

# Guía para la Estandarización de Datos Abiertos

Dirección de Gobierno en Línea

MinTIC

Gobierno de Colombia



# Tabla de Contenidos

- 0 Introducción
- 1 Estándares de datos abiertos
- 2 Compartir Datos en la Web: Recomendaciones del W3C
- 3 Casos de éxito de aplicación de estándares de datos abiertos
- 4 Recomendaciones y hoja de ruta

# Introducción

- ▶ Esta guía establece recomendaciones y lineamientos para la estandarización de datos abiertos en Colombia. La guía pretende dar los insumos para ayudar a servidores públicos con un perfil técnico, interesados en la apertura de datos en Colombia según las prácticas y estándares internacionales ya existentes.
- ▶ Con esta guía se busca contribuir a un mejor entendimiento de la importancia del uso de estándares, sus implicaciones y sus posibilidades.

## Estructura de la guía:

- ▶ Definición de estándares y estándares abiertos;
- ▶ Recomendaciones del World Wide Web Consortium en el estándar para compartir datos abiertos en la Web;
- ▶ Casos prácticos de aplicación de estándares de datos abiertos internacionales;
- ▶ Recomendaciones y propuesta de hoja de ruta para la aplicación de estándares de datos abiertos en Colombia;

# Datos Abiertos en Colombia

- ▶ Los datos abiertos son parte de la Estrategia de Gobierno En Línea (GEL) del MinTIC, bajo el eje temático TIC para Gobierno Abierto.
- ▶ Colombia ha sido pionera a nivel regional en datos abiertos. Una de las acciones de la iniciativa de datos abiertos ha sido la creación de una guía para la apertura de datos en Colombia y la introducción de la apertura de datos en los lineamientos del manual para Gobierno En Línea colombiano.
- ▶ Esta guía aporta insumos para la guía para la apertura de datos en Colombia. A continuación se presentan un conjunto importante de insumos que pueden ampliar el impacto de la apertura de datos de las entidades públicas colombianas a través del uso de estándares de datos abiertos.

# Convenio MinTIC - Banco Mundial

- ▶ El Gobierno de Colombia, a través de MinTIC ha establecido un convenio con el Banco Mundial para el apoyo y fomento de la innovación pública digital en Colombia.
- ▶ Uno de los ejes principales de la colaboración se ha centrado en la estrategia de datos abiertos, a través de un diagnóstico de la iniciativa de apertura de datos del país, del desarrollo de una nueva plataforma para la iniciativa de datos abiertos y, del desarrollo de esta guía para la apertura de datos en Colombia, entre otros.
- ▶ Los contenidos de esta guía han sido preparados en colaboración con MinTIC y a partir de los talleres organizados con servidores públicos involucrados en la iniciativa de datos abiertos del Gobierno. De las discusiones de esos talleres han resultado insumos clave para la diseñar y validar el enfoque y la importancia de los elementos que conforman esta guía.

# 1

## Estándares de datos abiertos

### Definiciones

Esta sección presenta las definiciones de algunos términos que utilizaremos frecuentemente en esta guía.

# 1.1 Definición de abierto y de datos abiertos

Según <http://opendefinition.org/>, **abierto** es algo que cualquier persona puede acceder, utilizar, modificar y compartir, gratuitamente y para cualquier fin (sujeto a requisitos que preserven la proveniencia y la apertura).

En ese contexto, la definición de **datos abiertos** nos dice que son aquellos datos que pueden ser “utilizados, modificados y compartidos por cualquier persona con cualquier fin.”

Fuente: <http://opendefinition.org/>

# 1.2 Estándar Técnico y Estándar Abierto

## Estándar Técnico

Un estándar técnico es una norma o un conjunto de requisitos establecidos en relación con un sistema técnico.

Normalmente, un estándar técnico es un documento formal que establece criterios, métodos, procesos y prácticas técnicas o de ingeniería.

Fuente: [https://en.wikipedia.org/wiki/Technical\\_standard](https://en.wikipedia.org/wiki/Technical_standard)

## Estándar Abierto

Los estándares abiertos son aprobados por comités formales abiertos a la participación de todos los actores interesados y operan con base en consensos. Un estándar abierto es de utilización gratuita, y está disponible públicamente. Estos estándares son desarrollados, aprobados y mantenidos de manera colaborativa con base en consensos.

Fuente: [https://en.wikipedia.org/wiki/Open\\_standard](https://en.wikipedia.org/wiki/Open_standard)



2

# Compartir Datos en la Web: Recomendaciones del W3C

## 2.1 Introducción

En esta sección se presenta la propuesta de estándar del World Wide Web Consortium (W3C) para compartir datos en la web. Inicialmente, se presenta qué es el W3C y cómo funcionan sus estándares.

En seguida, se exploran las recomendaciones del W3C para compartir datos en la web.

## 2.2 Estándares del World Wide Web Consortium (I)

Según, el World Wide Web Consortium (W3C), sus estándares tienen un conjunto de características comunes que justifican su amplia adopción por la industria de TI y por desarrolladores de aplicaciones web. Las características comunes son las siguientes:

- Los estándares son creados según un **proceso de decisión basado en consensos**. No existe un proceso unidireccional de decisiones para la creación de un nuevo estándar, sino que se ha diseñado un proceso de trabajo basado en decisiones consensuales para la definición de nuevos estándares.
- Para asegurar su adopción a una escala lo más amplia posible, los estándares **consideran aspectos como la accesibilidad, privacidad, seguridad e internacionalización**. Estos aspectos van más allá de la definición de características técnicas del uso de las TIC en la web y discuten también aspectos importantes para los usuarios como la accesibilidad de las tecnologías, los aspectos de seguridad y privacidad de la información o la aplicación del estándar a otros idiomas, sistemas de fechas y de numeración.
- Gracias a la gran visibilidad de estos estándares la participación de los distintos actores está garantizada, por lo que los estándares **reflejan perspectivas de distintas industrias y actores globales**.

## 2.2 Estándares del World Wide Web Consortium (II)

- Equilibran **velocidad, justicia, rendición de cuentas y calidad**. Las actividades para desarrollar los estándares son publicadas por adelantado, así como el alcance del trabajo y las condiciones para poder participar en su elaboración.
- Todas las decisiones y el material utilizado es registrado, publicado y fácilmente accesible.
- Las actividades para el desarrollo de estándares no son excluyentes, ni dominadas por ninguna persona, empresa o grupo de interés.
- Benefician de un compromiso de no cobranza de licencia de patentes por los participantes y se pueden bajar sin costes
- Son **estables**, una vez que el W3C busca asegurar la persistencia de su URI
- Benefician de una revisión profunda por grupos dentro y fuera del W3C, lo que ofrece garantía de actualización regular.
- Son mantenidos de una forma regular
- Son fortalecidos a través de pruebas de interoperabilidad.

## 2.3 Valores de los estándares del W3C

Aún según el W3C, los estándares que propone tienen una oferta de valor importante para los usuarios, empresas o entidades que los adopten:

- **Transparencia** - todo el proceso de creación, las discusiones técnicas y las actas de las reuniones son archivadas.
- **Relevancia** - la creación de nuevos estándares se inicia después del análisis de las necesidades del mercado, lo que incluye una fase de requerimientos.
- **Apertura** - cualquier persona puede participar: industria, personas individuales, gobiernos y academia, a una escala global.
- **Imparcialidad y consenso** - la justicia del proceso y la neutralidad del W3C es garantizada, con igual peso para cada participante.
- **Disponibilidad** - acceso gratis al texto del estándar y de sus traducciones, desde el momento de desarrollo a su versión final. Garantía de que tecnologías de la Web y del Internet pueden ser implementadas sin pago de derechos.
- **Mantenimiento** - proceso continuo de prueba, erratas, revisiones, acceso permanente y validación.

## 2.4 Mejores prácticas para compartir datos en la web (I)

El W3C está en este momento en el proceso de creación de un estándar que defina mejores prácticas para compartir datos en la web. Este estándar se concentra en la creación de un ecosistema sostenible para compartir datos, donde humanos y máquinas pueden descubrir y entender datos, facilitando la interacción entre las entidades que publican y las que consumen datos en la web. El estándar se detiene sobre un conjunto de retos sobre datos en la web y, sobre cada uno de esos se han identificado algunas mejores prácticas que contribuyen para su resolución.

