



Setting the stage for FAIR Digital Objects (FDOs) in research

Tilo Mathes 
Research Space

Christophe Blanchi 
DONA Foundation

Tibor Kálmán 
GWDG

Maggie Hellström 
ICOS Carbon Portal

Sven Bingert 
GWDG

Felix Bach 
FIZ Karlsruhe



Co-funded by
the European Union



Agenda for today

1. Intro to the FDO concept & specifications (10 mins)
2. Different ways to implement FDOs (10 mins)
3. Tools & workflows supporting FDOs in research practice (10 mins)
4. Open discussion: What do you think about FDOs? (20 mins)
5. Wrap-up: ways to engage further with FDOs (5 mins)

Intro to the FDO concept & specifications

Sven Bingert, GWDG

The Best Recipe

For a FDO add and mix

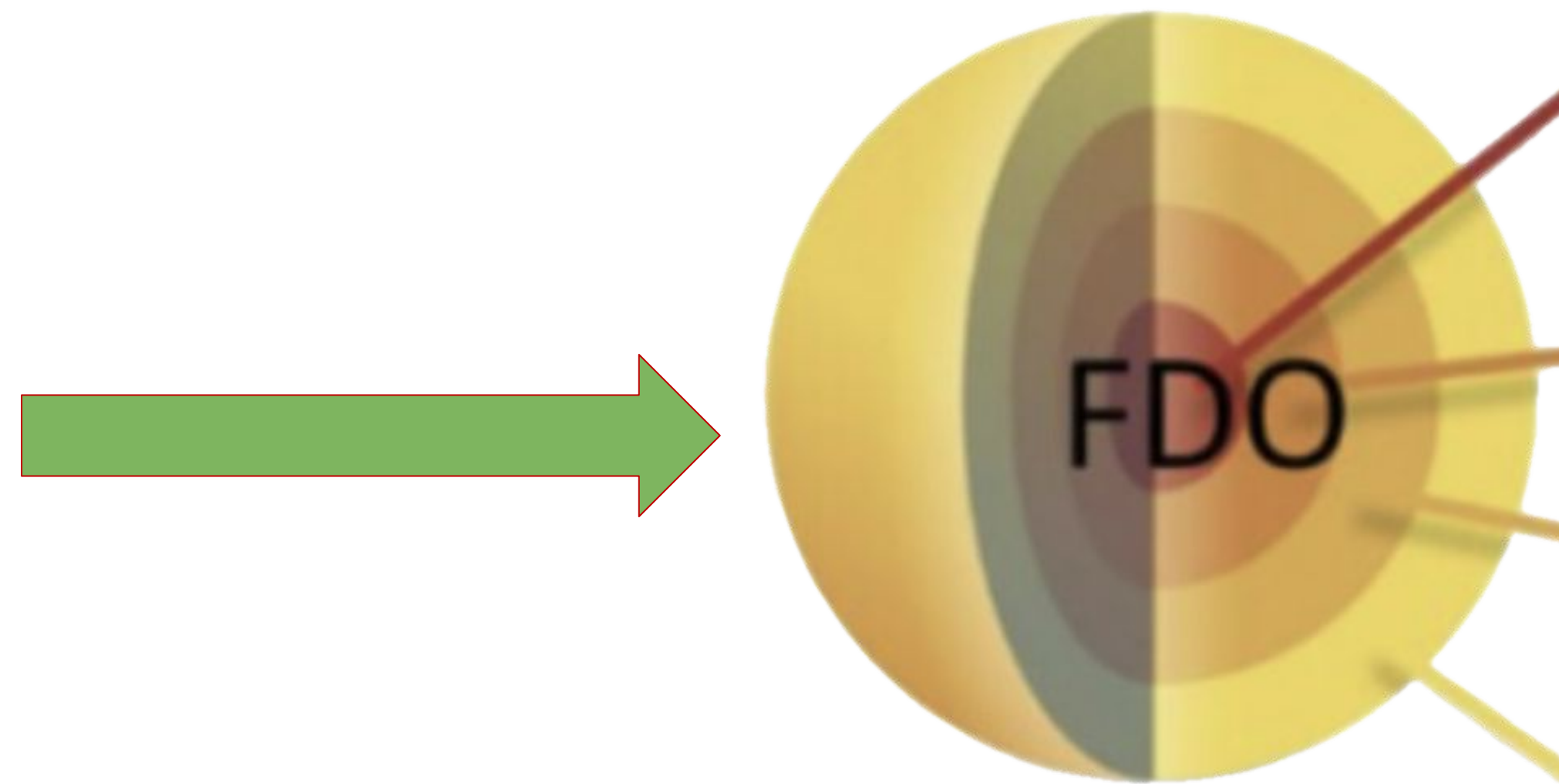
- *1 DATA*
- *some PIDs*
- *few Operations*
- *Infrastructure*
- *Lot of FAIRness*
- *some Tools*

Don't save on FAIRness !
Multiply the recipe if you have more data !

The Best Recipe

For a FDO add and mix

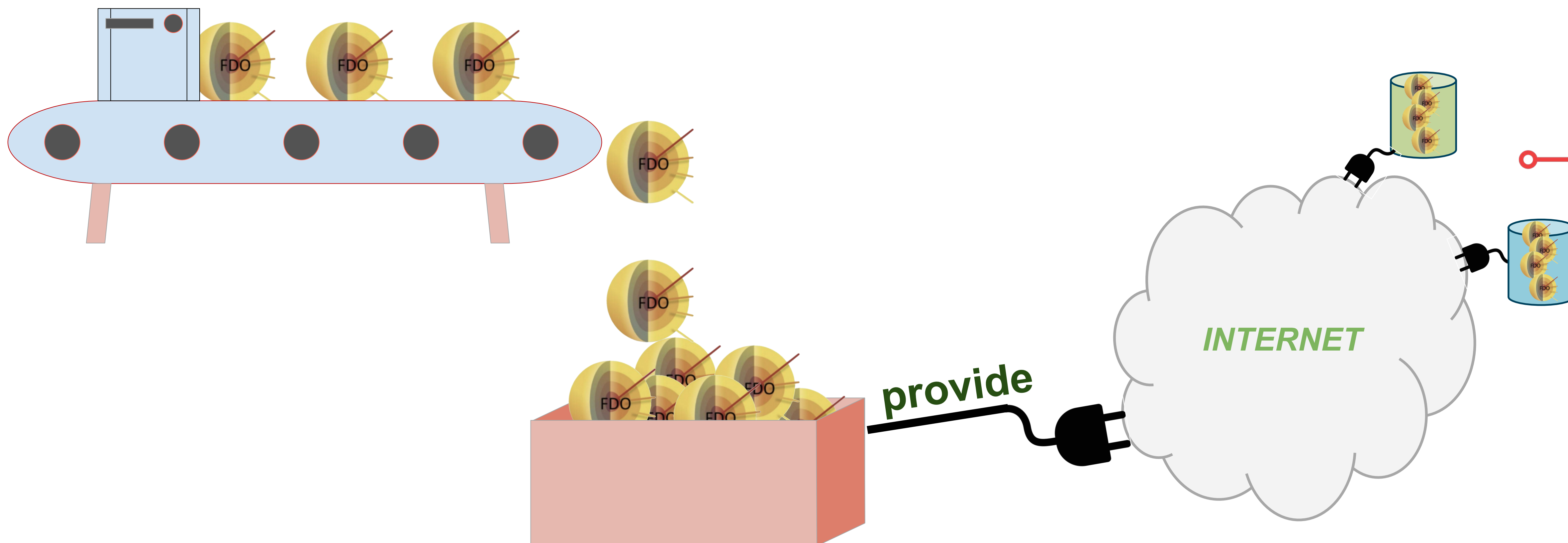
- *1 DATA*
- *some PIDs*
- *few Operations*
- *Infrastructure*
- *Lot of FAIRness*
- *some Tools*



