Fundamentos de Ingeniería Informática

Modulo III - Unidad 10 Bases de datos

Profesor: Héctor Molina García







Los datos son elementos crudos, desorganizados y sin refinar.

vs Información



La información es la organización e interpretación de esos elementos.



La información se clasifica en función de cómo se almacenan y organizan los datos a través de algún tipo de estructura y/o etiquetado.



La información se clasifica en función de cómo se almacenan y organizan los datos a través de algún tipo de estructura y/o etiquetado.

Estructurado

La información se organiza utilizando algún modelo de datos o esquema. Hay una definición precisa del significado de cada elemento.

Hay una estructura definida e identificable.

Databases



La información se clasifica en función de cómo se almacenan y organizan los datos a través de algún tipo de estructura y/o etiquetado.

Estructurado

La información se organiza utilizando algún modelo de datos o esquema. Hay una definición precisa del significado de cada elemento.

Hay una estructura definida e identificable.

Semiestructurado

La información no obedece a la estructura tabular de los modelos de datos asociados con bases de datos relacionales u otras formas de tablas de datos. Hay etiquetas u otros marcadores que incluyen cierta semántica.

Databases

XML, HTML, JSON



La información se clasifica en función de cómo se almacenan y organizan los datos a través de algún tipo de estructura y/o etiquetado.

Estructurado

La información se organiza utilizando algún modelo de datos o esquema. Hay una definición precisa del significado de cada elemento.

Hay una estructura definida e identificable.

Semiestructurado

La información no obedece a la estructura tabular de los modelos de datos asociados con bases de datos relacionales u otras formas de tablas de datos. Hay etiquetas u otros marcadores que incluyen cierta semántica.

Desestructurado

Los datos no estructurados son información que no está organizada de acuerdo con un modelo o esquema de datos preestablecido.

Databases

XML, HTML, JSON

Texto sin formato



Existen diferentes sistemas de almacenamiento en función del formato que queramos utilizar para almacenar los datos.

Archivos

Archivos de texto plano con información útil.
Por lo general, almacenan datos sin procesar.
La información es desordenada y desestructurada.



Existen diferentes sistemas de almacenamiento en función del formato que queramos utilizar para almacenar los datos.

Archivos

Archivos de texto plano con información útil.
Por lo general, almacenan datos sin procesar.
La información es desordenada y desestructurada.

Archivos distribuidos

Almacenamiento
masivo en archivos de
datos estructurados.
Se almacenan en un
entorno distribuido.
La información tiene
cierta estructura.



Existen diferentes sistemas de almacenamiento en función del formato que queramos utilizar para almacenar los datos.

Archivos

Archivos de texto plano con información útil.
Por lo general, almacenan datos sin procesar.
La información es desordenada y desestructurada.

Archivos distribuidos

Almacenamiento
masivo en archivos de
datos estructurados.
Se almacenan en un
entorno distribuido.
La información tiene
cierta estructura.

Base de datos

Almacenamiento en bases de datos SQL y NOSQL.





Una base de datos es una colección de información almacenada en una computadora o sistema informático en una forma a la que se puede acceder, recuperar y modificar fácilmente..

- Facilitar el almacenamiento de grandes cantidades de información.
- Facilitar la recuperación de información de forma rápida y flexible.
- Facilitar la organización permitiendo la vinculación de diferentes tipos de información
- Facilitar la impresión y distribución de información en una amplia variedad de formas.



