PiD fest CONFERENCE ON PERSISTENT IDENTIFIERS 2024

11–13 June, Prague, Czechia

Characteristics of PID Policies: A Comprehensive Approach to Supporting Assessment

Josefine Nordling

CSC

Natascha van Lieshout

SURF

Co-funded by the European Union

MINISTRY OF EDUCATION, YOUTH AND SPORTS

NTK
50°6'14.083"N, 14°23'26.365"E
Národní technická knihovna
National Library of Technology

NCRIS
National Research Infrastructure for Australia
An Australian Government Initiativ

ARDC
Australian Research Data Commons

Presentation outline

- Setting the scene
  - The FAIR-IMPACT project
  - EOSC PID Policy & actors
- Compliance assessment framework for PID policies
  - The tool
  - Criteria
  - Conceptual model
- Why assess PID policies
  - Showcasing the benefits involved
- Supporting PID policy assessment
  - Two fictional PID service providers

eosc | FAIR-IMPACT
Expanding FAIR solutions across EOSC

FAIR-IMPACT
A bird's eye view

Call HORIZON-INFRA-2021-EOSC-01-05

Enabling discovery and interoperability of federated research objects across scientific communities

Expanding FAIR solutions in Europe

Partly following up on FAIRsFAIR

EU funded project

Coordination and Support Action

10 million euro

36 months, starting 1 June 2022

28 partners and affiliate entities

From 10 EU member states: NL, FI, FR, DK, IT, DE, ES, NO, BE, RO

and the UK

FAIR-IMPACT overall objective

WHAT:
to realise a FAIR EOSC by supporting the implementation of FAIR-enabling practices across scientific communities and research outputs at a European, national, and institutional level;

HOW:

- identifying current and emerging components for enabling FAIR (practices, policies, tools & technical specifications);
- translating viable solutions, guidelines and frameworks that have been developed for one domain or research output and supporting their application in others;
- taking the next step in implementation by defining the support, governance, and coordination mechanisms required to ensure the continuous function of FAIR-enabling practices in the EOSC.

