Taller GBIF.ES: Estandarización, documentación y publicación de datos de biodiversidad en colecciones

16 – 17 de septiembre de 2026, Institut Botànic, Barcelona

Metadatos y licencias de uso de los datos

Francisco Pando
GBIF.ES

XARXA DE MUSEUS DE CIÈNCIES NATURALS DE CATALUNYA

institut botànic
Centre mixt
CSIC
Ajuntament de Barcelona

GBIF.ES

Elementos que hacen la ciencia abierta

- Metadatos
- Estándares
- Identificadores
- Licencias
- Citación
- El cómo se sirve el dato

Imagen generada con IA

Metadatos (la parte del contexto)

FORBES > INNOVATION

Data Is Worthless Without Context

Adam Blue Forbes Councils Member
Forbes Technology Council COUNCIL POST | Membership (Fee-Based)

https://www.forbes.com/councils/forbestechcouncil/2022/01/03/data-is-worthless-without-context/

Data, without context, is essentially useless. Raw data lacks meaning until processed and organized into information, which then allows for insights to be drawn and decisions made

Datos sin contexto no son datos

1. **¿Quién?**:
    1. ¿Quién creó los datos?
    2. ¿Quién es responsable de mantener los datos?
2. **¿Qué?**:
    1. ¿Cuál es el contenido o tema de los datos?
    2. ¿En qué formato están los datos?
3. **¿Cuándo?**:
    1. ¿Cuándo se crearon los datos?
    2. ¿Cuándo se actualizaron o modificaron por última vez?
4. **¿Dónde?**:
    1. ¿Dónde se recopilaron o crearon los datos?
    2. ¿Dónde están almacenados los datos?
5. **¿Por qué?**:
    1. ¿Por qué se crearon los datos?
    2. ¿Por qué son importantes o útiles?
6. **¿Cómo?**:
    1. ¿Cómo se recopilaron o generaron los datos?
    2. ¿Cómo están estructurados y organizados los datos?
7. **Calidad de los datos**:
    1. ¿Son precisos y fiables los datos?
    2. ¿Cuáles son las limitaciones y problemas con los datos?
8. **Acceso y uso**:
    1. ¿Quién puede acceder a los datos?
    2. ¿Existen restricciones en el uso de los datos?
9. **Relaciones**:
    1. ¿Cómo se relacionan estos datos con otros de datos?
    2. ¿Hay enlaces a información adicional relevante?

