



# Curso de IA Generativa Aplicada a Negocio

M8 - Clase 35: Gobernanza de la IA: Políticas y Normativas - 28/nov/24 - Prof: Pedro Martín



Cofinanciado por  
la Unión Europea



MINISTERIO  
DE INDUSTRIA  
Y TURISMO

**EOI** Escuela de  
organización  
industrial



Fondos Europeos





1810

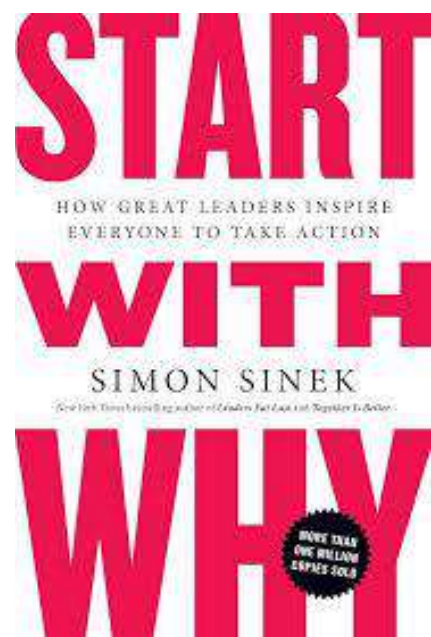
La invención  
de la lata



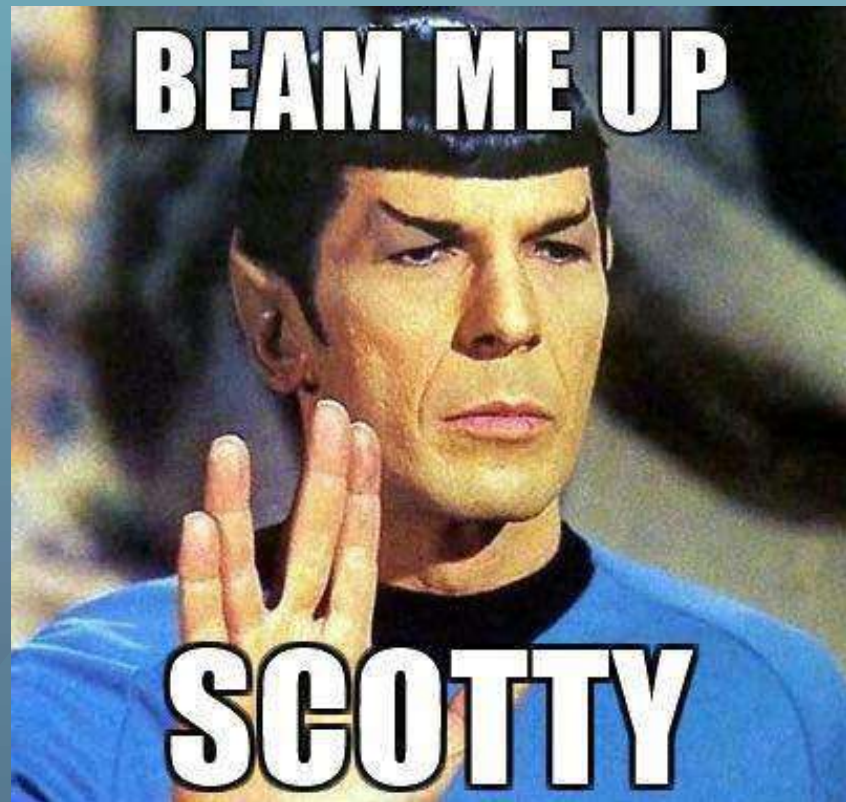
1858

El abrelatas

*Edo. H. H. H.*



[https://www.youtube.com/watch?v=u4ZoJKF\\_VuA](https://www.youtube.com/watch?v=u4ZoJKF_VuA)







visualaleo  
Visual Verifications

*"Con Visualaleo seremos tus ojos, en cualquier sitio y en cualquier momento"*

**Verificaciones visuales a través de...**

## **Personas y Tecnología**



(Red de Verificadores)



