

Métricas FAIR. Principios FAIR y su aplicación en ASIO

Diego López-de-Ipiña

MORElab research group, Universidad de Deusto

dipina@deusto.es



HĒRCULES



Temas a tratar

☐ FAIR Data

- Principios
- FAIR vs Open Data
- Implementing FAIR & FAIRmetrics
- FAIRness de ASIO-HERCULES

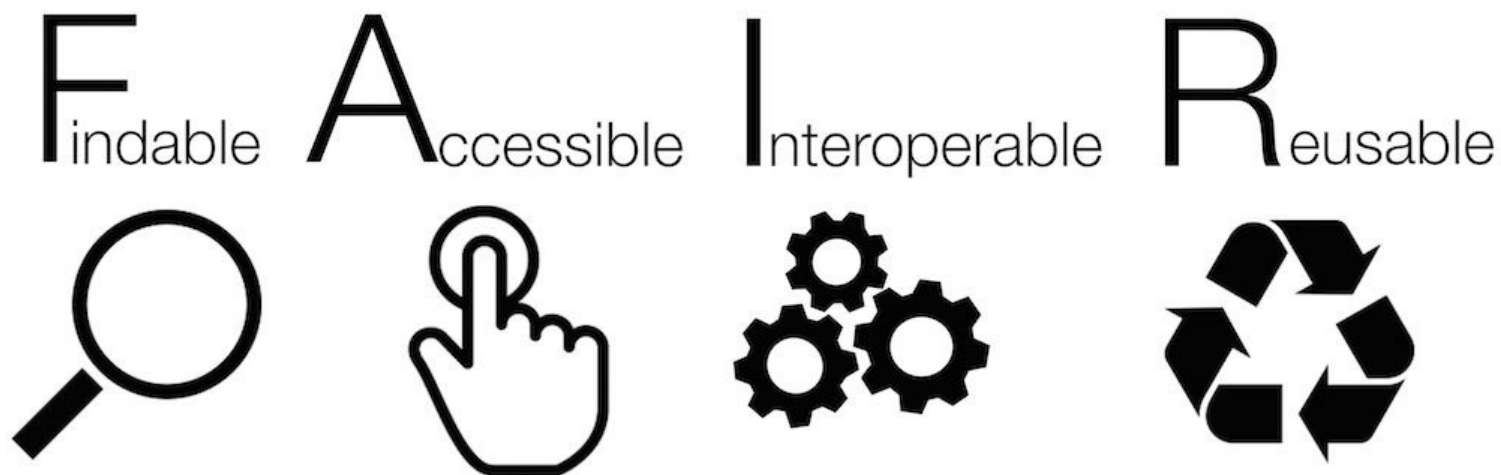
☐ Research Objects

- Definición
- Estándar RO-CRATE
- Ejemplos de uso



What is FAIR data?

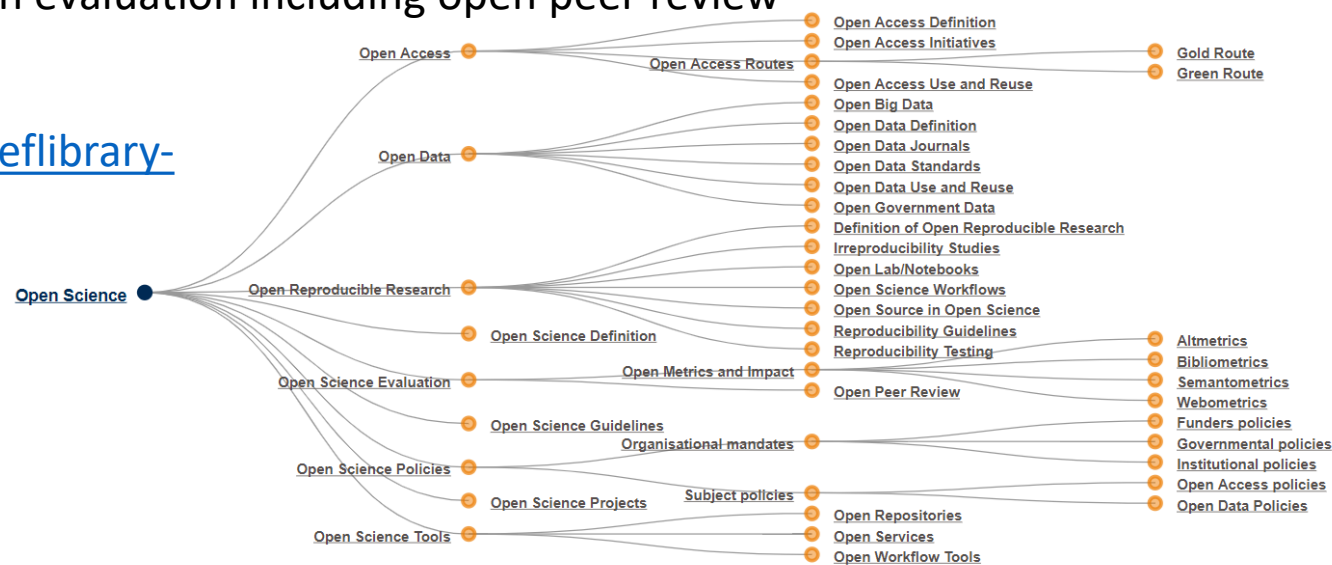
- ❑ A set of principles to ensure that data are shared in a way that enables & enhances reuse, by humans and machines