Existen diferentes tipos de bases de datos, dependiendo de:

Cómo se organiza la información

Bases de datos SQL

Bases de datos NoSQL (no solo SQL)

Cómo se almacena la información en el nivel físico

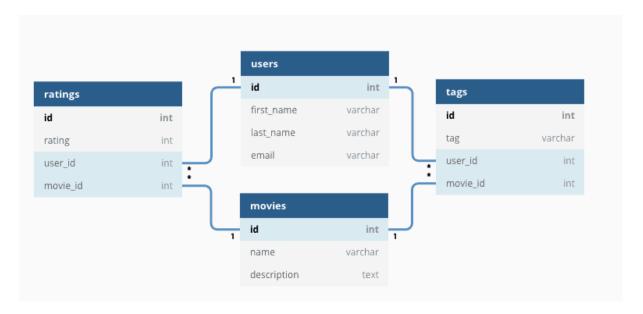
Centralizada

Distribuido

Nube



Una base de datos relacional (RDB) es una forma de estructurar información en tablas, filas y columnas. Un RDB tiene la capacidad de establecer vínculos (o relaciones) entre la información mediante la unión de tablas, lo que facilita la comprensión y la obtención de información sobre la relación entre varios puntos de datos.





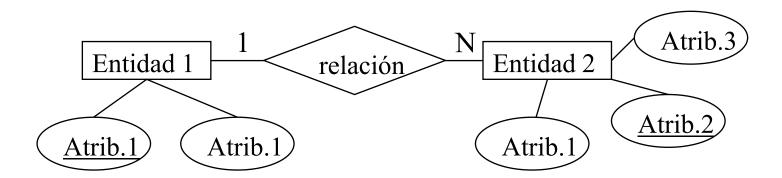
A menudo, una base de datos relacional se puede denominar base de datos SOL.

Las bases de datos relacionales se describen utilizando un diagrama entidadrelación que se compone de tres elementos básicos:

- **Entidad**: cosa, persona, lugar, unidad, objeto o cualquier elemento cuya información se almacene. Por ejemplo, los usuarios del diagrama anterior.
- **Atributo**: Es información específica (características) de una entidad. Son las propiedades de la entidad. Por ejemplo, el nombre de pila de un usuario.
- Clave primaria: Es un atributo especial que identifica cada registro en una entidad.
- Relación: son vínculos o relaciones, como su nombre indica, entre entidades.
 Por ejemplo, un usuario específico puede calificar varias películas de acuerdo con el diagrama anterior..



- Representación de un diagrama E-R
 - Las Entidades se representan con rectángulos
 - Sus Atributos con óvalos, aquellos que constituyen una clave primaria se subrayan.
 - Las Relaciones entre entidades con rombos. Los extremos de la relación se etiquetan para indicar el tipo de esta:
 - Un 1 en un extremo indica que la relación se establece entre una única entidad de ese extremo
 - Una N (o M) indica que la relación se establece con más de una entidad de ese extremo





TIPOS DE RELACIONES

- Dados dos conjuntos de entidades A y B, las relaciones entre ambos pueden ser:
 - Una relación uno a uno (1:1): Una entidad A se relaciona únicamente con una entidad B y viceversa. Esta relación no es muy común, porque a menudo una de las entidades se define como atributo de la otra.
 - EJ: la relación coche-matrícula. Cada coche tiene una única matrícula y cada matrícula pertenece únicamente a un coche
 - Relación uno a varios (1:N): Una entidad en A se relaciona con cero o muchas entidades en B. Pero una entidad en B se relaciona con una única entidad en A.
 - Ej: la relación entre Clientes y Pedidos. Un cliente puede realizar cualquier número de pedidos. Pero cada pedido es realizado únicamente por un cliente.
 - Una relación de varios a varios (N:M): Una entidad en A se puede relacionar con 0 o muchas entidades en B y viceversa.
 - Ejemplo: la relación entre Productos y Pedidos. Un solo pedido puede incluir varios productos. Por otro lado, un único producto puede aparecer en muchos pedidos.