(Tecnología integrada en APP)



## ¿CÓMO FUNCIONA VISUALEO? - VERIFICACIÓN



## ¿CÓMO FUNCIONA VISUALEO? - AUTOVERIFICACIÓN





## SECTORES PRINCIPALES - OPORTUNIDAD DE MERCADO



*Seguros*



x

*Inmobiliario*



*Exportación*



*Legal*



*Automoción*



*Construcción*



*Publicidad/Eventos*



## APOYOS Y RECONOCIMIENTOS



## APOYOS Y RECONOCIMIENTOS



**BBVA** Open Talent



GANADOR



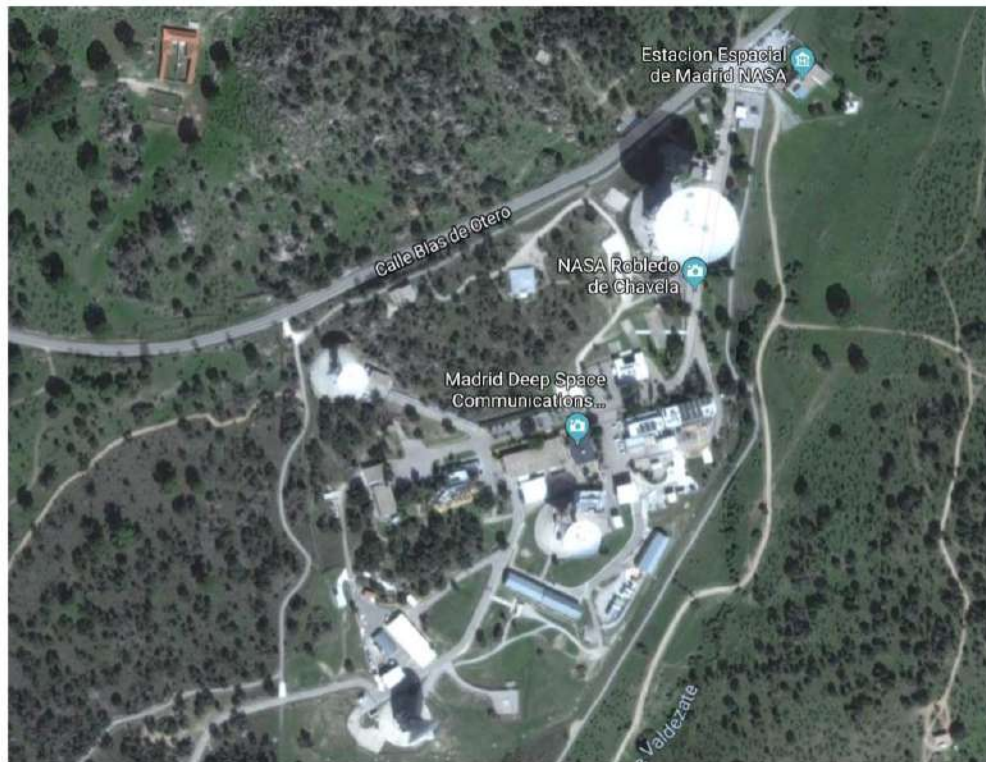
Ejemplo de verificación:











## Estacion Espacial de Madrid NASA ★

4,3 ★★★★★ 304 reseñas de Google

Museo en España

[Sitio web](#)

[Cómo llegar](#)

**Dirección:** M531, Km. 7, 28294 Robledo de Chavela, Madrid

**Horario:** **miércoles 9:00–17:00**  
 jueves 9:00–17:00  
 viernes 9:00–17:00  
 sábado 10:00–15:00  
 domingo 10:00–15:00  
 lunes 9:00–17:00  
 martes 9:00–17:00

