

Deliverable 2.2

Compliance Assessment Toolkit

Dissemination Level	PU
Due Date of Deliverable	30/04/2024, Month 23
Actual Submission Date	30/04/2024
Work Package	WP2 - Compliance Assessment Toolkit
Task	T2.2
Type	Report
Version	V 1.0
Number of Pages	p.1 – p.29

Deliverable Abstract

The report provides an overview of the beta version of the Compliance Assessment Toolkit, which was readied for release at the end of March 2024.

The information in this document reflects only the author's views and the European Community is not liable for any use that may be made of the information contained therein. The information in this document is provided "as is" without guarantee or warranty of any kind, express or implied, including but not limited to the fitness of the information for a particular purpose. The user thereof uses the information at his/her sole risk and liability. This deliverable is licensed under a Creative Commons Attribution 4.0 International License.



DELIVERY SLIP

Contribution	Name	Partner/Activity	ORCID
Lead author(s)	W Hugo	DANS	0000-0002-0255-5101
Contributor(s)	W Steinhoff	DANS	0000-0003-1106-8441
	N v. Lieshout	SURF	0000-0001-8025-8412
	M Buys	DataCite	0000-0001-7234-3684
	T Zamani	GRNET	0000-0003-1148-1738
	F van Rijsselberg	DANS	n/a
Reviewer(s)	A Märkälä	CSC	n/a
	J Kesaniemi	CSC	0000-0002-3770-0006
	M Gruenpeter	INRIA	0000-0002-9777-5560

DOCUMENT LOG

Issue	Date	Comment	Editor/Reviewer	ORCID ID
v.0.1	06-03-2024	Inception and first draft	W Hugo	0000-0002-0255-5101
v.0.2	14-03-2024	Monitoring & Testing	W Steinhoff	0000-0003-1106-8441
v.0.3	28-03-2024	First draft released for internal review, contributions from members	W Hugo	0000-0002-0255-5101
v.0.4	11-04-2024	First internal review	M Gruenpeter	0000-0002-9777-5560
v.0.5	18-04-2024	Second draft released for internal review	W Hugo	0000-0002-0255-5101
v.0.6	29-04-2024	Second internal review	J Kesaniemi	0000-0002-3770-0006
v.06	29-04-2024	Final draft	W Hugo	0000-0002-0255-5101
v.07	29-04-2024	Editorial review	A Märkälä	n/a
v.08	30-04-2024	Publication version	W Hugo	0000-0002-0255-5101

TERMINOLOGY

Terminology/Acronym	Definition
CAT	Compliance Assessment Toolkit, a service being developed in the FAIRCORE4EOSC project to assist with EOSC PID Policy compliance assessment
Compliance Assessment	The process of determining to what extent a service, object, organisation, or capabilities comply with a set of criteria, based on reproducible tests.
EOSC	The European Open Science Cloud
EOSC PID Policy	The policy being developed by the EOSC PID Policy Task Force to ensure a minimum standard of performance for the PID ecosystem in EOSC
FAIR	Principles based on community expectations in respect of research outputs - findable, accessible, interoperable, and reusable.
FAIR-IMPACT	An EU-funded project that has as its main objectives to identify practices, policies, tools and technical specifications to guide researchers, repository managers, research performing organisations, policy makers and citizen scientists towards a FAIR data management cycle. The focus

	will be on persistent identifiers (PIDs), metadata, ontologies, metrics, certification and interoperability.
FAIRCORE4EOSC	The FAIRCORE4EOSC project focuses on the development and realisation of core components for the European Open Science Cloud (EOSC).
GDPR	General Data Protection Regulation: Regulations aimed at protecting the right to privacy of individuals in the EU in a digital context.
HE	Horizon Europe funding programme
Landscape Assessment	A milestone report produced by FOARCCORE4EOSC WP2 to assess the scope of case studies that may influence a conceptual model for compliance assessment.
PID	Persistent identifier: generally expected to be unique, resolvable, and persistent, but many other features and performance aspects apply.
RDA	The Research Data Alliance, an international organisation developing standards, recommendations, and best practices in respect of research data management using voluntary contributions.
TRUST	The principles that describe the community expectations in respect of trustworthy repositories.
PID Stack	A PID Stack defines the actors in the ecosystem to deliver a specific PID service. For example, DataCite DOIs are based on the Handle Schema and system, and additionally require functioning infrastructure or capacity from DONA Foundation, the International DOI Foundation, and DataCite to work.
Intrinsic identifier	Identifiers that are bound to the designated object, they do not need a register, only agreement on a standard. They can be independently computed by anyone, using freely available software.

Table of Contents

1	Context	2
2	Status of Implementation.....	4
2.1	EOSC Compliance Assessment Toolkit (CAT)	4
2.2	Conceptual Model Status.....	4
2.3	CAT Minimum Viable Product and Implementation Status.....	4
2.4	Discussion of Deviations.....	6
2.5	CAT requirements overview.....	7
2.6	CAT Overview and the status of the development.....	11
2.6.1	CAT API	11
2.6.2	CAT User Interface (CAT UI).....	12
3	Conceptual Model Feedback and Community Engagement.....	15
4	Vocabularies and Registries.....	16
5	Acceptance Testing.....	18
6	Documentation.....	20
7	Challenges.....	21
7.1	Assessment data.....	21
8	Expected Improvements and Refinements.....	22
8.1	Additional Guidance and Best Practices Support.....	22
8.2	User Interface Improvements.....	22
8.3	Vocabulary Server and PID MetaResolver Integration	23
8.4	Automated Monitoring	23
	References	25
	Annexure A: Supplementary Materials	27
	Annexure B: Presentations and Community Engagements.....	29

List of Figures

Figure 1	– Example Assessment Screen.....	14
Figure 2	– Documentation Available to Users.....	20
Figure 3	– Example of UI Improvements.....	22

List of Tables

Table 1	– High-Level Roadmap of Compliance Assessment Toolkit MVP	5
Table 2	– Requirements Overview	7
Table 3	– Compliance Assessment Toolkit Requirements Summary - MVP.....	8
Table 4	– Compliance Assessment Toolkit Vocabularies and Registries.....	16
Table 5	– CAT UI Test Cases Overview	18

Executive Summary

Persistent Identifiers (PIDs) are foundational to establishment of the digital infrastructure in support of the broad research enterprise, and in recognition of this critical role, EOSC (The European Open Science Cloud) published a policy whereby the quality of PID provision and application can be standardised and improved. The FAIRCORE4EOSC Compliance Assessment Toolkit (CAT) is being developed to assist actors in the PID ecosystem with assessment of their alignment and compliance with the EOSC PID Policy, although its design allows it to be used for a wide variety of assessment purposes, including for example FAIR, GDPR compliance, and many more.

This report outlines the implementation of a beta release of the Compliance Assessment Toolkit (CAT), based on requirements and specifications documented in deliverable D.2.1. It documents the following:

- ② Status of implementation: progress with releases, including a minimum viable product (MVP) as a beta release, and a summary of the functions and capabilities that are available to end users and systems in the beta release, and a summary of deviations from original specifications;
- ② A summary on feedback sought and received from the community, both in respect of the conceptual model underpinning the CAT, and limited feedback already received in respect of the beta release;
- ② A review of the status of vocabulary implementation;
- ② Information on testing and documentation in support of the CAT.

There are some challenges to address during the final phase of operationalisation of the CAT, including (1) dealing with sensitivity around public availability of assessments and (2) whether such assessments are endorsed by EOSC or not. These are expected to be mitigated through a programme of engagement with actors in the PID ecosystem and the EOSC Association, as part of Task 2.3.