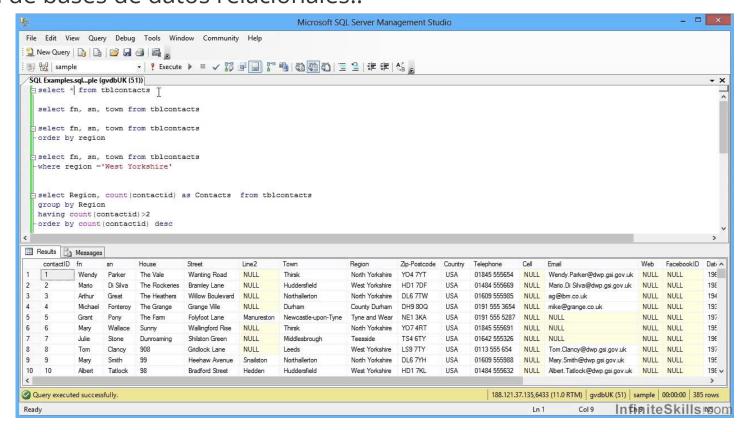


DataBase Management System (DBMS) es un software de aplicación diseñado para almacenar, recuperar, consultar y administrar datos. Los DBMS proporcionan algunas funciones que permiten la administración de una base de datos y sus datos:

- Definición: Creación, modificación y eliminación de definiciones que definen la organización de los datos.
- Gestión y explotación: Inserción, modificación, consulta y supresión de los datos.
- Administración: Registro y supervisión de usuarios, aplicación de la seguridad de los datos, supervisión del rendimiento, mantenimiento de la integridad de los datos, gestión del control de simultaneidad y recuperación de información que ha sido dañada por algún evento, como un error inesperado del sistema..



SQL (Structured Query Language) es un lenguaje de programación específico de dominio diseñado para administrar y recuperar información de sistemas de administración de bases de datos relacionales..





SQL (Structured Query Language) es un lenguaje de programación específico de dominio diseñado para administrar y recuperar información de sistemas de administración de bases de datos relacionales.

```
CREATE TABLE shop (
 article INT(4) UNSIGNED ZEROFILL DEFAULT '0000' NOT
NULL,
dealer CHAR(20)
                                   DEFAULT ''
                                                   NOT
NULL,
        DOUBLE (16,2)
                                  DEFAULT '0.00' NOT
price
NULL,
 PRIMARY KEY(article, dealer));
INSERT INTO shop VALUES
(1, 'A', 3.45), (1, 'B', 3.99), (2, 'A', 10.99), (3, 'B', 1.45),
(3,'C',1.69),(3,'D',1.25),(4,'D',19.95);
INSERT INTO shop (article, dealer, price) VALUES
(1, 'A', 3.45);
```

SELECT * FROM shop;						
+	+· +·	dealer		-	i	
0001 0001 0002 0003 0003 0004	 	A B A B C D		3.45 3.99 10.99 1.45 1.69 1.25 19.95		
+	+-		+-		-+	



SQL (Structured Query Language) es un lenguaje de programación específico de dominio diseñado para administrar y recuperar información de sistemas de administración de bases de datos relacionales.

```
CREATE TABLE shop (
                                                            SELECT * FROM shop;
 article INT(4) UNSIGNED ZEROFILL DEFAULT '0000' NOT
NULL,
 -dealer CHAR(20)
                                                            | article | dealer | price |
                                               NOT
                                 DEFAULT ''
NULL,
 price
       DOUBLE (16,2)
                               DEFAULT '0.00' NOT
                                                                 0001 | A
                                                                               1 3.45 I
NULL,
                                                                 0001 | B | 3.99 |
 PRIMARY KEY(article, dealer));
                                                                 0002 | A | 10.99 |
                                                                 0003 | B | 1.45 |
                                                                 0003 | C | 1.69 |
                                                                 0003 | D | 1.25 |
                                                                 0004 | D
                                                                             | 19.95 |
INSERT INTO shop VALUES
(1, 'A', 3.45), (1, 'B', 3.99), (2, 'A', 10.99), (3, 'B', 1.45),
                                                            +-----+
(3, 'C', 1.69), (3, 'D', 1.25), (4, 'D', 19.95);
INSERT INTO shop (article, dealer, price) VALUES
(1, 'A', 3.45);
                  Hay diferentes formas de insertar datos en una base de datos, pero
```

estas son las más comunes. Aunque, la sintaxis puede cambiar



según el DBMS.

OFF-EXAM CONTENT

Queremos crear una base de datos para una tienda online.

