

# Bases de Datos 1



Alejandra Lliteras  
Prof. Titular

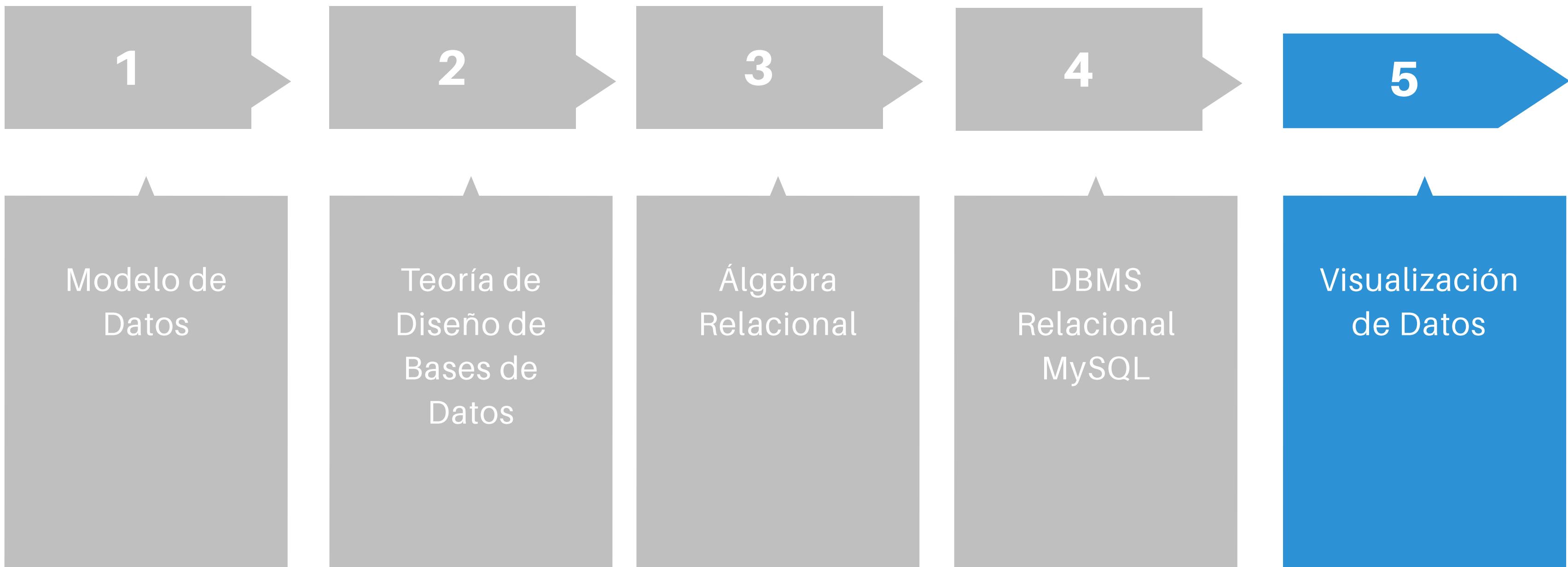


Federico Orlando  
Prof. Adjunto



# TEMAS GENERALES

## Bases de Datos 1



# **TEMAS Y SUBTEMAS**

## **Hoy veremos Datos Abiertos**

- 1.Definiciones**
- 2.Metadatos**
- 3.Principios**
- 4.Niveles de apertura**
- 5.Beneficios**
- 6.Licencias de Datos Abiertos**
- 7.Tecnologías para Datos Abiertos**
- 8.Portales de Datos Abiertos**
- 9.Marco Normativo**
- 10.Dataset**
- 11.Wikidata**

# Definiciones

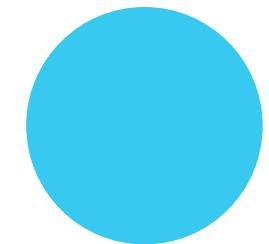
## Dato e información

Un dato es un elemento descontextualizado que puede dar origen a la generación de información. Son símbolos que describen hechos, condiciones, valores o situaciones. Un dato puede ser una letra, un número o cualquier símbolo que representa una cantidad, una medida, una palabra o una descripción.

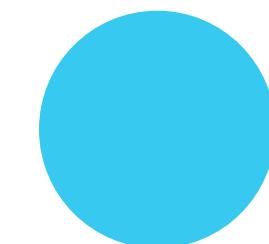
La información es un dato procesado con valor agregado, dotado de relevancia, utilidad e interpretación.

# Definiciones

## Datos públicos y abiertos



**Un dato público es cualquier dato generado en el ámbito gubernamental, o que se encuentra bajo su guarda**



**Los datos abiertos son aquellos de origen público o no a los que cualquier persona puede acceder, usar y compartir libremente. Sólo deben atribuirse y compartirse con la misma licencia con la que fueron publicados.**

# Definiciones

## Datos abiertos

No son datos abiertos:

- Archivos de datos generados en formatos no susceptibles de ser reutilizados por un sistema informático (imágenes de texto, PDFs o documentos bloqueados por algún sistema de protección).
- Datos estadísticos, agregados o resumidos.
- Memorias e informes de gestión anuales.
- Textos legislativos, ordenanzas, reglamentos, normativas individuales.

# **Metadatos**



# Definiciones

## Metadatos

- Ciertas características o atributos de los datos son definidos para facilitar su comprensión. Se las conocen como "metadatos".
- Suele decirse, por esa razón, que "los metadatos son datos sobre los datos".
- Los metadatos pueden describir cómo un dato es representado, cuáles son los rangos de valores aceptables, su relación con otros datos y cómo estos deben ser etiquetados. También pueden proporcionar otra información pertinente, como el administrador responsable, leyes y reglamentos asociados, y la política de gestión de acceso.

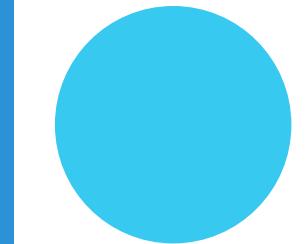
# Definiciones

## Metadatos

Un dato que no puede reutilizarse por falta de información,  
no está realmente abierto

# Definiciones

## Lenguajes de Metadatos



Existen lenguajes estándar para los metadatos, que se basan en estándares ISO para garantizar la interoperabilidad y una amplia reutilización:

- Dublin Core
- el vocabulario de catálogo de datos de la W3C (DCAT)
- la infraestructura para información espacial (INSPIRE) de la UE.

# Definiciones

## Lenguajes de Metadatos

**La infraestructura para información espacial (INSPIRE) de la UE se basa en la ISO - 19115, cuya última versión data de 2014.**

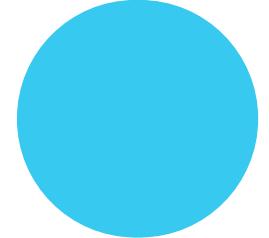
**Dublin Core se publicó en 2009 como estándar ISO - 15836 y es uno de los estándares de metadatos más conocidos y usados.**

**DCAT es un vocabulario RDF diseñado para facilitar la interoperabilidad entre catálogos de datos publicados en la Web . Se describe en clases y propiedades.**

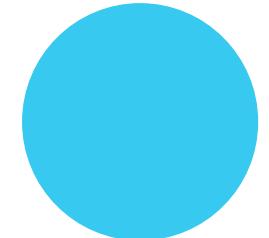
# **Principios**



# Open Data Charter



**Es una colaboración entre más de 170 gobiernos y organizaciones que trabajan para abrir los datos basándose en un conjunto de principios compartidos.**

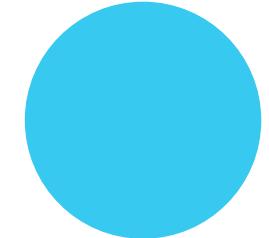


**Promueven políticas y prácticas que permiten a los gobiernos y a las organizaciones de la sociedad civil recopilar, compartir y utilizar datos bien gobernados, para responder de manera efectiva y responsable a las siguientes áreas de trabajo: anticorrupción, acción climática, género, inteligencia artificial y derechos de datos.**

# Principios globales según Open Data Charter

- Abiertos por defecto
- Oportunos y exhaustivos
- Accesibles y utilizables
- Comparables e interoperables
- Para mejorar la gobernanza y participación ciudadana.
- Para el desarrollo inclusivo y la innovación.

# Principios Open Data Government

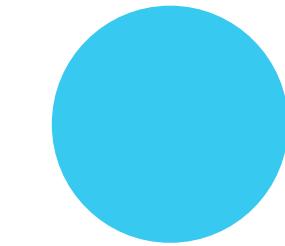


La iniciativa Open Government Data, definió 8 características que deben cumplir los datos “públicos” para ser “abiertos”

- Completos
- Primarios
- Oportunos
- Accesibles
- Procesables
- No discriminatorios
- No propietarios
- Licencia de uso libre

# Principios Open Data Government

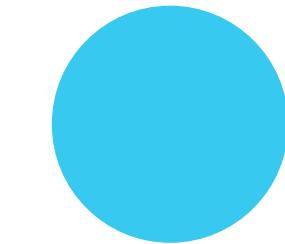
Completos



**Están disponibles y no tienen  
restricciones para su apertura**

# Principios Open Data Government

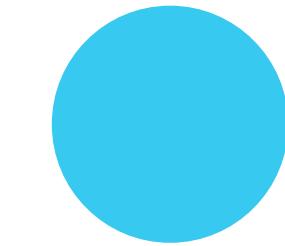
## Primarios



Son recolectados en la fuente de origen, con el nivel de especificidad más alto posible, no en forma agregada ni modificada

# Principios Open Data Government

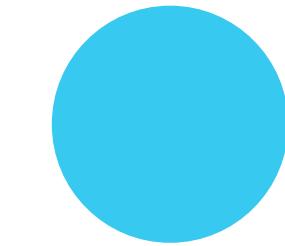
Oportunos



Están disponibles tan rápido como sea necesario para preservar su valor

# Principios Open Data Government

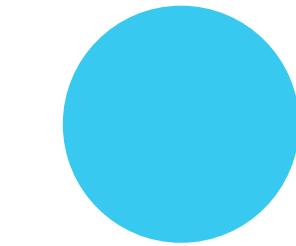
Accesibles



Están disponibles para el rango más amplio de usuarios y propósitos

# Principios Open Data Government

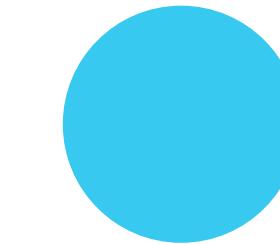
Procesables por  
máquinas



Están estructurados de forma tal que  
permitan un procesamiento  
automático

# Principios Open Data Government

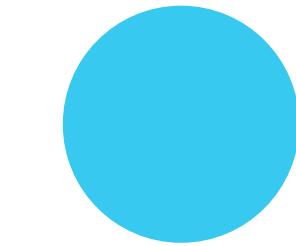
No discriminatorios



Están disponibles para cualquier ciudadano, sin requerir un registro

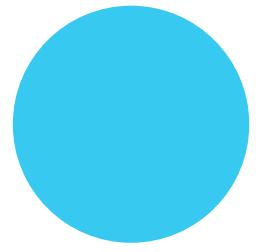
# Principios Open Data Government

No propietarios

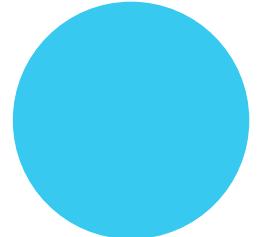


Están disponibles en un formato  
sobre el cual ninguna empresa de  
software tiene un control exclusivo

# Principios Open Data Government Licencias libres



**Los datos abiertos deben tener una licencia que especifique que se trata de datos abiertos y que defina bajo qué condiciones estos datos pueden ser reutilizados.**



**La licencia podría, además, detallar que las personas que usen los datos deben atribuirlos y que, en caso de que combinen los datos con otros datos, deben compartir esta información también como datos abiertos.**

# **Niveles de apertura de Datos**



# Niveles de apertura de datos

- Primer nivel: formato de publicación de datos que no permite ser reutilizado.
- Segundo nivel: se permite operar sobre los datos para agregar valor. Software propietario.