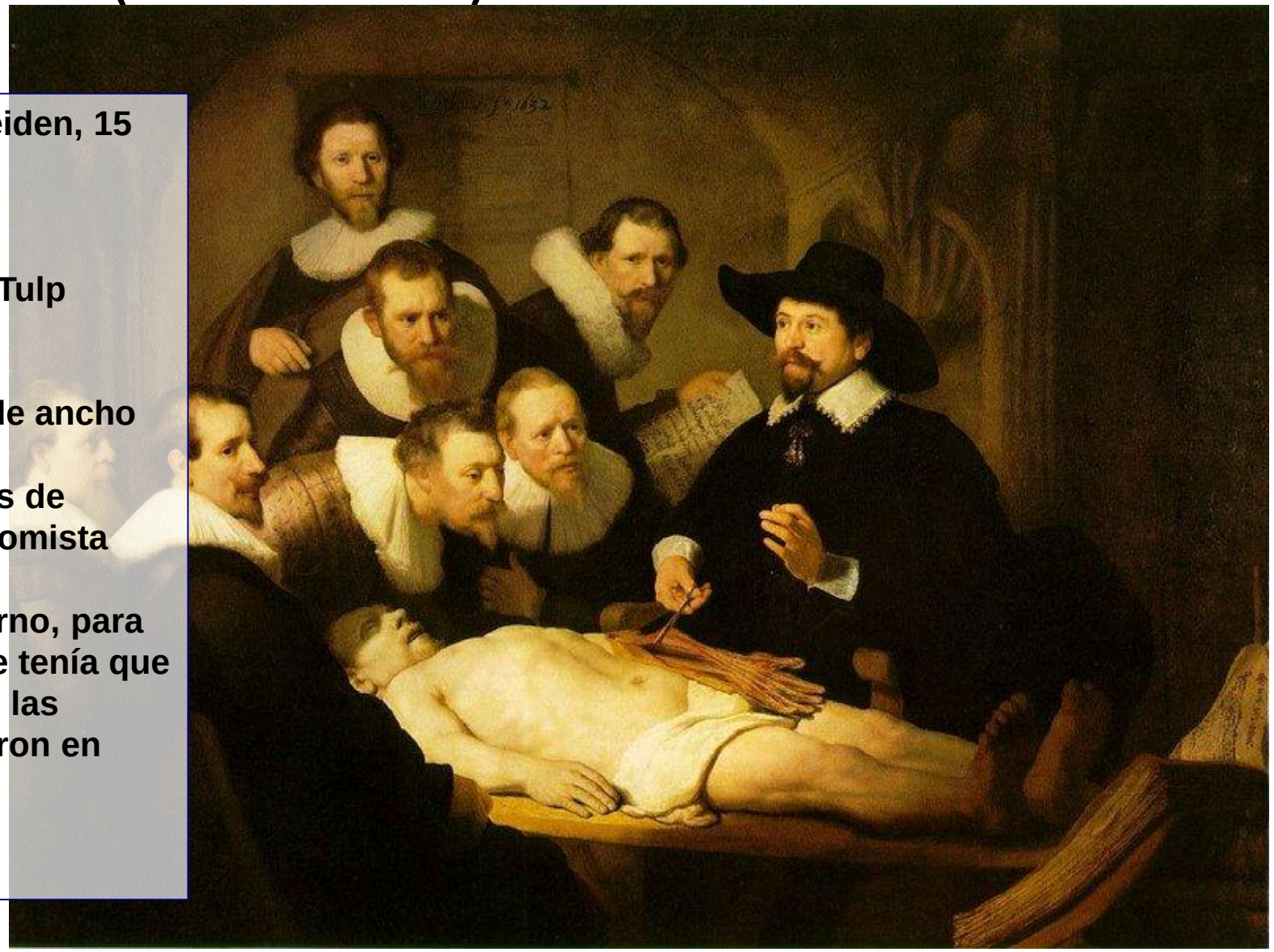
Conocimiento explícito e implícito

- El conocimiento –contexto si se prefiere– es algo muy difícil de aprehender pero vital para un uso óptimo de los datos
- La documentación proporciona contexto.
- Las tecnologías informáticas están dando acceso a datos e información a una escala inédita hasta ahora, haciendo que esta sea accesible sin importar la distancia o el ámbito donde la información fue generada.
- Este beneficioso fenómeno provoca sin embargo una descontextualización de los datos.
- Esta situación hace que una buena documentación [metadatos] sea más importante que nunca.

Hay que luchar contra el síndrome de “todo el mundo sabe eso” (y no se documenta)

Documentación de la información (metadatos)

- **Rembrandt Harmenszoon van Rijn (Leiden, 15 de julio de 1606 – † Amsterdam, 4 de octubre de 1669**
- **1632**
- **Lección de anatomía del Dr. Nicolaes Tulp**
- **Mauritshuis, La Haya)**
- **pintura al óleo**
- **169,5 centímetros de alto y 216,5 cm de ancho**
- **El evento puede fecharse el 16 de enero de 1632: la cofradía de cirujanos de Ámsterdam, de la que Tulp era el anatomista oficial de la ciudad, permitía sólo una disección pública al año, en invierno, para mejor conservación del cuerpo; y éste tenía que ser de un criminal ejecutado. Por ello, las clases de anatomía con disección fueron en el siglo XVII actos poco frecuentes y espectaculares**



“Los metadatos son un elemento fundamental en la generación del conocimiento”

…y en su descubrimiento y recuperación

Y en su uso

- Generar, descubrir, recuperar, usar

$\Rightarrow$ Intercambiar, integrar, reutilizar

$\Rightarrow$ estándares

Dublin Core

- The Dublin Core metadata element set se convirtió en norma ISO 15836/2003 en febrero de 2003.
- Creada por una comunidad de individuos de diferentes procedencias y disciplinas, de organizaciones de todo el mundo que incluyen tanto al sector público como al privado.
- Define 15 elementos básicos y generales para describir un recurso (un programa, una página Web, un mapa,. ..).

- http://dublincore.org/documents/dces/

Dublin Core Metadata Element Set

Contenido:

1. **Título:** el nombre dado a un recurso, habitualmente por el autor.
2. **Claves:** (Palabras clave)
3. **Descripción:** una descripción textual del recurso. Puede ser un resumen en el caso de un documento o una descripción del contenido en el caso de un documento visual.
4. **Fuente:** secuencia de caracteres usados para identificar unívocamente un trabajo a partir del cual proviene el recurso actual.
5. **Lengua:** lengua/s del contenido intelectual del recurso.
6. **Relación:** es un identificador de un segundo recurso y su relación con el recurso actual. Este elemento permite enlazar los recursos relacionados y las descripciones de los recursos.
7. **Cobertura:** espacial y/o temporal del contenido intelectual del recurso.

Propiedad Intelectual:

8. **Autor o Creador:** la persona o organización responsable de la creación del contenido intelectual del recurso.
9. **Editor:** la entidad responsable de hacer que el recurso se encuentre disponible.
10. **Otros Colaboradores:** una persona u organización que haya tenido una contribución intelectual significativa,
11. **Derechos:** son una referencia (p.e. URL) sobre términos y condiciones de acceso a un recurso.

Instanciación:

12. **Fecha:** una fecha de publicación del recurso en su forma actual.
13. **Tipo del Recurso:** la categoría del recurso.
14. **Formato:** es el formato usado para identificar el software y el hardware para mostrar el recurso.
15. **Identificador del Recurso:** secuencia de caracteres utilizados para identificar unívocamente un recurso. URL, URN, ISBN, DOI,…