¿Qué entidades necesitamos para almacenar los datos básicos?



Queremos crear una base de datos para una tienda online.

¿Qué entidades necesitamos para almacenar los datos básicos?

Clientes, Productos y pedidos.



Queremos crear una base de datos para una tienda online.

¿Qué entidades necesitamos para almacenar los datos básicos?

Clientes, Productos y pedidos.

¿Qué información necesitamos para almacenar los datos del cliente?



Queremos crear una base de datos para una tienda online.

¿Qué entidades necesitamos para almacenar los datos básicos?

Clientes, Productos y pedidos.

¿Qué información necesitamos para almacenar los datos del cliente?

- Cliente: Nombre, Apellidos, DNI, dirección de facturación, dirección de envío y más...
- Producto: Referencia, nombre, descripción, precio, unidades disponibles y más ...
- Orden: Order_id, fecha, client_id, product_id, estado y más ...



Queremos crear una base de datos para una tienda online.

¿Qué entidades necesitamos para almacenar los datos básicos?

Clientes, Productos y pedidos.

¿Qué información necesitamos para almacenar los datos del cliente?

- Cliente: Nombre, Apellidos, DNI, dirección de facturación, dirección de envío y más...
- Producto: Referencia, nombre, descripción, precio, unidades disponibles y más ...
- Orden: Order_id, fecha, client_id, product_id, estado y más ...

¿Qué relaciones necesitamos?



Queremos crear una base de datos para una tienda online.

¿Qué entidades necesitamos para almacenar los datos básicos?

Clientes, Productos y pedidos.

¿Qué información necesitamos para almacenar los datos del cliente?

- Cliente: Nombre, Apellidos, DNI, dirección de facturación, dirección de envío y más...
- Producto: Referencia, nombre, descripción, precio, unidades disponibles y más ...
- Orden: Order_id, fecha, client_id, product_id, estado y más ...

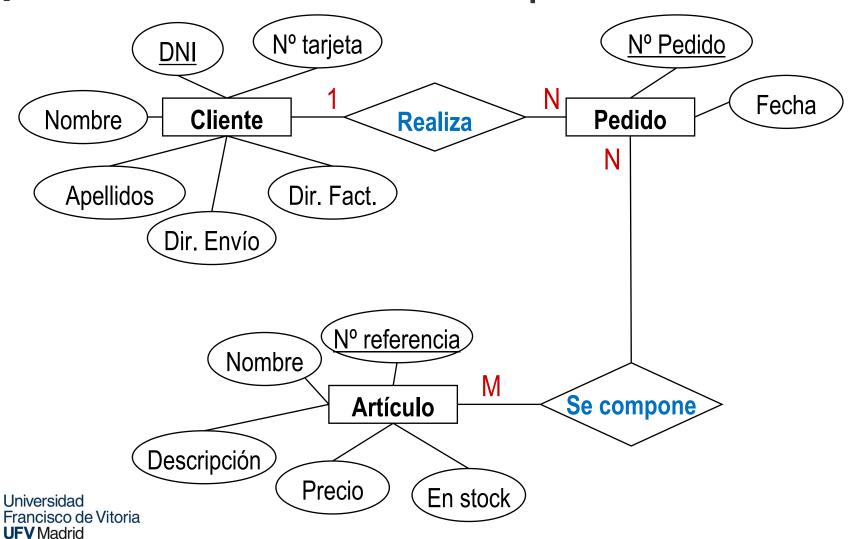
¿Qué relaciones necesitamos?

1:N → cliente:pedido

N:M → producto:pedido



Queremos crear una base de datos para una tienda online.



Queremos crear una base de datos para una tienda online.

Definición de la información

- Las diferentes entidades se convierten en tablas, donde cada atributo es un campo de la tabla.

Los campos tienen un tipo: varchar, int, double, char, etc.

Cada elemento concreto dentro de la tabla (fila) se llama registro.

Las claves primarias se crean utilizando campos que deben ser únicos para identificar cada elemento concreto.