New ideas and requirements¹, many suggested by early users of the CAT and via community engagement in collaboration with FAIR-IMPACT, will be addressed in the next (penultimate) release, including greater access to and visibility of guidance and best practices, assistance with selection of PID services, and the creation of automated services for monitoring of persistence and resolution. This last requirement will require closer integration with the PID Metaresolver², being developed in FAIRCORE4EOSC WP3, and a codesign session on how this can be done have been scheduled.

The annexures provide a record of community engagements, and of supplementary materials created by the activities reported here.

¹ https://drive.google.com/file/d/1bPkNqZVArVDT_RKAgZ-1HX_KxWNxX8go/view?usp=sharing

² <https://pidmr.devel.argo.gnet.gr/>

1 Context

Persistent Identifiers (PIDs) are foundational to establishment of the digital infrastructure in support of the broad research enterprise, and in recognition of this critical role, EOSC (The European Open Science Cloud) published a policy whereby the quality of PID provision and application can be standardised and improved³. The policy has a number of objectives and formulates a set of principles and attendant criteria whereby alignment or compliance with the objectives and principles can be determined. The policy envisages that PID services will be certified in future through an authority or governance structure associated with EOSC.

The FAIRCORE4EOSC Compliance Assessment Toolkit (CAT) is being developed to assist actors in the PID ecosystem with assessment of their alignment and compliance with the EOSC PID Policy. These actors include all layers of PID provision, management, and end use, and the policy defines the roles of such actors and the performance expectations for each of them. The CAT will play a crucial role in ensuring that

- 🕒 Assessment of policy compliance is performed in a consistent manner, and that assessments are based on documented tests and processes;
- 🕒 It is less likely that compliance will be claimed without evidence or evaluation by promoting the CAT to actors in the ecosystem;
- 🕒 Assessment is not judgmental, but based on a process of self-evaluation (supported by standardised and automated tests where applicable) and is supported by improvement through linking to guidance and best practices.

Deliverable D2.1 (“Compliance Assessment Specification”) (M12)⁴ provided two major inputs for the development of the CAT: definition of a Conceptual Model and the vocabularies that would be required to implement such a model, and a section detailing the requirements and specifications for implementation of the CAT. This publication included suggestions for reformulation and refinement of the policy principles, objectives, and criteria - primarily aimed at defining the actors responsible for implementation unambiguously, clearly stating whether compliance is mandatory or optional, and how compliance will be evaluated. It is supported by a provisional set of tests whereby compliance can be measured, and contains a preliminary set of guidance and best practice materials to aid with improvement of performance in cases of non-compliance.

A beta version of the CAT has been in operation since M18 (December 2023), and this report provides a record of the release. Since the publication of D2.1, WP2 in FAIRCORE4EOSC has engaged the community on numerous occasions to gain feedback on the model, the content envisaged for the assessment of the EOSC PID Policy, and on the beta version of the CAT. Engagements included work with the EOSC PID Policy Task Force, engagement with other Horizon Europe (HE) projects, and close collaboration with FAIR-IMPACT. An in-depth conversation about EOSC PID Policy assessment, best practices in support of it, and use of the CAT formed part of the recent EOSC Winter School in Thessaloniki⁵.

³ European Commission (2020). A Persistent Identifier (PID) policy for the European Open Science Cloud, doi: 10.2777/926037

⁴ Hugo, W., Steinhoff, W., Turner, D., Buys, M., & Zamani, T. (2023). D2.1 Compliance Assessment Specification. Zenodo. <https://doi.org/10.5281/zenodo.10067253>

⁵ <https://eosc.eu/eosc-focus-project/winter-school-2024/> - opportunity area 1 - PIDs

The benefits of these engagements are seen in the following, on which we report in the document:

1. Validation of the conceptual model
2. Confirmation of the way in which the original EOSC PID Policy will be encoded in the CAT
3. Suggestions for improvement to the CAT:
 - a. User interface improvements
 - b. Requests to add more comprehensive support for guidance and PID selection
 - c. Additional tests for persistence and resolution conformation
4. Extension to and refinement of the guidance, best practices and recommendations available to users of the CAT.

The suggested improvements to the CAT are included in extensions to the requirements and specifications for the CAT, and will be deployed in the first post-beta release (planned M26).

Functional compliance with the requirements and specifications are automatically verified using a test automation platform, and the report provides details on the coverage and implementation.

A significant challenge to deployment of a beta version that was not anticipated initially involves the data to be used for demonstration: until such time that EOSC, in one way or another, mandates the use of the CAT for EOSC PID Policy compliance assessment, it would be unwise to publish any real assessments of actors in the ecosystem. The relevant section below details our mitigation efforts.

2 Status of Implementation

2.1 EOSC Compliance Assessment Toolkit (CAT)

The EOSC Compliance Assessment Toolkit (CAT) is based on a conceptual model, implemented as a database and accessible via APIs, and supported by user interfaces via the APIs. The database has been populated with an encoding of the EOSC PID Policy, mapped to the model. The CAT supports the EOSC PID policy with services to encode, record, and query compliance with the policy. Though primarily aimed at machine-actionable operations, the API-based services will be complemented by user interfaces to broaden its use.

The PID policy provisions are linked to guidance and support information, which is currently being refined in collaboration with FAIR-IMPACT. These inputs will be deployed to the CAT as additional data, and the aim is to have it available at M24, after publication of a guidance deliverable being prepared by FAIR-IMPACT.

2.2 Conceptual Model Status

The CAT Conceptual model was developed and published as a part of the first deliverable for WP2 in FAIRCORE4EOSC (Deliverable D2.1 [1]). The model is has now grown to four interrelated sub-models, each representing an aspect of the PID compliance and service provision landscape:

1. The original compliance model, with enhancements suggested by the community and derived from additional case studies, describes the scope of entities, concepts, and relations required to encode and store all types of assessments, including the assessment of compliance with the EOSC PID Policy. This model establishes the relations between concepts such as principles, criteria, metrics, benchmarks, tests, guidance, and best practices, and how these are used to create assessments of the subjects (products and services) being evaluated. The concepts are captured and described in a [glossary](#)⁶, which is foundational to the CAT, and a [working representation](#)⁷ is available.
2. FAIRCORE4EOSC and FAIR-IMPACT jointly developed additional submodels (PID-specific) to
 - a. Accurately describe the **roles of actors in the ecosystem**: PID schemes and standards bodies, PID naming and namespace authorities, providers, managers, owners, and end users. This is also captured as a [vocabulary](#)⁸ and attendant [working representation](#)⁹.
 - b. Describe the scope of [identifiers and their relationships to resolvers](#)¹⁰. A working representation of this is available.
 - c. Describe the scope of user expectations of the PID ecosystem and how concepts such as workflows, use cases, best practices, and PID services relate to one another.

Community feedback on the conceptual model is described in a separate section.

2.3 CAT Minimum Viable Product and Implementation Status

The beta version of the FAIRCORE4EOSC Compliance Assessment Toolkit (CAT) has been deployed at project M22 with the basic features that are needed.

