

Introducción a los Metadatos



GOBIERNO DE LA
CIUDAD DE MÉXICO

AGENCIA DIGITAL DE
INNOVACIÓN PÚBLICA



Contenido

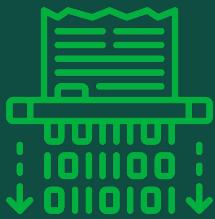
1. ¿Qué son los metadatos?
 2. Tipos de metadatos
 3. Ejemplos de metadatos
 4. Estándares de metadatos
 5. *Datálogo*: El inventario de Datos de la CDMX
 6. Primera fase del *Datálogo*: Formulario de metadatos de Sistemas de Información
-



01

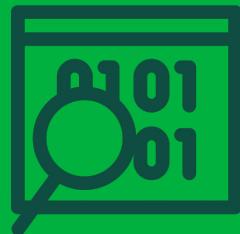
¿Qué son los
metadatos?

¿Qué son los **metadatos**?



Datos que describen la estructura
y las propiedades de otros datos.

¿Para qué **sirven**?

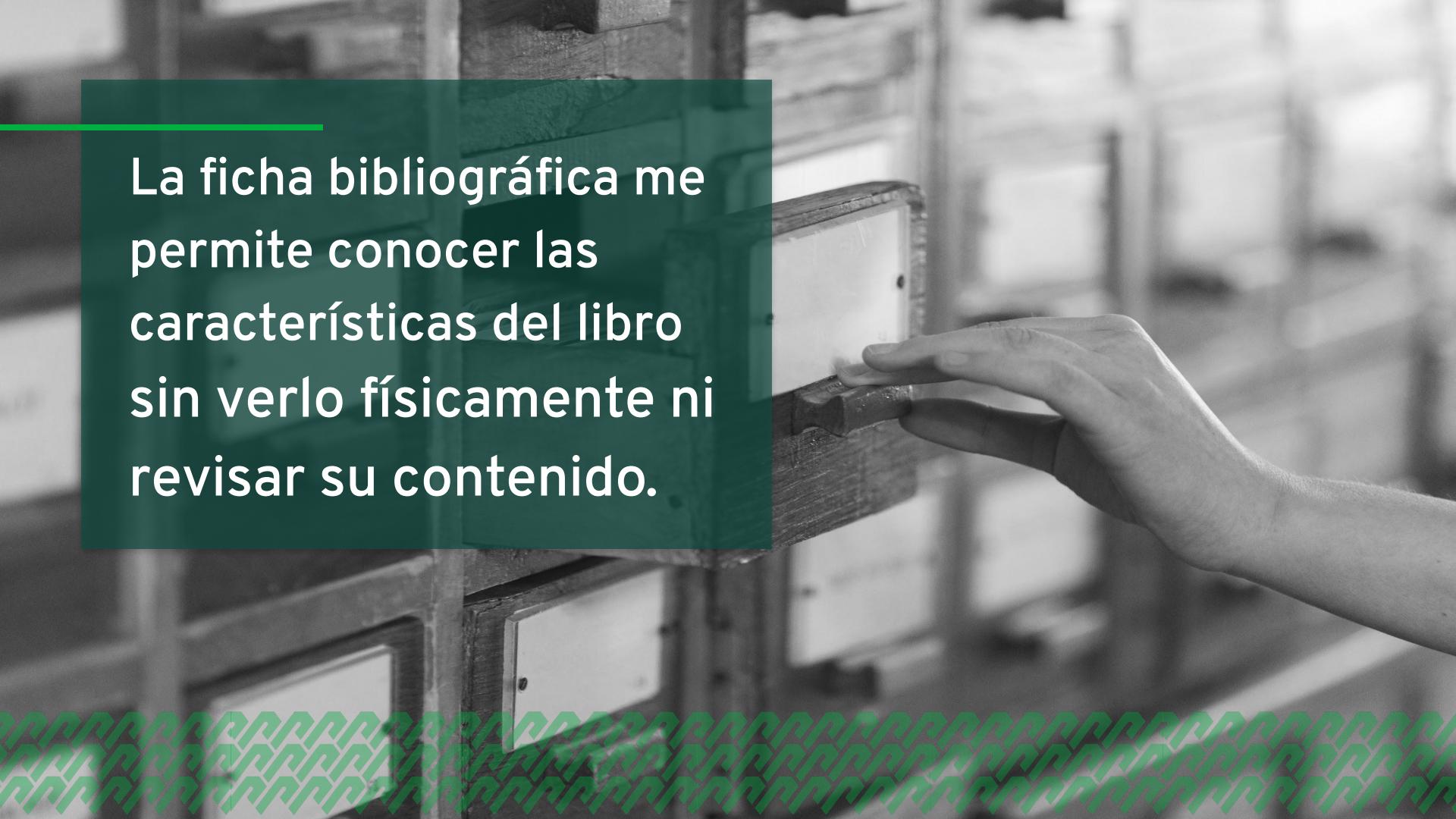


Para organizar, buscar, explorar
y entender otros datos.



Son análogos a las fichas bibliográficas: describen un libro y nos permiten ubicarlo en el estante.