Origins of FAIR Data

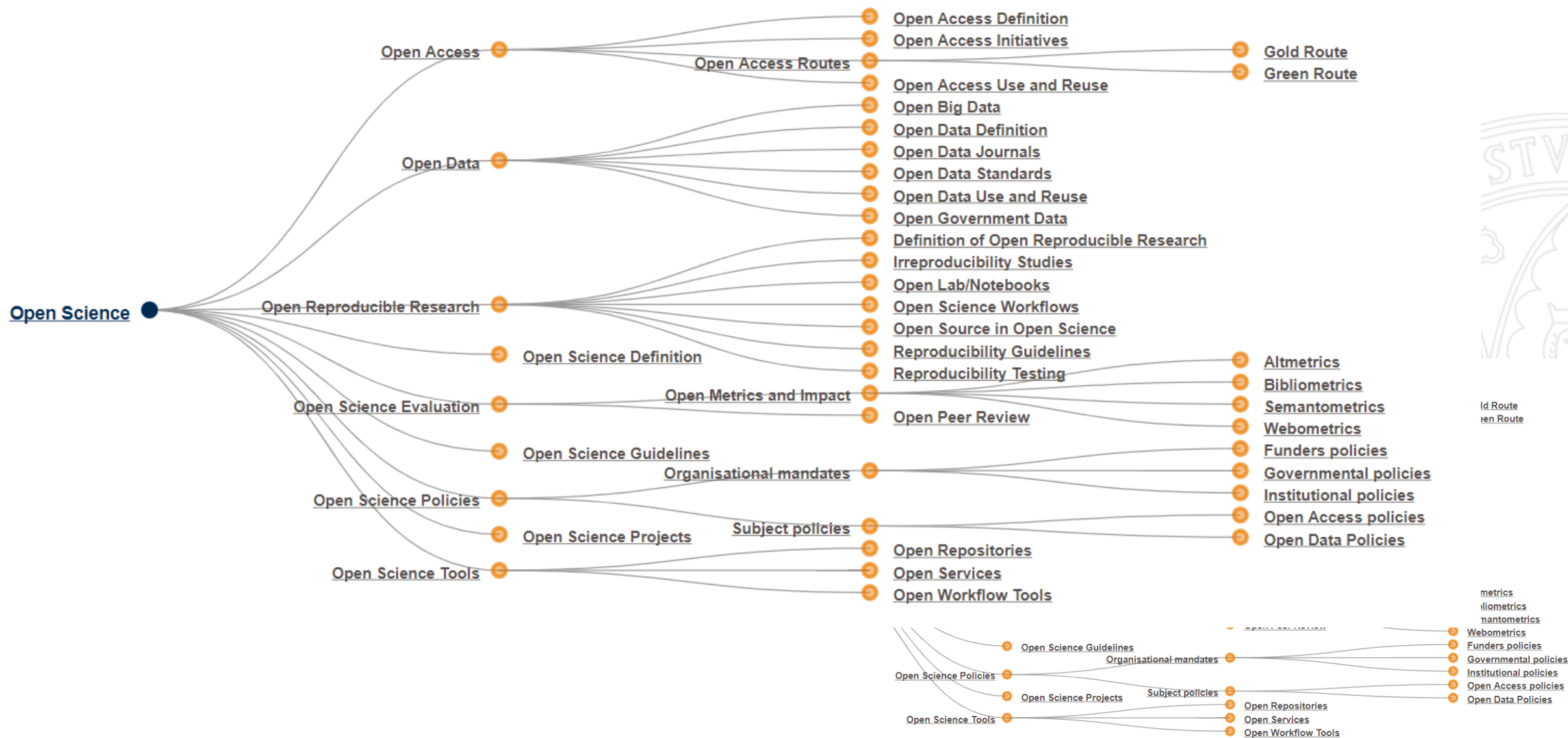
- ❑ Emerged from a workshop held in Leiden in 2014
- ❑ Come from life sciences but intended for all data
 - Issued by [FORCE11](#) community
- ❑ Echo previous principles on Open Research Data
 - Data underpinning scientific research results that has no restrictions on its access, enabling anyone to access it.
 - **Open Science** encompasses a variety of practices, usually including areas like open access to publications, open research data, open source software/tools, open workflows, citizen science, open educational resources, and alternative methods for research evaluation including open peer review

Beyond Open Access to publication: <https://blogs.uef.fi/ueflibrary-bors/different-ways-to-make-your-research-more-open/>



FONDO EUROPEO DE DESARROLLO REGIONAL (FEDER)

Una manera de hacer Europa



What FAIR means?

Findable

- F1. (meta)data are assigned a globally unique and eternally persistent identifier.
- F2. data are described with rich metadata.
- F3. (meta)data are registered or indexed in a searchable resource.
- F4. metadata specify the data identifier.

Accessible

- A1 (meta)data are retrievable by their identifier using a standardized communications protocol.
 - A1.1 the protocol is open, free, and universally implementable.
 - A1.2 the protocol allows for an authentication and authorization procedure, where necessary.
- A2 metadata are accessible, even when the data are no longer available.

Interoperable

- I1. (meta)data use a formal, accessible, shared, and broadly applicable language for knowledge representation.
- I2. (meta)data use vocabularies that follow FAIR principles.
- I3. (meta)data include qualified references to other (meta)data.

Reusable

- R1. meta(data) have a plurality of accurate and relevant attributes.
 - R1.1. (meta)data are released with a clear and accessible data usage license.
 - R1.2. (meta)data are associated with their provenance.
 - R1.3. (meta)data meet domain-relevant community standards.

FAIR Data Checklist

☐ Findable

- Persistent ID
- Metadata online

☐ Accessible

- Data online
- Restrictions where needed

☐ Interoperable

- Use standards, controlled vocabs
- Common (open) formats

☐ Reusable

- Rich documentation
- Clear usage licence

How FAIR are your data?

Findable

It should be possible for others to discover your data. Rich metadata should be available online in a searchable resource, and the data should be assigned a persistent identifier.

- ☐ A persistent identifier is assigned to your data
- ☐ There are rich metadata, describing your data
- ☐ The metadata are online in a searchable resource e.g. a catalogue or data repository
- ☐ The metadata record specifies the persistent identifier

Accessible

It should be possible for humans and machines to gain access to your data, under specific conditions or restrictions where appropriate. FAIR does not mean that data need to be open! There should be metadata, even if the data aren't accessible.

- ☐ Following the persistent ID will take you to the data or associated metadata
- ☐ The protocol by which data can be retrieved follows recognised standards e.g. http
- ☐ The access procedure includes authentication and authorisation steps, if necessary
- ☐ Metadata are accessible, wherever possible, even if the data aren't

Interoperable

Data and metadata should conform to recognised formats and standards to allow them to be combined and exchanged.

- ☐ Data is provided in commonly understood and preferably open formats
- ☐ The metadata provided follows relevant standards
- ☐ Controlled vocabularies, keywords, thesauri or ontologies are used where possible
- ☐ Qualified references and links are provided to other related data

Reusable

Lots of documentation is needed to support data interpretation and reuse. The data should conform to community norms and be clearly licensed so others know what kinds of reuse are permitted.

- ☐ The data are accurate and well described with many relevant attributes
- ☐ The data have a clear and accessible data usage license
- ☐ It is clear how, why and by whom the data have been created and processed
- ☐ The data and metadata meet relevant domain standards