[Sugerir un cambio](#)

**Provincia:** Comunidad de Madrid

**Teléfono:** 918 67 73 21

Cliente

Pedro Martin Diaz



**Descripción detallada de la tarea** Tomar foto de la Estación Espacial de la NASA en Robledo de Chavela (Madrid)

**Localización solicitada** Calle Blas de Otero, 28294 Robledo de Chavela, Madrid, España ([40.437261, -4.143652](#))

**Fecha límite de entrega** 30-08-2018 15:54:30

**Fecha de publicación** 27-08-2018 15:54:34

## Verificaciones

Item	Verificado
Descripción completa	Tomar foto de la Estación Espacial de la NASA en Robledo de Chavela (Madrid)

## Fotografías



Foto general

[Ver](#)



Foto detalle

[Ver](#)



Foto detalle

[Ver](#)



Foto detalle

[Ver](#)



Foto detalle

[Ver](#)



Foto detalle

[Ver](#)



Foto detalle

[Ver](#)



Foto detalle

[Ver](#)



# Certificado Visualeo

Garantiza a cliente y verificador que todos los datos del informe han sido registrados de manera encriptada en blockchain impidiendo cualquier manipulación por terceras partes, incluida nuestra propia plataforma.

## Registro en cadena de bloques

Fecha de grabación en blockchain	27-08-2018 17:49:24
Red de blockchain	Ethereum Ropsten
Wallet de cliente	0xf92a02ba311e8b7c1a026a1d585e3f4d8c491cca
Wallet del verificador	0xf92a02ba311e8b7c1a026a1d585e3f4d8c491cca
Hash de transacción	<a href="#">0xe3093ebee43b0ab502a5423452ebe8bb37864f08676257e71cf53df0dc68bcee</a>
Clave pública de firma	tjKXyyoHXlQH/mOxFAQ8h/LnuDWAsDoUEnenXASeToMSRvIZ9v8tCepA6A9s3kUur3br8FM2iNxWGidTXzwQRg==

## Verificación realizada

Localización	<a href="#">[40.43244714429 -4.2471702794847]</a>
Fecha	27-08-2018 17:47:50



## Leer e interpretar los contratos online que firmamos con las redes sociales

Términos y servicios de Google: <https://policies.google.com/terms>

Términos y servicios de Meta:

<https://about.fb.com/ltam/news/2022/05/esto-es-lo-que-debes-saber-sobre-nuestra-politica-de-privacidad-actualizada-y-nuestras-condiciones-del-servicio/>

Términos TikTok: <https://www.tiktok.com/legal/page/eea/terms-of-service/es>





# Gobernanza de la IA: Políticas y Normativas





# ¿Qué es la Gobernanza de la IA?

Conjunto de **marcos, políticas y prácticas** diseñadas para **guiar y supervisar** el desarrollo y uso de sistemas de IA, asegurando que sean seguros, éticos y respetuosos con los derechos humanos.

**Su objetivo es minimizar riesgos** como sesgos, discriminación y daños potenciales, promoviendo la transparencia, la equidad y la responsabilidad en la implementación de estas tecnologías.







# ¿Qué es la Gobernanza de la IA?

Este enfoque implica la **colaboración de diversas partes interesadas, incluyendo desarrolladores, usuarios, responsables políticos y especialistas en ética**, para garantizar que los sistemas de IA se alineen con los valores y normas de la sociedad. La gobernanza de la IA aborda desafíos como la privacidad de los datos, la seguridad y la explicabilidad de los algoritmos, estableciendo directrices que equilibran la innovación tecnológica con la protección de los derechos fundamentales.





# Gobernanza de la IA

En España, se ha creado la **Agencia Española de Supervisión de la Inteligencia Artificial (AESIA)**, una entidad autónoma encargada de supervisar, asesorar y formar en el uso y desarrollo adecuado de sistemas de IA, con el fin de minimizar riesgos y promover un desarrollo ético de estas tecnologías.