# Niveles de apertura de datos

● Tercer nivel: publicación de datos en formatos estructurados. Software libre. Es el estadio mínimo de apertura de datos.



● Cuarto y Quinto nivel: estándares de mayor calidad y homogeneidad en publicaciones.



# Niveles de apertura de datos

TEXTO		TABULARES	
Formato	Nivel de apertura	Formato	Nivel de apertura
PDF	Bajo	XLS	Bajo
DOC	Bajo	XLSX	Medio
DOCX	Medio	ODS	Medio
ODT	Medio	CSV	Alto
TXT	Alto		

GEOGRÁFICOS		ESTRUCTURADOS	
Formato	Nivel de apertura	Formato	Nivel de apertura
SHP	Medio	XML	Medio
KML	Alto	JSON	Alto
GEOJSON	Alto	HTML	Alto
GEOPACKAGE	Alto		

# **Beneficios**



# Beneficios

- Mejorar el proceso de toma de decisiones basadas en evidencia empírica e interoperabilidad.
- Combatir la corrupción y fakenews.
- Posibilitar el monitoreo, transparencia y la rendición de cuentas ciudadana.
- Crear y optimizar servicios, aplicaciones y proyectos de gran valor.
- Innovar, buscar soluciones y productos.

# Beneficios

¿Quiénes se benefician o aprovechan de los datos abiertos?:

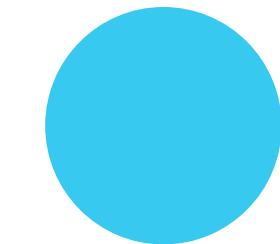
- Organizaciones de la sociedad civil.
- Sector Privado
- Sector académico
- Ciudadanía
- Periodismo
- Sector Público

# **Licencias de Datos Abiertos**

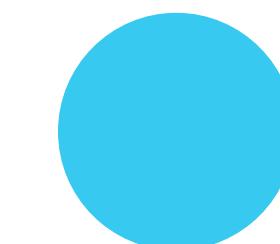


# Licencias de datos abiertos

## Definiciones



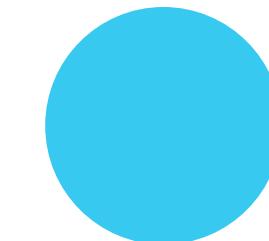
**Los datos abiertos necesitan estar abiertos tanto técnica como legalmente. Para hacerlo, tenemos que aplicarles una licencia.**



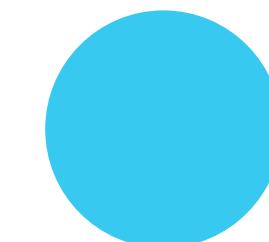
**Las licencias permiten que las personas sepan, sin ambigüedades, los usos a los que se pueden someter los datos en caso de reutilización y las condiciones que deberán tener en cuenta en materia de propiedad intelectual.**

# Licencias de datos abiertos

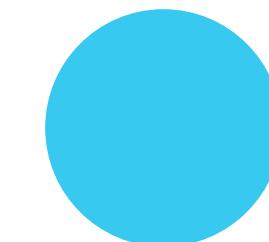
## Tipos



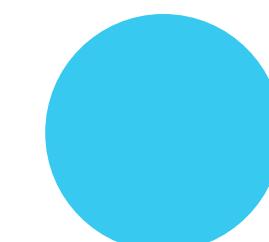
**Open Data Commons Public Domain Dedication and License o PDDL**



**Open Data Commons Attribution License**



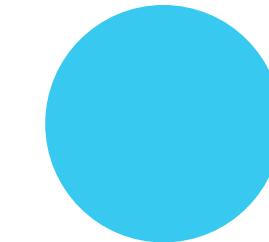
**Open Data Commons Open Database License (ODbL)**



**Creative Commons Atribución 4.0 Internacional**

# Licencias de datos abiertos

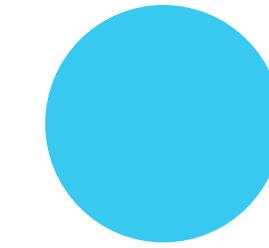
Open Data Commons  
Public Domain  
Dedication and  
License o PDDL



**Permite difundir, reutilizar o adaptar los datos sin restricción alguna.**

# Licencias de datos abiertos

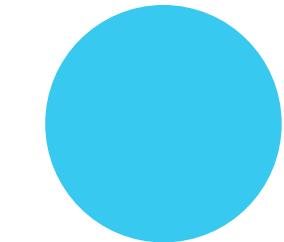
Open Data Commons  
Attribution License



**Exige la referencia a la autoría o fuente de los datos para la reutilización de la información.**

# Licencias de datos abiertos

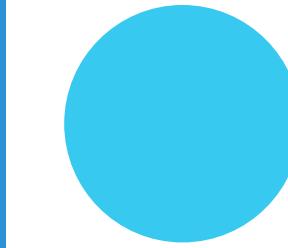
Open Data Commons  
Open Database  
License (ODbL)



**Permite la reutilización de los datos siempre que se reconozca la autoría de la información original y se mantenga la misma licencia en las obras derivadas, de las cuales se puede restringir su uso si, además, se distribuye una versión sin dichas restricciones.**

# Licencias de datos abiertos

Creative Commons  
Atribución  
4.0 Internacional



**Esta licencia posibilita no sólo copiar y redistribuir el material en cualquier medio o formato, sino también adaptar y transformarlo para cualquier propósito, incluso comercial.**

**Condiciones:**

- Atribución
- No comercial
- Sin obras derivadas
- Compartir igual.

# Licencias de datos abiertos

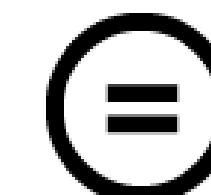
Creative Commons  
Atribución  
4.0 Internacional



**Atribución (Attribution):** En cualquier explotación de la obra autorizada por la licencia será necesario reconocer la autoría (obligatoria en todos los casos).



**No Comercial (Non commercial):** La explotación de la obra queda limitada a usos no comerciales.



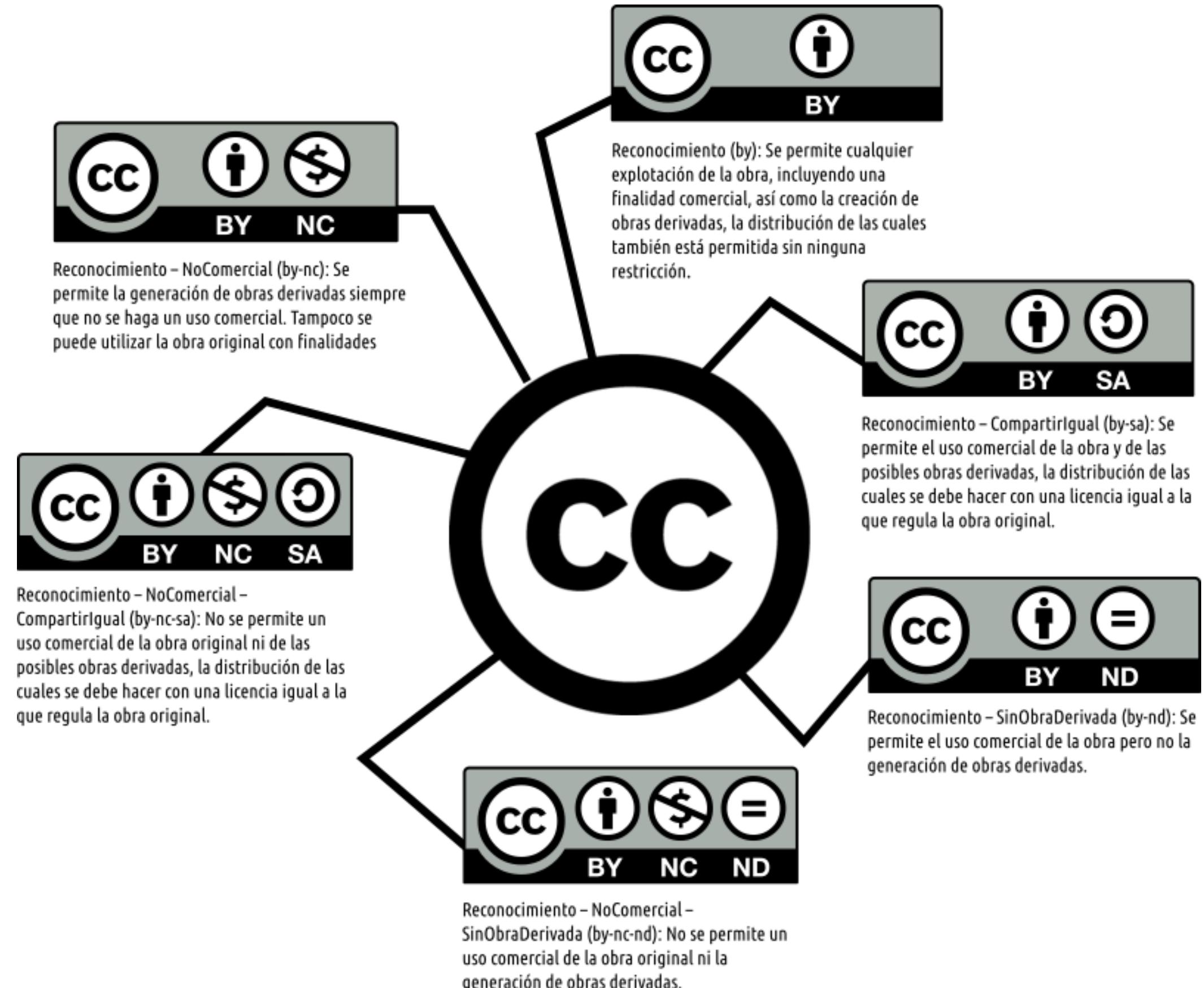
**Sin obras derivadas (No Derivate Works):** La autorización para explotar la obra no incluye la posibilidad de crear una obra derivada.



**Compartir Igual (Share alike):** La explotación autorizada incluye la creación de obras derivadas siempre que mantengan la misma licencia al ser divulgadas.

# Licencias de datos abiertos

Creative Commons  
Atribución  
4.0 Internacional

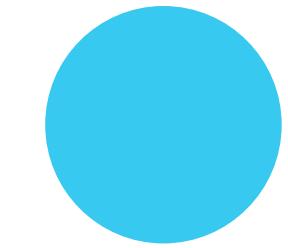


# **Tecnologías para Datos Abiertos**



# Tecnologías para datos abiertos

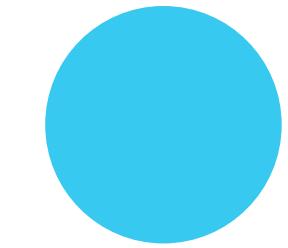
## Catálogos de datos



Es una lista de conjuntos de datos disponibles en una iniciativa de datos abiertos. Los elementos esenciales son: metadatos con información clara en cuanto a la licencia legal de uso y el acceso a las bases de datos propiamente dichas.

# Tecnologías para datos abiertos

## Plataforma



Constituye “una puerta de acceso” en línea para que los usuarios tengan disponibles todos los recursos de una iniciativa de datos abiertos. Incluye al catálogo de datos, otros servicios que forman parte del ecosistema de datos abiertos (IDE, APIs, etc) y un buscador que permite interactuar con ellos.

