

**RA5. Evalúa la importancia de los datos, así como su protección en una economía digital globalizada, definiendo sistemas de seguridad y ciberseguridad tanto a nivel de equipo/sistema como globales.**

#### **a) Establece la diferencia entre dato e información.**

En el contexto de la digitalización aplicada al sector productivo, distinguir con precisión entre **dato** e **información** es fundamental, ya que de esta diferencia depende la correcta gestión del conocimiento dentro de la organización.

El **dato** es la unidad básica. Se trata de un valor bruto, aislado y sin interpretación por sí mismo. Puede ser numérico, textual, temporal o incluso una señal recogida automáticamente por un sistema. Por ejemplo, una lectura de temperatura de una máquina, un identificador de cliente, una hora de registro o el número de unidades producidas en un turno concreto. El dato, en este estado, no permite tomar decisiones, ya que carece de contexto y significado operativo.

La **información**, en cambio, surge cuando esos datos se procesan, se organizan y se relacionan con otros datos dentro de un contexto determinado. Al añadir significado, los datos se transforman en información útil. Siguiendo el ejemplo anterior, conocer que la temperatura media de una máquina supera el umbral recomendado durante determinados turnos y que eso coincide con un aumento de fallos ya no es un simple conjunto de datos, sino información relevante para mantenimiento y gestión de la producción.

Desde el punto de vista del sector productivo, esta transformación es crítica. Las empresas no generan valor por almacenar grandes cantidades de datos, sino por convertirlos en información que apoye la toma de decisiones estratégicas, tácticas y operativas. Además, esta distinción también tiene implicaciones en seguridad y regulación: mientras los datos pueden parecer inofensivos de forma aislada, al convertirse en información pueden revelar patrones sensibles, por lo que requieren mayores medidas de protección.

#### **b) Describe el ciclo de vida del dato.**

El **ciclo de vida del dato** describe el conjunto de etapas por las que pasa un dato desde que se origina hasta que se elimina definitivamente. Comprender este ciclo es esencial en la digitalización del sector productivo, ya que permite gestionar los datos de forma eficiente, segura y conforme a la normativa vigente.

#### **Listado de etapas principales:**

1. Generación o captura
2. Almacenamiento
3. Procesamiento y preparación
4. Análisis y uso
5. Compartición y distribución

## 6. Archivado y eliminación

### 1. Generación o captura

En esta fase el dato se origina o se recoge a partir de distintas fuentes. En el sector productivo es habitual que los datos procedan de sensores industriales (IoT), sistemas de control de producción, aplicaciones empresariales (ERP, CRM), registros de actividad o interacciones de clientes. La calidad del dato empieza aquí: errores en la captura, mediciones defectuosas o datos incompletos afectarán a todo el ciclo posterior.

### 2. Almacenamiento

Una vez capturados, los datos deben almacenarse de forma estructurada y segura. Esto puede realizarse en bases de datos locales, servidores corporativos o infraestructuras en la nube. En esta etapa se definen aspectos críticos como el formato de almacenamiento, la redundancia, las copias de seguridad y los controles de acceso. Un almacenamiento adecuado garantiza que los datos estén disponibles cuando se necesiten y protegidos frente a pérdidas o accesos no autorizados.

### 3. Procesamiento y preparación

En esta fase los datos brutos se limpian, transforman y normalizan. Se eliminan duplicados, se corrigen errores y se integran datos procedentes de distintas fuentes. En el entorno productivo, este paso es clave para asegurar que los análisis posteriores reflejen la realidad del proceso industrial y no se basen en información sesgada o incorrecta.

### 4. Análisis y uso

Aquí los datos preparados se utilizan para generar información y conocimiento. Se aplican técnicas de análisis descriptivo, predictivo o prescriptivo, así como modelos de machine learning. El objetivo es apoyar decisiones como la optimización de procesos, la planificación de la producción, el mantenimiento predictivo o la mejora de la calidad del producto.

### 5. Compartición y distribución

Los datos o la información resultante pueden compartirse entre departamentos, con socios externos o con clientes. Esta etapa requiere políticas claras de seguridad y gobernanza, ya que implica controlar quién accede a qué datos y con qué finalidad. En una economía digital globalizada, la compartición debe cumplir además con normativas de protección de datos y acuerdos contractuales.