La ficha bibliográfica me permite conocer las características del libro sin verlo físicamente ni revisar su contenido.

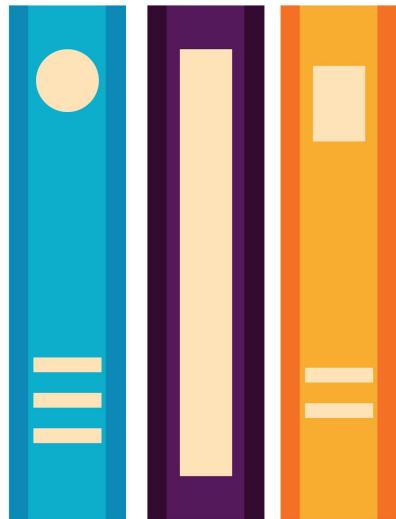


¿Qué me interesa de un **libro**?



¿Quién es el autor?
¿De qué trata?
¿En qué año se publicó
¿Cuántas páginas tiene?

Por ejemplo: ¿Cuál de los siguientes libros te interesaría leer?



Sin descripción e información contextual no podríamos decidir.

¡Es lo mismo con los datos!

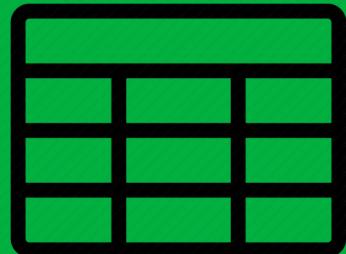


Metadata = Describen y
organizan datos



Ficha bibliográfica =
Describen y organizan libros

¿Qué me interesa de un **conjunto de datos**?



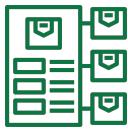
- ¿Cuándo se creó?
- ¿Qué campos contiene?
- ¿Qué Ente Público lo generó?
- ¿Qué periodo abarcan los datos?



02

Tipos de metadatos

¿Qué tipo de **metadatos** existen?