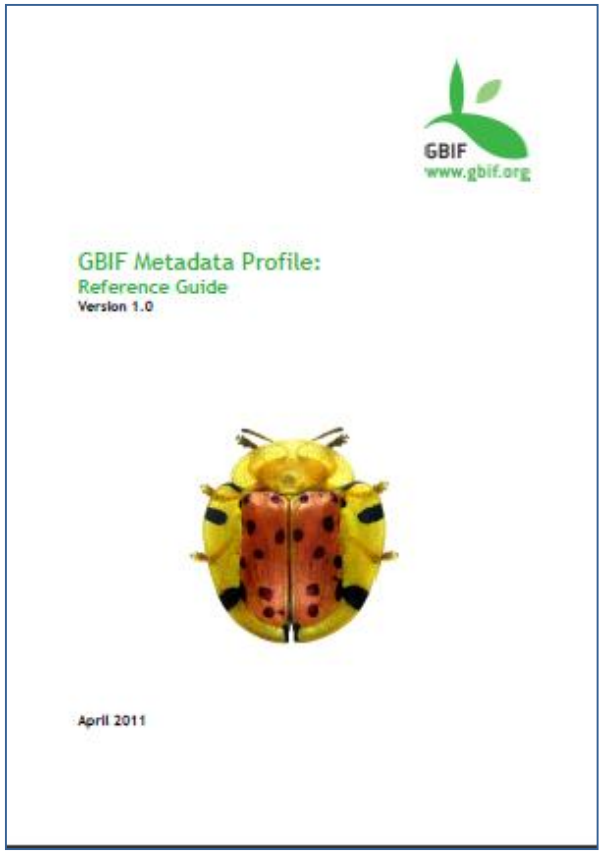
Metadatos de conjuntos de datos

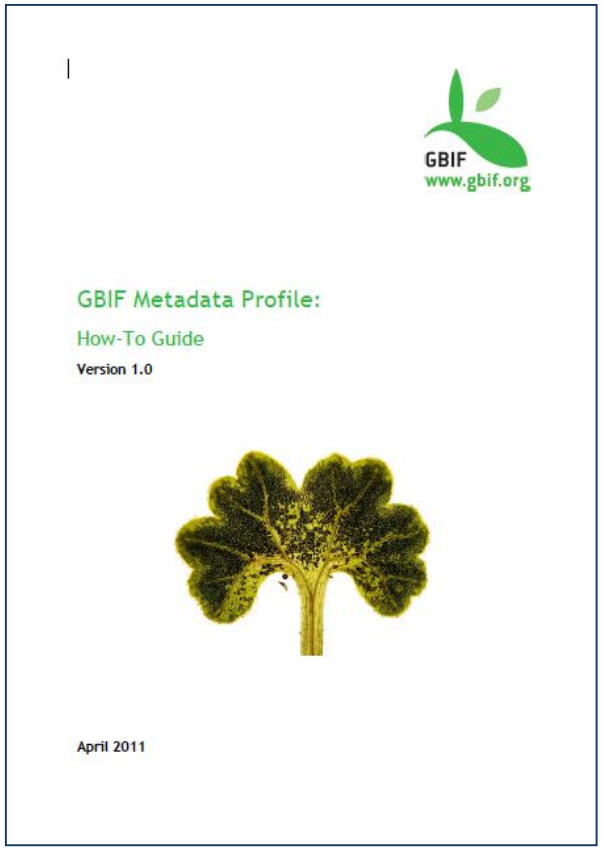
<u>GMP (GBIF Metadata profile):</u>

EML (Ecological Metadata Language)

- Dataset (Resource)
- Project
- People and Organisations
- Keyword Set (General Keywords)
- Coverage
  - o Taxonomic Coverage
  - o Geographic Coverage
  - o Temporal Coverage
- Methods
- Intellectual Property Rights

- Additional Metadata
  - Latimer Core (Natural Collections Descriptions Data)





https://www.gbif.org/sites/default/files/gbif_resource/resource-80640/gbif_metadata_profile_guide_en_v1.pdf

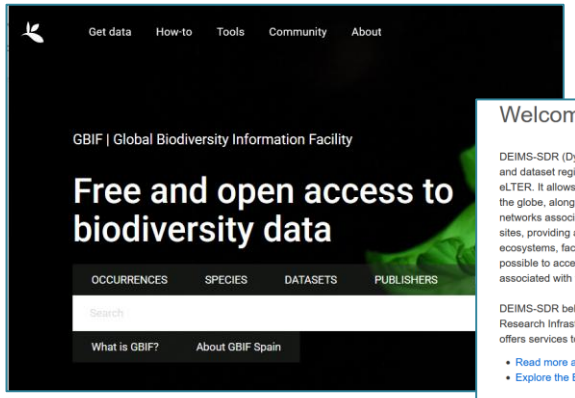
https://ipt.gbif.org/manual/en/ipt/latest/gbif-metadata-profile

https://ipt.gbif.org/manual/es/ipt/latest/gbif-metadata-profile

EML (Ecological Metadata Language)

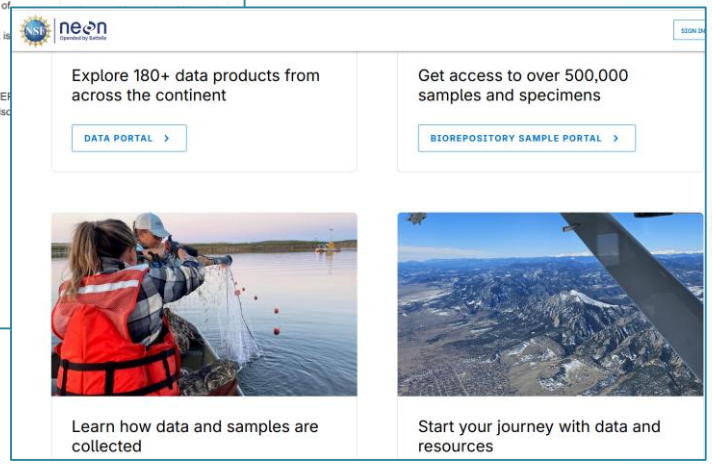
>> estándar de código abierto y basado en XML, diseñado para documenta datos de investigación en ciencias ambientales y ecológicas. Permite describir el contenido, alcance espacial y temporal, métodos y protocolos los datos