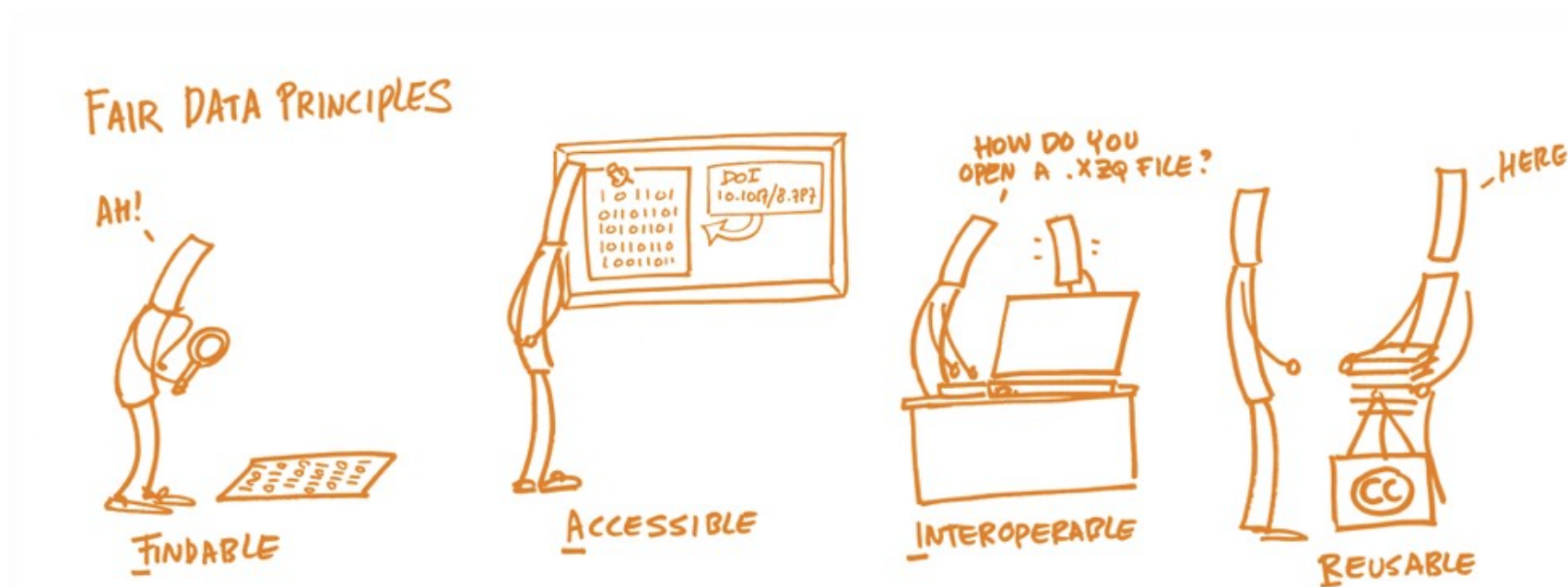
Una manera de hacer Europa

FAIR vs. Open Data

- ☐ FAIR data does not have to be open
- ☐ Data can be shared under restrictions & still be FAIR
- ☐ Making data FAIR ensures it can be found, understood and reused
- ☐ Open data is a subset of all the data shared

Implementing FAIR

- ❑ The principles do not specify particular technologies or implementations e.g. semantic web
- ❑ FAIR is not a standard to be followed or strict criteria – it's a spectrum / continuum
- ❑ Supporting FAIR will require investment in infrastructure, coordination across initiatives and engagement with research communities



Implementing FAIR: FAIRsharing

- ❑ **FAIRsharing** is a web-based, searchable portal of three interlinked registries, containing both in-house and crowdsourced manually curated descriptions of standards, databases and data policies, combined with an integrated view across all three types of resource.
 - By registering your resource on FAIRsharing, you not only gain credit for your work, but you increase its visibility outside of your direct domain, so reducing the potential for unnecessary reinvention and proliferation of standards and databases.

<https://fairsharing.org/educational/>

Researchers in academia, industry, government

Use FAIRsharing to identify and cite the standards, databases or repositories that exist for your discipline when creating a data management plan, releasing data or submitting a manuscript to a journal.

Journal publishers or organizations with data policy

Create in FAIRsharing your interrelated list of citable databases, repositories and relevant standards to recommend to your authors, users or their community; maintain and revise your recommendation over time.

Learned societies, unions and associations

Collaborate with FAIRsharing to raise awareness of standards, databases, repositories and data policies; mobilize your community to take action to promote registration, use and citation of key resources.



Developers and curators of resources

Make your standard, database or repository discoverable, by adding or claiming it in FAIRsharing; increase exposure and credit outside of your immediate community and promote adoption.

Research data facilitators, librarians, trainers

Use FAIRsharing to provide a foundation on which to create or enrich educational lectures, training and teaching material; and to plug into data management planning tools and other FAIR-supporting resources.

Funders and data policy makers

Recommend FAIRsharing to your awardees or community, to inform the development of their data management plan; and select the appropriate resources to recommend in your data policy.



Implementing FAIR: FAIR metrics

- ❑ Implement the FAIR principles into a **data assessment tool** so that every dataset which is deposited or reused from any digital repository can be assessed in terms of a score on the principles Findable, Accessible, Interoperable, and Reusable, using a **‘FAIRness’ scale from 1 to 5 stars**.
- ❑ Documentation: <https://www.slideshare.net/OSFair/osfair2017-workshop-outcomes-fair-metrics-starring-your-data-sets>
- ❑ “A design framework and exemplar metrics for FAIRness”, <https://dx.doi.org/10.1038%2Fdata.2018.118>

Comprobación del nivel FAIR en ASIO

- ☐ Sistema de comprobación automatizado del nivel FAIR cumplido por los recursos publicados durante el desarrollo del proyecto, que devuelve el nivel FAIR alcanzado, asociado a la versión del proyecto y la fecha en la que se ejecutó la comprobación.

- ☐ La red de ontologías ROH es FAIR significa que:
 1. Ofrece datos que sean Findable (Encontrables) a través de un identificador persistente e incluyendo metadatos
 2. Accessible (Accesible) a través del protocolo universal HTTP
 3. Interoperable usando vocabularios ampliamente adoptados y
 4. Reusable, se publican usando licencias de uso que promuevan la reusabilidad, como por ejemplo Creative Commons 4.0 BY-SA.

- ☐ La mejor manera de publicar datos FAIR es hacerlo mediante Linked Data, teniendo especial cuidado de generar datos y metadatos de alta calidad, mejorando así la reusabilidad de los datos para máquinas (y, como consecuencia y en última instancia, para humanos).

Sistema de Comprobación FAIRness Automatizado en HERCULES-ASIO

- ☐ La definición precisa de las métricas se ha realizado en base a las métricas definidas en el proyecto FAIR Metrics
- ☐ 14 métricas formuladas por FAIRmetrics
 - <https://github.com/FAIRMetrics>
- ☐ Front-end (interfaz de usuario) que facilite el lanzamiento del test suite y la generación de reportes que informen sobre el grado de cumplimiento de las métricas FAIR
 - <http://herc-as-front-desata.um.es/collections/15/evaluate/template>
- ☐ Tales reportes son archivados (históricos) para poder analizar el compromiso a lo largo del tiempo de ROH con los principios FAIR.