FAIR-IMPACT project design

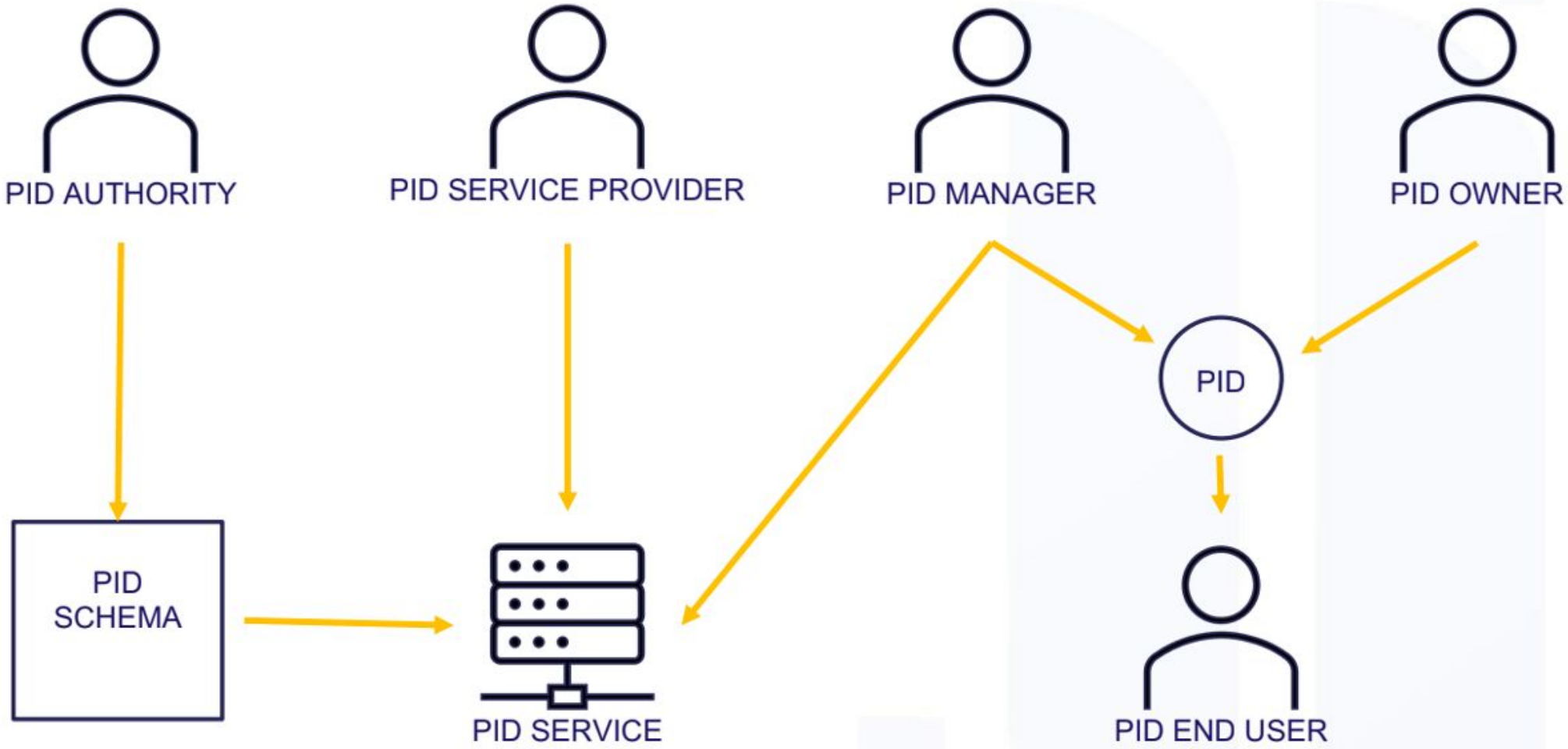
EOSC PID Policy guiding our work

- Published in 2020
- Authored by the EOSC FAIR WG and the EOSC Architecture WG
- Defines roles and components in PID infrastructures
- Written for senior decision makers for potential EOSC infrastructure and service providers
- Defines a set of expectations through definitions, guidelines & usage requirements