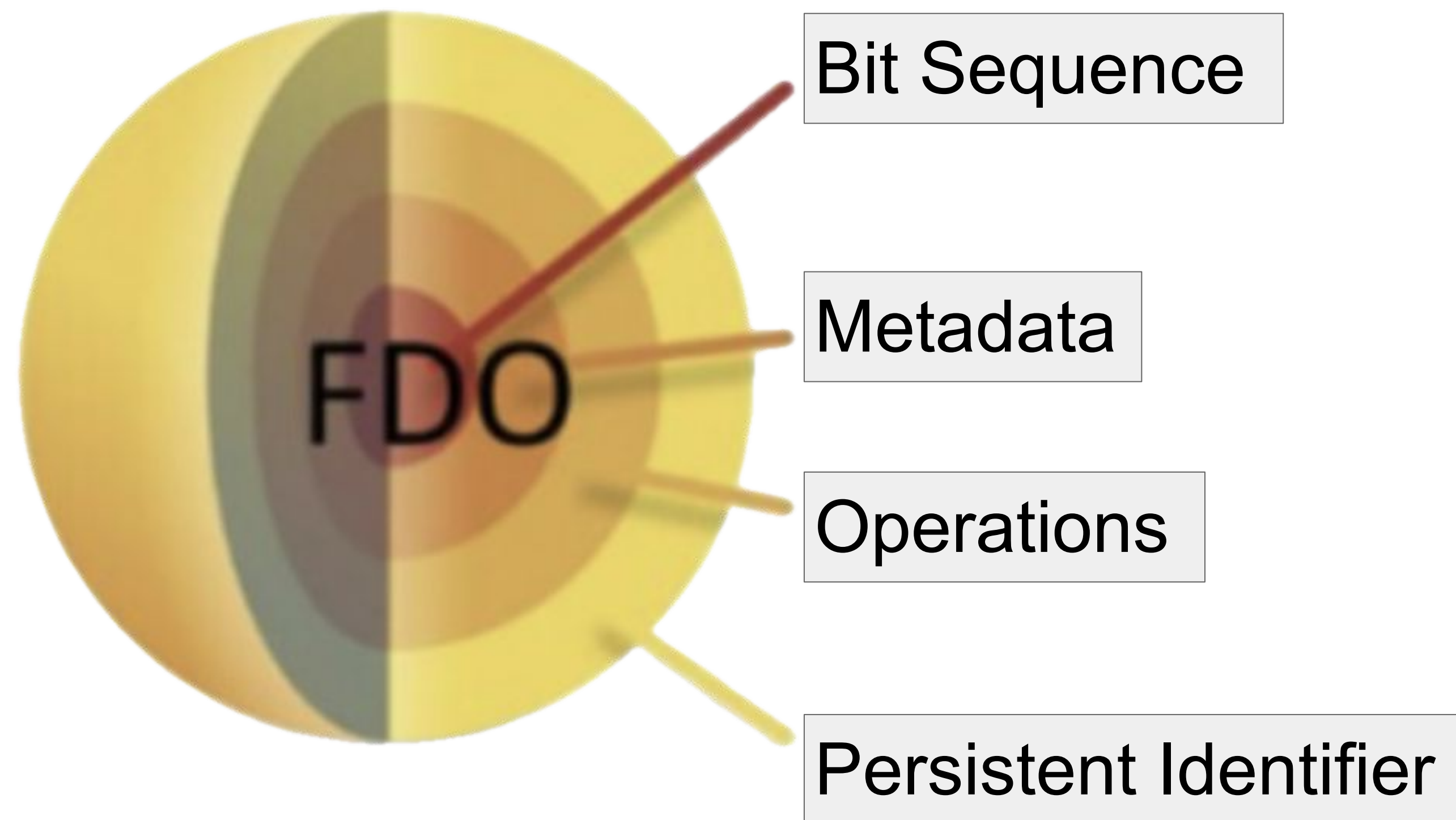
Don't save on FAIRness !
Multiply the recipe if you have more data !

FDO to be reused

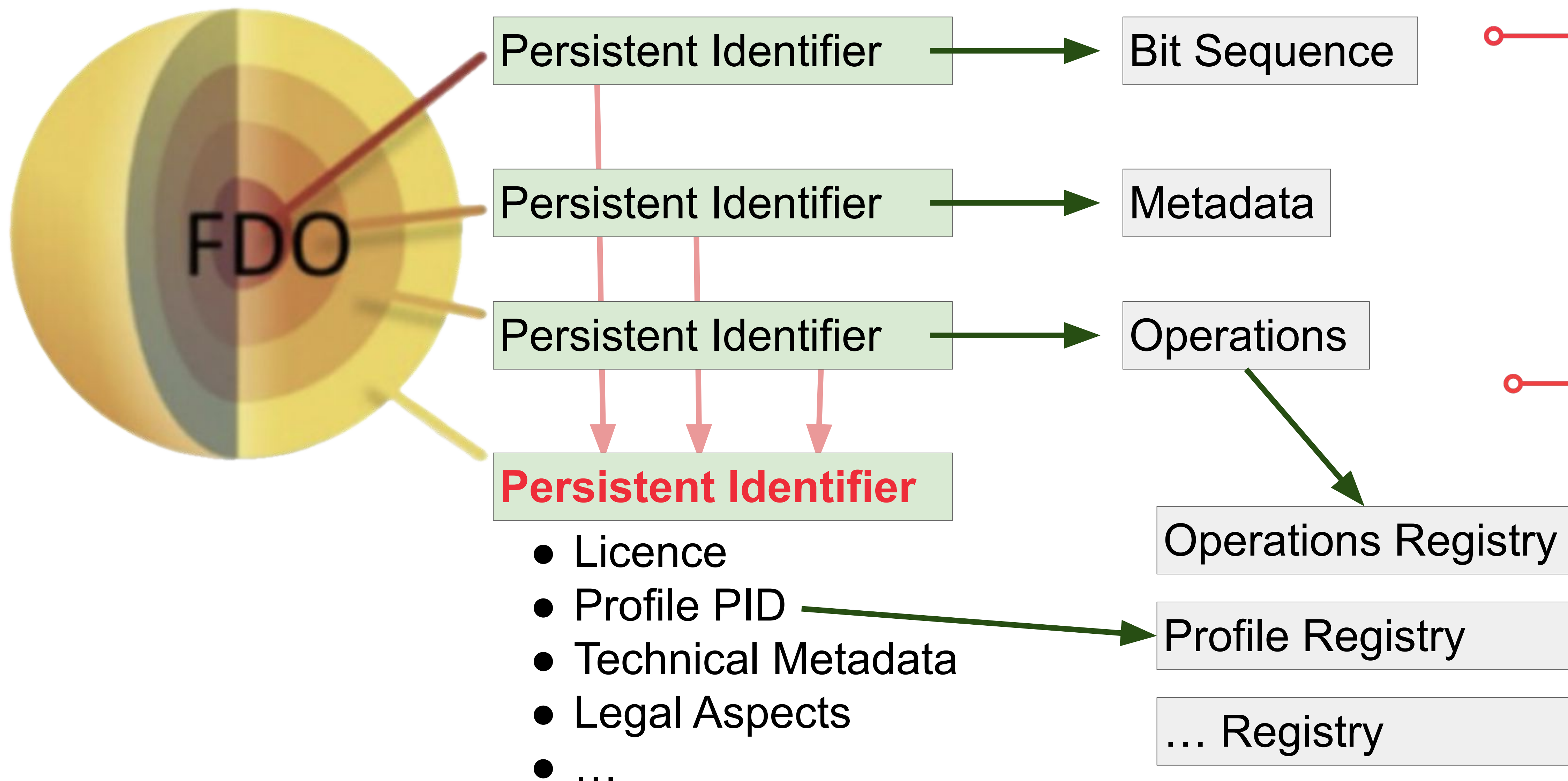
- FDO as output of workflows / processes
- FDOs are provided / accessible from the internet
- FDOs are ready for reuse



FDO basic ingredients

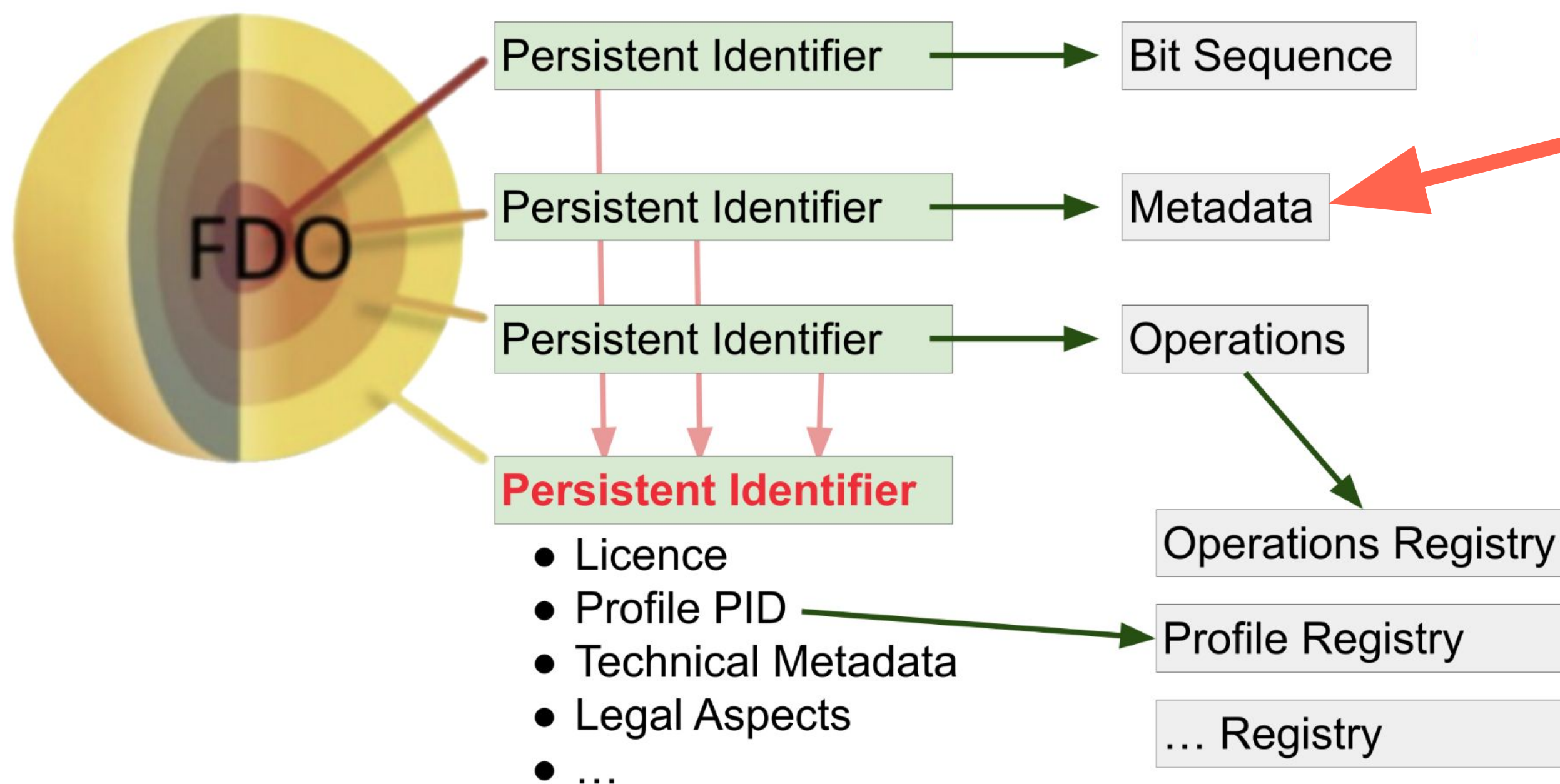


Machine Actionable FDO



Machine Actionable FDO

- Metadata is described by typed attributes
- Types are referenced by **Persistent Identifier**