⁶ <https://mscr-test.rahtiapp.fi/vocabularies/terminology/2a062144-1766-40e7-b8da-6a4b1d5f9f00>

⁷ <https://arrows.app/#/local/id=uVrxR2qchjpt8561cLuB>

⁸ <https://mscr-test.rahtiapp.fi/vocabularies/terminology/9c735525-960e-4f13-a74d-4eb23ea9c308>

⁹ https://arrows.app/#/local/id=E9w_I1RRHouOPHTBPzsK

¹⁰ <https://arrows.app/#/local/id=cBrJ6BR4d4QZOVzagGhv>

The table 1 shows, for each major functionality expected of the system, by which month in the project lifecycle it will be available, and at which level of maturity. Current status is indicated as follows: green: achieved; yellow - partly achieved; amber - not achieved. Deviations from the plan are discussed following the table. Technology Readiness Levels (TRL): 4-Concept | 6-Pilot | 7-Beta | 8-Production |9-In Operation

Table 1 – High-Level Roadmap of Compliance Assessment Toolkit MVP

Capabilities/ Functions	Description	MVP	M12	M18	M22	M26	M30	M34
Create and edit/ update vocabulary items	For a given vocabulary, add an element (label or term), or add/ update relations between vocabulary elements. or register a new relation type. Can be shared with MSCR/ EOSC vocabulary service		-	TRL3	TRL7	TRL8	TRL8	TRL8
List/ retrieve vocabulary elements and details	List the available vocabularies, elements of a vocabulary, and the details of a specific vocabulary element. Can be shared with MSCR/ EOSC vocabulary service		-	TRL3	TRL7	TRL8	TRL8	TRL8
Add a vocabulary	Add a vocabulary to the list of vocabularies. Can be shared with MSCR/ EOSC vocabulary service		-	TRL3	TRL7	TRL8	TRL8	TRL8
Add, edit, and delete an assessment	An API for managing an assessment record to the CAT database	V	TRL4	TRL6	TRL7	TRL8	TRL8	TRL8
Add, manage, and define user roles	Add and manage, disable users and their responsibilities in the CAT API services	V	TRL4	TRL6	TRL7	TRL8	TRL8	TRL8
Capabilities/ Functions	Description	MVP	M12	M18	M22	M26	M30	M34
Query the CAT database in respect of compliance status of actors	Actors include PID Scheme (Component), PID Authority (Role), PID Service Provider (Role), PID Service (Component), PID Manager		TRL4	TRL6	TRL7	TRL7	TRL8	TRL8

	(Role), PID Owner (Role), End User (Role), Compliance Monitoring (Role)							
Administrator View	Add, edit, delete vocabularies Add, edit, delete, and administer users and SSO	V	TRL4	TRL6	TRL7	TRL7	TRL8	TRL8
User Account View	Register a user, request user validation, associate with roles	V	TRL4	TRL6	TRL7	TRL7	TRL8	TRL8
Contributor View	Add, edit, and delete vocabulary elements	V	TRL4	TRL6	TRL7	TRL7	TRL8	TRL8
General End User View	List, query, and explore vocabularies. Explore compliance history and status of an actor, object, or service.		TRL4	TRL6	TRL7	TRL7	TRL8	TRL8
Evaluator/ Assessment System View	View past assessments performed. Add a new assessment (by reference, by value)	V	TRL4	TRL6	TRL7	TRL6	TRL8	TRL8
Dashboard/ Discovery	Summary of compliance assessments and status across actors, objects, and services for each type of assessment regime Link to listings and individual status/ history views		TRL4	TRL6	TRL7	TRL7	TRL8	TRL8

2.4 Discussion of Deviations

1. **Deviations in respect of vocabulary services:** initially, it was planned that the CAT would implement its own vocabulary server to support the application. It soon became clear that (1) other FAIRCORE4EOSC components, notably the MSCR and possibly the Type Registry, would also require managed vocabulary services, and (2) that it would be desirable to establish such a vocabulary service as a possible additional component that forms part of the EOSC Core. Discussions on the latter are ongoing. In the interim, FAIRCORE4EOSC has established a centralised vocabulary server for the project, hosted at CSC. More detail about this server is provided in the relevant section. The delay in implementation, due to the process of engagement within EOSC, is expected to be addressed by M26.
2. **General End User View:** This is impacted by the delay in availability of the vocabulary services since end users are not yet able to view vocabularies within CAT - although they can already do so via the vocabulary server. Target: M26.

3. **Evaluator/ Assessment System View:** It is not yet possible to add a new assessment by reference. The functionality that allows an external assessment to be imported exists, but UI functions to allow it must still be developed. Target - M26.
4. **Dashboard/ Discovery:** The dashboard and discovery service is ready for integration into CAT, but there are tasks that remain: specifically, the back-end data integration and automated indexing of the assessment corpus is not yet implemented - catalogue data for search and discovery is manually transferred in the current implementation.

2.5 CAT requirements overview

Table 2 provides an overview of the state of tickets generated from the requirements and specifications. The values in (a) and (b) show the number of tickets added or modified since the publication of the original specifications. As can be seen, very few tickets were deemed to be out of scope (d), and the major changes relate to new monitoring services for PIDs.

Table 2 – Requirements Overview

User Story - Landscape Analysis	New (a)	Restructuring (b)	Accepted (c)	Out-of scope (d)	Total (c+d)
Users of persistent identifiers want to assess the level of compliance of foundational actors in the PID ecosystem to determine the most appropriate service to use in the context of the EOSC PID Policy. (Foundational actors - schemes, authorities, service providers, and managers).	0	0	8	3	11
Providers of persistent identifier services need to be able to substantiate and publish the level of compliance with EOSC PID Policy. Such publications are citable and versioned.	0	0	14	0	14
User Story - External engagement	New (a)	Restructuring (b)	Accepted (c)	Out-of scope (d)	Total (c+d)
There is a general need expressed in international forums to disambiguate and align assessment of TRUST, and by extension those for FAIR. (The conceptual model and its vocabulary is designed to address this need). This project extends the	0	0	4	0	4
User Story - Landscape Analysis	New (a)	Restructuring (b)	Accepted (c)	Out-of scope (d)	Total (c+d)
approach proposed there to accommodate PID Policy compliance assessment.					
Institutional users of persistent identifiers want to obtain guidance on best practices and recommendations in terms of implementation. The guidance applies in two categories: <ul style="list-style-type: none"> • how to use or provide actionable persistent identifiers best, depending on the use case, 	1	0	8	0	9

<ul style="list-style-type: none"> and how to develop policy in respect of identifiers. 					
PID resolvability Monitoring The Persistence and Resolution Monitoring Service (PRM) is a service that monitors and logs the resolution of provided PIDs, determines the 'first-seen' and 'last-seen' date (using Memento if needed), and calculates the text vector of the landing page for both these dates. The results are stored in a log/database. To determine resolution and get related metadata, integration with the PIDMR is planned.	4	0	3	1	4
PID Properties Test API (PPT) The PID Properties Test API (PPT) is an application that should be able to supply PID persistence test information. It will use the monitor records in the PID Resolution Monitor (PRM) to generate responses on (1) resolution percentages, (2) median persistence age and (3) content drift.	3		3		3

All entries have been addressed in the Beta Release, including ones for which the current status indicates 'Updated' (an amendment to the requirements and specifications published in M12), except where status is indicated as 'In Progress'

Table 3 – Compliance Assessment Toolkit Requirements Summary - MVP

Requirement	Short description	Reference
WP2_T2_CAT_3.1 Manage User Roles	Placeholder user accounts for each defined role. Add, manage, and define user roles	CSCFC4EREQ-428
WP2_T2_CAT_3.2 Manage Identified User Roles	User registers as an 'Identified' user -role is assigned automatically, no further action required.	CSCFC4EREQ-429
WP2_T2_CAT_3.3 Manage Validated User Roles	User registers as a 'validated' user - admin needs to validate and assign a role. In order for a User to be assigned with this role, he should update his profile and make a validation request with the required fields	CSCFC4EREQ-430
WP2_T2_CAT_2.3 Add compliance assessments manually	Add compliance assessment manually for a PID owner	CSCFC4EREQ-426