- ▶ **Metadatos:** Cómo publicar metadatos legibles por humanos y máquinas?
- ▶ **Identificación de datos:** Cómo identificar conjuntos de datos y sus distribuciones?
- ▶ **Vocabularios de datos:** Cómo lograr la interoperabilidad semántica?
- ▶ **Datos sensibles:** Cómo respetar seguridad y privacidad?
- ▶ **Licencias de datos:** Cómo permitir y restringir el acceso a los datos?
- ▶ **Origen y Calidad de los datos:** Cómo agregar confianza en el ecosistema?
- ▶ **Acceso a los datos:** Opciones de acceso a los datos.
- ▶ **Formatos de datos:** Qué formatos de datos debemos utilizar?
- ▶ **Versionamiento de datos:** Cómo generar rastreabilidad de versiones y series de datos?
- ▶ **Conservación de datos:** Cómo se deben archivar los datos?
- ▶ **Retroalimentación:** Cómo involucrar a los usuarios para mejorar la calidad de los datos?
- ▶ **Enriquecimiento de los datos:** Cómo agregar valor a los datos?

## 2.4 Mejores prácticas para compartir datos en la web (II)

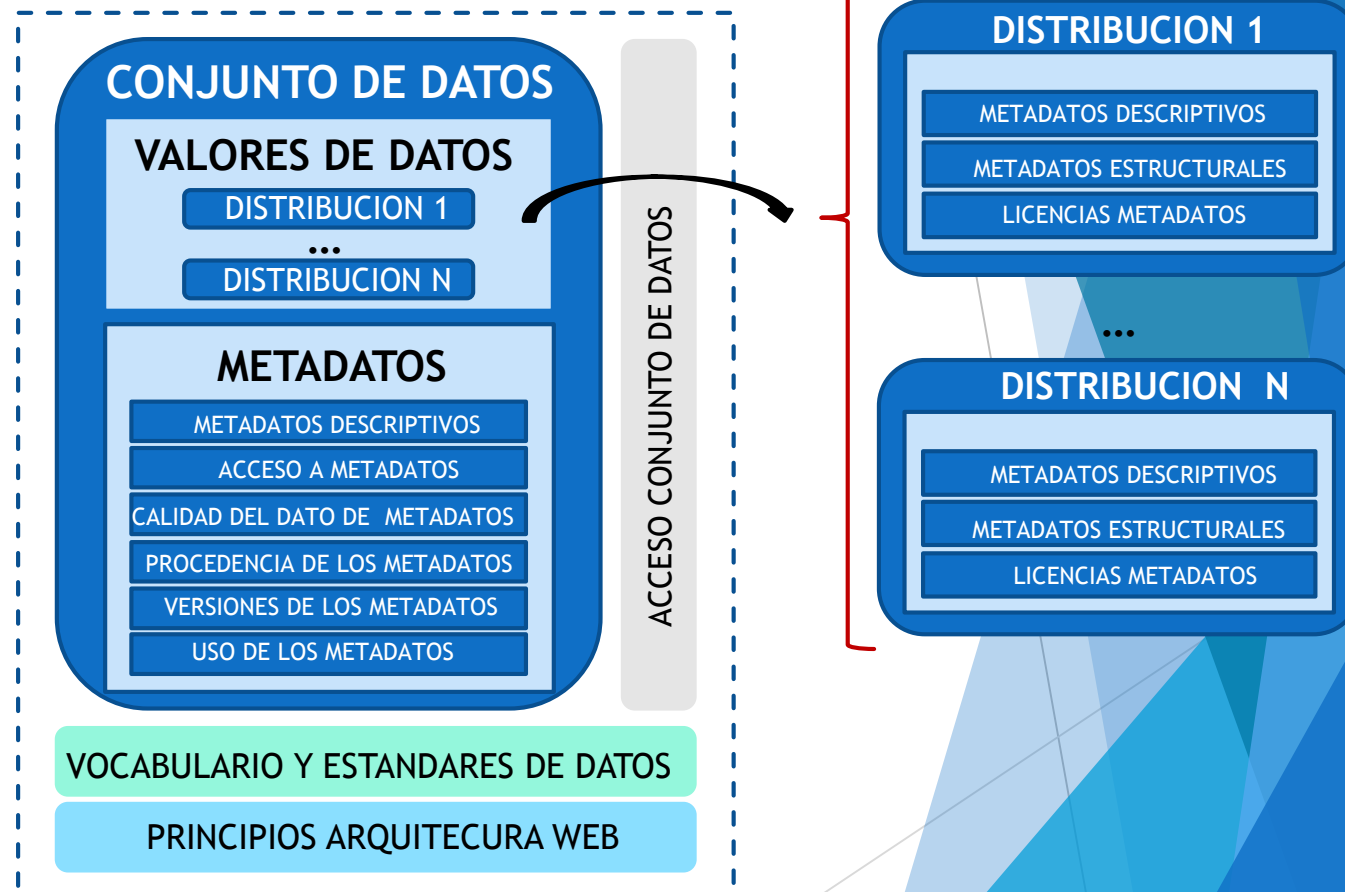
El diagrama ilustra una composición de los conjuntos de datos (valores y metadatos), así como los otros componentes relacionados con la publicación y utilización de conjuntos de datos.

Los valores de los datos corresponden a los datos.

Las distribuciones de datos corresponden a la publicación de los datos en un formato específico (conjunto de datos).

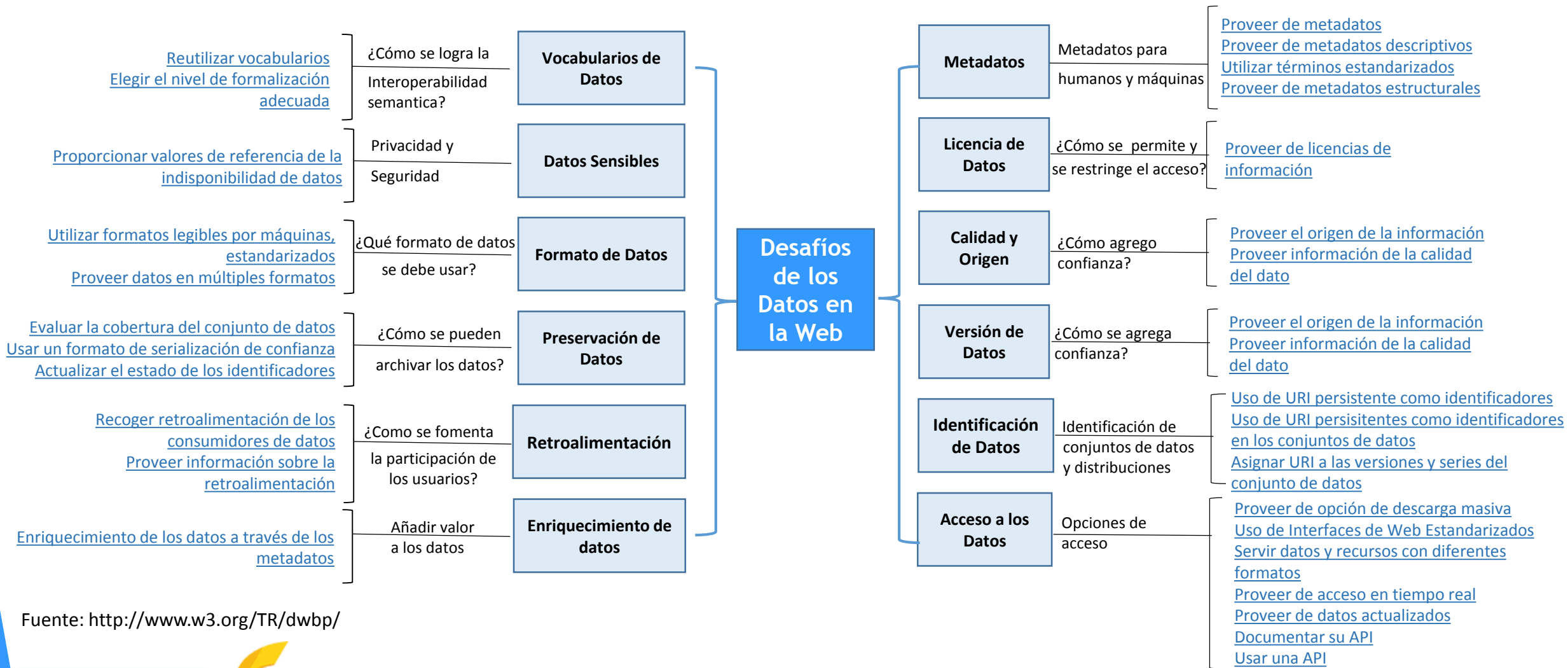
Los metadatos corresponden a la información adicional que describe el conjunto de datos y sus distribuciones.

Para permitir interoperabilidad entre conjuntos de datos es importante adoptar vocabularios y estándares de datos.



Fuente: <http://www.w3.org/TR/dwbp/>

## 2.4 Mejores prácticas para compartir datos en la web (III)



Fuente: <http://www.w3.org/TR/dwbp/>



## 2.5 Metadatos

### Desafíos

- ▶ ¿Qué tipos de metadatos se deben considerar cuando describimos datos en línea?
- ▶ ¿Cómo podemos abrir metadatos en lenguaje legible por máquinas?