Persistent Identifiers form the basis for the creation of FDOs !

CONFERENCE
ON PERSISTENT
IDENTIFIERS
2024

- **FDO Overview**

- **FDO Requirement Specification**

- **Implementation of Attributes, Types, Profiles and Registries**

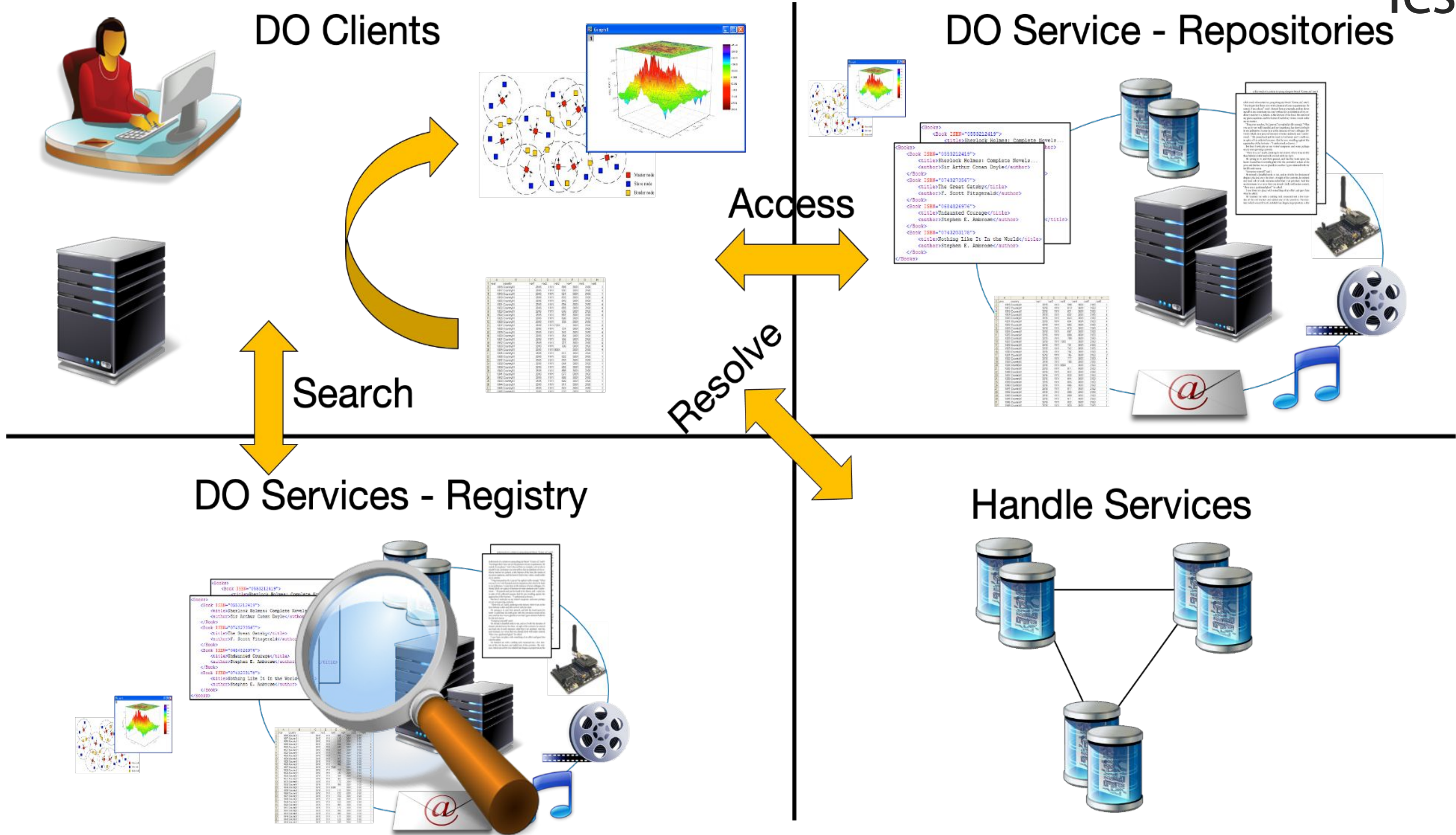
- Any more helpful documentation!



Different ways to implement FDOs

Christophe Blanchi, DONA Foundation & Felix Bach, FIZ Karlsruhe

Digital Object Architecture Handle/DOIP

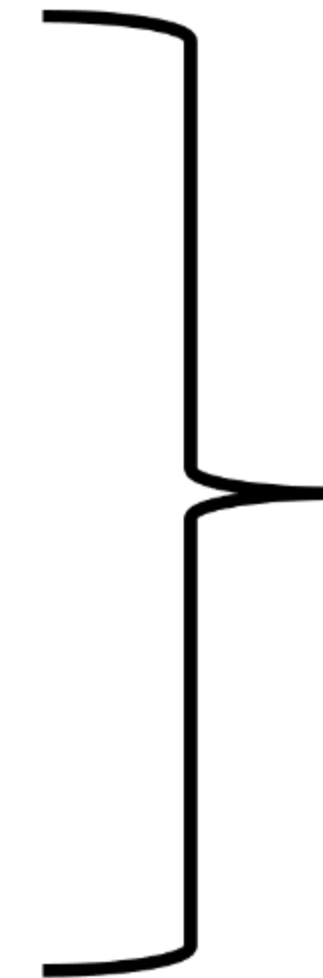


Using the Handle System and DOIP

- **Handle protocol (DO-IRP)** – Leverages the Global Handle Registry® to resolve any handle globally.
 - Create and administer globally unique identifiers known as handles.
 - Provides rapid secure resolution of a handle into a digital object structured as a handle record.
 - Provides distributed administration of handles and their records.
 - Has built in PKI and security.
- **DOIP** – Interoperable access to distributed, independently managed services and data sources.
 - DOIP is a protocol to interact with digital objects.
 - Leverages existing transport encryption standards such as TLS.
 - Provides built-in security leveraging the PKI of the Handle System™.
 - Leverages the Handle for typing of data and extensible operations, and for client authentication.

Digital Object Data Model

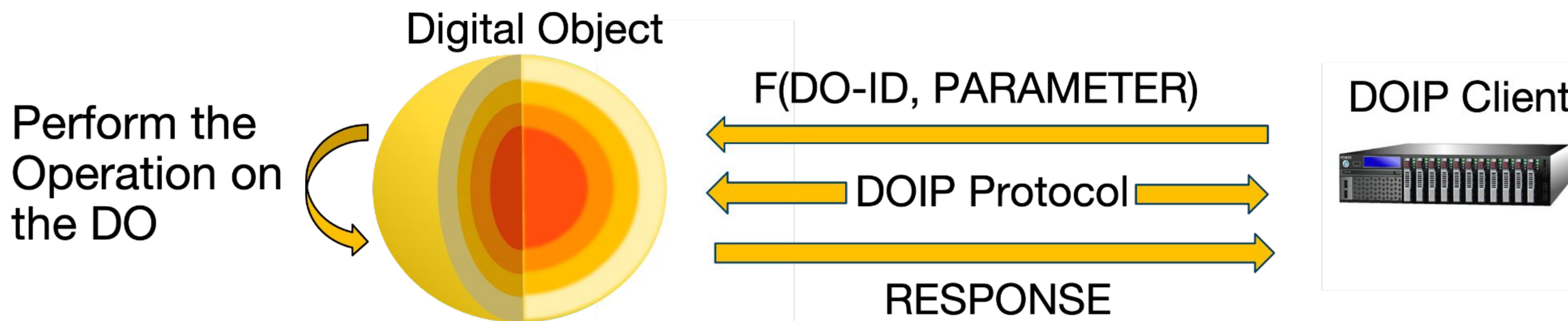
- **id:** the DO's identifier: A handle
- **type:** The type of the DO. (typically a handle)
- **attributes:** 0 or more key-value pairs.
 - **key:** ID (a Handle) for the type of the value
 - **value:** the an instance of the associated type.
- **elements:** 0 or more elements each consisting of:
 - **id:** ID of the element. Must be unique in the DO.
 - **type:** the type of the element.
 - **length(optional):** the length of the data element.
 - **attributes:** 0 or more key-value pairs.
 - **data:** the payload of the data element.
- **signature** Optional