Estrategia de implementación de métricas FAIR

1. Evaluación del estado de cumplimiento actual de las propiedades FAIR del prototipo ROH:

- Servicio de ejecución de evaluación de servicios FAIR: <https://fairsharing.github.io/FAIR-Evaluator-FrontEnd/#!/collections/new/evaluate>
- Resultado: <https://fairsharing.github.io/FAIR-Evaluator-FrontEnd/#!/evaluations/4065>

2. Análisis de la implementación disponible de las métricas en Ruby

- Cambios introducidos en los últimos meses lo que implica que el proyecto sigue activo: <https://github.com/FAIRMetrics/Metrics/tree/master/MetricsEvaluatorCode/Ruby/metrictests>
- Despliegue del back-end de métricas FAIR en máquina local para uso dentro del proyecto:
 - GNOSS ha realizado despliegue en el [servidor de desarrollo](#)
 - Se ha contactado con autores, concretamente mark.wilkinson@upm.es
 - Su recomendación ha sido que hagamos uso de la API disponible en https://ejp-evaluator.appspot.com/FAIR_Evaluator//
 - Sin embargo, es necesario tener nuestro propio despliegue para evitar dependencias de terceros

Estrategia de implementación de métricas FAIR

3. Diálogo con autores de FAIRsharing (Mark Wilkinson):

- Es razonable contribuir al proyecto existente en vez de reinventar la rueda
- Se ha contrastado que las métricas FAIRsharing funcionan. Análisis realizado con:
 - Simple recurso apuntando a un DOI (10.1016/j.future.2019.09.033):
<https://fairsharing.github.io/FAIR-Evaluator-FrontEnd/#!/evaluations/3236>
 - Tal como se ha mencionado ejecutado también contra la definición de la red de ontologías ROH: <https://fairsharing.github.io/FAIR-Evaluator-FrontEnd/#!/evaluations/4065>
- Es posible crear nuevas métricas específicas al dominio de gestión de datos de la investigación a través de la compleción de FAIR Maturity Indicator [template](#). A partir de esta plantilla:
 - Equipo de FAIRsharing se ofrece a implementarlo
 - Se pueden realizar extensiones a través de la provisión de un endpoint para tales nuevas métricas usando SmartAPI (<https://smart-api.info/>), extensión de OpenAPI

Estrategia de implementación de métricas FAIR

4. Decisión de estrategia de implementación de métricas FAIR

- Creación de un proxy API usando [OpenAPI](#), que será utilizada desde back-end HERCULES-ASIO para realizar validación de recursos (resources) y de la red de ontologías como tal
 - Tal API se comunico en primera instancia con API pública ya desplegada por FAIRsharing: https://ejp-evaluator.appspot.com/FAIR_Evaluator// y ofrece interfaz web Swagger
 - Incluye interfaz gráfica para introducir, basada en <https://github.com/FAIRsharing/FAIR-Evaluator-FrontEnd>:
 - Recurso a validar
 - Conjunto de métricas a ejecutar
 - Visualización amigable del grado de cumplimiento de las políticas FAIR
- Publicación de [issues](#) y [pull requests](#) con modificaciones a las métricas ya implementadas en el caso de encontrarse bugs o realizarse extensiones a las existentes
- [Clonar API pública de FAIRsharing](#) para asegurar no dependencia con recurso externo
- Generación de nuevas métricas específicas al ámbito ASIO a través de smartAPI y/o entrega de plantillas para FAIRsharing

Estrategia de implementación de métricas FAIR

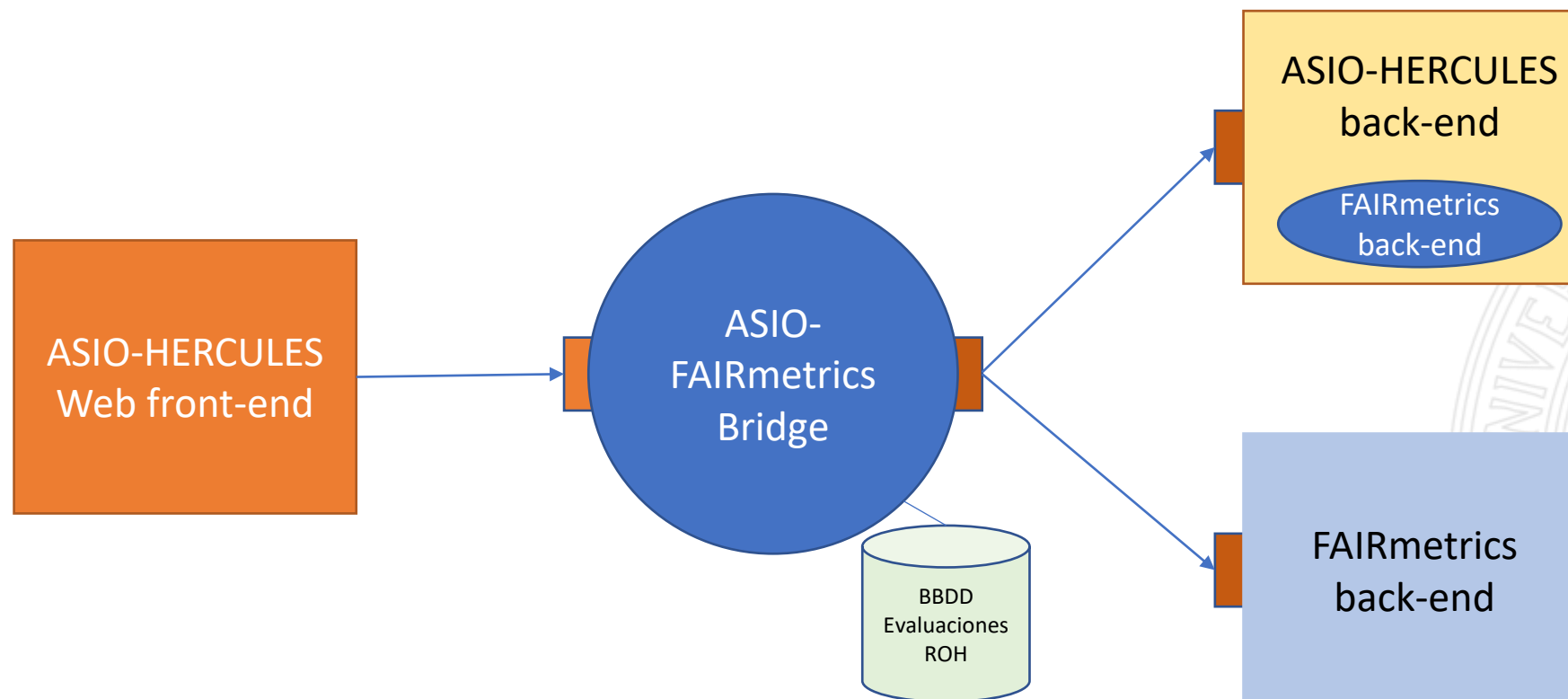
5. Solución desarrollada

- Se ha creado un proxy API (bridge) usando [OpenAPI](#), que será utilizada desde HERCULES-ASIO para realizar validación de recursos (resources) y de la red de ontologías como tal
 - [Incluye interfaz Swagger con 4 operaciones](#):
 - Recuperar colecciones de métricas (baterías de test disponibles)
 - Ejecución de colección de métricas seleccionada sobre recurso indicado
 - Recuperación de evaluaciones realizadas desde un ORCID
 - Recuperación de resultados de evaluación de ejecución de colección de métricas

6. Trabajo futuro

- Generación de recomendaciones para hacer ROH y recursos asociados (y por ende otras ontologías y recursos) compatible con políticas FAIR:
https://github.com/FAIRMetrics/Your_Path_To_FAIRness/tree/master/MI_Test_Tutorials
- Ejecución de políticas FAIR sobre test suite de recursos ROH y la red de ontologías
 - Adaptación de ROH para su cumplimiento integro de políticas FAIR

Arquitectura puente entre ASIO y FAIRmetrics



Bridge ASIO-FAIR Metrics código fuente

<https://github.com/HerculesCRUE/GnossDeustoBackend/tree/master/src/fair/bridge>

<http://herc-as-front-desata.atica.um.es/bridgeswagger/v1/ui>

GnossDeustoOnto/fair/bridge at x +

github.com/HerculesCRUE/GnossDeustoOnto/tree/master/fair/bri...