WP2_T2_CAT_2.2 Add compliance assessments manually	Add compliance assessment manually for a PID Manager	CSCFC4EREQ-425
WP2_T2_CAT-5.2 Administrator Dashboard for users	Administrator opens a dashboard-like view of users, able to edit, add, and remove users.	CSCFC4EREQ-435
WP2_T2_CAT-5.3 Administrator Dashboard for pending requests	Administrator opens a dashboard-like view of users, and filters on pending validation requests. Administrator can confirm or reject validation,	CSCFC4EREQ-436
WP2_T2_CAT-6.1 Add, edit, and delete registry information for the user	Add, edit, and delete registry information for the user provides a dashboard-like view that corresponds to the first page of a manual assessment to confirm, edit, and delete information about a specific service or object. Display an assessment history for an object.	CSCFC4EREQ-437
WP2_T2_CAT-6.2 A validated user requests or registers a new manual external assessment.	A validated user requests or registers a new manual external assessment. This assessment has the following steps and sub-components: <ul style="list-style-type: none"> ⦿ Provide or confirm information about the submitter (=logged-in user account) ⦿ Select the type of assessment (Scheme, Authority, Provider, Manager, ...) ⦿ Define whether the assessment should be private, or whether it is licenced under CC 4.0 BY. ⦿ Provide information about the assessment target, and a pointer to a validated assessment result. The assessment result needs to conform to the assessment exchange specification. 	CSCFC4EREQ-438
WP2_T2_CAT-6.3 A validated user requests or registers a new manual assessment.	A validated user requests or registers a new manual assessment. This assessment has the following steps and sub-components:	CSCFC4EREQ-439
Requirement	Short description	Reference
	<ul style="list-style-type: none"> ⦿ Provide or confirm information about the submitter (=logged-in user account) ⦿ Select the type of assessment (Scheme, Authority, Provider, Manager, ...) ⦿ Define whether the assessment should be private, or whether it is licenced under CC 4.0 BY. ⦿ Provide information about the assessment target by filling in the evaluation form. 	

	Support is provided by making benchmarking, guidelines, and best practices information available.	
WP2_T2_CAT-7.4 User selects 'My Objects and Services'.	User selects 'My Objects and Services'. The dashboard lists all services or objects associated with the account with assessment status. Subcases include: <ol style="list-style-type: none"> 1. Add a new object (navigates to 6.1) 2. Modify an existing object (navigates to 6.1) 3. Delete as existing object (verify dependencies and warn) 	CSCFC4EREQ-443
WP2_T2_CAT-7.2 User selects "Account Information".	User selects "Account Information". Displays information captured at registration, and allows modification.	CSCFC4EREQ-441
WP2_T2_CAT-5.2 Administrator Dashboard for users (updated¹¹)	Administrator opens a dashboard-like view of users, able to edit, add, and remove users.	CSCFC4EREQ-435
WP2_T2_CAT-5.3 Administrator Dashboard for pending requests (updated)	Administrator opens a dashboard-like view of users, and filters on pending validation requests. Administrator can confirm or reject validation, triggering a helpdesk ticket process	CSCFC4EREQ-436
WP2_T2_CAT-8.3 Detailed View (equivalent to 9.3) (in progress)	Listing/ Cards - all assessments matching the filter specification. Clicking a list entry opens a detailed view.	CSCFC4EREQ-446
WP2_T2_CAT-6.2 A validated user requests or registers a new manual external assessment. Import assessment	A validated user requests or registers a new manual external assessment. This assessment has the following steps and sub-components: <ol style="list-style-type: none"> 1. Provide or confirm information about the submitter (=logged-in user account) 2. Select the type of assessment (Scheme, Authority, Provider, Manager, ...) 	CSCFC4EREQ-438

¹¹ Indicates a revision of the original requirements and specifications in Deliverable D2.1.

	<ol style="list-style-type: none"> 3. Define whether the assessment should be private, or whether it is licenced under CC 4.0 BY. 4. Provide information about the assessment target, and a pointer to a validated assessment result. The assessment result needs to conform to the assessment exchange specification. 	
<p>WP2_T2_CAT-9.3</p> <p>Detailed Object or Service View</p>	Information about an object. The list of assessments per object	CSCFC4EREQ-449
<p>WP2_T2_CAT_2.3</p> <p>Add compliance assessments manually - For others</p>	<p>Ingest a remotely performed assessment for any of the above</p> <ol style="list-style-type: none"> 1. Selected Schemes 2. Selected Authorities 3. Selected MPAs 4. Selected Service Providers 	CSCFC4EREQ-427

2.6 CAT Overview and the status of the development

2.6.1 CAT API

The Compliance Assessment Toolkit (CAT): CAT API^{12,13}, is built on the [Quarkus](#)¹⁴ framework, providing a secure and efficient solution for user validation and assessment management. Through designated API endpoints, users can validate themselves or request validation in specific roles¹⁵, enabling access to features such as assessment creation and management, as well as the customisation of profile information, including name, surname, email, and ORCID ID.

To ensure the security of all interactions, CAT API is protected by the AAI Proxy, safeguarding each endpoint. Clients / Users can obtain specific roles that determine their access permissions, defining which endpoints they can interact with. The entire authorization process is handled through Keycloak¹⁶, ensuring a reliable authentication mechanism for a smooth user experience.

¹² <https://api.cat.argo.grnet.gr/swagger-ui/#/> (API definition)

¹³ <https://api.cat.argo.grnet.gr/> (API service endpoint)

¹⁴ <https://quarkus.io/>

¹⁵ Some roles in the PID ecosystem (authorities, schemes, providers, and multi-provider agencies) require validation by an administrator to guard against false or unauthorised compliance claims.

¹⁶ <https://www.keycloak.org/>

CAT API also provides administrators with efficient user management capabilities, allowing them to manage Users and Roles and to control the validation request process. Administrators can navigate through users' validation requests, approving or rejecting them as needed whilst at the same time they have the authority to assign specific roles to different users, defining access permissions based on organisational needs.

The main functionalities the API currently supports (based on the permissions a user has)

- ⦿ User Management
 - Validation Requests: The management of validation requests to assume one or more roles in the system.
- ⦿ Assessments: The management and creation of a manual assessment
- ⦿ Integrations: The API supports the integration with external systems. The first integration refers to Research Organization Registry (ROR)¹⁷ and it is used when a user is creating a validation request for a specific actor
- ⦿ Subjects: Management of subjects. The subject may be an object, entity or service, a web resource identified by the owner offered by an Authority, Provider, or Manager, for which the assessment will be completed.
- ⦿ Other:
 - Code lists, filters, and endpoints that help the client use the API.
 - Notifications sent to users through emails, when the user needs to act or be aware of updates.

2.6.2 CAT User Interface (CAT UI)

The Compliance Assessment Toolkit (CAT): CAT UI¹⁸ is a web-application, created using the React JavaScript framework. It also uses JS Typescript as a technology for a “clean” code, and of-course for making the code more scalable. The main goal of the CAT UI is to assist actors in the PID ecosystem with assessment of their compliance with the EOSC PID policy. The CAT UI supports different levels of authorization for registered users, separating the UI capabilities for administrators and simple users. Key features like user authentication using EOSC-AAI demo instances (through the use of Keycloak), automatic user registration, and basic user profile configurations are supported.