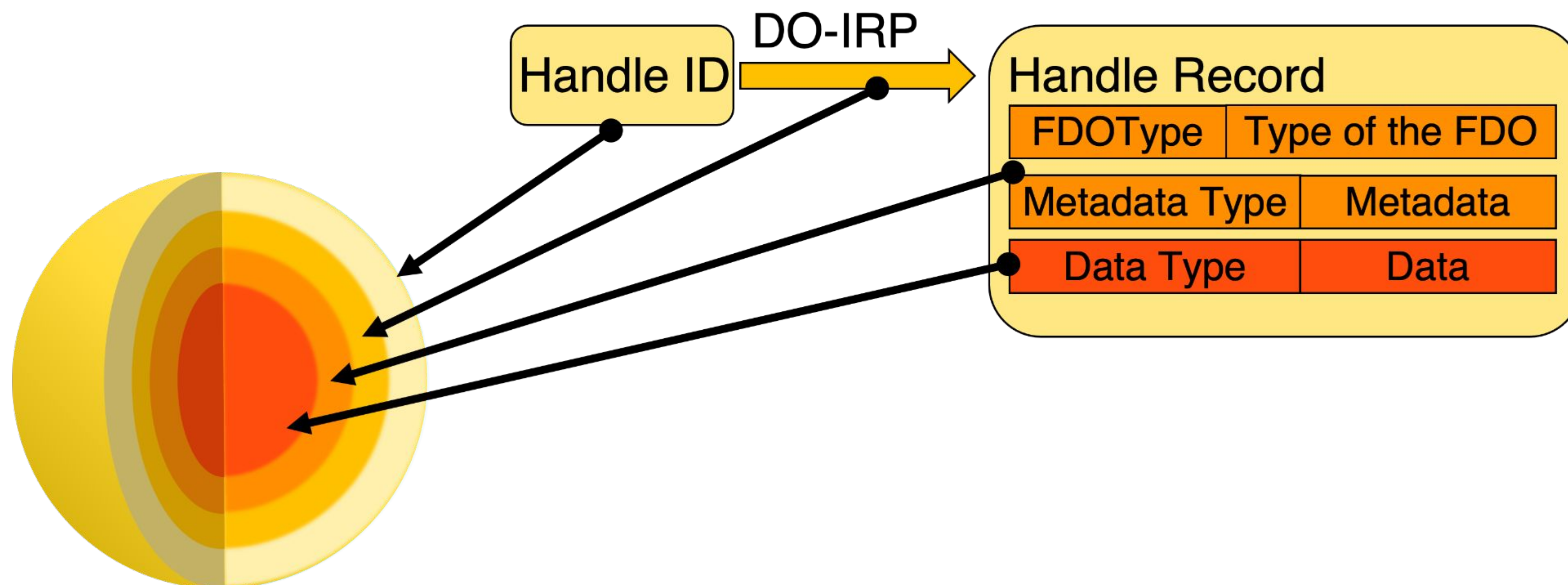
DOIP Protocol - Operations

- DO Operations are invoked on a DO by issuing a request/response over DOIP.
- Standard DO Operations include ListOperations, Create, Update, Delete, Search.
- New Operations can be created as a DO by anyone.
- An FDO Operation can be represented as a function F that expects the ID of the targeted FDO as well as any parameters required by the operation

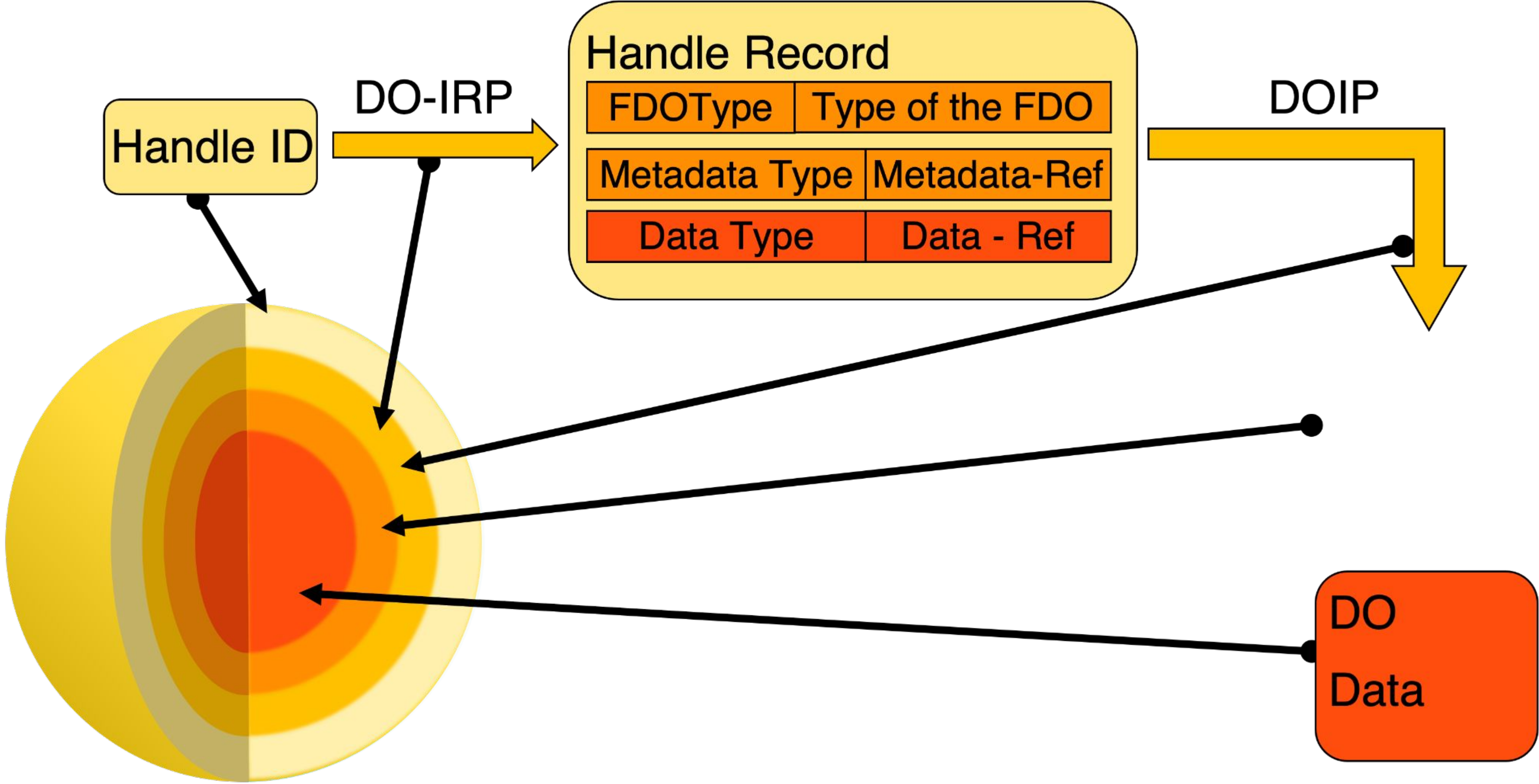
$$F (\text{DOID}, \text{PARAMETER}) = \text{RESPONSE}$$