### 6. Archivado y eliminación

Finalmente, los datos que ya no son necesarios para la operativa diaria pueden archivar para fines legales, históricos o de auditoría, o bien eliminarse de forma segura. La eliminación controlada es especialmente importante para cumplir la normativa de protección de datos y reducir riesgos de seguridad derivados del almacenamiento innecesario de información sensible.

### c) **Identifica la relación entre Big Data, análisis de datos, machine/deep-learning e inteligencia artificial.**

En la digitalización aplicada al sector productivo, estos conceptos están estrechamente relacionados y forman un **ecosistema tecnológico** orientado a extraer valor de los datos y automatizar la toma de decisiones. No son equivalentes, sino complementarios, y se entienden mejor como niveles conectados entre sí.

#### **Relación general entre los conceptos:**

- Big Data actúa como base de datos masivos.
- El análisis de datos transforma esos datos en conocimiento.
- El machine learning permite aprender automáticamente de los datos.
- El deep learning amplía esa capacidad mediante modelos más complejos.
- La inteligencia artificial integra todo lo anterior para simular comportamientos inteligentes.

#### **Big Data como punto de partida**

Big Data se refiere a la gestión de grandes volúmenes de datos, generados de forma continua y procedentes de múltiples fuentes (sensores, máquinas, usuarios, sistemas empresariales). En el sector productivo, sin Big Data no sería posible almacenar y manejar la enorme cantidad de datos generados por procesos industriales digitalizados. Big Data no “piensa” ni analiza por sí mismo, pero proporciona la materia prima necesaria.

#### **Análisis de datos como transformación del dato en conocimiento**

El análisis de datos utiliza técnicas estadísticas y matemáticas para explorar, resumir e interpretar los datos almacenados en sistemas Big Data. Permite responder preguntas como qué está ocurriendo, por qué ocurre y qué tendencias se observan. En una empresa productiva, este análisis ayuda a detectar ineficiencias, patrones de fallos o comportamientos anómalos en la producción.

#### **Machine learning como automatización del aprendizaje**

El machine learning va un paso más allá del análisis tradicional. En lugar de programar reglas explícitas, los algoritmos aprenden directamente de los datos históricos. En el sector productivo, esto permite crear modelos que predicen fallos de maquinaria, ajustan parámetros de producción o clasifican productos defectuosos de forma automática. Su eficacia depende directamente de la cantidad y calidad de datos disponibles, normalmente proporcionados por entornos Big Data.

#### **Deep learning como especialización del machine learning**

El deep learning es una rama del machine learning basada en redes neuronales profundas, capaces de manejar datos muy complejos como imágenes, sonidos o secuencias temporales extensas. En entornos productivos se utiliza, por ejemplo, para visión artificial en control de calidad, análisis de imágenes térmicas o reconocimiento avanzado de patrones. Requiere

grandes volúmenes de datos y alta capacidad de cálculo, lo que lo vincula directamente con Big Data y la computación en la nube.

### **Inteligencia artificial como marco global**

La inteligencia artificial engloba todas estas técnicas cuando se aplican para crear sistemas capaces de tomar decisiones, adaptarse y actuar de forma autónoma. En el sector productivo, la IA integra Big Data, análisis de datos y modelos de aprendizaje automático para optimizar procesos completos, desde la planificación de la producción hasta la logística o la atención al cliente.

En conjunto, la relación entre estos conceptos permite pasar de la simple acumulación de datos a sistemas productivos inteligentes, capaces de mejorar continuamente su rendimiento.

#### **d) Describe las características que definen Big Data.**

Big Data se define por un conjunto de características que reflejan los retos y oportunidades asociados al tratamiento masivo de datos en la economía digital. Estas características explican por qué los sistemas tradicionales de gestión de datos resultan insuficientes y por qué el sector productivo necesita nuevas arquitecturas tecnológicas para extraer valor de la información.

En primer lugar, Big Data se caracteriza por el enorme volumen de datos que se generan y almacenan de forma continua. En entornos productivos digitalizados, cada máquina, sensor o sistema de gestión aporta datos de manera constante, lo que provoca un crecimiento exponencial de la información disponible. Este volumen supera con facilidad la capacidad de los sistemas convencionales y obliga a utilizar soluciones distribuidas y escalables.