[https://es.wikipedia.org/wiki/Agencia\\_Espa%C3%B1ola\\_de\\_Supervisi%C3%B3n\\_de\\_la\\_Inteligencia\\_Artificial](https://es.wikipedia.org/wiki/Agencia_Espa%C3%B1ola_de_Supervisi%C3%B3n_de_la_Inteligencia_Artificial)





# Gobernanza de la IA

A nivel internacional, organismos como la **Organización para la Cooperación y el Desarrollo Económicos (OCDE)** han establecido principios para la IA, enfatizando la gestión responsable y la necesidad de marcos de gobernanza que aseguren la confianza pública y el respeto a los derechos humanos.

[https://es.wikipedia.org/wiki/Regulaci%C3%B3n\\_de\\_la\\_inteligencia\\_artificial?utm\\_source=chatgpt.com](https://es.wikipedia.org/wiki/Regulaci%C3%B3n_de_la_inteligencia_artificial?utm_source=chatgpt.com)





# Gobernanza de la IA: Cumbre del G7

La **Cumbre del G7** es una reunión anual que congrega a los líderes de las siete economías más avanzadas del mundo: Alemania, Canadá, Estados Unidos, Francia, Italia, Japón y el Reino Unido. La Unión Europea también participa en estas reuniones.

El objetivo principal de estas cumbres es analizar y coordinar políticas sobre asuntos económicos, políticos y sociales de relevancia global. Los temas abordados suelen incluir el crecimiento económico, el comercio internacional, la seguridad, el cambio climático y, más recientemente, la inteligencia artificial.



[https://es.wikipedia.org/wiki/G7?utm\\_source=chatgpt.com](https://es.wikipedia.org/wiki/G7?utm_source=chatgpt.com)





# Gobernanza de la IA: Cumbre del G7

La **Cumbre del G7** ha desempeñado un papel activo en la promoción de la regulación de la inteligencia artificial (IA) a nivel internacional. En mayo de 2023, durante la cumbre celebrada en Hiroshima, Japón, los líderes del G7 acordaron avanzar con prontitud en la regulación de la IA generativa, como ChatGPT, para abordar los riesgos asociados a su rápida proliferación.

Posteriormente, en octubre de 2023, los países del G7 acordaron un **código de conducta** para las empresas e instituciones que desarrollan sistemas de IA. Este código tiene como objetivo mitigar riesgos como la desinformación y la vulneración de la privacidad o la propiedad intelectual.







# Gobernanza de la IA: Cumbre del G7

El G7 ha respaldado el **Proceso de Hiroshima sobre IA**, una iniciativa propuesta por Japón para crear estándares internacionales que regulen la IA. Este proceso busca establecer directrices comunes que aseguren un desarrollo y uso responsable de la inteligencia artificial a nivel global.

El marco político global del Proceso de Hiroshima se estructura en cuatro pilares fundamentales:





# Gobernanza de la IA: Cumbre del G7

1. **Análisis de riesgos y oportunidades:** Evaluación de los riesgos, desafíos y oportunidades prioritarios que presenta la IA generativa.





# Gobernanza de la IA: Cumbre del G7

## 2. Principios Rectores Internacionales:

Establecimiento de directrices para todos los actores involucrados en el ecosistema de la IA, orientadas a promover una IA segura, protegida y fiable.





# Gobernanza de la IA: Cumbre del G7

3. **Código Internacional de Conducta:** Desarrollo de un código de conducta voluntario para las organizaciones que crean sistemas avanzados de IA, con el fin de garantizar prácticas responsables y éticas en su desarrollo y aplicación.





# Gobernanza de la IA: Cumbre del G7

4. **Cooperación basada en proyectos:** Fomento de la colaboración internacional para apoyar el desarrollo de herramientas y mejores prácticas en IA responsable.