Implementing FDO Using Handle



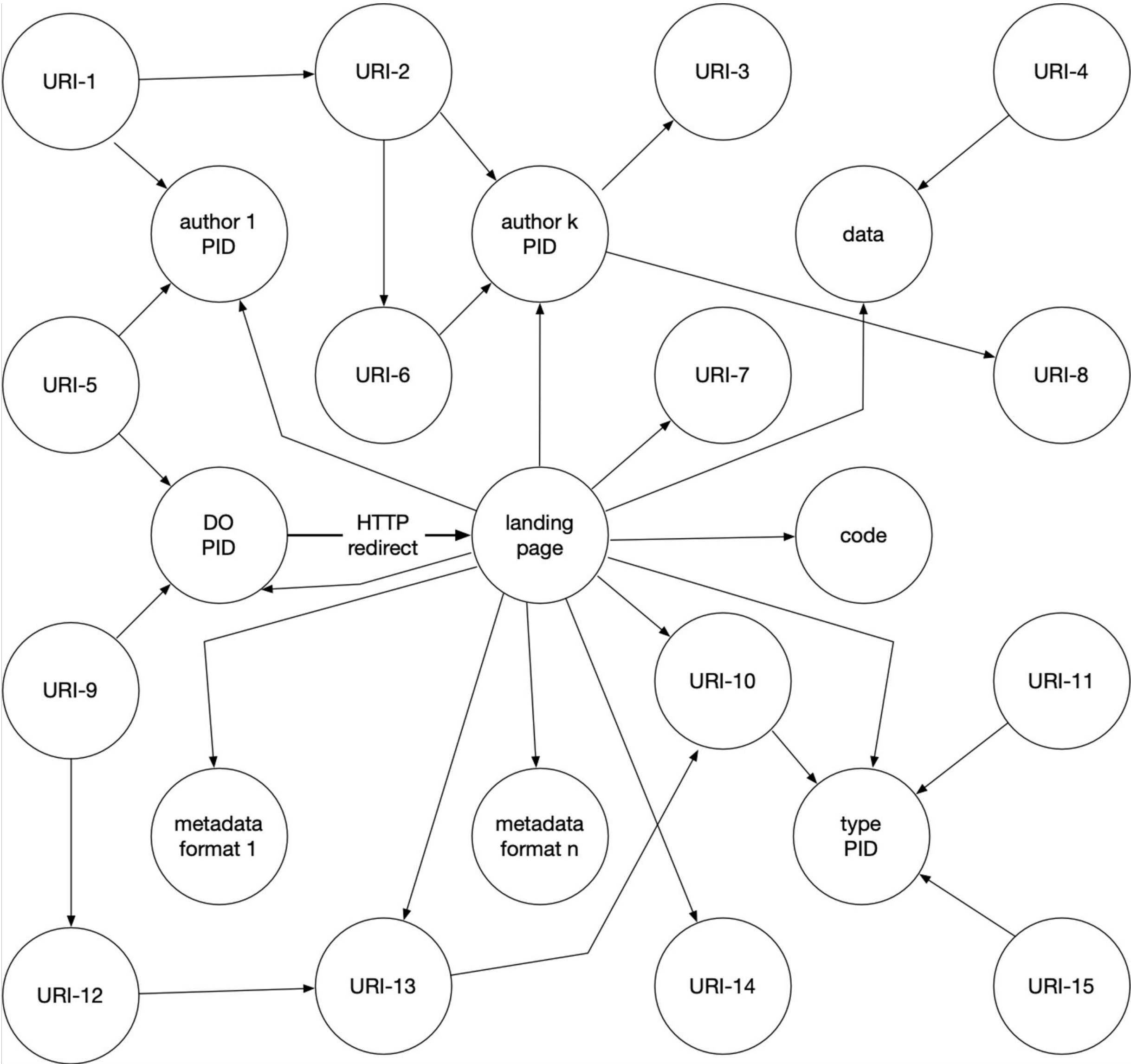
Implementing FDO Using Handle/DOIP



FDO - Signposting Flavour

- * Signposting, like FDOs has for purpose to associate identifiers metadata, data and access operations found on the web to make data more descriptive, findable, accessible, interoperable and reusable.
- * Signposting specifies approaches to create and make accessible relationships between distributed web resources thereby logically connecting identifiers, metadata, data and operation resulting in an FDO whose contents are distributed around the web.
- * The relationships between FDOs and Signposting may not be obvious because Signposting does not start from an abstract model but rather from a very concrete problem related to how research objects exist on the web.

FDO Resources: Someplace in the Web Graph

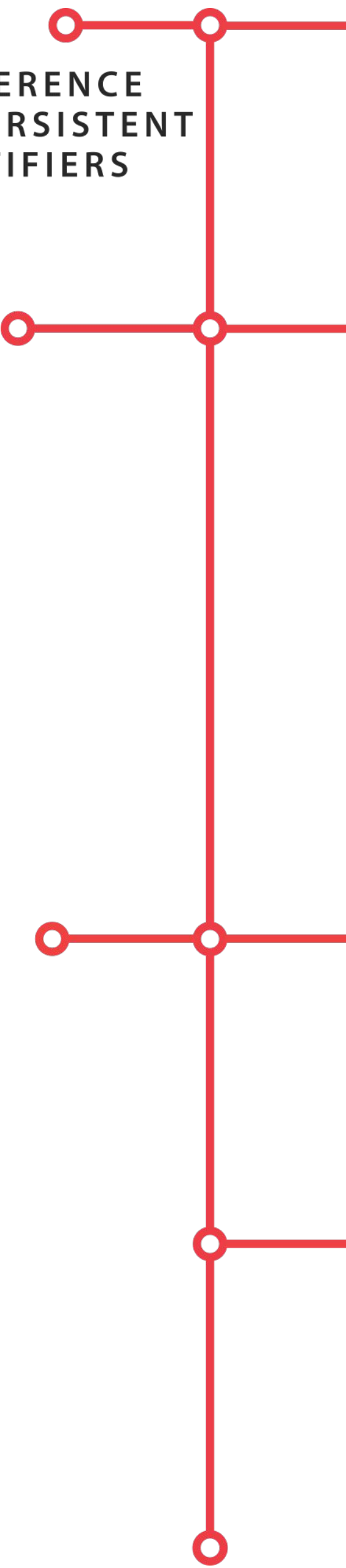
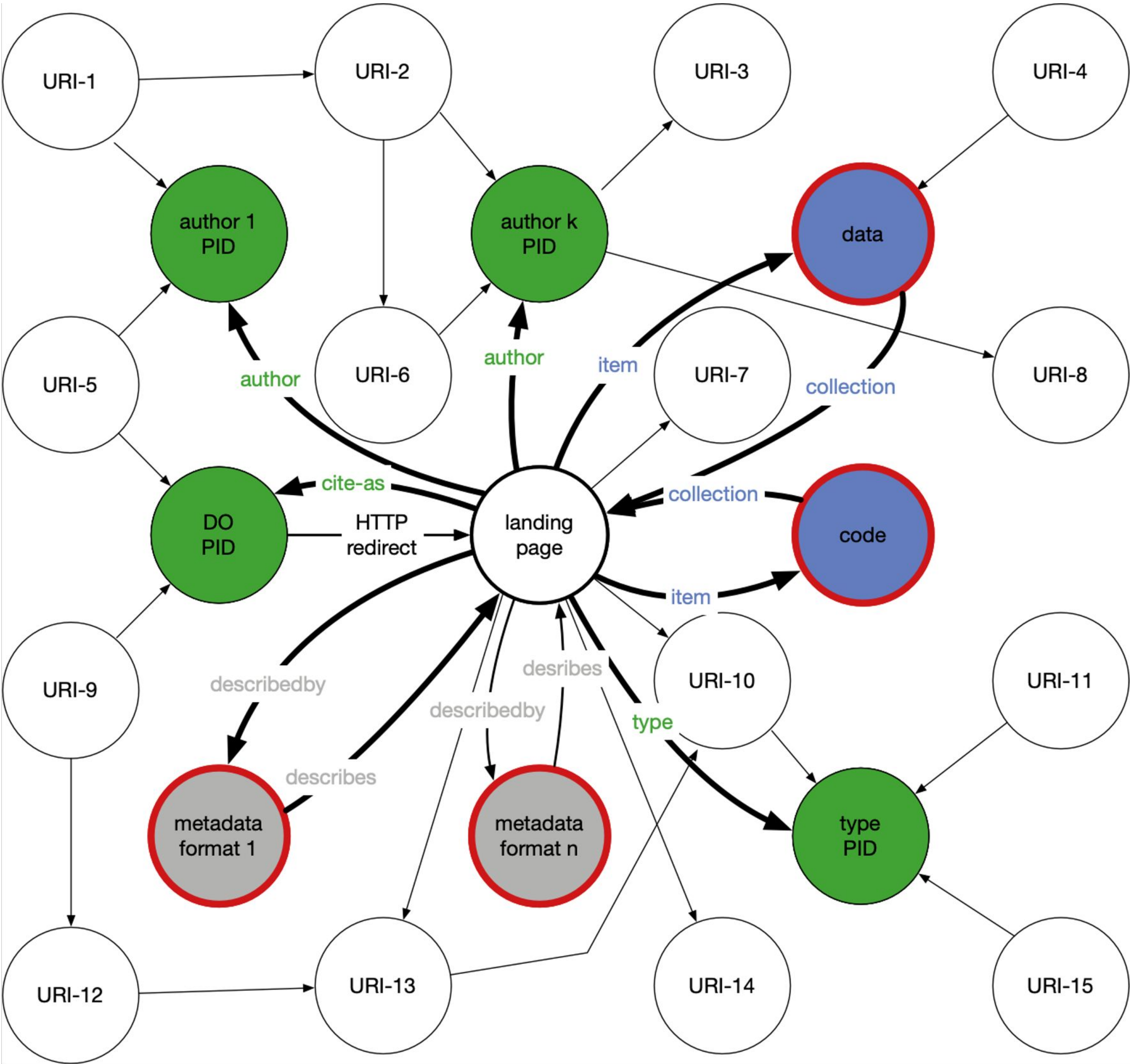


Signposting - A Graph of typed web links

Signposting proposes the provision of typed web links that lead from the landing page to other resources of the Digital Object.

