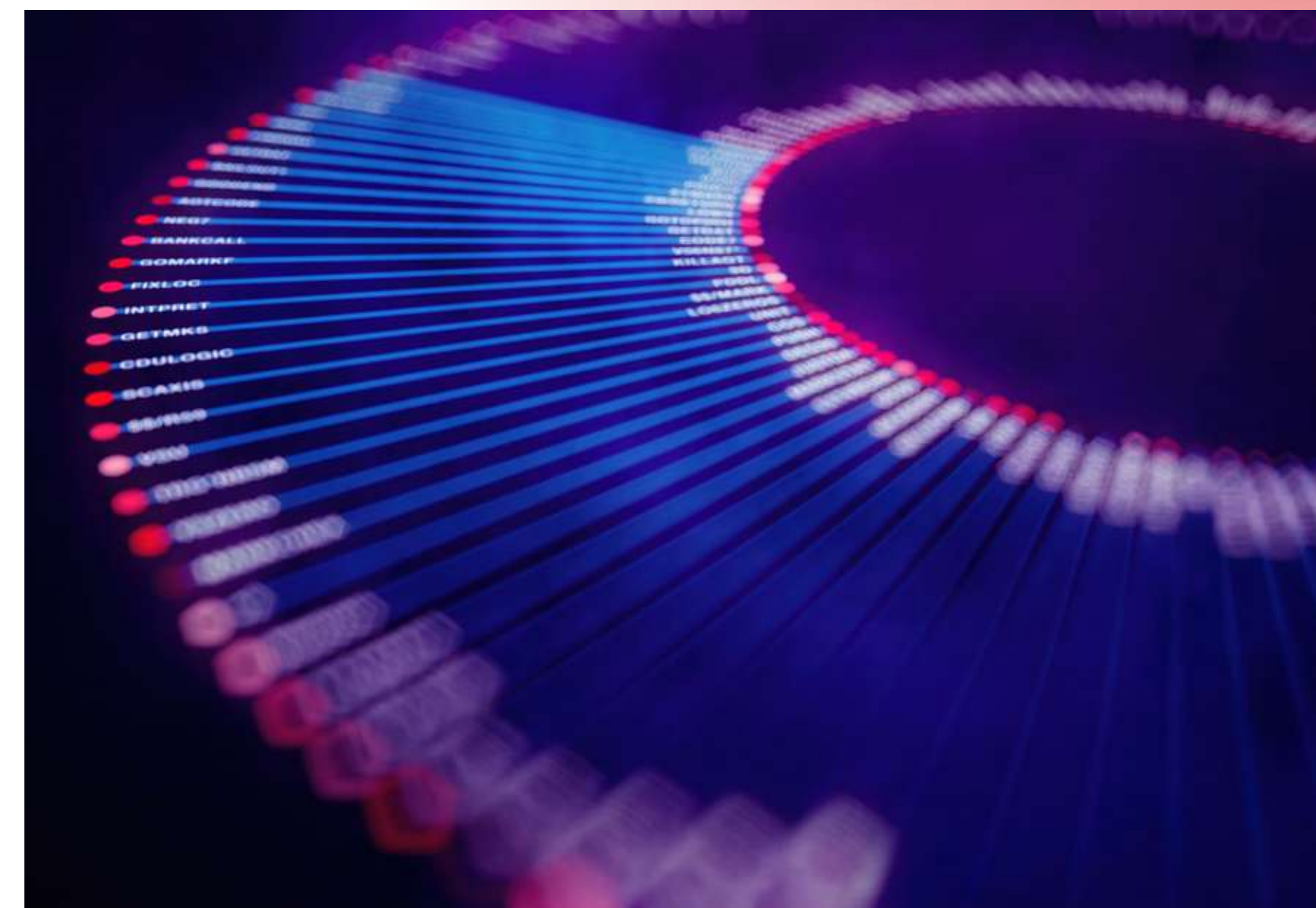
ferrovial

4.1.4.4. PANGEA

En Ferrovial han desarrollado un catálogo de datos "PANGEA" que permite de manera centralizada descubrir los diferentes dominios de datos que existen dentro de su organización, aunque éstos se encuentren almacenados de manera descentralizada en diferentes repositorios de datos (arquitectura "*data mesh*").

Los datos de carácter público, son visibles para toda la organización. Los datos de carácter sensible, son visibles solo para el conjunto de usuarios que tengan los roles adecuados (*data owners*) o a aquellos a los que se haya concedido permiso explícitamente a través de los procesos que gobiernan el acceso a los datos. Los algoritmos de Inteligencia Artificial, suponen un activo más dentro del catálogo.

Tanto los datos como los algoritmos se pueden describir mediante atributos que faciliten subúscue y descubrimiento.



El acceso a un dominio de datos concreto o a un algoritmo se implementa a través de diferentes interfaces en función del tipo de consumidor al que se quiere dar acceso. En este sentido, no es lo mismo pensar en usuarios externos que podrían llegar a consumir datos vía *APIs*, que en usuarios de negocio internos (acceso *self-service* o auto-descubrimiento), que en *Data Scientists* (acceso a los datos *raw*). El tipo de acceso y el procedimiento adecuado para cada uno de ellos debe adaptarse a cada tipo de usuario.

En organizaciones centralizadas, el acceso a los datos y algoritmos de Inteligencia Artificial resulta más sencillo. Pero en organizaciones muy descentralizadas, con datos esparcidos geográficamente en distintos países con políticas de protección de datos diversas, el acceso al dato por parte de un usuario y, especialmente el consumo cruzado de datos entre diferentes dominios es un reto que determina el diseño de la plataforma tecnológica.

4.2. Reto 6: Calidad y seguridad de los datos

4.2.1. Descripción del problema

El 50% piensan que la seguridad, privacidad y calidad de los datos están entre las 3 principales barreras para escalar el uso de la Inteligencia Artificial

El desarrollo de algoritmos de Inteligencia Artificial implica necesariamente el tratamiento de datos masivos, dentro de los cuales pueden estar incluidos datos personales y/o datos de carácter sensible. Estos datos son necesarios para nutrir al algoritmo de la capacidad de toma de decisiones en determinados contextos. Para algunas empresas, el acceso seguro a esos datos y la protección adecuada de los mismos es uno de los factores a tener en cuenta en el diseño de este tipo de soluciones.

Un ejemplo para ilustrarlo: si se quiere lanzar una campaña para los clientes y se quiere segmentar para adecuar mejor la oferta, es probable que haya que usar un algoritmo de IA que permita etiquetar a toda la base de datos de usuarios y extraer ciertas conclusiones que permitan posteriormente ajustar la campaña. Por ejemplo: los usuarios entre 20 y 30 años residentes en este distrito tienen un gasto medio de X € mientras que los residentes en tal otro tienen un gasto medio de Y €... Por lo tanto, puedo lanzar una oferta a unos más agresiva que a otros.

¿Quién desarrolla ese algoritmo? ¿Qué hace con esos datos de carácter sensible el algoritmo? ¿Copiarlos a unos servidores alojados en otro país? ¿Qué regulación cumple? ¿Existe algún contrato con garantías respecto a nuestros datos? ¿Cómo podemos protegerlos?, etc. Son algunas de las preguntas a resolver.

Otro de los factores indispensables a la hora de escalar el uso de IA es la calidad del dato. La mayoría de los modelos de Inteligencia Artificial “aprenden” de manera desatendida a partir de un histórico de datos ya existente. Esto significa que, si los datos base no son de calidad, el algoritmo final tampoco lo será. Por lo tanto, a la hora de desarrollar una solución de este tipo, hay que analizar previamente si los datos de partida son lo suficientemente buenos como para que los algoritmos creados cumplan unos mínimos umbrales de aceptación asumibles por el caso de uso al que están destinados.

4.2.2. Soluciones

4.2.2.1. Programas de gobierno de datos. Framework de calidad

Los programas de gobierno del dato, se ocupan, entre otras cosas, de establecer una serie de procedimientos y buenas prácticas para garantizar la seguridad y calidad de los datos en una organización. El *framework* DAMA es un estándar internacional en gestión eficiente de datos que puede servir a cualquier organización como punto de partida a la hora de implementar procesos que aseguren la calidad y la seguridad de su información.



Si nos centramos en la parte específica de calidad del dato, se puede tomar como referencia el siguiente documento de carácter público, compartido por DAMA en septiembre, 2020: **Dimensions for Data Quality** donde se explican algunas de las métricas más importantes a tener en cuenta a la hora de medir la calidad de un dato o de un dominio de datos: *Completeness, Accessibility, Availability, Compliance with laws, regulations or standards, Interpretability, Latency, Redundancy, Traceability, etc.*

Sin embargo, cada organización debe adaptar el estándar a sus políticas internas específicas y a la manera en la que están organizados sus datos y los responsables de los mismos.

Por ejemplo, si se quiere tener una visión homogénea y estándar de la calidad de los datos de la organización, hay que llevar a cabo las siguientes actividades:

1. Identificar qué personas o roles van a ser responsables de definir los procesos de calidad homogéneos.
2. Definir cómo se van a implantar sobre los datos existentes o nuevos (plan de despliegue del *framework* de calidad).
3. Validar cómo se va a adaptar el estándar a cada caso de uso (umbrales aceptables en función del contexto o actividad de negocio).
4. Planificar cómo se van a automatizar los procesos de medición de calidad
5. Definir cómo se van a clasificar en función de su criticidad y remediar los problemas de calidad de los datos.

4.2.2.2. Programas de gobierno de datos. Framework de seguridad

La seguridad de los datos es el proceso de mantener la confidencialidad, integridad y disponibilidad de los datos de una organización de una manera consistente con la estrategia de riesgo de la organización.

Un *framework* de seguridad, contiene una serie de procesos documentados que definen políticas y procedimientos en torno a la implementación y gestión continua de los controles de seguridad de la información. Estos marcos son un modelo para gestionar el riesgo y reducir las vulnerabilidades.



Un *framework* de seguridad recoge diversas actividades que toda organización debería implementar para proteger sus datos:

1. Definición de una arquitectura de datos segura (on premise, on cloud o híbrida).
2. Elaboración de un plan de respuesta ante posibles ataques.
3. Elaboración de un plan de recuperación ante incidentes.

Al igual que en el caso de los procesos de calidad, es bueno usar un *framework* ya conocido como punto de partida y adaptarlo a las necesidades específicas de cada organización, al tipo de datos a tratar y al uso final que se vaya a hacer del dato. Los ISO 27000 Series, suelen ser un estándar habitual en muchas organizaciones para diseñar sistemas seguros.

Pero más allá de los *frameworks* de seguridad habituales, la gestión de datos exige el cumplimiento de la regulación vigente en cada país o zona donde se aplique.

INDESIA

Framework Data & IA
Industrial Marco de
referencia de datos e
inteligencia artificial para el
sector industrial español

INTRODUCCIÓN

ÁMBITO DEL VALOR PARA NEGOCIO

ÁMBITO DE LA CIENCIA DE DATOS

ÁMBITO DE LOS DATOS

ÁMBITO DE LA ORGANIZACIÓN

ÁMBITO DE PERSONAS Y CULTURA

ÁMBITO DE LA SOSTENIBILIDAD

CONCLUSIONES

Pero más allá de los *frameworks* de seguridad habituales, la gestión de datos exige el cumplimiento de la regulación vigente en cada país o zona donde se aplique.

Así en Europa existe un marco regulatorio (GDPR), de carácter obligatorio, que las organizaciones deben implementar para proteger la seguridad y privacidad de la información personal de los ciudadanos de la UE. Los requisitos de la GDPR incluyen controles para restringir el acceso no autorizado a los datos almacenados y medidas de control de acceso, como privilegios mínimos, accesos basados en roles y autenticación multifactor. La recogida de datos a través de un interfaz web o móvil, también exige el cumplimiento de una política de privacidad para garantizar la seguridad y el uso legal de los datos personales recogidos y cumplir con los siguientes principios:

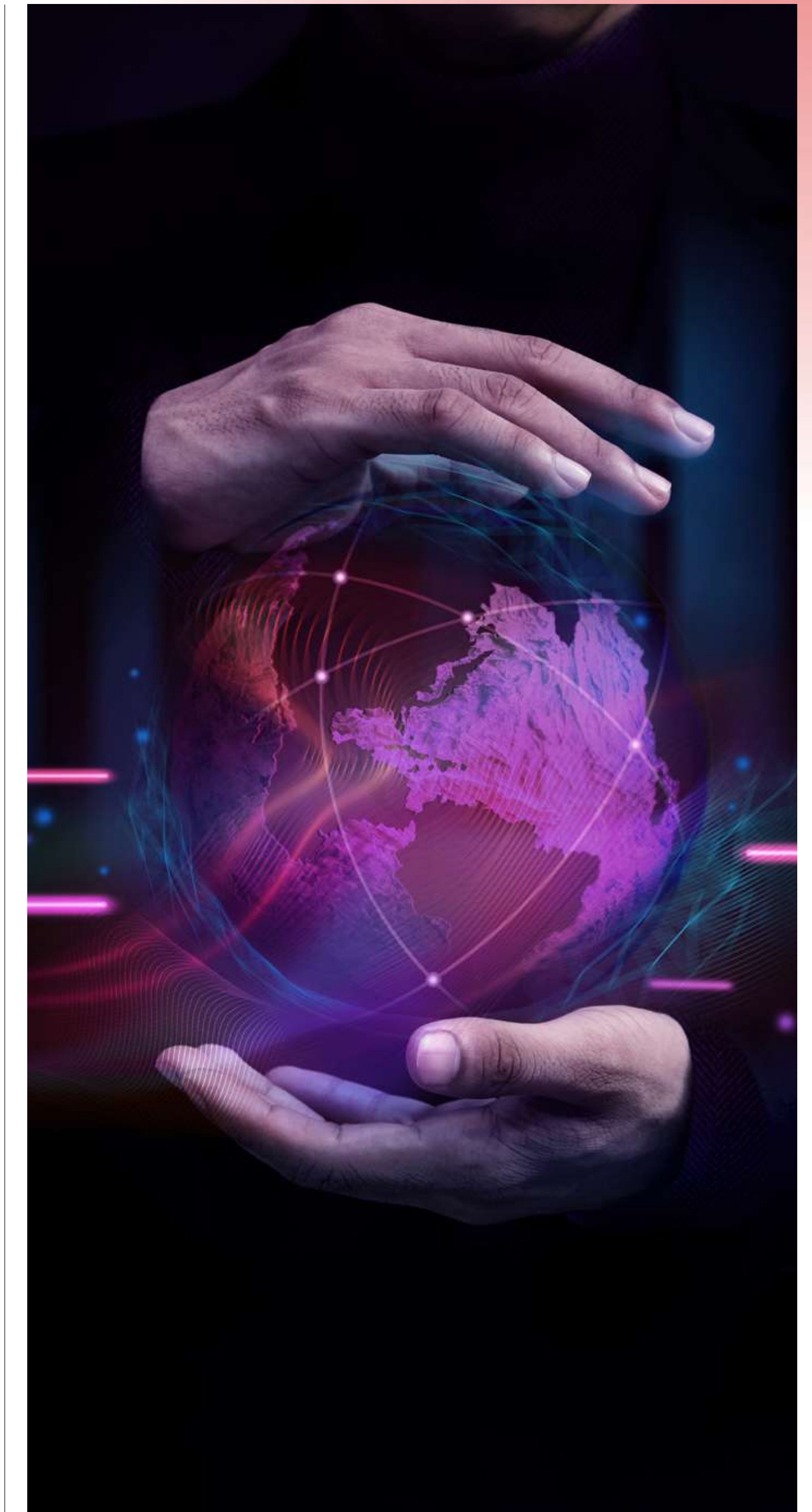
Principio de transparencia lealtad y licitud. Se necesita el consentimiento expreso del usuario para poder realizar el tratamiento de datos personales.

Principio de minimización. Solo se recabarán los datos imprescindibles para poder llevar a cabo los servicios solicitados por el usuario.

Principio de limitación del plazo. Los datos personales se destruirán una vez se haya cumplido la finalidad para la que fueron recabados.

Principio de confidencialidad e integridad. Se debe asegurar que se cumplen y toman todas las precauciones para que terceros no autorizados no puedan acceder a los datos.

Se puede leer más información aquí: [Política de privacidad web 2021-2022. Cumple el RGPD | Grupo Atico34 \(protecciondatos-lopdc.com\)](#)



4.2.3. Niveles de madurez

4.2.3.1. Nivel 1. Fundacional. No hay gobierno

No existe un modelo de gobierno definido.

4.2.3.2. Nivel 2. Acercamiento. Gobierno a nivel teórico y de documentación

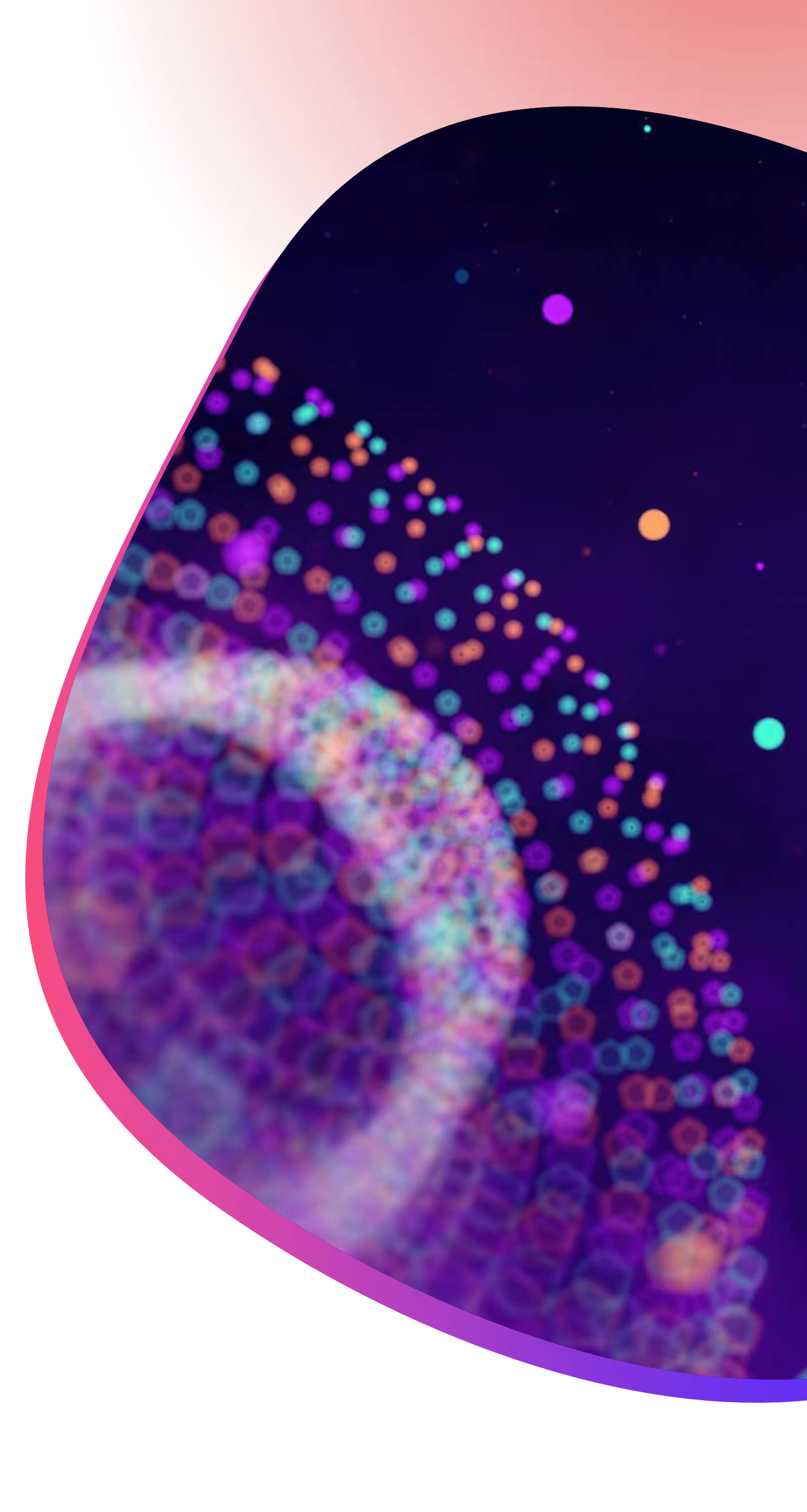
Además de las prácticas, comienza a ser clara la importancia de contar con modelos de gobernanza que abarquen de principio a fin el proceso de implementación de la RAI. Sin embargo, estos modelos de gobernanza no suelen implementarse en esta etapa.

4.2.3.3. Nivel 3. Aspiracional. Implementación del modelo de gobierno

Se definen e implementan modelos de gobierno que comprenden todo el ciclo de desarrollo de los casos y esas directivas o políticas forman parte de la estrategia de puesta en producción y monitorización (MLOps).

4.2.3.4. Nivel 4. Maduro. Gobierno y ética de IA a nivel corporativo

En esta etapa, es de esperar que la visión sobre los modelos de gobierno incorpore una visión ética y de Inteligencia Artificial responsable. Adicionalmente a este nivel de madurez el gobierno suele ser público y conocido por proveedores y colaboradores para asegurar que cualquier iniciativa que involucre datos o casos de terceras partes también cumplen el modelo de gobierno definido.



4.2.4.

CASOS DE ÉXITO



4.2.4.1. Modelo de gobierno

Para asegurar una transformación digital exitosa es necesario gobernar dichos activos digitales mediante la implantación de un marco de gobernanza que se ajuste a la organización de la compañía, al negocio y a los planes estratégicos de la misma. Esto permitirá que la organización libere por completo el potencial de sus datos.

Repsol está formado por una gran diversidad de negocios y funciones que ofrecen una rica heterogeneidad en la gestión de sus datos. Todos los días se toman decisiones con ellos, pequeñas y grandes y, lo llevan haciendo así desde hace mucho tiempo.

Desde el programa digital y, específicamente, el Hub de datos, analítica e IA se acompaña, refuerza y facilita su capacidad de explotar los datos mediante técnicas analíticas avanzadas e inteligencia artificial, garantizando su unicidad, calidad y seguridad. También se garantiza el uso de los mismos de acuerdo a principios éticos y respetuosos con su confidencialidad /privacidad, las diferentes regulaciones de aplicación y acuerdos contraídos con terceras partes con las



que nos relacionamos en el ejercicio de nuestras actividades.

El equipo de *Data Governance for Analytics* busca garantizar que los datos de una organización puedan usarse de la manera más flexible y efectiva posible para lograr su máximo valor en conformidad con las políticas de Repsol y con terceras partes.

El modelo de gobierno de los datos para analítica comenzó sus primeros esbozos en el 2019 con una clara vocación de convertirse en un mecanismo que permitiera la “democratización del dato” como una palanca de valor a los negocios/ casos de uso en la aceleración del uso/ re-uso de los datos con las adecuadas garantías para su escalabilidad y sostenibilidad en el tiempo.

El desarrollo del modelo está estructurado por releases, que desarrollan y diseñan los mecanismos que lo soportan al mismo tiempo que lo aplican ponen en funcionamiento en los casos específicos. En release 1.X, se han centrado en los siguientes ejes fundamentales:

4.2.4.

CASOS DE ÉXITO

La construcción de un modelo federado distribuido de gobierno del dato sencillo, sin jerarquías, basado en la colaboración y que distribuye las responsabilidades del dato en toda la compañía donde nuestra figura protagonista es el *data steward*.

¿Cómo se ha desplegado?

- Con un marco de gobierno transversal y estandarizado que se acciona por negocio o función corporativa, y por cada caso de uso, donde se establecen los roles y responsabilidades alrededor del dato.

La **información** es un activo de la compañía y nuestra capacidad de descubrimiento y explotación nos orienta hacia la eficiencia y la diferenciación.



El **data governance for analytics** busca garantizar que los **datos** de una organización puedan usarse de la manera más **flexible** y efectiva posible para lograr su máximo valor en conformidad con las políticas de Repsol y con terceras partes.

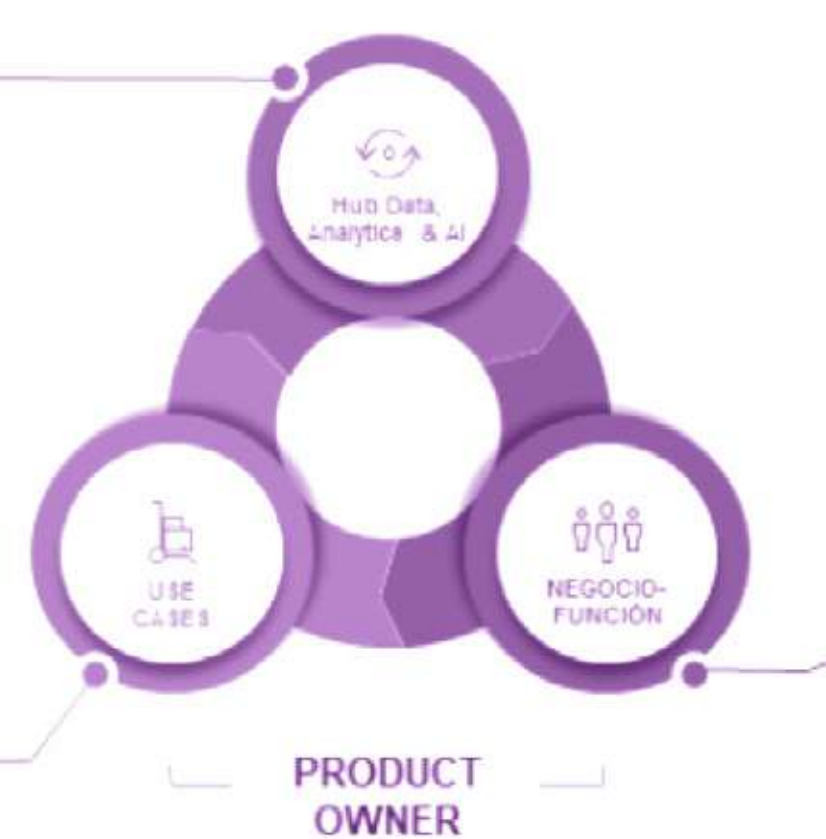
El modelo federado se basa en la **coordinación centralizada (HuB)** junto con una **ejecución descentralizada (NEGOCIOS)**.
La interlocución HUB-Negocio es altamente colaborativa, basada en los datos y **centrada en casos de uso**.

DATA GOVERNANCE FOR ANALYTICS OFFICE

- Facilitar los **mecanismos y directrices** basados en principios y mejores prácticas para el eficaz ejercicio del **gobierno del dato** de las funciones y los negocios.
- Coordinación transversal de los negocios y las funciones para la explotación y democratización del dato

PRODUCT OWNER/ TECHNICAL LEAD

- **Generador de demanda** de ingesta de fuentes y datos que requiere el caso de uso
- Metadadado de los datos a ingestar



NEGOCIO

El ejercicio del **gobierno del dato del negocio** Se realiza a través de 2 roles:

DATA CHAMPION

- Proporcionar la **visión de gestión y gobierno del dato del negocio**, definir directrices a aplicar dentro de su alcance
- Punto focal del negocio en referencia a datos

DATA STEWARD

- Persona de negocio con un conocimiento de los datos asociados a uno o varios dominios.
- Definir las reglas de calidad de los datos (y métricas)
 - Identificar y definir los diferentes aspectos que pueden afectar a los datos en cuanto al tratamiento de los mismos
 - Autorización de ingesta y distribución de los datos a casos de uso

El modelo se se pone en funcionamiento con la **demanda real de datos por un caso de uso que lo requiere ingestar y/o consumir**

4.2.4.

CASOS DE ÉXITO

- Con un **modelo de relación** con y entre las diferentes funciones con responsabilidad directa o de apoyo en el ámbito del dato.
- Con una forma de trabajar basado en **mesa de trabajo y colaboración**:
- Mesas de trabajo al que se han definido con un propósito compartido, donde hay unos participantes permanentes y se adapta según la necesidad o idiosincrasia incorporando otros participantes.
- Se acuerda una periodicidad de dichas mesas y atendiendo a la cantidad de temas a tratar y/o resolver para adaptar su convocatoria, siempre con una agenda clara a discutir.

Adicionalmente, se mantienen mesas de trabajo con las funciones transversales para definir y acordar principios a desplegar en el gobierno de los datos, ejemplo de ello es la coordinación constante con el equipo de Ciberseguridad, Compliance y Legal.

- La construcción de una metodología que incorpora los aspectos de

gobernanza del dato en el flujo de trabajo de cada caso de uso para facilitar el consumo del dato de acuerdo con el propósito, acelerar la identificación, ingestay disponibilidad de datos en el ciclo de vida del caso de uso y permitir su reutilización ágil en otros casos.

- **La construcción de la plataforma ARiA** pensada “*governance by design*” con sus principales componentes.

Esta es la explicación de algunos de estos componentes:

- Diccionario de datos **My Catalog**, con un acceso universal a todos los usuarios de la empresa donde se ha hecho de Repsol el principio “*a la vista de suficientes ojos, los errores son evidentes*” de Wikipedia, que permita su mejora continua con la colaboración de todos.

My Catalog integra los aspectos de gobierno negocio y técnicos, proporcionándonos la orquestación de todos los procesos (vía API) de recogida de datos en la plataforma y garantizando que todo lo recopilado

en la plataforma analítica está catalogado y gobernando.

Integra un cuadro de mando de acceso universal con las principales métricas de gobierno, datos y la funcionalidad de exportación fácil de los datasets y datos con su metadato.

- Herramientas para **búsqueda** de información catalogada puesta a disposición.
- **My Request**. Automatización de la gestión de autorizaciones de acceso a datos a los lagos, borrados y catalogación masiva.



CASOS DE ÉXITO



- **Descubrimiento y linaje físico.** En cuanto a descubrimiento de datos Repsol es partner estratégico con Microsoft para el desarrollo del servicio Purview

Actualmente, hay una monitorización activa de datos asociados a GDPR (datos personales) y PCI/DSS mediante la creación de reglas basadas en algoritmos de expresiones regulares para el descubrimiento y clasificación de los activos de datos.