Queremos crear una base de datos para una tienda online.

Definición de las cardinalidades de las relaciones

La relación 1 a muchos (1:N) se incluye para describir cuando un cliente tiene muchos pedidos, pero cada pedido específico solo está relacionado con un cliente.

Esta relación incluirá la clave principal del lado de la entidad 1 en el lado de la entidad N.

El campo DNI (clave primaria del cliente) se incluye como campo de la tabla de pedidos. Este nuevo campo se denomina clave externa.

UNA CLAVE EXTERNA es un campo (o colección de campos) en una tabla, que hace referencia a la CLAVE PRIMARIA en otra tabla.



Queremos crear una base de datos para una tienda online.

Definición de las cardinalidades de las relaciones

La relación muchos a muchos (N:M) se incluye para describir cuándo un pedido se compone de algunas unidades de producto.

Esta relación generará una nueva tabla en la Base de datos. Puede incluir otros atributos como unidades, precio, etc.

Order id	Reference
12324	364834034843784

Esta tabla incluirá las CLAVES PRINCIPALES de las tablas de la relación.







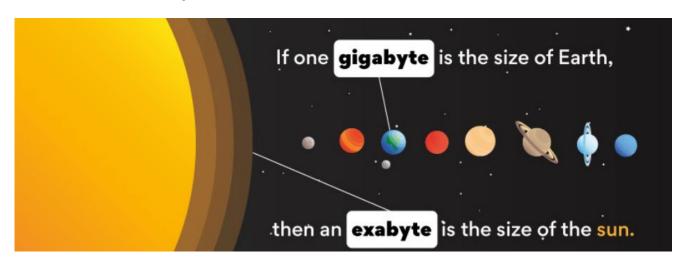
Definición de "Big Data"

El término "Big Data" describe el gran volumen de datos, tanto estructurados como no estructurados, que se crean diariamente por la sociedad actual.





Desde la escritura cuneiforme, el sistema de escritura más antiguo conocido hasta la fecha, hasta los centros de datos modernos, la raza humana siempre ha recopilado información. Además, se predice que para 2030 nuestra civilización generará varios yottabytes de información por año..



El yottabyte es actualmente la mayor unidad reconocida de almacenamiento de datos para dispositivos y servicios en la nube.







Leyes y ética

Evolución de las leyes de datos en España

1978: Constitución Española. Art. 18.4 donde se garantiza el derecho de las personas al honor y la intimidad personal y familiar.

1992: LORTAD. Es la Ley Orgánica de Protección del Tratamiento Automatizado de los Datos de Carácter Personal (no vigente en la actualidad).

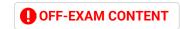
1994: Reglamento que desarrolla determinados aspectos de la LORTAD (este Reglamento sigue vigente a pesar de la derogación de la LORTAD).

1995: Directiva comunitaria relativa a la protección de las personas físicas en lo que respecta al tratamiento de datos personales y a la libre circulación de estos datos. La LOPD española se deriva de esta Directiva.

1999: Reglamento de Medidas de Seguridad (RMS). Especifica las medidas de seguridad técnicas y organizativas que se deben adoptar para los ficheros que contengan datos de carácter personal (11 de junio de 1999).

1999: LOPD - Ley Orgánica de Protección de Datos de Carácter Personal (adaptación de la antigua LORTAD a la Directiva Comunitaria de 1995). La ley sólo se aplica a los datos personales de las personas físicas, no de las personas jurídicas (empresas) (13 de diciembre de 1999).





LOPD - Ley Orgánica de Protección de datos

La Ley Orgánica 15/1999, de 13 de diciembre, de Protección de Datos de Carácter Personal (LOPD) fue una ley orgánica española que garantizaba y protegía el tratamiento de datos personales, las libertades públicas y los derechos humanos fundamentales, y especialmente del honor y la intimidad personal y familiar..

- Regular el tratamiento de los datos y ficheros, de carácter personal, independientemente del soporte en el que sean tratados.
- Los derechos de los ciudadanos sobre ellos y las obligaciones de quienes los crean o tratan.