### Mejores prácticas

- ▶ Proveer metadatos utilizables por usuarios humanos o por aplicaciones informáticas
- ▶ Proveer descripción de los metadatos
- ▶ Proveer parámetros locales (fecha, hora, formatos de números, idioma) en los metadatos
- ▶ Proveer metadatos sobre el esquema y la estructura de las distribuciones

## 2.5 Metadatos

### Catalogación de datos abiertos

- ▶ Los datos deben ser encontrables y entendibles tanto por personas como por máquinas, así la descripción de la información reutilizable servirá para clasificar los datos que las entidades publican y así facilitar la búsqueda y descubrimiento de la información.
- ▶ Cualquier información reutilizable -catálogo, conjuntos de datos y distribuciones- debe llevar asociado unos metadatos mínimos que describen detalladamente los recursos de información que se permiten reutilizar. Tan importante como los datos expuestos para la reutilización, son los metadatos que proporcionan información descriptiva sobre los recursos publicados. Para la descripción de esta información se toma como referencia el estándar DCAT.
- ▶ Además de la información, comprensible para las personas, incluida en el catálogo de datos y las fichas de los conjuntos de datos publicados en un portal de datos abiertos, es necesario identificar los diferentes elementos de información reutilizable mediante referencias únicas, unívocas, fiables y persistentes en el tiempo, utilizando para ello un esquema de URIs (Identificadores de Recursos Uniformes -del inglés, Uniform Resource Identifiers o URIs-). Estos identificadores constituyen los valores que tomarán determinados metadatos utilizados para describir cada recurso de información, por ejemplo, la procedencia, el tipo de información o la cobertura geográfica del dataset.

## 2.5 Metadatos

### El estándar DCAT

- ▶ El conjunto de metadatos seleccionados para describir cada uno de los recursos de información, si bien es una decisión de los gestores de la iniciativa de reutilización, debe conducir a la máxima expresividad posible de la información proporcionada sobre los recursos para los reutilizadores. Esto incrementará la calidad descriptiva de los recursos publicados y aumentará la potencialidad de reutilización. Sin embargo, en ocasiones no es posible contar con la metainformación suficiente para describir adecuadamente dichos recursos, por tanto es necesario establecer un vocabulario mínimo de metadatos que siempre puedan (y deban) ser descritos y publicados de forma constante en el proceso de publicación de nuevos datasets.
- ▶ En las normas o guías para la apertura de datos se determinan aquellos metadatos que son de obligatoria publicación, así como los recomendados u opcionales. Los metadatos se dividen en:
  - ▶ **Obligatorios:** aquellos que, por su especial relevancia y también por cumplimiento con la norma establecida, será obligado cumplimentar.
  - ▶ **Recomendados:** aquellos que, si bien no serían obligatorios para cumplir con la legislación vigente, igualmente es muy recomendable que se proporcionen por su relevancia a la hora de clasificar los datos adecuadamente.
  - ▶ **Opcionales:** aquellos que, aun siendo igualmente recomendables, puede que no siempre estén disponibles.

## 2.5 Metadatos

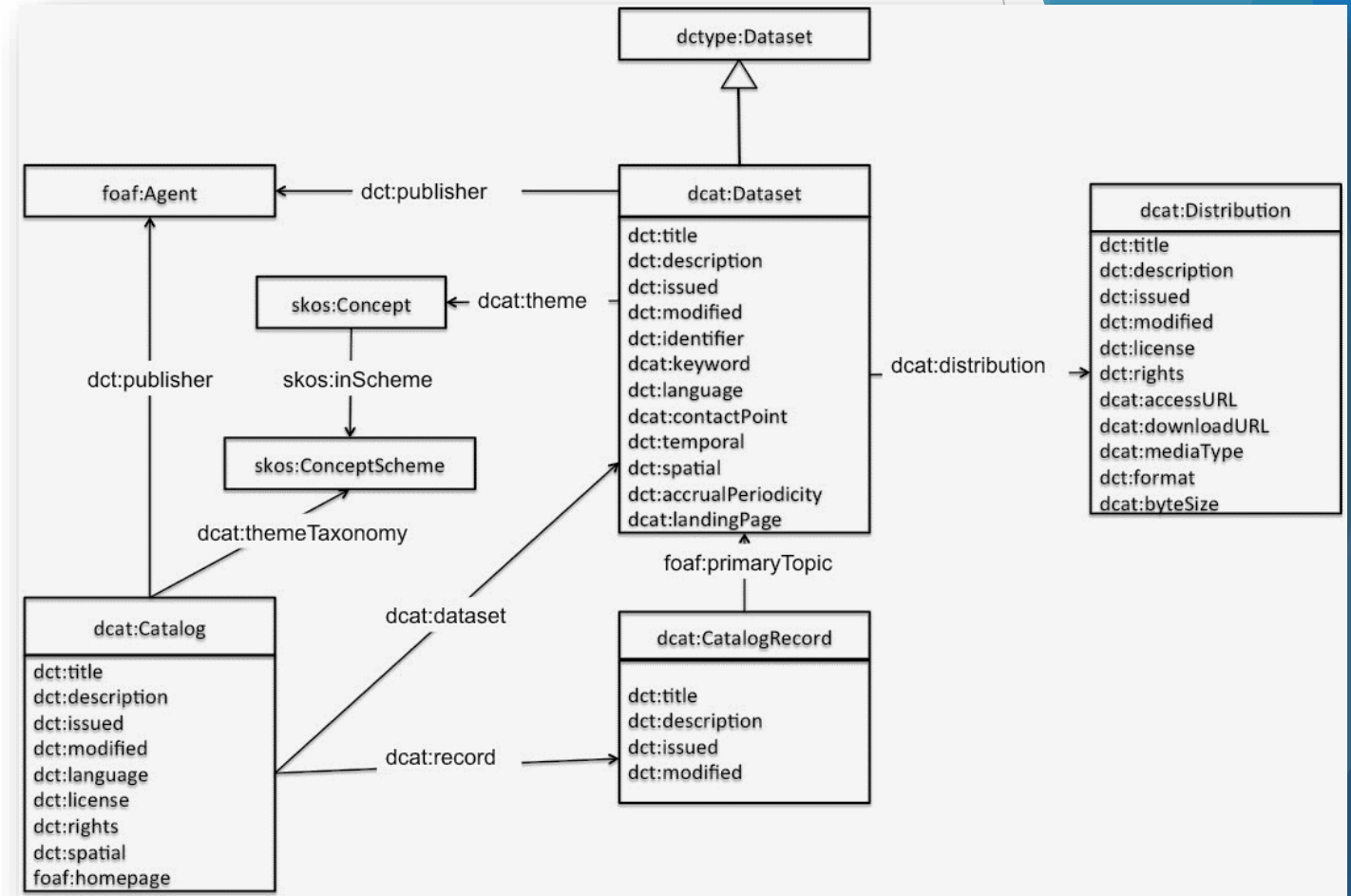
# Implementación del estándar DCAT

- ▶ Siguiendo el paradigma de la Web semántica (Linked Data), cualquier recurso, sea un objeto físico o lógico, debe de ser descrito de dos formas complementarias: en formato legible para las personas -habitualmente en documentos HTML- y en un formato procesable por las máquinas, a través de metadatos expresados mediante RDF.
- ▶ La forma legible para las personas estará conformada por las diferentes fichas de los conjuntos de datos y sus distribuciones que son accesibles a través de la sección de catálogo del portal de datos abiertos.
- ▶ La representación en RDF será manejada por los servidores que gestionan los URI que los identifican a través de mecanismos de negociación de contenido en función de la cabecera del agente de usuario que realiza la petición.
- ▶ La representación semántica de los recursos de información reutilizables (catálogo, datasets y distribuciones) se basa en el vocabulario DCAT, desarrollado por el World Wide Web Consortium (W3C) y que permite la estandarización en la definición de catálogos de documentos y recursos de información y su interoperabilidad. DCAT es un vocabulario en RDF diseñado para facilitar la interoperabilidad entre catálogos de datos publicados en la Web.
- ▶ Especificación de W3C, Resource Description Framework (RDF): <http://www.w3.org/RDF/>
- ▶ Especificación DCAT: [http://www.w3.org/2011/gld/wiki/Data\\_Catalog\\_Vocabulary](http://www.w3.org/2011/gld/wiki/Data_Catalog_Vocabulary)

## 2.5 Metadatos

### DCAT - Metadatos

- ▶ DCAT recomienda el uso de todos los metadatos del vocabulario, aunque cada iniciativa de datos abiertos adopta los requeridos según su estrategia de gestión, adaptando el estándar a sus necesidades según la disponibilidad de metadatos esperada.
- ▶ En la siguiente imagen se detalla brevemente el significado de los metadatos más utilizados según el estándar.



