GEOGRAMA

www.geograma.com

Los principios FAIR para mejorar los conjuntos de datos de alto valor

y su importancia para los usuarios de Información Geográfica

Diciembre 2023

Versión 1.0

GEOGRAMA

Agenda

- Por qué usamos datos
- HVD
- Cómo mejorar el uso de los datos
- Datos JUSTOS
- Herramientas
- Por qué usamos la IG

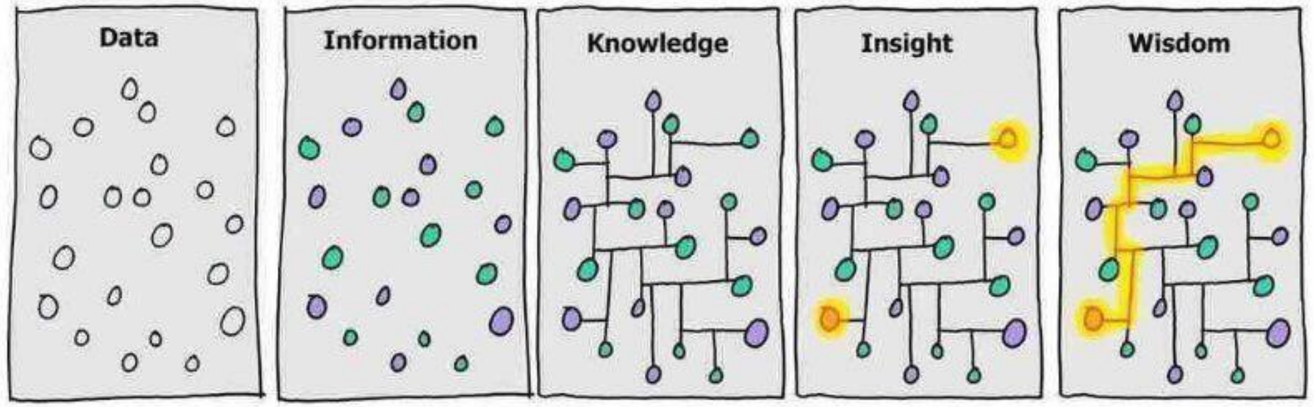
GEOGRAMA

Para qué usamos los datos

"¿Dónde está la vida que hemos perdido viviendo? ¿Dónde está la sabiduría que hemos perdido en el conocimiento? ¿Dónde está el conocimiento que hemos perdido en la información?" T. S. Eliot.

"Los datos no son información, la información no es conocimiento, el conocimiento no es comprensión, la comprensión no es sabiduría." Clifford Stoll.

"Lo que necesitamos es sabiduría, porque la información… no es igual al conocimiento, y el conocimiento no es igual al entendimiento, y el entendimiento no es igual a la sabiduría…" Koehn



GEOGRAMA

Conjuntos de datos de alto valor

Los conjuntos de datos de alto valor (HVD) corresponden a información del sector público que, **si estuviera disponible en formato abierto y reutilizable, generaría beneficios considerables para la economía, la sociedad y el medio ambiente**.

Actualmente:

- **Geoespaciales.** Parcelas agrícolas y datos básicos de referencia (direcciones, unidades administrativas, parcelas catastrales, nombres geográficos y edificios)
- **De observación de la Tierra y medioambiente.** Informes medioambientales, otros datos de valor (Hidrografía, Lugares protegidos, Elevaciones, Geología, Cubierta terrestre, Ortoimágenes, Zonas sujetas a ordenación, a restricciones o reglamentaciones y unidades de notificación, Regiones biogeográficas, Recursos energéticos, Instalaciones de observación del medio ambiente, Hábitats y biotopos, Uso del suelo, Recursos minerales, Zonas de riesgos naturales, Rasgos geográficos oceanográficos, Instalaciones de producción e industriales, Regiones marinas, Suelo y Distribución de las especies.)
- **Meteorológicos**
- Estadísticos
- De empresas
- **De movilidad.** Infraestructuras de transporte

Planificados:
Pérdida climática, gobierno y sector público, salud, justicia y derecho, lenguaje, energía y finanzas.

https://administracionelectronica.gob.es/
INSPIRE Geoportal (europa.eu)

GEOGRAMA

Cómo mejorar el uso de los datos

- A menudo no se conoce al cliente, o persona usuaria
- Aunque se conozca,
  - No se sabe para qué se usa
  - Ni con qué frecuencia
  - Ni qué transformaciones hace
- Esto ocurre con datos abiertos y con datos restringidos

FAIR

GEOGRAMA

F. Fáciles de encontrar

Los datos y metadatos pueden ser encontrados por la comunidad después de su publicación, mediante herramientas de búsqueda.