The CAT UI follows a list of design patterns, tips and best practices for the flow of information in order to solve common problems in user interface design. This leads to a dashboard-like UI with interfaces that are both consistent and efficient. The UI supports and makes use of:

- ⦿ Consistent tables, with pagination, and sorting features
- ⦿ Filters in specific views that help the user find what they want quickly, saving time and reducing frustration
- ⦿ Consistent notification messages for the interaction with the API

¹⁷ <https://ror.org/>

¹⁸ <https://cat.argo.gmet.gr/>

- ⦿ Autocomplete on input forms and more refined UI layouts
- ⦿ Wizard-like forms
- ⦿ Tooltips, help messages, and information are key features of the UI
- ⦿ TypeScript declarations and updated assessment schemas are also used

The UI consists of the following main views:

- ⦿ Administrator Dashboard: This is the place where administrators can manage users, assessments and validation requests.
- ⦿ Assessments per actor: This is a public view where visitors can be informed about the latest assessments made by the different actor types
- ⦿ User Validations: A user can manage as many requests for becoming an actor as he wants.
- ⦿ User Profile: The place where the user can manage its validations, assessments and profile data.
- ⦿ Management of subjects: The validated user can easily manage its own subjects. He can create, update, delete, search the subjects he owns. The subjects are also integrated in the assessment process in the wizard in step 2.
- ⦿ Management of assessments: The validated user can easily manage its own assessments for the different actors. He can create, update, delete, search or even export the assessment he/she wants. For the creation of new assessments, the UI supports a wizard step by step component. The wizard guides the dashboard user through-out the necessary steps for creating a valid assessment request. An algorithm that follows the conceptual model and the combination of principles, criteria, and metrics is also part of the UI. In the latest version of the UI dynamic features that allow the display of principles, criteria, and metrics for the assessment creation view are key to the user support.

The deployed beta version of the CAT UI implements the basic features to validate the conceptual model, principles and metrics. Next to provide feedback on the basic features, the early end-users can also provide feedback on future developments to advance the UI assessment process and the user experience.

In order to support the processes a dedicated documentation site is provided¹⁹.

¹⁹ <https://fc4e-cat.github.io/fc4e-cat-doc/docs/intro>

C2 - Sensitive Metadata ✖

Sensitive PID Kernel Metadata requires users to be authenticated - evidence is provided.

C3 - Ownership ✔

The PID owner SHOULD maintain PID attributes.

C15 - Type Information *

The PID Authority SHOULD provide information on the referenced object's fundamental type and management policy in a machine-actionable way.

C12 - Persistence - Authority *

The PID owner SHOULD maintain PID attributes.

C34 - Persistence Mean *

PID Services SHOULD aim for a persistence median that is acceptable to and aligns with community and dependency expectations.

C35 - Resolution Percentage *

The PID owner SHOULD maintain PID attributes.

C28 - Certification *

PID Authorities and Services MUST agree to be certified with a mutually agreed frequency in respect of policy compliance.

C2: Sensitive Metadata Optional

Sensitive PID Kernel Metadata requires users to be authenticated - evidence is provided.

part of principle P2: Secure ⓘ

● ● ● ●

complete - result: ✖

FAIL

T2.1 Secure - Encrypted ⏪ Ⓞ

Are the API services offered, encrypted using https?

Yes No

T2.2 Sensitive - Indication ⏪ Ⓞ

“Is sensitive PID Kernel Metadata defined and supported?”

Yes No

Evidence:

<https://www.grnet.gr> remove

URL: Add

T2.3 Secure - Encrypted ⏪ Ⓞ

“Is sensitive PID Kernel Metadata encryption supported?”

Yes No

Evidence:

URL: Add

T2.4 Secure - Access ⏪ Ⓞ

Is Access to Sensitive PID Kernel Metadata password protected?

Yes No

Evidence:

URL: Add

Figure 1 – Example Assessment Screen

Applicable criteria are clearly shown, with current compliance status. Users can track test result status in the top right, and guidance is available for each test.

Funded by the European Union

FAIRCORE4EOSC has received funding from the EU's Horizon Europe research and innovation programme under Grant Agreement no. 101057264.

14

3 Conceptual Model Feedback and Community Engagement

The Conceptual Model published in D2.1 [1] was reviewed with the community on a number of occasions (Annexure B), and is approaching maturity with few changes expected in the future. The most important feedback and validation was obtained from two sources:

- in-depth review of CAT at the EOSC Winter School,
- feedback received from the EOSC PID Policy Task Force.

The only amendment identified in these engagements required criteria to be aligned with the nature of the PID Stack²⁰: Intrinsic identifier metadata is less editable (or not editable at all), and this will remove the need for some criteria to be mandatory. These changes will be incorporated into the revised requirements and specifications.

The feedback also emphasised the importance of guidance and best practice, and while this is accommodated in the model, additional UI views will be developed to assist with guidance and PID Stack selection.

²⁰ A PID Stack defines the actors in the ecosystem to deliver a specific PID service. For example, DataCite DOIs are based on the Handle Schema and system, and additionally require functioning infrastructure or capacity from DONA Foundation, the International DOI Foundation, and DataCite to work.

4 Vocabularies and Registries

The CAT requires the implementation of a number of vocabularies and registries²¹, and these are listed below together with links to the current working definitions of the resources. The legend defines the source of the vocabulary or registry content as anticipated in D2.1. In the interim, we have identified a need for additional vocabularies and registries, and these are listed as additions to the table.

Table 4 – Compliance Assessment Toolkit Vocabularies and Registries

Vocabulary/ Registry	Implementation	Typical Values
Concepts and Relations	Ontology	The ontology of the compliance graph
Motivation	Registry	{TRUST, FAIR, CARE, PID Policy, ...}
Motivation Typology	Vocabulary	{Community expectation, mandatory, rules of engagement, digital systems architecture, ...}
Criterion Typology (Imperative)	Vocabulary	{Mandatory, Optional, Conditional}
Roles	Vocabulary	{PID Scheme (Component), PID Authority (Role), PID Service Provider (Role), PID Service (Component), ...}
Cases	Registry	A registry of use cases, typified using the PID Use Case Typologies
PID Use Case Ontology (Work in Progress)	Ontology with registries and vocabulary	A classification of demand for PID services - use cases in terms of features and performance requirements (governance, cost, scale, ...)
PID Stack Ontology (Work in Progress)	Ontology with registries and vocabulary	A classification of the features, attributes, and characteristics of PID Services.
Maturity/ Level of Compliance	Vocabulary	{Ad-hoc, aware, implementation phase, managed and measured, optimised}
Principles	Registry	The principles included in a 'motivation'
Criteria	Registry	FAIR criteria, CTS criteria, PID policy provisions, ...
Metrics Typology	Vocabulary	Types of metrics (qualitative, quantitative, and elaboration of these, how repeatable they are)
Metrics	Registry	An inventory of registered metrics
Metric Predictability	Vocabulary	Indication of the extent to which independent assessors will reach similar results given the same input.
Test Typology	Vocabulary	Types of tests (automated, manual, machine actionable, ...), and possibly whether tests are standardised or not
Tests (Methods)	Registry	Registry/ inventory of test methods
Standards (Benchmarks)	Registry	Registry of standards and benchmarks - links to Information on how to interpret a test result
Recommendations (Work in Progress)	Registry	A registry of recommendations as annotations from published work, linked to criteria
Guidance (Work in Progress)	Registry	A registry of guidance as annotations from published work, linked to criteria
Best Practices (Work in Progress)	Registry	A registry of best practices as annotations from published work, linked to criteria

²¹ Vocabularies: concepts or abstractions, generally used by systems to refer to these concepts unambiguously. Registries contain instances of the same concept (for example a list of 'Authorities' or 'PID Stacks'.