Search or jump to...

Pull requests Issues Marketplace Explore

Learn Git and GitHub without any code!

Using the Hello World guide, you'll start a branch, write comments, and open a pull request.

Read the guide

HerculesCRUE / GnossDeustoOnto Private

Unwatch 5 Star 0 Fork 0

Code Issues 0 Pull requests 0 Actions Projects 0 Wiki Security Insights

Branch: master GnossDeustoOnto / fair / bridge /

Create new file Upload files Find file History

dipina Updated README.md to include details on how to launch bridge includin... Latest commit ed22a5f 2 hours ago

..		
codegen_server	Publicada carpeta fair/ que contiene implementación de bridge/README...	2 hours ago
.gitignore	Publicada carpeta fair/ que contiene implementación de bridge/README...	2 hours ago
Pipfile	Publicada carpeta fair/ que contiene implementación de bridge/README...	2 hours ago
Pipfile.lock	Publicada carpeta fair/ que contiene implementación de bridge/README...	2 hours ago
README.md	Updated README.md to include details on how to launch bridge includin...	2 hours ago
api_spec.yml	Publicada carpeta fair/ que contiene implementación de bridge/README...	2 hours ago
app.py	Publicada carpeta fair/ que contiene implementación de bridge/README...	2 hours ago
config.py	Publicada carpeta fair/ que contiene implementación de bridge/README...	2 hours ago
openapi.yaml	Publicada carpeta fair/ que contiene implementación de bridge/README...	2 hours ago

README.md

bridge

Enlace entre la API del servidor público de FAIR Metrics y el futuro *backend* de ROH.

Pasos para la ejecución de FAIRness de recursos

1. Identificar y seleccionar batería de tests a realizar

- Se puede usar una de las baterías existentes: <https://fairsharing.github.io/FAIR-Evaluator-FrontEnd/#!/collections>
- O definirse una nueva a través de interfaz: <https://fairsharing.github.io/FAIR-Evaluator-FrontEnd/#!/collection/new>

2. Instalación y lanzamiento del puente ASIO-FAIRmetrics

- Configuración para selección endpoint FAIRmetrics con el que conectarse
- Acceso al front-end Swagger

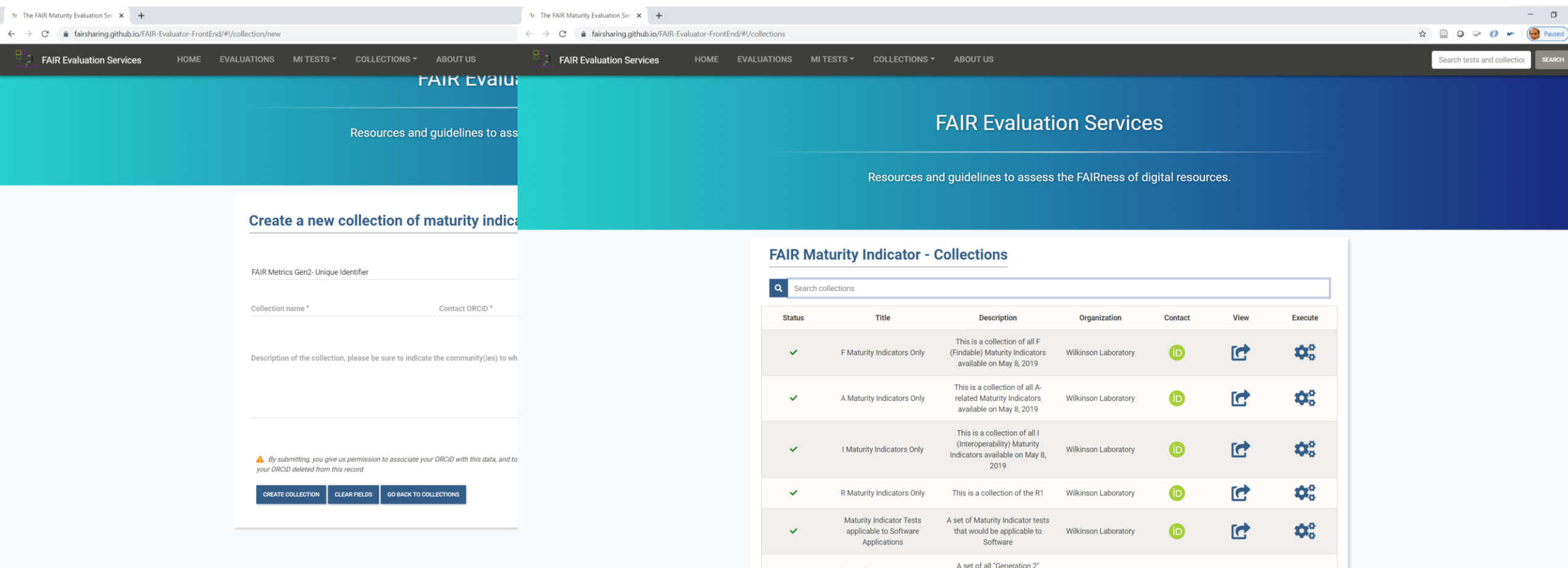
3. Ejecución de operaciones sobre API bridge ASIO-FAIRmetrics

- Ejecución de los métodos del API con recursos externos, ej. [10.1109/ACCESS.2019.2952321](https://doi.org/10.1109/ACCESS.2019.2952321)
- O con la propia red de ontologías de Hércules: <http://purl.org/roh>

FONDO EUROPEO DE DESARROLLO REGIONAL (FEDER)

Una manera de hacer Europa

Pasos para la ejecución de FAIRness de recursos (paso 1)