## 2.5 Metadatos

### DCAT - Metadatos de catálogo (I)

METADATO	DESCRIPCIÓN	PROPIEDAD
Clase Catálogo	Colección de metadatos sobre datasets	dcat:Catalog
Nombre	Breve título o nombre dado al catálogo de datos.	dct:title
Descripción	Resumen descriptivo del catálogo de datos.	dct:description
Identificador	Referencia URI para identificar el catálogo.	dct:identifier
Fecha de creación	Fecha de publicación inicial del catálogo.	dct:issued
Fecha de actualización	Fecha en la que se modificó por última vez el catálogo (se añade, elimina o modifica un documento o recurso de información).	dct:modified
Idioma(s)	Idioma(s) en el(los) que se proporciona la información del catálogo.	dct:language
Página web	Dirección web de acceso al catálogo de datos (acceso para el público).	foaf:homepage
Órgano publicador	Entidad que publica el catálogo.	dct:publisher
Cobertura geográfica	Ámbito geográfico cubierto por el catálogo.	dct:spatial
Temáticas	Totalidad de materias incluidas en el catálogo.	dcat:themeTaxonomy
Términos de uso	Referencia a los términos de uso generales del catálogo.	dct:license
Derechos	Derechos de reutilización del catálogo (no datasets)	dct:rights
Documento(s) y recurso(s) de información	Lista de cada uno de los documentos y recursos de información del catálogo.	dcat:dataset
Registro	Registro del catálogo que es parte del propio catálogo	dcat:record
Tamaño del catálogo	Número total de documentos y recursos de información inventariados en el catálogo.	dct:extent

## DCAT - Metadatos de catálogo (II)

METADATO	DESCRIPCIÓN	PROPIEDAD
Clase dataset	Colección de datos disponibles para la reutilización	dcat:Dataset
Nombre	Nombre o título del documento o recurso de información.	dct:title
Descripción	Descripción detallada del documento o recurso de información.	dct:description
Temática(s)	Temática o materia primaria del documento o recurso de información.	dcat:theme
Etiqueta(s)	Etiqueta(s) textual(es) que permiten categorizar libremente el documento o recurso de información.	dcat:keyword
Identificador	URI que identifica al documento o recurso de información.	dct:identifier
Fecha de creación	Fecha de creación del documento o recurso de información.	dct:issued
Fecha de última actualización	Última fecha conocida en la que se modificó o actualizó el contenido del documento o recurso de información.	dct:modified
Frecuencia de actualización	Periodo de tiempo aproximado entre actualizaciones del documento o recurso de información, si hubiera.	dct:accrualPeriodicity
Idioma(s)	Idioma(s) en el(los) que se encuentra la información del documento o recurso de información.	dc:language
Organismo que expone y publica los datos	Organismo que publica el documento o recurso de información.	dct:publisher
Contacto	Información de contacto (mail o vcard) que puede ser utilizada para indicar incidencias sobre el dataset o envío de comentarios.	dcat:contactPoint
Condiciones de uso	Recurso que describe las condiciones de uso o licencia específica aplicable al propio documento o recurso de información.	dct:license
Cobertura geográfica	Ámbito geográfico cubierto por el documento o recurso de información.	dct:spatial
Cobertura temporal	Fecha de inicio, fin y la duración del periodo cubierto por el documento o recurso de información.	dct:temporal
Vigencia del recurso	Fecha de validez de un documento o recurso de información o en la que se estima una modificación o actualización de su contenido.	dct:valid
Recurso(s) relacionado(s)	Enlaces a recursos relacionados con el documento o recurso de información (información sobre los propios datos, material audiovisual, etc.).	dct:references
Normativa	Normativa relativa al documento o recurso de información. Es un enlace a un documento legal.	dct:conformsTo
Distribución(es)	Referencia a los recursos que identifican los volcados del documento o recurso de información en sus posibles formatos.	dcat:distribution

## 2.5 Metadatos

### DCAT - Metadatos de las distribuciones

METADATO	DESCRIPCIÓN	PROPIEDAD
Clase Distribución	Representa la disponibilidad específica de un dataset	dcat:Distribution
Identificador	URI que identifica a la distribución.	dct:identifier
Nombre	Breve título o nombre dado a la distribución.	dct:title
URL de acceso	URL que permite el acceso al volcado o consulta de los documentos o recursos de información.	dcat:accessURL
Formato	Formato en que se encuentra representado el documento o recurso de información.	dcat:mediaType
Tamaño	Tamaño aproximado del documento o recurso de información.	dcat:byteSize
Información adicional sobre el formato	Enlace(s) a documento(s) relacionado(s) con el formato, donde se indica el formato, el esquema utilizado para su representación u otra información técnica sobre cómo acceder a los documentos o recursos de información.	dct:relation

Un catálogo de datos abiertos contiene un número determinado de datasets que a su vez, proporcionan acceso a un conjunto de distribuciones de datos reutilizables:





# Referencias para la implementación de DCAT

- ▶ La iniciativa del Reino Unido, publica una guía para usuarios, publicadores y gestores con los detalles técnicos sobre el soporte de DCAT.  
[http://guidance.data.gov.uk/dcat\\_fields.html](http://guidance.data.gov.uk/dcat_fields.html)
- ▶ El proyecto Open Data de USA, publica el documento Project Open Data Metadata Schema v1.1 que define las normas de aplicación del estándar para los datos albergados en *agency.gov/data*.  
<https://project-open-data.cio.gov/v1.1/schema/>

# Referencias para la implementación de DCAT

- ▶ La Norma Técnica de Interoperabilidad (NTI), que rige en España para el desarrollo de la Ley sobre reutilización de la información del sector público, establece los principios para llevar a cabo una metodología de exposición de información pública reutilizable, desde la identificación, selección y descripción de la información, la elección de formatos adecuados y como aplicar unos términos de uso apropiados. En el Anexo III de la NTI se describen los distintos metadatos asociados con el catálogo y los documentos y recursos de información incluidos en él.

<http://datos.gob.es/saber-mas?q=node/2676>

- ▶ La especificación DCAT-AP -perfil de aplicación para portales Open Data en Europa- basada en DCAT, cuyo uso básico es hacer posible la interoperabilidad entre conjuntos de datos albergados en diferentes portales Open Data, cumple con el objetivo de la identificación de los elementos esenciales y atributos del DCAT en el contexto europeo; definición de los vocabularios a utilizar en el proyecto y determinación del mínimo necesario de metadatos para el intercambio entre los diferentes portales del continente. DCAT-AP también distingue entre metadatos de uso obligatorio, recomendado y opcional.

<https://joinup.ec.europa.eu/node/69559>

## 2.6 Identificación de los datos

### Desafíos

- ▶ ¿Cómo podemos diseñar URIs y gestionarlos para asegurar su persistencia?
- ▶ ¿Cómo podemos proveer reutilizaciones del mismo URI para los mismos conjuntos de datos?

### Mejores prácticas

- ▶ Utilizar URIs persistentes para identificar conjuntos de datos
- ▶ Utilizar URIs persistentes como identificadores dentro de los conjuntos de datos
- ▶ Atribuir URIs a versiones y series de conjuntos de datos

## 2.6 Identificación de la información reutilizable (URI)

- ▶ Los Identificadores de Recursos Uniforme (URI) ofrecen la capacidad de identificación de forma unívoca de cualquier recurso de información a través de la Web y garantizan la persistencia, lo que permite establecer identificadores universales para cualquier pieza de información que se desee exponer. Esencialmente, son direcciones web que identifican cualquier elemento que se representa en la Web.
- ▶ Un esquema de URIs se comporta como una guía de estilo de referencia a la hora de diseñar identificadores para los elementos de información reutilizables. En este esquema se indicarán las partes que conforman los URI en una plataforma de datos abiertos. Los URIs siguen patrones claramente definidos, lo que permite ofrecer identificadores semánticos e intuitivos, coherencia para la ampliación del sistema y facilidad para la interoperabilidad.