Ya están trabajando en la extensión de está funcionalidad a otros aspectos de negocio e implementación de nuevas funcionalidades (p.e. auto catalogado, linaje de transformaciones y lógicas, etc.).

En cuanto a linaje físico de los datos:

Se dispone de la capacidad de trazar el movimiento desde la recogida hasta la disponibilización para su consumo. Incluyendo los tratamientos físicos básicos (*joint*, etc) de los *dataset* contenidos en la plataforma.

- **My Monitor App.** Se ha desarrollado una App que permite el monitoreo (pasivo y activo) de gobierno de los casos de uso y de la plataforma.

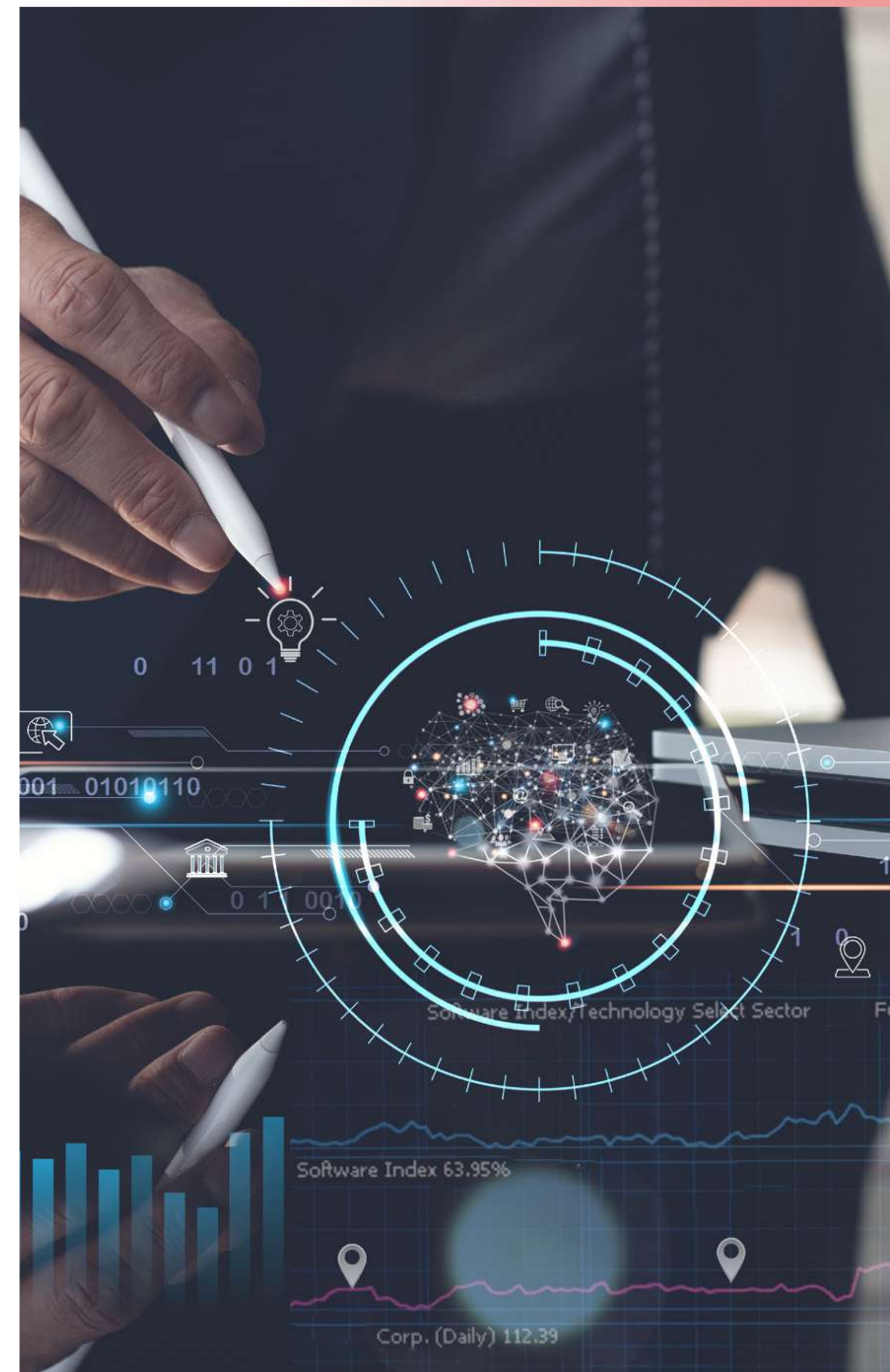
Estos son algunos de los componentes y de los paneles de esta aplicación:

- Monitorización de los casos de uso y datos asociados en la plataforma: monitoreo completo de los casos de uso, fuentes de datos, dataset y datos que consumen, la re-utilización de datos real y potencial atendiendo a su catalogación y principales métricas (por negocio, medias, modas, etc).
- **Monitorización de calidad del metadato**
- **Monitorización de volúmetrías** por entornos, zonas, fuentes y datasets

CASOS DE ÉXITO

- **Monitorización de permisos:** de los casos de uso y privilegiados
- **Monitorización de las operaciones** por entorno, zona, operación, usuario con alertado de operación "no autorizada"

Ya se está trabajando en la *release* 2.0 y se están retando los paradigmas para redefinir qué hacer y cómo hacerlo en lo relativo a gobernanza del dato, incluyendo nuevas capacidades de protección de datos "en profundidad" como **palanca de democratización del dato.**



4.2.4.

CASOS DE ÉXITO

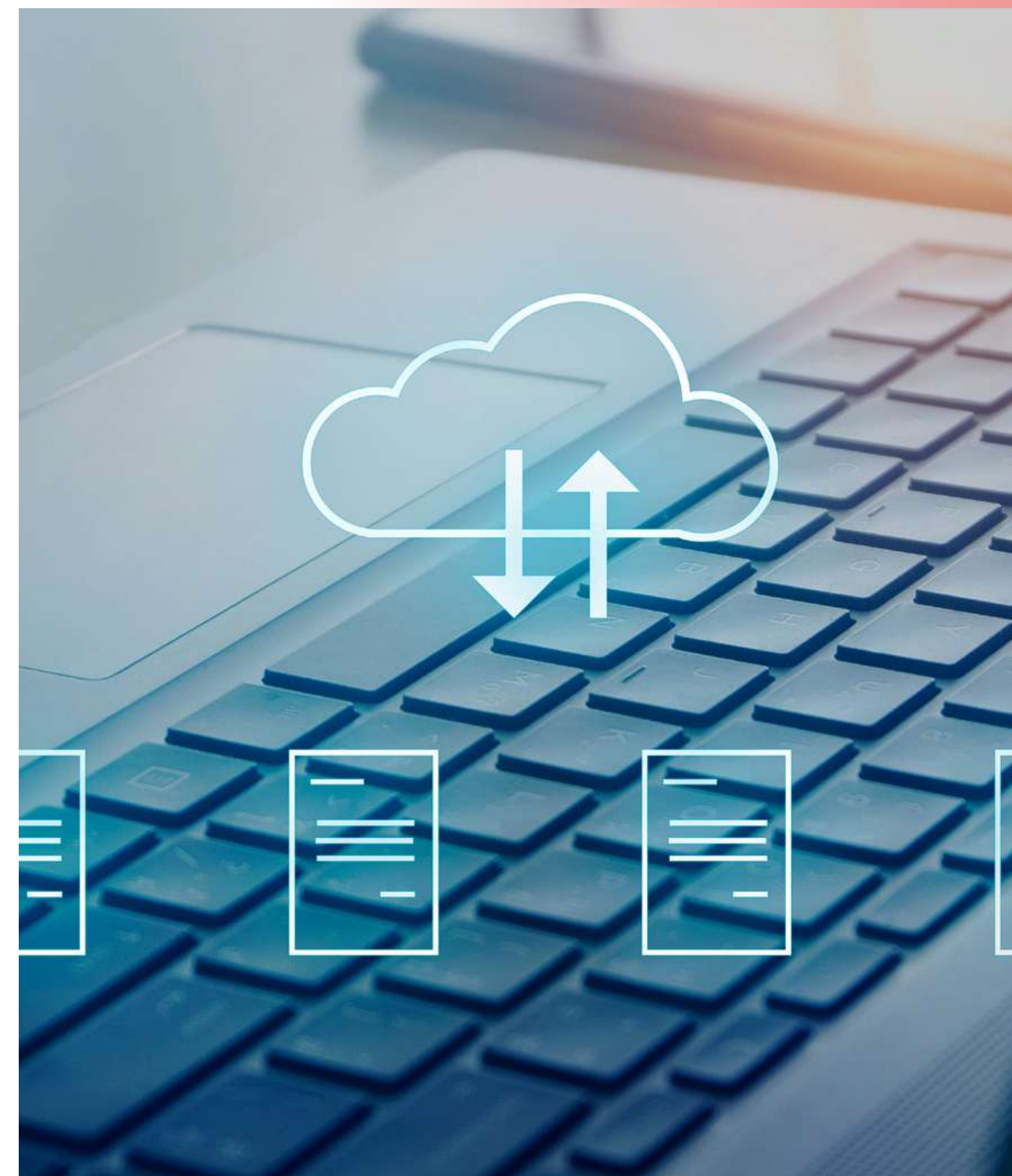


4.2.4.2. Thor

Telefónica Kernel asegura la calidad, consistencia e integridad de los datos que almacena, consulta o genera, a través de Thor. Thor es una solución que provee un framework y una capacidad en nuestras distintas operaciones, para validar las extracciones de datos de acuerdo a la versión del URM (Unified Reference Model) tanto durante el desarrollo, antes de la puesta en producción, como en la ingesta y puesta en producción en Kernel.

La funcionalidad de Thor es la siguiente:

- La funcionalidad de Thor como herramienta integrada en Kernel, desplegada en cada una de las instancias de Kernel de los países del grupo Telefónica, permite la validación de la calidad durante el proceso de desarrollo y productivo de la transformación de los datos a un modelo standard así como la calidad de estos en el proceso productivo, no solo a nivel de formato homogéneo sino en la explotación en sí de los datos.
- Provee de validaciones simples y complejas de acuerdo a la definición standard del URM y las especificaciones de los productos soportados por Telefónica Kernel.
- Incluye la validación de todos los datasets del URM, en cada una de sus versiones: 1.226 datasets, 22.664 variables, 37.000 validaciones



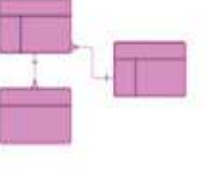



CASOS DE ÉXITO

Como decíamos, y en resumen, Thor es una solución que provee un framework y una capacidad en los países para validar las extracciones de datos de acuerdo a la versión del URM en el desarrollo, antes de la puesta en producción, así como en la ingesta y puesta en producción en Kernel.

Asegura la calidad la consistencia y la integridad de los datos generados

 **37.000** data validations available with THOR

	Validación de datasets <ul style="list-style-type: none">Entity(dataset)/versión vs. URMNº filasValidación de cabeceras (estructura y campos)	<ul style="list-style-type: none">Nombres de las carpetas de almacenamiento(durante desarrollo)Ejecución del dato vs. El intervalo
	Validación de campos <ul style="list-style-type: none">Estadísticas de campos y variables (max, min, total, etc.)Tipo de datos vs modelo URM	<ul style="list-style-type: none">Conteo de nulosValidación de Nulos
	Validación referencial y de integridad <p>Validación de dimensionales. (campos vs. Entidades de dimensionales)</p> <ul style="list-style-type: none">Conteo de error% filas con errorMuestras de datos con error	<p>Validaciones Referenciales (campos vs. El mismo campo en otra entidad)</p> <ul style="list-style-type: none">Conteo de error% de filas con errorMuestras de datos con error
	Business Validation <p>Validaciones específicas de producto que aseguran el sentido con el negocio</p>	

4.2.4.

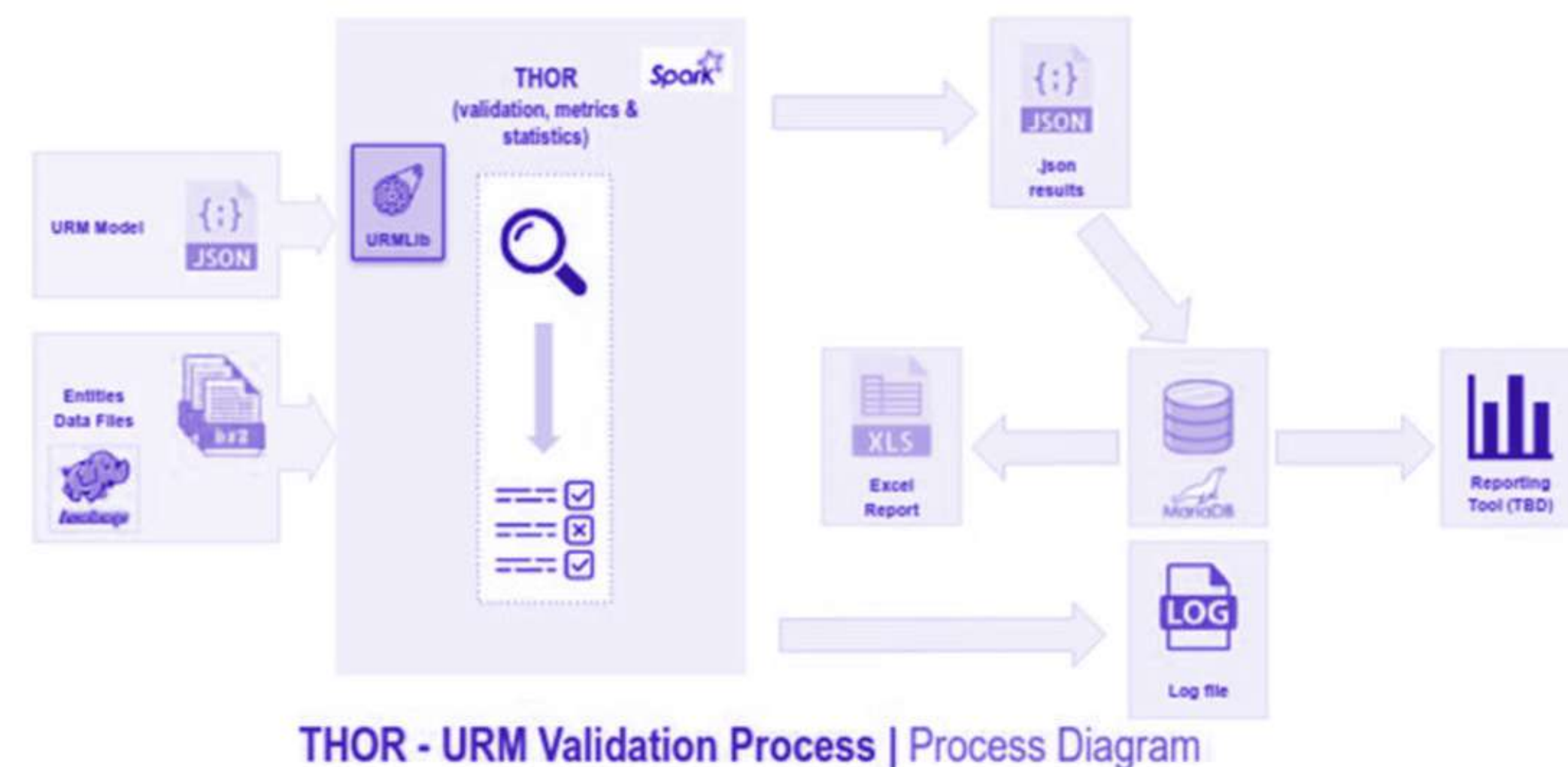
CASOS DE ÉXITO

El flujo de trabajo para el proceso de validación de URM es el siguiente:

- Lectura del modelo URM en formato Json/ avro.
- Lectura de los *datasets* de datos en el *storage*.
- Ejecución del motor de validaciones en Spark.
- Salida de los resultados de las validaciones en formato Json.
- Inserción en una base de datos, MariaDB or PostgreSQL, de los resultados para la visualización de gráficas.
- Generación automática de un excel con los resultados de las validaciones, así como con muestras de los errores.

Después de la ejecución los resultados se consolidan en un proceso separado:

- Exportación de los ficheros Json al storage del nodo Global.
- Generación de un resumen ejecutivo en una herramienta de visualización.

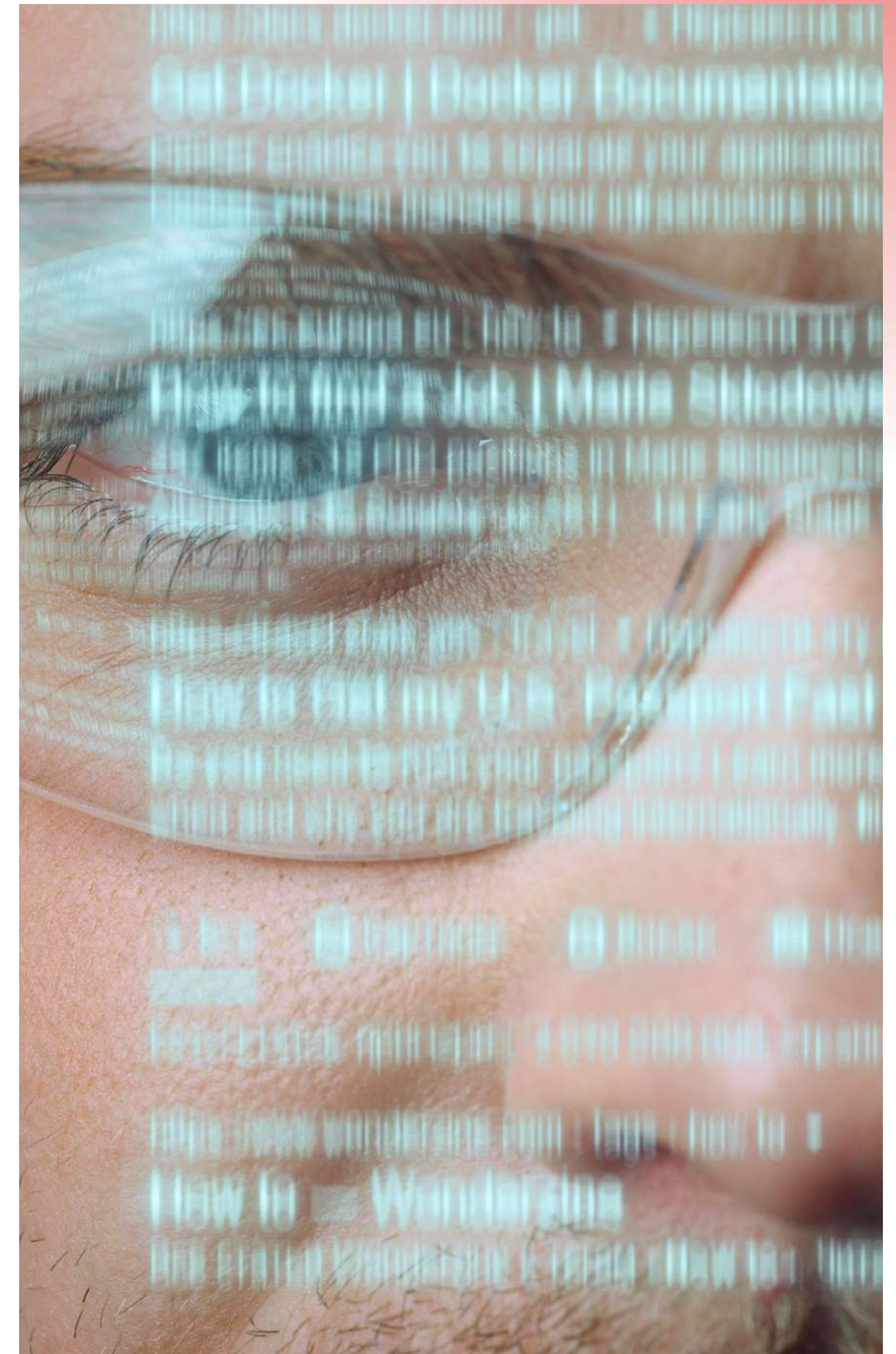


CASOS DE ÉXITO

ferroviario

4.2.4.3. Modelo de gobierno corporativo

Para velar por la calidad de los datos, desde el equipo de gobierno del dato de Ferrovial se ha diseñado un *framework* de calidad del dato que define de una manera estándar cómo debe medirse ésta, cómo se puede hacer un seguimiento de las métricas o *DQIs* que representan la calidad de un dato y cómo hacer seguimiento de planes de remediación para ir mejorando de manera continua la calidad de los datos que son consumidos dentro de la organización.



4.2.4.

CASOS DE ÉXITO

Como la calidad de un dato o un algoritmo, puede ser algo subjetivo, es necesario, además de establecer un formato común de medición, complementarlo con niveles de tolerancia o umbrales que permitan ofrecer confianza para su uso en determinados contextos o casos de uso. Un mismo dato puede

tener un nivel de calidad suficiente para un determinado caso de uso, pero no para otro. Por eso, habrá que tener en cuenta quién va a consumir ese dato y para qué se va a usar a la hora de medir la calidad de un dominio de datos.

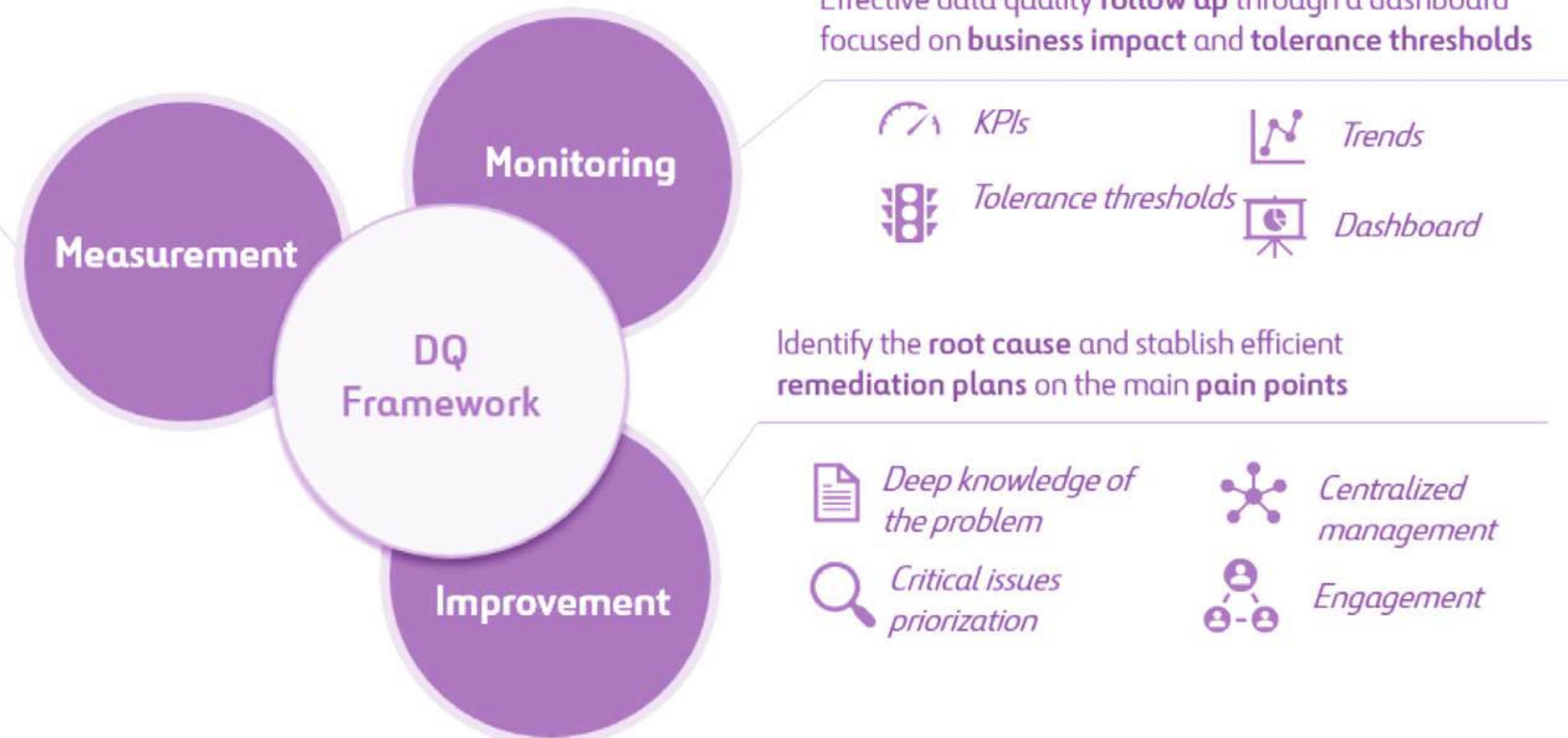
PRINCIPLES OF DQ FRAMEWORK What does the Data Quality program consist of?

ferrovial

DQ priorities must align with business strategy. Before launching a DQ program, it is beneficial to understand business needs, identify pain points, and start to build consensus about the drivers and priorities for DQ improvement.

Quantity whether the data is fit for the purposes to which data consumers want to apply it

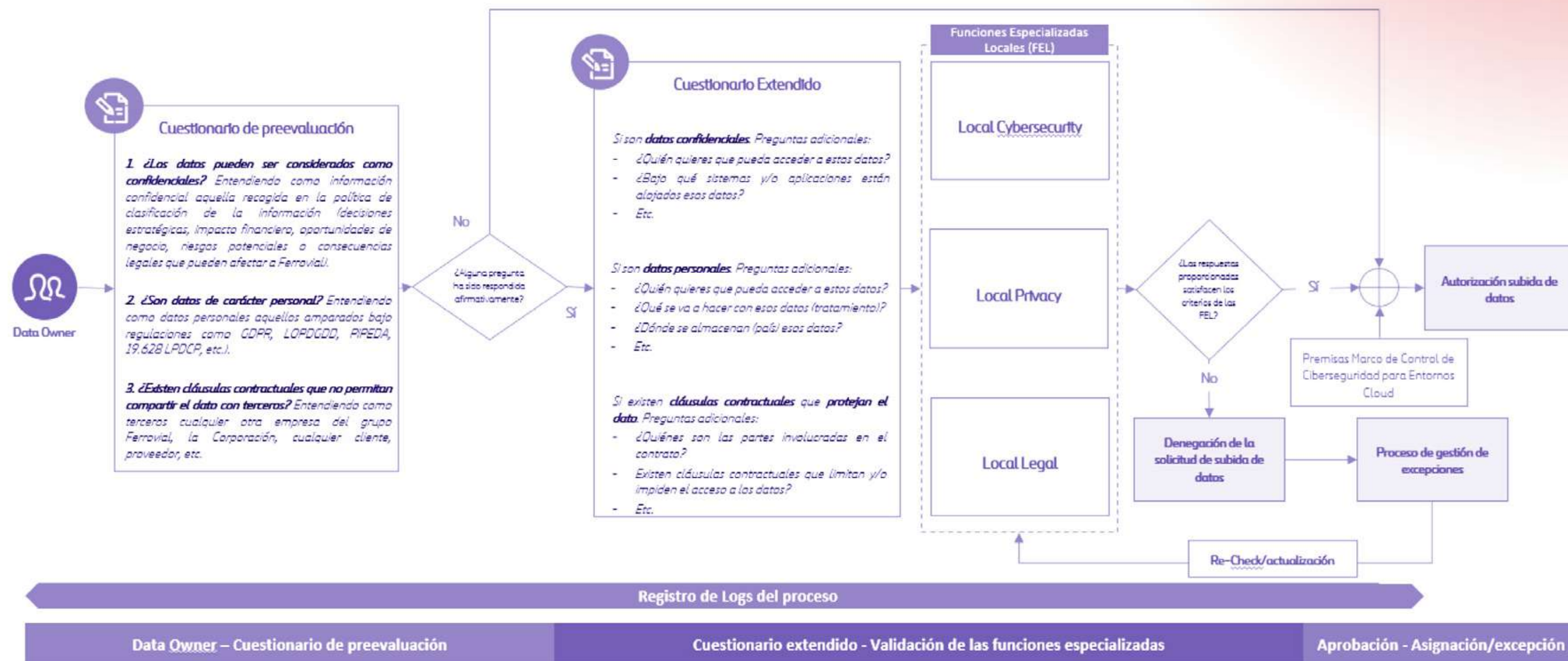
-  Focus on critical data
-  Standard data quality dimensions
-  Objective and uniform measurement
-  Lifecycle management
-  Sharing of knowledge



4.2.4.

CASOS DE ÉXITO

Proceso de acceso al dato Enfoque metodológico



ferroviario

En cuanto a la seguridad y privacidad de los datos, también se ha definido un *framework* o marco común que permite identificar qué tipo de acciones se deben llevar a cabo sobre un conjunto de datos para garantizar su privacidad y protección teniendo en cuenta dónde están alojados, la regulación asociada a esa región, el tipo de dato (sensibilidad, criticidad) y los acuerdos contractuales que pueden aplicarle en función del caso de uso del que se trata.

Algunos casos de uso pueden requerir la participación de especialistas en ciberseguridad, temas legales y oficina de protección de datos.

Los escenarios en los que la propiedad del dato es compartida, pueden ser especialmente complejos a la hora de abordar el tipo de acciones que se deben llevar a cabo. Sin embargo, contar con un *framework* común facilita la toma de decisiones.

CASOS DE ÉXITO

AIRBUS

4.2.4.4. Data governance model & strategy

La estrategia de datos de Airbus tiene como objetivo construir la continuidad digital a lo largo del ciclo de vida de sus productos y servicios: desde el diseño de los aviones, helicópteros y satélites hasta su producción y reciclaje al final de su servicio 20 años después.

Para lograr esta estrategia, han establecido una estructura de gobierno de datos en todo el grupo, con más de un centenar de actores de gobierno de datos desplegados en cada división y cada función de negocio.



La definición de la política de datos permite la creación de valor de los datos, protegiendo al mismo tiempo a la empresa. Dicha definición es llevada a cabo por una unidad centralizada de *Chief Data Officer* que depende directamente del CIO, con una coordinación muy fuerte con los departamentos de Legal y Seguridad, en línea con la estrategia de la empresa.