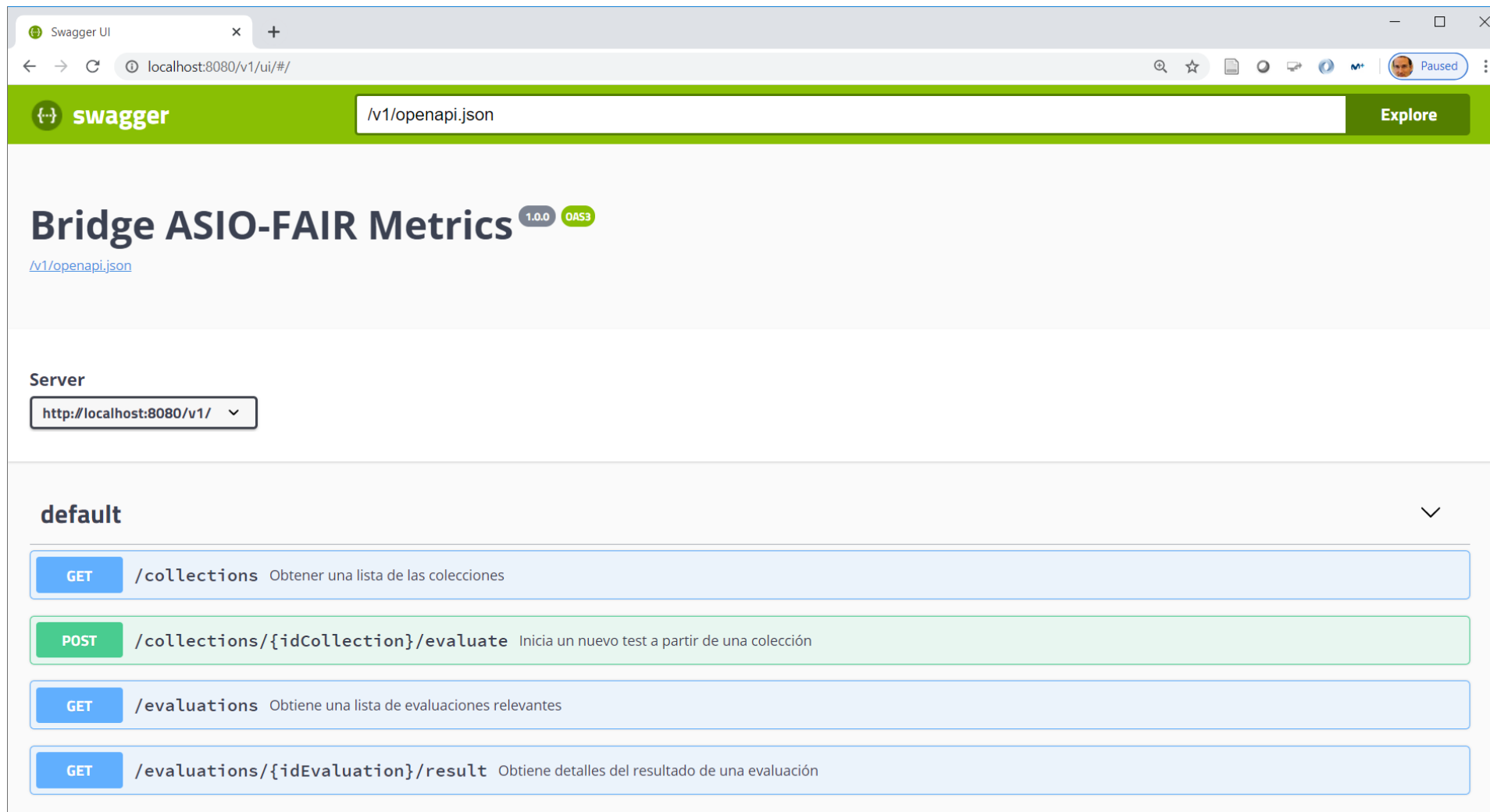


The screenshot displays the FAIR Evaluation Services web application. The top navigation bar includes links for HOME, EVALUATIONS, MI TESTS, COLLECTIONS, and ABOUT US. The main header area features the text "FAIR Evaluation Services" and "Resources and guidelines to assess the FAIRness of digital resources." Below this, the "Create a new collection of maturity indicators" form is visible, containing fields for "FAIR Metrics Gen2- Unique Identifier", "Collection name *", "Contact ORCID *", and a "Description of the collection, please be sure to indicate the community(ies) to which the indicators belong". A warning message states: "By submitting, you give us permission to associate your ORCID with this data, and to your ORCID deleted from this record". At the bottom of the form are buttons for "CREATE COLLECTION", "CLEAR FIELDS", and "GO BACK TO COLLECTIONS".

On the right side, the "FAIR Maturity Indicator - Collections" table is shown, listing various maturity indicator collections. The table has columns for Status, Title, Description, Organization, Contact, View, and Execute. The collections listed are:

Status	Title	Description	Organization	Contact	View	Execute
✓	F Maturity Indicators Only	This is a collection of all F (Findable) Maturity Indicators available on May 8, 2019	Wilkinson Laboratory	ID	View	Execute
✓	A Maturity Indicators Only	This is a collection of all A-related Maturity Indicators available on May 8, 2019	Wilkinson Laboratory	ID	View	Execute
✓	I Maturity Indicators Only	This is a collection of all I (Interoperability) Maturity Indicators available on May 8, 2019	Wilkinson Laboratory	ID	View	Execute
✓	R Maturity Indicators Only	This is a collection of the R1	Wilkinson Laboratory	ID	View	Execute
✓	Maturity Indicator Tests applicable to Software Applications	A set of Maturity Indicator tests that would be applicable to Software	Wilkinson Laboratory	ID	View	Execute
✓	All Maturity Indicators Tests	A set of all "Generation 2"	Wilkinson Laboratory	ID	View	Execute

Ejecución de puente Bridge ASIO-FAIR Metrics (paso 2)



The screenshot shows the Swagger UI for the Bridge ASIO-FAIR Metrics API. The browser address bar indicates the URL is `localhost:8080/v1/ui/#/`. The Swagger logo and the file `/v1/openapi.json` are visible in the top bar. The main title is **Bridge ASIO-FAIR Metrics** with version `1.0.0` and a `OAuth3` security scheme. The server URL is set to `http://localhost:8080/v1/`. Under the **default** tab, four API endpoints are listed:

- GET** `/collections`: Obtener una lista de las colecciones
- POST** `/collections/{idCollection}/evaluate`: Inicia un nuevo test a partir de una colección
- GET** `/evaluations`: Obtiene una lista de evaluaciones relevantes
- GET** `/evaluations/{idEvaluation}/result`: Obtiene detalles del resultado de una evaluación

Ejecución de colección de métricas en puente bridge ASIO-FAIR Metrics (paso 3)

http://herc-as-front-desata.atica.um.es/bridgeswagger/v1/ui/#/default/collections_id_evaluate_post

FON

Swagger UI

localhost:8080/v1/ui/#/default/collections_id_evaluate_post

POST /collections/{idCollection}/evaluate Inicia un nuevo test a partir de una colección

Parameters Cancel

Name	Description
idCollection * required integer (path)	ID de la colección sobre la que ejecutar el test
resource * required string (query)	URI/Identificador del recurso sobre el que ejecutar el test
orcid * required string (query)	ORCID ID de la persona que ejecuta el test
title * required string (query)	Título de la prueba

Execute Clear

Responses

Curl

```
curl -X POST "http://localhost:8080/v1/collections/5/evaluate?resource=http%3A%2F%2Fpurl.org%2Froh&orcid=0000-0001-8055-6823&title=Test%20to%20measure%20FAIRness%20in%20roh.owl%20with%20bridge%20ASIO-FAIRmetrics" -H "accept: application/json"
```