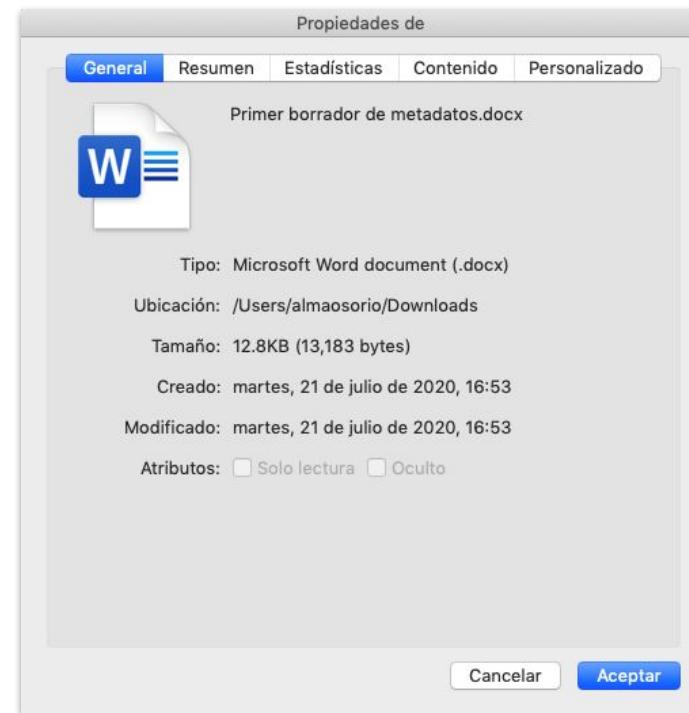
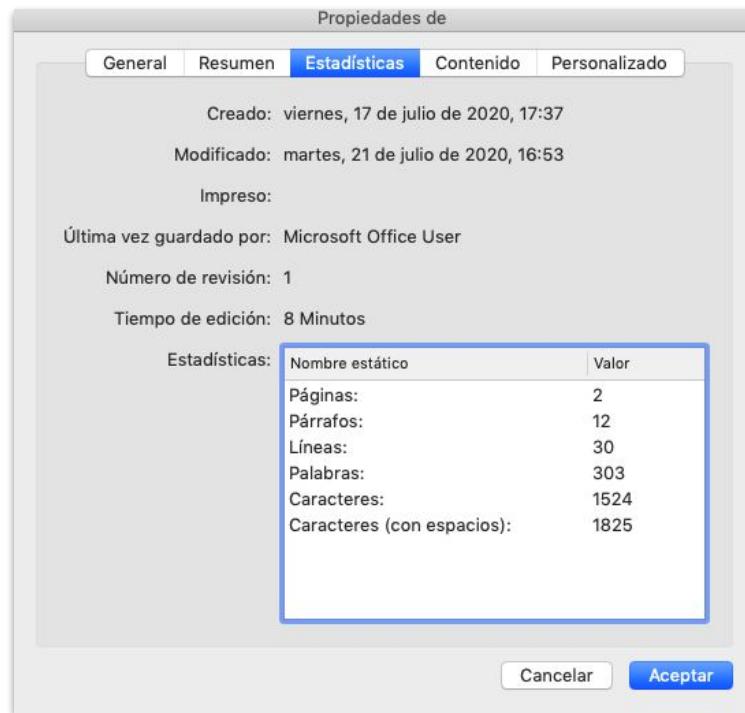
Descriptivos	Estructurales	Administrativos	De referencia
			
Información sobre el contenido de los datos. Sirve para identificación y búsqueda.	Información que describe la estructura o formato de un recurso.	Información sobre la administración o gestión de un recurso.	Información sobre el contenido y la calidad de los datos estadísticos.
<ul style="list-style-type: none">■ Autor■ Título■ Tema■ Palabras clave	<ul style="list-style-type: none">■ Formato■ Tamaño■ Lenguaje	<ul style="list-style-type: none">■ Fechas de creación, modificación y acceso.■ Permisos y licencias	<ul style="list-style-type: none">■ Información conceptual■ Información metodológica



03

Ejemplos de
metadatos

Ejemplo de metadatos



Más ejemplos



[Metadatos del banco mundial](#)

[Metadatos del Data.gov](#)

[Metadatos del gobierno de españa](#)

[Metadatos de Google](#)

[Metadatos del INEGI](#)



04

Estándares de Metadatos

¿Para qué sirven?



Para tener una forma homogeneizada y ordenada de sistematizar y documentar los metadatos de alguna organización en específico, para poder buscar y localizarlos..

¿Qué pasa si todos hacemos nuestros metadatos de forma **distinta**?



No podemos integrar y sistematizar los datos.

Es decir: no podemos usar los metadatos como herramienta para organizar y buscar información.