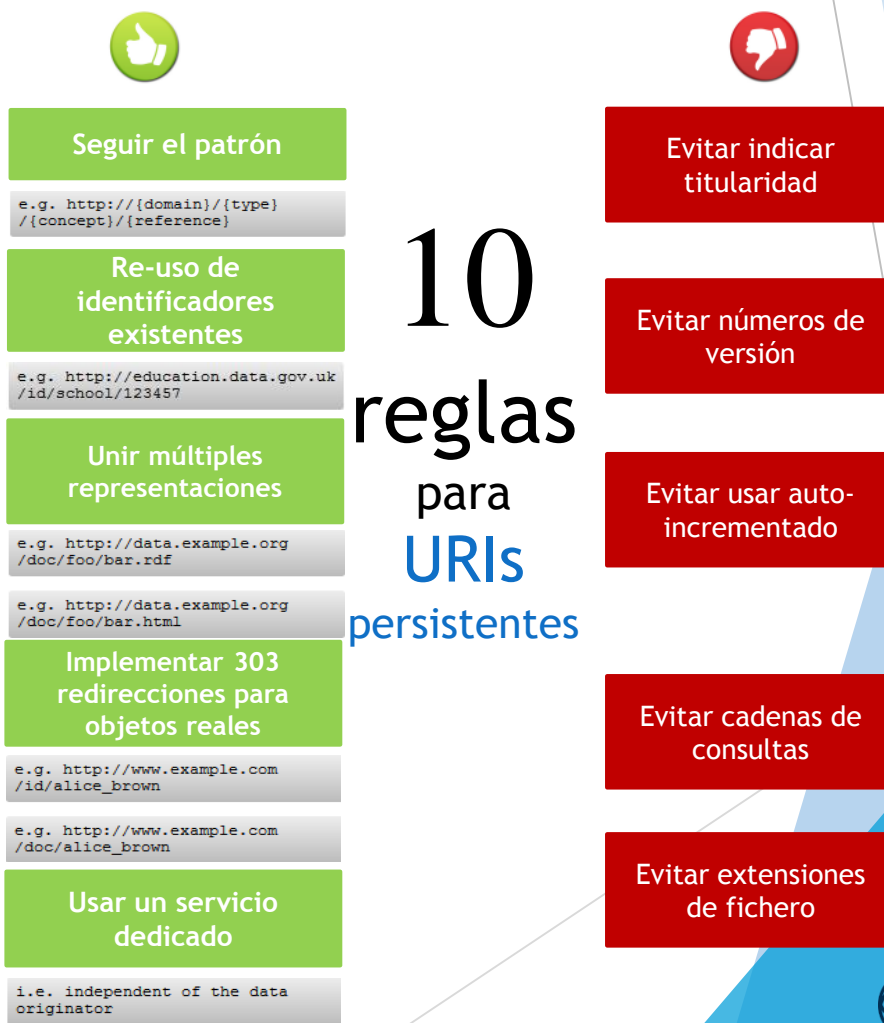
# 2.6 Identificación de la información reutilizable (URI)

## Implementación de URI (I)

- ▶ En la literatura disponible sobre el tratamiento de URIs existe una serie de recomendaciones técnicas que se deben seguir para la implementación de un esquema que se resumen en el gráfico siguiente:

Fuente: Study on Persistent URIs de Phil Archer:  
<http://philarcher.org/diary/2013/uripersistence/>

### Las 10 recomendaciones para la implementación de un esquema



## Implementación de URI (I)

Para mantener la coherencia y el mantenimiento posterior del esquema de URI se sugiere aplicar las siguientes reglas para normalizar las distintas partes que componen los URI:

- ▶ Seleccionar identificadores alfanuméricos cortos únicos, que sean representativos, intuitivos y semánticos.
- ▶ Usar siempre minúsculas, salvo en los casos en los que se utilice el nombre de la clase o concepto. Habitualmente, los nombres de las clases se representan con el primer carácter de cada palabra en mayúsculas.
- ▶ Eliminar todos los acentos, diéresis y símbolos de puntuación. Como excepción puede usarse el guión (-).
- ▶ Eliminar conjunciones y artículos en los casos de que el concepto a representar contenga más de una palabra.
- ▶ Puede usarse el guion (-) como separador entre palabras.
- ▶ Evitar en la medida de lo posible la abreviatura de palabras, salvo que la abreviatura sea intuitiva.

Los términos que componen los URI deberán ser legibles e interpretables por el mayor número de personas posible, por lo que se utilizará el idioma del país o cualquiera de las lenguas oficiales.

## Implementación de URI (III)

El patrón para diseñar los URIs se compone de elementos, cada uno de ellos con un significado específico, que juntos componen la ruta de un URI. Estos son: sector, carácter de la información, tipo de representación, dominio o temática y los conceptos o recursos específicos.

Dentro de la composición de un URI se especifican por el siguiente orden:

▶ `http://{base}/{carácter}[/{sector}][/{dominio}][/{concepto}][.{ext}]`

La **base** de los URI incluirá información básica sobre la procedencia de los datos, que representará un espacio dedicado por parte del promotor de la iniciativa de datos abiertos para albergar su plataforma de publicación de datos. En el caso de la iniciativa de datos abiertos del Gobierno de Colombia, la base de los identificadores es:

▶ <http://www.datos.gov.co/>

## Referencias para la implementación de URIs

- ▶ El documento de mejores prácticas para la construcción de URIs editado por W3C es la referencia a seguir para su implementación en iniciativas *Government Linked Data*.

[http://www.w3.org/2011/gld/wiki/223\\_Best\\_Practices\\_URI\\_Construction](http://www.w3.org/2011/gld/wiki/223_Best_Practices_URI_Construction)

- ▶ La Norma Técnica de Interoperabilidad publica la guía de aplicación de la norma donde se detallan los esquemas de URIs a utilizar en la implementación de las iniciativas de datos abiertos de la Administración Pública española

<http://datos.gob.es/saber-mas?q=node/2676>

- ▶ El portal [legislation.gov.uk](http://www.legislation.gov.uk) implementado por Archivos Nacionales del Reino Unido, describe el esquema de URIs usado para el acceso a la API del portal.

<http://www.legislation.gov.uk/developer/uris>

- ▶ Estudio de URIs persistentes: relación de buenas prácticas y casos de uso de implementación de URIs:

<http://philarcher.org/diary/2013/uripersistence/>



## Identificadores fundamentales

- ▶ En una iniciativa de datos abiertos es posible encontrar patrones distintos de URI, en función del **carácter** de la información que representan. La mayoría de las iniciativas disponen, al menos cuatro elementos básicos: el portal Web con información relativa a la iniciativa, el catálogo de los conjuntos de datos disponibles para su reutilización con sus fichas descriptivas asociadas y las distribuciones de los datasets.
- ▶ A modo de ejemplo, se relacionan a continuación los URIs relativos al carácter de la información y su interpretación:

URI	Interpretación
<a href="http://datos.gov.co/portal">http://datos.gov.co/portal</a>	Identificador del sitio Web del portal de datos abiertos de Colombia
<a href="http://datos.gov.co/catalogo">http://datos.gov.co/catalogo</a>	Identificador del catálogo de datos de la iniciativa Open Data de Colombia
<a href="http://datos.gov.co/catalogo/conjunto-de-datos">http://datos.gov.co/catalogo/conjunto-de-datos</a>	Identificador de un conjunto de datos incluido en el catálogo de datos abiertos de Colombia
<a href="http://datos.gov.co/datos/distribucion-en-CSV-de-conjunto-de-datos">http://datos.gov.co/datos/distribucion-en-CSV-de-conjunto-de-datos</a>	Identificador de una distribución específica, por ejemplo, un volcado de datos en formato CSV, de un conjunto de datos del catálogo de datos abiertos de Colombia

# Identificadores de temas de conjuntos de datos (I)

- ▶ La relación de un recurso de información con un **sector** adecuado, proporcionará al usuario la confianza de conocer el tipo de información que está manejando. Cada documento o recurso de información utilizará el tema más representativo o, en su defecto, alguno que se pueda considerar común a las características de la información representada.
- ▶ Para la implementación de este conjunto de identificadores es necesario definir una taxonomía o clasificación que servirá de base común para componer el esquema de URI y para la categorización de los catálogos de recursos de información pública y sus registros, según los metadatos especificados en base al estándar DCAT.
- ▶ A continuación se propone un catálogo de identificadores y URIs asociados para sectores primarios cuyo uso puede servir de inspiración en la iniciativa de datos de Colombia.

## Identificación de temas de conjuntos de datos (II)

SECTOR	IDENTIFICADOR	URI
Medio Rural Incluye: Agricultura, Ganadería, Pesca y Silvicultura.	medio-rural-pesca	<a href="http://datos.gov.co/kos/sector-publico/sector/medio-rural-pesca">http://datos.gov.co/kos/sector-publico/sector/medio-rural-pesca</a>
Salud Incluye: Sanidad.	salud	<a href="http://datos.gov.co/kos/sector-publico/sector/salud">http://datos.gov.co/kos/sector-publico/sector/salud</a>
Sector público Incluye: Presupuestos, Organigrama institucional, Legislación interna, Función pública.	sector-publico	<a href="http://datos.gov.co/kos/sector-publico/sector/sector-publico">http://datos.gov.co/kos/sector-publico/sector/sector-publico</a>
Seguridad Incluye: Protección civil, Defensa.	seguridad	<a href="http://datos.gov.co/kos/sector-publico/sector/seguridad">http://datos.gov.co/kos/sector-publico/sector/seguridad</a>
Sociedad y bienestar Incluye: Participación ciudadana, Marginación, Envejecimiento Activo, Autonomía personal y Dependencia, Invalidez, Jubilación, Seguros y Pensiones, Prestaciones y Subvenciones.	sociedad-bienestar	<a href="http://datos.gov.co/kos/sector-publico/sector/sociedad-bienestar">http://datos.gov.co/kos/sector-publico/sector/sociedad-bienestar</a>
Transporte Incluye: Comunicaciones y Tráfico.	transporte	<a href="http://datos.gov.co/kos/sector-publico/sector/transporte">http://datos.gov.co/kos/sector-publico/sector/transporte</a>
Turismo Incluye: Alojamientos, Hostelería, Gastronomía.	turismo	<a href="http://datos.gov.co/kos/sector-publico/sector/turismo">http://datos.gov.co/kos/sector-publico/sector/turismo</a>
Urbanismo e infraestructuras Incluye: Saneamiento público, Construcción (infraestructuras, equipamientos públicos).	urbanismo-infraestructuras	<a href="http://datos.gov.co/kos/sector-publico/sector/urbanismo-infraestructuras">http://datos.gov.co/kos/sector-publico/sector/urbanismo-infraestructuras</a>