# Tecnologías para datos abiertos



 Andino

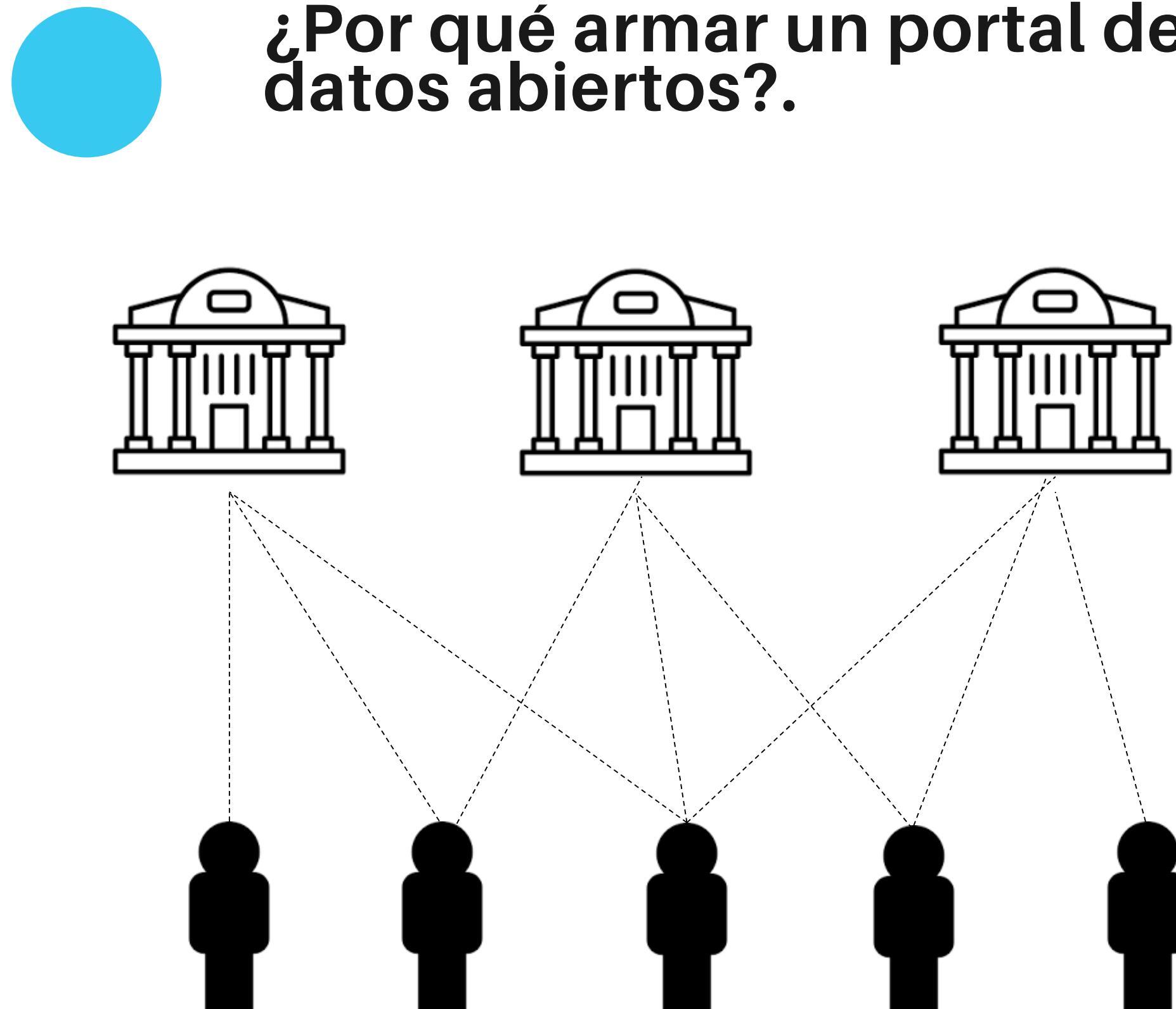
The Andino logo icon is a dark blue graphic element consisting of three interconnected circles or nodes.

# **Portal de Datos Abiertos**



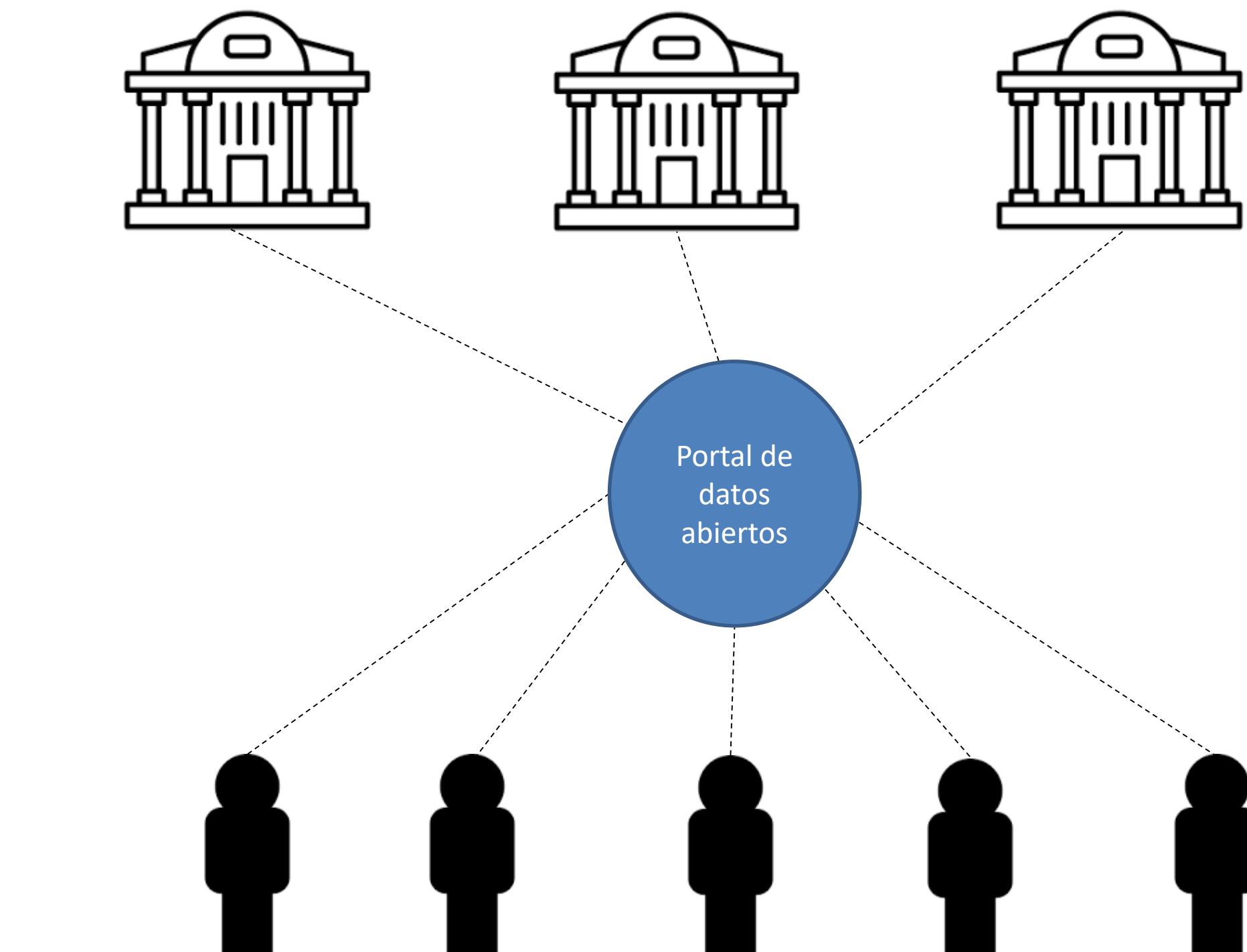
# Portal de datos abiertos

¿Por qué armar un portal de  
datos abiertos?.



# Portal de datos abiertos

¿Por qué armar un portal de  
datos abiertos?.



# Portales

Argentina

[https://datos.gob.ar.](https://datos.gob.ar)

**datos.gob.ar**

Datasets Series Organizaciones APIs Acerca ▾

## Datos Argentina

Ponemos a tu alcance datos públicos en formatos abiertos para que puedas usarlos, modificarlos y compartirlos. Estos datos son tuyos. Podés crear visualizaciones, aplicaciones y grandes herramientas con ellos.

1245  
DATASETS

037  
ORGANIZACIONES  
CON DATOS

¿Qué dataset buscás? 



Agropecuario, pesca y  
forestación



Asuntos internacionales



Ciencia y tecnología



Economía y finanzas



Educación, cultura y deportes



Energía



Gobierno y sector público



Justicia, seguridad y legales



Medio ambiente



Población y sociedad



Regiones y ciudades



Salud

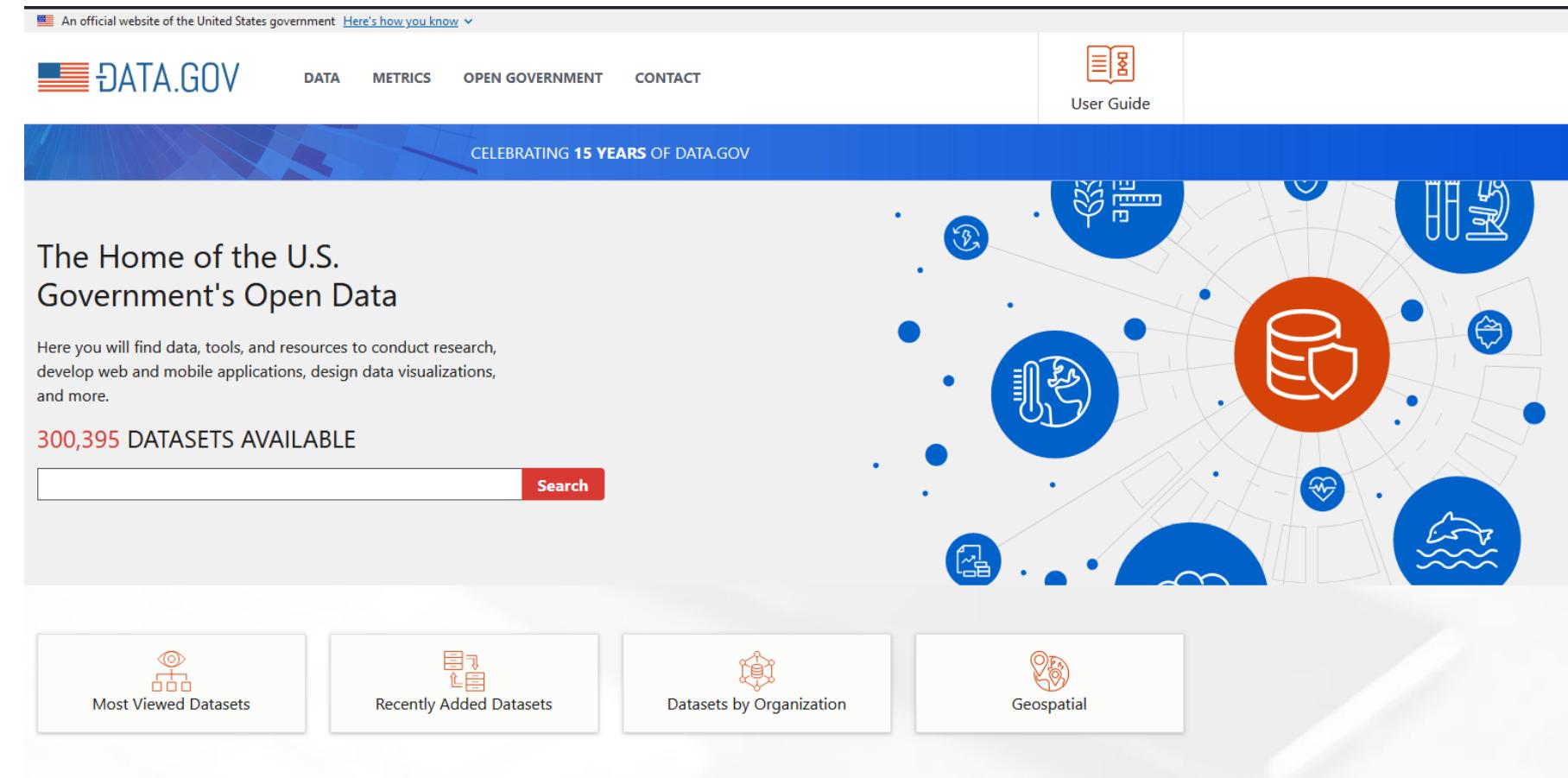


Transporte

# Portales

## Estados Unidos

[https://data.gov/.](https://data.gov/)



