FAIRness Evaluation, Validation and Advising with

FAIReva
evaluator, validator & advisor

Fernando Aguilar (IFCA-CSIC)

iBERGRiD

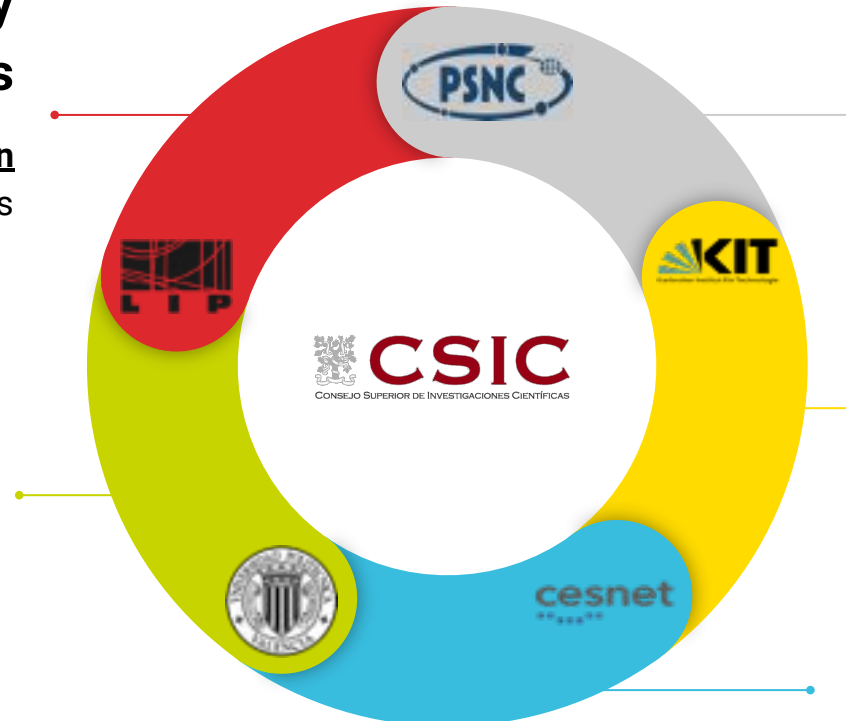# EOSC-SYNERGY in a nutshell

Promote EOSC High Quality Services

Software quality as a service, **FAIRness evaluation** and quality certification badges

Thematic Services Integration

10 thematic services addressing 4 scientific areas (Earth Observation, Environment, Biomedicine and Astrophysics)

22 partners in 10 countries
(ES, PT, FR, UK, DE, NL, CZ, SK, PL and BR)

Skills development

Environment for tutorials with a dedicated MOOC platform, courses methodology and a Hackaton as a service platform

Capacity Expansion at the Infrastructure level

Integration of services and resources from the RIs of the consortium partners

Alignment at the Policy Level

Collaboration with regional projects on landscaping activities, gap analysis and contribution to EOSC policies

Overview of the FAIR Principles

- **F**indable
- **A**ccessible
- **I**nteroperable
- **R**eusable

Refer to three types of entities: data (digital object), metadata (information about that digital object), and infrastructure.

FAIR Indicators

Findable

- ❏ F1:   (Meta) data are assigned globally unique and persistent identifiers
- ❏ F2:   Data are described with rich metadata
- ❏ F3:   Metadata clearly and explicitly include the identifier of the data they describe
- ❏ F4:   (Meta)data are registered or indexed in a searchable resource

Accessible

- ❏ A1:   (Meta)data are retrievable by their identifier using a standardised communication protocol
- ❏ A1.1:   The protocol is open, free and universally implementable
- ❏ A1.2:   The protocol allows for an authentication and authorisation where necessary
- ❏ A2:   Metadata should be accessible even when the data is 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 the FAIR principles
- ❏ I3:   (Meta)data include qualified references to other (meta)data

Reusable

- ❏ R1:   (Meta)data are richly described with 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 detailed provenance
- ❏ R1.3:   (Meta)data meet domain-relevant community standards