4.2.4.

CASOS DE ÉXITO

Esta política de datos es aplicada por una red de Data Officers en las distintas líneas de negocio, que son los responsables de los datos de las distintas áreas. Dichos Data Officers se apoyan en los Data Custodians, que son los garantes del conocimiento de los datos, de las normas de cumplimiento de los datos y de la calidad de los mismos en cada uno de los equipos de las áreas. Todo ello en estrecha coordinación central con la oficina del Chief Data Officer.

El procesamiento lo llevan a cabo los analistas de datos y científicos de datos. Gran parte de estos analistas y científicos

de datos han sido formados gracias al marco de aprendizaje a medida desarrollado para Airbus en colaboración con una prestigiosa organización educativa, que ha permitido que en la actualidad haya más de mil doscientos empleados cualificados en las distintas áreas del negocio asumiendo estas tareas.

Además, los analistas están permanentemente soportados por los equipos de IT para garantizar la disponibilidad e integridad de los datos cuando son requeridos.





5.1. Reto 7. Modelo organizativo

5.1.1. Descripción del problema

Falta de una estructura organizativa claramente definida y con un modelo de sourcing apropiado a la ambición y el nivel de madurez

Uno de los puntos clave para llegar a ser una compañía data-driven, que permita la correcta explotación de la información e implementar los principios de gobierno, es establecer un modelo organizativo capaz de gestionar el dato adecuadamente y generar un cambio cultural en toda la empresa.

Esta nueva estructura organizativa debe facilitar la adopción de los datos en la cultura corporativa, así como, democratizar y gobernar el uso de los mismos, apoyar a las áreas de negocio en la ejecución de proyectos de datos e Inteligencia Artificial, dotar de datos confiables en tiempo y forma a las áreas usuarias, o coordinar los esfuerzos necesarios para homogeneizar los criterios de compañía en torno a los datos.

Diseñar, y finalmente implantar, la estructura organizativa más adecuada requiere responder una serie de preguntas que ayuden a definir de dicho modelo y a asentar unas buenas prácticas:

- **¿Quién orienta y ejecuta el desarrollo y la gestión del dato?** Definición de la estructura organizativa que representa la forma en que se toman las decisiones, se ejecutan las actividades de datos e IA y se guía la gestión del dato.
- **¿Quién y cómo toma las decisiones relacionadas con el dato y la analítica?** Creación de los órganos de dirección del modelo de gobierno que representan un liderazgo interfuncional para la toma de decisiones y aprobaciones que sirven al marco operativo de gestión de datos para asegurar su evolución y adecuación.
- **¿Qué roles se establecen, qué funciones tienen y cómo se relacionan?** Definición de los roles y funciones de la gestión del dato para respaldar las actividades de gestión de datos a nivel tanto de negocio como técnico.
- **¿Quién asume cada rol?** Asignación de roles de acuerdo a las necesidades y dimensionamiento del número de recursos internos necesarios para la ejecución de las funciones del dato en la compañía.
- **¿Qué competencias se necesitan?** Abordar estas cuestiones no es algo sencillo, requiere de una reflexión profunda y de un apoyo al más alto nivel de la compañía. A continuación, se describen los distintos modelos de referencia que existen, y referencias de otras organizaciones.

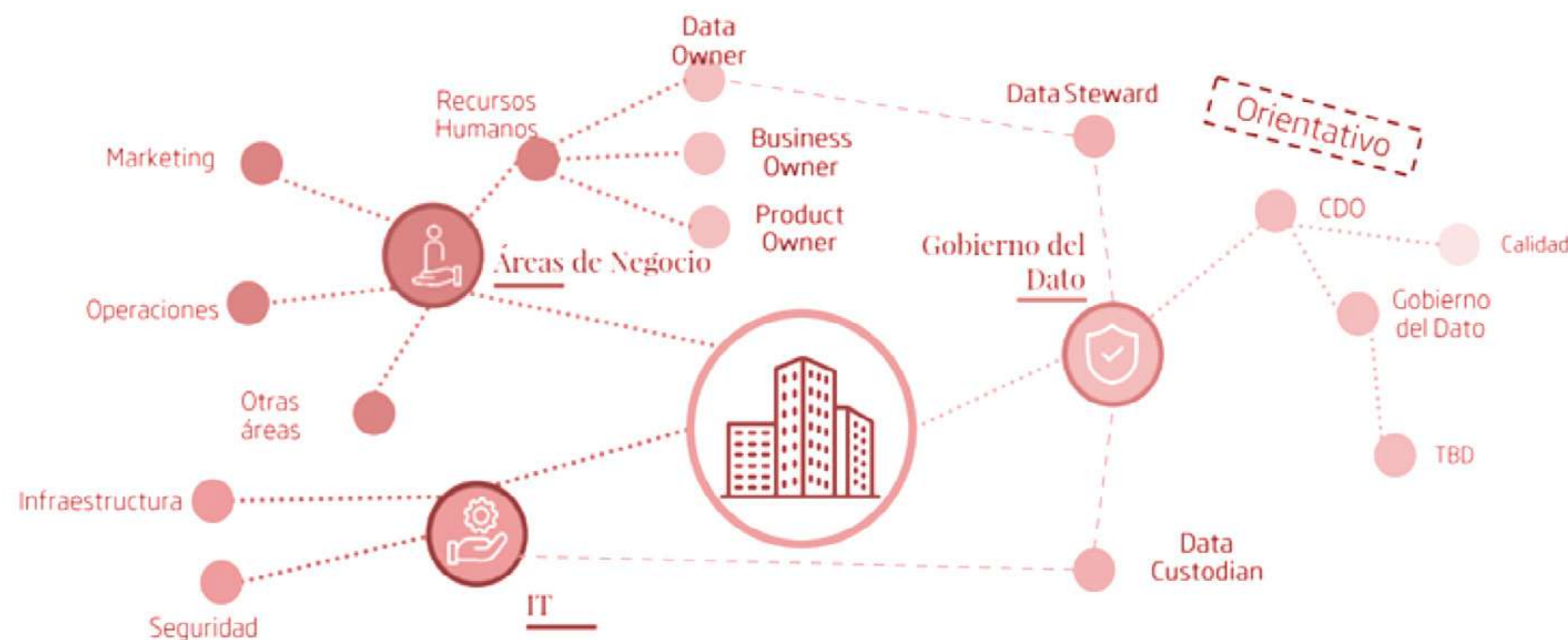


5.1.2. Soluciones

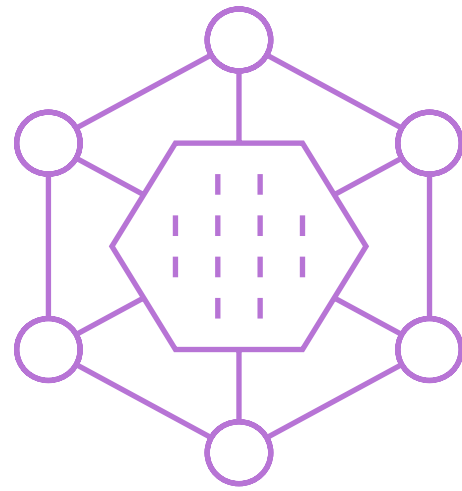
5.1.2.1. El líder de datos, CDO

Los distintos modelos organizativos que definen cómo debe ser la gestión del dato y la Inteligencia Artificial dentro de las empresas tienen en común la aparición de un CDO, *Chief Data Officer*, en los organigramas. El CDO estará ubicado más o menos cerca del CEO, en función del valor estratégico que se otorgue a los datos en la empresa. En cualquier caso, se trata de una figura cuyo principal objetivo debe ser transversal a la empresa y que representa una gran parte del negocio.

La nueva estructura organizativa debe contemplar el área de datos e inteligencia artificial y debe estar liderada por el CDO, figura del líder de datos, como actor principal en la ejecución de la estrategia del dato de la compañía, sirviendo este de mediador con la dirección de la compañía y velando por el alineamiento hacia los objetivos estratégicos del negocio.



Introducir una oficina del dato corporativa liderada por el CDO, permitirá sinergia de los proyectos para todas las áreas y contar con diferentes actores que guíen la evolución del dato de forma centralizada junto al negocio, abordando las iniciativas analíticas desde una perspectiva transversal.

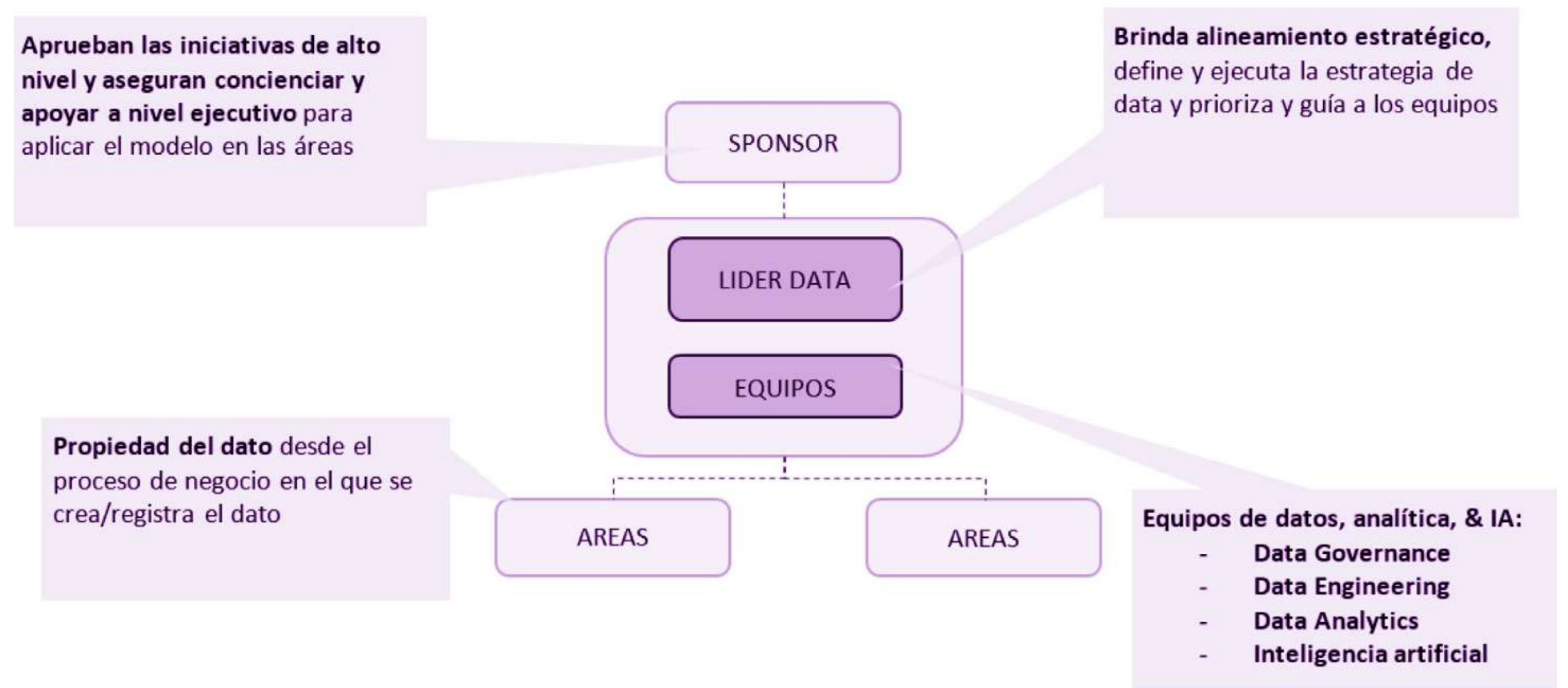


5.1.2.2. La oficina del dato

La creación de una oficina del dato velará porque todos los desarrollos, independientemente del área que lo realice, cumplan unos estándares, niveles de calidad y gobernanza adecuados ayudando además a impulsar y potenciar las habilidades de explotación proporcionando apoyo a las áreas.

La centralización de las funciones técnicas de desarrollo y la guía en cuanto a la gestión del dato en una sola unidad, considerando la sinergia con roles de negocio, creará la base de conocimiento de datos e inteligencia artificial en la compañía.

Los componentes organizacionales a implantar para el adecuado desarrollo y gestión del dato y su analítica deben considerar como mínimo las piezas descritas en el siguiente modelo organizativo de alto nivel.



INDESIA

Framework Data & IA
Industrial Marco de
referencia de datos e
inteligencia artificial para el
sector industrial español

INTRODUCCIÓN

ÁMBITO DEL VALOR PARA NEGOCIO

ÁMBITO DE LA CIENCIA DE DATOS

ÁMBITO DE LOS DATOS

ÁMBITO DE LA ORGANIZACIÓN

ÁMBITO DE PERSONAS Y CULTURA

ÁMBITO DE LA SOSTENIBILIDAD

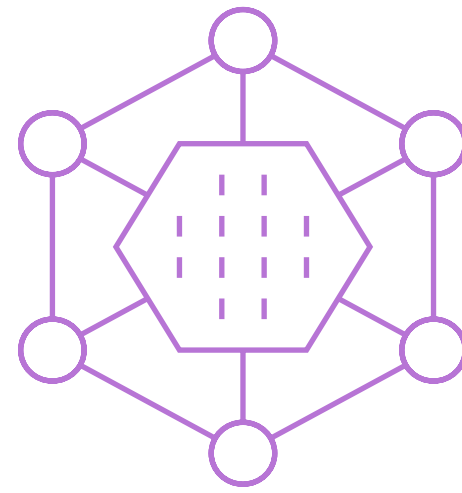
CONCLUSIONES

La misión de los equipos de la oficina del dato es democratizar el uso del dato y maximizar el valor dentro de la organización en el camino para conseguir ser una compañía Data Driven garantizando el gobierno del dato, asegurando su calidad, animando y facilitando la generación de casos de uso entre las áreas. Así mismo serán los encargados de disponibilizar los datos para analítica descriptiva, desarrollar la analítica avanzada y modelos de inteligencia artificial.

La oficina del dato debe contar con capacidad en materia de analítica de datos y estadística, la puesta en funcionamiento de modelos analíticos y plataformas, la programación, el desarrollo de metodologías y procesos, la creación de plataformas de Inteligencia de negocio, el conocimiento de la ingeniería de datos, la arquitectura empresarial, la calidad y gobierno del dato, la gestión y visualización de datos, la seguridad y en resumen aspectos relacionados con las tecnologías IoT. Por lo tanto, hay que incorporar al equipo nuevos perfiles profesionales como: Data Architect, Data Scientist, Data Governance, ML Engineer, Data Engineer, Data Steward, Data Visualization, Data Translator, Citizen Data Scientist, IoT Specialist, Industry 4.0, Data Analyst.

Resolver cuál es el mejor modelo de la oficina del dato, dimensionamiento de recursos, nivel de centralización, y su relevancia dentro de la organización empresa será necesario hacer una reflexión profunda de las necesidades del negocio con la dirección, diseñar el modelo, y ejecutarlo de acuerdo a un plan.





5.1.3. Niveles de madurez

5.1.3.1. Fragmentado

Es un modelo departamental, los datos se utilizan por cada área sin un marco de gobierno específico, los equipos adquieren conocimientos específicos y *know-how* sin compartirse con otras áreas. Es un enfoque “isla” que no aprovecha el *know-how* adquirido o reutiliza soluciones de una forma fluida.

5.1.3.2. Centralizado CoE

Los Centros de Excelencia de datos, CoE, son una estrategia utilizada para gestionar todas las iniciativas de forma integral en la empresa, no solo las habilidades técnicas necesarias, sino los recursos de negocio necesarios para ejecutar los casos de uso que más impactan los objetivos estratégicos de la compañía.



5.1.3.3. Hub & Spoke

El modelo *Hub and Spoke* se refiere a un método de organización en el que existe una base centralizada de gobierno, metodología y conocimiento. Toda la estrategia de datos e IA se promueve y gobierna en el *hub* y se distribuye hacia las áreas de la compañía, los recursos de negocio gestionan gran parte de los casos de uso siguiendo las reglas marcadas por el *hub*.

5.1.3.4. Descentralizado

En organizaciones más maduras la gestión y gobierno de datos descentralizada puede ser una buena solución ya que permite una mayor agilidad y control a las áreas de la compañía sobre sus casos de uso, y por lo tanto extraer mayor beneficio en menor tiempo. Pero, no hay que confundirlo con un modelo fragmentado, los modelos descentralizados permiten el gobierno de datos e inteligencia artificial por las áreas de negocio bajo una estrategia común corporativa de procedimientos, recursos y conocimiento.

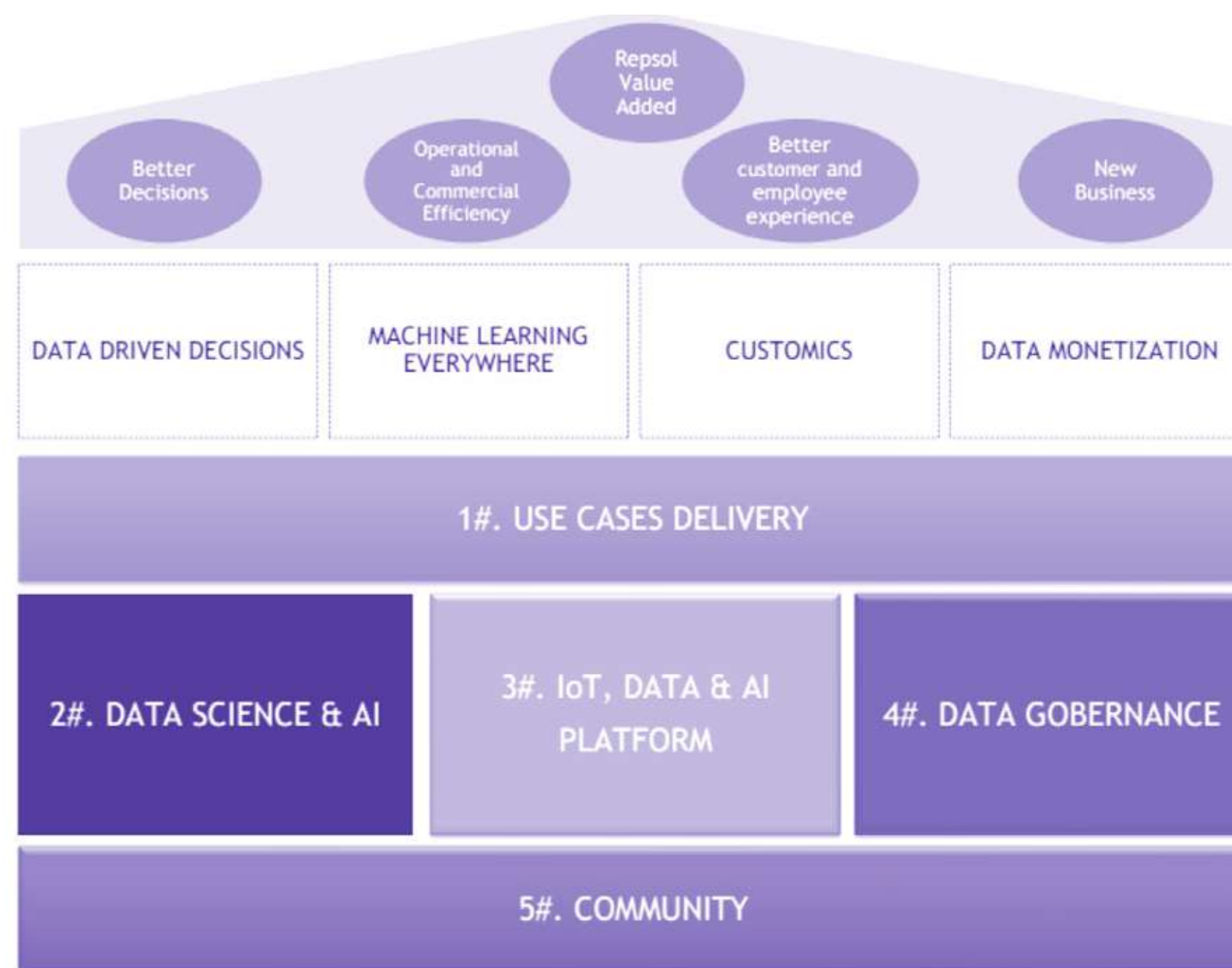
5.1.4.

CASOS DE ÉXITO



5.1.4.1. HuB de Data & Analytics

El HuB de Data & Analytics se creó en 2018 con la misión de ayudar a Repsol a convertir sus datos en una fuente de ventaja competitiva, acelerando y escalando el desarrollo de casos de uso analítico a través de la ejecución de las siguientes funciones:



CASOS DE ÉXITO

- **Soporte a la ideación y *delivery* de casos analíticos.** Ayudar a los negocios a identificar oportunidades en las que los datos puedan generar valor para Repsol y facilitar el desarrollo de los casos de uso analíticos, ayudando a los *Product Owners* en las decisiones relacionadas con la gestión de los datos y el desarrollo de modelos, garantizando que las soluciones desarrolladas son viables y escalables para Repsol.
- **Conocimiento y experiencia en *Data Science*.** Si los datos y los algoritmos son core para el negocio hay que ser capaces de hacer *ownership* de los algoritmos que nuestros *partners* nos ayudan a desarrollar. Esto significa conocer su funcionamiento para poder monitorizar su performance, mantenerlos y aplicar una mejora continua.
- **Gobierno y calidad de los datos.** Las empresas se benefician de la gobernanza de los datos porque garantiza que los datos sean coherentes y confiables. La gestión de datos o gobierno de los datos es la gestión general de la disponibilidad, facilidad de uso, integridad y seguridad de los datos utilizados en una empresa. Un buen programa de gobernabilidad de datos incluye un órgano de gobierno o consejo, un conjunto definido de procedimientos y un plan para ejecutar estos procedimientos.
- **Comunidad para crear cultura *data driven*.** Como HUB que es Repsol, en su esencia está el generar comunidad que gire en torno a un interés común: los datos, la analítica y su capacidad de transformar el negocio, la industria de la energía y la sociedad en su conjunto.

5.1.4.

CASOS DE ÉXITO

El HuB de Data & Analytics está formado por más de 65 profesionales del dato divididos en los siguientes perfiles:

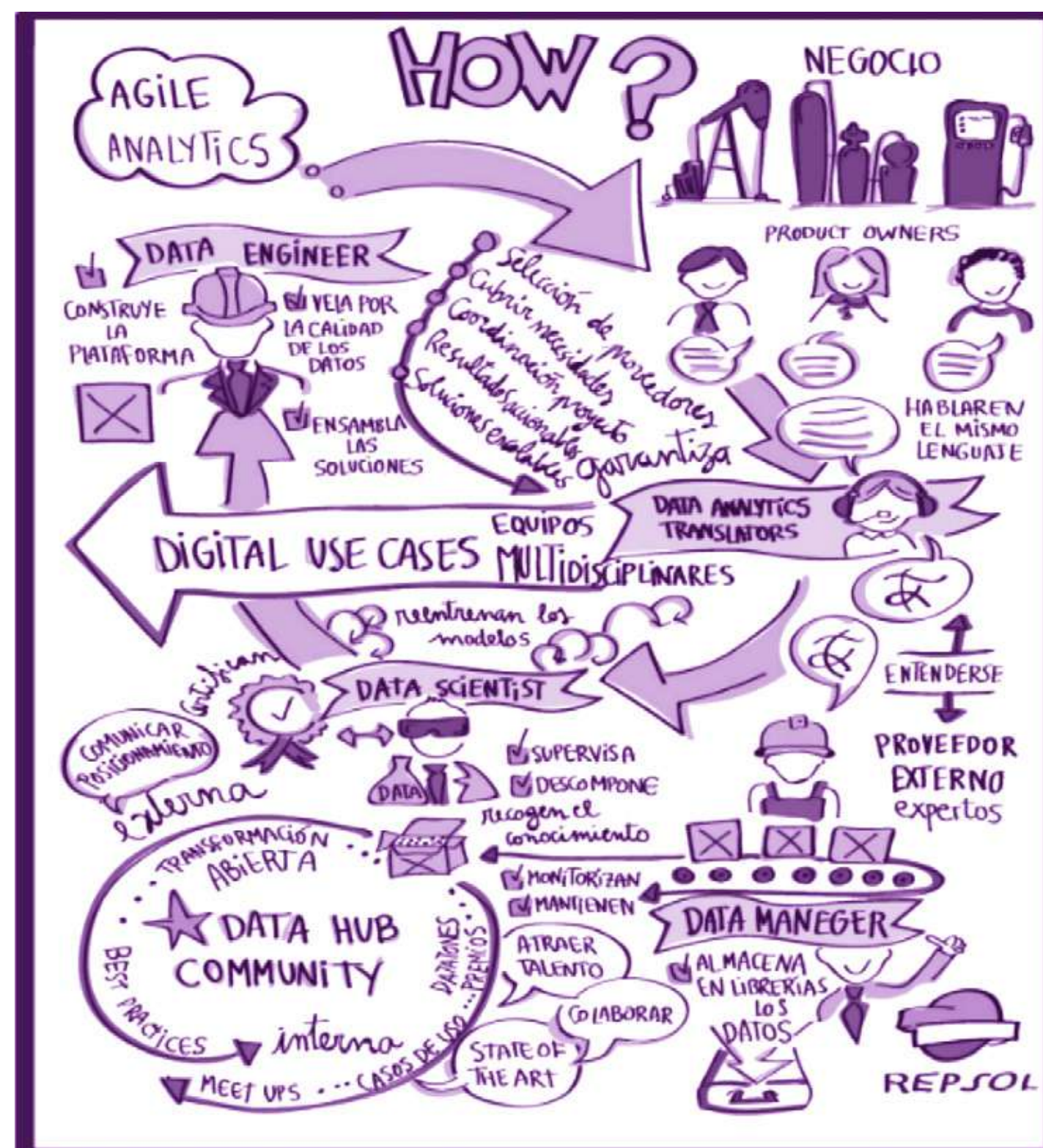
- **Data analytics translators**
 - Ayuda a las unidades de negocio a identificar y priorizar iniciativas analíticas que les ayuden a cumplir sus objetivos.
 - Acompañan al *Product Owner* en el desarrollo de los casos de uso analíticos haciendo de bisagra entre los intereses de Repsol y las recomendaciones del proveedor.
 - Conectan los casos de uso con los recursos del HuB.
 - Aseguran que el caso cumple con el estándar metodológico exigido a los casos de uso analíticos.
- **Data scientists**
 - Definen la metodología analítica y asesoran sobre las mejores técnicas a desarrollar en cada caso.
 - Realizan el *Quality Assurance* mediante *peer review* a los *vendors* con los que se esté trabajando.
 - Recogen el conocimiento de cómo se ha realizado el modelo participando en su desarrollo.
 - Monitorizan la calidad del modelo y lo someten a un proceso de mejora continua.
- **Data engineers**
 - Definir y consensuar políticas de gobierno de información.
 - Desarrollar catálogos de información.
 - Implementar medidas para mejorar la calidad de los datos.
- **Data managers**
 - Definir y consensuar políticas de gobierno de información.
 - Desarrollar catálogos de información.
 - Implementar medidas para mejorar la calidad de los datos.
- **Data community builders**
 - Crear una comunidad interna de gente interesada en los temas analíticos para fomentar compartir el conocimiento.
 - Definir e implementar planes de formación y comunicación tanto interna como externa.
 - Conectar con la comunidad exterior realizando el *scouting* de proveedores que permita saber quién sabe de cada tema.



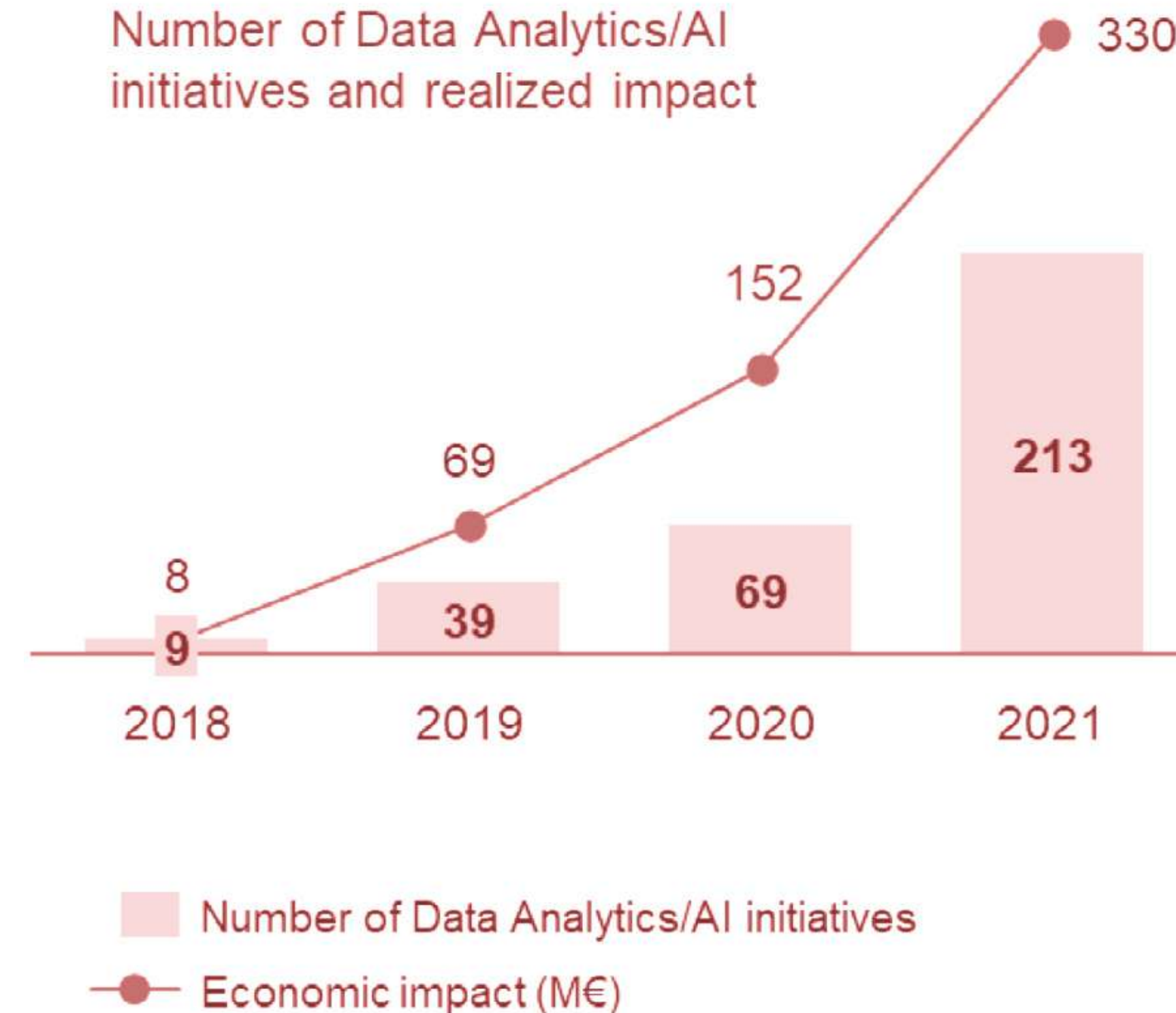
5.1.4.