Si los metadatos se colectan de manera estandarizada con un mismo formato:



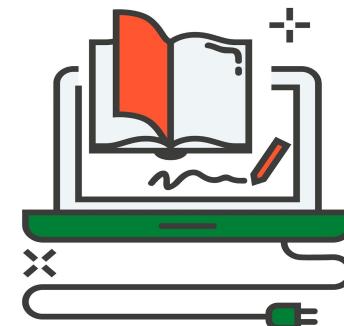
Podemos hacer
comparaciones y tomar
decisiones.

Título: El Quijote
Autor: Miguel de Cervantes
Año: 1605
Disciplina: Literatura

Título: Estadística Aplicada
Autor: Jesús Rodríguez
Año: 2009
Disciplina: Matemáticas



Título, Autor, Año y Disciplina,
son los metadatos.



Principales estándares internacionales de metadatos:

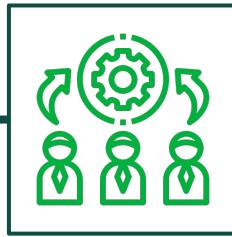


DCAT

DDI

Dublin Core

Schema



Dudas y/o comentarios



05

Datálogo: el inventario de datos de la CDMX



Documentar los **metadatos** de la APCDMX de forma estandarizada nos ayudará a construir el **Inventario de Datos de la CDMX**.

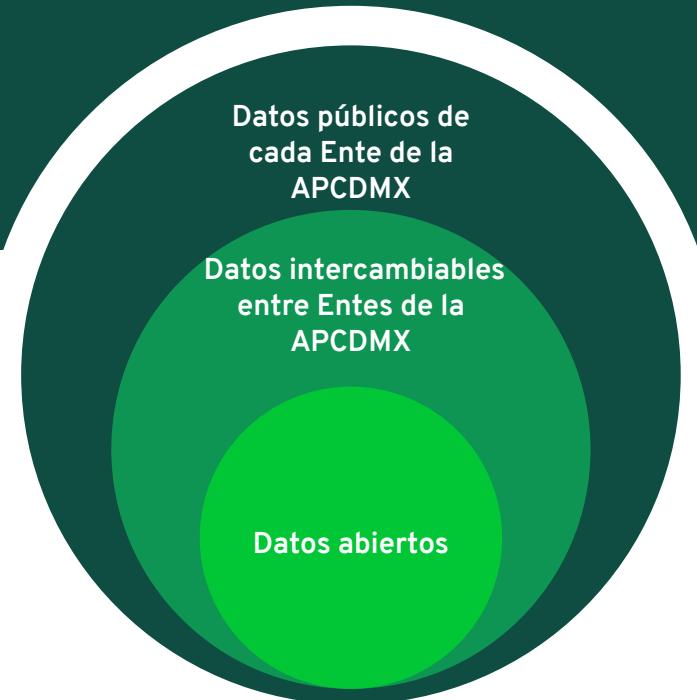
¿Qué metadatos recogerá el Inventario de Datos?



Los de todos los *datos públicos*

Datos públicos =
Propiedad del pueblo

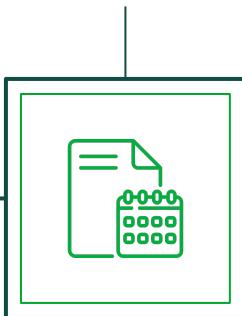
Datos públicos
≠ datos abiertos



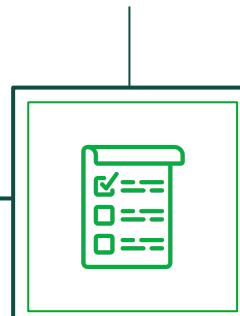
Fases del Inventario de Datos

La recopilación de los metadatos de la APCDMX se realizará de manera gradual:

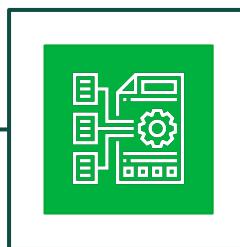
1. Sistemas de Información



3. Conjuntos de datos



2. Bases de Datos



4. Capas geográficas



Datálogo CDMX



06

Primera fase: Sistemas de Información

Sistema de Información



- Conjunto integrado de componentes de recursos de tecnologías de información* que tiene como objeto recolectar, almacenar y/o procesar datos.
 - Suelen consultar o escribir información en una base de datos.
-

Ejemplos de sistemas de información:

- Sistema de Información, Evaluación y Seguimiento (SIEV).
- Sistema Único de Información Delictiva (SUIDF).
- Sistema Integral de Administración de Personal (SIAP).

*Recursos de Tecnologías de Información: Los conjuntos de datos, bases de datos, las aplicaciones y/o sistemas de información digitales, la tecnología (hardware, software, sistemas operativos, sistemas manejadores de bases de datos, redes, y demás aplicables) e instalaciones (recursos para alojar y dar soporte a los sistemas de información).

Formulario de metadatos para Sistemas de Información:

Campos solicitados:



1	Datos de quien responde (nombre, Ente público donde trabaja, nivel, puesto, correo electrónico y teléfono de contacto).
2	Nombre completo y acrónimo del Sistema
3	Para qué se utiliza el sistema
4	Quiénes son los usuarios
5	Cómo se desarrolló
6	Quién le da mantenimiento

Formulario de metadatos para Sistemas de Información:

Campos solicitados:



7	Donde se albergan el software y/o bases de datos del sistema.
8	Tiene bases de datos y cuántas
9	Nombre de las bases de datos
10	Usa Servicios web y/o APIs
11	Datos de la persona encargada de configurar y asegurar el funcionamiento correcto del sistema (nombre, área responsable, nivel, cargo, correo electrónico y teléfono de contacto).

Instrucciones generales



1. Los Ol recibirán por correo electrónico una liga para responder el Formulario.
2. Deberán acompañarse de alguien que conozca del Sistema.
3. Dar “enviar”.
4. Registrar un Formulario por cada Sistema.

¿Qué sistemas de información debo registrar?

Todos aquellos que son propiedad del Ente Público donde trabajas. Es decir, aquellos que el Ente Público desarrolló, solicitó a un proveedor que fueran desarrollados o por los que paga una licencia.

¿Llenar el formulario de metadatos es igual a abrir mis datos?

NO

Los metadatos documentan las **características** de los datos.

El llenado de alguno de los formularios no implica acceder o compartir el sistema, la base y/o el conjunto de datos.



Dudas y/o comentarios

Dirección de Coordinación de Proyectos de Datos

 politicadedatos@cdmx.gob.mx



GOBIERNO DE LA
CIUDAD DE MÉXICO

AGENCIA DIGITAL DE
INNOVACIÓN PÚBLICA

Contacto



Residuos - guardar estas
ppt para otro momento

Un ejemplo



The screenshot shows a Mac OS X desktop environment with a window titled "Debating deliberative democracy [WorldCat.org]". The URL in the address bar is http://www.worldcat.org/oclc/49875309&referer=brief_results. The page displays the following information about the book:

Debating deliberative democracy
by [James S Fishkin](#); [Peter Laslett](#)

(not yet rated)

Type: Book; English
Publisher: Malden, MA : Blackwell, ©2003.
Editions: [4 Editions](#)
ISBN: 1405100427 9781405100427 1405100435 9781405100434
OCLC: 49875309
Related Subjects: [Deliberative democracy](#), [Representative government and representation](#), [Democracy](#), [More Subjects ...](#)
Citations: [Cite this Item](#) | [Export to EndNote](#) | [Export to ReWorks](#)

Get It
[Google Preview](#)
[Search my library](#)

Save It
[Add This Page to Favorites](#)
Save to: [\[New List\]](#)

Add to It
[Review this Item](#)
[Edit Public Notes](#)
[Tag this Item](#)

Share It
[Link to this Page](#)
[E-mail this Item](#)
[BOOKMARK](#)