F1. Asignarles un identificador único y persistente a los datos y los metadatos

F2. Describir los datos con metadatos de manera prolija

F3. Registrar/Indexar los datos y los metadatos en un recurso de búsqueda

F4. En los metadatos se debe especificar el identificador de los datos que se describen.

GEOGRAMA

A. Accesibles

Los datos y metadatos están accesibles y por ello pueden ser descargados por otros investigadores utilizando sus identificadores.

A1 Los datos y los metadatos pueden ser recuperados por sus identificadores mediante protocolos estandarizados de comunicación

A1.1 Los protocolos tienen que ser abiertos, gratuitos e implementados universalmente

A1.2 El protocolo debe de permitir procedimientos para la autentificación y la autorización (por si fuera necesario).

A2 Los metadatos deben de estar accesibles, incluso cuando los datos ya no estuvieran disponibles.

GEOGRAMA

I. Interoperables

Tanto los datos como los metadatos deben de estar descritos siguiendo las reglas de la comunidad, utilizando estándares abiertos, para permitir su intercambio y su reutilización.

I1. Los datos y los metadatos deben de usar un lenguaje formal, accesible, compartible y ampliamente aplicable para representar el conocimiento

I2. Los datos y los metadatos usan vocabularios que sigan los principios FAIR

I3. Los datos y los metadatos incluyen referencias cualificadas a otros datos o metadatos

GEOGRAMA

R. Reutilizables

Los datos y los metadatos pueden ser reutilizados por otros investigadores, al quedar clara su procedencia y las condiciones de reutilización.

R1. Los datos y los metadatos contienen una multitud de atributos precisos y relevantes

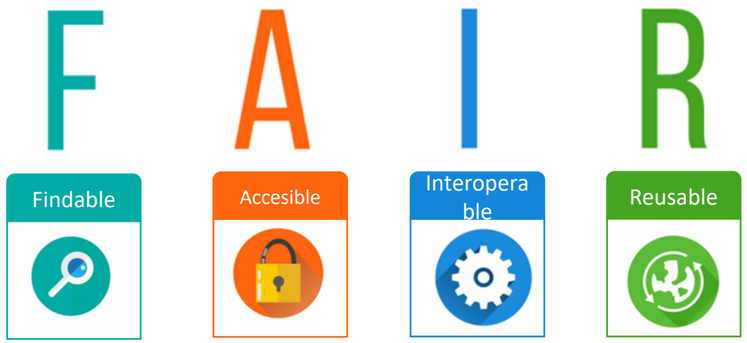
R1.1. Los datos y los metadatos se publican con una licencia clara y accesible sobre su uso y reutilización

R1.2. Los datos y los metadatos se asocian con información sobre su procedencia

R1.3. Los datos y los metadatos siguen los estándares relevantes que usa la comunidad del dominio concreto

https://datos.gob.es/es/noticia/principios-fair-buenas-practicas-para-la-gestion-y-administracion-de-datos-cientificos

FAIR



Harmonised data

- common data models
- harmonised vocabularies
- standard encodings

Harmonised services

conformant to well-known international standards

Harmonised metadata

for data and services documentation

Unique and persistent identifiers

for unambiguous reference to harmonised resources

GEOGRAMA
digital maps

FAIR

QUÉ QUEREMOS

QUÉ HEMOS HECHO

F

- **Catálogos de datos** (ayudan a buscar, descubrir y comprender los activos de datos)
- **Descripción de metadatos rica y estándar** (para determinar fácilmente si un activo es adecuado para el uso previsto).

- Crear **metadatos legibles** por máquina (descripciones detalladas)
- Seguir **estándares** (GeoDCAT-AP, ISO19139)
- Publicar a través de Catálogos Nacionales de Datos en el **Portal Europeo de Datos**

A

- **Condiciones de licencia claras y sencillas** (se reducen al mínimo los requisitos de registro y autenticación)
- **Normas y formatos** conocidos y **abiertos** para compartir datos

- Condiciones de **licencia estándar**
- **Servicios** de datos basados en **estándares** (OGC GEOAPI, WFS...)
- **Consultas** de datos **habilitadas**, para que los usuarios descarguen "justo lo que necesitan".

FAIR

QUÉ QUEREMOS

QUÉ HEMOS HECHO

I

- **PIDS para la identificación inequívoca** por humanos y máquinas
- **Especificaciones** de datos
- **Vocabularios controlados** (interoperabilidad semántica).