CASOS DE ÉXITO

Desde la creación del HuB de Data & analytics Repsol ha conseguido escalar el uso de los datos y la analítica a todos los negocios y una gran parte de los procesos, desarrollando en 4 años más de 200 iniciativas con un retorno superior de 300 millones lo que le ha valido el reconocimiento externo como una de las compañías que más ha avanzado en su camino de convertirse en una compañía verdaderamente data driven.



Number of Data Analytics/AI initiatives and realized impact





Dónde poner en el organigrama el *Chief Data Officer*

Dado que *big data* e Inteligencia Artificial se han convertido en temas críticos en el mundo de los negocios, no sorprende que el CDO (*Chief Data Officer*) haya logrado abrirse camino hasta un asiento adicional en el comité de dirección. Cada vez más organizaciones, tanto del sector público como privado, consideran los datos como un activo estratégico, y por ello, las empresas están nombrando CDOs.

¿Una empresa se debe hacer las siguientes preguntas antes de nombrar un CDO?:

- ¿A qué distancia debe estar el CDO del CEO? CEO-1 o CEO-n?
- Si es CEO-1, ¿cómo se relaciona el CDO con los demás directivos, en particular, con el CIO y el CTO?
- Si es CEO-n, ¿a qué área debe reportar el CDO? ¿Al CIO, COO, CMO, CFO, el Director de Transformación o el Director Digital?



Para aprovechar todo el potencial de los datos, el CDO está mejor ubicado en un área cuya misión es transversal a la empresa y que representa una gran parte del negocio. De esta manera, la creación de valor no se limita a un área específica (por ejemplo, marketing), y el valor es relevante para el negocio. De lo contrario, se crea un sesgo hacia la creación de valor solo a partir de datos en un área específica o en un área que realmente no importa.

5.1.4.

CASOS DE ÉXITO

Hay otras alternativas que también aseguran la aplicación y relevancia: el Director de Transformación y el Director Digital. Si bien, por naturaleza, estos dos roles tienen un papel temporal (aunque durante varios años), funcionan de manera inter-organizacional y tienen la misión de adaptar su negocio al mundo digital, del cual los datos forman una parte fundamental.

Hay que tener en cuenta que tener el CDO reportando directamente al CEO no es necesariamente adecuado para todas las organizaciones en todo momento. Requiere un nivel de madurez digital y de datos y es probable que esté

reservado para las organizaciones con más visión de futuro que realmente saben y aceptan el hecho de que tienen que adaptarse al mundo digital de una manera basada en datos.

En una organización que acaba de empezar con Inteligencia Artificial, también es adecuado poner el CDO por debajo de un área con mucho empuje, que muchas veces es marketing, pero dependiendo del tipo de organización podría ser en otras áreas. En este caso hay que tener en cuenta las ventajas y desventajas de cada caso se ve en la tabla abajo:

Office	Pros	Cons
CMO	Marketing and sales provide use cases with direct impact	Usually focused on consumer, forgetting the business customers, missing value in other areas
CFO	Financial ledger requires high-quality data	Less business focused, and financial management doesn't need big data
CIO	Technological problems are solved according to company standards	Governed by technological criteria, not business
CTO	Take advantage of the latest technological innovations	Driven by new technology, rather than by business
CSO (security)	Good for security and privacy of customer data	Less focus on business
CRO (resources)	Cost savings go directly to the bottom line	Driven by efficiency, rather than by growth

CASOS DE ÉXITO

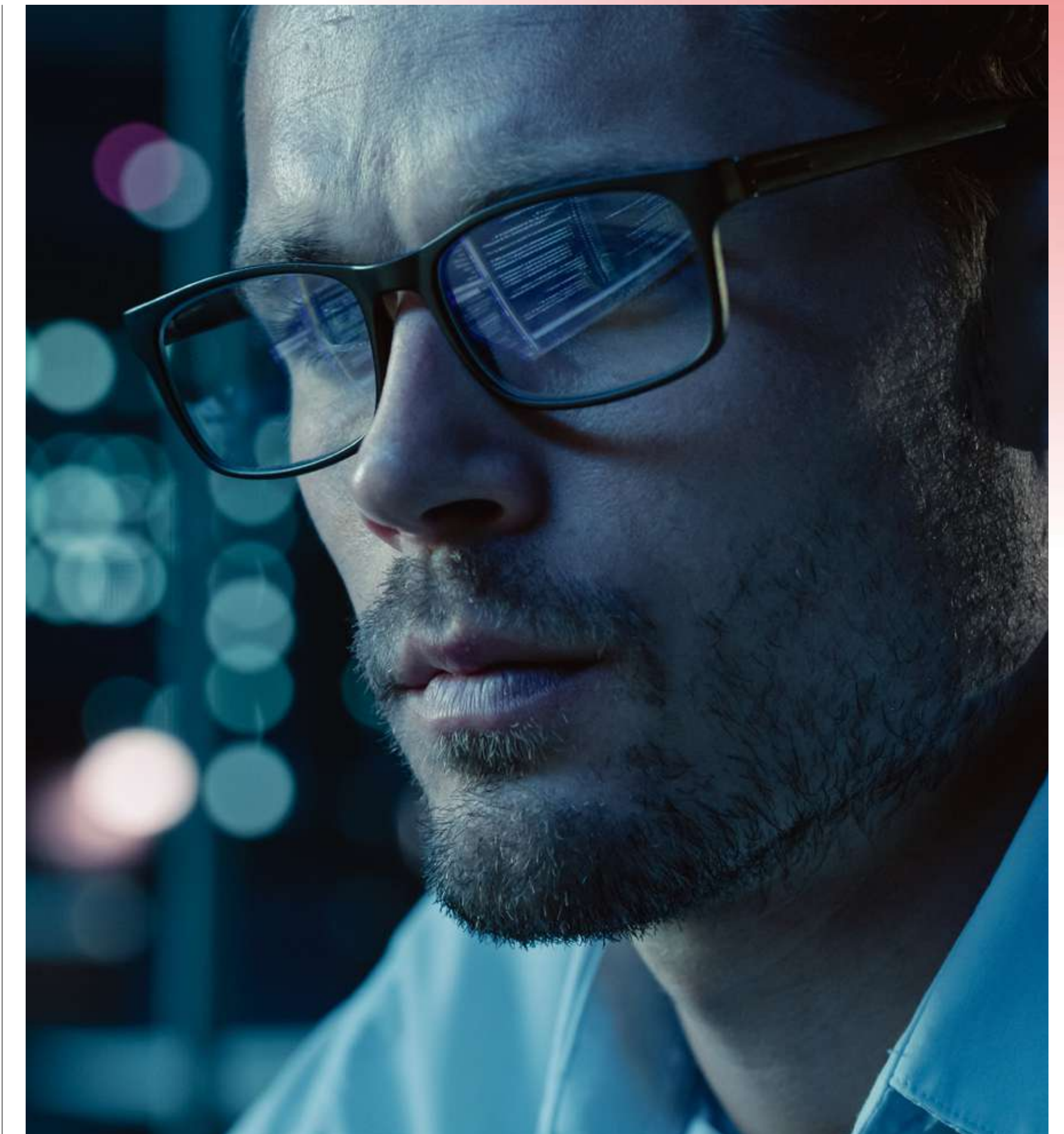


TECNICAS REUNIDAS

Digital Innovation HUB, definición del modelo de gobierno corporativo y matriz de responsabilidades

En Técnicas Reunidas se aplica un modelo híbrido en base a un reparto de roles y responsabilidades que fomenta la segregación de funciones, para potenciar que las áreas de negocio puedan definir y lanzar de forma ágil sus líneas de innovación, entre las que se incluyen las de datos e inteligencia artificial montadas sobre estándares de arquitecturas y modelos previamente homologados por el área de sistemas de información y el Comité de Dirección para la innovación digital.

Todo este modelo se recoge en el marco metodológico de constitución del HUB de innovación, que define de forma detallada esta matriz de segregación de funciones y los



diferentes comités que la conforman (operativo, dirección, inversión, ...).

Dentro de las funciones de este HUB están el potenciar y gobernar el ciclo de vida de las iniciativas de datos e inteligencia artificial dentro de Técnicas Reunidas generando un ecosistema de innovación gobernado, que filtre y priorice las iniciativas para ser capaces de generar casos de uso cualificados.

5.1.4.

CASOS DE ÉXITO



El marco presupuestario definido da cobertura al HUB de innovación (ecosistema, miembros fijos, comités, etc) y la previsión de dedicaciones de usuarios, por áreas de negocio, para cubrir las fases iniciales de ideación, exploración y cualificación. Hasta llegar a tener casos de uso documentados, los cuales se presentan al comité de dirección para su aprobación, bajo un modelo definido en el marco metodológico que identifica entre otras cosas los KPIs que justifican el lanzamiento o escalado del proyecto. Es decir, hay un presupuesto aprobado para idear y cualificar casos de uso, para su ejecución y escalado se debe pasar aprobación individual, dependiendo del importe de la iniciativa deberá ser aprobado por un comité u otro.

Respecto al modelo de ejecución se evalúa caso a caso, considerando la fase o grado de madurez de este, se define el equipo descentralizado asignado de proyecto y roles. Pudiendo ser asumido internamente o requerir colaboración con proveedores externos, en este último caso pueden

aplicar los contratos marcos o acuerdos estratégicos vigentes o bien requerir abrir un proceso de licitación conforme a los procedimientos internos de contratación. Parte de las funciones del HUB es establecer estos acuerdos estratégicos por proveedores tecnológicos de primer nivel buscando la optimización de las inversiones y la co-innovación.

Un proyecto estratégico en curso en Técnicas Reunidas, para consolidar estas directrices de gobierno en el ámbito de innovación digital, es el proyecto TRDat@, que define la arquitectura analítica de referencia, con el objetivo de que con la incorporación gradual de casos de uso se vayan implantando por dominios (áreas de negocio) los modelos de datos que se puedan poner a disposición de las áreas de negocio para que desarrollen tanto sus propios modelos analíticos tradicionales (autoconsumo) como modelos más avanzados basados en inteligencia artificial, pero en ambos casos partiendo de un modelo de datos gobernado.

5.1.4.

ferrovial

Modelo Hub & Spoke. Gobierno del dato centralizado, equipos de desarrollo descentralizados.

Identificación de roles y responsabilidades en torno al dato dentro de la organización

En Ferrovial, dada la naturaleza de esta organización, ha apostado por un modelo descentralizado "Hub & Spoke".

CASOS DE ÉXITO



En este modelo, cada unidad de negocio puede incorporar ciertas capacidades analíticas y perfiles específicos internos y externos, para ganar autonomía y velocidad a la hora de implementar proyectos de datos & IA.

5.1.4.

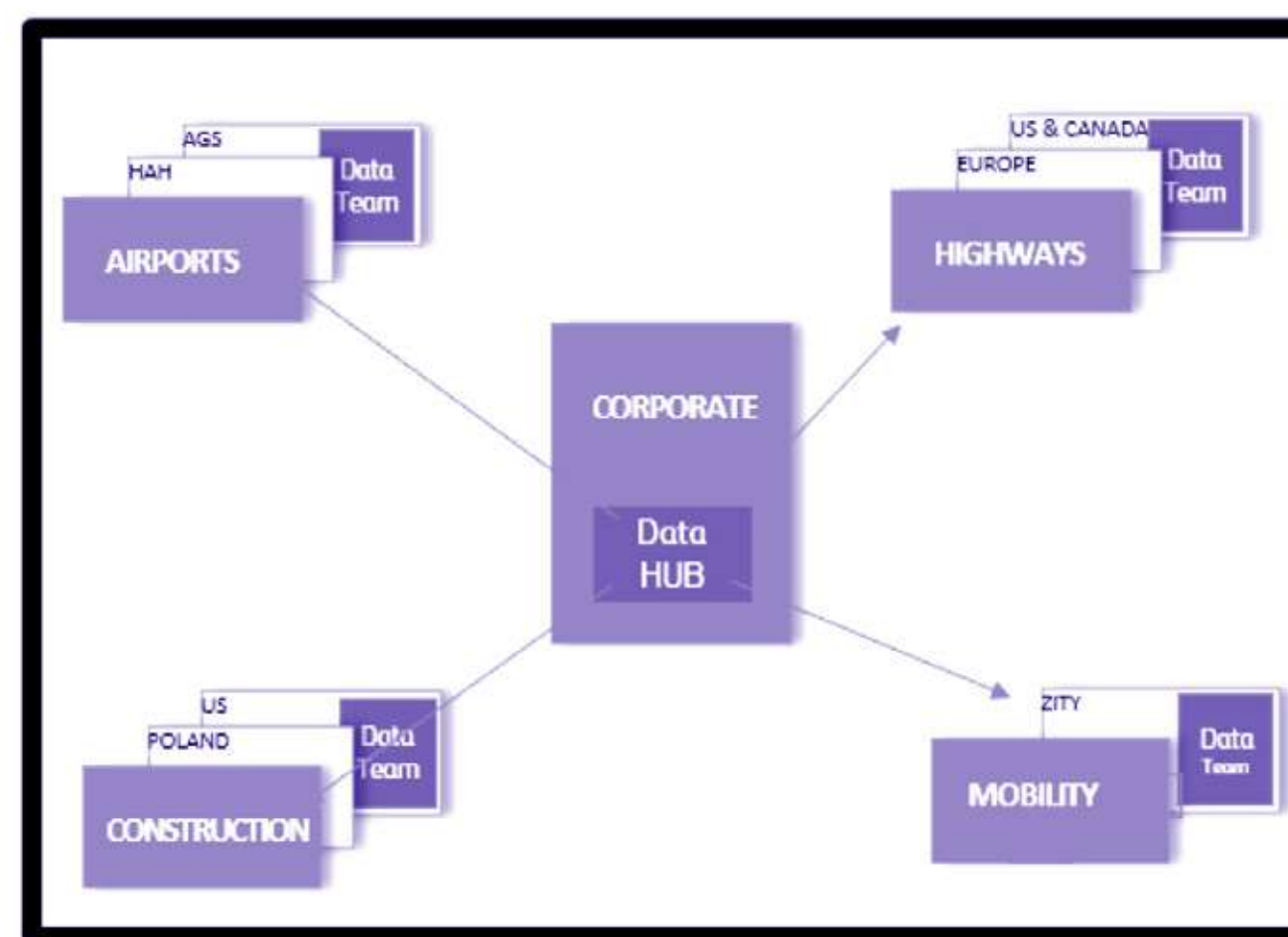
CASOS DE ÉXITO

El rol del CDO en este caso, es un rol facilitador que posee ciertas capacidades de ejecución apoyadas en un HUB de datos & IA centralizado, pero que está centrado fundamentalmente en definir los procedimientos de gobierno y el desarrollo de capacidades técnicas comunes para agilizar la captura de valor dentro de la organización.

Este es un modelo que combina capacidades internas y externas que se van adaptando a las necesidades concretas del portfolio de casos de uso.

La relación entre diferentes unidades organizativas se establece a través de un Data Council, donde hay representantes de cada unidad de negocio.

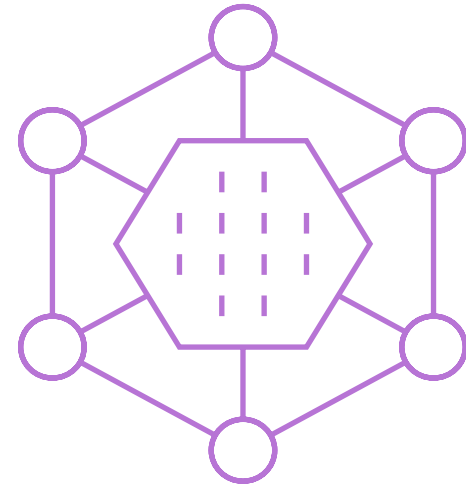
HUB & SPOKE



ÁREAS DE RESPONSABILIDAD



- Cada unidad de negocio o filial puede tener sus propios equipos de datos descentralizados, con dimensiones y skills variables. Aquéllas que no lo tienen, aprovechan las capacidades del Data HUB Corporativo.
- Gobierno común de datos a nivel de empresa.
 - Creación de **Data Council**
 - **Nivel de madurez** en torno al dato
 - Evaluación de **oportunidades** en torno al dato
 - Identificación de **dominios de datos** estratégicos
 - Definición de **políticas** de gobierno del dato
 - Implantación de **procesos** de gobierno
 - Creación de una **plataforma** de gobierno común
 - Exploración de oportunidades de monetización
- Capacidades centrales (Data Hub) de creación de productos digitales y algoritmos de IA basados en datos.



5.2. Reto: 8: Modelo de Sourcing y Financiación

5.2.1. Descripción del problema

La gran mayoría de empresas comienzan su viaje de datos identificando casos de uso de datos e inteligencia artificial con poco de criterio. Se centran en arrancar y empezar a generar “algo”. La necesidad de recursos no es una cuestión que se tenga presente inicialmente, no hay visibilidad real de la dimensión de los retos, y se hace mayor foco en encontrar aquellas iniciativas que parecen tener mayor sentido para la organización, es decir, el viaje comienza con falta de estrategia y gobierno, “se hace como se puede”.

La falta de visión sobre la dimensión de los retos en el ámbito de los recursos y la financiación supone un gran riesgo para el desarrollo de las iniciativas, y que los casos de uso no lleguen a prosperar. Por lo tanto, para afrontar con mayores garantías los casos de uso es necesario responder cuestiones fundamentales como:

¿Con que recursos se van a desarrollar los casos de uso?

Fuentes de recursos necesarios para el desarrollo de casos de uso y su industrialización: humanos e infraestructura.

¿Cuál va a ser el modelo de financiación? ¿Quién va a pagarlo?

Financiación propia de los departamentos, áreas de negocio o corporativa centralizada. Financiación externa mediante acuerdos con clientes o proveedores, ayudas.

Es importante abordar estos aspectos y empezar a producir resultados tangibles de los datos e inteligencia artificial. En otras palabras, hay que generar valor a la compañía en un plazo razonable. Poder escalar los casos de uso es una cuestión esencial y si no se pone el suficiente foco en desarrollar una buena estrategia de sourcing y financiación no lo conseguiremos, y corremos el riesgo de fracasar en nuestros objetivos corporativos.

INTRODUCCIÓN

ÁMBITO DEL
VALOR PARA NEGOCIO

ÁMBITO DE
LA CIENCIA DE DATOS

ÁMBITO DE
LOS DATOS

ÁMBITO DE
LA ORGANIZACIÓN

ÁMBITO DE
PERSONAS Y CULTURA

ÁMBITO DE
LA SOSTENIBILIDAD

CONCLUSIONES

5.2.2. Soluciones:

5.2.2.1. Modelo de sourcing

Cuando hablamos de recursos para el desarrollo y escalado de casos de uso, fundamentalmente nos estamos refiriendo a recursos humanos e infraestructura y la compañía debe marcar una estrategia en varios ejes:

- **Organizacional:** esto abarca la definición de un modelo de operación en términos de recursos humanos y la estructura organizativa, definiendo que roles y de que áreas participarán en el desarrollo de cada caso de uso.
- **Capacidades:** establecer un plan de recursos específicos en el área de datos e inteligencia artificial, y esto abarca la captación de nuevo talento STEM y la capacitación interna mediante programas de up-skilling y re-skilling.
- **Ecosistema:** los recursos pueden venir de fuentes externas, como proveedores tecnológicos, strat-up, universidades, instituciones, o socios. Por lo que debemos trabajar en tener un ecosistema potente dentro de la organización que nos permita escalar rápidamente. Según el caso habrá que estudiar el mejor modelo de ejecución. Cuando se decida subcontratar a un proveedor es necesario evaluar los riesgos y seleccionar solo aquellos *partners* que demuestren tener experiencia y capacidad para la buena ejecución del caso de uso.
- **Plataforma de datos:** es necesario definir una arquitectura de plataforma *big data/ Analytics* corporativa que se ajuste de forma óptima a las necesidades del negocio en el medio y largo plazo, así como desarrollar una hoja de ruta tecnológica para la implementación de las piezas requeridas para la implementación de casos de uso. La infraestructura y el gobierno de la misma debe crecer de forma gradual.



5.2.2.2. Modelo de financiación

La pregunta clave ¿quién paga? Los recursos financieros a la hora de desarrollar los proyectos de datos e inteligencia artificial son el principal obstáculo en él escalado.

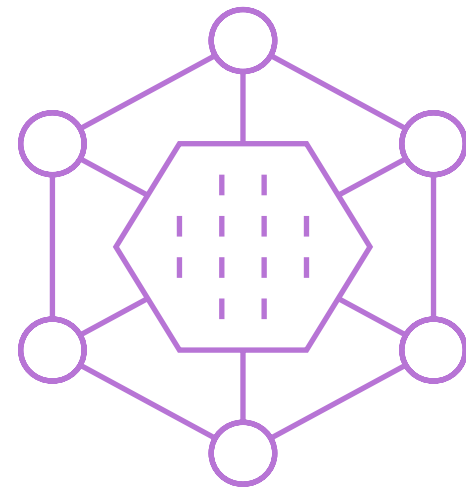
Financiación privada interna: los costes los asumimos íntegramente en la organización. La empresa debe contemplar en sus presupuestos tanto el CapEx como el OpEx para poder abordar los proyectos de datos e inteligencia artificial. La distribución de los presupuestos estará de acuerdo con los niveles de madurez de la organización:

- **Nivel 1 Fundacional:** departamental.
- **Nivel 2 Acercamiento:** corporativa.
- **Nivel 3 Aspiracional:** híbridos.
- **Nivel 4 Maduro:** negocios.

Financiación privada externa o mixta: crear un ecosistema de innovación abierta que nos permita definir modelos de co-desarrollo con clientes, proveedores, otras empresas o instituciones para abordar los costes del proyecto y marco de explotación del activo digital.

Financiación pública, subvenciones: activar un radar de convocatorias y ayudas para el desarrollo de proyectos con financiación española o europea, de forma individual o en consorcio.

La financiación de los proyectos de innovación lleva asociado un nivel muy alto incertidumbre, por lo que debemos definir metodologías que nos permitan minimizar costes. Los primeros pasos no suelen ser muy demandantes en términos económicos, pero nuestro modelo de financiación debe estar preparado para abordar las pruebas de concepto y prototipado, sin tener visibilidad clara del retorno. Una vez se dispone del primer MVP, mínimo producto viable, es cuando empezaremos a poner en producción la solución y por lo tanto empezaremos a obtener cierto retorno hasta que alcancemos el punto muerto, *break even point*, y esperamos los beneficios.



5.2.3. Niveles de madurez

5.2.3.1. Nivel 1 Fundacional. Departamental

En este primer grado de madurez cada departamento es el que dedica parte de su presupuesto anual a financiar iniciativas de IA. Seguramente en fases de madurez baja también a nivel de medición de impacto y ambición de puesta en producción. Independientemente del desglose de financiación (interna, externa, OpEx, CapEx) lo que caracteriza este nivel de madurez es que se reduce al ámbito de un área de negocio concreta que suele ser la pionera o más necesitada de resultados haciendo uso de IA.

5.2.3.2. Nivel 2 Acercamiento. Corporativa

Una vez varias áreas de negocio han demostrado sus resultados y madurez en la puesta en práctica de iniciativas de IA es normal que a nivel corporativo se dediquen ciertas partidas presupuestarias a la financiación de estas iniciativas. Ya sea

financiando desde el área de digital las primeras fases o MVP para reducir incertidumbre y replicar mejores prácticas dentro de lo posible o mediante la financiación de programas cross de adopción y reconocimiento de la IA como se ha mencionado en otras secciones.

5.2.3.3. Nivel 3 Aspiracional. Híbridos

El modelo de madurez híbrido es cuando el nivel 1: departamental y nivel 2: corporativo, coexisten durante un tiempo en la empresa. Es decir, la financiación y el sourcing sigue siendo a nivel departamental pero ya existe un departamento digital a nivel corporativo que valida y participa en todos los proyectos de IA para asegurar su correcta ejecución. El modelo híbrido pueden incluir muchas distintas maneras de financiar la innovación que quedará repartida entre el negocio y corporación según la madurez del caso o la potencial capacidad de replicarlo en otras áreas de negocio.

5.2.3.4. Nivel 4 Maduro. Negocios

La financiación de las iniciativas de IA forma parte de un programa que está incluido dentro del plan estratégico de la empresa. Es decir, cada área de negocio dedica una parte de su presupuesto a financiar iniciativas que hacen uso de la IA pero que en ningún caso se diferencia de la manera de medir, planificar y ejecutar cualquier otra iniciativa del presupuesto del negocio. Las iniciativas de IA están asumidas como iniciativas de negocio y su proceso ya no se diferencia en base a la tecnología a usar. El mayor nivel de madurez es cuando la IA ya no es algo novedoso ni diferencial, es parte de cada iniciativa dentro de un plan de transformación.

INTRODUCCIÓN

ÁMBITO DEL
VALOR PARA NEGOCIO

ÁMBITO DE
LA CIENCIA DE DATOS

ÁMBITO DE
LOS DATOS

ÁMBITO DE
LA ORGANIZACIÓN

ÁMBITO DE
PERSONAS Y CULTURA

ÁMBITO DE
LA SOSTENIBILIDAD

CONCLUSIONES

5.2.4.

CASOS DE ÉXITO



5.1.4.1. HuB de Data & Analytics

Para que una iniciativa sea considerada digital debe cumplir 3+1 premisas



- **¿Por qué?** Que promueva transformaciones profundas en los procesos y actividades del negocio.
- **¿Cómo?** Que se desarrolle utilizando metodologías "ágiles", que permiten adaptar la forma de trabajo a las condiciones de la iniciativa flexibilizando el proceso de desarrollo y procurando respuestas más inmediatas, orientadas a la construcción de productos mínimos viables en plazos de 15-20 semanas.



- **¿Qué?** Que se apoye en tecnologías digitales innovadoras, aprovechando al máximo sus posibilidades. (por ejemplo, *big data*, *Omnichannel*, *RPA*, *Blockchain*...)
- **Impacto económico.** Que además tenga un impacto económico relevante que contribuya a alcanzar los 800 millones de euros de flujo de caja establecidos en el Plan Estratégico como objetivo para 2022.



¿Cómo identificar una iniciativa digital?

Para saber si nuestra iniciativa puede considerarse digital debemos responder positivamente a las 3+1 premisas



Creemos que en la era digital tenemos que trabajar desde la colaboración para ser mejores, compartiendo en todas nuestras iniciativas unos valores comunes a tener siempre presentes y con unos objetivos y prioridades claros. Para la transformación digital de los negocios se han definido siete palancas digitales de creación de valor:

- Sin fallos y siempre seguro.
- Planificación y programación inteligente E2E (End to End).
- Excelencia en las operaciones digitalizadas.
- Desarrollo de activos optimizados digitalmente.
- Mejora de la experiencia del cliente en cualquier canal o dispositivo, a través de los datos.
- Innovación en nuevos modelos de negocio.
- Organización más ágil y flexible.

5.2.4.

CASOS DE ÉXITO

Dentro del Plan Estratégico 2021-2025 Repsol ha ratificado su apuesta por el despliegue de tecnologías digitales como habilitador de la estrategia de los negocios. En los primeros pasos de la digitalización cada uno de los diferentes negocios y áreas corporativas visualizaron la contribución de la digitalización para el desarrollo de su estrategia, materializado en un portafolio digital de Compañía. En la actualidad, más de 370 iniciativas ya están en marcha.

Por otro lado, convencidos de la importancia de la transformación digital para los negocios y con el objetivo de facilitar la puesta en marcha de iniciativas digitales, desde el Programa Digital se asume el riesgo inicial de los proyectos, financiando las fases iniciales de los proyectos, hasta llegar al MVP. Solo, en el caso de que dicho MVP sea considerado

técnicamente exitoso por el negocio y su impacto económico positivo, pasa a ser financiado por él.

El Plan Estratégico recoge, dos objetivos transversales para el Programa digital:

- La contribución positiva en materia de sostenibilidad, ayudando a Repsol a conseguir su objetivo de ser Net Zero en emisiones para 2050.
- Promover la transformación de formas de trabajo de Repsol y el desarrollo de sus empleados en capacidades digitales. En cuanto a resultados, a cierre de 2020 el Programa Digital ha generado un impacto económico recurrente de 333M€/año, con una inversión media anual de 160M€/año (incluidos CapEx y OpEx) en 2018-2022.