Otra característica esencial es la velocidad. Los datos no solo son muchos, sino que se generan y deben procesarse a gran rapidez. En numerosos procesos industriales, la utilidad del dato depende del tiempo: detectar un fallo, ajustar un parámetro o reaccionar ante una anomalía requiere análisis casi en tiempo real. Big Data permite manejar flujos continuos de datos y obtener respuestas con baja latencia.

La variedad también define Big Data. Los datos ya no son exclusivamente numéricos o estructurados, sino que incluyen textos, imágenes, vídeos, señales de sensores, registros de actividad y datos procedentes de múltiples sistemas. Esta diversidad aumenta la complejidad del análisis, pero al mismo tiempo enriquece la visión global del proceso productivo.

A estas características se suma la veracidad, entendida como la fiabilidad y calidad de los datos. En Big Data no todos los datos son precisos o completos, por lo que es necesario evaluar su consistencia y reducir el ruido. En el sector productivo, decisiones basadas en datos poco fiables pueden provocar errores operativos o pérdidas económicas.

Finalmente, Big Data se justifica por el valor que puede extraerse de los datos. El objetivo último no es almacenar información masiva, sino transformarla en conocimiento útil que mejore la eficiencia, reduzca riesgos o genere nuevas oportunidades de negocio. Sin esta capacidad de

generar valor, Big Data perdería su sentido estratégico dentro de la digitalización del sector productivo.

### **e) Describe las etapas típicas de la ciencia de datos y su relación con el proceso.**

La ciencia de datos se estructura como un proceso continuo y cíclico cuyo objetivo es transformar datos en conocimiento accionable para la empresa. En el sector productivo, estas etapas no son independientes, sino que están estrechamente conectadas y se retroalimentan en función de los resultados obtenidos.

#### **Etapas principales del proceso de ciencia de datos:**

1. Comprensión del problema y del contexto empresarial
2. Obtención y comprensión de los datos
3. Preparación y limpieza de los datos
4. Análisis exploratorio
5. Modelado y construcción de modelos
6. Evaluación de resultados
7. Despliegue, comunicación y mejora continua

#### **1. Comprensión del problema y del contexto empresarial**

El proceso comienza con la definición clara del problema que se quiere resolver. En una empresa productiva, esto implica entender los objetivos del negocio, los procesos industriales y las decisiones que se desean apoyar con datos. Sin esta fase, el análisis puede generar resultados técnicamente correctos pero irrelevantes desde el punto de vista operativo.

#### **2. Obtención y comprensión de los datos**

En esta etapa se identifican las fuentes de datos disponibles y se recopila la información necesaria. También se analiza su estructura, volumen y calidad inicial. En entornos productivos, los datos pueden proceder de sensores, sistemas de producción, bases de datos corporativas o proveedores externos, lo que exige una visión global del ecosistema de datos.

#### **3. Preparación y limpieza de los datos**

Esta es una de las fases más críticas y costosas del proceso. Los datos se limpian, se corrigen errores, se eliminan valores inconsistentes y se integran distintas fuentes. En el sector productivo, una preparación deficiente puede provocar modelos inexactos que no reflejen el comportamiento real de los procesos industriales.

#### **4. Análisis exploratorio**

El análisis exploratorio permite descubrir patrones, tendencias y relaciones entre variables. Se utilizan técnicas estadísticas y visualizaciones para comprender el comportamiento de los datos antes de aplicar modelos complejos. Esta fase ayuda a validar supuestos y a orientar el tipo de modelos que se utilizarán posteriormente.

## 5. Modelado y construcción de modelos

Aquí se aplican algoritmos estadísticos, de machine learning o deep learning para resolver el problema planteado. En la empresa productiva, estos modelos pueden servir para predicción de fallos, optimización de recursos o control de calidad. La elección del modelo depende tanto del tipo de datos como del objetivo empresarial.

## 6. Evaluación de resultados

Los modelos se evalúan para comprobar su precisión, fiabilidad y utilidad real. Se comparan resultados, se validan con datos nuevos y se analiza su impacto en la toma de decisiones. En el sector productivo, esta evaluación es clave para evitar decisiones erróneas basadas en modelos poco robustos.

## 7. Despliegue, comunicación y mejora continua

Finalmente, los resultados se integran en los sistemas de la empresa y se comunican a los responsables de la toma de decisiones. El proceso no termina aquí: los modelos deben actualizarse y mejorarse a medida que cambian los datos y el entorno productivo, cerrando así un ciclo de mejora continua.