Entities (Synthetic Data)	Registry	A registry of entities (institutions) and object types in the ecosystem. These are individual organisations offering PID services based on a PID Stack.
PID Stacks (Work in Progress)	Registry	A list of all the end-user services (PID Stacks), as well as the entities involved in providing specific roles in the PID stack.
Subject of Assessment (Synthetic Data)	Registry	This will often be a combination of a PID Stack and an actor in the ecosystem (role). As a result, one needs to register the actors in the ecosystem and define their role in a specific PID Stack, and point to the service or object being assessed.

Legend: The source of authoritative content for vocabulary and registry entries

EOSC PID TF	FAIRCORE4EOSC	FAIR-IMPACT	ADDITIONS after D2.1
-------------	---------------	-------------	----------------------

These vocabularies and registries will be deployed to the FAIRCORE4EOSC Vocabulary Server, hosted at CSC for the interim.

5 Acceptance Testing

The [beta version of CAT](#)²² can be found in GitHub and a [version of record](#)²³ has been deposited in Software Heritage.

To ensure that all functionality and user interface components will remain functional with each (snapshot) release of the CAT, the CAT UI is tested with the front-end testing tool (cypress.io). Cypress.io is an end-to-end (E2E) testing framework. The testing methodology verifies the working order of a software product in a start-to-finish process.

Cypress scripts test the various pages of the CAT UI using actual user inputs. These tests cover different aspects of the user experience. Currently all forms in the application are filled out according to their “happy flow”. A happy flow describes the default scenario for the user without exceptions or error conditions. Besides this, the required field restrictions and other error handling on these forms are also tested. All scripts test the pages in isolation (it describes a single page and tries to stay in the context of that page).

Before creating the CAT UI tests per single page, an inventory of the required fields, including their default values was made. We have identified six different pages in the CAT UI. These are the following: Admin page, Assessment page, Sign-in page, Profile page, Subject page and Validation page.

Test cases were described for each page. The number of test cases varied per page, but there were always at least two. Each test case consists of a number of expected and executable steps.

These steps are captured in the form of Cypress tests in the JSON format, which can be automatically tested by the test framework. The tests are also described in the comments of the pull requests (PR) made on GitHub (see ‘PR Number’ in Table 5).

Table 5 provides an overview of the created and merged pull requests per page up to now. It also lists the test cases names per page to get a brief impression of what is being tested. A complete overview of the CAT UI test cases including their testing steps are available from the FC4E project wiki²⁴.

Table 5 – CAT UI Test Cases Overview

Test file/page	PR number	PR status	Test cases
Admin	148	<input checked="" type="checkbox"/> Merged	<ul style="list-style-type: none"> - Accept a validation request - Declines a validation request

²² https://archive.softwareheritage.org/browse/origin/?origin_url=https://github.com/FC4E-CAT/fc4e-CAT -UI

²³

https://archive.softwareheritage.org/browse/directory/dc20d687e8a93a835f61b25cfda13558d43fad6a/?origin_url=https://github.com/FC4E-CAT/fc4e-CAT-UI&revision=dad92fec0be3b0e9ca133c800c84673e419c3950&snapshot=eb6058676239c17cbef68191afe932361a941fdf

²⁴ <https://drive.google.com/file/d/1Zdq1MHfCCKji2RnkgnnsXzJ8WPQ92Yuu/view?usp=sharing>

Test file/page	PR number	PR status	Test cases
Assessment	135 & 152	<input checked="" type="checkbox"/> Merged	<ul style="list-style-type: none"> - Greets with create assessment - Brings me back to assessments - Has the correct actor option - Can go to step 2 and go back - Makes the submitter fields read only - Create an assessment
Sign in	145	<input checked="" type="checkbox"/> Merged	<ul style="list-style-type: none"> - Greets with sign in - Requires username - Requires password - Logs in
Profile	146	<input checked="" type="checkbox"/> Merged	<ul style="list-style-type: none"> - Checks an identified user profile information - Does not allow validations without updating details - Does not show assessments/subjects without validation - Links to validations - Links to assessments - Links to subjects - Requires the required fields to be filled out - Requires the name to be longer than 3 characters - Does not require an ORCID id - Updates the personal details
Subject	147	<input checked="" type="checkbox"/> Merged	<ul style="list-style-type: none"> - Deletes a subject - Edits a subject - Greets with create new subject - Closes the modal - Creates a subject
Validation	148	<input checked="" type="checkbox"/> Merged	<ul style="list-style-type: none"> - Creates a validation - Creates a validation with a custom organisation

6 Documentation

The first release of the [Documentation Site](#)²⁵ is available, editable and generated from the Github repository. It contains guides, tutorials, and other resources designed to help users understand and effectively use the toolkit. Documenting the component in an open Source repository enables the community to participate in this endeavour and provides a mechanism towards sustainability. At the end of the project with the PROD release, a versioned copy will be archived in Zenodo. The code repository for the component, the documentation and the API endpoints are regularly archived in Software Heritage²⁶.

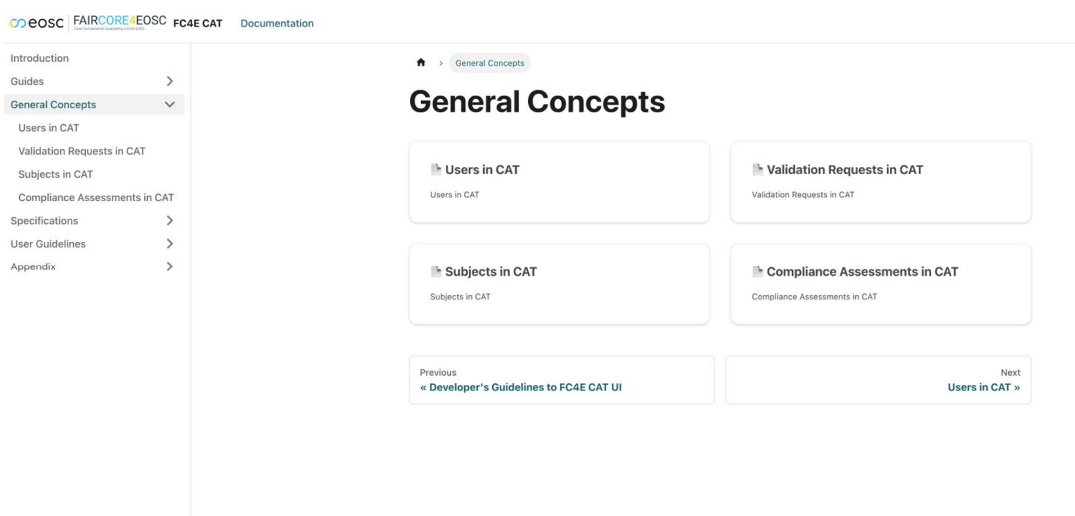


Figure 2 – Documentation Available to Users

To assist end users with the different types of metrics and tests in the creation of an assessment, a number of demonstration sites have been created. Examples are available at the following URLs.