# Comunidad Europea

# Portales

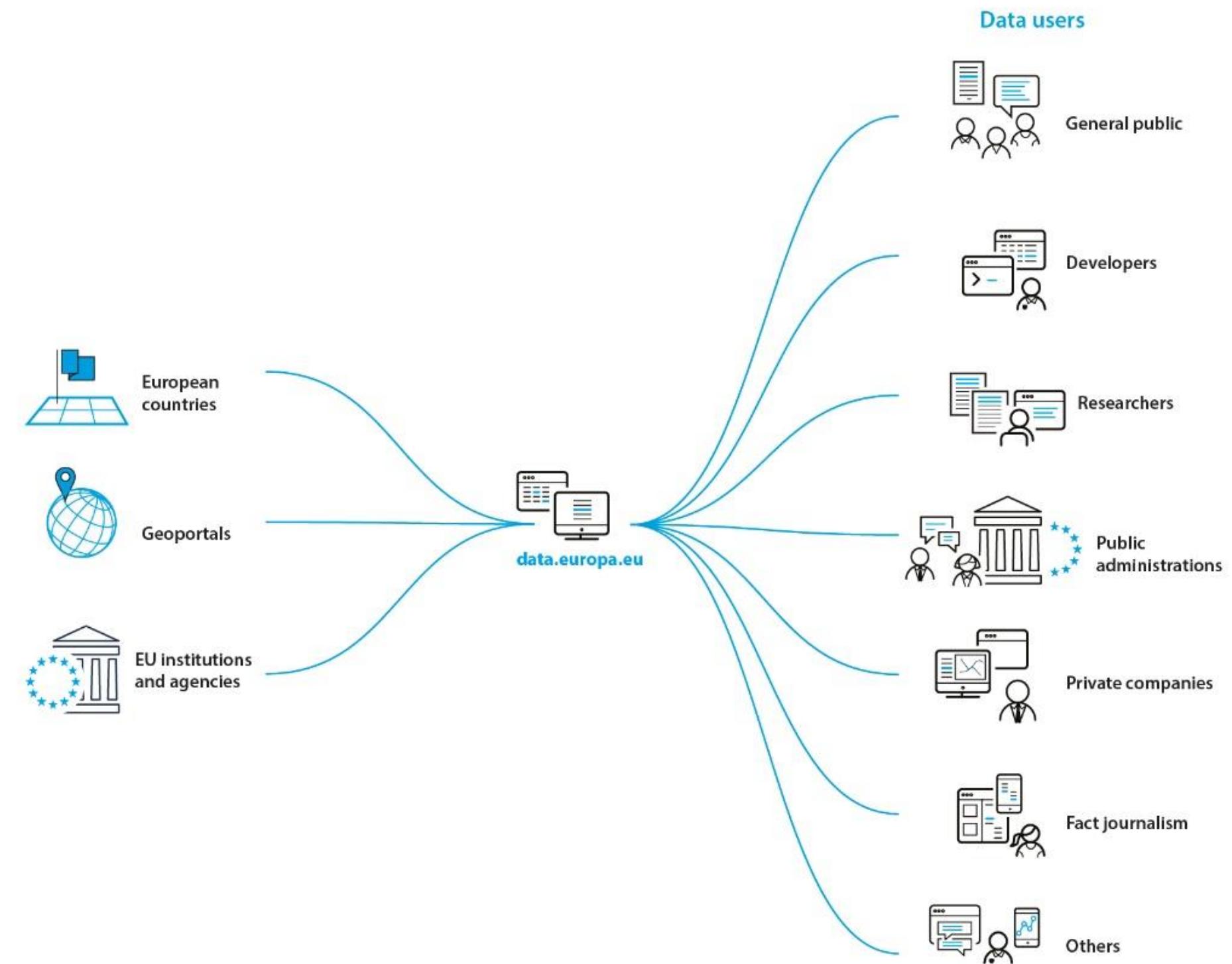
[https://data.europa.eu/en.](https://data.europa.eu/en)

The screenshot shows the homepage of data.europa.eu. At the top, there is a navigation bar with the European Union flag, 'Log in', 'English', and a search bar. Below the navigation bar, the text 'European data' and 'data.europa.eu - The official portal for European data' is displayed. A blue banner at the bottom features the text '10th anniversary of Open Data Maturity Report' and 'Celebrating a decade of open data developments in Europe'. A yellow button labeled 'Find out more >' is visible. At the bottom of the page, there is a footer with statistics: '1 757 010 Datasets', '189 Catalogues', and '35 Countries'.

# Portales

# Comunidad Europea

https://data.europa.eu/en.



# **Marco Normativo**



# Argentina

2016 ----- 2016 ----->

**Decreto 117/16**

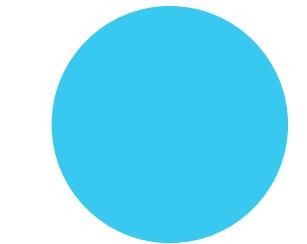
Plan Nacional de Apretura de Datos

**Ley 27275**

Derecho de Acceso a la Información  
Pública

# Marco normativo

Argentina



Estos son los decretos y resoluciones que enmarcan la política de apertura de datos de la Administración Pública Nacional.

● Plan Nacional de Apertura de Datos: El decreto 117/2016 reglamenta plazos tanto para la publicación de los primeros conjuntos de datos de nuestro portal. También reglamenta los tiempos para detallar los activos de datos de los distintos organismos de la Nación y su cronograma de publicación.

● Régimen de Acceso a la Información Pública: Ley 27275 de Derecho de Acceso a la Información Pública.

# Marco normativo

Argentina

Límites legítimos de la apertura de datos

Protección de seguridad, privacidad y propiedad.

- Artículo 43 Habeas Data de la Constitución Nacional.
- Ley 25326 Protección de datos personales.
- Ley 11723 Propiedad Intelectual
- Ley 27275 Acceso a la información pública.
- Ley 17622. Secreto Estadístico
- Ley 11.683. Procedimiento Fiscal

# Unión Europea



# Marco normativo

## Comunidad Europea

- 2003 - Directiva 2003/98/CE relativa a la reutilización de la información del sector público.
- 2011 - La Decisión de la Comisión 2011/833/UE sobre la reutilización de los documentos de la Comisión estableció las bases para el antiguo Portal de Datos Abiertos de la UE, gestionado por la Oficina de Publicaciones de la UE (2012-2021).
- 2013 - Directiva 2013/37/UE por la que se modifica la Directiva 2003/98/CE relativa a la reutilización de la información del sector público.

# Comunidad Europea

# Marco normativo

- 2019 - Directiva 2019/1024 sobre datos abiertos y reutilización de la información del sector público refunda las Directivas 2003/98/CE (Directiva ISP) y 2013/37/UE por las que se establece un marco jurídico de requisitos mínimos para los Estados miembros en materia de accesibilidad de los recursos de datos públicos.
- 2020 - Comunicación (COM/2020/66) sobre una estrategia europea de datos destinada a crear un mercado único de datos en beneficio de las empresas, los investigadores y las administraciones públicas para garantizar la competitividad global de Europa

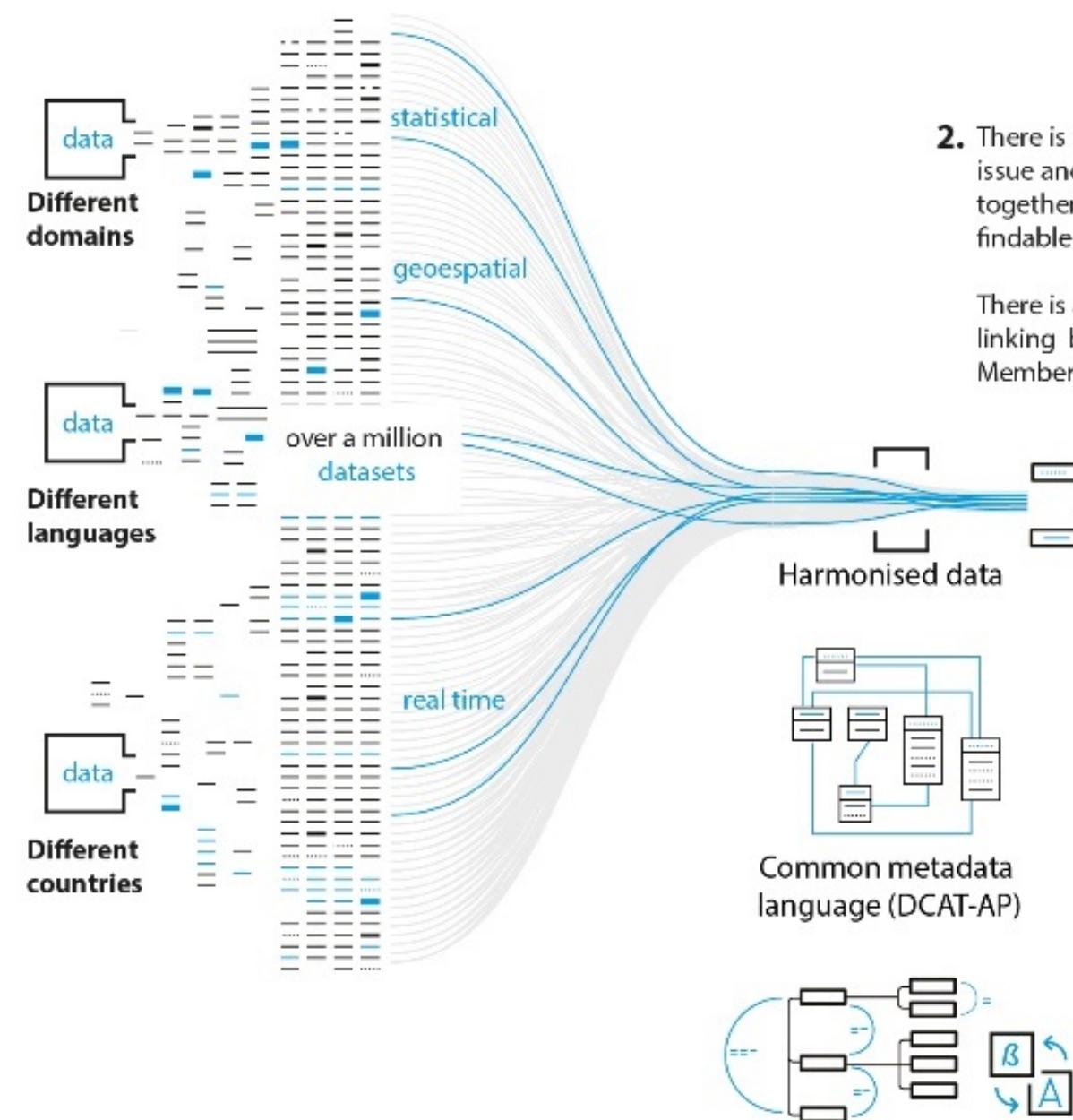
# Marco normativo

## Comunidad Europea

- 2022 - La ley de Gobernanza de Datos establece un marco para fomentar el intercambio voluntario de datos más allá de los datos abiertos.
- 2023 - Entrada en vigencia de la regulación de implementación de la Comisión sobre conjuntos de datos de alto valor, aquellos que tienen un alto potencial de impacto económico y social.

## Unique features of data.europa.eu

- 1.** The **challenge** is to present a heterogeneous and large amount of data coming from different countries and domains in a way that makes sense.

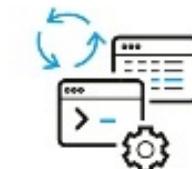


- 2.** There is still a big data comparability issue and a need to link the data together to make them more findable and understandable.

There is also a need to improve linking between EU data and Member States data.



- 3.** We provide access to over a million datasets, which requires full automation, hindering the checking of quality at a sufficient level of detail.