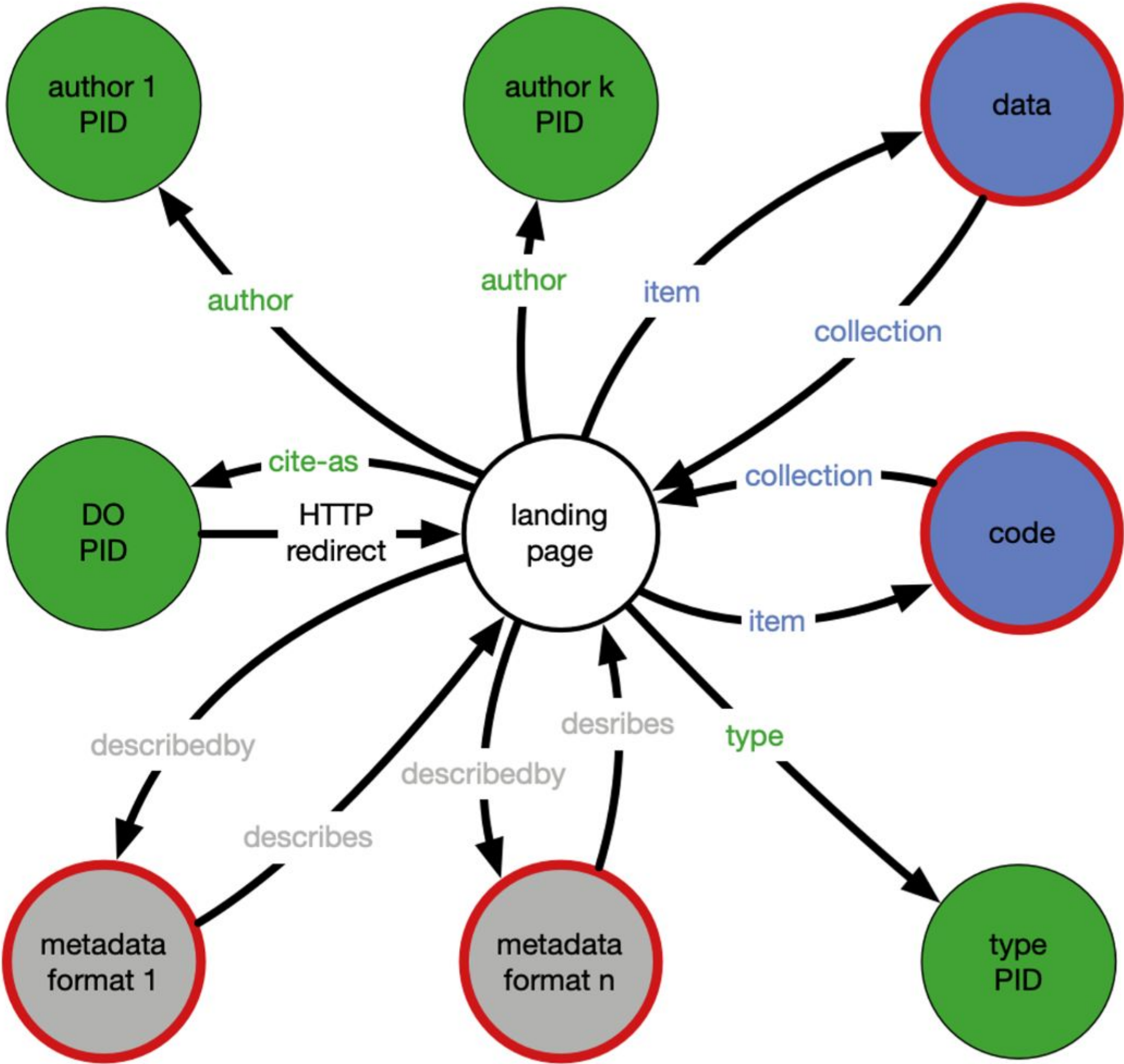
- **“item”** links are used to link to content resources
- **“describedby”** links are used to link to metadata resources;
- **“cite-as”** to link to the PID of the Digital Object,
- **“author”** to link to identifiers of authors,
- **“type”** to link to URIs that convey what kind of Digital Object is involved,
- **“license”** to link to license URIs, etc.

FDO Resources: Someplace in the Web Graph



FDO Resources: Someplace in the Web Graph

-  **DIGITAL OBJECT**
Data, code and other research outputs
-  **IDENTIFIERS**
Persistent and unique (PIDs)
-  **METADATA**
Contextual documentation
-  **STANDARDS & CODE**
Open, documented formats



Signposting & FAIR Signposting

Use standards-based REST/HATEOAS approach:

- Typed web links (1) available via landing page, content resources, metadata resources
- Typed links are a small subset of IANA-registered relation types (2) defined in formal specifications
- FAIR Signposting (3) is a concrete Implementation Guideline for Signposting aimed at uniformity across systems

(1) RFC8288 – Web Linking ; <https://tools.ietf.org/html/rfc8288>

(2) IANA Link Relations ; <https://www.iana.org/assignments/link-relations/>

(3) FAIR Signposting ; <https://signposting.org/FAIR/>

Accessing the Typed Links

Typed links can be conveyed:

- by-value - in the HTTP Link header
- by-value - in HTML `<link>` element
- by-reference - by using a Link Set document (JSON) - discoverable via a link with the `linkset` relation type
- Approaches can be combined but should not be contradictory

FDO - FAIR Signposting relationship

FAIR Signposting exposes the topology of an FDO on the Web using basic, off-the-shelf technology

- It's not less than that
- It's not more than that

Tools & workflows supporting FDOs in research practice

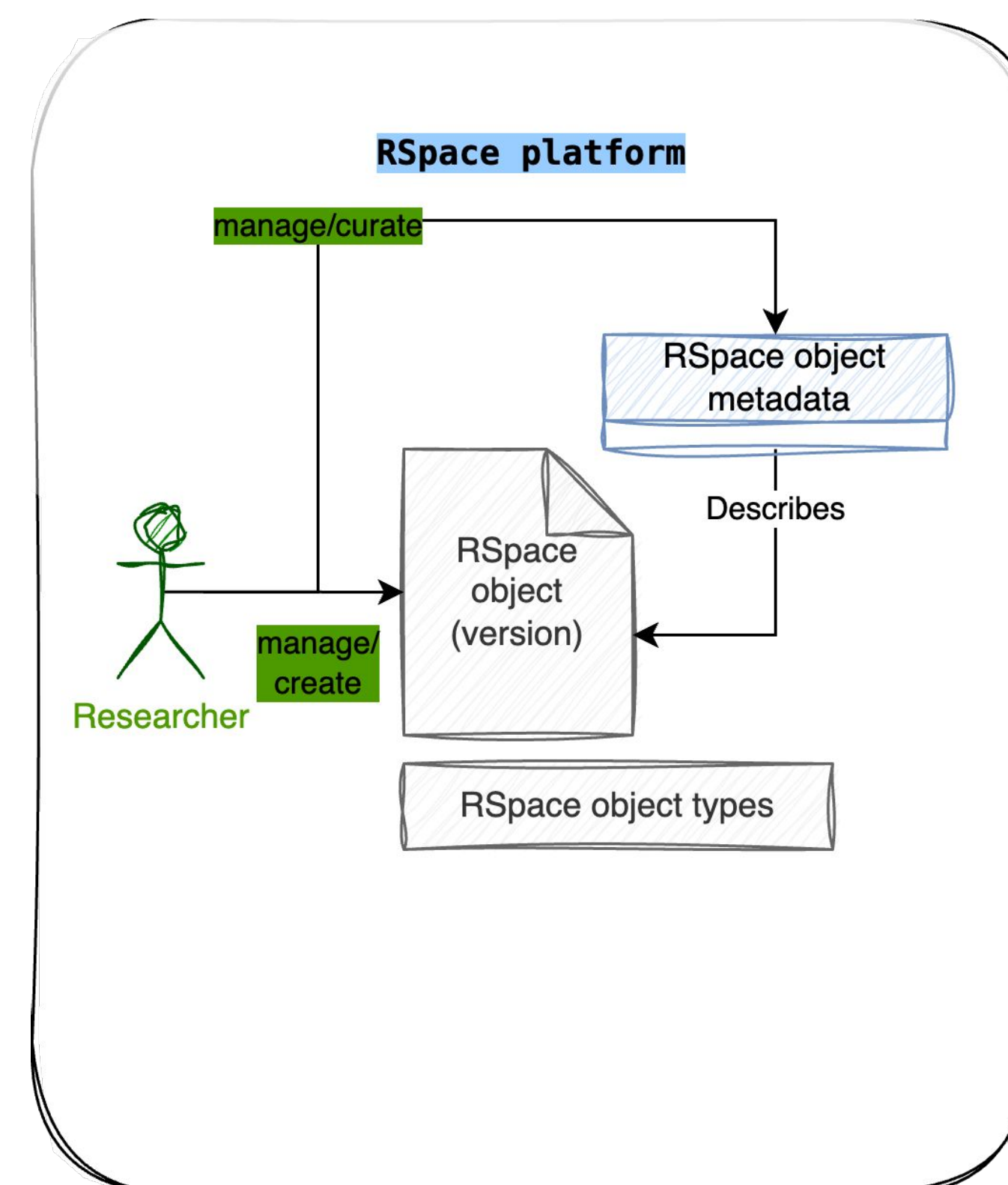
Tilo Mathes, Research Space

Digital workflows enable FDO creation

When researchers use digital tools,
they typically ...

create
import
export
... **and describe**

digital research objects.