LOPD en la Unión Europea

El GRDP (Reglamento General de Protección de Datos) es una regulación en la legislación de la Unión Europea (UE) sobre protección de datos y privacidad en el Espacio Económico Europeo (EEE). Esta ley es obligatoria desde el 25 de mayo de 2018 en todos los países miembros de la Unión Europea.

La LOPD, como ley nacional (España), se integró en el RGPD, que unificó toda la normativa europea de protección de datos bajo un mismo paraguas legal. Para que todos los ciudadanos europeos tengan los mismos derechos y garantías con respecto a la información personal.







LOPD en la Unión Europea

La LOPD-GDD (Ley Orgánica de Protección de Datos y Garantía de los Derechos Digitales) es la nueva ley nacional integrada en el GRPD, que unificó toda la normativa europea de protección de datos bajo un mismo paraguas legal. Esta ley amplía la normativa definida en el RGPD.







Cómo implantar la nueva LOPD-GDD en una empresa?

LOPD en la Unión Europea

Cualquier tipo de empresa o negocio que trate datos sensibles de terceros, deberá cumplir con todos y cada uno de los requisitos establecidos en la nueva normativa de la Ley de Protección de Datos Personales y Garantías de los Derechos Digitales. La LOPD-GDD se aplicará cuando se administren los siguientes tratamientos:

- Tratamiento de datos de empresarios individuales y profesionales liberales.
- Operaciones comerciales.
- Uso de sistemas de vigilancia.
- Sistemas de exclusión publicitaria.
- Canales de comunicación y denuncias.
- Sistemas de información crediticia.





LOPD en la Unión Europea

Hay algunos puntos importantes en la nueva ley:

Protección de menores: El consentimiento de un menor sólo será válido cuando sea mayor de catorce años, siendo necesaria la autorización del padre, madre o tutor si no lo fuere.

Control de datos personales: Para evitar el uso de datos personales para uso comercial sin consentimiento previo, la LOPD-GDD establece que el control de los datos personales recae directamente en el usuario, requiriendo siempre su consentimiento para utilizarlos.

Privacidad de los empleados: Está prohibido realizar grabaciones en las zonas destinadas al descanso de los trabajadores, aseos y otros lugares destinados al ocio.



Laws and ethics

LOPD en la Unión Europea

Hay algunos puntos importantes en la nueva ley:

Derecho al olvido: establece el derecho a suprimir datos en redes sociales y otros servicios equivalentes.

Datos de personas fallecidas: En caso de fallecimiento, cualquier familiar vinculado a la persona fallecida podrá solicitar el acceso, rectificación o supresión de los datos compartidos.

Información clara sobre el uso de los datos: Las empresas deben informar a los usuarios de forma clara, sencilla y concisa sobre el posible uso de los datos personales que se les han facilitado.

Las empresas podrían ser multadas con hasta 20 millones de euros.







¿Qué pasa con la ética?





After Uber, Tesla incidents, can artificial **intelligence** be trusted?

Apr 9, 2018 | News Stories

The reliability of self-driving cars and other forms of artificial intelligence is one of several factors that affect humans' trust in AI, **machine learning** and other technological advances, write two Missouri University of Science and Technology researchers in a recent journal article. "Trust is the cornerstone of ...





La ética, también llamada filosofía moral, la disciplina que se ocupa de lo que es moralmente bueno y malo y moralmente correcto e incorrecto. El término también se aplica a cualquier sistema o teoría de valores o principios morales..

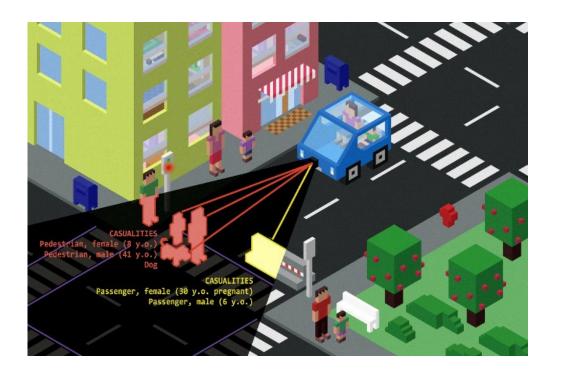
¿Por qué necesitamos ética cuando estamos creando software?