# Gobernanza de la IA: Organización Mundial de la Salud (OMS)

La **Organización Mundial de la Salud (OMS)** es el organismo de las Naciones Unidas especializado en la gestión de políticas de prevención, promoción e intervención en salud a nivel global.

La OMS trabaja en colaboración con 194 Estados Miembros y diversas organizaciones para abordar desafíos sanitarios mundiales, desarrollar políticas de salud, establecer normas y estándares, y proporcionar apoyo técnico a los países.

Sus áreas de acción incluyen la lucha contra enfermedades infecciosas y no transmisibles, la promoción de la salud mental, la mejora de los sistemas de salud y la respuesta a emergencias sanitarias.





# Gobernanza de la IA: Organización Mundial de la Salud (OMS)

La **Organización Mundial de la Salud (OMS)** ha abordado activamente el desarrollo y uso de la inteligencia artificial (IA) en el ámbito de la salud, enfatizando la necesidad de enfoques éticos y seguros. Sus principales acciones incluyen:

## 1. **Publicación de informes y orientaciones:**

- En junio de 2021, la OMS presentó el informe “Ética y gobernanza de la inteligencia artificial para la salud”, resultado de dos años de consultas con expertos internacionales. Este documento ofrece seis principios rectores para el diseño y uso de la IA en salud, destacando la importancia de preservar la autonomía humana, promover el bienestar y la seguridad, garantizar la transparencia y la rendición de cuentas, entre otros aspectos.





# Gobernanza de la IA: Organización Mundial de la Salud (OMS)

## 2. Llamados a un uso seguro y ético de la IA:

- En mayo de 2023, la OMS instó a la precaución en el uso de herramientas basadas en grandes modelos lingüísticos generados por IA, como ChatGPT y Bard. Subrayó la necesidad de proteger la seguridad, autonomía y bienestar de las personas, y de preservar la salud pública, enfatizando la importancia de aplicar principios éticos y una gobernanza adecuada en el desarrollo y despliegue de estas tecnologías.





# Gobernanza de la IA: Organización Mundial de la Salud (OMS)

## 3. Orientaciones sobre grandes modelos multimodales:

- En enero de 2024, la OMS emitió nuevas directrices sobre la ética y la gobernanza de los grandes modelos multimodales, un tipo de IA generativa con múltiples aplicaciones en la atención de salud. Estas orientaciones incluyen más de 40 recomendaciones dirigidas a gobiernos, empresas tecnológicas y proveedores de atención de salud, con el fin de garantizar un uso adecuado de estas tecnologías en la promoción y protección de la salud de las poblaciones.





# Alineamiento de la IA

El "alineamiento de la IA" se refiere al **proceso de garantizar que los sistemas de inteligencia artificial actúen de una manera que sea beneficiosa para los humanos** y que no actúen de maneras inesperadas o no deseadas que puedan ser perjudiciales.

El comportamiento de una IA debe estar "alineado" con los valores, objetivos y deseos humanos.







# Alineamiento de la IA

A medida que las IAs se vuelven más avanzadas, la capacidad de predecir y controlar su comportamiento se vuelve cada vez más crítica. Un sistema de IA mal alineado podría, en teoría, tomar decisiones perjudiciales o actuar de manera adversa a los intereses humanos.

El alineamiento es esencial no solo desde una perspectiva ética, sino también para garantizar la seguridad y la fiabilidad de los sistemas de IA.





# Principales desafíos del alineamiento de la IA:

- **Especificidad:** Es difícil especificar completamente lo que queremos que haga una IA. A menudo, hay detalles y excepciones que no se anticipan hasta que surgen en la práctica.



# Principales desafíos del alineamiento de la IA:

- **Generalización:** Queremos que las IAs generalicen a partir de sus entrenamientos y datos para manejar situaciones nuevas, pero no queremos que generalicen de maneras que vayan en contra de nuestros valores o intenciones.