FAIR EVA - Evaluator, Validator & Advisor

1. FAIR indicators technical implementation.
   a. Starting from RDA
2. Modular, Scalable, Flexible
   a. Generic implementation OAI-PMH
   b. Plugins
3. Not only evaluate, but also validate and **advise**
4. Target: data producers, repository/data portal admins, funders.
5. Python. API + Web Interface
6. Stand-alone - Docker

DEMO

FAIR Assessment tools - FAIR EVA

- ## FAIR EVA functionality
- Comply with **FAIR Data principles**:
    1. **Data**: use a proper format
    2. **Metadata**: community standard. Machine-actionable (JSON, XML, RDF...)
    3. **PIDs**: Persistent Identifier (e.g. DOI). Provided by an accepted authority.
    4. **Repository/Data service**: indexed and machine-actionable.

- ## Integration: Different types of repositories/data portals

FAIR EVA
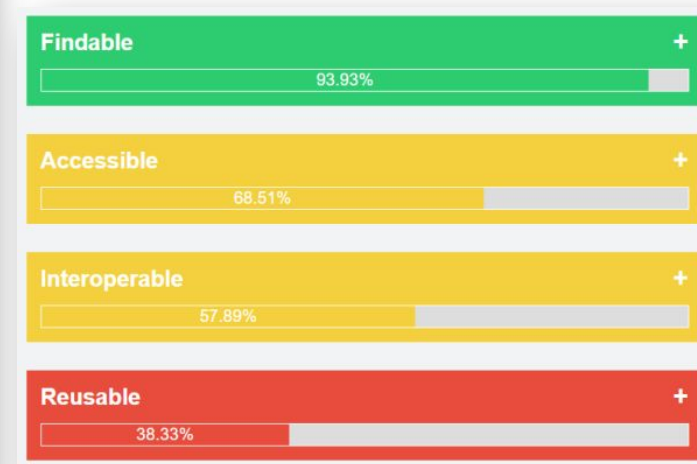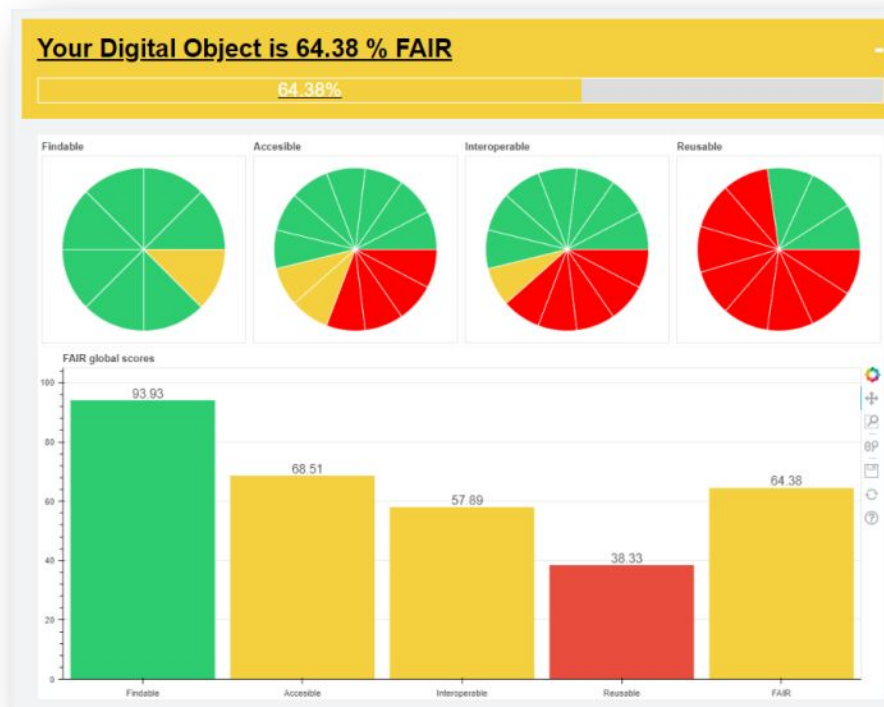
Technical Implementation examples