Request URL

```
http://localhost:8080/v1/collections/5/evaluate?resource=http%3A%2F%2Fpurl.org%2Froh&orcid=0000-0001-8055-6823&title=Test%20to%20measure%20FAIRness%20in%20roh.owl%20with%20bridge%20ASIO-FAIRmetrics
```

Server response

Code	Details
------	---------

FONDO EUROPEO DE DESARROLLO REGIONAL (FEDER)

Una manera de hacer Europa

Validación FAIRness publicación con puente ASIO-FAIRmetrics

Swagger UI

Not secure | herc-as-front-desa.atica.um.es/bridgeswagger/v1/ui/#/default/collections_id_evaluate_post

Apps | Bookmarks | Importados de la b... | XPS L501X no dupli... | Justificación de Hor... | Por dónde puedo c... | Por dónde puedo c... | ICT-based Solution... | Getting Ready for H... | Beginners Websites | CouchDB: The Defi... | Full Hierarchy - sch... | U-tad. Centro Unive... | Other bookmark

Servers

http://herc-as-front-desa.atica.um.es/fairmetrics_bridge/v1/

default

GET /collections Obtener una lista de las colecciones

POST /collections/{idCollection}/evaluate Inicia un nuevo test a partir de una colección

Parameters

Name

Description

idCollection * required

integer (path)

ID de la colección sobre la que ejecutar el test

5

resource * required

string (query)

URI/Identificador del recurso sobre el que ejecutar el test

http://purl.org/roh

orcid * required

string (query)

ORCID ID de la persona que ejecuta el test

0000-0001-8055-6823

title * required

string (query)

Título de la prueba

Latest test

Execute

Clear

Responses

Curl

```
curl -X POST "http://herc-as-front-desa.atica.um.es/fairmetrics_bridge/v1/collections/{idCollection}/evaluate?resource=http://purl.org/roh&orcid=0000-0001-8055-6823&title=LatestTest" -H "accept: application/json"
```

Request URL

http://herc-as-front-desa.atica.um.es/fairmetrics_bridge/v1/collections/{idCollection}/evaluate?resource=http://purl.org/roh&orcid=0000-0001-8055-6823&title=LatestTest

Server response

Code

Details

200

Response body

FONDO EUROPEO DE DESARROLLO REGIONAL (FEDER)

Una manera de hacer Europa

Validación FAIRness Red de Ontologías Hercules (ROH)

The FAIR Maturity Evaluation Service

fairsharing.github.io/FAIR-Evaluator-FrontEnd/#/evaluations/5095

FAIR Evaluation Services HOME EVALUATIONS MI TESTS COLLECTIONS ABOUT US

We are back online ! Thank you for your patience ! If you notice any unexpected failures in the tests, please report them to mark.wilkinson@upm.es

Latest test

Summary:

Description: FAIR Metrics Evaluation: Latest test; Tested identifier: <http://purl.org/roh>; generated by <https://orcid.org/0000-0001-8055-6823>


Resource: <http://purl.org/roh>

Collection: 5

Observations: Ran 20 tests (18 succeeded, 2 failed).

JSON response: https://w3id.org/FAIR_Evaluator/evaluations/5095.json

Tests passing and failing



FAIR METRICS GEN2- UNIQUE IDENTIFIER

Status: Success

Principle tested: F1

Description: Metric to test if the metadata resource has a unique identifier. This is done by comparing the GUID to the patterns (by regexp) of known GUID schemas such as URLs and DOIs. Known schema are registered in FAIRSharing (https://fairsharing.org/standards/?q=&selected_facets=type_exact:identifier%20schema)

Metric test created on: May 8, 2019 by Mark D Wilkinson (updated on Feb 10, 2020).

Test executed on: Feb 10, 2021

Test results

```
INFO: TEST VERSION 'Myst-1.3.3:Test-0.2.2'
SUCCESS: Found an identifier of type 'uri'
```

FAIR METRICS GEN2 - IDENTIFIER PERSISTENCE

This application is driven by the FAIRmetrics and the FAIRsharing groups. We recognize the support of the DBCLS BioHackathon series during which much of the back-end code was prototyped. Licensed under MIT.

FONDO EUROPEO DE DESARROLLO REGIONAL (FEDER)

Una manera de hacer Europa

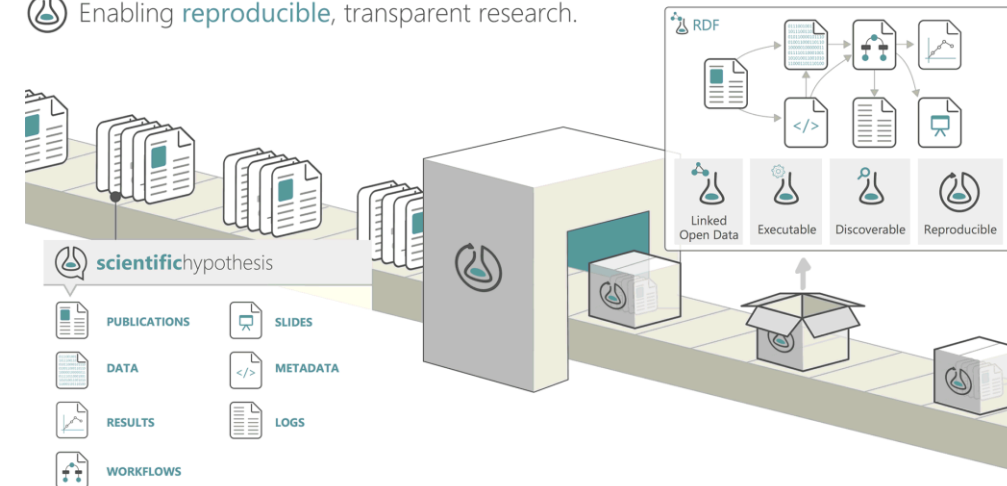
Science advances on a foundation of trusted discoveries. Reproducing an experiment is one important approach that scientists use to gain confidence in their conclusions.

Marcia McNutt, Editor-in-Chief of Science

Research Objects

- ❑ An **emerging approach to the publication, and exchange of scholarly information on the Web.**
- ❑ Research Objects aim to **improve reuse and reproducibility** by:
 - Supporting the publication of more than just PDFs, making data, code, and other resources first class citizens of scholarship
 - Recognizing that there is often a need to publish collections of these resources together as one shareable, cite-able resource.
 - Enriching these resources and collections with any and all additional information required to make research reusable, and reproducible!

 Enabling **reproducible**, transparent research.



Research Objects: going beyond PDFs

- ❑ The **goal** of research objects is to **improve the potential for understanding and reuse of research** by making sure that the information that is needed to make a published resource useful is associated with it and shared as a whole.
- ❑ As part of one research investigation, you might for example have:
 - Slides hosted on [slideshare](#),
 - Code in a [GitHub](#) repository,
 - Data in [figshare](#),
 - Data in [ArrayExpress](#).
 - **A research object not just a publication but everything else that goes into, and supports an investigation**
 - Initiatives and approaches to bring content together in a **machine-readable mechanism** so that it can be more easily shared and exchanged.