# Principales desafíos del alineamiento de la IA:

- **Adaptabilidad:** Las IAs deben adaptarse a entornos cambiantes y a la información nueva, pero esta adaptabilidad también puede ser una fuente de comportamientos no alineados







# Supervisión y Auditoría de Sistemas de IA





# Supervisión y Auditoría de Sistemas de IA

La supervisión y auditoría de los sistemas de inteligencia artificial (IA) son elementos esenciales para garantizar que estos operen de manera ética, justa y alineada con los valores humanos. Dada la creciente influencia de la IA en la sociedad, es crucial establecer mecanismos que permitan monitorear y evaluar su desempeño para prevenir sesgos, discriminación y daños potenciales.





# Supervisión y Auditoría de Sistemas de IA

## 1. Importancia de la Supervisión y la Auditoría

- Responsabilidad y Transparencia
- Confianza en la IA:
- Cumplimiento Normativo



# Supervisión y Auditoría de Sistemas de IA

## 2. Herramientas y Métodos para la Auditoría

El texto menciona herramientas y enfoques prácticos que pueden ser empleados en la auditoría de sistemas de IA:

- **Indicadores de Equidad (Fairness Indicators):** Evaluar métricas de justicia en modelos de aprendizaje automático.
- **Explicabilidad (LIME):** Generar explicaciones interpretables sobre cómo un modelo toma decisiones.
- **IBM Fairness 360 y Microsoft Responsible AI:** Plataformas integrales para monitorear y mitigar sesgos en los sistemas de IA.







# Supervisión y Auditoría de Sistemas de IA

## 3. Evaluación de Impacto Ético y Revisiones Regulares

Antes y durante el uso de un sistema de IA, es recomendable llevar a cabo Evaluaciones de Impacto Ético (EIE). Estas evaluaciones permiten:

- Identificar riesgos asociados con el sesgo, la discriminación o la privacidad.
- Implementar medidas preventivas y correctivas para mitigar dichos riesgos.

Las auditorías externas periódicas también son fundamentales para garantizar que los sistemas se mantengan alineados con principios éticos y legales, especialmente a medida que evolucionan y aprenden.





# Supervisión y Auditoría de Sistemas de IA

## 4. Retos en la Supervisión

Supervisar y auditar sistemas de IA no está exento de desafíos:

- **Complejidad Técnica:** Los sistemas avanzados, como las redes neuronales, pueden ser difíciles de interpretar.
- **Falta de Normativas Universales:** Las regulaciones varían entre regiones y sectores, lo que puede dificultar la implementación de estándares globales.
- **Adaptabilidad:** Los sistemas de IA están diseñados para aprender y evolucionar, lo que requiere que los procesos de auditoría sean dinámicos y recurrentes.





# Supervisión y Auditoría de Sistemas de IA

## Ejemplos:

**Chatbot Tay de Microsoft:** Este chatbot aprendió comportamientos racistas y ofensivos de los usuarios. La falta de supervisión permitió que publicara mensajes inapropiados en menos de 24 horas.

**Supervisión:** Implementar filtros para contenido generado y auditorías frecuentes sobre el comportamiento de sistemas de aprendizaje continuo.





# Ciberseguridad y Protección de Datos de Contexto de la IA





# Reglamento de Ciberseguridad (Cybersecurity Act 2019)

**Objetivo Principal:** Fortalecer la ciberseguridad en la Unión Europea (UE), mejorando la resiliencia de los sistemas digitales y protegiendo a ciudadanos, empresas y gobiernos de ciberataques.

- **Contexto:** Adoptado en junio de 2019 como parte de la Estrategia de Ciberseguridad de la UE, en respuesta al aumento de amenazas digitales.