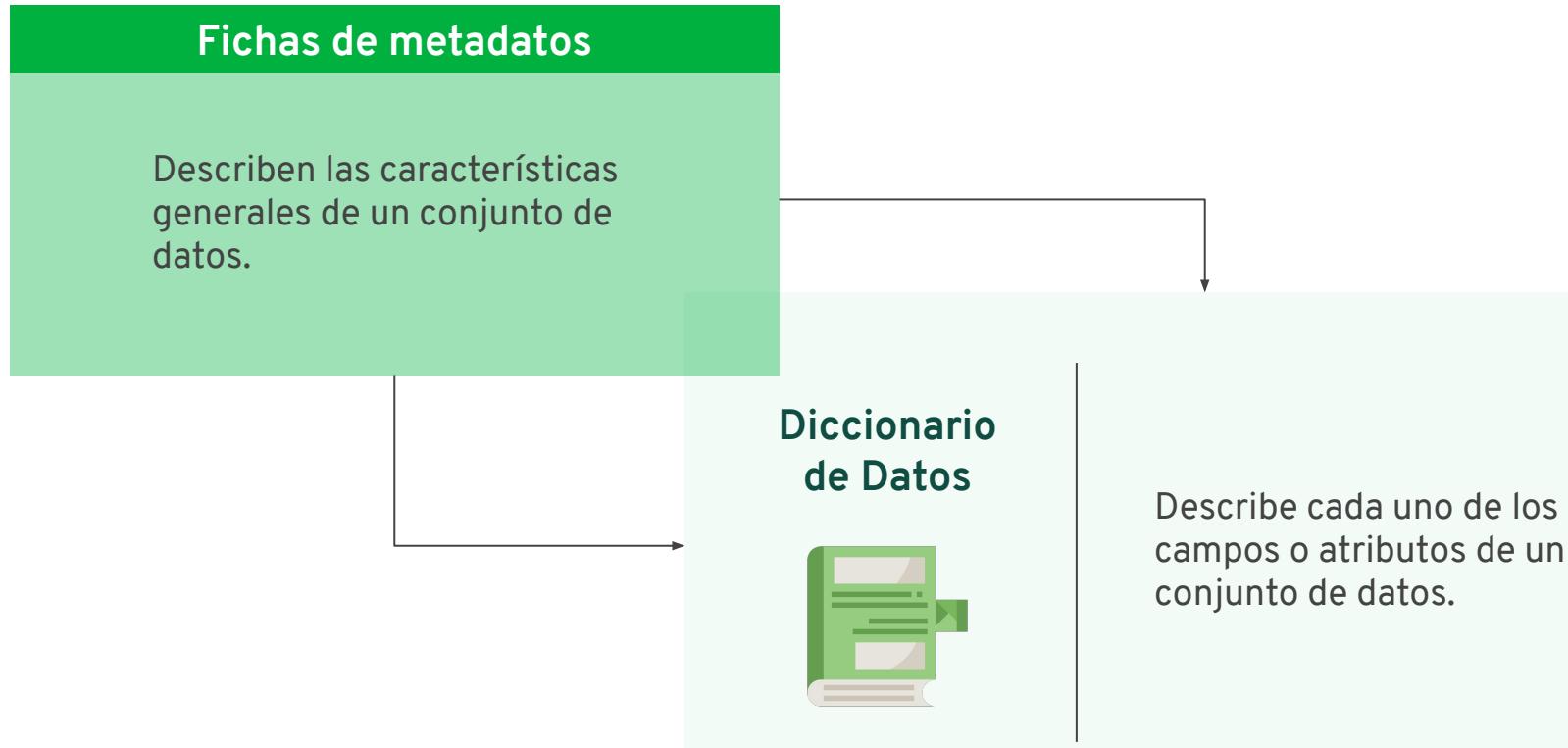
To the right of the book details, there is a thumbnail image of the book cover for "Debating Deliberative Democracy" by James S. Fishkin and Peter Laslett.