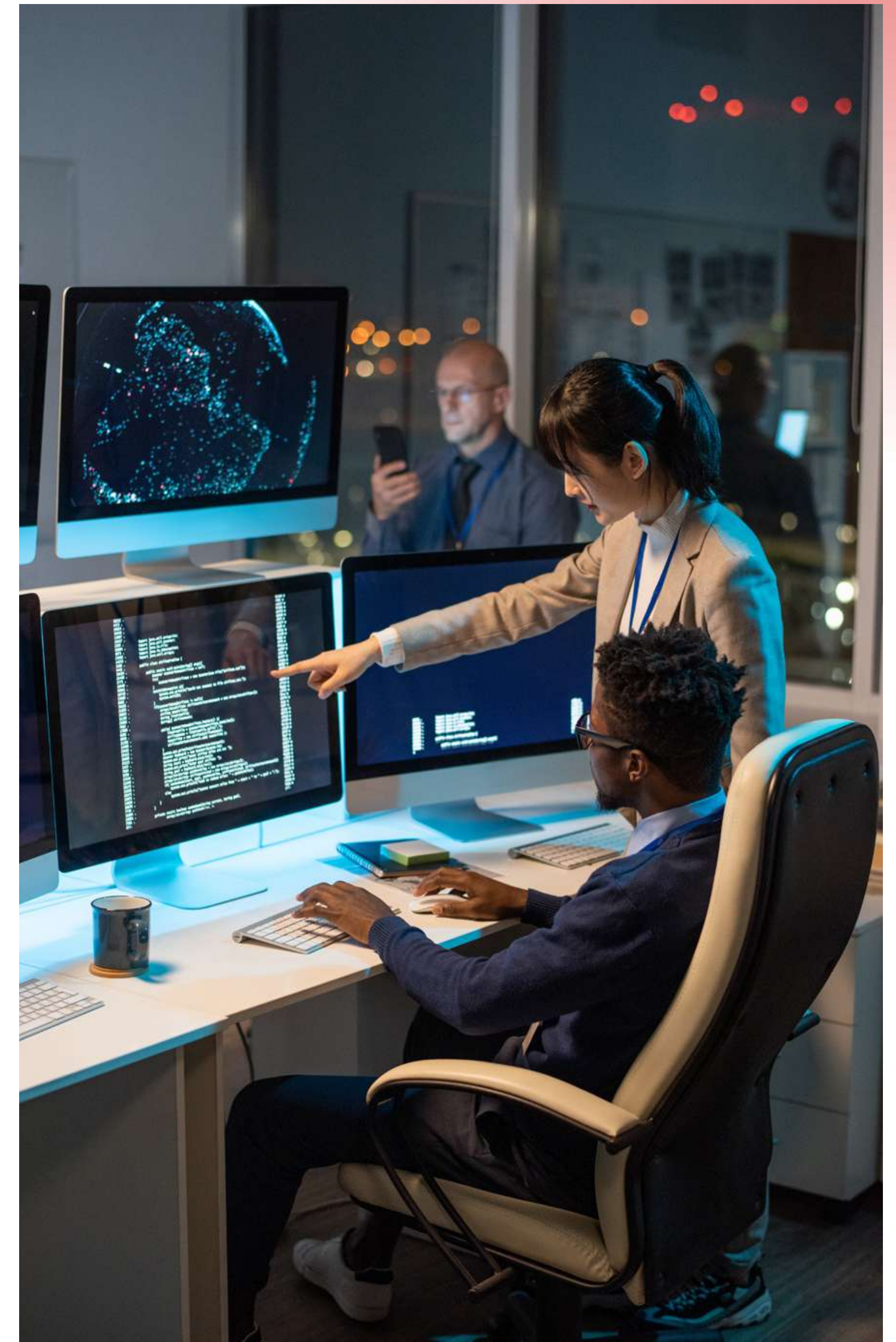
5.2.4.

CASOS DE ÉXITO



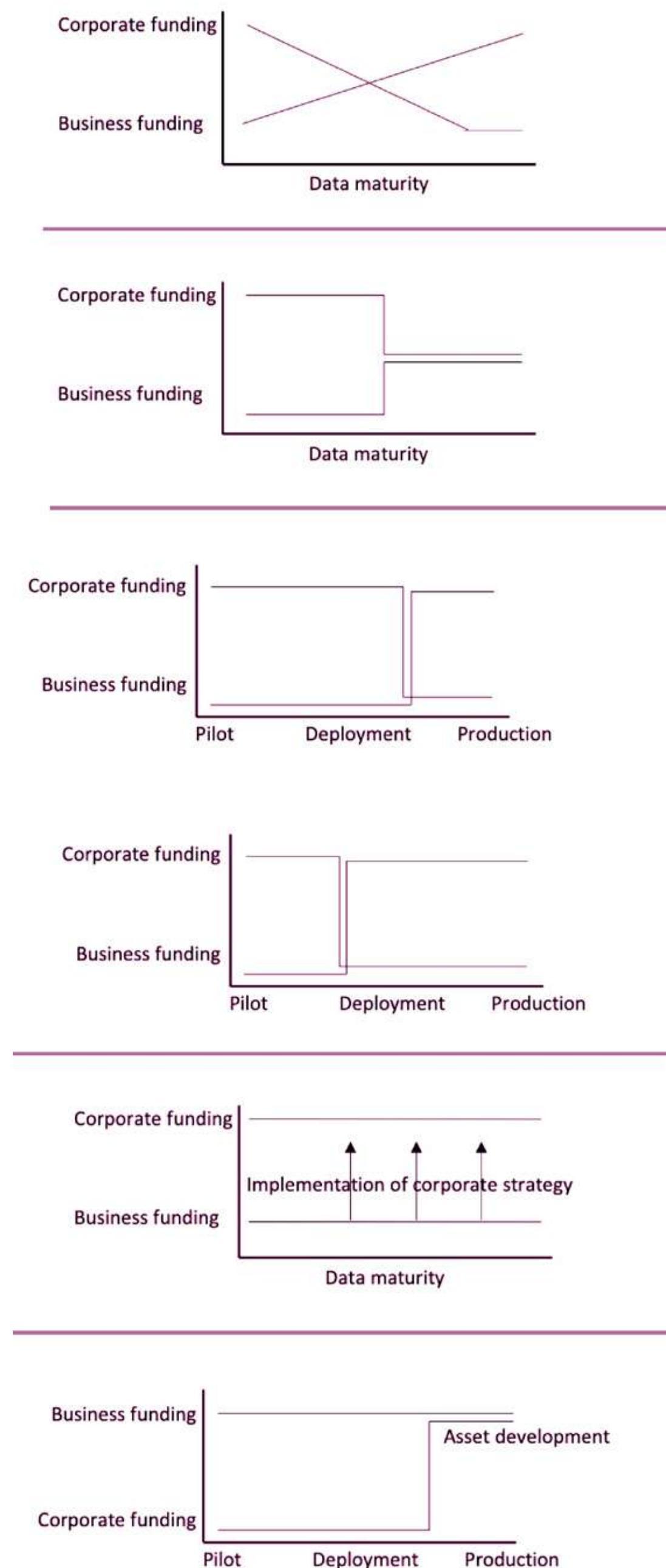
5.2.4.2. Distintos modelos de financiación

Una gran empresa tiene a su disposición muchas maneras de financiar las actividades y recursos asociados a la aplicación del *big data* y la Inteligencia Artificial. La metodología de innovación abierta en sí ya identifica distintas fuentes como la inversión propia, capital riesgo, financiación pública, etc. Pero existe menos experiencia publicada sobre la financiación interna.



5.2.4.

CASOS DE ÉXITO



Muchas grandes empresas son multinacionales con una organización corporativa y varias o muchas unidades de negocio. Entonces, ¿quién financia la inversión en IA? ¿La corporación? ¿Las unidades de negocio? Normalmente, cuando hablamos de innovación muchas veces empieza la corporación con la financiación y con el tiempo y madurez de las iniciativas, las unidades de negocio se encargan de su financiación, poniéndola en sus presupuestos anuales. Este modelo genérico se visualiza en la figura como el primer modelo. Pero en la práctica hay distintas variedades de este modelo, que son los otros modelos que se ven en la imagen:

- **Modelo 2 (izquierdo, medio).** La corporación empieza a financiar, y después de unos años, se financia a medias entre la corporación y la unidad de negocio.
- **Modelo 3 (izquierdo, abajo).** La corporación financia hasta el despliegue, pero la operación la financia la unidad de negocio.
- **Modelo 4 (derecho, arriba).** La corporación financia hasta el piloto, pero el despliegue y la operación la financia la unidad de negocio.
- **Modelo 5 (derecho, medio).** La corporación incentiva el uso de datos e IA financiando un uso específico de interés para el Grupo (no tanto para las unidades de negocio individuales).
- **Modelo 6 (derecho, abajo).** No siempre es la corporación quien lidera la innovación. De vez en cuando una unidad de negocio genera la primera innovación con datos e IA, la financia y después la corporación entra para generalizarlo y prepararlo para llevarlo a las otras unidades de negocio.

Fuente del gráfico. ¿Cuándo es adecuado trabajar con terceros o no?: capítulo 15 de A Data-Driven Company: 21 lessons for large organizations to create value from AI

5.2.4.

CASOS DE ÉXITO

ferroviario

5.2.4.3. Híbrido (negocios + corporativo)

En el caso de Grupo Ferrovial, la fuente de financiación es diversa y depende en muchas ocasiones del nivel de madurez y el grado de incertidumbre de cada caso de uso.

- **Casos de uso experimentales, con un grado de madurez bajo y un nivel de incertidumbre muy alto** se financian habitualmente por Corporación o se buscan fuentes de financiación externas vinculadas a subvenciones públicas o privadas orientadas al ámbito de Innovación.
- **Casos de uso de IA, con una aplicación directa al negocio (corto/medioplazo) pero con un nivel de madurez en cuanto a desarrollo bajo**, se financian por Corporación, como una inversión que permite ganar madurez e ir eliminando incertidumbres, antes de usar la IA en el día a día de una actividad del negocio.
- **Casos de uso de IA, con una aplicación directa al negocio y con un nivel de madurez alto** (soluciones ya probadas con un nivel bajo de incertidumbre), las financia habitualmente cada unidad de negocio y continúa su financiación en la etapa madura de operación.
- **Otros:** en algunos casos, se buscan también **partners estratégicos** que puedan estar interesados en una co-creación de ciertos casos de uso. En estos casos, también pueden participar financiando parte de esos desarrollos.





6.1. Reto 9. Conocimiento y capacidades

6.1.1. Descripción del problema

El 56% de los ejecutivos piensa que “la falta de capacidades analíticas” y los posibles casos de uso” es el principal desafío de la adopción de IA.

El principal activo de las empresas son las personas. Los equipos multidisciplinares y de alto rendimiento son los que consiguen no solo sacar adelante los proyectos, sino que mantienen el know-how y el ADN de la compañía para que siga siendo un ecosistema de valor. Los empleados, sobre todo los que llevan más tiempo, tienen un conocimiento del negocio profundo, aunque en muchas ocasiones les faltan las herramientas para poder dar nuevas respuestas a preguntas existentes apoyadas en la Inteligencia Artificial.

En la actualidad, universidades, escuelas y hasta los propios institutos están lanzando campañas, programas o incluso títulos oficiales para formar a los profesionales en las disciplinas relacionadas con la Inteligencia Artificial. Es labor de los centros educativos, pero también del tejido empresarial que los profesionales sepan enfocar estos *skills* adquiridos en los casos de uso reales. En muchas ocasiones este gap entre conocimientos y aplicación práctica representa un problema para que los profesionales puedan incluir la Inteligencia Artificial en la resolución de problemas desde el principio.

INDESIA

Framework Data & IA
Industrial Marco de
referencia de datos e
inteligencia artificial para el
sector industrial español

INTRODUCCIÓN

ÁMBITO DEL VALOR PARA NEGOCIO

ÁMBITO DE LA CIENCIA DE DATOS

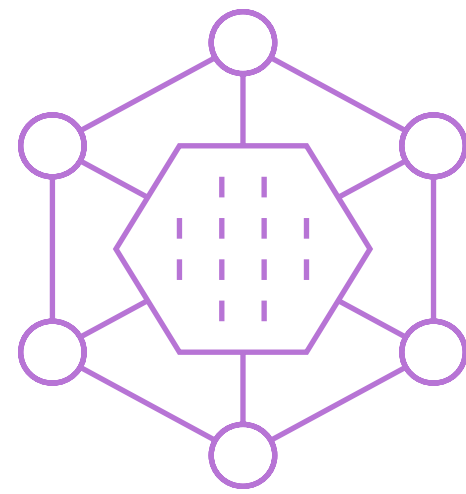
ÁMBITO DE LOS DATOS

ÁMBITO DE LA ORGANIZACIÓN

ÁMBITO DE PERSONAS Y CULTURA

ÁMBITO DE LA SOSTENIBILIDAD

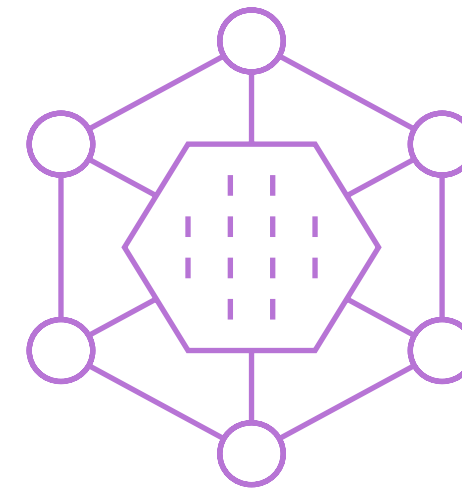
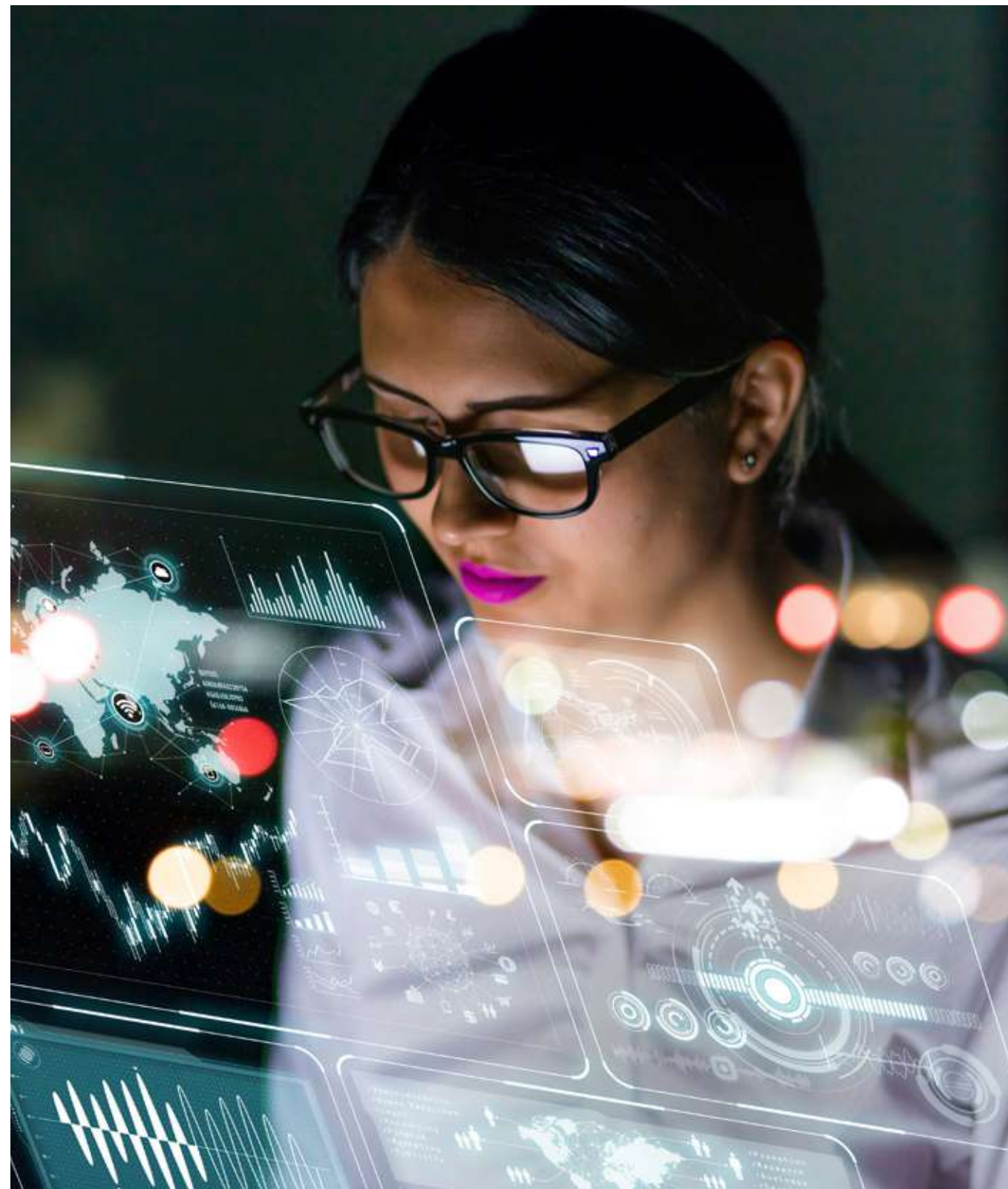
CONCLUSIONES



Además del conocimiento en Inteligencia Artificial, no hay que olvidar los desafíos que supone la gestión de este tipo de proyectos. La implementación de un *framework end to end* a la hora de seleccionar los casos de uso, establecer prioridades, crear los *backlogs* de trabajo e iterar para ir midiendo los resultados son puntos clave para poder medir el retorno de valor. En muchas ocasiones, la propia selección del modelo de IA a aplicar puede condicionar completamente los resultados por lo que hay que adoptar estrategias de 'fallar lo antes posible' para poder ser ágiles en la corrección de los parámetros existentes.

Por tanto, las compañías tienen que apostar por la formación de los profesionales que conocen bien el negocio en las disciplinas que apliquen en su sector relacionadas con la inteligencia artificial. Igualmente atraer a perfiles con menos experiencia, pero con sólidos *skills* técnicos puede suponer que surjan también nuevas preguntas a las que dar nuevas respuestas.





6.1.2. Soluciones

6.1.2.1. Programas de Re-skilling y Up-skilling analíticos para toda la organización

VentureBeat estima que más del 85% de los proyectos de Inteligencia Artificial nunca llegarán a desplegarse en un entorno 100% productivo y estable. Por otro lado, solo en 2018 más de 2 millones de puestos de STEM quedaron libres en los Estados Unidos. Sin embargo, existen numerosos casos de éxito de aplicación de la Inteligencia Artificial en grandes firmas mundiales. Por tanto, estamos en un momento donde la importancia de la IA convive con la dificultad de llevar a cabo este tipo de proyectos. Según las estadísticas, como parece lógico, podemos decir que las empresas que son 'nativas digitales' tienen mayores casos de éxito que aquellas que están en un proceso de transformación.

Teniendo en cuenta el contexto, ¿cuáles son las claves para llevar a cabo este cambio? La formación es una de las palancas más importantes de cambio sobre todo para aquellas sociedades que están en proceso de transición. En muchos casos, se tiene la creencia que IA es únicamente un dominio tecnológico por lo que basta enfocar la estrategia en formación en torno a plataformas o soluciones software lo cual es un error y genera frustración cuando no se logran resultados. La adopción de la Inteligencia Artificial en los pilares de la estrategia de la compañía pasa por varias claves donde algunas de ellas son formativas.

INDESIA

Framework Data & IA
Industrial Marco de
referencia de datos e
inteligencia artificial para el
sector industrial español

INTRODUCCIÓN

ÁMBITO DEL VALOR PARA NEGOCIO

ÁMBITO DE LA CIENCIA DE DATOS

ÁMBITO DE LOS DATOS

ÁMBITO DE LA ORGANIZACIÓN

ÁMBITO DE PERSONAS Y CULTURA

ÁMBITO DE LA SOSTENIBILIDAD

CONCLUSIONES



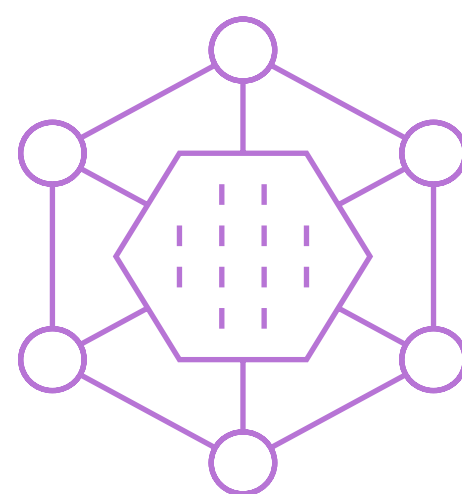
Lo primero de todo pasa por establecer una cultura en torno al dato: es lo que se denomina crear empresas data driven o hacer que la organización tome decisiones basadas en datos. Para ello, no basta solo con una estrategia tecnológica, debe existir también un framework único en donde se identifican roles, modelo de gobierno o la propia calidad de los datos. El éxito de este modelo depende en gran parte de que existan 'guardianes' de las fuentes de datos lo cual incrementará el éxito de los modelos de IA, reducirá el tiempo de procesamiento y por tanto una eficiencia en costes. También es importante que los equipos tengan un conocimiento de las diferentes soluciones y cómo la IA está intrínseca en las mismas.

Si bien no es necesario conocer al detalle el funcionamiento de los algoritmos, es importante igualmente un entendimiento de qué son los modelos de IA, qué ventajas tienen, en qué condiciones no aportan resultados y por tanto un paso más que permita desmitificarlos.

Finalmente, es importante una buena esponsorización desde la dirección de este proceso de cambio, pero no basta solo con eso. Transmitir objetivos donde la IA tiene un peso importante ayuda a traccionar su adopción, pero también es importante que los individuos sientan que forman parte del 'todo'. Para ello, impulsar al mismo tiempo una estrategia Botom-Up donde se generen comunidades y exista una difusión de los logros, lecciones aprendidas o retos ayuda a acelerar el proceso.

Desde el punto de vista formativo, para llevar a cabo el proceso de cambio se requiere de la generación de una estrategia de Re-Skilling y Upskilling. Para ello, son importantes los siguientes aspectos a tener en cuenta:

- Disponer de un mapa claro de los *skills* existentes en la *workforce* de la compañía. Para ello es necesario identificar dicha información en herramientas como las *skills matrix*.
- Diseñar cuál o cuáles son los perfiles que se desean generar, teniendo en cuenta que no es únicamente un proceso de formación en aspectos tecnológicos. Para ello se pueden agrupar los *skills target* según su naturaleza y adecuar los perfiles 'tipo' que se quieran generar.
- Como el proceso de transformación es 360°, las dinámicas de entrevistas o encuestas ayudan también a entender los intereses de los empleados en cuanto a su crecimiento profesional, identificar sus puntos fuertes, etc. Esto ayudará a optimizar los resultados posteriores.
- Todo proceso de cambio requiere de un plan y por tanto será necesario identificar fases, recursos necesarios o incluso ver la forma en la que poder ejecutar dicho plan teniendo en cuenta que la compañía sigue teniendo que cumplir objetivos y por ellos los empleados están asignados. Dicho plan puede incluir una fase inicial de selección de proveedor, herramientas de formación, etc
- Crear comunidades, si puede ser con perfiles heterogéneos, sobre áreas de interés relacionadas con IA. Si es posible, nombrar embajadores o facilitadores en las mismas que ayuden a su vez a otros miembros en el proceso de cambio.
- Si bien el proceso de transformación es interno, es importante tener en cuenta e incluir en los planes de *upskilling* Instituciones Educativas de referencia con las que poder colaborar. En muchos casos, la combinación de perfiles con experiencia del equipo junto con la llegada de perfiles junior en prácticas genera sinergias interesantes y de mucho valor donde ambas partes aprenden mutuamente y sienten que aportan valor.



6.1.3. Niveles de madurez

6.1.3.1. Nivel 1. Fundacional. Divulgación y concienciación

No existe un plan formativo como tal, si lo hay tiene pocas o ninguna habilidad específica de IA. Normalmente se resuelve con campañas de concienciación o divulgación de alto nivel sobre la importación de los datos y la necesidad de explotarlos. En el mejor de los casos hay formación concreta de seguridad e iniciativas individuales o departamentales de planes de acercamiento a AI principalmente hacia capacidades analíticas básicas o enfoques predictivos adicionales.

6.1.3.2. Nivel 2. Acercamiento. Formación externa selectiva

Tras la fase de concienciación y divulgación a alto nivel algunos departamentos más avanzados o con mayor necesidad de negocio solicitan, normalmente a recursos humanos, planes de formación concretos. En este punto se suele optar por seleccionar a los roles más avanzados o con mayor presión por obtener resultados y se les ofrece una formación externa ya sea en un curso reglado por una universidad o academia digital o, si las peticiones son muchas, se contrata a una empresa o formador externo para dar cursos personalizados a grupos selectos dentro de la empresa. Esta formación da lugar a la conciencia sobre la necesidad de nuevos roles, funciones y procesos.

6.1.3.3. Nivel 3. Aspiracional. Academy Inhouse

Como es imposible escalar a muchos departamentos o roles, las empresas más avanzadas optan por crear una academia o plan formativo dedicado dentro de la empresa. Son programas de aprendizajes complejos que buscan mejorar el conocimiento de manera transversal dentro de todas las empresas. El objetivo suele ser un porcentaje alto de la organización y se suele intentar llegar al mayor número de roles posible para democratizar el conocimiento y alcanzar a la mayor población posible dentro de la empresa.

6.1.3.4. Nivel 4. Maduro. MooC 100% todos los empleados

El último nivel y el que muestra mayor madurez es cuando la formación en Inteligencia Artificial y ética es parte obligatoria de la formación de todos los empleados de la empresa. Además de las formaciones y estrategias anteriores se implementan planes formativos a gran escala para asegurar que se alcanza a todas las personas independientemente de su función y tiempo que lleven en la empresa.

Se suelen usar cursos MOOC (Massive Open Online Courses) para que, en cualquier momento, cualquier empleado pueda formarse con formaciones que no requieren presencialidad o estar conectados a un horario concreto y de contrastada efectividad basada en haber formado previamente a miles o cientos de miles de personas. Estos cursos no pueden ser una mera colección de videos que el empleado puede ver o no ver, tienen que incluir lecturas, ejercicios, trabajos revisados por compañeros (peer review) o un proyecto final que es evaluado, como suele ser el estándar en cualquier MOOC de calidad.



6.1.4.1. Data School

Esta escuela del dato nace como un programa de aprendizaje que mejora el conocimiento y las habilidades analíticas de los empleados cubriendo tres niveles en función del conocimiento y habilidades, desde *reskilling*, adquiriendo nuevas habilidades al *upskilling*, mejorando las competencias existentes.

- En un primer nivel el programa **Data 4Everyone** ofrece un conjunto de cursos de corta duración abiertos a todos los empleados, disponibles en la plataforma de Repsol Workday Learning, como punto de partida para ampliar el conocimiento sobre la aplicación práctica del dato.

- En un segundo nivel, con el ambicioso objetivo de alcanzar en 2023 más de 800 Citizen Data Scientist, se ha diseñado con ISDI el **mini Máster Data 4Business**, de aproximadamente 100 horas, donde los alumnos adquieren competencias analíticas que les permitan una mejor explotación de los datos en su entorno, profundizando en temáticas como estadística aplicada a la analítica de datos, modelos en Python, *big data* o la importancia del gobierno de los datos (legislación y normativa).
- En un tercer nivel formativo, orientado a la mejora de personas con un mayor conocimiento (por ejemplo, los Data Scientist), se ofrecen **formaciones adicionales para un perfil más especializado**, junto con una comunidad del dato que permite la interacción.

Repsol *Data School* cubre todo el ámbito y presencia de Repsol, tanto a nivel nacional como internacional, estando actualmente localizado en 30 países. La formación, en los diferentes niveles, está disponible tanto en español como en inglés.

El programa fue lanzado a finales de 2020 cubriendo los niveles anteriormente mencionados.

Actualmente el mini Máster, cuenta con más de 20 ediciones programadas hasta final de 2023, si bien, en función de resultados podría ser extendido. Repsol Data School avanza siguiendo el plan previsto:

CASOS DE ÉXITO

- Han pasado más de 1000 empleados a nivel de sensibilización, Data 4Everyone.
- Ya han finalizado tres ediciones del mini Máster Data4 Business y del programa Data4 Digital (programa que fue precursor de la escuela) con cerca de 100 alumnos. A finales de 2021 ya habían pasado por este programa más de 200 empleados en todo el mundo.
- Repsol quiere ser una organización Data Driven, donde el análisis de datos permita tomar mejores decisiones, para así generar valor y eficiencia en todos los procesos, al tiempo que ofrece la oportunidad de aumentar competencias profesionales y adaptarlas a los nuevos retos del entorno.
- Data4 Everyone que es el conjunto de cursos de corta duración abiertos a todos los empleados tendrá continuidad a través de nuevos cursos que permitan entender cómo puede ayudar el conocimiento de los datos en los objetivos del Plan Estratégico.



- Los mini Máster, Data Business, continuarán hasta 2023, actualmente con una previsión de 800 alumnos, si bien se está evaluando su extensión y crecimiento.

Permanentemente, se irán incorporando nuevas formaciones, con mayor nivel de contenidos, con el objetivo de continuar el desarrollo profesional de los empleados.

El concepto de cultura Data Driven o cultura de los datos, será parte del Índice de Cultura, que sirve de barómetro global de la evolución cultural en Repsol.

CASOS DE ÉXITO



6.1.4.2. Power BI Corner

El Power BI Corner nace con la intención de ser una palanca de transformación para ser una Data Driven Company.

Si el objetivo es que los negocios tomen decisiones basadas en datos, éstos deben ser fiables y fácilmente accesibles en tiempo y forma (dispositivos móviles, equipos corporativos, etc.) y la aplicación seleccionada para ello fue Power BI.

En Repsol se ha comprobado que la forma más rápida de adopción de esta nueva aplicación de reporting es el

autoconsumo. Pero si se quiere que el autoconsumo funcione en la organización, y que los negocios sean independientes de los departamentos de IT tradicionales, es necesario acompañar a las personas en su aprendizaje de las nuevas tecnologías y aplicaciones.