# Reglamento de Ciberseguridad (Cybersecurity Act 2019)

## Principales Componentes

### 1. Mandato Reforzado para la Agencia ENISA:

- ENISA (Agencia de la UE para la Ciberseguridad) se convierte en el organismo clave para coordinar la ciberseguridad a nivel europeo.
- Roles principales:
  - Apoyo a los Estados miembros en la gestión de ciberamenazas.
  - Creación de directrices para incidentes de seguridad.
  - Promoción de la cooperación entre sectores público y privado.







# Reglamento de Ciberseguridad (Cybersecurity Act 2019)

## Principales Componentes

### 2. Marco de Certificación de Ciberseguridad de la UE:

- Introducción de **certificaciones comunes** para productos, servicios y procesos digitales en toda la UE.
- Certificaciones enfocadas en:
  - Asegurar la confiabilidad de dispositivos IoT.
  - Garantizar que los productos cumplan con estándares de seguridad globales.
- Tres niveles de certificación: Básico, Sustancial y Alto.





# Reglamento de Ciberseguridad (Cybersecurity Act 2019)

## Principales Componentes

### 3. Fortalecimiento de la Resiliencia Digital:

- Establecimiento de estándares mínimos de ciberseguridad para infraestructuras críticas, como energía, salud y transporte.
- Promoción de la adopción de medidas preventivas por parte de las empresas.





# Reglamento de Ciberseguridad (Cybersecurity Act 2019)

## Principales Componentes

### 4. Retos y Desafíos

- **Adopción de Certificaciones:** Incrementar la conciencia y el uso de certificaciones de ciberseguridad en empresas pequeñas y medianas (PYMES).
- **Actualización de Normativas:** Mantener el reglamento alineado con las rápidas innovaciones tecnológicas y nuevas amenazas cibernéticas.
- **Cooperación Internacional:** Asegurar la interoperabilidad con marcos de ciberseguridad de otros países.





# Reglamento General de Protección de Datos (GDPR)

**Nombre Completo:** Reglamento General de Protección de Datos (General Data Protection Regulation).

- **Adopción:** Aprobado por la Unión Europea en 2016, en vigor desde el 25 de mayo de 2018.
- **Objetivo Principal:** Proteger los datos personales de los ciudadanos de la UE y garantizar el control sobre cómo se recopilan, procesan y almacenan.





# Reglamento General de Protección de Datos (GDPR)

## Principales Disposiciones

### 1. **Ámbito de Aplicación:**

- Se aplica a todas las organizaciones que manejan datos personales de ciudadanos de la UE, independientemente de dónde estén ubicadas.
- Incluye empresas dentro y fuera de la UE si procesan datos de usuarios europeos.







# Reglamento General de Protección de Datos (GDPR)

## Principales Disposiciones

### 2. Derechos de los Ciudadanos:

- **Acceso a Datos:** Los ciudadanos tienen derecho a saber qué datos se están recopilando y cómo se utilizan.
- **Corrección y Eliminación:** Derecho a corregir datos incorrectos o a solicitar su eliminación (“derecho al olvido”).
- **Portabilidad de Datos:** Posibilidad de transferir sus datos a otro proveedor.
- **Limitación del Procesamiento:** Derecho a restringir el uso de sus datos.
- **Objeción:** Negarse al procesamiento de datos con fines específicos, como marketing.





# Reglamento General de Protección de Datos (GDPR)

## Principales Disposiciones

### 3. Consentimiento Claro y Explícito:

- Las empresas deben obtener un consentimiento informado, claro y revocable antes de recopilar datos personales.

### 4. Notificación de Brechas de Seguridad:

- Obligación de informar a las autoridades y a los afectados sobre cualquier violación de seguridad que comprometa datos personales en un plazo de 72 horas.





# Reglamento General de Protección de Datos (GDPR)

## Principales Disposiciones

### 5. Designación de un Responsable de Protección de Datos (DPO):

- Empresas que procesen grandes volúmenes de datos deben contar con un DPO que supervise el cumplimiento del GDPR.