<https://tests.cat.argo.grnet.gr/galid/> and <https://tests.cat.argo.grnet.gr/badid/>

²⁵ <https://fc4e-cat.github.io/fc4e-cat-doc/docs/intro>

²⁶

<https://archive.softwareheritage.org/swh:1:dir:08f1da7dbbbca398a7807d51e51b45667e15f5df;origin=https://github.com/FC4E-CAT/fc4e-cat-doc;visit=swh:1:snp:718fe4e7f884999f228170ae87ff07b38fc4e932;anchor=swh:1:rev:5f33d090fb88ff93d37c06779e999e87ac2c6fde>

7 Challenges

7.1 Assessment data

A significant challenge is experienced in respect of the data that will be included in beta versions (and possibly initial production versions) of CAT. The main reason for this extends to the mandate and authority of the CAT to publish EOSC PID Policy compliance assessments on behalf of EOSC, which was the intended positioning of the CAT. In short:

1. EOSC does not yet provide an established governance structure or mechanism to provide legitimacy to compliance assessments issued on their behalf. Note that this does not only affect the CAT: there are several other sets of compliance expectations (Rules of Participation, interoperability framework, ...) that potentially also need to be evaluated formally.
2. It would be unwise to publish any assessments of actual PID Stack performance against the policy provisions without EOSC endorsement of some kind, even if the participants agree to such an assessment. Such endorsement could possibly form part of the sustainability aspects of the CAT on completion of the project, but at present there is no clear indication of how this might work in practice.

The specific objective is to ***avoid or contextualise claims of EOSC PID Policy Compliance without endorsement by EOSC.***

To address the situation in the interim, WP2 has taken three mitigation steps:

1. It has started construction of a [demonstration database](#) that covers all the actors in the ecosystem with synthetic assessment data, as well as synthetic actor definitions. This database will be used in the beta release and will serve as a demonstrator database until further notice (see 'Sample Data' - Annexure A).
2. A programme of engagement with major actors (providers, authorities, multi-provider agencies) will be started to prepare them for eventual release of operational data, and to complete their assessments over the remainder of the project. It is expected that all of these assessments will eventually be public, since public compliance information about the major PID Stacks is in the interest of both end users and the providers of the services.
3. Other actors (managers, owners) will be assisted via collaboration with FAIR-IMPACT to complete assessments, and since these assessments can be made public at the discretion of the submitter, these could be made public without EOSC endorsement, provided they are clearly indicated as such.

8 Expected Improvements and Refinements

8.1 Additional Guidance and Best Practices Support

Engagements with the end-user community via FAIR-IMPACT and during the EOSC Winter School has stressed the importance of guidance and best practices. The CAT originally intended such guidance and best practice to be linked to tests during the assessment process, but this has two deficiencies: (1) end users can only access the guidance when they are executing an assessment, and (2) selection of PID stacks and services that meet expectations (including alignment or compliance with the policy, but not limited to it) cannot easily be determined.

The FAIR-IMPACT project is refining and structuring PID best practices and guidance with EOSC PID Policy compliance in mind, and the resulting Knowledge Base can be incorporated into CAT once it becomes available (expected M24). The mechanism for doing so is relatively simple, and does not require extensive additional development work:

1. The guidance, best practices, and its mapping to PID stacks can be encoded as a catalogue of information, and indexed using Elasticsearch using performance elements and policy compliance as one or more facets.
2. The ‘Dashboard and Search’ component that forms part of the CAT can be implemented for any catalogue and index, and configured to reflect a different set of facets. By duplicating the component in CAT, clearly defining the distinction between a catalogue of assessments against the policy, and a catalogue of best practices matched with PID Stacks, the wider interest of the community will be served.

It is in the interest of the CAT to implement such a best practices and guidance catalogue - it will broaden the user base for the component significantly, and drive users towards the web presence of the CAT.

8.2 User Interface Improvements

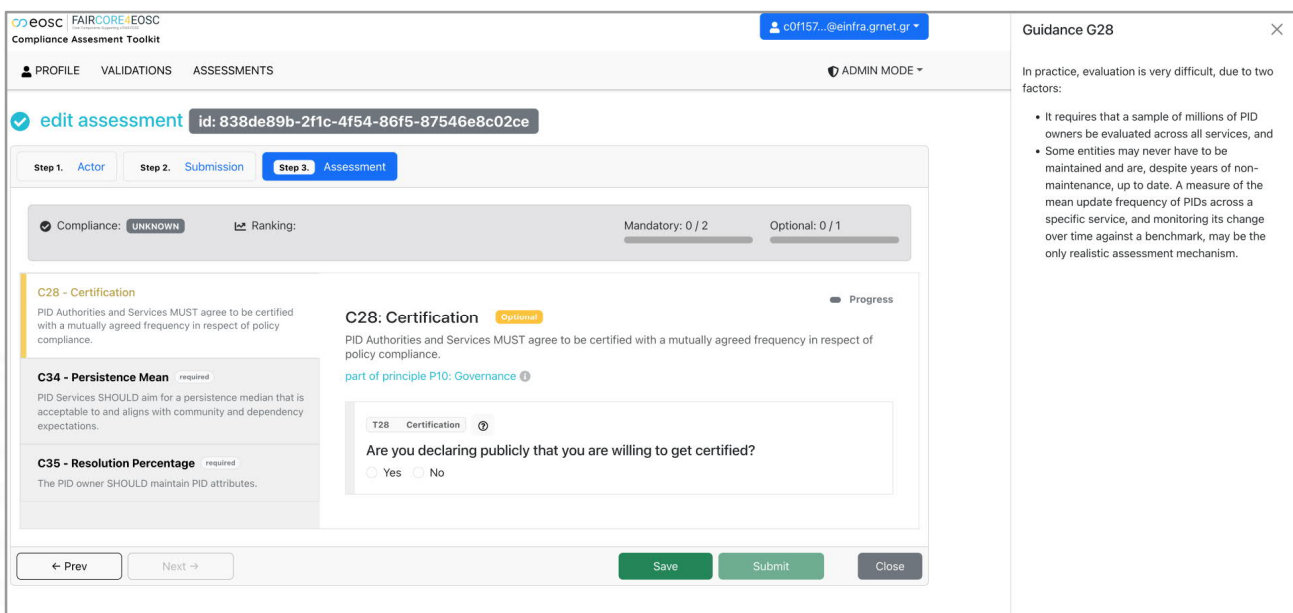


Figure 3 – Example of UI Improvements

Guidance is made more prominently available. Example of improvements being incorporated into the next release.

The current beta release of the CAT meets its quality expectations in the sense that the use cases, executed from an end user perspective using automated tests, all pass. WP2 does, however, envisage two sets of refinements to the user interface, as follows:

1. It has been agreed to employ a user experience expert to provide feedback on structuring, layout, and process flow in the application. This feedback will be complemented with user experience feedback gathered by FAIR-IMPACT at the IDCC in Edinburgh in February 2024: these refinements are expected to be included by M26;
2. The scope of information provided to end users will be extended in the UI, including the following aspects:
 - a. Providing links to documentation, example policies for managers, and other resources developed by FAIR-IMPACT.
 - b. Improving scope and visibility of guidance (Figure 3).
 - c. Deployment of an additional catalogue of knowledge-base information, also gathered by FAIR-IMPACT, that describes the available PID services and stacks in terms of functions, attributes, and characteristics of the services, which will in time include EOSC PID Policy compliance.

User interface improvements are documented in detail elsewhere (see Annexure A).