https://www.gbif.org

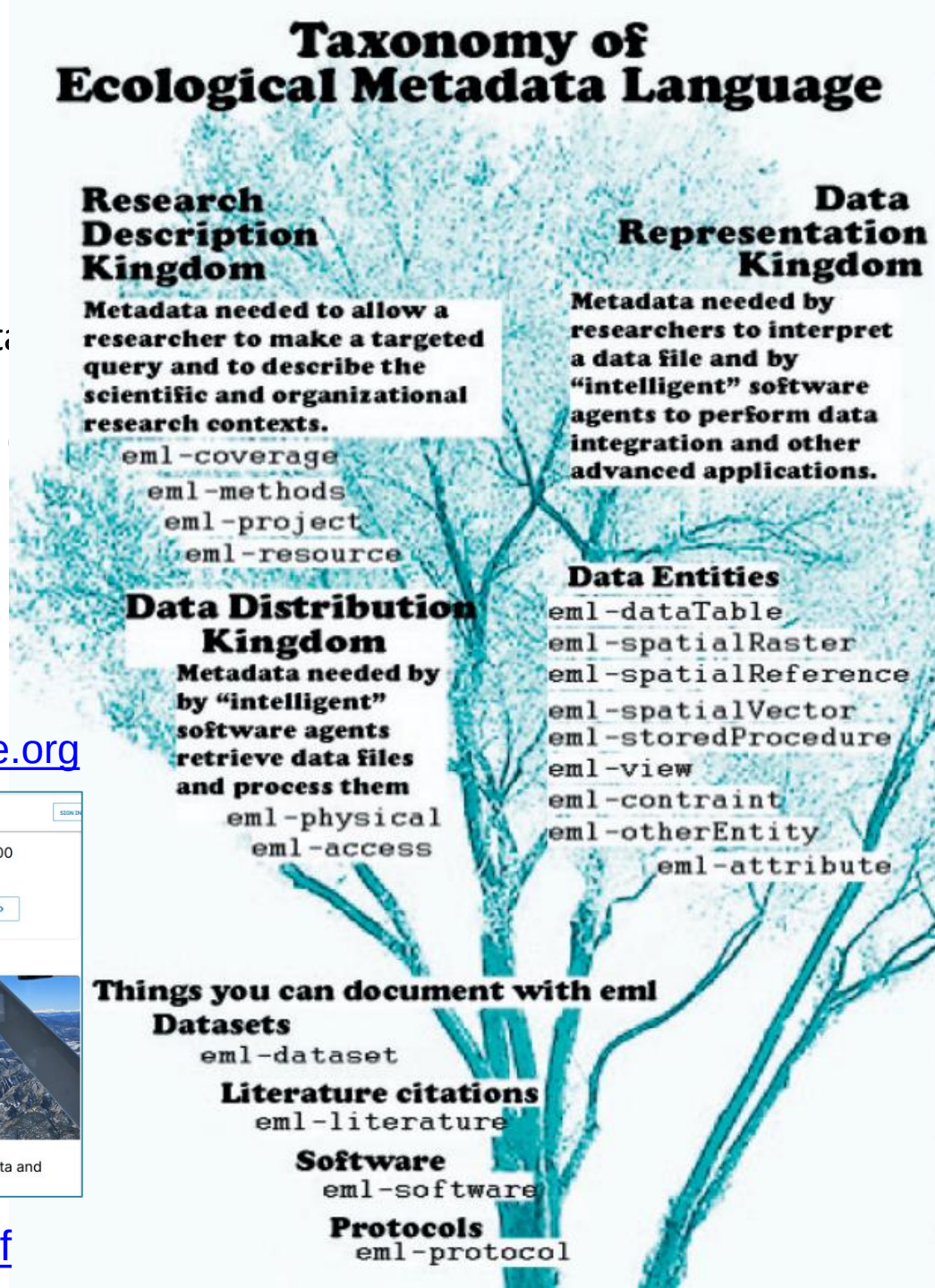


https://deims.org

www.neonscience.org



https://eml.ecoinformatics.org/
https://lternet.edu/wp-content/uploads/2010/12/EMLHandbook-2.pdf



Metadatos en GBIF en el mundo real: IPT

ipt.gbif.es/resource?r=geranium_in_america

Introduction

Inicio Gestión de recursos Administración Acerca de

...propietario de los derechos de este trabajo es Real Jardín Botánico (CSIC). Esta obra está bajo una licencia Creative Commons de Atribución/Reconocimiento (CC-BY 4.0).

- Descripción
- Registros
- Versiones
- ¿Cómo referenciar?
- Derechos
- Registro GBIF
- Palabras clave
- Contactos
- Cobertura geográfica
- Cobertura taxonómica
- Cobertura temporal
- Métodos de muestreo
- Datos de la colección
- Referencias bibliográficas
- Metadatos adicionales

Registro GBIF

Este recurso ha sido registrado en GBIF con el siguiente UUID: 26d72d3b-4544-4645-aa56-27aa8a669c6f. Real Jardín Botánico (CSIC) publica este recurso y está registrado en GBIF como un publicador de datos avalado por GBIF Spain.