### **f) Describe los procedimientos de almacenamiento de datos en la cloud/nube.**

El almacenamiento de datos en la nube se basa en una serie de procedimientos técnicos y organizativos que permiten gestionar grandes volúmenes de información de forma flexible, escalable y segura. En el sector productivo digitalizado, estos procedimientos son esenciales para garantizar la disponibilidad de los datos y su correcta explotación a lo largo de todo su ciclo de vida.

En primer lugar, los datos se almacenan en infraestructuras distribuidas, lo que significa que no residen en un único servidor físico, sino que se reparten entre distintos centros de datos. Este enfoque mejora la tolerancia a fallos y asegura la continuidad del servicio incluso ante incidencias técnicas. Para las empresas productivas, esto se traduce en una mayor fiabilidad de los sistemas de información críticos.

Otro procedimiento clave es el uso de distintos modelos de almacenamiento según el tipo de dato. La nube permite combinar bases de datos relacionales para información estructurada, bases de datos NoSQL para datos semiestructurados y sistemas de almacenamiento de objetos para grandes volúmenes de datos no estructurados. Esta flexibilidad facilita adaptar el almacenamiento a las necesidades reales de cada proceso productivo.

La escalabilidad automática es otro elemento fundamental. En la nube, la capacidad de almacenamiento puede aumentar o disminuir según la demanda, sin necesidad de realizar inversiones previas en infraestructura física. Esto resulta especialmente importante en entornos productivos con picos de generación de datos, como campañas de producción intensiva o sistemas de monitorización continua.

Desde el punto de vista de la seguridad, los procedimientos de almacenamiento en la nube incluyen el cifrado de los datos tanto en reposo como durante su transmisión. Además, se aplican sistemas de control de acceso basados en identidades y roles, que determinan qué usuarios o aplicaciones pueden acceder a cada conjunto de datos. Estas medidas reducen el riesgo de accesos no autorizados y fugas de información.

Finalmente, el almacenamiento en la nube incorpora mecanismos de copia de seguridad, versionado y recuperación ante desastres. Los datos se replican de forma automática y se mantienen copias en distintas ubicaciones geográficas. En el sector productivo, esto garantiza que la información crítica pueda recuperarse rápidamente y que los procesos no se vean interrumpidos por pérdidas de datos o fallos graves del sistema.

### **g) Describe la importancia del cloud computing.**

El cloud computing es un elemento central en la digitalización del sector productivo, ya que actúa como infraestructura habilitadora para la gestión de datos, el análisis avanzado y la automatización de procesos. Su importancia no se limita al ámbito tecnológico, sino que tiene un impacto directo en la organización, la estrategia y la competitividad de las empresas.

En primer lugar, el cloud computing permite a las empresas acceder a recursos tecnológicos avanzados sin necesidad de realizar grandes inversiones iniciales en infraestructura propia. Esto reduce las barreras de entrada a la digitalización y facilita que empresas de cualquier tamaño puedan utilizar sistemas de almacenamiento, análisis de datos o inteligencia artificial que antes solo estaban al alcance de grandes corporaciones.

Otra aportación clave es la escalabilidad y flexibilidad. La nube permite adaptar de forma rápida los recursos de computación y almacenamiento a las necesidades reales del negocio. En el sector productivo, donde la demanda puede variar en función de campañas, estacionalidad o cambios en la producción, esta capacidad de ajuste evita tanto el sobredimensionamiento como la falta de recursos críticos.

El cloud computing también favorece la integración y el acceso global a la información. Los datos y aplicaciones pueden ser utilizados desde distintas ubicaciones y por diferentes departamentos, lo que mejora la colaboración y la coordinación dentro de la empresa y a lo largo de la cadena de valor. En una economía globalizada, esta accesibilidad es esencial para operar de forma eficiente y competitiva.

Desde el punto de vista de la innovación, la nube acelera el desarrollo y despliegue de nuevos servicios digitales. Las empresas pueden probar soluciones, analizar datos o entrenar modelos de machine learning con rapidez, reduciendo el tiempo necesario para llevar una idea desde el diseño hasta su aplicación en el proceso productivo.

Por último, el cloud computing contribuye a la resiliencia y continuidad del negocio. Las infraestructuras en la nube incorporan mecanismos avanzados de redundancia, recuperación

ante desastres y alta disponibilidad. Esto reduce el riesgo de interrupciones graves y garantiza que los sistemas productivos y de gestión de datos sigan funcionando incluso ante fallos técnicos o incidentes de seguridad.