Goals

Extended Principles

Core Principles

Mechanisms

Findable	Accessible	Interoperable	Reusable
★ Execution	★ Dependencies	★ Graceful Degradation	
★ Checklists	★ Versioning	★ Provenance	
★ Identity	★ Aggregation	★ Annotation	
DOI	OAI-ORE	PROV	
ORCID	Git	ISA	
PURLs	ZIP	Puppet	MIM
		Open Annotation	

Research Objects: principles

□ Identity:

- Use globally unique identifiers as names for things. For example DOI's for publications or ORCID IDs for researchers. This is for two reasons
 1. So that we can talk unambiguously about things.
 2. So that people can find those things.

□ Agregation

- Use some mechanism of aggregation to associate things that are related or part of the broader investigation, study, etc.
 - Core to the value of Research Objects - providing the supporting artefacts that make the research potentially useful to someone else.

□ Annotation:

- Provide additional metadata about those things, how they relate to each other, where they came from, when etc.
 - This helps people discover data that are relevant and potentially useful.

Research Objects: benefits

- ☐ A way to refer to (and cite!) a bundle of research artefacts as a whole.
- ☐ A mechanism to associate human and machine-readable metadata with those research artefacts. Anything that can understand that metadata can therefore benefit.
- ☐ A graceful degradation of understanding.

Research Objects: modelling ROs with RO-CRATE

- ❑ RO-Crate is a community effort to establish a lightweight approach to packaging research data with their metadata.
 - It is based on schema.org annotations in JSON-LD.
- ❑ The RO-Crate effort brings together practitioners from very different backgrounds, and with different motivations and use-cases. Among our core target users are:
 1. **researchers** engaged with computation and data-intensive, workflow-driven analysis;
 2. **digital repository managers** and infrastructure providers;
 3. **individual researchers** looking for a straight-forward tool or how-to guide to “FAIRify” their data;
 4. **data stewards** supporting research projects in creating and curating datasets.
- ❑ RO-CRATE metadata specification: <https://www.researchobject.org/ro-crate/1.1/>
 - Reference: Eoghan Ó Carragáin; Carole Goble; Peter Sefton; Stian Soiland-Reyes (2019): **A lightweight approach to research object data packaging** *Bioinformatics Open Source Conference (BOSC2019)* <https://doi.org/10.5281/zenodo.3250687>

Research Objects: tools to ROs with RO-CRATE

- ❑ [Describe](#) interactive desktop application to create, update and export RO-Crates for different profile
- ❑ More tools at: <https://www.researchobject.org/ro-crate/implementations.html>

Research Objects: RO structure

- ❑ The structure an RO-Crate MUST follow is:

```
<RO-Crate root directory>/  
| ro-crate-metadata.json      # RO-Crate Metadata File MUST be present  
| ro-crate-preview.html      # RO-Crate Website homepage MAY be present  
| ro-crate-preview_files/    # MAY be present  
| | [other RO-Crate Website files]  
| [payload files and directories] # 0 or more
```

- ❑ Documented in <https://www.researchobject.org/ro-crate/1.1/metadata.html>

Research Objects: RO metadata

- ❑ Documented in <https://www.researchobject.org/ro-crate/1.1/metadata.html>
- ❑ The **Root Data Entity** is a Dataset that represent the RO-Crate as a whole; a Research Object that includes the Data Entities and the related Contextual Entities.
- ❑ The primary purpose for RO-Crate is to gather and describe a set of **Data entities** in the form of (linked with the root entity with hasPart property):
 - Files
 - Directories
 - Web resources
- ❑ The data entities can be further described by referencing **contextual entities** such as persons, organizations and publications.
 - Contextual entities however primarily exist outside the digital sphere (e.g. People, Places) or are conceptual descriptions that primarily exists as metadata, like GeoCoordinates and ContactPoint.

Research Objects: example 1

```
{ "@context": "https://w3id.org/ro/crate/1.1/context",
  "@graph": [
    {
      "@type": "CreativeWork",
      "@id": "ro-crate-metadata.json",
      "conformsTo": {"@id": "https://w3id.org/ro/crate/1.1"},
      "about": {"@id": "."}
    },
    {
      "@id": "./",
      "@type": [
        "Dataset"
      ],
      "hasPart": [
        {
          "@id": "cp7glop.ai"
        },
        {
          "@id": "lots_of_little_files/"
        }
      ]
    },
    {
      "@id": "cp7glop.ai",
      "@type": "File",
      "name": "Diagram showing trend to increase",
      "contentSize": "383766",
      "description": "Illustrator file for Glop Pot",
      "encodingFormat": "application/pdf"
    },
    {
      "@id": "lots_of_little_files/",
      "@type": "Dataset",
      "name": "Too many files",
      "description": "This directory contains many small files, that we're not going to describe in detail."
    }
  ]
}
```

```
<RO-Crate root>/
|   ro-crate-metadata.json
|   cp7glop.ai
|   lots_of_little_files/
|       | file1
|       | file2
|       | ...
|       | file54
```

Research Objects: example 2

```
{ "@context": "https://w3id.org/ro/crate/1.1/context",
  "@graph": [

    {
      "@type": "CreativeWork",
      "@id": "ro-crate-metadata.json",
      "conformsTo": {"@id": "https://w3id.org/ro/crate/1.1"},
      "about": {"@id": "./"}
    },
    {
      "@id": "./",
      "identifier": "https://doi.org/10.4225/59/59672c09f4a4b",
      "@type": "Dataset",
      "datePublished": "2017",
      "name": "Data files associated with the manuscript:Effects of facilitated family case conferencing for ...",
      "description": "Palliative care planning for nursing home residents with advanced dementia ...",
      "license": {"@id": "https://creativecommons.org/licenses/by-nc-sa/3.0/au/"}
    },
    {
      "@id": "https://creativecommons.org/licenses/by-nc-sa/3.0/au/",
      "@type": "CreativeWork",
      "description": "This work is licensed under the Creative Commons Attribution-NonCommercial-ShareAlike 3.0 Australia License. To view a copy of this license, visit http://creativecommons.org/licenses/by-nc-sa/3.0/au/ or send a letter to Creative Commons, PO Box 1866, Mountain View, CA 94042, USA.",
      "identifier": "https://creativecommons.org/licenses/by-nc-sa/3.0/au/",
      "name": "Attribution-NonCommercial-ShareAlike 3.0 Australia (CC BY-NC-SA 3.0 AU)"
    }
  ]
}
```

Documentación de contextual entities:

- <https://www.researchobject.org/ro-crate/1.1/contextual-entities.html>

FONDO EUROPEO DE DESARROLLO REGIONAL (FEDER)

Una manera de hacer Europa

Red Ontologías Hércules – ROH

Diego López-de-Ipiña

MORElab research group, Universidad de Deusto

dipina@deusto.es



HĒRCULES

MORE lab

ICT for good