- Los humanos tenemos sesgos que incluimos en la información que creamos y analizamos.
- Durante el diseño y desarrollo de software y datos aportamos conocimiento inconsciente, debemos preguntarnos si hemos tenido en cuenta suficientes ejemplos.
- Ideologías, expresiones, información técnica validada, luz natural/artificial, casos atípicos, etc..





https://www.moralmachine.net/



Option 1

In this case, the self-driving car with a sudden brake failure will swerve and drive through a crosswalk in the other lane. This will result in the death of an elderly woman, two athletes and a child.

Please note that affected pedestrians are complying with the law when crossing with the green signal

Option 2

In this case, the self-driving car with sudden brake failure will continue forward and crash into a concrete barrier. This will result in the death of a child, a pregnant woman (passengers) and a dog."

https://www.eeworldonline.com/this-mit-game-lets-you-choose-who-lives-and-dies-in-a-self-driving-car-wreck/

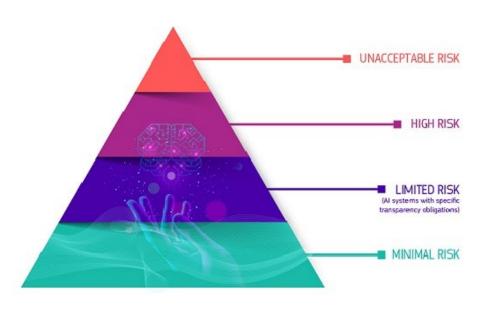






Cómo estamos tratando de crear ¿Software o datos éticamente?

Nuevo marco regulatorio sobre IA (Unión Europea).



- Riesgo inaceptable: un conjunto muy limitado de usos especialmente perjudiciales de la IA que contravienen los valores de la UE porque violan los derechos fundamentales. Puntuación social para los gobiernos (Esto está sucediendo en China).
- Alto riesgo: Un número limitado de sistemas de IA definidos en la propuesta que crean un impacto adverso en la seguridad de las personas o sus derechos fundamentales (protegidos por la Carta de los Derechos Fundamentales de la UE) se consideran de alto riesgo.
- Infraestructura.
- Educación.
- Seguridad.
- Servicios públicos.
- Controles de inmigración o de líneas fronterizas.
- Riesgo limitado: Para determinados sistemas de IA se imponen requisitos específicos de transparencia, por ejemplo, cuando existe un riesgo claro de manipulación (por ejemplo, mediante el uso de chatbots). Los usuarios deben ser conscientes de que están interactuando con una máquina.
- Riesgo mínimo: Todos los demás sistemas de IA pueden desarrollarse y utilizarse sujetos a la legislación vigente sin obligaciones legales adicionales.
- La gran mayoría de los sistemas de IA utilizados actualmente en la UE entran en esta categoría.





Hay reglas en otros países y estados.

- GDPR: Article 22 empowers individuals with the right to demand an explanation of how an automated system made a decision that affects them.
- Algorithmic Accountability Act 2019: Requires companies to provide an assessment of the risks posed by the automated decision system to the privacy or security and the risks that contribute to inaccurate, unfair, biased, or discriminatory decisions impacting consumers
- California Consumer Privacy Act: Requires companies to rethink their approach to capturing, storing, and sharing personal data to align with the new requirements by January 1, 2020.
- Washington Bill 1655: Establishes guidelines for the use of automated decision systems to protect consumers, improve transparency, and create more market predictability.
- Massachusetts Bill H.2701: Establishes a commission on automated decision-making, transparency, fairness, and individual rights.
- Illinois House Bill 3415: States predictive data analytics determining creditworthiness or hiring decisions may not include information that correlates with the applicant race or zip code.