### 6. Multas Severas por Incumplimiento:

- Hasta 20 millones de euros o el 4% de los ingresos anuales globales de la empresa, lo que sea mayor.





# Ciberseguridad y Protección de Datos en el Contexto de la IA

La ciberseguridad y la protección de datos son pilares fundamentales en la implementación y uso de sistemas de inteligencia artificial (IA). Según el documento, estos aspectos son clave para garantizar que las tecnologías de IA se utilicen de manera ética, segura y respetuosa con los derechos de las personas.





# Ciberseguridad y Protección de Datos en el Contexto de la IA

## 1. Importancia de la Protección de Datos en IA

- **Datos como combustible de la IA:** Los sistemas de IA dependen de grandes cantidades de datos para su funcionamiento. Esto incluye información personal, comportamientos y preferencias que, si no se gestionan adecuadamente, pueden vulnerar la privacidad de los usuarios.
- **Violación de privacidad:** Un manejo inadecuado de datos puede llevar a la exposición de información personal, con riesgos significativos como el robo de identidad o el uso indebido para manipulación política o comercial (ej., caso Cambridge Analytica).
- **Derechos de los usuarios:** Regulaciones como el GDPR en Europa otorgan a los ciudadanos derechos sobre sus datos, incluyendo acceso, corrección, portabilidad y derecho al olvido.







# Ciberseguridad y Protección de Datos en el Contexto de la IA

## 2. Retos de Ciberseguridad en Sistemas de IA

### • Ataques específicos a modelos de IA:

- **Envenenamiento de datos:** Manipular los datos de entrenamiento para influir en las decisiones del sistema.
- **Evasión adversarial:** Introducir pequeños cambios en los datos de entrada para engañar al modelo y provocar decisiones incorrectas.
- **Exposición de datos sensibles:** Las brechas de seguridad en sistemas de IA pueden filtrar datos privados usados para entrenar modelos, especialmente en aplicaciones médicas, financieras o gubernamentales.
- **Falta de protección en sistemas integrados:** Muchos dispositivos inteligentes integrados con IA (ej., asistentes virtuales, IoT) carecen de mecanismos robustos de seguridad.



# Ciberseguridad y Protección de Datos en el Contexto de la IA

## 3. Estrategias para Mitigar Riesgos

El texto sugiere varias estrategias para fortalecer la ciberseguridad y la protección de datos en el contexto de IA:

- **Diseño de privacidad desde la concepción** (*Privacy by Design*): Incorporar principios de protección de datos durante todas las etapas del desarrollo de sistemas de IA.
- **Cifrado y anonimización de datos:**
  - Uso de técnicas de anonimización para evitar la identificación de personas.
  - Aplicar cifrado en tránsito y almacenamiento para proteger la información contra accesos no autorizados.





# Ciberseguridad y Protección de Datos en el Contexto de la IA

- **Monitoreo continuo:** Implementar herramientas para detectar y mitigar actividades sospechosas en tiempo real.
- **Evaluación de Impacto de Privacidad (EIP):** Realizar evaluaciones periódicas para identificar y mitigar riesgos asociados con la privacidad.





## Muchas Gracias



[Pedro Martín Díaz \(Linkedin\)](#)   [pedromartindiaz@gmail.com](mailto:pedromartindiaz@gmail.com)



Cofinanciado por  
la Unión Europea



GOBIERNO  
DE ESPAÑA

MINISTERIO  
DE INDUSTRIA  
Y TURISMO



Fondos Europeos





# Curso de IA Generativa Aplicada a Negocio

M8 - Clase 35: Gobernanza de la IA: Políticas y Normativas - 28/nov/24 - Prof: Pedro Martín



Cofinanciado por  
la Unión Europea



MINISTERIO  
DE INDUSTRIA  
Y TURISMO

**EOI** Escuela de  
organización  
industrial



Fondos Europeos