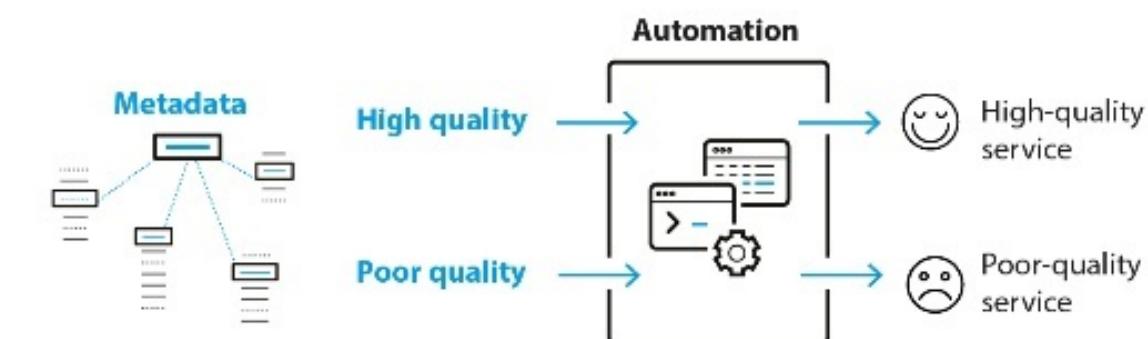
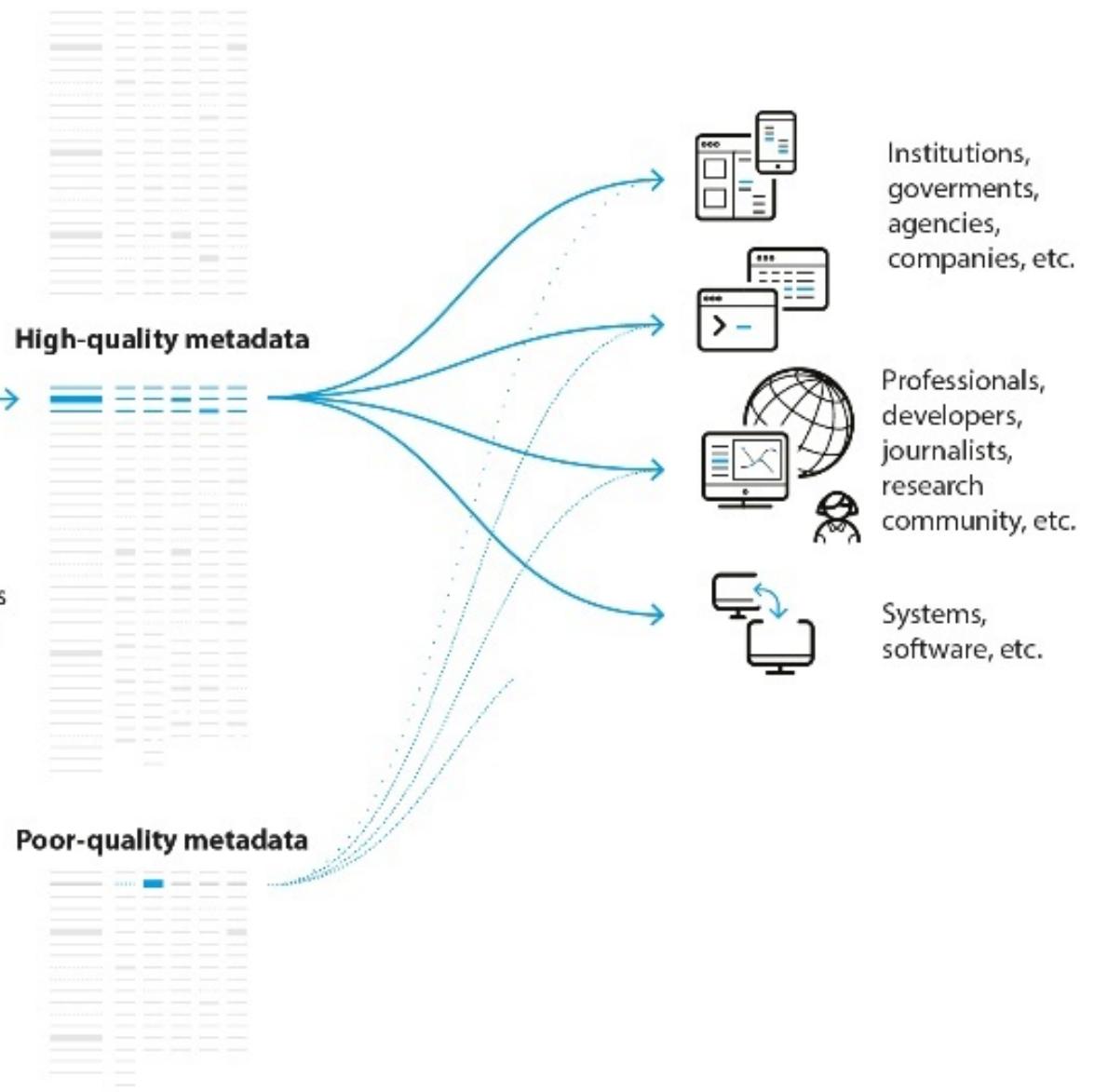


## Multilingual EU reference data

Data become harmonised by being categorised into domains via data themes. They are described with the same metadata language (DCAT-AP), harmonised and linked using multilingual EU reference data.

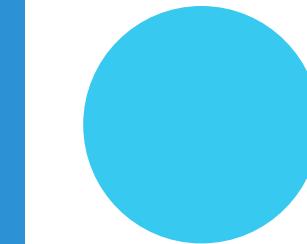
## From quantity to quality

Current priorities are to improve the **quality of metadata**, to allow for interoperability, and the relevance and quality of the data.



# Marco normativo

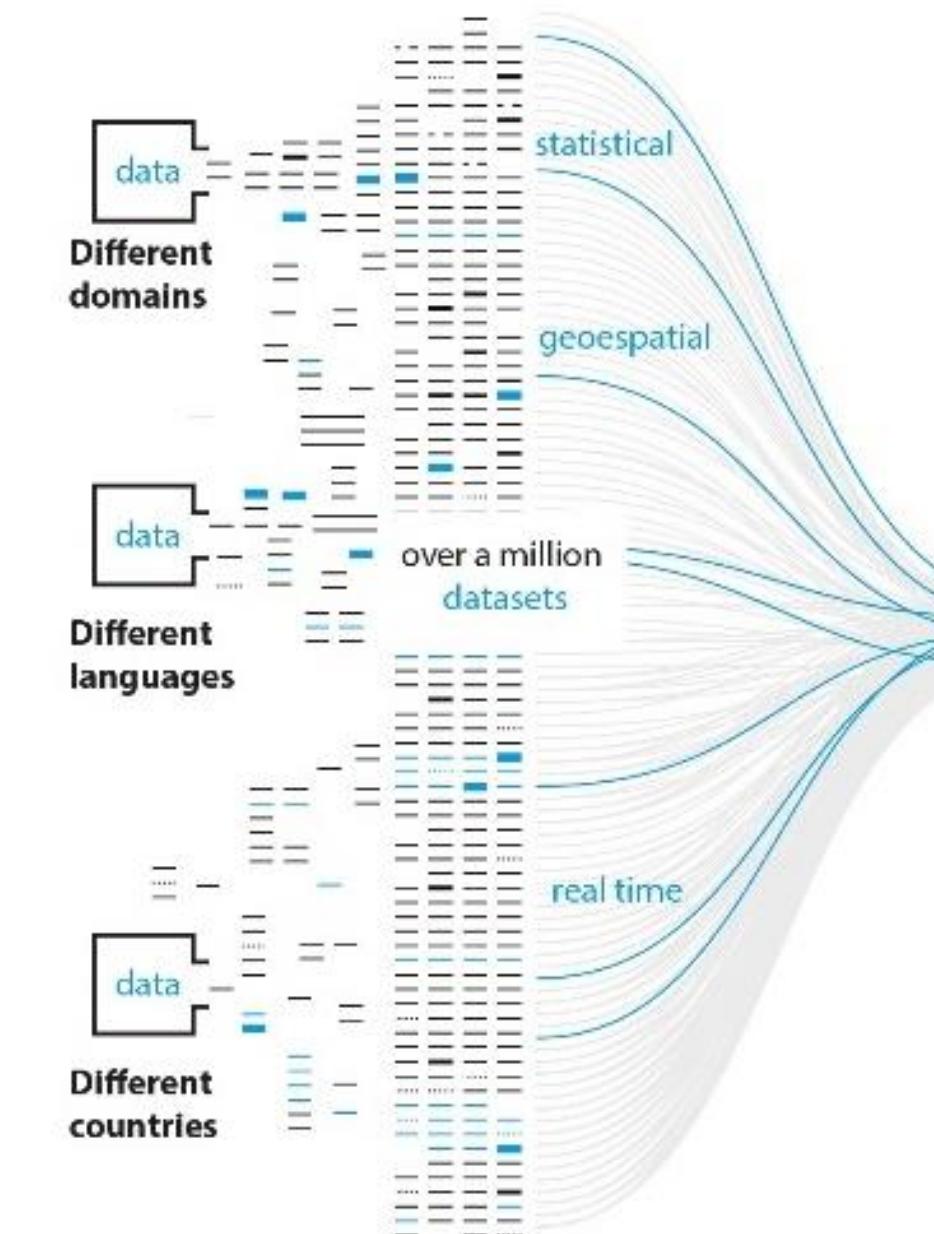
Comunidad Europea



## ¿Qué tipos de datos y en qué formato se producen en la Unión Europea?.

### Unique features of data.europa.eu

1. The **challenge** is to present a heterogeneous and large amount of data coming from different countries and domains in a way that makes sense.



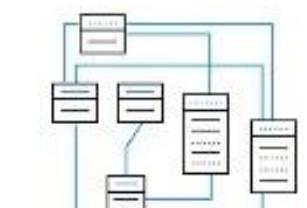
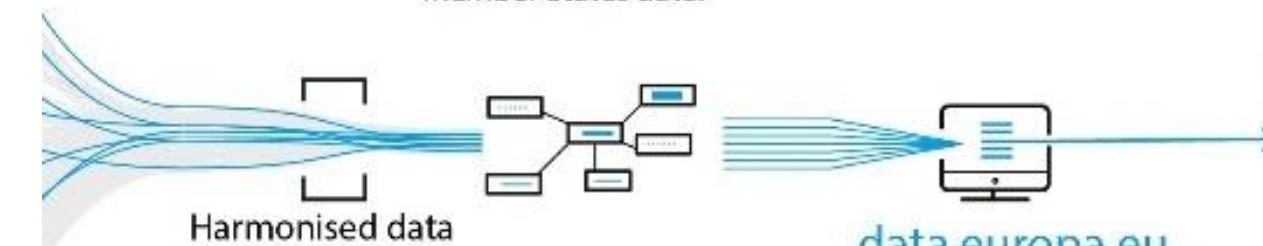
# Marco normativo

Comunidad Europea

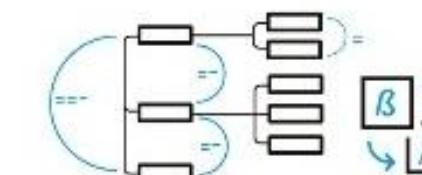
## ¿Cómo compatibilizamos los datos?

2. There is still a big data comparability issue and a need to link the data together to make them more findable and understandable.

There is also a need to improve linking between EU data and Member States data.

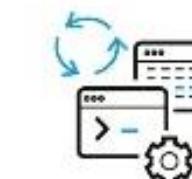


Common metadata language (DCAT-AP)



Multilingual EU reference data

Data become harmonised by being categorised into domains via data themes. They are described with the same metadata language (DCAT-AP), harmonised and linked using multilingual EU reference data.



3. We provide access to over a million datasets, which requires full automation, hindering the checking of quality at a sufficient level of detail.