Ingredients for FDOs are most likely already there in digital research tools

digital objects

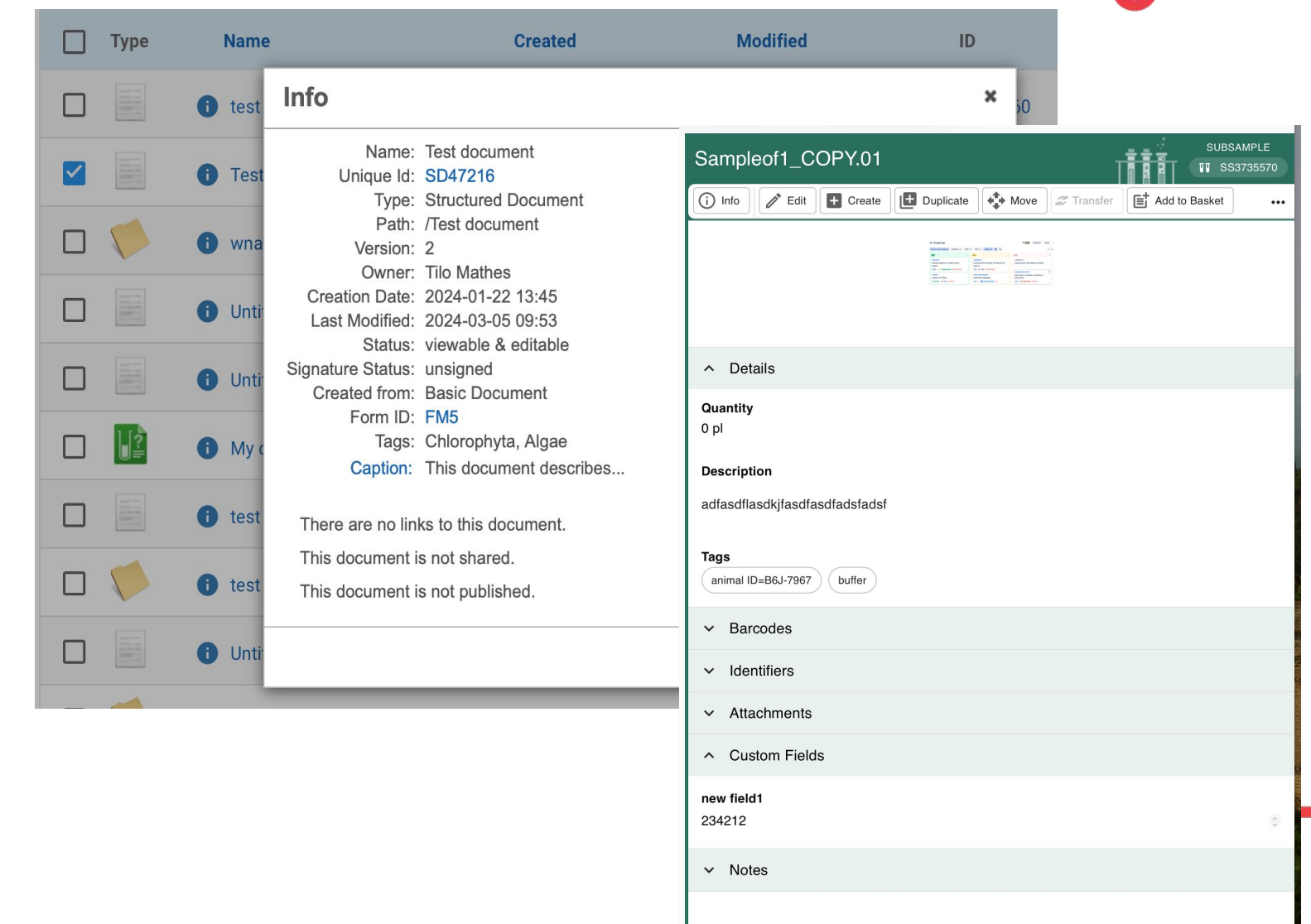
metadata

persistent identifiers

interoperability

Describing digital research objects: Metadata

- Along with the digital objects, metadata flows through a connected digital research ecosystem
- This is an opportunity for digital research tools to **preserve** and/or **improve** research object metadata in the workflows they manage



Challenges for quality metadata:

- Standardization AND specificity
- Avoiding overhead for researchers and curators
- Transfer of metadata between services

- RSpace objects=linked data (versioned objects, documents, ontologies, ...)
- Datacite metadata schema
- Specificity through custom metadata (key value pairs in tags, custom fields in samples, ...)

Resolving digital research objects: PIDs

- Like metadata, persistent identifiers of research objects can flow with them through digital research tools
- Moreover, the creation/curation of PIDs can be integrated into workflows provided by such tools

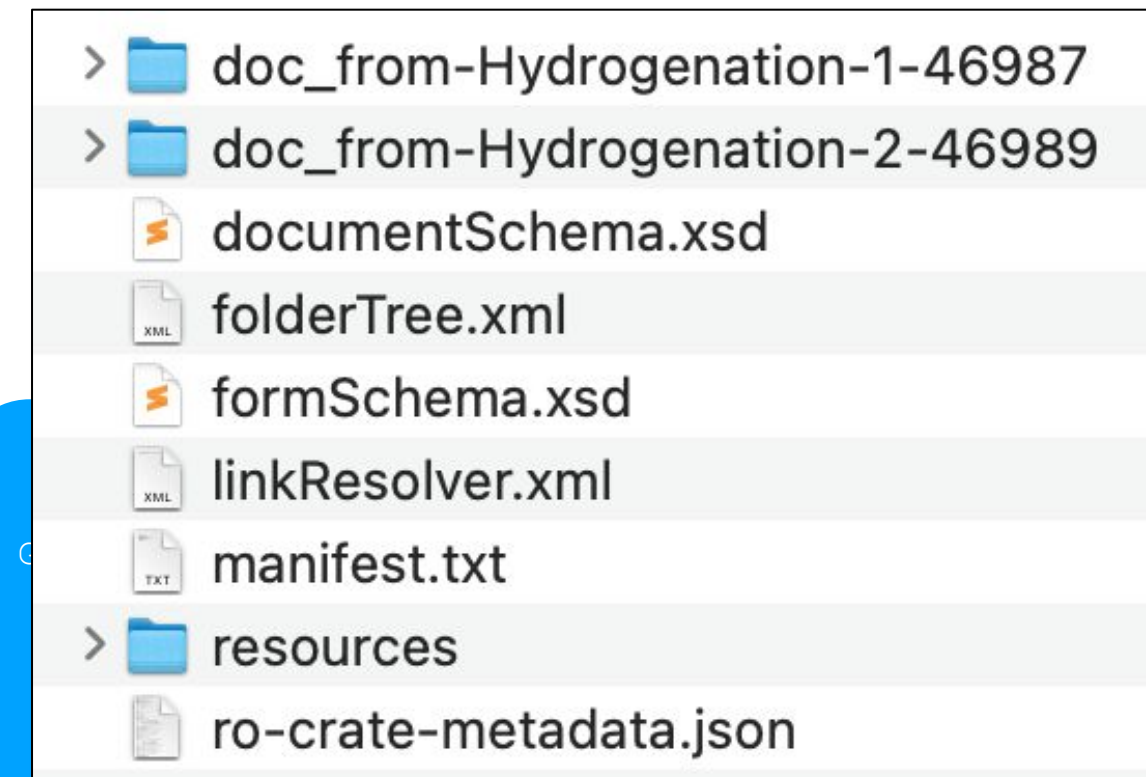
Challenges around PIDs:

- **PID workflows** in digital research tools
- **Avoiding overhead** for researchers and curators
- **Enriching metadata** with PIDs and vice versa

Object type	PID	RSpace
Actor	ORCID	✓
Affiliation	ROR	✓
Physical sample	IGSN (DOI)	✓
DMP	DOI	✓
ELN document	DOI	✓ via repository export
Project	RAID, DOI	⌚ Planned
Instrument	PIDINST	⌚ Planned