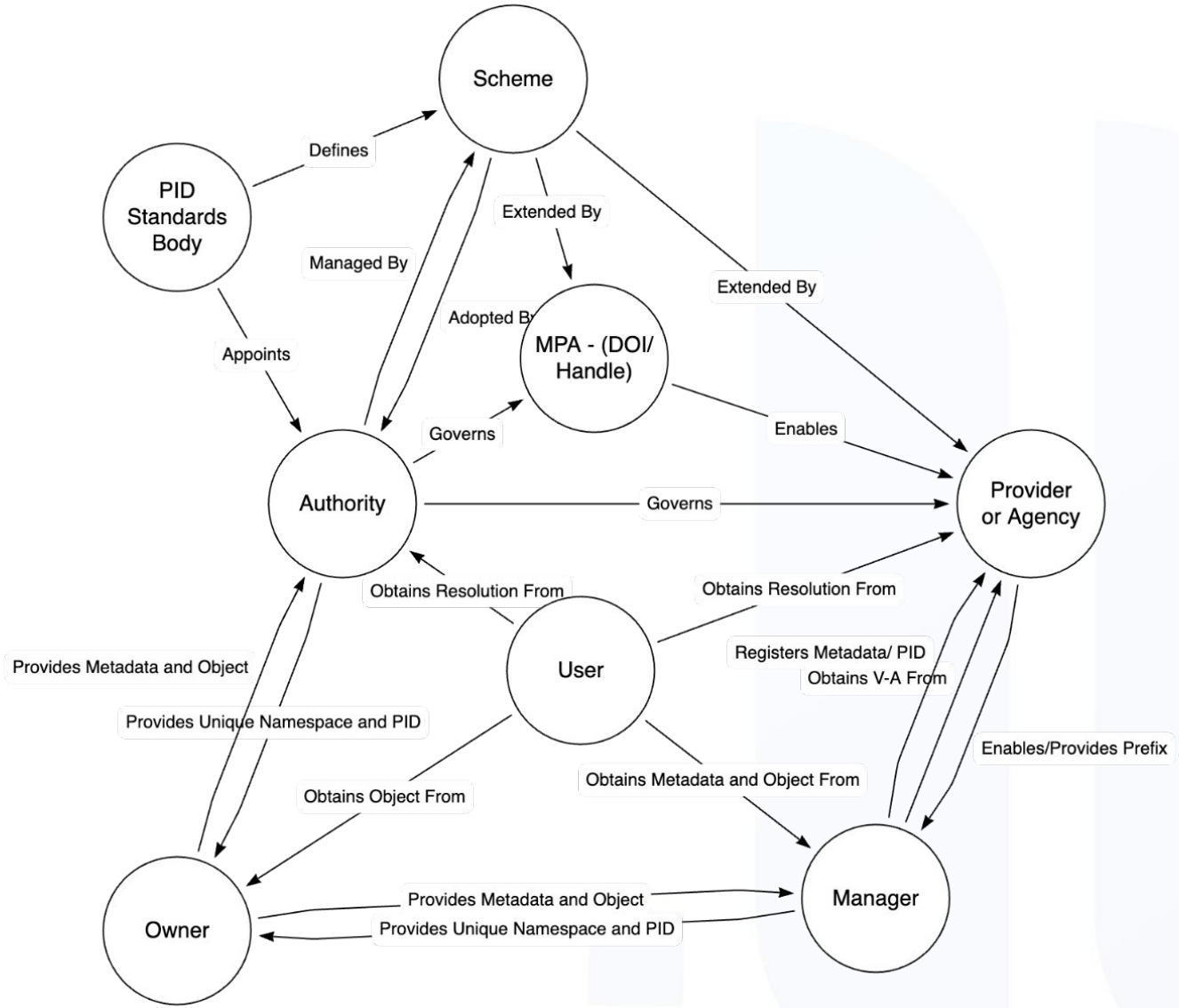
EOSC PID Policy: Guiding principles

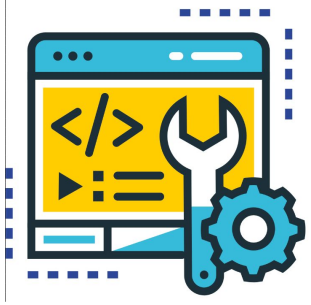
1. Viable, trusted PID infrastructure suitable for the long-term
2. Wide use case coverage
3. PIDs as the preferred method for reference
4. Enable a FAIR research environment
5. Interoperable PID ecosystem
6. Technological independence for flexibility
7. Supports mature and established PID practices, technologies, providers etc.
8. Sustainability & persistence
9. Innovative tools and services
10. Regular updates/review of policy

EOSC PID Policy: Roles and Components



Roles


Scheme
Defines
PID Standards Body
Extended By
Managed By
Extended By
Adopted By
MPA - (DOI/ Handle)
Appoints
Governs
Enables
Authority
Governs
Provider or Agency
Obtains Resolution From
Obtains Resolution From
Provides Metadata and Object
User
Registers Metadata/ PID
Obtains V-A From
Provides Unique Namespace and PID
Enables/Provides Prefix
Obtains Object From
Obtains Metadata and Object From
Owner
Provides Metadata and Object
Provides Unique Namespace and PID
Manager



CAT
EOSC Compliance Assessment Toolkit

Objectives:

1. Allow consistent and unambiguous encoding of assessment principles, objectives, criteria, metrics and tests using a vocabulary developed for the toolkit
2. Enable the recording of PID policy compliance for a range of important actors in the ecosystem. Some assessments are made by the administrators of the CAT on behalf of the community, while the majority of service providers and managers will be able to conduct self-assessments.

Description:

The FAIRCORE4EOSC Compliance Assessment Toolkit (CAT) will assist actors in the PID ecosystem with assessment of their compliance with policy. The toolkit is by design capable of accommodating a wide variety of compliance assessment use cases but will initially focus on PID compliance only.