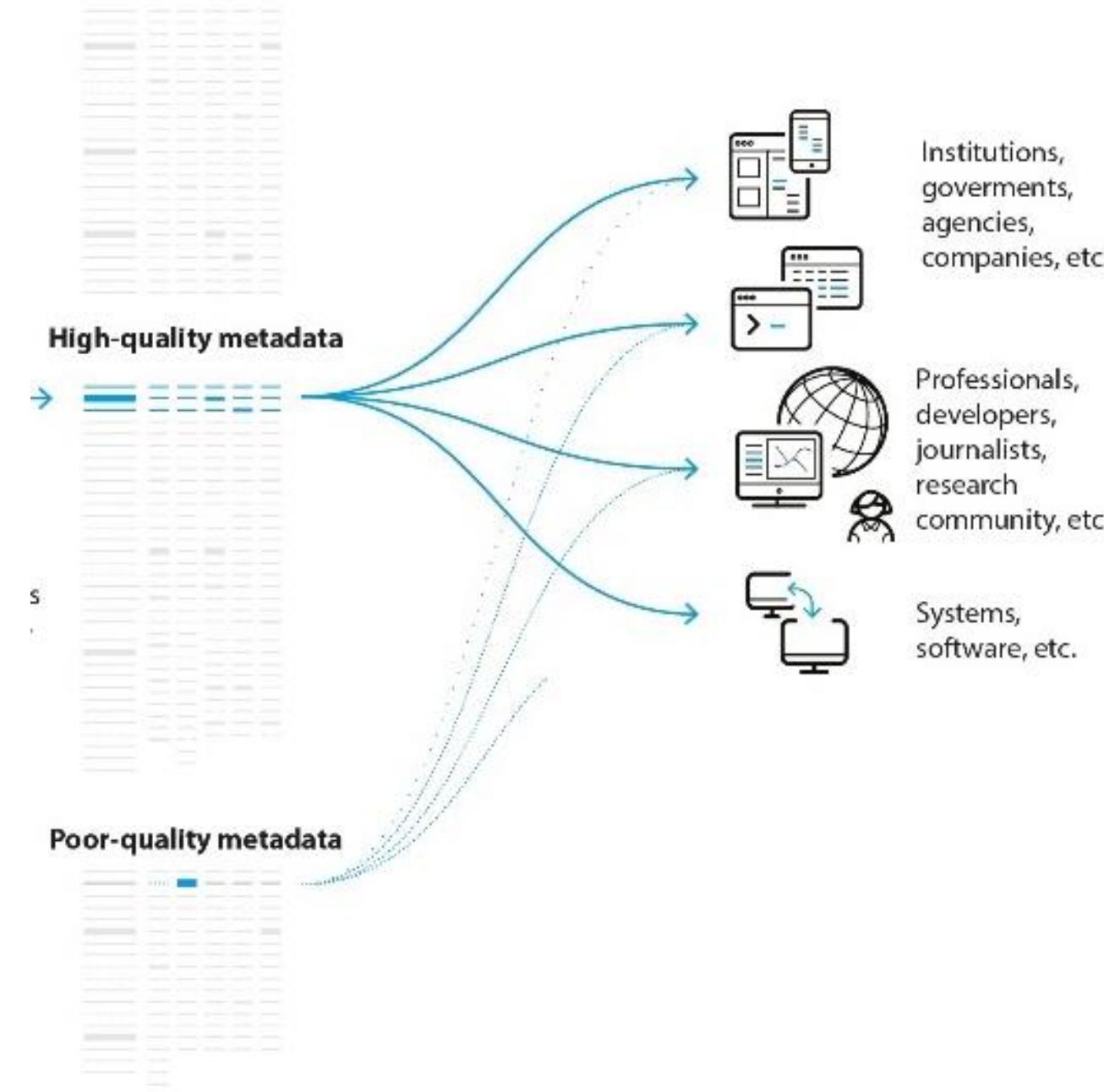
# Comunidad Europea

# Marco normativo

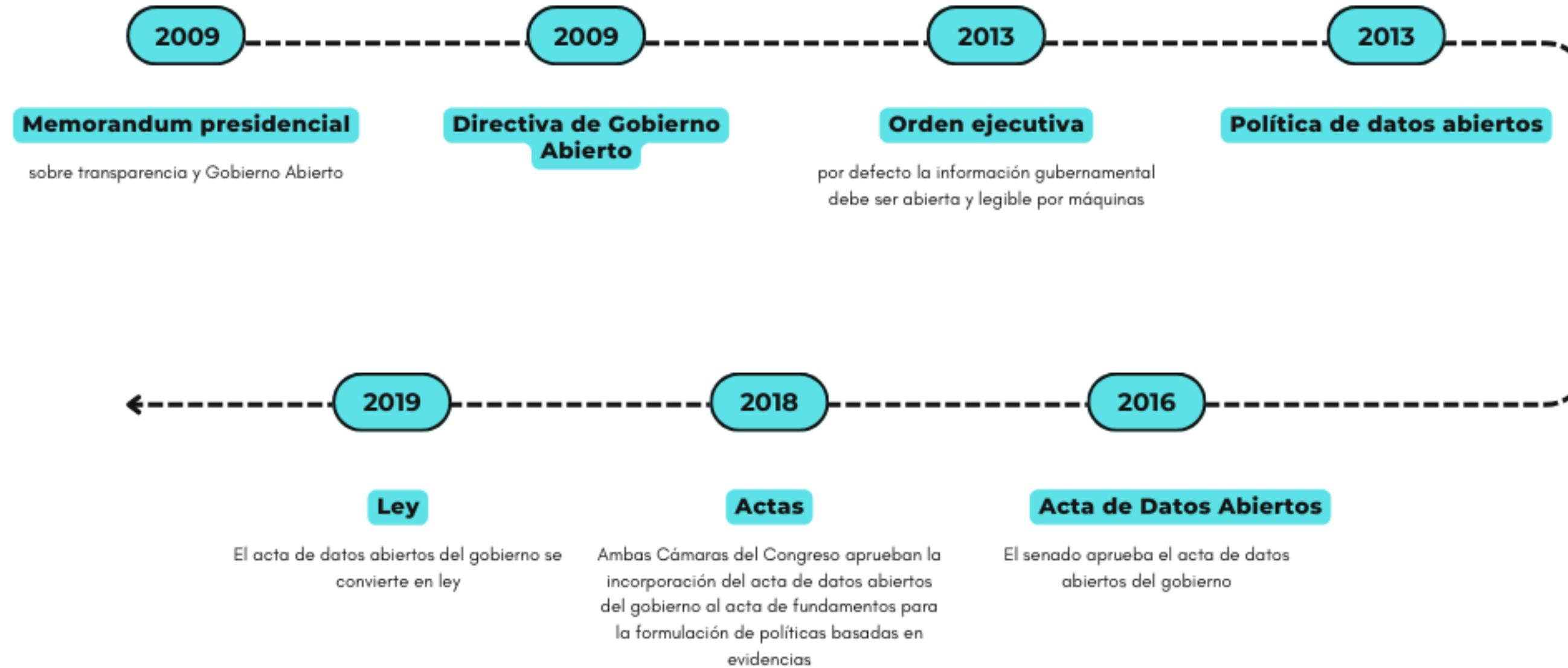
## ¿Cómo se utilizan los datos?

### From quantity to quality

Current priorities are to improve the **quality of metadata**, to allow for interoperability, and the relevance and quality of the data.



# Estados Unidos



# Marco normativo

Estados Unidos

- 2009 - Memorandum presidencial sobre transparencia y Gobierno Abierto.
- 2009 - Directiva de Gobierno Abierto.
- 2013 - Orden ejecutiva: por defecto la información gubernamental debe ser abierta y legible por máquinas.
- 2013 - Política de datos abiertos.
- 2016 - El senado aprueba el acta de datos abiertos del gobierno.

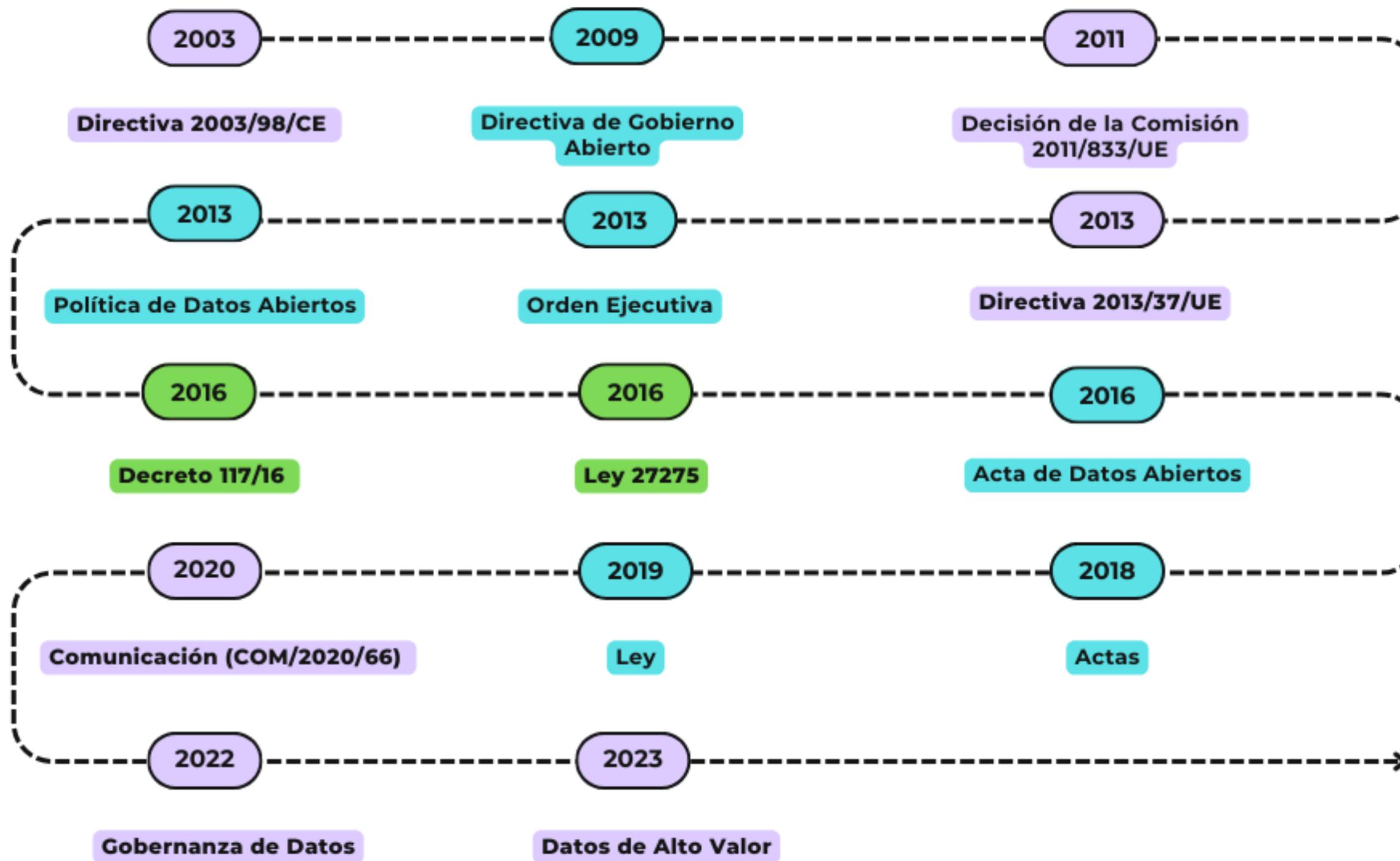
# Marco normativo

## Estados Unidos

- 2018 - Ambas Cámaras del Congreso aprueban la incorporación del acta de datos abiertos del gobierno al acta de fundamentos para la formulación de políticas basadas en evidencias.
- 2019 - El acta de datos abiertos del gobierno se convierte en ley.

# Global

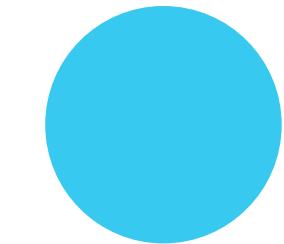
Unión Europea  
Estados Unidos  
Argentina



# **Infraestructura Nacional de Datos Públicos**



# Infraestructura Nacional de Datos Públicos (INDAP)



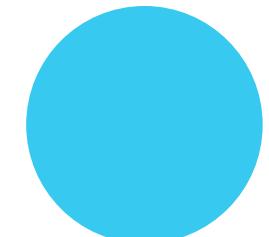
Tiene como objetivo principal nuclear los distintos mecanismos referidos al diseño, planificación, ejecución y monitoreo de la estrategia nacional de apertura de datos del Sector Público Nacional y a todos aquellos actores involucrados en su implementación; como así también generar espacios de diálogo entre éstos, la sociedad civil y el sector académico.