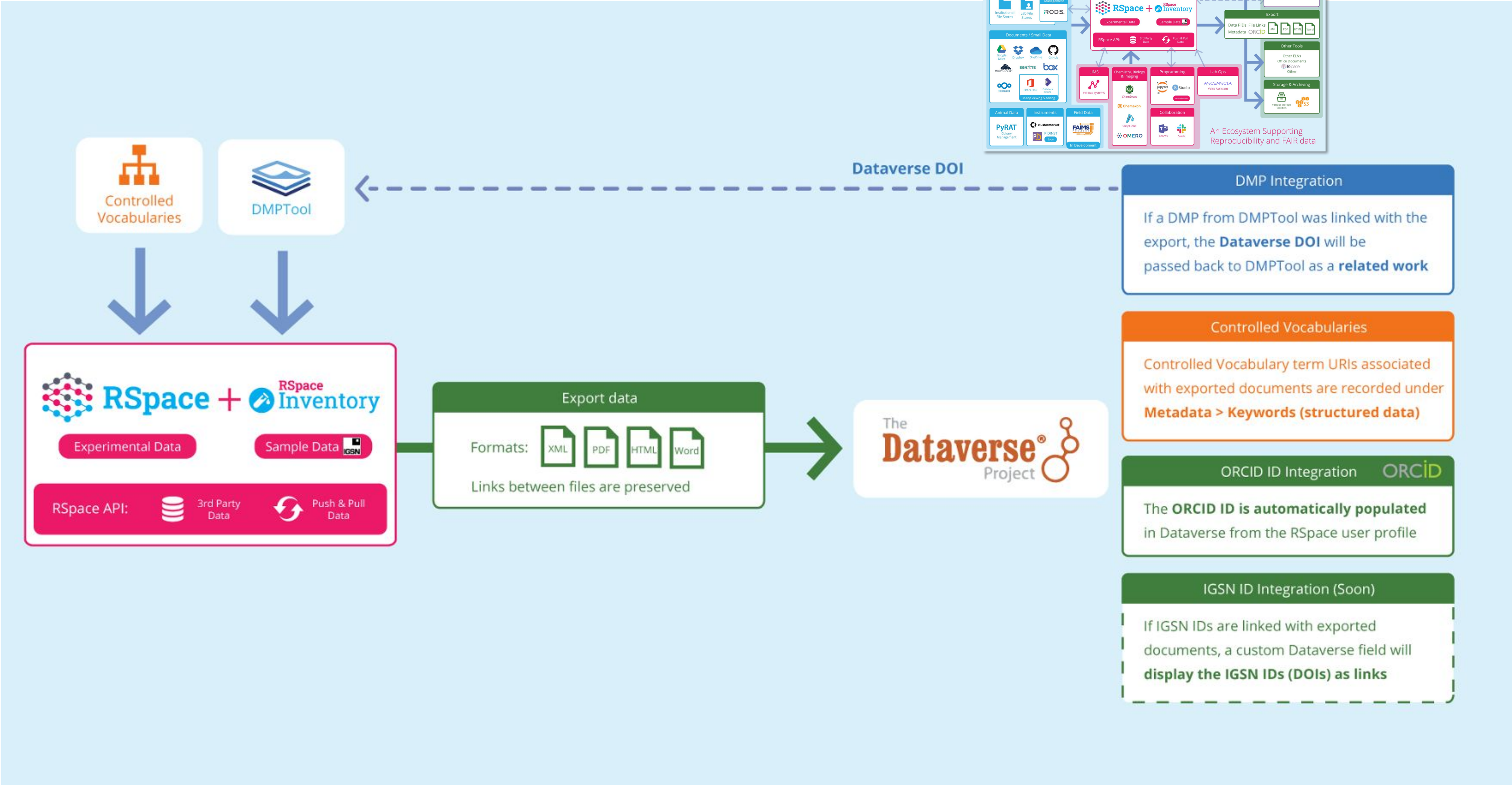
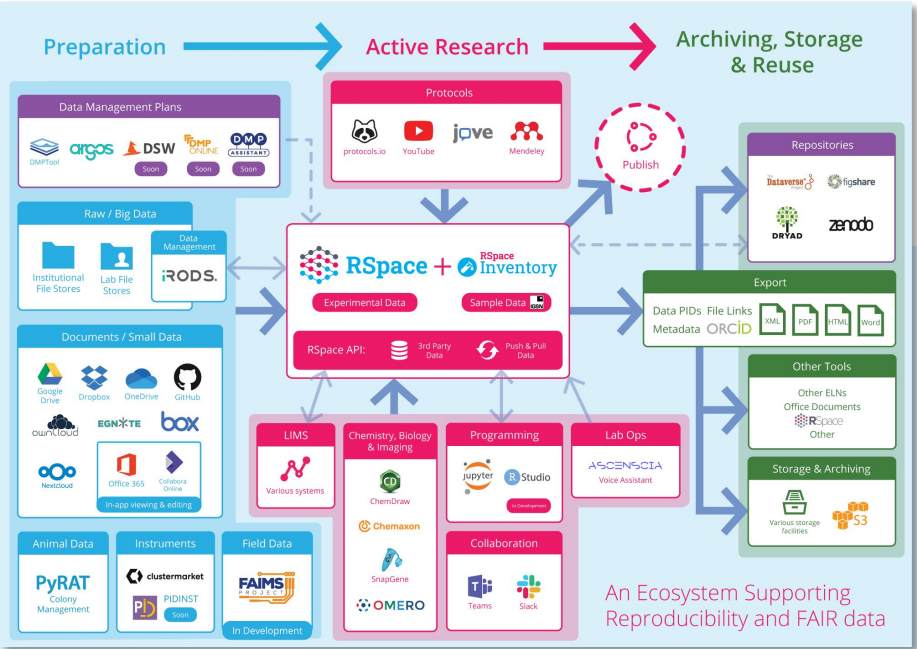
Packaging metadata and PIDs: RO-crate

- The RO-crate format allows to describe a set of digital assets as a single object, yet with explicit metadata on all digital assets
- Digital research tools can create RO-crate.JSON files, e.g. in an export workflow, that contain standardized metadata and PIDs to enable creation of FDOs
- An alternative to API based interoperability between tools



```
"description" : "A Form used to create an RSpace Structured Document",  
"dateCreated" : "06:12:23:07:07:10",  
"dateModified" : "06:12:23:07:10:15",  
"@id" :  
"/doc_from-Hydrogenation-2-46989/doc_from-Hydrogenation-2-46989_form.xml",  
"@type" : "File"  
}, {  
  "description" : "The RSpace user who exported this data",  
  "alternateName" : "Tilo Mathes",  
  "@id" : "https://orcid.org/0000-0002-9694-748x",  
  "email" : "tilo.mathes@researchspace.net",  
  "givenName" : "Tilo",  
  "familyName" : "Mathes",  
  "@type" : "Person"  
}, {  
  "name" : "RSpace",  
  "url" : "https://www.researchspace.com/",  
  "@id" : "https://ror.org/038xqyz77",  
  "email" : "info@researchspace.com",  
  "@type" : "Organization"
```


Example: Metadata enrichment in a repository export workflow



Example: Data management workflow integrating multiple tools/services

Yoda:iRODS:RSpace

Semi-automatic metadata
enrichment

in an electronic laboratory
notebook workflow

integrated into institutional
RDM infrastructure

Researcher selects data to add to
documentation in RSpace

Data is uploaded directly to iRODS for
storage and discoverability

Metadata/PIDs of RSpace document is
passed to iRODS (author, experiment
ID, project ID, tags, instrument, ...)

Yoda metadata (e.g., ePIC,) is passed
back to RSpace along with iRODS file
ID

RSpace document receives metadata
rich link to data on remote location

*Inspired by **Fabian Monheim** (Leibniz Institute on Aging (FLI), Jena)

FLI uses **iRODS**, **Yoda**, and **RSpace** in their research data ecosystem

A suggested FDO test run in RSpace:

Increase the findability and reusability of samples with FDOs

Sample management in RSpace

- User and workflow centric design
- Rich metadata (DataCite schema + custom fields)
- PID support: IGSN
- Sample landing pages managed by RSpace instance
- Indexed in DataCite commons

Planned PID and metadata related improvements

- Generalized RO-crate export
- Enrich metadata with instrument identifiers (PIDINST)
- Enrich metadata with project identifiers (RAID)

Next: creating FDOs for RSpace samples

The screenshot displays the RSpace interface for a sample named 'Chlorella 2'. The top bar is green with the sample name and a 'SAMPLE' label with the ID 'SA262298'. Below the bar is a toolbar with buttons: Info, Edit, Create, Duplicate, Move, Transfer, and Add to Basket. The main content area is divided into sections: Overview, Identifiers, and a Preview Image section. The Overview section shows the sample name, owner (Rob Principal), and sample template (Sample). The Identifiers section shows a table with one identifier: 10.82316/f4b4-eb43, Type: IGSN, State: Draft. Below the table are buttons for PREVIEW, PUBLISH, and DELETE. A message at the bottom states: 'This IGSN ID is a Draft. Metadata can be specified, but no information is publicly available. See IGSN Documentation for details'.

Chlorella 2

SAMPLE SA262298

Info Edit Create Duplicate Move Transfer Add to Basket

Overview

Name Chlorella 2 SA262298

Owner Rob Principal

Sample Template Sample IT163840

Preview Image Tap to view at full resolution, scroll to exit (or pinch on mobile)

Identifiers

+ Create a new Identifier

Hide all Identifiers associated with this item.

To update any details, press Edit first.

Identifier	Type	State	Hide
10.82316/f4b4-eb43	IGSN	Draft	^

PREVIEW PUBLISH DELETE

This IGSN ID is a Draft. Metadata can be specified, but no information is publicly available. [See IGSN Documentation for details](#)

Some questions tool providers may have around creating FDOs

Which FDO **implementation** is most suitable for their use case(s)?

Should research tools create FDOs **directly or as part of another service?**

What are the common requirements for an FDO creation workflow that will work with future FDO implementations?

(How) Do tools share the management of an FDO?

...

Open discussion: What do you think about FDOs?

Moderator: Tibor Kálmán, GWDG

The floor is open!

Some starting-point questions, if needed:

- * Will FDOs fulfill expectations on supporting more efficient research data processing?
- * Can FDOs alleviate concerns related to autonomous analyses of data by AI and ML?
- * Are FDOs safe to use?

Wrap-up: ways to engage further with FDOs

Maggie Hellström, ICOS Carbon Portal

Learn more!

Some examples of organisations and projects working with FDOs

- * The FDO Forum  
- * Research Space 
- * RDA FAIR Digital Object Fabric Interest Group  
- * FDO One Project (with international Testbed) 

Browse documentation

- * Public documents of the FDO Forum 
- * FAIR Digital Object Framework 
- * Slides & recordings from the 2024 FDO Summit conference 

Thank you for joining us today!