Palabras clave

Occurrence; Specimen; Botany; Plant Sciences; Taxonomy; Species distribution; America; inventoryRegional; taxonomicAuthority; Occurrence

Contactos

Carlos Aedo
Proveedor De Contenido • Proveedor De Los Metadatos • Autor • Originador • Punto De Contacto
Researcher
Real Jardin Botanico -CSIC
Claudio Moyano 1
28014 Madrid
Madrid
ES
aedo@rjb.csic.es

Francisco Pando
Proveedor De Los Metadatos • Originador • Punto De Contacto
Researcher
Real Jardin Botanico -CSIC
Claudio Moyano 1
28014 Madrid
Madrid
ES
pando@rjb.csic.es

Magnoliopsida
Geraniales
Geraniaceae
Geranium (Geranio)
Geranium, Robertianum, Erodioidea

1787-01-01 / 2012-01-01

...phic and taxonomic gaps identified during herbarium revision work

Specimens of 128 herbaria were studied. The herbarium acronyms (and number of studied specimens) containing the higher number of relevant specimens were: MO (1493); NY (1181); MA (774); GH (666); F (630); MEXU (504); COL (469); DAO (440); CONC (413); BM (258) Herbarium material was supplemented by specimens collected during nine field trips, which allowed knowing better the variability of characters within populations, and refining species distribution ranges. These expeditions took place in Argentina, Bolivia, Chile, Costa Rica, Ecuador, Peru, United States and Venezuela, and yielded 235 herbarium specimens of Geranium belonging to 46 species.

Studied specimens are listed in sufficient detail to allow revision and act checking. These data are made publicly available. Previous taxonomic sources studied and listed in the monograph. Database details (scientific names, dates, geographic coordinates and places checked for consistency using controlled vocabularies. Further validation procedures (geographic coordinate format, coordinates within country/provincial boundaries, absence of ASCII anomalous characters in the dataset) were performed with DARWIN_TEST

https://ipt.gbif.es/

Metadatos en imágenes

- EXIF
  - En TIFF,JPEG y PSD files
  - No soportado por PNG o GIF y
  - Estandarizado en XML para JPEG 2000 según ISO TC42 (12234-1.4.)
  - Date & time, camera settings (camara model, orientation, aperture…), thumbnail
  - Usado por cámaras fotográficas.
  - version 2.3 Abril 2010

- IPTC-IIM *(Information Interchange Model)*
  - *En JPEG, TIFF*
  - *No soportado por GIF, PNG*
  - http://www.iptc.org
  - *"IPTC headers"* ampliamente adoptado por software para edición

- XMP *(Extensible Metadata Platform)*
  - http://www.adobe.com/products/xmp
  - XML y RDF
  - Información de procesos por los q la imagen pasa. Dispositivos y software
  - Dublin core +Elemenos extensibles + propios
  - En PDF, JPEG, JPEG 2000, GIF, PNG, HTML, TIFF, Adobe Illustrator, PSD, MP3, MP4,…
  - "*IPTC Core Schema for XMP*"

Metadatos geográficos

- ***ISO 19115:2006 "Geographic Information-Metadata"*** :
  - Norma internacional de metadatos, perteneciente a la familia de normas Internacionales ISO 19100 para la Información Geográfica desarrollada por el Comité Técnico 211 (*Geomatic/Geographic Information*) dentro de la Organización de Estandarización Internacional (ISO).
  - Define un modelo, estableciendo un conjunto común de terminología, definiciones y procedimientos de aplicación para los metadatos.
  - Ha sido adoptada como Norma Europea por CEN/TC287 y como Norma Española por AEN/CTN148.

Un poco más allá

- En la actualidad la norma ISO 19115 es la Norma Internacional de Metadatos, pero es:
  - Muy amplia (409 elementos).
  - Muy voluminosa (140 páginas).
  - Muy compleja (descripción items, uso de UML, terminología,etc).
- Muchos de los metadatos definidos en ISO 19115, no son útiles para las necesidades de los organismos cartográficos.
- La iniciativa Dublín Core es muy pobre para crear metadatos relacionados con Información Geográfica.
- Surge la necesidad de crear un Perfil de Metadatos: Conjunto mínimo de metadatos, recomendados para la descripción de los recursos relacionados con la Información Geográfica, basado en Normas de metadatos.
- Se define "**NEM**" como el **Perfil de Metadatos para España**.

Núcleo Español de Metadatos (NEM)

- En el año 2002 el Consejo Superior Geográfico creó el GT IDEE para implementar una IDE en España.

- GT IDEE es un grupo abierto formado por productores de datos geográficos digitales de referencia y temáticos de ámbito nacional, regional y local.

- GT IDEE creó un Subgrupo de Trabajo de Metadatos (SGT2) cuya misión era investigar, analizar en inventariar la situación de cada uno de los organismos cartográficos de España en materia de metadatos.

- Como resultado, una gran deficiencia en materia de metadatos :
  - Falta de conocimiento.
  - Falta de herramientas para la creación de metadatos.