## Identificación de temas de conjuntos de datos (III)

SECTOR	IDENTIFICADOR	URI
Ciencia y tecnología Incluye: Innovación, Investigación, I+D+i, Telecomunicaciones, Internet y Sociedad de la Información.	ciencia-tecnologia	<a href="http://datos.gov.co/kos/sector-publico/sector/ciencia-tecnologia">http://datos.gov.co/kos/sector-publico/sector/ciencia-tecnologia</a>
Comercio Incluye: Consumo.	comercio	<a href="http://datos.gov.co/kos/sector-publico/sector/comercio">http://datos.gov.co/kos/sector-publico/sector/comercio</a>
Cultura y ocio Incluye: Tiempo libre.	cultura-ocio	<a href="http://datos.gov.co/kos/sector-publico/sector/cultura-ocio">http://datos.gov.co/kos/sector-publico/sector/cultura-ocio</a>
Demografía Incluye: Inmigración y Emigración, Familia, Mujeres, Infancia, Mayores, Padrón.	demografia	<a href="http://datos.gov.co/kos/sector-publico/sector/demografia">http://datos.gov.co/kos/sector-publico/sector/demografia</a>
Deporte Incluye: Instalaciones deportivas, Federaciones, Competiciones.	deporte	<a href="http://datos.gov.co/kos/sector-publico/sector/deporte">http://datos.gov.co/kos/sector-publico/sector/deporte</a>
Economía Incluye: Deuda, Moneda y Banca y finanzas.	economia	<a href="http://datos.gov.co/kos/sector-publico/sector/economia">http://datos.gov.co/kos/sector-publico/sector/economia</a>
Educación Incluye: Formación.	educacion	<a href="http://datos.gov.co/kos/sector-publico/sector/educacion">http://datos.gov.co/kos/sector-publico/sector/educacion</a>
Empleo Incluye: Trabajo, Mercado laboral.	empleo	<a href="http://datos.gov.co/kos/sector-publico/sector/empleo">http://datos.gov.co/kos/sector-publico/sector/empleo</a>

## Identificación de temas de conjuntos de datos (IV)

SECTOR	IDENTIFICADOR	URI
Hacienda Incluye: Impuestos.	hacienda	<a href="http://datos.gov.co/kos/sector-publico/sector/hacienda">http://datos.gov.co/kos/sector-publico/sector/hacienda</a>
Industria Incluye: Minería.	industria	<a href="http://datos.gov.co/kos/sector-publico/sector/industria">http://datos.gov.co/kos/sector-publico/sector/industria</a>
Legislación y justicia Incluye: Registros.	legislación-justicia	<a href="http://datos.gov.co/kos/sector-publico/sector/legislacion-justicia">http://datos.gov.co/kos/sector-publico/sector/legislacion-justicia</a>
Medio ambiente Incluye: Meteorología, Geografía, Conservación fauna y flora.	medio-ambiente	<a href="http://datos.gov.co/kos/sector-publico/sector/medio-ambiente">http://datos.gov.co/kos/sector-publico/sector/medio-ambiente</a>
Vivienda Incluye: Mercado inmobiliario, Construcción (viviendas).	vivienda	<a href="http://datos.gov.co/kos/sector-publico/sector/vivienda">http://datos.gov.co/kos/sector-publico/sector/vivienda</a>
Energía Incluye: Fuentes renovables	energia	<a href="http://datos.gov.co/kos/sector-publico/sector/energia">http://datos.gov.co/kos/sector-publico/sector/energia</a>

## Identificadores de cobertura geográfica (I)

- ▶ Los identificadores correspondientes a los recursos geográficos a nivel nacional, o territorial permitirán identificar y definir la cobertura geográfica de un recurso de información. Para resolver estos identificadores se puede optar por reutilizar un servicio externo como *GeoNames*.
- ▶ Cada característica en *GeoNames* está representado como un sitio de recursos identificados por un identificador URI estable. Este identificador URI proporciona acceso comprensible por las personas, a un Wiki en la página HTML o una descripción de los recursos RDF utilizando el vocabulario *GeoNames*.

GeoNames

## Identificadores de cobertura geográfica (II)

- ▶ A modo de ejemplo, la identificación de recursos geográficos correspondientes a la división administrativa de Colombia, tendría la siguiente forma :
  - ▶ Departamento de Boyacá: <http://www.geonames.org/3688536>
  - ▶ Departamento del Valle del Cauca: <http://www.geonames.org/3666313>
  - ▶ Municipio de Cali:  
<http://www.geonames.org/3687926>
  - ▶ Municipio de Bogotá:  
<http://www.geonames.org/3688689>
- ▶ Un listado de la división administrativa de Colombia se puede encontrar en:  
<http://www.geonames.org/CO/administrative-division-colombia.html>
- ▶ Como alternativa al uso del vocabulario Geonames, es posible desarrollar una taxonomía propia de recursos geográficos del país que sirva como referencia común a todos los organismos publicadores.
- ▶ En el contexto colombiano, las entidades referencia en materia de manejo de recursos geográficos son DANE e IGAC. Si alguna de estas dos entidades utiliza una codificación oficial expresada en forma de URIs, se aconseja utilizarla.

## Identificadores de cobertura geográfica (III)

- ▶ Para identificar un organismo gestor de catálogos y publicador de datos se puede desarrollar una taxonomía basada en las referencias definidas en algún Directorio Común de Entidades o Unidades Orgánicas, gestionado por el Gobierno de Colombia. Cada entidad estará referenciada como organismo colaborador del Directorio Común, por tanto están disponibles los identificadores unívocos de los organismos radicados en el territorio, al menos a nivel departamental y municipal. Es factible desagregar todos los entes de la administración, con lo que sería posible identificar publicadores a nivel orgánico.
- ▶ Los URI que identifican a cada entidad dependen del identificador {ID} alfanumérico unívoco del organismo, unidad orgánica u oficina -a incluir en la parte final del URI- y tendrá la forma siguiente:
- ▶ <http://datos.gov.co/recurso/sector-publico/org/Organismo/{ID}>

ORGANISMO	URI
MinTIC	<a href="http://datos.gov.co/recurso/sector-publico/org/Organismo/M01300012">http://datos.gov.co/recurso/sector-publico/org/Organismo/M01300012</a>
MinEducación	<a href="http://datos.gov.co/recurso/sector-publico/org/Organismo/M01300026">http://datos.gov.co/recurso/sector-publico/org/Organismo/M01300026</a>



## 2.7 Formatos de Datos

### Desafíos

- ▶ ¿Qué tipos de formatos de datos deben ser considerados cuando se publican datos en la web?

### Mejores prácticas

- ▶ Utilizar formatos de datos estándar legibles por máquina
- ▶ Proveer datos en múltiples formatos

## 2.8 Vocabularios de Datos

### Desafíos

- ▶ ¿Cómo podemos utilizar los vocabularios de datos ya existentes con vista a una interoperabilidad semántica?
- ▶ ¿Cómo se puede diseñar un nuevo vocabulario?

### Mejores prácticas

- ▶ Utilizar términos estándar para proveer datos y metadatos
- ▶ Reutilizar vocabularios
- ▶ Seleccionar el nivel correcto de formalización

## 2.9 Licencias de Datos

### Desafíos

- ▶ ¿Cómo podemos hacer con que las licencias de datos sean legibles por máquina?
- ▶ ¿Cómo hacer con que la información de licencias sobre datos publicados en la Web sea publicada?

### Mejores prácticas

- ▶ Proveer información sobre la licencia de utilización de los datos abiertos

## 2.10 Origen de los Datos

### Desafíos

- ▶ ¿Cómo se puede publicar el origen de los datos en la web?

### Mejores prácticas

- ▶ Proveer información sobre la proveniencia de la información

## 2.11 Calidad de los Datos

### Desafíos

- ▶ ¿Por qué es importante publicar información sobre la calidad de los datos?