La implementación, administración y supervisión del programa estará en la órbita de la Dirección Nacional de Gobierno Abierto, siendo la Dirección de Datos Abiertos la autoridad de aplicación y alcanzando a todo el Sector Público Nacional.

# **Hacia donde vamos**



# Hacia donde vamos



Actualmente en la comunidad de datos abiertos se esta discutiendo el reemplazo del actual estándar en que se comparten los datos de forma estructurada el CSV por formatos que permitan compartir datos en crudo, como ser imágenes, audios, videos, etc.

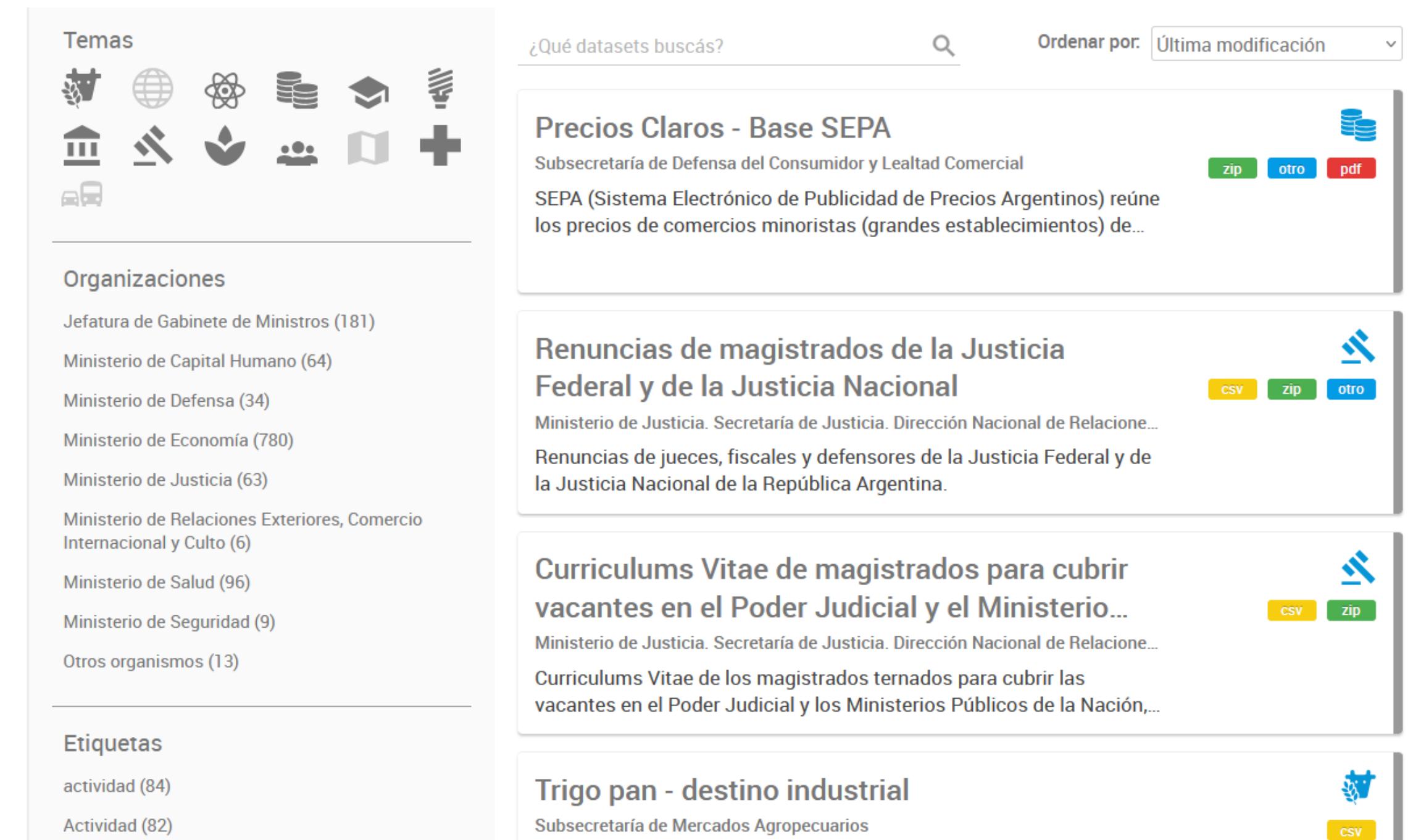
Uno de las alternativas que estaría teniendo buena aceptación es la apificación.

# **Dataset**



# Dataset

Veamos un dataset de  
<https://datos.gob.ar/dataset>



The screenshot shows the search interface for datasets on the datos.gob.ar platform. At the top, there's a search bar with placeholder text "¿Qué datasets buscás?", a magnifying glass icon, and a dropdown menu set to "Ordenar por: Última modificación". To the left, there are sections for "Temas" (Topics) with icons for agriculture, global network, science, finance, education, energy, government, infrastructure, environment, people, maps, and a plus sign; "Organizaciones" (Organizations) listing various government ministries and agencies with their counts; and "Etiquetas" (Tags) for "actividad" and "Actividad", each with a count of 84 and 82 respectively. The main area displays three dataset cards:

- Precios Claros - Base SEPA**  
Subsecretaría de Defensa del Consumidor y Lealtad Comercial  
SEPA (Sistema Electrónico de Publicidad de Precios Argentinos) reúne los precios de comercios minoristas (grandes establecimientos) de...  
[zip](#) [otro](#) [pdf](#)
- Renuncias de magistrados de la Justicia Federal y de la Justicia Nacional**  
Ministerio de Justicia. Secretaría de Justicia. Dirección Nacional de Relacione...  
Renuncias de jueces, fiscales y defensores de la Justicia Federal y de la Justicia Nacional de la República Argentina.  
[csv](#) [zip](#) [otro](#)
- Curriculums Vitae de magistrados para cubrir vacantes en el Poder Judicial y el Ministerio...**  
Ministerio de Justicia. Secretaría de Justicia. Dirección Nacional de Relacione...  
Curriculum Vitae de los magistrados ternados para cubrir las vacantes en el Poder Judicial y los Ministerios Públicos de la Nación,...  
[csv](#) [zip](#)

Each dataset card includes a small thumbnail icon and download links for CSV, ZIP, and other formats.

# Dataset

Veamos un dataset de  
<https://datos.gob.ar/dataset>

Datos Argentina / Dataset

**Maíz - Producción para alcohol etílico**  
Dirección de Bioenergía, Secretaría de Alimentos y Bioeconomía 

Producción estimada de maíz destinada a la elaboración de alcohol etílico para utilizar como biocombustible (bioetanol) y otros. Estos datos, correspondientes al año 2018, son provisорios y están expresados en toneladas.

**Recursos del dataset**

 Producción de maíz destinada a la elaboración de alcohol etílico  
Datos estimados de la producción de maíz destinada a la elaboración de alcohol etílico para utilizar como biocombustible (bioetanol) y otros.... [CONSULTAR](#)  
[DESCARGAR](#)

 Serie de tiempo  
Evolución de la producción estimada de maíz destinada a la elaboración de alcohol etílico para utilizar como biocombustible (bioetanol) y otros.... [CONSULTAR](#)  
[DESCARGAR](#)

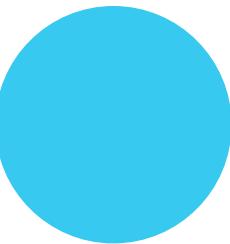
**Información adicional**

Temas	   
Etiquetas	alcohol biocombustibles bioeconomía bioenergía bioetanol
	<a href="#">Mostrar más</a>
Licencia	Creative Commons Attribution 4.0
Frecuencia de actualización	Mensualmente
Mantenedor	Dirección de Bioenergía 
Fecha de publicación	7 de diciembre de 2018
Fecha de actualización	24 de octubre de 2019
Página de referencia	<a href="https://datos.agroindustri...">https://datos.agroindustri...</a>

**Actividad reciente**  
Se actualizó el dataset [Maíz - Producción para alcohol etílico](#). Hace 2 días.

# Dataset

Veamos un dataset de  
<https://datos.gob.ar/dataset>



Datos Argentina / Dataset / Recurso

## Producción de maíz destinada a la elaboración de alcohol etílico

Datos estimados de la producción de maíz destinada a la elaboración de alcohol etílico para utilizar como biocombustible (bioetanol) y otros. Datos provisorios, expresados en toneladas.

[DESCARGAR](#)

[CONSULTAR EN EL  
PORTAL DE ORIGEN](#)

### Campos de este recurso

Título de la columna	Tipo de dato	Descripción
cod_pais	Texto (string)	cod_pais
nom_pais	Texto (string)	nom_pais
año	Texto (string)	año
mes	Texto (string)	mes
insumo	Texto (string)	insumo
cod_uni_med	Texto (string)	cod_uni_med
nom_uni_med	Texto (string)	nom_uni_med
destino_bioetanol		destino_bioetanol
destino_otros		destino_otros
total		total

# Dataset

Veamos un dataset de  
<https://datos.gob.ar/dataset>

## Información adicional

Responsable	Dirección de Bioenergía, Secretaría de Alimentos y Bioeconomía
Formato	<a href="#">CSV</a>
Temas	Agropecuario, pesca y forestación Economía y finanzas Energía Medio ambiente
Frecuencia de actualización	Mensualmente
Fecha de creación	16 de julio de 2018
Último cambio	27 de agosto de 2019

## Otros recursos de este dataset

 csv	Serie de tiempo Evolución de la producción estimada de maíz destinada a la elaboración de alcohol etílico para utilizar como biocombustible (bioetanol) y otros....	<a href="#">CONSULTAR</a>	<a href="#">DESCARGAR</a>
---	--	---------------------------	---------------------------

# Dataset

Veamos un dataset de  
<https://datos.gob.ar/dataset>

```
1 cod_pais,nom_pais,año,mes,insumo,cod_uni_med,nom_uni_med,destino_bioetanol,destino_otros,total
2 32,Argentina,2018,Enero,Maiz,t,Toneladas,127018,21016,148033
3 32,Argentina,2018,Febrero,Maiz,t,Toneladas,120435,16379,136814
4 32,Argentina,2018,Marzo,Maiz,t,Toneladas,122314,21572,143885
5 32,Argentina,2018,Abril,Maiz,t,Toneladas,136903,21795,158698
6 32,Argentina,2018,Mayo,Maiz,t,Toneladas,130278,20467,150745
7 32,Argentina,2018,Junio,Maiz,t,Toneladas,129047,19831,148878
8 32,Argentina,2018,Julio,Maiz,t,Toneladas,119416,18102,137518
9 32,Argentina,2018,Agosto,Maiz,t,Toneladas,123428,25861,149289
10 32,Argentina,2018,Septiembre,Maiz,t,Toneladas,121535,19187,140722
11 32,Argentina,2018,Octubre,Maiz,t,Toneladas,128307,18153,146460
12 32,Argentina,2018,Noviembre,Maiz,t,Toneladas,126599,17963,144562
13 32,Argentina,2018,Diciembre,Maiz,t,Toneladas,128486,18708,147194
14 32,Argentina,2019,Enero,Maiz,t,Toneladas,116238,19991,136229
15 32,Argentina,2019,Febrero,Maiz,t,Toneladas,112369,18731,131100
16 32,Argentina,2019,Marzo,Maiz,t,Toneladas,116267,21687,137954
17 32,Argentina,2019,Abril,Maiz,t,Toneladas,134030,27699,161729
18 32,Argentina,2019,Mayo,Maiz,t,Toneladas,126180,21838,148018
19 32,Argentina,2019,Junio,Maiz,t,Toneladas,122734,22854,145588
20
```