- En Noviembre de 2004, GT IDEE tomó la decisión de crear un Subgrupo de Trabajo del Núcleo Español de Metadatos (SGT NEM).

Do

1 INTRODUCCIÓN

1.1 Metadatos 1
1.2 Normas de Metadatos....3
1.2.1 ISO 19115 “Geographic Information-Metadata....4
1.2.2 Dublin Core Metadata....5
1.3 El Núcleo Español de Metadatos 6
1.4 El ámbito en los Metadatos 7

2 EL NUCLEO ESPAÑOL DE METADATOS (NEM)

2.1 Elementos del Núcleo Español de Metadatos....8
2.2 Tabla resumen de los elementos del NEM 10

3 DESCRIPCIÓN DE LA CALIDAD

3.1 Introducción....12
3.2 Ábito de la Calidad 13
3.3 Metadatos para describir la calidad....14
3.3.1 Componentes Cualitativas de la calidad 14
3.3.2 Componentes Cuantitativas de la calidad 16

4 BIBLIOGRAFÍA....23

5 ANEXO DICCIONARIO DE DATOS

5.1 Introducción....24
5.2 Descripción de elementos Principales del NEM....26
5.3 Descripción de Tipos de Datos 44
5.4 Enumeraciones y Listas de Códigos 123

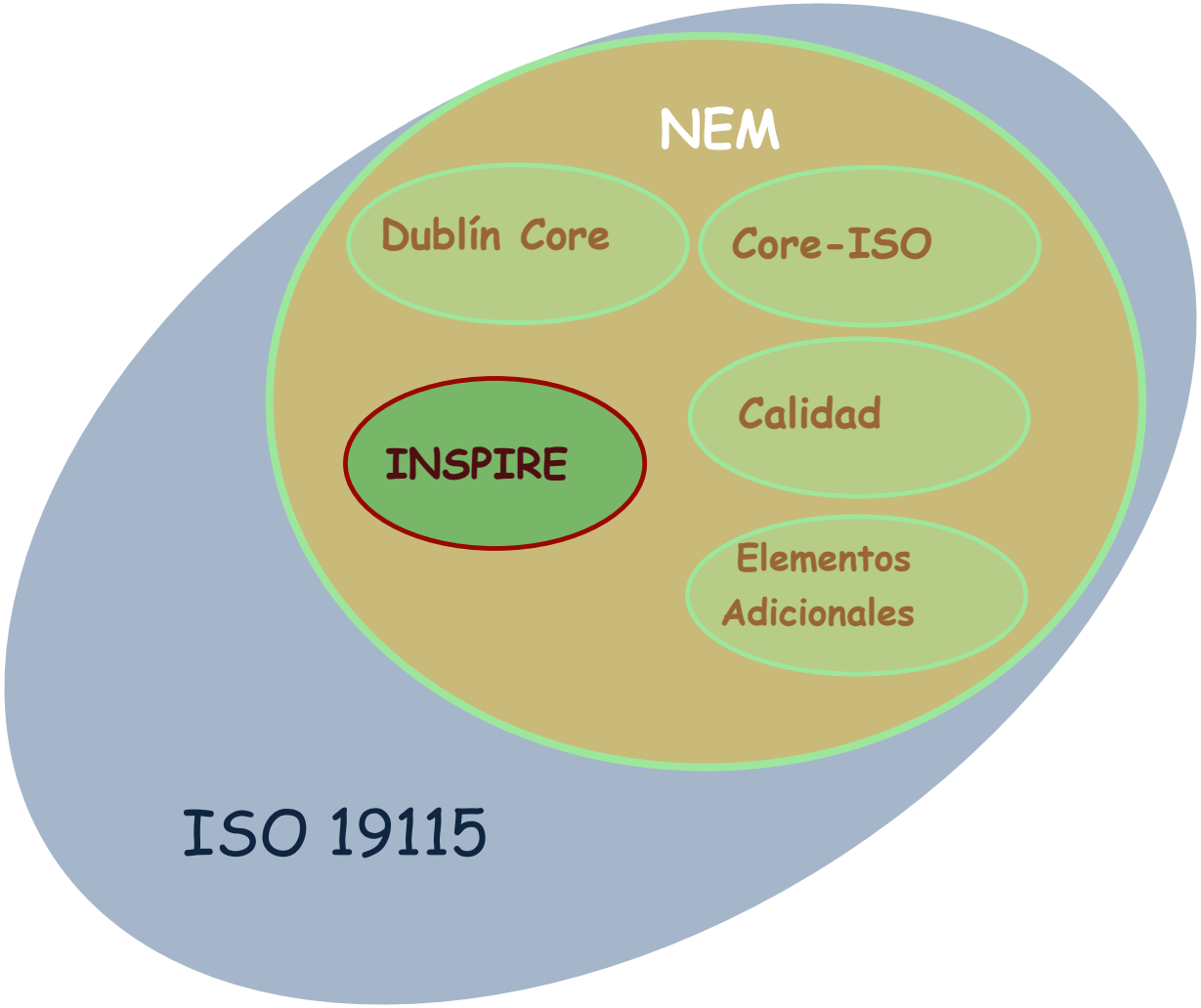
Infraestructura de Datos Espaciales Española		CONSEJO SUPERIOR GEOGRÁFICO	
SGT NEM	Núcleo Español de Metadatos	2005-01-10	Página 1 de 129