### Mejores prácticas

- ▶ Proveer información, en los propios metadatos, sobre los procesos de aseguramiento de la calidad que han seguido los datos para su publicación, facilita significativamente el proceso de selección de los datos, lo que aumenta las posibilidades de reutilización. Independientemente de las peculiaridades específicas de dominio, la calidad de los datos debe ser documentada y los problemas de calidad conocidos deben mencionarse expresamente en los metadatos.

## 2.12 Datos Sensibles

### Desafíos

- ▶ ¿Cómo se puede publicar información respetando los derechos de los usuarios?

### Mejores prácticas

- ▶ Los datos sensibles tienen gran importancia para garantizar la privacidad de los individuos en situaciones donde la apertura de datos puede afectar la seguridad de las personas. La privacidad puede incluir: nombres, direcciones, correos electrónicos, números de identificación de las personas, placas de vehículos, huellas, números de tarjetas de crédito, fechas de nacimiento, lugar de nacimiento, información genética, registros de salud.
- ▶ Proveer referencias de datos no abiertos, o abiertos bajo diferentes restricciones, con referencia de cómo se puede acceder a los datos.
- ▶ Los proveedores de datos deben identificar todos los datos sensibles, evaluar la exposición al riesgo y determinar la utilización de los datos, el público-objetivo de los datos y las políticas de utilización. Además de eso, deben obtener aprobaciones y determinar medidas de seguridad necesarias para proteger a los datos. Estas medidas de seguridad pueden incluir autenticación y utilización de HTTPS.

## 2.13 Acceso a los datos

### Desafíos

- ▶ ¿Qué tipo de acceso se debe proporcionar datos en la web?
- ▶ ¿Qué requerimientos deben ser considerados para la apertura de datos en la web?

### Mejores prácticas

- ▶ Permitir la descarga de forma masiva de conjuntos de datos
- ▶ Utilizar interfaces Web estándar
- ▶ Proveer datos y recursos con diferentes formatos
- ▶ Proveer datos en tiempo real
- ▶ Proveer datos actualizados
- ▶ Utilizar un API para proveer datos
- ▶ Documentar el API

## 2.14 Versionamiento de datos

### Desafíos

- ▶ ¿Cómo se puede mantener la trazabilidad y gestionar los cambios realizados en los conjuntos de datos?

### Mejores prácticas

- ▶ Proveer información sobre el versionamiento
- ▶ Proveer histórico de versiones de los conjuntos de datos
- ▶ Evitar cambios del API que rompan el código de los desarrolladores
- ▶ Comunicar cambios a los desarrolladores
- ▶ La identificación y versionado de archivos se implementa utilizando un esquema de URIs. Consultar la explicación detallada en en slide 26.



## 2.15 Conservación de datos

### Desafíos

- ▶ ¿Cómo archivar datos de manera confiable?

### Mejores prácticas

- ▶ Evaluar la cobertura del conjunto de datos
- ▶ Utilizar un formato de creación de series confiable para archivar datos
- ▶ Actualizar el estado de los identificadores, conectando los recursos archivados con los actuales

## 2.16 Agregar valor a los datos

### Desafíos

- ▶ ¿Cómo agregar valor a los datos abiertos?

### Mejores prácticas

- ▶ Agregar valor a los datos abiertos se puede realizar a través de la generación de nuevos metadatos

## 2.17 Retroalimentación

### Desafíos

- ▶ ¿Cómo se puede recoger la retroalimentación de los consumidores de datos?

### Mejores prácticas

- ▶ Recoger feedback de los usuarios
- ▶ Proveer información sobre el feedback de los usuarios

## 2.18 Oportunidades del uso de estándares de Datos Abiertos

En el taller sobre estandarización de datos abiertos se discutieron, desde una perspectiva general, las oportunidades del uso de estándares en datos abiertos. Los resultados de esa discusión fueron los siguientes:

- ▶ Permitir integración de sistemas
- ▶ Compartir datos entre organizaciones
- ▶ Permitir comparar organizaciones
- ▶ Ampliar tamaños de mercados para desarrolladores
- ▶ Escalar aplicaciones
- ▶ Reutilizar código de proyectos TI
- ▶ Interoperabilidad
- ▶ Re-uso de los datos
- ▶ Hacer indicadores
- ▶ Transformación - crear otros datos
- ▶ Internacionalización
- ▶ Creación de herramientas para uso del público

3

# Casos de éxito de aplicación de estándares de datos abiertos



# CitySDK\*

## Datos abiertos en ciudades

- ▶ CitySDK fue un proyecto financiado por la Comisión Europea que tenía como objetivo facilitar el acceso a datos abiertos por desarrolladores en ocho ciudades europeas.
- ▶ El proyecto ha trabajado en los temas de participación ciudadana, turismo y movilidad, desarrollando o adoptando estándares para la apertura de datos en las ciudades.
- ▶ CitySDK ha sido desarrollado para crear un mercado más amplio para aplicaciones y servicios basados en datos abiertos, facilitando la transferencia y la reutilización de servicios entre ciudades.
- ▶ En muchos casos, los datos de la ciudad tienen una gran variedad de orígenes, como la Alcaldía a las empresas de transporte, pero también los Gobiernos Regionales y Nacionales.
- ▶ Un desarrollador trabajando en varias ciudades con CitySDK sabe que puede reutilizar su código de ciudad en ciudad, ahorrando tiempo y dinero.

\*Más información sobre CitySDK en el siguiente enlace: <http://www.citysdk.eu/about-the-project/>

54

# CitySDK\*

## Datos abiertos en ciudades

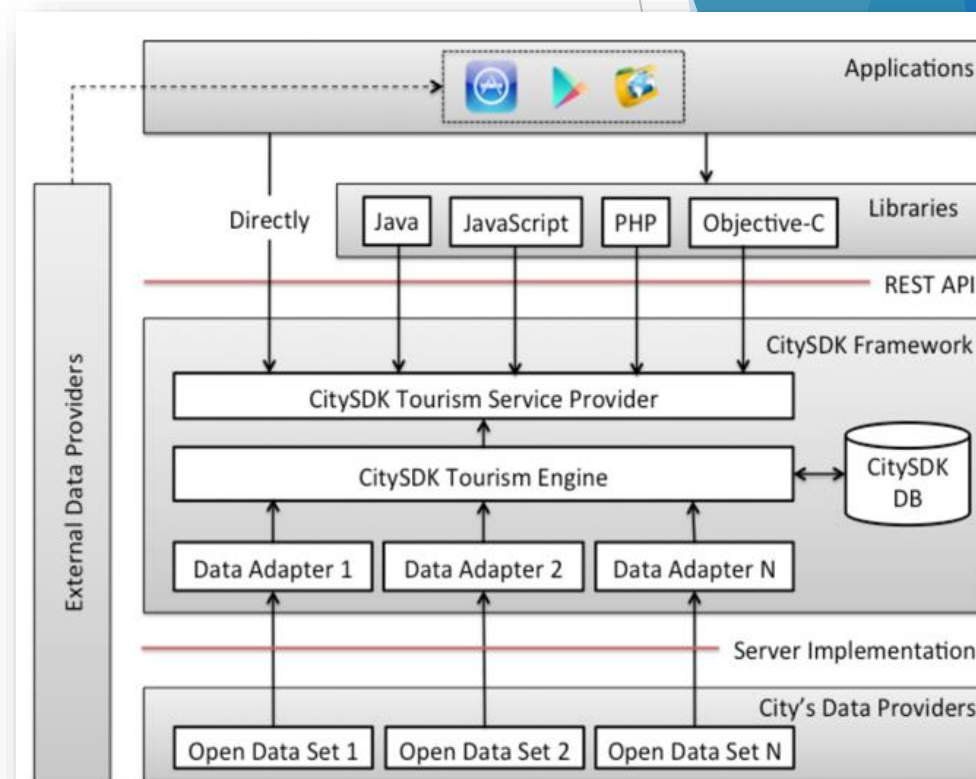
La propuesta de valor de CitySDK para los desarrolladores es la siguiente:

- ▶ Información sobre fuentes de datos relevantes en las áreas de Participación Ciudadana, Movilidad y Turismo.
- ▶ Mercados más amplios con necesidades mínimas de adaptación de las apps
- ▶ APIs agnósticas en relación a la tecnología
- ▶ Recursos como librerías de código, apps y plataformas disponibles como ejemplo
- ▶ Intercambio conocimiento y experiencias con la comunidad de desarrolladores que utiliza CitySDK

\*Más información sobre CitySDK en el siguiente enlace: <http://www.citysdk.eu/about-the-project/>