# wikidata



# Wikidata



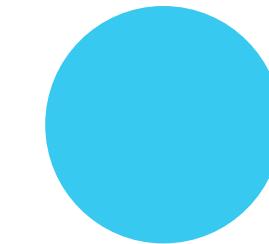
**Es una base de conocimiento secundaria, libre, colaborativa, multilingüe y abierta que cualquiera puede leer y editar.**

**Actúa como un repositorio para de datos estructurados que es utilizado en los proyectos de Wikimedia, como puede ser wikipedia, aunque también puede brinda soporte a otros sitios.**

**Su contenido esta disponible bajo una licencia libre.**

# Wikidata

Libre

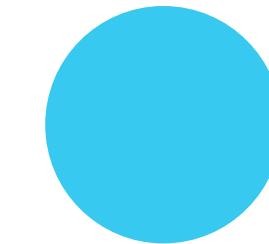


**Los datos en Wikidata están publicados bajo la licencia Creative Commons Dedicación al Dominio Público 1.0, lo que permite la reutilización de los datos en muchos escenarios diferentes.**

**Puede copiar, modificar, distribuir y presentar los datos, incluso con fines comerciales, sin necesidad de pedir permiso.**

# Wikidata

Colaborativa

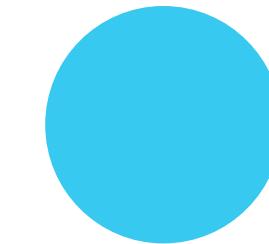


**Los datos son introducidos y mantenidos por los editores de Wikidata, quienes deciden las reglas de creación y gestión de contenidos.**

**Los bots automatizados también ingresan datos en Wikidata.**

# Wikidata

# Multilingüe



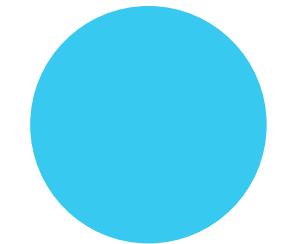
**La edición, el consumo, la navegación y la reutilización de los datos es totalmente multilingüe.**

**Los datos introducidos en cualquier idioma están disponibles de forma inmediata en todos los demás idiomas.**

**La edición en cualquier idioma es posible y fomentada.**

# Wikidata

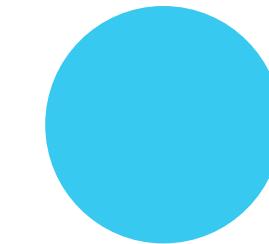
## Base de conocimiento secundaria



Wikidata no solo registra declaraciones, sino también sus fuentes y conexiones a otras bases de datos. Esto refleja la diversidad de conocimiento disponible y respalda la noción de verificabilidad.

# Wikidata

## Funcionamiento



El repositorio Wikidata consta principalmente de elementos, cada uno conteniendo una etiqueta, una descripción y una lista de alias. Los elementos son identificados de manera única por un prefijo Q seguido por un número.

Las declaraciones describen características detalladas de cada elemento y constan de una propiedad y un valor. Las propiedades en Wikidata tienen un prefijo P seguido de un número.

# Wikidata

## Ejemplo

### Jeffrey David Ullman (Q92794)

American computer scientist

Jeffrey D. Ullman | Jeffrey Ullman | Jeff Ullman



► In more languages

#### Statements

instance of

human



► 2 references

+ add value

sex or gender

male

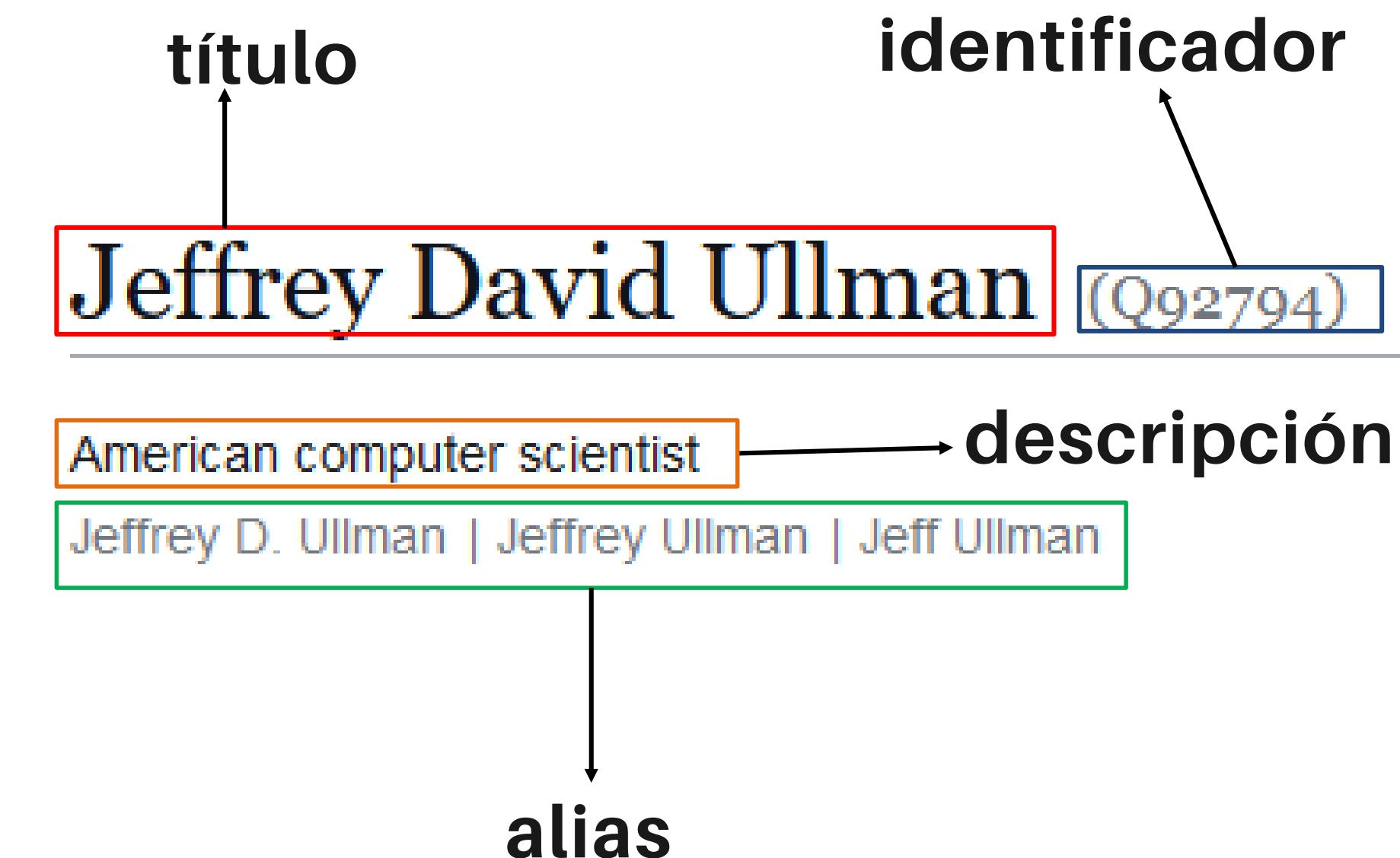


► 1 reference

+ add value

# Wikidata

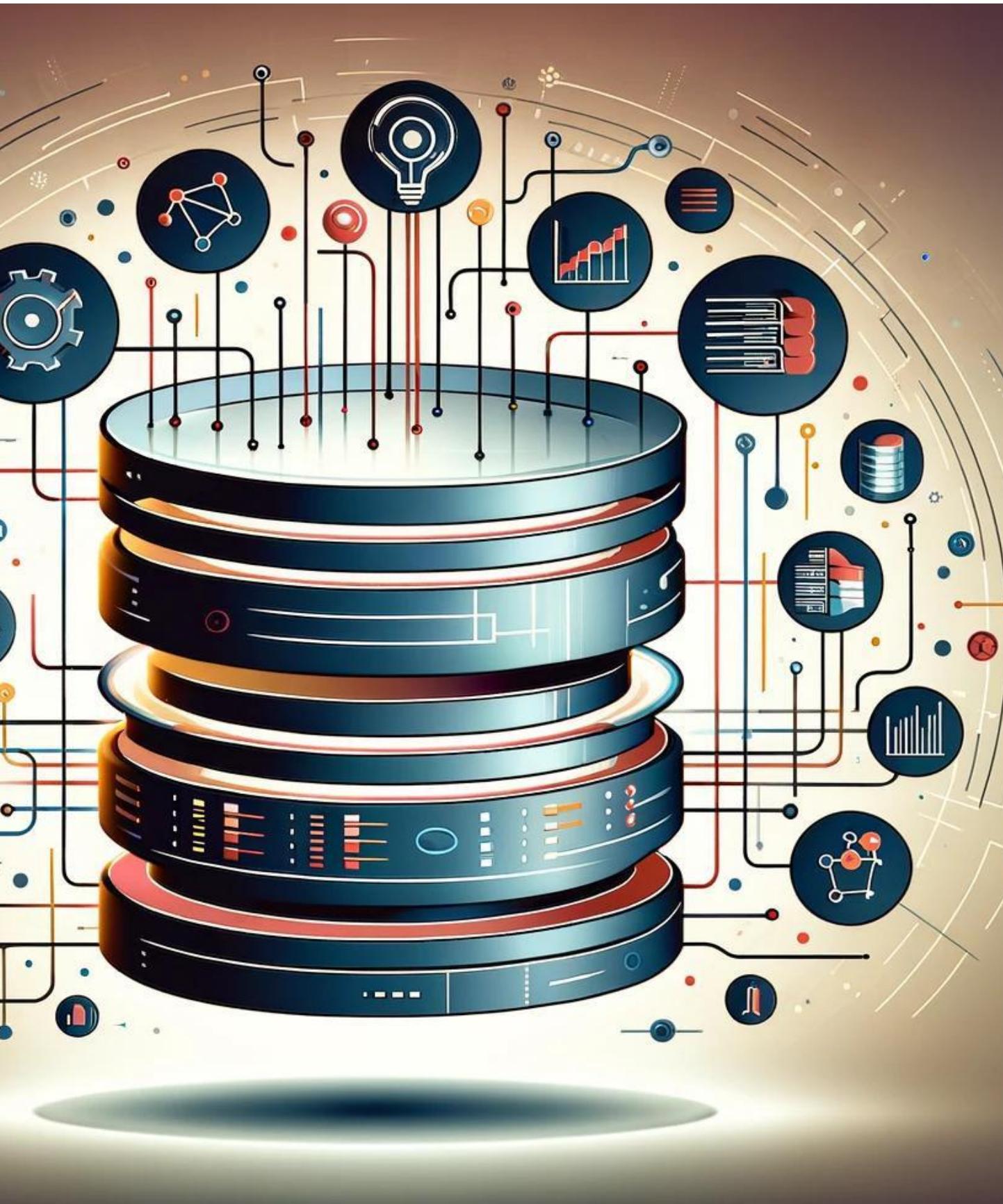
## Ejemplo



# Wikidata

## Ejemplo





# Bibliografía de la clase

# Bibliografía

- <https://datos.gob.ar/acerca/seccion/Glosario>
- <https://opendatacharter.org/who-we-are/>
- <https://creativecommons.org.ar/licencias/>
- <https://creativecommons.org>
- [https://es.m.wikipedia.org/wiki/Archivo:Las seis licencias Creative Commons.png](https://es.m.wikipedia.org/wiki/Archivo:Las_seis_licencias_Creative_Commons.png)
- <https://www.w3.org/TR/vocab-dcat/>
- <https://www.youtube.com/watch?app=desktop&v=IBqx4GDuo74>
- <https://devinit.org/wp-content/uploads/2018/01/Metadata-for-open-data-portals.pdf>
- <https://www.wikidata.org/?uselang=es>
- [https://www.wikidata.org/wiki/Wikidata:Main\\_Page](https://www.wikidata.org/wiki/Wikidata:Main_Page)
- <https://www.argentina.gob.ar/indap>



Importante!

**Los slides usados en las clases teóricas de esta materia, no son material de estudio por sí solos en ningun caso.**