IDEE

Núcleo Español de Metadatos
(NEM v1.0)

Título	Núcleo Español de Metadatos (NEM v1.0)
Identificador	SGTNEM_2005_01
Autor	Subgrupo de Trabajo del Núcleo Español de Metadatos
Fecha	2005-01-10
Tema	Descripción del Núcleo Español de Metadatos
Estado	Borrador
Objetivo	Circulación en el SGT NEM para su modificación si procede y aprobación
Descripción	Documento descriptivo detallado del Núcleo Español de Metadatos, conjunto mínimo de metadatos recomendado en España, extraído como perfil informal de la norma ISO19115, que incluye los elementos del Dublin Core Metadata
Contribuciones	Javier Nogueras Iso (Universidad de Zaragoza) José Lázaro Amaro (Gobierno de la Rioja) Marisa López, Manuel Madrid y Michael Gould (Universitat Jaume I) José Masó y Aleix Zabala (Universidad Autónoma de Barcelona) Susana Fontana Ruiz (Trabajos Catastrales, SA) Antonio F. Rodríguez Pascual (Instituto Geográfico Nacional) Alejandra Sánchez Maganto (Instituto Geográfico Nacional)
Fuente	ISO19115 y Dublin Core Metadata Initiative
Difusión	Miembros del GT IDEE
Idioma	Español (Es)
Documentos relacionados	ISO19115 y Dublin Core Metadata Initiative
Periodo de validez	Hasta la próxima versión

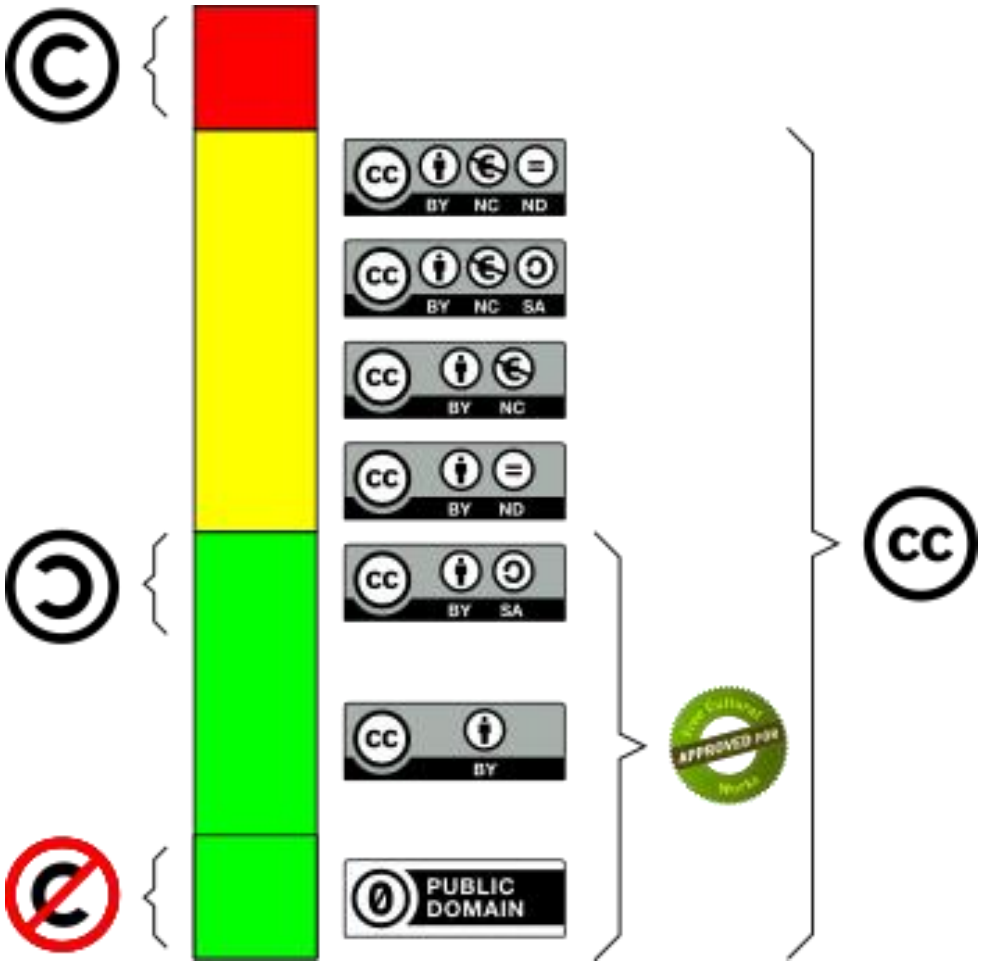
Elementos del NEM 1.1



Metadatos, en resumen

- Los metadatos dan contexto para usar mejor los datos
- Los metadatos se descubre y agregan porque van en formatos estándar
- Los metadatos se crean en origen
- Lo importante de los metadatos es tenerlos, convertirlos a distintos estándares es sencillo

IPR y Copyright



https://ikusimakusi.eus/2018/cc-traffic-light-3-0/CC0

Derechos de propiedad intelectual (Intelectual Property Rights, IPR)