# CitySDK para Turismo

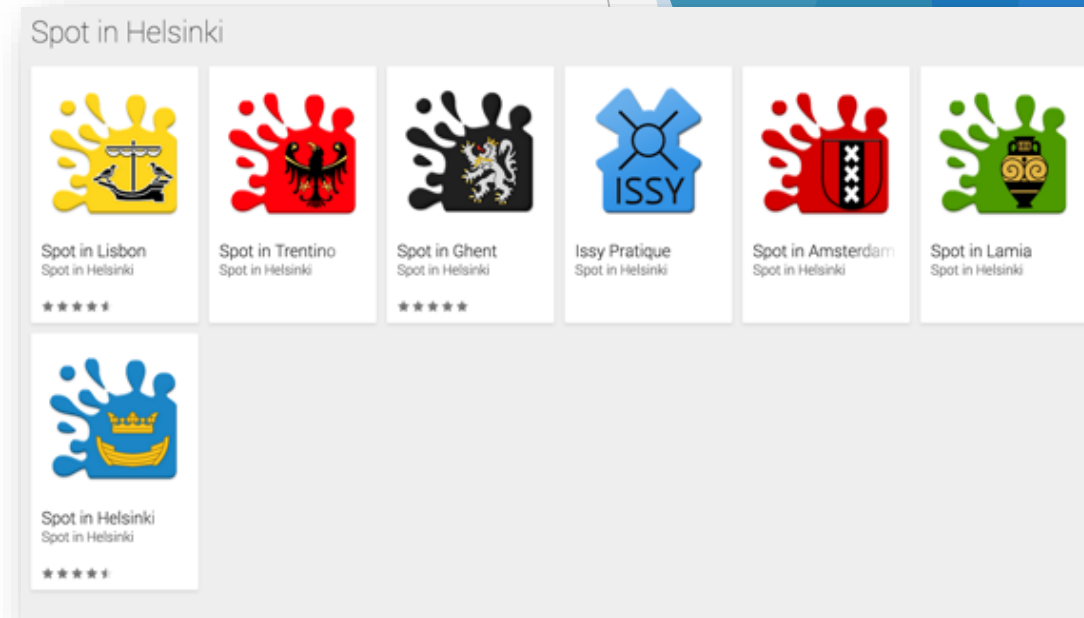
- ▶ En el dominio de turismo, CitySDK trabajó en la creación de una API para compartir información turística, generando un estándar para compartir estas informaciones con base en el estándar para de puntos de interés del W3C.
- ▶ El proyecto ha desarrollado un API que recoge datos provenientes de los sistemas de información de la ciudad y los comparte según el estándar definido por el proyecto.
- ▶ La API de CitySDK permite acceso a puntos de interés, rutas y eventos:
  - ▶ **Puntos de Interés** - permite acceder a todos los puntos de interés de la ciudad
  - ▶ **Eventos** - permite acceder a todos los eventos de la agenda de la ciudad
  - ▶ **Rutas** - descripción de un grupo de Puntos de Interés





# City SDK: Spot in ...

- ▶ Un caso de éxito del proyecto ha sido la aplicación Spot In.
- ▶ Spot In ha sido desarrollada inicialmente en Helsinki con el objetivo de compartir información turística para usuarios de teléfonos Android.
- ▶ Los desarrolladores de Spot In han participado en la creación del estándar de CitySDK y, en pocos meses, han logrado crear la misma aplicación para las ciudades de Lisboa (Portugal) , Trentino (Italia), Ghent (Bélgica), Issy les Moulineaux (Francia), Amsterdam (Holanda) y Lamia (Grecia).
- ▶ La transferencia del app de ciudad en ciudad ocurrió sin grandes dificultades, porque se recurrió al uso del estándar de CitySDK para compartir los datos.
- ▶ Este es un claro ejemplo de ampliación del mercado de los desarrolladores a través del uso de estándares para la apertura de datos abiertos.



Fuente: <https://play.google.com/store/apps/developer?id=Spot+in+Helsinki>

# CitySDK para Participación Ciudadana

- ▶ El estándar de participación ciudadana de CitySDK está basado en un estándar ya existente: Open311.
- ▶ Este estándar define una API para publicitar problemas reportados por los ciudadanos en la ciudad. Este tipo de aplicaciones es cada vez más común en ciudades y la utilización de estándares como Open311 permite la integración de soluciones y bajar los costes de implementación.
- ▶ El estándar le permite a cada ciudad establecer cuales son las categorías de problemas, permite el envío de nuevos problemas con indicación de localización y archivos media; permite buscar problemas por categoría o fecha y permite hacer seguimiento del estado de los problemas reportados.
- ▶ La información sobre los problemas reportados por los ciudadanos se encuentra disponible en los formatos XML y JSON.
- ▶ Ciudades como Nueva York y Chicago han adoptado este estándar.
- ▶ Un listado de todas las ciudades que utilizan Open311 está disponible en [http://wiki.open311.org/GeoReport\\_v2/Servers/](http://wiki.open311.org/GeoReport_v2/Servers/)

# Proyecto Citadel

El proyecto Citadel, financiado por la Comisión Europea, tenía como objetivo crear lineamientos para la apertura de datos en ciudades. Este proyecto ha generado un conjunto de aplicaciones que están disponibles en línea y pueden ser utilizadas libremente.

- **Open Data Commons:** una filosofía de integración de componentes de sistemas de información en ciudades para la apertura y integración de datos con bajo coste.
- **Herramienta de generación de apps:** que permite la creación de apps móviles con

gran facilidad a través del uso de tecnología HTML5

- **Citadel converter:** que permite convertir datasets desde cualquier formato para un formato compatible con las apps del generador de apps.



<http://www.citadelonthemove.eu>

4

# Recomendaciones y hoja de ruta

# City SDK:

## Recomendaciones para la implementación exitosa

Del proyecto CitySDK ha resultado un conjunto de recomendaciones para una implementación exitosa que pueden ser tenidas en cuenta para la utilización de estándares de datos abiertos en Colombia. A continuación se proponen los siguientes pasos para una implementación en Colombia:

- 1) **Familiarizar al equipo con CitySDK y revisar referencias** - el primer paso es compartir información, casos de éxito y referencias entre pares. El equipo debe tener noción de la importancia del uso de estándares desde el principio.
- 2) **Presentar beneficios de CitySDK internamente** - el público interno es el primer cliente.
- 3) **Evaluar los sistemas, procesos y fuentes de datos** - para planificar la implementación de estándares, es importante tener un diagnóstico de los sistemas, procesos de negocio y fuentes de datos existentes en la organización.
- 4) **Definir y planificar la integración de la API y la implementación** - tener un plan de proyecto para la integración del estándar.
- 5) **Implementar y testear**
- 6) **Capacitar** - capacitación interna para los proveedores de datos y entrenamiento de los posibles usuarios de los datos.
- 7) **Documentar el proceso** - generar documentación de todo el proceso y de cómo se han interpretado los conjuntos de datos a la luz del estándar.
- 8) **Establecer términos del servicio y licencia de datos** - estos puntos son muy importantes para la reutilización de los datos (como se presentó en la sección 2 de esta Guía).
- 9) **Lanzamiento**
- 10) **Mantener y desarrollar**

# Hoja de Ruta

#	Recomendación	T1	T2	T3	T4	Responsables
1	Definir un conjunto de taxonomías necesarias para implementar el mecanismo de identificación de recursos de información. Será necesario implementar:					
1.1	- Taxonomía de sectores					MinTIC + AGN + DANE
1.2	- Taxonomía de órganos publicadores					MINTIC + AGN + DANE
1.3	- Taxonomía de la división administrativa del territorio de Colombia, como alternativa al uso del vocabulario Geonames.					MINTIC + AGN + DANE
2	Implementar un esquema de URIs habilitando mecanismos de negociación de contenido para de-referenciar el acceso a los recursos disponibles.					MinTIC + Proveedor Plataforma DA
3	Implementar el vocabulario de propiedades de DCAT sobre la plataforma de datos abiertos					MinTIC + Proveedor Plataforma DA
4	Incluir en la Guía para la apertura de datos, las indicaciones pertinentes para la correcta identificación y descripción de recursos de información.					MinTIC
5	Difusión y sensibilización sobre buenas prácticas sobre catalogación e identificación de datos abiertos					MinTIC + Entidades publicadoras

# Recomendaciones Datos Abiertos

A continuación se relatan una serie de recomendaciones necesarias para implementar la estrategia de catalogación e identificación de recursos de información:

1. Definir un conjunto de taxonomías necesarias para implementar el mecanismo de identificación de recursos de información. Será necesario implementar:
  - ▶ Taxonomía de sectores
  - ▶ Taxonomía de órganos publicadores
  - ▶ Taxonomía de la división administrativa del territorio de Colombia, como alternativa al uso del vocabulario *Geonames*.
2. Implementar un esquema de URIs habilitando mecanismos de negociación de contenido para de-referenciar el acceso a los recursos disponibles.
3. Implementar el vocabulario de propiedades de DCAT sobre la plataforma de datos abiertos
4. Incluir en la Guía para la apertura de datos, las indicaciones pertinentes para la correcta identificación y descripción de recursos de información.