- **Modelos de datos bien documentados** (INSPIRE)
- Registro de **listas de códigos** para mejorar la interoperabilidad semántica
- **Identificadores únicos** para datos y metadatos

R

- **Metadatos enriquecidos:** Calidad de los datos (exactitud, integridad, conformidad)
- **Especificaciones de datos**
- Datos disponibles en **múltiples formatos abiertos** estándar
- **Servicios** basados en **normas**

- **Metadatos enriquecidos** para datos y servicios, liberados con una **licencia** transparente, adheridos a normas y publicados en **EDP**
- Datos disponibles en **múltiples formatos**
- **Modelos de datos bien documentados**: especificaciones de datos INSPIRE
- Uso de **codificaciones alternativas INSPIRE** (GeoPackage, GeoJSON).
- Los servicios **API del OGC,** facilitan a personas y máquinas el descubrimiento, la comprensión y la interacción con los datos.

GEOGRAMA
digital maps

FAIR -Herramientas para la evaluación automatizada

FAIRassist.org

FAIRassist.org

Help you discover resources to measure and improve FAIRness.

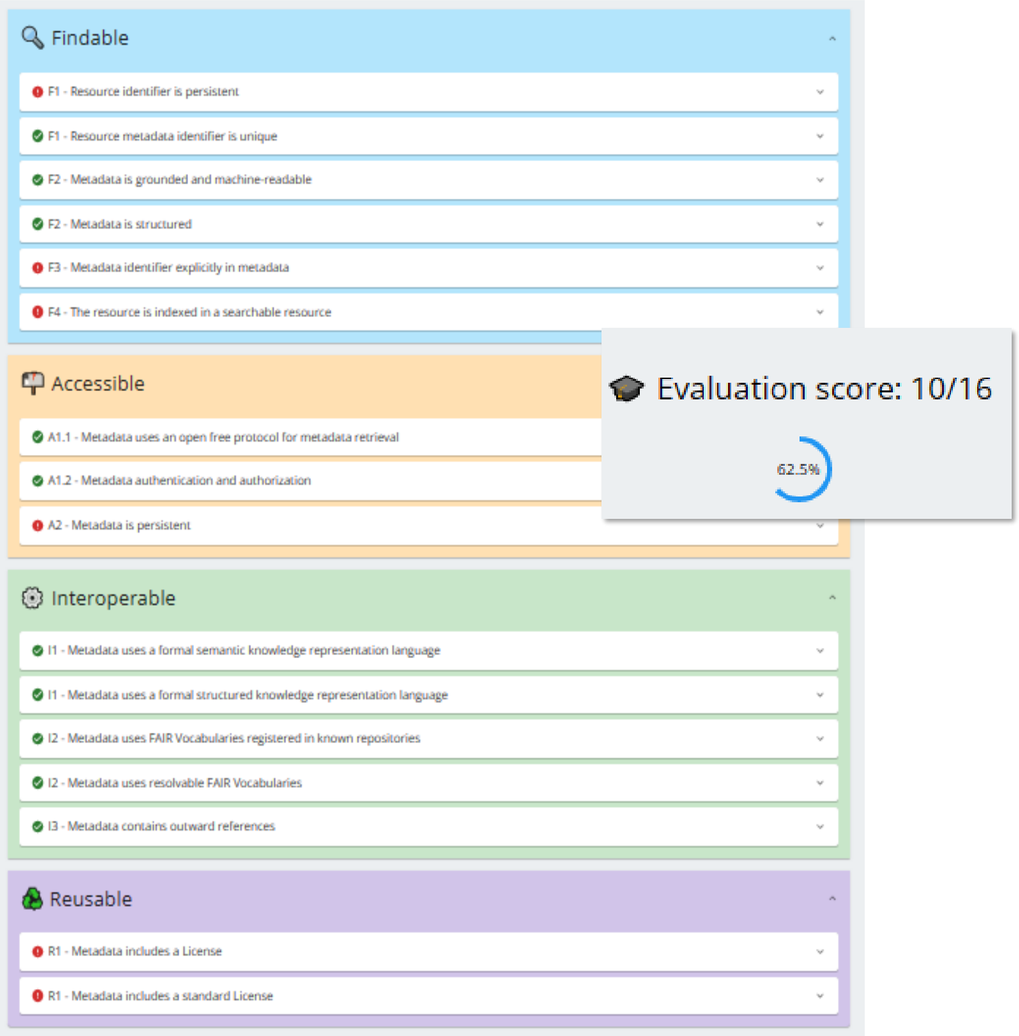
FAIRassist is the new, under development, educational component of the well established FAIRsharing resource.

FAIRassist is being designed to **offer personalised guidance to all stakeholders to discover standards and repositories in FAIRsharing**, which should be used to make data FAIR, as well as signpost other resources that enable FAIRness.