Es por esta razón por la que se creó Power BI Corner, que es un grupo de consultores expertos en Power BI, en formación y en atención a usuarios, cuyos principales objetivos son:

- **Motivación**

Las personas necesitan motivación para adoptar grandes cambios en su forma de trabajar. Por ello, el equipo del Power Bi Corner busca motivar constantemente a las personas. Mostrarles los grandes avances que pueden obtener gracias a una curva de aprendizaje básico muy rápida. Mostrar que la aplicación es “asequible” y las ventajas que tiene el ser “independiente” en sus propios desarrollos, y que nunca caminarán solos en su aprendizaje.

Esta motivación busca evitar el abandono en la adopción de las nuevas aplicaciones y aportar valor al sistema de reporte de los negocios.

- **Aprendizaje**

Una de las funciones del Power Bi Corner es dar formación a todos los niveles. En un principio, las formaciones se crearon calendarizadas, y las personas interesadas se apuntaban a las sesiones. Posteriormente, según fue creciendo la implicación de los negocios, las formaciones se daban a demanda, ajustando los contenidos de las mismas a las necesidades de los asistentes, y a ser posible, con los datos de su propio departamento.

CASOS DE ÉXITO

De esta forma, al terminar la sesión, los asistentes ya se llevaban un boceto de informe sobre sus propios datos, cumpliendo el objetivo de la formación, y la motivación. El modelo ha crecido hasta disponer de sesiones online y cursos on line prediseñados para varios niveles de conocimiento.

- **Resolución de dudas técnicas**

Basado en bookings, disponemos de un servicio de consultas con un equipo experto de Power BI. Las personas pueden reservar su cita de 30 minutos con antelación. La cita, presencial o en remoto, permite al usuario contar con la ayuda de un consultor experto de Power BI, que le resuelve la duda técnica sobre su propio informe.

En estas citas se pueden resolver:

- Dudas técnicas.
- Dudas de modelo de datos.
- Optimización.
- Visualización.
- Publicación y compartición de informes.

Hasta la llegada de la pandemia, el Power Bi Corner consistía en una ubicación física en las oficinas centrales. Esto permitía

crear un punto de referencia para la comunidad de usuarios. Además, al desplazar a las personas, saliendo de su puesto de trabajo, permitía “aislarles” de su día a día y centrar su atención en el aprendizaje de la aplicación.

Con la pandemia, las citas han pasado a ser todas ellas mediante reuniones de Teams, y las comunicaciones de novedades, dudas, comunidad, se realizan todas basadas en Teams.

En resumen, pasamos de un modelo tradicional en el que el negocio solicitaba un informe y los departamentos de IT lo desarrollaban basados en unos requerimientos cerrados a un proceso ágil, en el que los usuarios son independientes para desarrollar sus propios informes, e ir descubriendo nuevas posibilidades de análisis según van construyendo.

6.1.4

CASOS
DE ÉXITO

6.1.4.3. Comunidad de expertos en Workplace + el plan de promocionar profesionales del dato en el año 2015 y la compra de una consultora de *big data*

Reto cultural – Comunidad de expertos e interesados en Machine Learning y *big data*

Las comunidades de aprendizaje y/o desarrollo profesional son un instrumento que se ha venido usando en el grupo Telefónica desde hace más de 10 años. Se pusieron en marcha desde la organización de Recursos Humanos para promover la iniciativa de los empleados en el desarrollo de sus propias habilidades y capacidades, y en paralelo con otras iniciativas de formación lideradas por las áreas.

En el contexto de las comunidades de conocimiento, se fomenta el contacto entre perfiles similares que realizan sus actividades en diferentes áreas del negocio, el intercambio de buenas prácticas, técnicas o aplicaciones y se impulsan actividades formativas internas o externas para lo que se cuenta con un pequeño presupuesto anual.

Las comunidades se mantienen activas mientras existe una masa crítica de participantes que la impulsan y al menos un voluntario que se encargue de la organización de las actividades de la comunidad.

En el caso de la comunidad de Machine Learning y *big data*, ha estado activa con diferentes denominaciones desde el 2014, buscando ser un lugar de encuentro entre empleados con experiencia y/o interés en desarrollar su carrera en el campo de Machine Learning y *big data*. Con más de 1000 miembros es, con diferencia, la comunidad más grande en Telefónica.

La organización de las actividades de la comunidad se realiza a través de un grupo de personas que actúan como embajadores de las diferentes áreas dentro de la empresa, que facilitan un trato cercano y la canalización de oportunidades tanto en sentido ascendente como descendente. Las comunicaciones entre miembros se realizan a través de un grupo en Teams, y otro en Workplace, exprimiendo al máximo las capacidades que nos ofrecen estas herramientas para crear el entorno propicio la conexión personal entre profesionales.

A modo de ejemplo, en el último año la comunidad ha organizado los siguientes tipos de actividades:

- **Keynotes con expertos externos**, con el objeto de acercar nuevos casos de uso, tecnologías y enfoques en el ámbito del aprendizaje automático a los miembros de la comunidad.
- **Formación, tanto interna** (revisión de técnicas novedosas por parte de expertos, presentación de casos de uso, ...) como externa (tanto a medida/en directo, como MOOCs). Anualmente se confecciona un plan de formación teniendo en cuenta los inputs de los miembros. En el plan de formación se busca mapear los intereses de formación de los miembros y las habilidades de los mismos en el contexto del plan tecnológico de la compañía, y en base a ello definir iniciativas de formación interna y externa que se acometerá en el marco de la comunidad en el año.
- **Competiciones (datathons)**, enfocados en la resolución de problemas de negocio por medio del aprendizaje automático. Promueven la conexión entre profesionales (normalmente la participación es en grupos), facilitan el acceso a datasets habitualmente no disponibles y proporcionan un entorno seguro para la puesta en práctica de habilidades adquiridas en procesos de formación.
- **Biblioteca de publicaciones clave en el ámbito.**
- **Colaboración en el programa de doctorado industrial**, para fomentar los estudios de tercer ciclo para la especialización perfiles en el área de Machine Learning entre otras.

- **Reconocimiento a los mejores perfiles de Data Science de la compañía**, como mecanismo para identificar y distinguir a los referentes en el campo, y dar a conocer sus trayectorias para que puedan servir como modelo a otros empleados en el desarrollo de su carrera profesional.

Todo ello como una iniciativa autoorganizada por y para empleados, y con el apoyo del equipo de personas y de la dirección del área de CDO global de Telefónica.





6.1.4.4. Programa smart factory y Programa de arquitecturas híbridas

El sector industrial, como otros muchos sectores de actividad, está atravesando una época de cambios impulsada por el uso y análisis de los datos como base de la toma de decisiones. Esta transformación digital implica un cambio importante en los procesos productivos, desde la recogida de datos con la sensorización y robots colaborativos hasta el análisis de los mismos utilizando técnicas de *Business Intelligence*, que permiten optimizar y mejorar la productividad de los procesos. El experto en Industria 4.0 debe abarcar una gran cantidad de disciplinas y campos, al menos a nivel conceptual, de cara a poder implantar proyectos que transforman el tejido de las empresas. El programa plantea un recorrido de principio a fin por todo el ciclo de vida del proyecto, profundizando en los distintos aspectos de la tecnología: sensores, datos, algoritmos y estrategias de implantación e innovación. La primera edición se lanzó en 2019.

El contenido está orientado a un formato muy práctico bajo los principios de learning by doing donde los alumnos implementan un proyecto piloto end to end. Los perfiles de los alumnos son normalmente profesionales interesados en

la Industria 4.0 con una experiencia contrastada en procesos industriales, especialistas en logística, producción, gestores de innovación o jefes de proyecto entre otros.

En cuanto a los objetivos, el programa se centra en:

- Conocer la interrelación y mutuo beneficio de la Ciencia de Datos (Data Science), Industria 4.0, IoT (*Internet of Things*) y big data.
- Conocer, dominar y gestionar un proyecto basado en datos con casos prácticos de Ciencia de Datos (*Data Science*) y *big data* en el ámbito industrial.
- Profundizar en las diversas arquitecturas de gestión de datos de principio a fin en procesos de industria 4.0.
- Prototipado de sensores de IoT y adquisición de datos en tiempo real
- Transformación almacenamiento de los datos
- Arquitecturas y servicios cloud más relevantes para estos proyectos
- Aprender las técnicas y modelos de Machine Learning tanto supervisados como no supervisados
- Aprender técnicas de visualización avanzada (Visual Data Mining) y manejar herramientas para dicho fin, como PowerBI, en el ámbito de la Industria 4.0.



TECNICAS REUNIDAS

6.1.4.5. Formación en datos e IA. Captación de talento STEM mediante Hackathons con estudiantes.

Técnicas Reunidas tiene diversas iniciativas orientadas al upskilling de sus trabajadores, así como la captación de talento STEM.

En particular, pone a disposición de toda su plantilla la plataforma de cursos

eLearning relacionados con los datos y la inteligencia artificial en la plataforma de formación corporativa, nacida de la colaboración con un proveedor líder de contenidos técnicos en el sector de la ingeniería y construcción.

A través de este portal, se ofrece un amplio rango de materias y contenidos formativos en formato online, con acceso libre, ilimitado y 24/7, entre los que se encuentran cursos dirigidos a la difusión de las nuevas tecnologías como, por ejemplo, "Inteligencia Artificial", "Impacto de la Inteligencia Artificial en la empresa", "Metodologías de ideación y gestión de proyectos de Inteligencia Artificial" o "Introducción al *big data*". Siendo este último, uno de los cursos mejor valorados en la compañía, y con más éxito de demanda.

Otra de las iniciativas que se llevan a cabo son competiciones de innovación en la que estudiantes universitarios de toda España trabajan durante 24 horas sobre uno de los retos de negocio para ofrecer soluciones innovadoras relacionadas con datos e Inteligencia Artificial. Se apuesta por el talento de los jóvenes universitarios y, a través de estas iniciativas se quiere contribuir a impulsar su desarrollo y crecimiento.



Impacto de la
Inteligencia Artificial
en la empresa
Categoría: Nuevas tecnologías



Metodologías de
ideación y gestión de
proyectos de
Inteligencia Artificial
Categoría: Nuevas tecnologías



Inteligencia artificial
Categoría: Nuevas tecnologías



ferrovial

6.1.4.6. Formación externa selectiva y plan de desarrollo para perfiles técnicos en vías de definición. Power BI Corner para perfiles de negocio

Desde Ferrovial impulsan el desarrollo de capacidades de dos maneras: por un lado, financiando programas formativos para perfiles técnicos, de manera que puedan desarrollar sus capacidades en función de su rol y motivación; y por otro, desarrollando manuales y guías de usuario para perfiles de negocio que deseen comenzar a familiarizarse con el desarrollo de soluciones de analítica descriptiva para ganar autonomía e independizarse de los equipos de IT, dándoles soporte y apoyándolos cuando lo necesitan. (Power BI Corner).

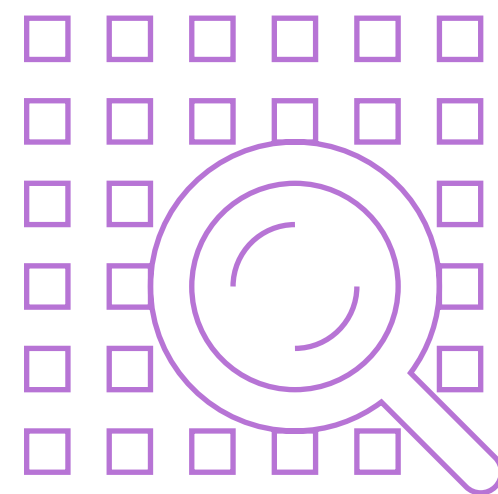
AIRBUS

6.1.4.7. Digital Academy, orientada a desarrollar y profundizar en las siguientes capacidades digitales (Advanced Analytics, A, - Cybersecurity, Robotics, IoT, Mixed reality and smart wearable, Cloud & DevOps...)

AIRBUS tiene una estructura de academias internas para desarrollar las competencias y perfiles definidos en base a las necesidades de la compañía y con enfoque en el sector aeroespacial. Concretamente, para todas las actividades y perfiles relacionados con Indesia, existe la Digital Academy que tiene la misión de identificar las competencias requeridas para los entornos digitales y IT, definir las en base a las actividades que se realizará por parte de los empleados que las tengan en sus jobs y crear planes de formación y desarrollo para poder llegar el nivel de experiencia requerido en cada puesto.

De la misma manera, se definen los Jobs de acuerdo a la misión y actividades a realizar por cada perfil, dotándolos de planes de formación y desarrollo ("Learning Paths") que permitan a los empleados obtener el nivel de conocimientos y autonomía necesarios para su día a día.

La Digital Academy está formada por un equipo de profesionales y expertos transnacionales, ubicados en los principales países donde AIRBUS tiene sede, lo que permite tener una cercanía y conocimiento óptimos del negocio. Entre su portfolio, tiene definidos 17 jobs, que a su vez contienen las necesarias competencias que se definen y priorizan de acuerdo a la Competence Strategy que se revisa anualmente de acuerdo a las prioridades de la Compañía. Estos Jobs y Competencias, están soportados por 15 Learning Paths divididos en perfiles de "iniciación" y "experto", que permiten un amplio camino de desarrollo gracias a soluciones como cursos on-line, presenciales, seminarios, "on the job training", tutorizaciones, etc.



6.2. Reto 10: Adopción y gestión del cambio

6.2.1. Descripción del problema

Uno de los primeros problemas a abordar es el de que la cultura de compañía no reconoce la necesidad de cambiar los procesos de toma de decisiones.

La cultura organizacional es la suma de las prácticas, creencias, símbolos, ritos, valores y expectativas acerca de lo que se considera adecuado, que priman mayoritariamente en una institución. Es “cómo se hacen las cosas aquí”, y se traduce en prácticas concretas y cotidianas, negativas o positivas. Por lo general, todas las compañías públicas o privadas se enfrentan con

obstáculos y fuertes resistencias para cambiar cualquier aspecto de la cultura organizacional existente. La mayor parte de las veces éstos se relacionan con el escepticismo y con la dificultad y el tiempo necesarios para transformar culturas, hábitos y prácticas de trabajo, cuyos actores principales no ven la necesidad del cambio.

Habitualmente la toma de decisiones en las empresas ocurre en base al instinto y la intuición de las personas involucradas en esa tarea, ya sean directivos, mandos medios o personal operacional. La construcción y establecimiento de espacios de datos relevantes para el negocio, permite basar esta toma de decisiones sobre información más precisa, tangible y contrastada con la realidad del negocio empresarial.

Con lo cual uno de los factores clave a tener en cuenta al momento de plantearse la creación de espacios de datos en la compañía, es contemplar que es necesario un meditado, planificado y meticuloso cambio cultural a todos los niveles, especialmente en los niveles donde se realizan las decisiones más trascendentes para el futuro de las compañías. Crear un clima para el cambio, comprometer y habilitar a toda la organización para el cambio, pero muy especialmente implantar y mantener la transformación generada son los pilares para que las resistencias y los obstáculos a los que se enfrentan las compañías puedan ser removidos y aceleran la migración de la toma de decisiones intuitiva a una toma de decisiones basada en la información que genera la propia organización, sus colaboradores y clientes.

INDESIA

Framework Data & IA
Industrial Marco de
referencia de datos e
inteligencia artificial para el
sector industrial español

INTRODUCCIÓN

ÁMBITO DEL VALOR PARA NEGOCIO

ÁMBITO DE LA CIENCIA DE DATOS

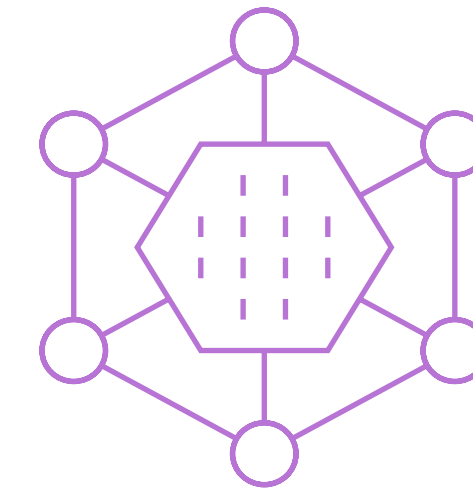
ÁMBITO DE LOS DATOS

ÁMBITO DE LA ORGANIZACIÓN

ÁMBITO DE PERSONAS Y CULTURA

ÁMBITO DE LA SOSTENIBILIDAD

CONCLUSIONES



6.2.2. Soluciones

Cuando la empresa plantea la adopción de implementar una estrategia de toma de decisión basada en los datos que recoge de sus clientes, se encuentra con la oposición de la cultura organizacional como hemos comentado en el apartado anterior.

Para facilitar la adopción y el cambio de mentalidad necesarios para implantar cambios de tal calado en la compañía, planteamos el modelo de Kotter, creado por John Kotter, experto escritor y profesor de la Escuela de Negocios de Harvard, que lleva muchos años mejorando su proceso de 8 pasos hacia al cambio en su libro "Liderando el cambio" ("Leading Change"), publicado inicialmente en 1995.

Paso 1: crear sentido de urgencia

Para que ocurra el cambio, es necesario que toda la empresa realmente lo desee. Desarrolle un sentido de urgencia alrededor de la necesidad de cambio. Esto puede ayudar a despertar la motivación inicial para lograr un movimiento. No es simplemente mostrar una estadística de ventas o hablar de la creciente competencia. Es abrir un diálogo honesto y convincente acerca de lo que está pasando en el mercado y con su competencia, y cómo puede contribuir el disponer de datos sobre sus clientes y competidores a construir una ventaja competitiva a la organización. Si mucha gente empieza a hablar acerca del cambio que propone, la urgencia puede construirse y alimentarse a sí misma.

Qué hacer:

- Identificar potenciales amenazas y desarrollar escenarios que muestren lo que podría suceder en el futuro.
- Examinar oportunidades que deben ser o podrían ser explotadas, si disponemos la información que nos pueden proporcionar los datos.
- Iniciar debates honestos y dar razones convincentes para hacer a la gente pensar y hablar.
- Solicitar el apoyo de clientes para reforzar sus argumentos, si fuera necesario.

Kotter sugiere que, para que el cambio ocurra con éxito, el 75% de los directores y gerentes de la compañía deben “comprar” el cambio. En otras palabras, debe trabajar duro en este primer paso e invertir mucho tiempo y energía en la construcción de una urgencia, antes de pasar al siguiente paso.

Paso 2: formar una coalición poderosa

Convencer a la gente de que el cambio es necesario. Esto a menudo implica un fuerte liderazgo y soporte visible por parte de la alta dirección de la organización. Gestionar el cambio no es suficiente. También tiene que liderarlo. Se pueden encontrar líderes del cambio dentro de la empresa. Para liderar el cambio, se debe

reunir una coalición o equipo de personas influyentes cuyo poder proviene de una variedad de fuentes, incluyendo los puestos que ocupan, status, experiencia e importancia política.

Una vez formada, su “coalición” necesita trabajar como equipo, en la continua construcción de la urgencia y del impulso en torno a la necesidad del cambio.

Qué hacer:

- Identificar los verdaderos líderes de su organización.
- Pedir un compromiso emocional.
- Trabajar en equipo en la construcción del cambio.
- Identificar áreas débiles dentro del equipo y asegurarse de que tiene una buena mezcla de personas de diferentes departamentos y diferentes niveles de la empresa.

Paso 3: crear una visión para el cambio

Al empezar a pensar en un cambio, probablemente habrá muchas grandes ideas y soluciones dando vueltas. Hay que vincular esos conceptos con una visión general que la gente pueda entender y recordar fácilmente.

Una visión clara puede hacer entender a todos el por qué se les pide que hagan algo. Cuando las personas ven por sí mismas lo que están tratando de lograr, las directivas que les son dadas cobran más sentido.

Qué hacer:

- Determinar los valores que son fundamentales para el cambio.
- Elabore un breve resumen que capture “lo que ve” como futuro de la organización.
- Cree una estrategia para ejecutar esa visión.
- Asegúrese de que su coalición pueda describir la visión en 5 o menos minutos.
- Practique su “declaración de la visión” a menudo.

Paso 4: comunicar la visión

Lo que se haga con la visión después de crearla determinará su éxito. El mensaje posiblemente encuentre fuertes competencias en las comunicaciones diarias dentro de la empresa,

por lo que debe comunicar frecuentemente y con fuerza, e incluirla dentro de todo lo que haga.

No hay que limitarse a llamar a reuniones extraordinarias para comunicar su visión. Hay que hablar de la visión cada vez que pueda y de la importancia de basar las decisiones de la empresa en datos tangibles e información fiable, no en el instinto empresarial de cada área. Use diariamente la visión para tomar decisiones y resolver problemas. Cuando se mantenga fresca en la mente de todos, ellos recordarán la visión y actuarán en respuesta a ella.

También es importante “hacer camino al andar”. Lo que se hace es mucho más importante – y creíble – que lo que usted dice. Demostrar el tipo de comportamiento que se espera de los demás, argumentando las decisiones sobre datos y usando la información disponible.

Qué hacer:

- Hablar a menudo de su visión de cambio y por qué es importante acometerlo.
- Responder abierta y honestamente a las preocupaciones y ansiedades de la gente.
- Aplicar su visión en todos los aspectos operativos, desde la formación hasta la evaluación de la performance.
- Predicar con el ejemplo.



Paso 5: eliminar los obstáculos

Si se siguen estos pasos y se llega a este punto en el proceso de cambio, es porque se ha hablado de la visión y se ha construido la adhesión a ella desde todos los niveles de la organización. Lo ideal es que su gente quiera ocuparse del cambio y conseguir los beneficios que ha estado promoviendo.

Pero, ¿hay alguien que se resista al cambio? ¿Existen procesos o estructuras que están en el camino?

Poner en marcha la estructura para el cambio y comprobar constantemente las barreras que existan. La eliminación de los obstáculos puede potenciar a las personas que se necesita para ejecutar esta visión y poder ayudar a avanzar en el cambio.

Qué hacer:

- Identificar personas nuevas que sean líderes del cambio y cuyas funciones principales sean hacer el cambio.
- Mirar la estructura orgánica, puestos, y sistemas de recompensas para asegurarse de que están en consonancia con la visión.
- Reconocer y recompensar a la gente que trabaja para el cambio.
- Identificar a las personas que se resisten al cambio y ayudarlas a ver que lo necesitan.
- Adopte medidas para eliminar las barreras (humanas o no).



Paso 6: asegurar triunfos a corto plazo

Nada motiva más que el éxito. Hay que dar a la empresa el sabor de la victoria en una fase temprana del proceso de cambio. Dentro de un breve período de tiempo (que podría ser un mes o un año, dependiendo del cambio), van tener que lograr resultados palpables por su gente. De otra manera, la gente crítica y negativa podría dañar el proceso.

Crear metas a corto plazo y no sólo un objetivo a largo plazo. Hay que desear que cada pequeño logro sea posible, con poco margen para el fracaso. El equipo de trabajo para el cambio puede tener que trabajar muy duro para llegar a los objetivos, pero cada "victoria" a corto plazo puede ser muy motivadora para todo el personal.

Qué hacer:

- Buscar proyectos de éxito asegurado, que pueda implementar sin la ayuda de aquellos que sean críticos del cambio.
- No elegir metas tempranas que sean costosas. Hay que poder justificar la inversión de cada proyecto.
- Analizar cuidadosamente los pros y contras de cada proyecto. Si no hay éxito en la primera meta, se puede dañar enteramente la iniciativa de cambio.
- Reconocer el esfuerzo de las personas que le ayudan a alcanzar los objetivos.

Paso 7: construir sobre el cambio

Kotter sostiene que muchos proyectos de cambio fallan porque se declara la victoria muy tempranamente. El cambio real sucede muy profundamente.

Las victorias tempranas son solo el comienzo de lo que se necesita hacer para lograr los cambios a largo plazo.

Alcanzar un incremento en las ventas de un producto usando la nueva información obtenida de los datos que se tienen está muy bien. Pero si se puede incrementar en 10 productos, eso significa que el nuevo sistema de toma de decisiones funciona. Para llegar a esa décima victoria, deberá seguir buscando mejoras.

Cada victoria proporciona una oportunidad para construir sobre lo que salió bien y determinar qué se puede mejorar.

Qué hacer:

- Después de cada victoria, analizar qué salió bien y qué se necesita mejorar.
- Fijarse más metas para aprovechar el impulso que se ha logrado.
- Aprender sobre Kaizen, la idea de la mejora continua.
- Mantener ideas frescas sumando más agentes y líderes del cambio.

Paso 8: anclar el cambio en la cultura de la empresa

Por último, para lograr que cualquier cambio se implante, éste debe formar parte del núcleo de la organización. La cultura organizacional a menudo determina qué hacer, por lo que los valores detrás de su visión deben mostrarse en el día a día.

Hacer que los esfuerzos continuos para garantizar el cambio se vean en todos los aspectos de la organización. Esto ayudará a darle un lugar sólido al cambio en la cultura de la organización. También es muy importante que los líderes de su empresa sigan apoyando el cambio. Esto incluye el personal existente y los nuevos líderes que vayan sumando. Si pierde el apoyo de estas personas, podría terminar donde empezó.

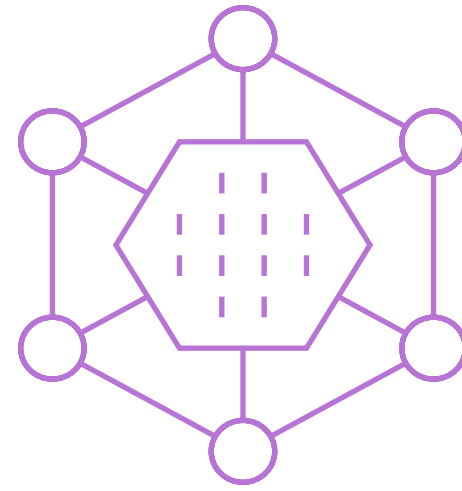
Qué hacer:

- Hablar acerca de los avances cada vez que se dé la oportunidad. Contar experiencias de éxito sobre procesos de cambio, y repetir otras experiencias que oiga.
- Incluir los ideales y valores del cambio cada vez que contrate y entrene gente nueva.
- Reconocer públicamente los principales miembros de su coalición de cambio original, y asegurarse de que el resto del personal, nuevos y viejos, se acuerden de sus contribuciones.

- Crear planes para sustituir a los líderes principales del cambio, a medida que éstos se vayan. Esto ayudará a asegurar que el cambio no se ha perdido o quede olvidado.

Con este método, igualmente se debe aplicar el sentido común y realizar un cambio de mentalidad en las esferas ejecutivas de decisión para que los cambios permeen entre todos los colaboradores de la organización y se lleve a cabo de forma efectiva a todos los niveles.

Crear un sentido de urgencia, contratar poderosos líderes de cambio, construir una visión y comunicar de manera eficaz, eliminar los obstáculos, crear triunfos a corto plazo, y construir sobre el cambio. Si se hacen estas cosas, pueden ayudar a hacer del cambio parte de la cultura organizacional.



6.2.3. Niveles de madurez

6.2.3.1. Nivel 1. Fundacional. Early adopters

No Existe cierta conciencia sobre la necesidad de aplicar Inteligencia Artificial y tener espacios de datos compartidos, pero solo en ciertos departamentos o áreas de negocio. Los distintos grupos aplicando IA no suelen compartir buenas prácticas o incluso hablar entre ellos. No existe una idea clara sobre los atributos culturales necesarios para fomentar el uso de IA en términos de colaboración, intercambio, experimentación, innovación y mentalidad de crecimiento que

normalmente favorecen la construcción y el uso de sistemas, productos o servicios de IA.

A nivel de liderazgo en la organización, es posible que los top executives no sean conscientes de la importancia de su papel en la creación de capacidades organizativas de Inteligencia Artificial, o de cómo apoyar los esfuerzos de los *early adopters*.

Incluso puede haber cierto escepticismo sobre el valor de la IA. Es posible que la mayoría de los empleados no estén al tanto de las iniciativas existentes y si lo están, puede percibir la IA más como una amenaza que como un valor.

6.2.3.2. Nivel 2. Acercamiento. Comunidad

Se hace evidente la necesidad de datos y todo lo relacionado con datos (tener datos de calidad, compartirlos y frenar silos, tener las herramientas y los medios para procesar grandes volúmenes de datos, etc.) pero solo alrededor de ciertas comunidades de conocimiento o grupos de negocio concretos. Existen comportamientos iniciales de adopción y gestión del cambio, pero es posible que aún no sean efectivos o solo funcionen en ciertas áreas de la organización muy concienciadas del valor inmediato de negocio.

Se empieza a crear la necesidad de tener equipos multidisciplinares donde se une ingeniería de datos, científicos de datos y especialistas de IA. A nivel de liderazgos en la organización el apoyo a IA empieza a ser visible. Los líderes comunican el valor de la IA y brindan los medios para abordar los esfuerzos de la IA, como PoC o Pilotos, pero siempre dentro de comunidades o áreas de negocio limitadas.

Existen algunas iniciativas de adopción de IA enfocadas en ayudar a los empleados a comprender el valor de la IA y disminuir la resistencia.



6.2.3.3. Nivel 3. Aspiracional. Agentes del cambio

Las iniciativas de gestión de datos empiezan a alcanzar a toda la organización. Esto incluye plataformas de datos y tecnologías y ecosistemas en la nube. Los datos se consideran una pieza fundamental para definir las iniciativas de IA con éxito y existen agentes del cambio que se encargan de contagiar esta sensación más allá de las comunidades o negocios más avanzados. El empoderamiento y la colaboración efectiva de los empleados son una realidad que genera ideas y fomenta la innovación. Se crean y sostienen equipos multidisciplinares en toda la organización promovidos por aquellos empleados que fueron early adopter y ahora pueden compartir sus experiencias de éxito y errores de los que aprender.

Los errores y la toma de riesgos se aceptan comúnmente como una forma de aprender, crecer e innovar a través de la IA.

A nivel de liderazgo, hay una clara participación de C-Level. Los líderes están presentes, activos y apoyan visiblemente a la IA. Se puede decir que existe un enfoque estructurado para la adopción de la IA y las herramientas y los planes de gestión del cambio se utilizan de forma eficaz.