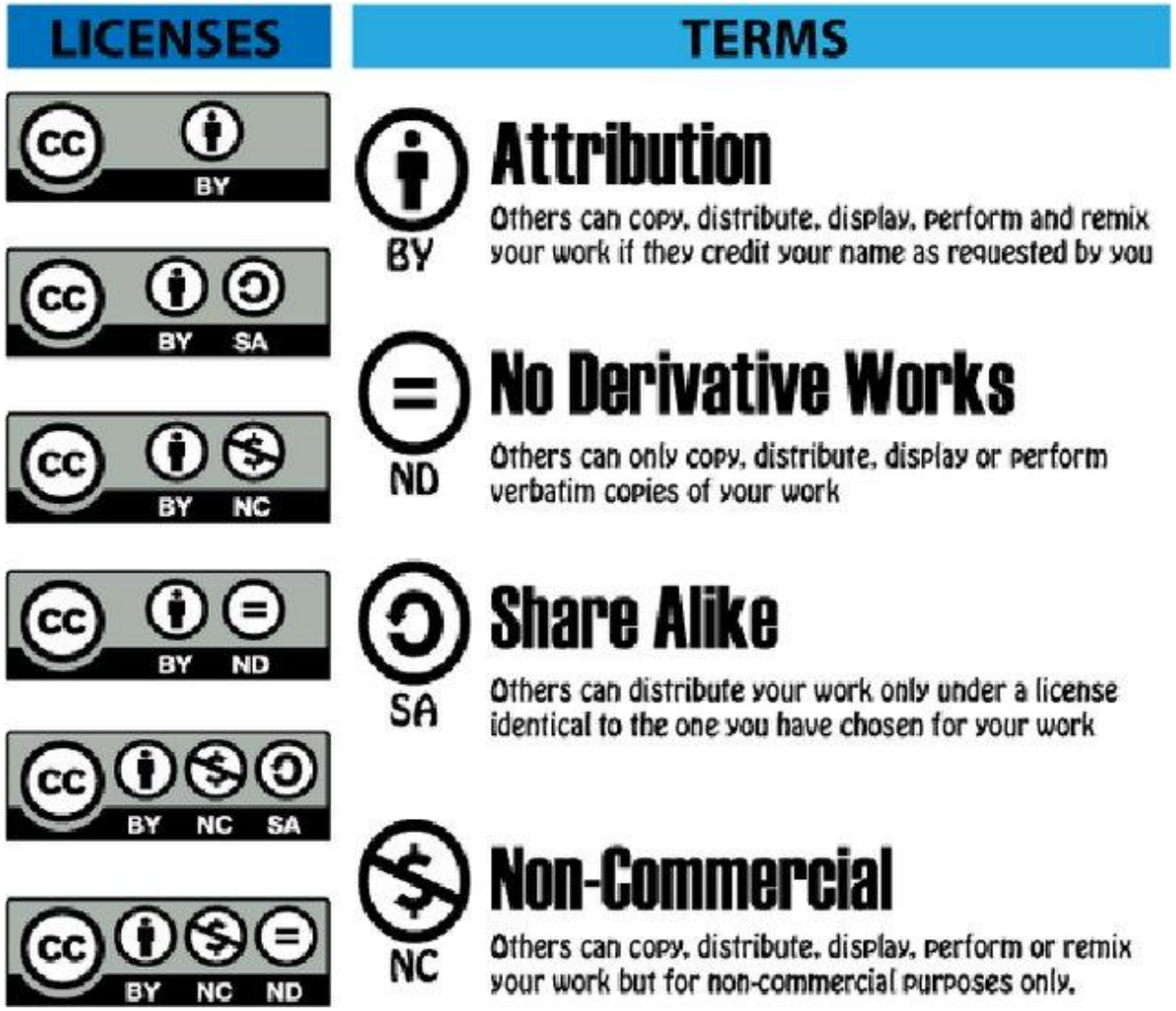
- Derechos morales: Su finalidad proteger a la persona del autor a través de su obra. Son irrenunciables e inalienables. (p.ej. Reconocimiento de autoría; respeto a la integridad de la obra)
- Derechos de carácter patrimonial: se pueden ceder o transferir: Explotación (reproducción, distribución, comunicación, acceso, transformación…) y compensatorios (p.ej. por ceder d. de reproducción)

- Si hay una licencia –aunque sea CC, o incluso CC0– tiene que poderse identificar quien la concede. Eso se hace bien con indicación de ©, bien indicando el autor
- Una base de datos (=juego de datos) puede tener un tipo de licencia y su contenido, o parte de su contenido, otro (p.ej., una BBDD con imágenes)

Licencias

creative commons

https://creativecommons.org



Francisco Pando
Unidad de Coordinación de GBIF.ES
CSIC
Joaquín Costa 22, 28002 Madrid

pando@gbif.es

https://creativecommons.org/licenses/by/4.0/deed.es

GBIF-ES es el Nodo Nacional de Información en Biodiversidad patrocinado por el Ministerio de Ciencia, Innovación y Universidades, gestionado por el Consejo Superior de Investigaciones Científicas (CSIC). https://www.gbif.es/ https://www.gbif.es/ , https://datos.gbif.es, https://elearning.gbif.es

MINISTERIO DE CIENCIA, INNOVACIÓN Y UNIVERSIDADES

CSIC
CONSEJO SUPERIOR DE INVESTIGACIONES CIENTÍFICAS

GBIF.ES

https://mutis.rjb.csic.es