FAIRassist is under development, and it will implement a phased rollout of its content, working with and for the community. As an initial step, we are collecting and describing existing resources for the assessment and/or evaluation of digital objects against the FAIR principles, which are aspirational. The focus is on *manual questionnaires*, *checklists* and *automated tests* that help users understand how to achieve a state of "FAIRness", and how this can be measured and improved. This is not intended to be a comprehensive list of all groups, projects and organizations that tackle FAIRness or FAIRification. If your resource is missing, you can submit it via the "Register resource" button below.

Register a new resource

Resource	Execution Type	Key Features	Organisation	Target Objects	Reading Material
5 Star Data Rating Tool	Manual - questionnaire	Based on rating systems and maturity models	CSIRO OzNome	Datasets	
AutoFAIR	Semi-automated	A portal for automating FAIR assessments for bioinformatics resources	Department of Computer Information Systems, Faculty of ICT, University of Malta	Bioinformatics resources	Published Article
CLARIN Metadata Curation Dashboard	Automated	Checks validity of metadata descriptions, schemas, existence of semantic references (to a concept registry) and correctness of links to data and landing pages.	CLARIN ERIC	Datasets	Documentation
Data Stewardship Wizard	Predictive; based on a manually filled questionnaire	Helps researchers to design a data stewardship process for a project aiming for the highest reasonable FAIR data.	ELIXIR NL and ELIXIR CZ	All digital objects	Published Article
Do I-PASS for FAIR	Manual - checklist	A self assessment tool to measure the FAIR-ness of an organization. This self-assessment tool is a simple instrument, presented in an editable PDF form. By answering the questions and evaluating the level (beginner, intermediate, or advanced) at which you assess the performance of your organization, you will be able to define the actual FAIR-ness. In addition you can define a Road Map to become a FAIR Enabling Research Organization using the information in de more advances level(s).	LCRDM Task group "FAIR Enabling Research Organization".	Organizations	
F-UJI	Automated	The REST API support a programmatic assessment of the FAIRness of research data objects based on a set of core metrics developed by the FAIRsFAIR project. The metrics specification is available at https://doi.org/10.5281/zenodo.3775793	FAIRsFAIR (PANGAEA)	Datasets	Documentation
FAIR Data Self-					



FAIR -Herramientas para la evaluación automatizada

FAIR	Métricas test	Issues
Findable	Resource identifier is persistent	El identificador único del recurso de metadatos no es persistente y no está contenido en los propios metadatos.
	Metadata identifier explicitly in metadata	
	The resource is indexed in a searchable resource	El recurso no está indexado en los motores de búsqueda: Buscador DuckDuckGo, API de búsqueda personalizada de Google y buscador Bing.
Accesible	Metadata is persistent	N/A
Reusable	Metadata includes a License	Esta información se pierde en la transformación entre Metadatos INSPIRE a Geo-DCAT-AP, realizada por data.europa.eu.
	Metadata includes a standard License	

```
</gmd:resourceConstraints>
<gmd:resourceConstraints>
    <gmd:MD_LegalConstraints>
        <gmd:useConstraints>
            <gmd:MD_RestrictionCode codeList="http://standards.iso.org/iso/19139/resources/gmxCodelists.xml#MD_RestrictionCode" codeListValue="otherRestrictions
        </gmd:useConstraints>
        <gmd:otherConstraints>
            <gmx:Anchor xlink:href="https://creativecommons.org/licenses/by/4.0/">CC BY 4.0</gmx:Anchor>
        </gmd:otherConstraints>
    </gmd:MD_LegalConstraints>
</gmd:resourceConstraints>
```

GEOGRAMA

Datos JUSTOS

- Justificable (encontrables): Fácilmente localizables y su existencia y ubicación deben ser justificables.
- Universales (accesibles): Accesibles universalmente, es decir, disponibles para todos los que tengan derecho a acceder a ellos.
- Sincronizables (interoperables): Capaces de integrarse y trabajar de manera sincronizada con otros sistemas de datos. Para máquinas.
- Transferibles (reutilizables): Enfatizando la reusabilidad, los datos deben ser transferibles para su uso en diferentes contextos y aplicaciones.
- Organizados (encontrables y reutilizables): Los datos deben estar bien organizados para facilitar su búsqueda y acceso.
- Sostenibles (Reutilizables):Mantenibles y útiles a largo plazo, lo que también respalda su reusabilidad.

GEOGRAMA

Herramientas

- Encontrable(CKAN, GeoNetwork, Home Page (doi.org))
- Interoperabilidad (Validador INSPIRE, FAIRASSIST.ORG, R3gistry, DCAT-AP validator (europa.eu))
- Acccesibilidad – APIs (Geoserver, Home – pygeoapi)
- Reutilización (Asistente de licencia | data.europa.eu, Office, Open Office)

Dudas y preguntas

GEOGRAMA

www.geograma.com | geograma@geograma.com | +34 945 13 13 72

PRIVATE & CONFIDENTIAL