04

Estándar de **metadatos** de la Ciudad de México

¿En qué consiste el **estándar de metadatos** de la CDMX?



Fichas de metadatos de la APCDMX: campos básicos en común

Información básica del contenido



1	Título de la base de datos y/o base de datos
2	ID de la base de datos y/o base de datos (autogenerado)
3	Descripción
4	Tema
5	Palabras clave
6	Idioma

Fichas de metadatos de la APCDMX: campos básicos en común

Información temporal



7	Cobertura temporal
8	Frecuencia de actualización
9	Última fecha de actualización

Información geográfica



10	Cobertura geográfica
11	Desagregación geográfica
12	Última fecha de actualización

Fichas de metadatos de la APCDMX: campos básicos en común

Información metodológica



13	Especificaciones técnicas
14	Material de apoyo

15	Diccionario de datos
----	----------------------

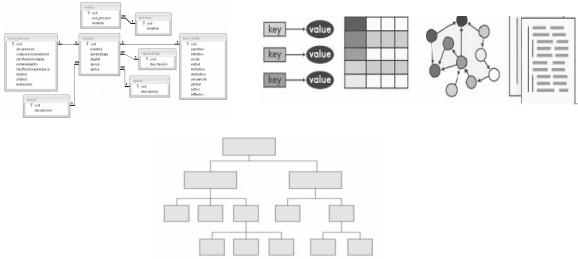
Información básica de contacto



16	Propietario
17	Nombre de contacto
18	Teléfono de contacto

Fichas de metadatos de la APCDMX:

Bases de datos



Bases de datos albergadas en un sistema informático.

Ejemplo: Relacionales (SQL), No relacionales (NoSQL), jerárquicas, de red.

Conjunto de datos



Conjuntos de datos en archivos electrónicos.

Ejemplos: Excel, Google Spreadsheets, CSV.

Datos geográficos



Conjuntos de datos que representan puntos, líneas o planos geográficos.

Ejemplos: shapefile, GEOJSON, KMZ, KML.

Algunos campos de la Ficha Estándar de metadatos para bases de datos

Nombre del campo	Obligatoriedad	Especificaciones	Ejemplo
Nombre del Sistema informático que contiene a la base de datos	Obligatorio	Nombre completo del Sistema informático en el que se encuentra la base de datos. Incluir siglas, abreviaturas o acrónimos; así como periodo y/o año de referencia.	<i>Sistema Integral de Gestión de Riesgos Ambientales (SIGRA) 2010-2020</i>
Sistema gestor de base de datos (SGBD)	Obligatorio	EEI Sistema gestor de base de datos (SGBD) es un software que permite definir, manipular, extraer, administrar y utilizar la información que contienen las bases de datos. Puede ser: -Microsoft Access -Microsoft SQL Server -MySQL...	MySQL
Tipo de modelo de base de datos	Obligatorio	El tipo de modelo de la bases de datos es la forma en la que se estructuran los datos en el Sistema informático. Puede ser: -Relacional -Red...	<i>Relacional</i>

Algunos campos de la Ficha Estándar de metadatos para conjuntos de datos (Excel o csv)

Nombre del campo	Obligatoriedad	Especificaciones	Ejemplo
Tamaño del conjunto de datos	Obligatorio	Volumen del archivo del conjunto de datos en unidades de almacenamiento de información. Se debe incluir la unidad de medida en el que se encuentra contenido el archivo: Kilobyte, Megabyte (MB), Gigabyte (GB) o Terabyte (TB).	2.87 GB
Número de campos o atributos	Obligatorio	Número de campos o atributos (también llamados variables) que conforman el conjunto de datos (columnas).	15
Número de registros u observaciones	Obligatorio	Número de registros que contiene el conjunto de datos (filas).	2345