6.2.3.4. Nivel 4. Madurez. Citizen Data Scientist

Existe una cultura de datos sólida y consolidada en la que capturar, compartir y utilizar datos en todos los niveles es parte de las organizaciones en Business as Usual. La organización se ha transformado para tener una mentalidad de toma de decisiones basada en datos hasta el punto de que un alto porcentaje de personas que trabajan con datos se consideran Citizen Data Scientist. Citizen Data Scientist o ciudadano científico de datos es una persona que trabaja en campos distintos a la estadística y el análisis pero que crea o genera modelos que incorporan análisis predictivo o prescriptivo de manera natural en su día a día.

A nivel global la organización adquiere una mentalidad de crecimiento e innovación donde la experimentación interactiva y la mejora continua se han convertido en parte del ADN de la empresa.

Los líderes de la organización, a todos los niveles, continúan apoyando y comunicando el valor de la IA como una forma de reforzar la adopción.

CASOS DE ÉXITO



6.2.4.1. HackIA

HackIA es un proceso que permite identificar ideas de negocio o retos de negocio que pueden ser resueltos mediante el uso de datos, de analítica, de algoritmos y de inteligencia artificial. Es también un instrumento de aceleración de productos “Data” basado en la filosofía de hackathon, con la ayuda de empresas colaboradoras especializadas en analítica avanzada.

Dentro de HackIA, se consideran aquellos retos analíticos donde se aplique analítica avanzada e Inteligencia Artificial, por tanto, casos que impliquen descubrimiento de insights, predicción, prescripción o que necesiten servicios cognitivos. No se consideran los casos de analítica descriptiva. En líneas generales, se trata de casos con alta incertidumbre y con alto potencial de valor, es decir, probar hipótesis con mucha incertidumbre minimizando el riesgo. Todo ello debe estar soportado por una sponsorización clara y fuerte de negocio.

Pone foco por tanto en:

- Acelerar todas las fases del proceso de construcción de un caso de analítica avanzada.
- Controlar la inversión a lo largo de todo el proceso, con la filosofía fail soon, minimizando el riesgo con puntos de salida en su ciclo de vida.
- Ampliar el ecosistema abriéndolo a nuevos partners, especialistas de garantía en Data e Inteligencia Artificial.
- Impulsar la cultura del dato (Data Driven) a toda la organización, con la visibilidad que aporta el formato de Hackathon.

Con este formato, se logra la realización de un Producto Mínimo Viable (MVP) acotado, evitando alcances muy grandes que aumenten la inversión y el riesgo.

El proceso se divide en las siguientes etapas:

- **Set-up:** se establece un proceso de registro continuo de casos y mediante revisión se acuerdan los que participan en cada edición de HackIA. La fase de Set-Up realiza la función de puerta de entrada, según los criterios indicados.

INDESIA

Framework Data & IA
Industrial Marco de
referencia de datos e
inteligencia artificial para el
sector industrial español

INTRODUCCIÓN

ÁMBITO DEL
VALOR PARA NEGOCIO

ÁMBITO DE
LA CIENCIA DE DATOS

ÁMBITO DE
LOS DATOS

ÁMBITO DE
LA ORGANIZACIÓN

ÁMBITO DE
PERSONAS Y CULTURA

ÁMBITO DE
LA SOSTENIBILIDAD

CONCLUSIONES

- **Preparación:** en esta fase, se arranca de lleno con los casos seleccionados. Su objetivo es preparar los casos analíticos lo máximo posible para dotar de la mayor agilidad a la siguiente fase, así como preparar el entorno técnico.
- **Semana HackIA / evento:** con la premisa del caso analítico preparado de la fase anterior, esta fase se centra en la extracción del máximo valor a través del trabajo en equipo en formato Hackathon, durante 5 días. Al finalizar la semana los equipos exponen el resultado y compiten por el premio
- **Post HackIA:** en base a los resultados y aprendizajes de la semana HackIA, se completará el detalle del alcance del MVP y se elaborará un Business Case de alto nivel que permita una rápida aprobación y continuidad del caso.
- **MVP HackIA:** su principal objetivo es conseguir un Mínimo Producto Viable acorde a las necesidades de negocio en un periodo de 8-12 semanas aprox. Si es aprobado el MVP daría paso al despliegue final y escalado.

Como clave de HackIA, las distintas fases del proceso involucran a los negocios de cada caso con los roles de Sponsor y Product Owner como factores diferenciales del éxito de este proceso. Junto a ellos se dota de equipos internos y externos que, con ayuda de datos y plataforma, proveen las capacidades necesarias para resolver los retos planteados.





TECNICAS REUNIDAS

6.2.4.2. Plan estratégico de innovación digital para fortalecer el intraemprendimiento y desarrollar la innovación abierta en nuevas tecnologías, datos, e IA.

Dentro del plan de transformación digital de Técnicas Reunidas tiene el objetivo de desarrollar y fortalecer el ecosistema de innovación digital durante los próximos años. Durante el primer año se enfocó en el desarrollo de la innovación interna, intraemprendimiento, y posteriormente se incorporó la innovación abierta. El objetivo es consolidar la cultura de innovación y que siga creciendo su ecosistema como pilar en el desarrollo de iniciativas de datos e inteligencia artificial.

Técnicas reunidas tiene identificadas las áreas de oportunidad de mayor potencial y las macro-tendencias de ideación y ejes estratégicos. Así mismo, también se han creado los órganos de gobierno internos para la ejecución del plan de Transformación Digital de la compañía, pero para acelerar y permeabilizar en toda la organización con la cultura de innovación digital se está desarrollando una plataforma colaborativa de innovación que permite, entre otros:

- Evangelizar la cultura digital.
- Plantear desafíos digitales.
- Promover la búsqueda de soluciones.

- Activar programas de gamificación.
- Articular el ecosistema de innovación cerrada y abierta.
- Capturar y priorizar ideas.
- Gestionar el ciclo de vida de casos de uso de innovación.
- Maximizar el valor de los proyectos.

El proceso de innovación de Técnicas Reunidas está diseñado como un proceso dinámico y sistemático durante todo el ciclo de vida. Debe identificar nuevas oportunidades e ideas, desarrollar y hacer pruebas de concepto, desarrollo de MVPs, pruebas y validación con usuarios, definición de casos de negocio, escalado, e industrialización. Está basado en tres grandes fases:

- **Generación de ideas**

Esta fase debe dar respuesta a la definición de retos, ideación, descubrimiento, conceptualización, priorización, y desarrollo de ideas, que permita generar un mapa de iniciativas.

- **Incubación**

Las ideas, previamente definidas y priorizadas, deben ser sometidas a testeos y validaciones de acuerdo a las opciones de implementación definidas en el marco metodológico en función de la naturaleza del proyecto o idea.

- **Escalado**

Una vez definido el caso de negocio y determinado el go/no go, el marco metodológico define los diferentes escenarios del proyecto.

En definitiva, el programa de innovación pretende mejorar la cultura corporativa de Técnicas Reunidas y la disposición al cambio en un proceso continuo de mejora.

6.1.2

CASOS DE ÉXITO

ferrovial

6.2.4.3. Data communities

Para poder promover una cultura Data Driven dentro de la organización y que cada vez haya más personas pensando en el valor que pueden obtener en el uso de datos e IA, se han creado las Global Data Communities. Tienen como objetivo, por un lado, dar a conocer las iniciativas de datos ya existentes y su impacto en el negocio y, por otro, compartir retos entre diferentes departamentos, de manera que se facilite el intercambio de ideas y se despierten nuevas inquietudes en la organización.

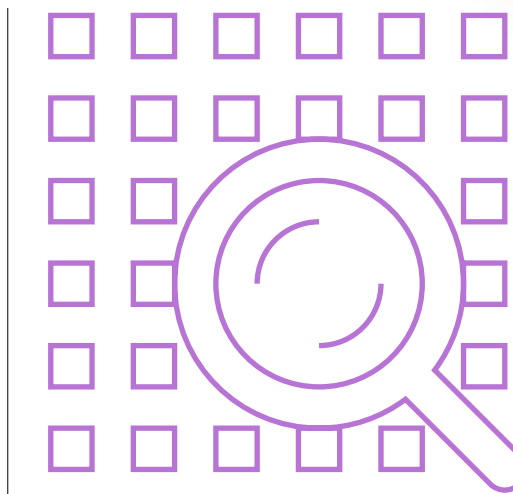
Para que estos eventos sean más abiertos, se invita a esas sesiones a ponentes o expertos en distintas materias en torno al dato y a la Inteligencia Artificial, con el fin de enriquecer las sesiones e intercambiar ideas con reconocidos profesionales del sector.





07

ÁMBITO DE LA SOSTENIBILIDAD



7.1. Reto 11: Ética de la Inteligencia Artificial

7.1.1. Descripción del problema

La inteligencia artificial se usa cada vez más porque brinda enormes oportunidades para los negocios, la sociedad y para nuestras vidas. Sin embargo, recientemente se ha expresado una preocupación por el uso de la IA, en particular en relación con la posible discriminación (sesgo, discriminación, paridad predictiva), la falta de interpretabilidad de las conclusiones algorítmicas (explicabilidad, problema de la caja negra) y la falta de transparencia de los datos personales utilizados.

Hay varias características porque hay que considerar el impacto ético de la Inteligencia Artificial. Un sistema de IA:



- **Siempre tiene un margen de error.** ¿Es ese margen aceptable socialmente? Si un sistema de IA hace una diagnosis de una enfermedad grave con una exactitud de 85%. ¿Es suficiente? ¿Cuál es el impacto para el 15%?
- **Puede tomar decisiones autónomas.** ¿Con qué problemas nos sentimos cómodos que un sistema de IA tome la decisión final? ¿Cuándo es necesario un humano?
- **Puede amplificar sesgos no deseados.** Si un sistema de IA para la contratación de personal sugiere a los candidatos a considerar y rechazar, existe el riesgo de que el sistema se base en el género típico de las profesiones tradicionales (enfermera, técnico, ...).
- **Puede ser difícil de entender por las personas.** Los algoritmos más potentes de IA se basan en el aprendizaje profundo que son redes con

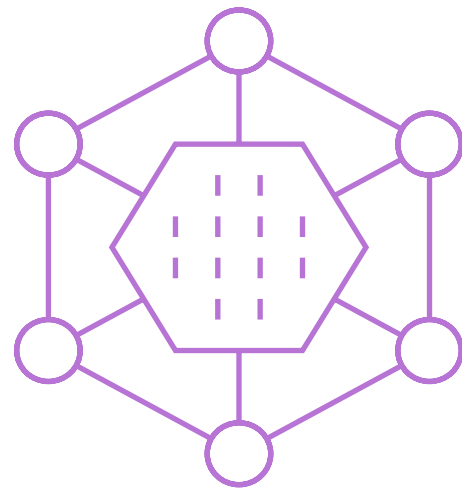
nodos, relaciones, pesos y capas tan complejas que nadie puede entender cómo funcionan.

Estos factores son relevantes porque la IA se está integrando en todos los aspectos de la vida y su uso será aún más masivo en el futuro. Por todo eso, es importante considerar los aspectos sociales y éticos durante todo el ciclo de vida de un sistema de IA. En el sector industrial hay algunos aspectos específicos.

Para evitar estos problemas éticos y sociales, muchas organizaciones han adoptado lo que llamamos principios de inteligencia artificial, o códigos éticos que especifican que el uso de esta tecnología tiene que ser justo, transparente, explicable, con privacidad y seguridad, respetando los derechos humanos, etc. De hecho, hoy en día ya hay muchos estudios que analizan todos estos principios y lo intentan clasificar en distintas categorías.

Pero el reto no es solo la definición de los principios, sino también aterrizarlos en el día en día en una organización concreta. Se está hablando de metodologías como el uso responsable de la IA desde el diseño, que ayudan con el proceso de implementar estos principios de. Estas metodologías contienen como ingredientes: los propios principios, un conjunto de preguntas y recomendaciones que garantizan que todos los principios de la inteligencia especial sean debidamente examinados en todo el ciclo de vida del producto, herramientas que ayudan a responder a algunas de las preguntas, formación y concienciación tanto técnica como no técnica, y por fin un modelo de

Gobernanza. En otro informe de IndesIA describimos en detalle el reto del uso ético de la IA y cómo aplicarlo en una organización.



7.1.2. Soluciones

Para tratar este problema ético de la IA se puede aplicar una metodología denominada “el uso responsable de la IA desde el diseño”. Considera los aspectos claves del impacto del uso de la IA durante todo el ciclo de vida del sistema, desde su diseño hasta su operación y mantenimiento.

Aplicar la ética a la inteligencia artificial consiste en:

- Decidir como organización qué principios éticos adoptar.
- Implementar estos principios éticos en la organización.

7.1.2.1. Seleccionar los principios adecuados

En los últimos años ha habido mucho sobre cuáles son los principios éticos de la IA. Gobiernos de todo el mundo como la Comisión Europea, EE. UU, China han definido sus principios. Instituciones internacionales como la OECD, UNESCO, WEF también han propuesto principios éticos de IA. ¿Cómo puede una organización concreta llegar a la elección adecuada de unos principios?





Hay varias consideraciones que pueden ayudar:

- Elegir aquellos principios que son accionables por la organización. La transparencia y la robustez de los sistemas de IA son accionables por una organización. Al contrario, el futuro del trabajo o la relación entre personas y robots es más la responsabilidad de un gobierno.
- El sector de la organización. Una empresa de servicios suele tener otros riesgos de la IA que una empresa industrial. Para la primera son más relevantes los principios de transparencia y no-discriminación, mientras que el sector industrial son más relevantes principios alrededor de la robustez, autonomía, y seguridad física.
- Existen principios éticos relacionados con aspectos específicos de la IA y otros que cubren aspectos para cualquier sistema digital. Explicabilidad, autonomía y no-discriminación (sesgos) están muy relacionados con la IA mientras que principios sobre la privacidad, seguridad, y robustez son mucho más aplicables. Es por eso importante decidir si los principios de IA se enfocan en los aspectos específicos de la IA o si son más amplios. Por ejemplo, los principios del HLEG de la CE cubren todos los aspectos.

7.1.2.2. Implementarlos en la organización

Una metodología puede ayudar a implementar los principios. Consiste en:

- Los propios principios.
- Formación y concienciación.
- Un cuestionario con varias preguntas y sugerencias para cada principio.
- Herramientas técnicas.
- Modelo de gobernanza.

Otro aspecto importante a tener en cuenta es el posible daño que puede causar un error del sistema de IA. No es lo mismo el impacto de un error en la recomendación de una película y un diagnóstico erróneo de cáncer. Y por eso, las actuaciones para conseguir un uso ético serán distintas según el posible daño. El daño se puede cuantificar hasta cierto punto pensando en:

- La gravedad del error.
- La escala del daño.
- La probabilidad de que ocurra el daño.

7.1.3. Niveles de madurez

Hay distintos niveles de madurez en las organizaciones cuando hablamos del uso ético de la IA.

7.1.3.1. Nivel 1. No ética

En el nivel más bajo, la organización no considera aspectos éticos de los sistemas IA en las decisiones relevantes. Solo se consideran aspectos de negocio. La mayoría de las empresas en España se encuentran en este nivel.

7.1.3.2. Nivel 2. Principios

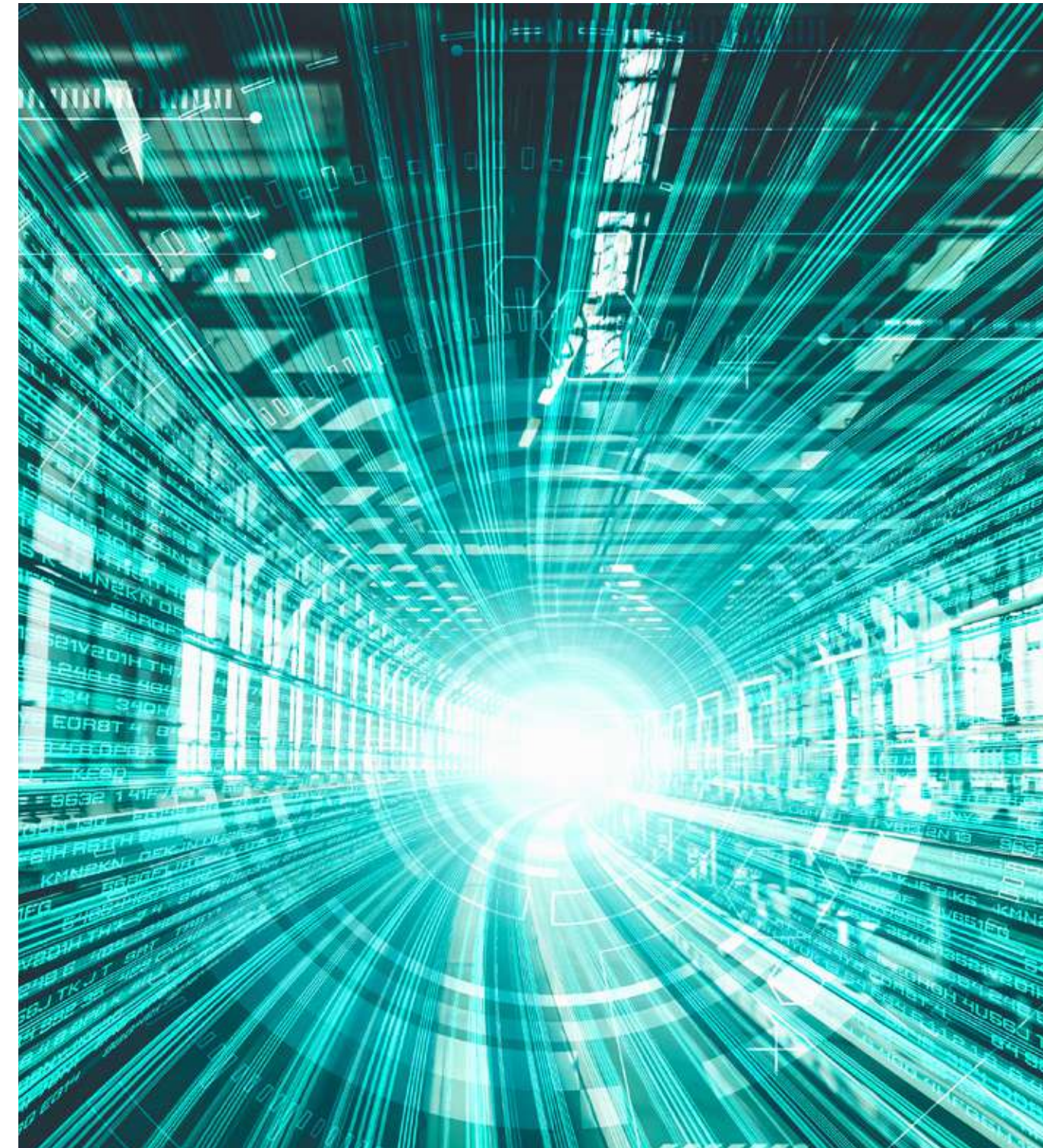
En este segundo nivel, la organización ha definido y publicado los principios éticos que quiere respetar en las decisiones sobre el uso de la IA para su negocio. Pero, no ha definido cómo implementarlo en su organización, es decir, lo deja al criterio de cada uno para respetar estos principios y no tiene manera de verificar si lo está haciendo bien o no. Muchas de las empresas grandes (IBEX) y las internacionales se encuentran en este nivel.

7.1.3.3. Nivel 3. Herramientas

En este nivel, aparte de los principios éticos, existen herramientas que ayudan a los empleados a tratar los aspectos éticos del uso de la IA. Puede existir, por ejemplo, un cuestionario con preguntas éticas que se tiene que completar para cada producto o proyecto que usa IA. O herramientas para detectar y/o mitigar sesgos en los datos o los algoritmos. Pero aún falta una coordinación completa en la organización que asegure el cumplimiento ético.

7.1.3.4. Nivel 4. Gobierno

El nivel más maduro implica que la ética de la IA está integrada en toda la organización con claras responsabilidades, roles y procesos. Requiere una conversación transversal de la organización para acordar cómo se relacionan los distintos departamentos implicados. Existe un modelo de gobernanza con un modelo operativo que deja claro en cada momento qué hay que hacer, quién es el responsable, y cómo actuar en caso de problemas (es decir, cómo y cuándo escalar). En este nivel, las organizaciones tienen roles específicos y comités éticos.



7.1.4.1. Principios éticos

Con el fin de guiar a la empresa en su aplicación de la IA y el *big data* en todas las líneas de negocio, Telefónica ha publicado sus «Principios de IA». Los principios incluyen:

- **IA justa.** Se busca asegurarse de que las aplicaciones de tecnología de IA produzcan resultados justos. Esto significa que no deben conducir a impactos injustos discriminatorios por razón de la raza, el origen étnico, la religión, el sexo, la orientación sexual, la discapacidad o cualquier otra condición de las personas. Se aplica la tecnología para reducir al mínimo la probabilidad de que los conjuntos de datos de formación que usemos creen o refuercen un sesgo o una discriminación injusta. Al optimizar un algoritmo de aprendizaje automático para obtener precisión en términos de falsos positivos y negativos, se considera el impacto del algoritmo en el dominio específico.



- **IA transparente y explicable.** Hay que ser explícito con el tipo de datos, personales o no, que utilizan los sistemas de IA, así como la finalidad para la que se van utilizar estos datos. Cuando las personas interactúen directamente con un sistema de IA, seremos transparentes con ello. Cuando los sistemas de IA toman o respaldan decisiones, se toman las medidas técnicas y organizativas necesarias para garantizar un nivel de comprensión adecuado para el área de aplicación. En cualquier caso, si las decisiones afectaran significativamente a la vida de las personas, hay que asegurarse de que entendemos la lógica que hay detrás de las conclusiones. Esto también será de aplicación cuando se use tecnología de terceros.

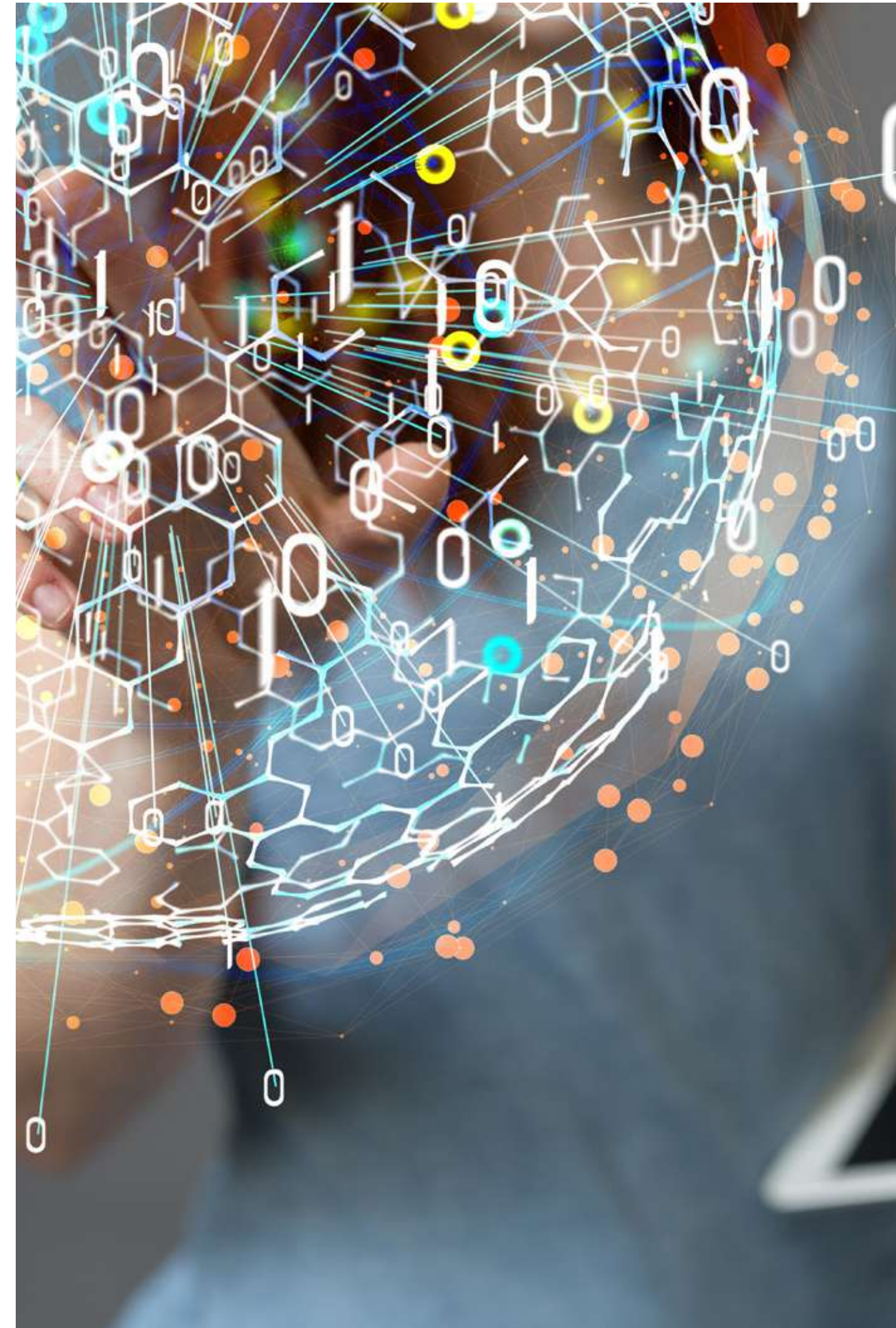
- **IA centrada en las personas.** La IA debería estar al servicio de la sociedad y generar beneficios tangibles para las personas. Telefónica es consciente del hecho de que la implementación de la IA en sus productos y servicios no debe, en ningún caso, provocar un impacto negativo en los derechos humanos o en el logro de los Objetivos de Desarrollo Sostenible de la ONU.
- **Privacidad y seguridad por diseño.** Al crear sistemas de IA que están alimentados por datos, la privacidad y los aspectos de seguridad son una parte inherente del ciclo de vida del sistema. Esto garantiza al máximo el respeto al derecho de las personas a la privacidad y a sus datos personales. Obsérvese que los datos

utilizados en los sistemas de IA pueden ser personales o anónimos/agregados. Obsérvese también que este principio se puede aplicar de manera más general, no solo a los sistemas de inteligencia artificial, y que Telefónica ya cuenta con procesos para garantizar la privacidad y la seguridad adecuadas.

- **Trabajar con socios y terceros.** Telefónica se compromete a verificar la lógica y los datos utilizados por los proveedores para garantizar que se respetan estos principios.

Estos principios también se basan en un amplio consenso entre las comunidades de expertos, así como en aspectos específicos de la industria de las telecomunicaciones.

CASOS DE ÉXITO



7.1.4.2. Principios éticos, metodología y recursos

Los sistemas de IA son el producto de muchas decisiones diferentes tomadas por quienes los desarrollan y despliegan. Desde los objetivos finales del sistema hasta la forma en que las personas interactúan con los sistemas de IA, tenemos que orientar estas decisiones de forma proactiva hacia resultados que sean beneficiosos y equitativos para la sociedad. Esto significa mantener a las personas en el centro de las decisiones de diseño del sistema y respetar valores como la equidad, la confiabilidad y la protección, la privacidad y la seguridad, la inclusión, la transparencia y la responsabilidad.

INTRODUCCIÓN

ÁMBITO DEL VALOR PARA NEGOCIO

ÁMBITO DE LA CIENCIA DE DATOS

ÁMBITO DE LOS DATOS

ÁMBITO DE LA ORGANIZACIÓN

ÁMBITO DE PERSONAS Y CULTURA

ÁMBITO DE LA SOSTENIBILIDAD

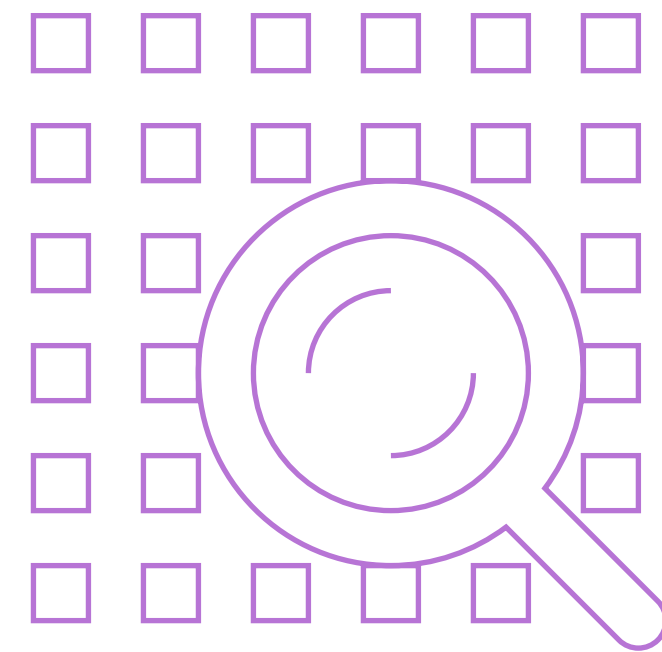
CONCLUSIONES

Microsoft ha hecho público el estándar de **IA Responsable**²² que expone nuestras ideas sobre cómo construir sistemas de IA para defender estos valores y ganarse la confianza de la sociedad. Proporciona una orientación específica y práctica para nuestros equipos que va más allá de los principios de alto nivel que han dominado el panorama de la IA hasta la fecha.

El estándar detalla los objetivos y resultados concretos que los equipos que desarrollan sistemas de IA deben procurar garantizar. Estos objetivos ayudan a desglosar un principio tan amplio como el de la “responsabilidad” en sus elementos clave, como las evaluaciones de impacto, la gobernanza de los datos y la supervisión humana. Cada objetivo se compone de un conjunto de requisitos, que son los pasos que los equipos deben dar para garantizar que los sistemas de IA cumplan los objetivos a lo largo del ciclo de vida del sistema. Por último, el estándar asigna las herramientas y prácticas disponibles a los requisitos específicos para que los equipos de Microsoft que la apliquen dispongan de recursos que les ayuden a tener éxito.