Key Definitions

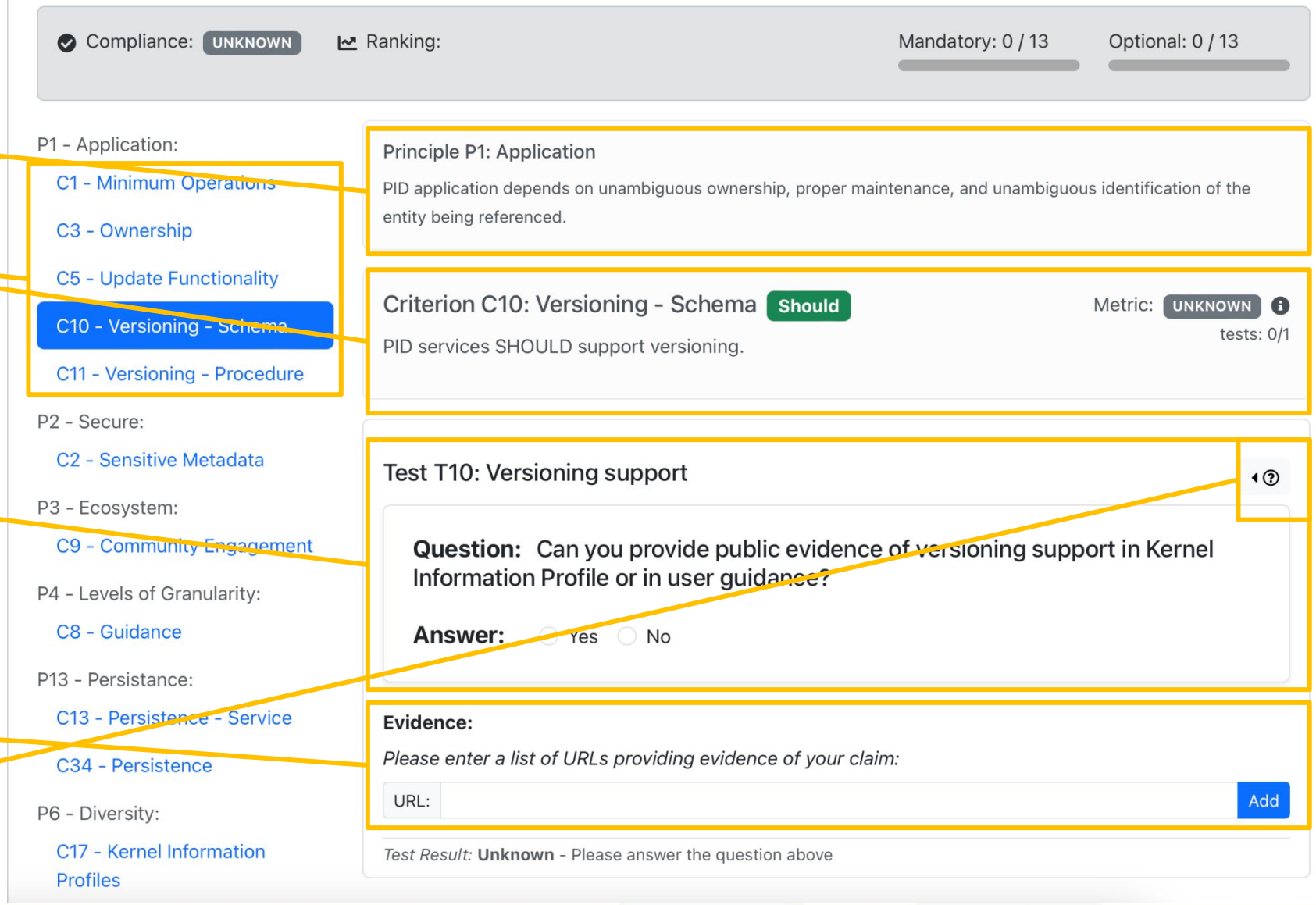
Principle

Criteria

Test

Evidence

Guidance



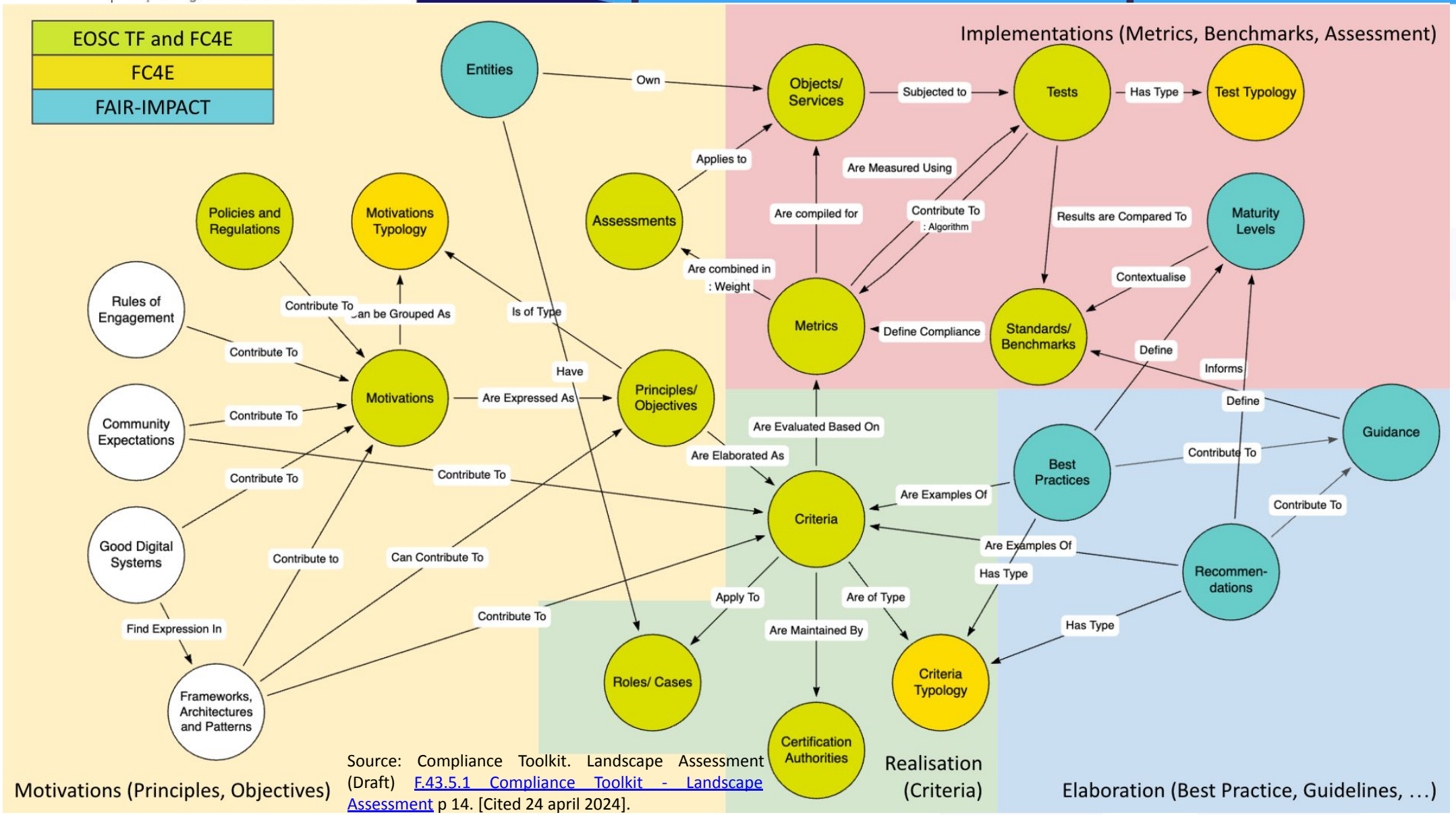
Mapping Principles and Criteria to Roles

#	Criterion	Imperative	Scheme	Authority	Service Provider	Manager	Owner
C1	Minimum Operations	SHOULD			✓		
C2	Sensitive Metadata	MAY		✓	✓		
C3	Ownership	MUST		✓	✓		
C4	Maintenance	SHOULD					✓
C5	Update Functionality	MUST			✓	✓	
C6	Ownership Transfer	SHOULD				✓	
C7	Resolution Integrity	MUST				✓	
C8	Guidance	SHOULD			✓		
C9	Community Engagement	SHOULD			✓		
C10	Versioning - Schema	SHOULD			✓		
C11	Versioning - Procedure	SHOULD			✓	✓	
…	…	…	…	…	…	…	…

6/12 CAT criteria for PID Managers

#	Criterion	Description	Imperative
C5	Update Functionality	The PID Manager **MUS**T provide the functionality required to maintain PID attributes. (Also mapped to the Actor: Service)	MUST
C6	Ownership Transfer	The PID Manager **SHOULD** provide policies and contractual arrangements for transfer of ownership should the owner no longer be able to assume responsibilities in compliance with the policy.	SHOULD
C7	Resolution Integrity	The PID Manager **MUST** maintain the integrity of the relationship between entities and their PIDs, in conformance to a PID Scheme defined by a PID Authority.	MUST
C11	Versioning - Procedure	PID Services and PID Managers **SHOULD** have clear versioning policies. (Also mapped to the Actor: Service)	SHOULD
C14	Resolution Authenticity	PID Manager **MUST** ensure that the entity remains linked to the PID. In case that the entity being identified is deleted or ceases to exist, tombstone information needs to be included in the PID attribute set.	MUST
C16	Digital Representation	Physical and conceptual entities **MUST** be represented via a digital representation (e.g. landing page, metadata, attribute set, database index) to have a presence in the digital landscape.	MUST

Conceptual Model for Compliance Assessment



Source: Compliance Toolkit. Landscape Assessment (Draft) F.43.5.1 Compliance Toolkit - Landscape Assessment p 14. [Cited 24 april 2024].