Algunos campos de la Ficha Estándar de metadatos para datos geoespaciales

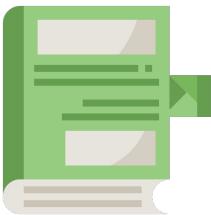
Nombre del campo	Obligatoriedad	Especificaciones	Ejemplo
Formato	Obligatorio	Formato en que el conjunto de datos geoespaciales se codifica y almacena en un archivo informático. Puede ser: - .shp - .dwg - .dxf - .kml - .kmz - GeoJson - GeoRSS...	.kml
Código EPSG/SRID/ITRF	Opcional	El EPSG, SRID e ITRF son códigos que hacen referencia a un sistema de referencia espacial, definidos por coordenadas del sistema local, regional o global que se utiliza para localizar entidades geográficas.	EPSG 4326
Geometría	Obligatorio	Tipo de geometría utilizado en los datos geoespaciales. Puede ser: -Punto -Línea -Polígono -No aplica	Polígono



05

Un metadato muy
especial: **el diccionario
de datos**

¿Qué es un **diccionario de datos**?



Es un tipo de metadato que enlista:

- los nombres
- las definiciones
- las características

De:

- Cada uno de los campos, variables o atributos de un conjunto de datos.

¿Para qué sirve un diccionario de datos?

Para entender e interpretar un conjunto de datos y/o base de datos.

¿Qué contiene un diccionario de datos?

Información básica sobre los campos o variables que contiene un conjunto y/o base de datos.

¿Qué tipos de diccionarios de datos existen?

Activos: Forman parte de un sistema gestor de base de datos, se generan y se actualizan automáticamente.

Pasivos: Se elaboran de forma manual, usualmente en un formato csv o excel.

¿Cómo se ve un diccionario de datos?

- ❑ Ejemplo: dataset1

id_us	username	nombre_us	edad_us	sexo_us
7889	pedro_lop_ju	Pedro López	15	M
8392	pricess_su	Ana Suárez	18	F
0286	pepetono_2	Juan Pérez	32	M

¿Cómo se ve un diccionario de datos?

□ Ejemplo: Diccionario del dataset1

Etiqueta	Título campo	Descripción	Tipo	Reglas de cálculo	Valores válidos
id_us	Número de identificación de usuario	Identificador único que el sistema asigna al usuario momento de hacer una cuenta	Número entero	Campo autogenerado aleatoriamente por el sistema	Código numérico a 4 dígitos
username	Nombre de usuario	Nombre que el usuario elige al hacer una cuenta	Texto	No aplica	Secuencia alfanumérica única de hasta 15 caracteres
nombre_us	Nombre del usuario	Nombre completo que el usuario que registra al darse de alta	Texto	No aplica	Secuencia alfanumérica de hasta 100 caracteres
edad_us	Edad del usuario	Edad del usuario	Número entero	Fecha actual- Fecha de nacimiento	0-99
sexo_us	Sexo del usuario	Sexo del usuario	Boolean	Si es mujer	mujer=TRUE; hombre=FALSE

Ejemplo de un diccionario de datos



[https://databank.worldbank.org/source/
world-development-indicators#](https://databank.worldbank.org/source/world-development-indicators#)

Estándar de diccionario de datos de la APCDMX

1	Nombre de la base y/o conjunto de datos
2	Nombre de la tabla de la base de datos u hoja del conjunto de datos
3	Etiqueta del campo o atributo
4	Nombre del campo o atributo
5	Descripción del campo o atributo
6	Tipo del campo o atributo
7	Valores que permite el campo

Estándar de diccionario de datos de la APCDMX

8	Clasificación de la información (confidencialidad)
9	Especificación del campo
10	Reglas de cálculo
11	Cambia el dato con el tiempo
12	Permite celdas vacías o campos vacíos
13	Cómo se registran las celdas o campos vacíos

¿Cómo luce el estándar de diccionario de datos de la CDMX?



Mira cómo luce este Diccionario de datos de una base de datos de SEMOVI el cual se hizo usando la plantilla estándar.

¡Hagamos un diccionario de datos!
Entremos al Portal de Datos Abiertos de la Ciudad.