| Indicator Code	EN_title	EN_tech
RDA-F1-01M	Metadata is identified by a persistent identifier	Search within a predefined list of potential metadata terms to identify the metadata (dc.identifier.uri and dc.identifier.doi) if any information is available.
RDA-A1-04M	Metadata is accessed through standardised protocol	Checks the protocol to access metadata (OAI-PMH, API..)
RDA-I1-01M	Metadata uses knowledge representation expressed in standardised format	Checks metadata terms including controlled vocabulary information. So far, it checks GEONAMES, Library of Congress Subject Headings and ORCID, but it's being extended.
RDA-R1-01M	Plurality of accurate and relevant attributes are provided to allow reuse	Depending on the metadata schema used, checks that at least the mandatory terms are filled (75%) and the number of terms are high (25%)
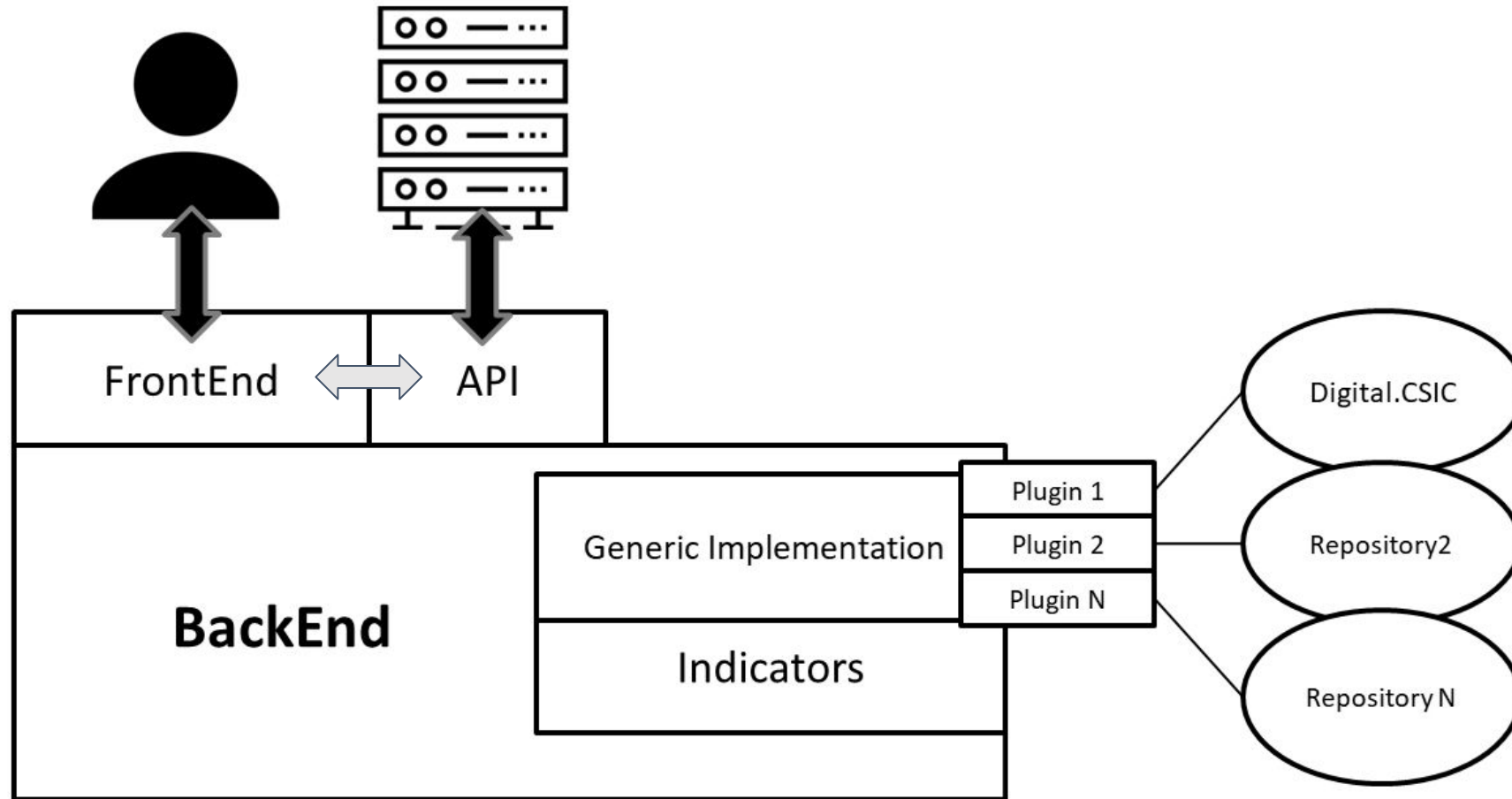
FAIR EVA - DIGITAL.CSIC plugin flow
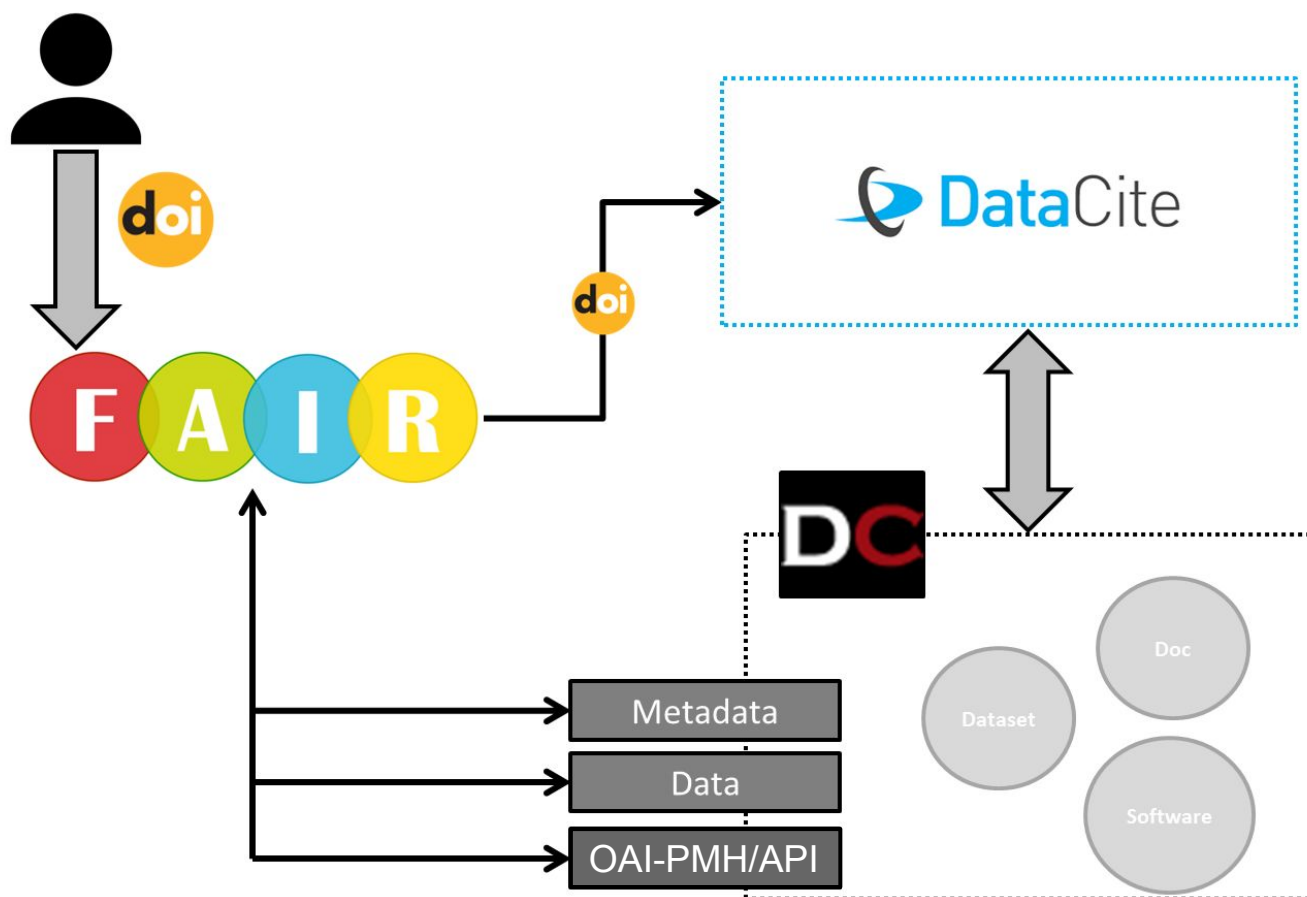
Semantics
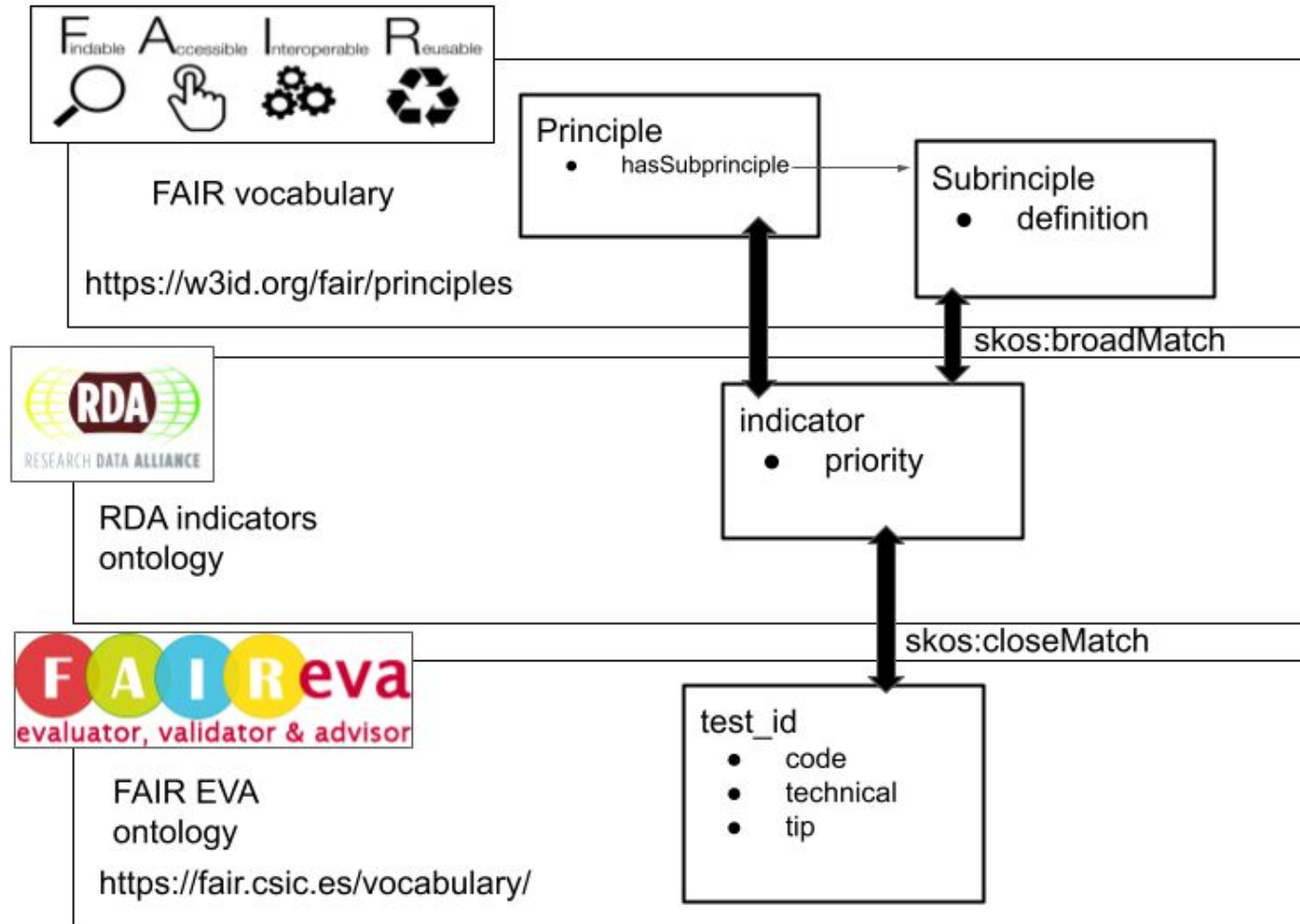
Thank you!

Fernando Aguilar (IFCA-CSIC)