#### **h) Identifica los principales objetivos de la ciencia de datos en las diferentes empresas.**

La ciencia de datos se ha convertido en un componente estratégico para las empresas, ya que permite transformar grandes volúmenes de datos en conocimiento útil que impacta directamente en la eficiencia, la innovación y la competitividad. En el sector productivo, los objetivos de la ciencia de datos están alineados con la optimización de procesos, la mejora del rendimiento y la anticipación de problemas o oportunidades.

Uno de los objetivos más relevantes es la **optimización de procesos productivos**. Mediante la analítica de datos y modelos predictivos, las empresas pueden identificar cuellos de botella, reducir desperdicios, mejorar la planificación de la producción y maximizar el rendimiento de la maquinaria. Esto no solo incrementa la eficiencia operativa, sino que también reduce costes y mejora la sostenibilidad.

Otro objetivo clave es la **mejora de la toma de decisiones estratégicas**. La ciencia de datos permite a los directivos contar con información precisa y basada en evidencia sobre tendencias de mercado, demanda de productos, comportamiento de los clientes o rendimiento de la cadena de suministro. Esto convierte las decisiones empresariales en acciones más fundamentadas y menos sujetas a incertidumbre.

La **personalización de productos y servicios** es otro enfoque en el que la ciencia de datos tiene un papel destacado. Analizando el comportamiento de clientes y usuarios, las empresas pueden ajustar sus ofertas a necesidades específicas, anticipar preferencias y mejorar la experiencia de cliente, lo que aumenta la fidelización y la ventaja competitiva.

La ciencia de datos también se aplica a la **detección de riesgos y fraudes**. Mediante algoritmos de análisis avanzado y modelos de machine learning, las empresas pueden identificar anomalías en procesos, prever fallos de equipos o detectar comportamientos fraudulentos, evitando pérdidas económicas y daños reputacionales.

Finalmente, un objetivo transversal es **incrementar la eficiencia operativa global**. La integración de datos provenientes de distintas áreas permite una visión holística de la empresa, facilitando la coordinación, reduciendo redundancias y potenciando la innovación continua. Así, la ciencia de datos no solo apoya decisiones puntuales, sino que se convierte en un motor de mejora constante en todos los niveles de la organización.

#### **i) Valora la importancia de la seguridad y su regulación en relación con los datos.**

La seguridad de los datos es un elemento crítico en la economía digital, especialmente en empresas productivas que dependen de información sensible para operar y tomar decisiones



estratégicas. Proteger los datos no solo implica evitar accesos no autorizados, sino garantizar la **confidencialidad, integridad y disponibilidad** de la información a lo largo de todo su ciclo de vida. Sin una seguridad adecuada, la empresa se expone a riesgos de fraude, sabotaje industrial, fugas de información confidencial o pérdidas económicas significativas.

A nivel regulatorio, la protección de datos está respaldada por normativas estrictas como el **Reglamento General de Protección de Datos (RGPD)** en la Unión Europea, que define obligaciones claras para el tratamiento, almacenamiento y transferencia de datos personales. Estas regulaciones obligan a las empresas a implementar políticas de seguridad robustas, a documentar procedimientos de gestión de información y a garantizar la trazabilidad de los datos. El cumplimiento legal no solo evita sanciones, sino que también genera confianza en clientes, proveedores y socios comerciales.

Los sistemas de seguridad abarcan tanto medidas tecnológicas como organizativas. A nivel de equipo y sistema, se utilizan firewalls, cifrado, autenticación multifactor, sistemas de detección de intrusiones y actualizaciones constantes de software. A nivel global, se implementan estrategias de ciberseguridad que incluyen planes de contingencia, monitoreo continuo, análisis de riesgos y formación del personal. Estas acciones combinadas aseguran que la información crítica de la empresa esté protegida frente a amenazas internas y externas.

En el contexto de la digitalización productiva, la ciberseguridad también tiene un impacto directo en la continuidad del negocio. La interrupción de sistemas de control industrial, pérdida de datos de producción o manipulación de información puede traducirse en paradas de línea, defectos en los productos o retrasos en la entrega. Por ello, la seguridad de los datos no es solo un requisito legal o técnico, sino un factor estratégico que protege la reputación, la eficiencia y la competitividad de la empresa.