La necesidad de este tipo de orientación práctica es cada vez mayor. La IA forma parte cada vez más de nuestras vidas y, sin embargo, nuestras leyes se están quedando atrás. No se han puesto al día con los riesgos únicos de la IA ni con las necesidades de la sociedad. Aunque vemos signos de que la acción regulatoria en Europa en materia de IA se está acelerando, también reconocemos nuestra responsabilidad de actuar como empresa. Creemos que debemos trabajar para garantizar que los sistemas de IA sean responsables por diseño.

Nuestro estándar de IA responsable refleja cientos de aportaciones de Microsoft, desarrollados por profesionales de Microsoft. Es un paso importante para nuestra práctica de la IA responsable porque es mucho más concreta: establece enfoques prácticos para identificar, medir y mitigar los daños con antelación, e impone requerimientos a los equipos para que adopten controles que aseguren un buen uso y protejan al usuario de un mal uso.



7.2. Reto 12: Algoritmos verdes

7.2.1. Descripción del problema

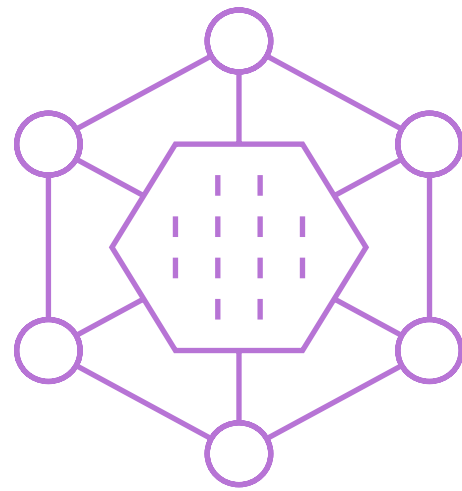
Cómo podemos desarrollar modelos analíticos y algoritmos que nos ayuden a reducir las emisiones?

La Inteligencia Artificial debe estar al servicio del ser humano, y eso implica necesariamente que dicha tecnología no debe ser dañina, y uno de los riesgos existentes es el impacto medioambiental que ésta puede originar a nuestro planeta.

Un análisis de 2018 dirigido por Dario Amodei y Danny Hernández del laboratorio de investigación OpenAI con sede en California, cuya misión es asegurar que la Inteligencia general Artificial beneficie a toda la humanidad, y con los que Microsoft colabora, reveló que la computación utilizada en varios de los modelos de IA se había duplicado cada 3,4 meses desde 2012 — lejos de la Ley de Moore que está en 18 meses — lo que representa un aumento de x300.000. El coste computacional en modelos hiper paramétricos tiene una relación lineal con la cantidad de datos, Epochs y parámetros que están involucrados en el entrenamiento, y es aún mayor en el uso posterior durante la fase de inferencia.

Todo esto supone un reto en un mercado muy dinámico y donde la creciente preocupación por dotar de soluciones sostenibles en el ámbito tecnológico se traslada cada vez con más fuerza al ámbito de la Inteligencia Artificial.

Muchas empresas se han enfocado en el desarrollo de soluciones de sostenibilidad utilizando técnicas de Inteligencia Artificial, denominadas en inglés Green by AI. Y es un área que está proporcionando grandes avances en áreas como la predicción meteorológica o el desarrollo de modelos que mejoren la productividad de la agricultura y su desarrollo ecosostenible. Pero también es importante que esos desarrollos apoyados en Inteligencia Artificial y otros no tan enfocados en resolver problemas de sostenibilidad partan de un desarrollo sostenible por defecto, lo que se ha venido a denominar Green in AI. A continuación, abordaremos este reto, ¿cómo podemos desarrollar soluciones que nos ayuden a reducir las emisiones teniendo en cuenta criterios de sostenibilidad en su desarrollo?



7.2.2. Soluciones

Las emisiones y el coste de la electricidad tienen una alta correlación y causalidad, si ahorramos energía, ahorramos costes a nuestros clientes, pero también emitimos menos carbono y equivalentes a la atmósfera.

Para ello desde nuestro punto de vista se debe seguir una estrategia sustentada en los siguientes pilares:

- En primer lugar, disponer de **herramientas de establecimiento y medición de métricas de costes monetarios, computacionales, energéticos y medioambientales**, y establecer cuadros de mando que permitan identificar áreas de mejora.

- **Costes energéticos, computacionales y monetarios:** soluciones como Energetic Training Cost In AzureML Studio UI²³ para medir el coste energético en el entrenamiento de modelos, estas métricas permiten medir la energía consumida por una GPU en Azure Monitor y calcular posteriormente su coste monetario.
- **Costes medioambientales:** con soluciones como el *Emissions Impact Dashboard*²⁴ que permite comparar el impacto de las infraestructuras on-premise con soluciones en la nube.
- **Impulsar el uso de algoritmos verdes** desde el diseño, capaces de reducir durante la fase de entrenamiento el consumo energético.
- **Implementando soluciones carbón aware**²⁵ que permitan escoger qué funcionalidades estarán disponibles para cada aplicación según la huella de carbono que vaya a provocar.
- **Mejorando las características de autoescalado** apoyándose en la nube, e implementando capacidades para mover cargas de

computación entre regiones o centros de datos que estén alimentados por energías renovables.

- **Potenciando estándares industriales como Green Software Foundation** a nivel software o LEET para las infraestructuras de computación (en general, pero también en el caso particular de infraestructura de entrenamiento e inferencia de sistemas IA).

²³ <https://techcommunity.microsoft.com/t5/green-tech-blog/charting-the-path-towards-sustainable-ai-with-azure-machine/ba-p/2866923>

²⁴ <https://www.microsoft.com/en-us/sustainability/emissions-impact-dashboard>

²⁵ https://docs.microsoft.com/en-us/learn/modules/sustainable-software-engineering-overview/9-demand-shaping?WT.mc_id=green-9537-cxa



- **Establecer sistemas de configuración de la demanda**, sabemos por ejemplo que navegar por redes 4g/5g es entre 4 y 7 veces más contaminante en carbono que hacer a través de redes WiFi, por lo tanto, utilizar sistemas de IA que sean capaces de detectar este comportamiento podría permitirnos que ese sistema IA configurara la demanda y ofreciera una experiencia ECO tal y como hacemos cuando utilizamos nuestros electrodomésticos o nuestros coches.

- **Establecer programas de formación** a los ingenieros de software y científicos de datos para que tomen en consideración los impactos que produce la IA y que bajo nuestro punto de vista deben centrarse en:

- **Carbono:** la eficiencia en el uso del carbono consiste en minimizar la cantidad de carbono emitida por unidad de trabajo.
- **Electricidad:** crear aplicaciones que sean eficientes energéticamente hablando. La electricidad se mide en potencia por unidad de tiempo, menor potencia durante menos tiempo, menor consumo eléctrico. Menor número de ciclos de computación durante menos tiempo, menor consumo energético.

- **Intensidad de carbono:** sería la cantidad de emisiones de carbono por unidad de energía. Hay varios servicios^{26 27 28} que permiten obtener datos en tiempo real con respecto a la intensidad del carbono actual de distintas redes eléctricas. Algunos proporcionan estimaciones futuras de la intensidad del carbono, y otros, la intensidad del carbono marginal.

- **Carbono incorporado:** Al calcular la huella de carbono total de los equipos que ejecutan los sistemas de IA, se ha de tener en cuenta la huella de carbono que emite durante su uso y el carbono incorporado en el equipo. Al alargar la vida útil de los dispositivos reducimos las emisiones por unidad de tiempo. Normalmente el hardware se retira porque se estropea o tiene dificultades para soportar las cargas de trabajo más modernas. El software no puede ayudar con el primer factor, pero, si se centra en crear aplicaciones que se puedan ejecutar en hardware más antiguo, se puede ayudar con el segundo. En el caso de los sistemas de IA, hay un movimiento centrado en datos²⁹ que estudia cómo diseñar sistemas de IA con un menor número de datos (small data) para sistemas industriales y que permite reducir el carbono incorporado además de otros principios de sostenibilidad.

²⁶ <https://www.watstime.org/>

²⁷ <https://api.electricitymap.org/>

²⁸ <https://carbonintensity.org.uk/>

²⁹ From Model-Centric to Data-Centric <https://m.youtube.com/watch?v=06-AZXmwHjo>

- **Proporcionalidad energética:** el uso mide la cantidad de recursos que utiliza un equipo y normalmente se representa en forma de porcentaje. Un equipo inactivo tiene un porcentaje de uso bajo y no se utiliza; un equipo que se ejecute a su máxima capacidad tiene un porcentaje alto y se utiliza de forma completa. La proporcionalidad energética mide la relación entre la potencia consumida en un sistema y la velocidad a la que se realiza el trabajo (su uso). Si el consumo de energía global es proporcional al uso del equipo, se dice que es energéticamente proporcional. La relación entre el consumo de energía y el uso no es lineal. Debido a esta relación, cuanto más se use un equipo, más eficiente será para convertir la electricidad en operaciones computacionales. La ejecución del trabajo en el menor número posible de servidores con la tasa de uso más alta maximiza su eficiencia energética. Aquí también intervienen el uso de energía en un estado estático del equipo o los ciclos de reloj.
- **Eficiencia de la red:** una red es una serie de conmutadores, enrutadores y servidores. Todos los equipos de una red consumen electricidad y tienen carbono incorporado. La cantidad de datos emitida por las redes depende del tamaño de los datos, la distancia que recorren, la eficiencia de

los dispositivos, la intensidad de carbono de la región de cada dispositivo y el protocolo de red entre otros. Así que en un sistema de AI donde se utilicen todos estos sistemas es imprescindible que se tome en consideración que según un estudio de The Shift Project de 2019 cada GB transmitido emite 1,22kg de carbono y se estima en el mismo estudio que en redes móviles se duplica esa cifra.

- **Modelado de la demanda:** es una estrategia similar al desplazamiento de la demanda, pero, en lugar de trasladar la demanda a otra región u hora diferente, la demanda se modela para que coincida con el suministro existente. Un buen ejemplo de ello es el software de videoconferencia. En lugar de transmitir contenido a la máxima calidad posible en todo momento, a menudo la demanda se moldea mediante la reducción de la calidad del vídeo para priorizar el audio. En el caso de los sistemas de IA, el modelado de la demanda podrá (previa notificación al usuario) reducir sus funcionalidades para que el consumo energético sea menor.

- **Optimización:** La clave del éxito de la optimización consiste en elegir un criterio de medición que proporcione señales claras sobre dónde concretar mejor los esfuerzos de la optimización. Por ejemplo, ¿merece la pena dedicar dos semanas a reducir los megabytes de la comunicación de la red si las consultas de la base de datos emiten 10 veces más carbono? Se ha de valorar el esfuerzo necesario para la descarbonización frente a las posibles recompensas que esta tiene. Como sucede con la sostenibilidad desde un punto de vista más amplio, algunos sectores serán más difíciles de descarbonizar que otros.



7.2.3. Niveles de madurez

7.2.3.1. Nivel 1. Fundacional

Nociones bajas sobre cómo habilitar mecanismos de medida del impacto medioambiental en el uso de algoritmos en los sistemas de inteligencia artificial. Sistemas estáticos de medición y escasos conocimientos en el ámbito del uso de la tecnología para diseñar sistemas de IA eficientes desde por defecto.

7.2.3.2. Nivel 2. Acercamiento

La necesidad de tener un enfoque sostenible en el desarrollo de sistemas de inteligencia artificial se vuelve clara. Los esfuerzos iniciales para establecer principios que impulsen cómo se construye una IA sostenible comienzan a tener importancia en el equipo directivo.

Las organizaciones en esta etapa comienzan a definir/implementar prácticas, generalmente en forma de pautas, que brindan un contexto adecuado para la implementación.

7.2.3.3. Nivel 3. Aspiracional

Conciencia del impacto medioambiental de la IA no solo a nivel organizacional o empresarial, sino también a nivel de personas y sociedad. Por lo tanto, queda más claro como diseñar un sistema de IA para que cumpla los principios definidos. Más allá de los principios y prácticas, las organizaciones en esta etapa se dan cuenta de la necesidad de tener también herramientas (a nivel técnico) que puedan ayudarlas a implementar, comprender/explicar, proteger y controlar cómo de eficiente y sostenible son los sistemas de IA.

7.2.3.4. Nivel 4. Maduro

Se establece una cultura de IA sostenible en todos los ámbitos de la empresa, ingeniería, ventas y atención al cliente y se infunde un enfoque sostenible en el proceso de creación de sistemas de IA. La pregunta “¿Cómo de sostenible son mis sistemas de IA?” está siempre presente. Alta conciencia del impacto de la IA en el medioambiente y la sociedad en general.



7.2.4.1. Algoritmos verdes

En febrero de 2020 Microsoft liberó una nueva librería denominada DeepSpeed³⁰ + ZeRO con el objetivo de mejorar la escala, la velocidad, el coste y la facilidad de uso para entrenar modelos hiper paramétricos de más de 100 millones de parámetros. Como ya se sabe, apenas el 12% de los modelos que se entrenan llegan a estar en producción, lo que los llevó a centrarse en los costes asociados al entrenamiento de modelos de IA. Gracias a la librería DeepSpeed + ZeRO se pueden entrenar modelos 10 veces más grandes, 5 veces más rápido con un coste 5 veces inferior en términos de energía y por tanto de emisiones con cambios mínimos en el código. Además, Microsoft ha introducido la capacidad de identificar los niveles de precisión durante el entrenamiento de los modelos de IA para evitar el consumo de energía y por tanto emisiones de CO₂ en aquellos casos cuya precisión sea inferior a umbrales establecidos.

La inferencia en el edge también es importante, el uso que hacemos de los modelos ya entrenados en el edge es otra área donde Microsoft ha puesto mucho esfuerzo. Se calcula que el 80%-90% de los costes y/o emisiones de carbono están relacionados con la fase de inferencia. Para solucionar este problema Microsoft se centró en investigar en esta área y publicó en marzo 2021 un nuevo paper, AsyMo inferencia escalable y eficiente en CPUs³¹ móviles asimétricos para aprendizaje profundo con resultados que muestran un 46% de mejora del rendimiento con un 37% de mejora de la eficiencia energética para modelos convolucionales en comparación con modelos optimizados en Tensor Flow.

Se ha avanzado en la mejora de los COGS por utilización de GPU, llegando a acuerdos con sus proveedores hardware lo cual les permite optimizar el consumo energético y por lo tanto reducir las emisiones mediante la gestión de la energía estática y dinámica.

³⁰ <https://www.microsoft.com/en-us/research/blog/zero-deepspeed-new-system-optimizations-enable-training-models-with-over-100-billion-parameters/>

³¹ <https://www.microsoft.com/en-us/research/publication/asymo-scalable-and-efficient-deep-learning-inference-on-asymmetric-mobile-cpus/>

7.2.4

CASOS
DE ÉXITO

7.2.4.2. Smart Energy

En el área industrial se está constantemente buscando e invirtiendo en soluciones digitales para reducir el consumo energético y las emisiones de CO₂ en los Centros Industriales de Repsol, y para esto, se están valiendo de nuevas tecnologías basadas en modelos de simulación y Machine Learning que permiten mejorar constantemente las herramientas, mientras se basan en datos con mejor calidad y más fáciles de comprender para la toma de decisiones.

Actualmente el objetivo en su plan de descarbonización es reducir, en el 2025, un aproximado de 55.000 toneladas de CO₂ equivalentes, para lo cual se trabaja en proyectos que le acercan cada vez más a esa meta. Estos son algunos de los proyectos:

- **HEAT-X.** Implementar una rigurosa monitorización en los trenes de precalentamiento de crudo que permita la toma anticipada de decisiones para reducir al máximo su ensuciamiento. Esta monitorización evalúa el desempeño térmico e hidráulico de cada cambiador, y permite identificar el momento idóneo para realizar la limpieza de los intercambiadores. También predice la temperatura de entrada al horno, con lo que podemos modificar la mezcla de los crudos para maximizar la recuperación de calor, evitar la pérdida de la producción, y con esto, reducir las emisiones de CO₂.
- **LEO.** Con esta iniciativa se busca optimizar energéticamente la operación de una unidad con el objetivo de reducir emisiones. Teniendo en cuenta las variables que afectan al consumo energético, buscando una simbiosis entre el Machine Learning y el simulador de Primeros Principios que, funcionando en conjunto, aportan mayor volumen y calidad de datos, permitiendo así ejecuciones más ágiles y escenarios cada vez más precisos.
- **MATRIX.** Optimización de las redes de utilities (Hidrógeno, Fuel Gas y Vapor) con el objetivo de reducir consumos con la consecuente disminución en la huella de carbono. Las redes de utilities están interrelacionadas, por lo que casi cualquier movimiento en una de ellas afecta a las otras. Esto lleva a tener que hacer ajustes casi continuos en cada una de ellas para estar siempre en el punto óptimo. Mediante el uso de simuladores y algoritmos de optimización en tiempo real hemos desarrollado un optimizador predictivo que permite tomar decisiones comunes en las tres redes al mismo tiempo, manteniendo así el óptimo en cada momento.



AIRBUS

7.2.4.3. Sustainable IT / Zero Emissions Next Generation Aircraft

El objetivo de Airbus es reducir, de cara a 2030, en un 63% sus emisiones industriales Scope 1 (emisiones directas de nuestros sites) y Scope 2 (emisiones indirectas de servicios que no se producen dentro de nuestros sites pero que Airbus adquiere para el correcto funcionamiento de los mismos como por ejemplo electricidad, sistemas de aire acondicionado, calefacción, etc.). Este nuevo objetivo, que se convierte en todavía más ambicioso con respecto al 40% anunciado en 2019, ha sido definido de acuerdo con el SBTi - Science Based Target Initiative - y está en línea con "1.5°C pathway" y "net-zero future".

El objetivo global tiene en cuenta cinco aspectos materiales y define la reducción para cada uno de ellos partiendo de un baseline de los datos recogidos en 2015:

- Energía y CO₂

INTRODUCCIÓN

ÁMBITO DEL VALOR PARA NEGOCIO

ÁMBITO DE LA CIENCIA DE DATOS

ÁMBITO DE LOS DATOS

ÁMBITO DE LA ORGANIZACIÓN

ÁMBITO DE PERSONAS Y CULTURA

ÁMBITO DE LA SOSTENIBILIDAD

CONCLUSIONES

- Residuos: reducción del 20% en la cantidad de residuos producidos en las plantas de Airbus
- Emisiones: 0% de incremento en emisiones de componentes volátiles orgánicos (VOC)
- Agua: 50% de reducción en consumo de agua potable

Estos son algunos ejemplos de proyectos que a través de soluciones de analítica avanzada contribuirán a conseguir estos objetivos:

- **BeSmart! Electricity & water saving.** El objetivo del proyecto es la identificación de anomalías en el consumo de agua y electricidad en los sites de Airbus para la emisión de alertas. La detección de anomalías en cuanto al consumo de agua se implementa con un modelo "tree-based" o árbol de decisión. En cuanto a la detección de anomalías para los componentes eléctricos se emplea un Autoencoder desarrollado con Pytorch. Las predicciones y alertas de los modelos son expuestos y pueden ser consultados a través de una API REST.
- **Transport route optimization.** Con el objetivo de mejorar la eficiencia en nuestra cadena de suministro y reducir las emisiones de CO2 este proyecto se encarga de optimizar las rutas de nuestros proveedores y sistemas de logística interna. A través de métodos metaheurísticos como Tabu Search y Evolution Strategies (ESs) la iniciativa es capaz de crear una serie de Dashboards que muestran indicadores sobre las rutas menos eficientes y se ofrecen opciones para su optimización.
- **Frugal.** En Airbus somos conscientes de la energía que se requiere para entrenar nuestros modelos de Machine Learning y por ello estamos comprometidos con desarrollar los algoritmos de un modo más eficiente. Una de las principales iniciativas que tenemos en nuestro roadmap es IAFRUGAL: conseguir implementar

y ejecutar nuestros algoritmos con menos datos y menos recursos de computación. La idea es intentar usar los modelos más simples con el menor número de parámetros posible pero obviamente intentando conseguir el nivel de performance requerido por el modelo. Los modelos más simples son más sencillos de explicar y también más sencillos de embeber en nuestros sistemas de vuelo. En relación a la reducción del volumen de datos estamos también investigando modelos innovadores que requerirán menos datos y explorando cómo se pueden combinar sistemas de inteligencia artificial data-driven con Symbolic AI y/o modelados físicos.

- **Digital twin of the globe.** Los satélites de alta resolución de Airbus que orbitan nuestro planeta nos permiten recoger una gran cantidad y variedad de imágenes. Con estas imágenes y a través del entrenamiento y ejecución de algoritmos de inteligencia artificial podemos detectar un crecimiento excesivo y desproporcionado de algunos cultivos como el del aceite de palma. En áreas con amenaza de deforestación esto puede indicar que la vegetación nativa está siendo sustituida por aquella que permite la producción de este tipo de aceites. Así mismo los datos e imágenes que capturamos desde el espacio nos permiten detectar áreas de contaminación y presentan un gran potencial a la hora de aconsejar a los sectores agrícolas

INDESIA

Framework Data & IA
Industrial Marco de
referencia de datos e
inteligencia artificial para el
sector industrial español

INTRODUCCIÓN

ÁMBITO DEL VALOR PARA NEGOCIO

ÁMBITO DE LA CIENCIA DE DATOS

ÁMBITO DE LOS DATOS

ÁMBITO DE LA ORGANIZACIÓN

ÁMBITO DE PERSONAS Y CULTURA

ÁMBITO DE LA SOSTENIBILIDAD

CONCLUSIONES

sobre dónde y cómo establecer sus plantaciones con la correlación de datos recibidos desde el satélite e información del suelo. Desde Airbus creemos firmemente que la inteligencia artificial ofrece un gran valor añadido a la hora de establecer un modus operandi de colaboración entre la propia inteligencia artificial y los humanos, la IA no es más inteligente que los propios humanos pero ofrece la capacidad de extraer información sobre grandes áreas geográficas en menos tiempo (algo que no era posible anteriormente) y representar esta información permitiéndonos realizar análisis y extraer conclusiones.

- **Trajectory optimisation in air traffic management powered by AI.** Los primeros estudios realizados por Airbus y su aplicación en proyectos reales nos han demostrado que podemos aplicar la inteligencia artificial para optimizar los planes de vuelo combinando datos de tráfico aéreo, información meteorológica y modelos de performance tanto de nuestros aviones como de nuestros helicópteros. Mejorando también

la sincronización de los despegues y aterrizajes de nuestras aeronaves comerciales a través de algoritmos de ML podemos incrementar su eficiencia y reducir por lo tanto el consumo de fuel.

- **ZEROe aircraft.** Más allá del corto plazo Airbus está explorando nuevas vías en las que la inteligencia artificial tendrá un papel clave para conseguir un mundo más sostenible. Mientras nuestros ingenieros trabajan para que los aviones de cero emisiones se conviertan en una realidad, la tecnología está siendo ya un factor determinante para entender cómo las celdas de combustible en estos aviones del futuro se comportan a través de modelos virtuales de estrés que nos permiten proyectar los sistemas reales en condiciones extremas. Estas tecnologías digitales son particularmente útiles en entornos de ingeniería tan complejos ya que permiten reducir los riesgos para producir de forma más rápida, barata y por supuesto más segura.

08 CONCLUSIONES



La Inteligencia Artificial es una de las tecnologías más disruptivas y que más impacto van a tener en los próximos años, no solo en los negocios, si no en el día a día de nuestras vidas y nuestra sociedad.

Los grandes players digitales, como Amazon, Facebook, Google, Netflix etc. han sido los primeros en demostrar cómo estas tecnologías van a ser estratégicas para el desarrollo futuro casi de cualquier industria.

Las empresas líderes de otros sectores como la banca, los medios o el retail, ya empezaron hace tiempo a analizar e implementar estrategias para obtener el máximo valor de estas tecnologías.

Además, todos los países más avanzados han comenzado también a desarrollar sus propias estrategias nacionales de Inteligencia Artificial para promover e impulsar el uso de estas tecnologías.

La Inteligencia Artificial junto con otras tecnologías habilitadoras como el Internet de las Cosas, la conectividad 5G, la computación en el Edge o los gemelos digitales son la clave para el desarrollo de un nuevo paradigma en el mundo industrial que conocemos como Industria conectada o Industria 5.0.

INDESIA

Framework Data & IA
Industrial Marco de
referencia de datos e
inteligencia artificial para el
sector industrial español

INTRODUCCIÓN

ÁMBITO DEL VALOR PARA NEGOCIO

ÁMBITO DE LA CIENCIA DE DATOS

ÁMBITO DE LOS DATOS

ÁMBITO DE LA ORGANIZACIÓN

ÁMBITO DE PERSONAS Y CULTURA

ÁMBITO DE LA SOSTENIBILIDAD

CONCLUSIONES

La aplicación de la IA a casos de uso industriales cambia totalmente el enfoque de los posibles retornos o resultados a la hora de usar las nuevas plataformas. Esto se debe a que los operarios de las industrias pueden obtener outputs donde el proceso o autómatas hace una predicción sobre un KPI clave industrial o una recomendación en cuanto a ciertos ajustes del proceso.

Precisamente los entornos industriales, por la gran cantidad de datos que se pueden recoger en sus procesos, las grandes inversiones que se hacen en sus activos, y la importancia como clave competitiva de la eficiencia y la automatización de sus procesos, facilitan la captura de valor de la Inteligencia Artificial en todo su potencial. Cada proceso industrial de nuestras cadenas de valor es susceptible de ser optimizado en base a datos, analítica e inteligencia artificial.

La automatización de las plantas, el mantenimiento predictivo, la maximización de la calidad, la optimización de la logística y la cadena de suministro, el desarrollo de gemelos digitales o la eficiencia energética son solo algunos ejemplos de los múltiples casos de uso que aplican la cadena de valor de cualquiera de nuestras industrias.

Además, la inteligencia artificial va a contribuir no solo a mejorar la competitividad, sino que también va a tener un alto impacto en la sostenibilidad de la industria, permitiendo transicionar hacia

una industria baja en carbono maximizando la eficiencia energética, apoyando el desarrollo de la economía circular o mediante el desarrollo de nuevos materiales más amigables con el medio ambiente entre otros.

La Inteligencia Artificial supone por tanto una gran oportunidad para el tejido industrial español de mejorar nuestra competitividad y sostenibilidad liderando el uso e implementación de estas tecnologías.

Sin embargo, escalar estas tecnologías no es sencillo y las empresas deben enfrentarse a una serie de retos que deben superar y que detallamos en este trabajo dando algunas claves de cómo poder superar estos retos paso a paso:

- **El primer paso que debe afrontar cualquier empresa es la identificación de cómo la IA puede aportar valor.** Parece importante que las empresas desarrollen workshops y benchmarks hasta conformar una agenda de los principales casos de uso que podrían implementar. También es relevante que identifiquen cómo van a medir y cuantificar el impacto que tendrán estas iniciativas para evitar el riesgo de poner foco en iniciativas muy llamativas desde el punto de vista de comunicación, pero con poco impacto.



- **El segundo paso es entender la dificultad o complejidad analítica que requiere cada caso** con el objetivo de ir creciendo paulatinamente de los más sencillos, aquellos basados en analítica descriptiva, a los más complejos, que requieren técnicas de Inteligencia Artificial más avanzadas (Deep learning, reinforcement learning etc.). Es necesario también reflexionar sobre cómo se van a operar en el futuro esos modelos ya que van a requerir un reentrenamiento constante y que herramientas se va a implementar para su operación (ML Ops). La clave es tratar de evitar desarrollar modelos que nunca llegan a ponerse en producción o que al poco tiempo

INTRODUCCIÓN

ÁMBITO DEL VALOR PARA NEGOCIO

ÁMBITO DE LA CIENCIA DE DATOS

ÁMBITO DE LOS DATOS

ÁMBITO DE LA ORGANIZACIÓN

ÁMBITO DE PERSONAS Y CULTURA

ÁMBITO DE LA SOSTENIBILIDAD

CONCLUSIONES

dejan de usarse por no haber nadie en la organización que entienda cómo funcionan.

- **El tercer paso es ser capaz de acceder a los datos para poner en marcha esos casos de uso** y garantizar que los datos tienen la calidad necesaria. Las empresas deben definir una estrategia de datos que incluya el desarrollo de una plataforma de datos, un programa de gobierno de datos y deben tener identificado los datos del ecosistema que podrían interesarle tener acceso para participar en los nuevos espacios de datos compartidos que van a surgir en Europa.
- **El cuarto paso es definir la estrategia para acceder a los recursos que van a permitir desarrollar este tipo de iniciativas.** Es necesario decidir qué presupuesto se va a asignar a este tipo de iniciativas, si se va a hacer caso a caso individualmente en cada negocio o si va a haber un presupuesto centralizado, y qué retorno se le va a exigir a este tipo de iniciativas. También es clave definir la estrategia de sourcing para el desarrollo de estas iniciativas, decidir si se van a afrontar completamente con personas externas a la organización o se van a necesitar recursos propios, entender las dificultades de captar y retener talento analítico actualmente en el mercado y sobre todo definir una estrategia de formación y capacitación de personal propio para garantizar las habilidades necesarias en el futuro.

- **El quinto paso es afrontar la gestión del cambio que va a ser necesaria** para que los modelos lleguen a implementarse y esta cultura data driven se adopte en la organización, el desarrollo de una comunidad de datos es clave para que la organización adopte una cultura de toma de decisiones en base a los datos y captura de todo su valor.
- **El último paso es garantizar que todo lo anterior se hace de una manera ética y sostenible,** para lo que será necesario definir unos principios éticos de actuación, siendo clave en el ámbito industrial principios como la seguridad y la robustez de las soluciones y también el uso de algoritmos verdes que garanticen que el desarrollo de modelos de Inteligencia artificial causa el menor impacto en el medio ambiente.

Los datos y la Inteligencia Artificial son una gran oportunidad y a la vez un gran reto para las empresas del ámbito industrial. La industria española tiene que aunar esfuerzos y para impulsar la aceleración de los procesos de digitalización de las empresas de nuestra cadena de valor ya que solo de este modo conseguiremos transformar el sector industrial, posicionar a España como referente en el uso de datos e inteligencia artificial y desarrollar una nueva economía del dato que genere crecimiento económico y fomente el empleo.