Motivation for Using the CAT

Barriers to Using the CAT

Snapshots

PID policies, strategies, roadmaps, and guidelines are in the early stages and under rapid development

Novel Framework

Need to understand the novel framework and theory of the CAT

New Tech

Need to understand how to physically use the CAT

Unknown Repercussions

Fear of sharing a “failed” assessment leading to loss of support or funding in future

Guinea Pig

Hesitant to be “the first” by which everyone else is measured

Lower the Barriers

Gal ID

Home | About | Community Forum | Documentation | Fees | PID Policy | Contact Us

As open as possible, as closed as necessary.

The Genovia Associated Library Identifier (Gal ID) is a Persistent Identifier (PID) service developed in 2001 by the Genovia National Library (GNL) for knowledge preservation and to help combat link rot and content drift in an increasingly digital world. Originally a national infrastructure, Gal ID is now a well-established actor in the international Academic and Heritage landscape.

# of Minted PIDs	# of Resolvable PIDs	Uptime
30126	29738	99.9%

DISCLAIMER: This website is entirely a work of fiction meant to serve as an example PID service provider for assessing compliance with the Compliance Assessment Toolkit

Home | Documentation | Policies | Contact Us

The most dependable PID service ever.

Welcome to the newly launched Babadook Archiving Department ID (BAD ID) Persistent Identifier service. BAD ID is one of the few projects awarded with a prestigious 2-year Great Depression Grant (GDG), despite this past year of austerity. We bring innovative concepts for the FAIRification of data, solving the decades-old problem of link rot and content drift.

Governance Structure

Samuel Vanek

President

Validation Request id: 30

Requestor

User id: [redacted]
User name: [redacted]
User surname: [redacted]
User email: [redacted]

Organisation

Id: 00xbe3815 - [ROR]
Name: DEMO Consultants
Website: http://www.demobv.nl/

Text

Roles

User role in organisation: advisor
User requests as Actor with: PID Service Provider

Status

Created on: [redacted]

✔ APPROVED

Approved on: [redacted]
Approved by: [redacted]

Compliance: UNKNOWN Ranking: Mandatory: 0 / 13 Optional: 0 / 13

P1 - Application:
- C1 - Minimum Operations
- C3 - Ownership
- C5 - Update Functionality
- C10 - Versioning - Schema
- C11 - Versioning - Procedure

P2 - Secure:
- C2 - Sensitive Metadata

P3 - Ecosystem:
- C9 - Community Engagement

P4 - Levels of Granularity:
- C8 - Guidance

P13 - Persistance:
- C13 - Persistence - Service
- C34 - Persistence

P6 - Diversity:
- C17 - Kernel Information Profiles

Principle P1: Application

PID application depends on unambiguous ownership, proper maintenance, and unambiguous identification of the entity being referenced.

Criterion C10: Versioning - Schema Should

Metric: UNKNOWN
tests: 0/1

PID services SHOULD support versioning.

Test T10: Versioning support

Question: Can you provide public evidence of versioning support in Kernel Information Profile or in user guidance?

Answer: ○ Yes ○ No

Evidence:

Please enter a list of URLs providing evidence of your claim:

URL: Add

Test Result: **Unknown** - Please answer the question above

A Starting Point

Some previously raised points of discussion:

- How are you publishing and maintaining all this information? A website? Zenodo?
- Is technical documentation sufficient evidence of commitments?
- Where should specific elements be documented? Does this need consistency?
  - One big massive document vs smaller documents that reference each other?
  - How does versioning of the documents themselves work?
- Should be we set a minimum standards on technical requirements like encryption or is anything enough?
- Shouldn’t alignment with other stakeholders/policy be mandatory for compliance?
  - How do we assess who needs to align with what?
- Should we start requiring services to maintain support teams?
- How vague vs specific should criteria be?

Try it yourself!

Gal ID

https://tests.cat.argo.grnet.gr/galid/

CAT demo

https://cat.argo.grnet.gr/

BAD ID

https://tests.cat.argo.grnet.gr/badid

Thank you! Questions?