8.3 Vocabulary Server²⁷ and PID MetaResolver Integration


The vocabulary server will be integrated with the CAT by M26, based on the deployment at CSC. This is an interim arrangement until such time that a more permanent hosting location can be agreed within EOSC, since such a vocabulary server is a critical component of EOSC-wide services. It is likely to become even more important with the creation of EOSC nodes in future.

8.4 Automated Monitoring

The need to objectively verify resolvability and persistence was identified in D2.1, and while not originally in scope for the CAT, this requirement is foundational to assessment of persistent identifier services and it was agreed to develop this as a service supporting CAT.

Persistence and resolvability, if maintained indefinitely, is a remedy for ‘Reference Rot’. Reference rot, link rot, and content drift are all related phenomena, or problems, that occur in the context of web resources that are dereferenced by a PID to a URL via a link resolver²⁸, and is increasingly seen as a major issue²⁹.

These phenomena are described as follows:

-  **Reference Rot:** Reference rot refers to the phenomenon where hyperlinks or URIs (Uniform Resource Identifiers) included in web resources become obsolete or broken over time. This can happen due to various reasons such as websites being moved or deleted, changes in URL structure, or the resources

²⁷ <https://mscr-test.rahtiapp.fi/vocabularies>

²⁸ Klein M, Van de Sompel H, Sanderson R, Shankar H, Balakireva L, et al. (2014) Scholarly Context Not Found: One in Five Articles Suffers from Reference Rot. PLoS ONE 9(12): e115253. <https://doi.org/10.1371/journal.pone.0115253>

²⁹ Eve, M. (2024) 'Digital Scholarly Journals Are Poorly Preserved: A Study of 7 Million Articles', Journal of Librarianship and Scholarly Communication. 12(1) doi: 10.31274/jlsc.16288

themselves being modified or removed. This is the combination of two problems involved in using URI references; link rot and content drift.

- **Link Rot:** Link rot is a subset of reference rot, specifically referring to the situation where the resource identified by a URI may cease to exist and hence a URI reference to that resource will no longer provide access to referenced content.
- **Content Drift:** Content drift refers to the gradual modification or alteration of the content of a web resource over time. This can happen due to updates, revisions, or edits made by the authors or administrators of the resource. At the end of the URI, content drift can lead to discrepancies to such an extent that it ceases to be representative of the content that was originally referenced.
- 🕒 In summary, reference rot encompasses both link rot and content drift, as it deals with the degradation of the integrity and reliability of web resources over time, whether through broken hyperlinks or changes in the content itself.

To be able to assess these known issues involved in PID resolution, we will create services that run automated tests for two interrelated results: (1) Median persistence for a specific collection of PIDs (time it takes for 50% of the corpus to become unresolvable), and (2) resolution percentage (percentage of PIDs in a corpus that resolves to some content without an error). A third test result, the degree of content drift in a corpus is not under consideration yet.

The PID compliance monitoring has been split into two services, which are described below. Their technical specification documents have been created using the FAIRCORE4EOSC Technical Specifications template³⁰ and are available from the project wiki.

- 🕒 The first specification describes the PID Resolution Monitor (PRM)³¹. This is a service that monitors and logs the resolution outcomes of a sample of PIDs within a given PID Stack.
- 🕒 The second specification is for creating a PID Properties Test API (PPT)³², which retrieves results from the PRM to calculate a persistence median and a resolution percentage for a corpus of PIDs.

The expected improvements and refinements have been documented in a next iteration of the CAT Requirements and Specifications documentation³³.

³⁰ https://drive.google.com/file/d/1c5l2o6_2OBOFHh7PVnIX2Jg5ZOGuNrur/view?usp=drive_link

³¹ https://drive.google.com/file/d/1c8CIKGO8cZLNNsL7GpZUjYgS_dxsTNbX/view?usp=drive_link

³² https://drive.google.com/file/d/1c8JRm-b3hrqL6O0S-eOHI5TuQ5lm4vvP/view?usp=drive_link

³³ https://drive.google.com/file/d/1cACIJULQ2yMSTeBSYFLDL6NBKjxvVfl7/view?usp=drive_link

References

No	Description/Link
R1	https://drive.google.com/file/d/1bPkNqZVArVDT_RKAgZ-1HX_KxWNxX8qo/view?usp=sharing
R2	https://pidmr.devel.argo.grnet.gr/
R3	European Commission (2020). A Persistent Identifier (PID) policy for the European Open Science Cloud, doi: 10.2777/926037
R4	Hugo, W., Steinhoff, W., Turner, D., Buys, M., & Zamani, T. (2023). D2.1 Compliance Assessment Specification. Zenodo. https://doi.org/10.5281/zenodo.10067253
R5	https://eosc.eu/eosc-focus-project/winter-school-2024/ - opportunity area 1 - PIDs
R6	https://mscr-test.rahtiapp.fi/vocabularies/terminology/2a062144-1766-40e7-b8da-6a4b1d5f9f00
R7	https://arrows.app/#/local/id=uVrxR2qchjpt8561cLuB
R8	https://mscr-test.rahtiapp.fi/vocabularies/terminology/9c735525-960e-4f13-a74d-4eb23ea9c308
R9	https://arrows.app/#/local/id=E9w_11RRHouOPHTBPzsk
R10	https://arrows.app/#/local/id=cBrJ6BR4d4QZOVzagGhv
R11	Indicates a revision of the original requirements and specifications in Deliverable D2.1.
R12	https://api.cat.argo.grnet.gr/swagger-ui/#/ (API definition)
R13	https://api.cat.argo.grnet.gr/ (API service endpoint)
R14	https://quarkus.io/
R15	Some roles in the PID ecosystem (authorities, schemes, providers, and multi-provider agencies) require validation by an administrator to guard against false or unauthorised compliance claims.
R16	https://www.keycloak.org/
R17	https://ror.org/
R18	https://cat.argo.grnet.gr/
R19	https://fc4e-cat.github.io/fc4e-cat-doc/docs/intro
R20	A PID Stack defines the actors in the ecosystem to deliver a specific PID service. For example, DataCite DOIs are based on the Handle Schema and system, and additionally require functioning infrastructure or capacity from DONA Foundation, the International DOI Foundation, and DataCite to work.
R21	Vocabularies: concepts or abstractions, generally used by systems to refer to these concepts unambiguously. Registries contain instances of the same concept (for example a list of 'Authorities' or 'PID Stacks'.
R22	https://archive.softwareheritage.org/browse/origin/directory/?origin_url=https://github.com/FC4E-CAT/fc4e-cat-ui
R23	https://archive.softwareheritage.org/browse/directory/dc20d687e8a93a835f61b25cfda13558d43fad6a/?origin_url=https://github.com/FC4E-CAT/fc4e-CAT-UI&revision=dad92fec0be3b0e9ca133c800c84673e419c3950&snapshot=eb6058676239c17cbef68191afe932361a941fdf
R24	https://drive.google.com/file/d/1Zdg1MHfCCKji2RnkgnnsXzJ8WPO92Yuu/view?usp=sharing
R25	https://fc4e-cat.github.io/fc4e-cat-doc/docs/intro
R26	https://archive.softwareheritage.org/swh:1:dir:08f1da7dbbbca398a7807d51e51b45667e15f5df;origin=https://github.com/FC4E-CAT/fc4e-cat-doc;visit=swh:1:snp:718fe4e7f884999f228170ae87ff07b38fc4e932;anchor=swh:1:rev:5f33d090fb88ff93d37c06779e999e87ac2c6fde
R27	https://mscr-test.rahtiapp.fi/vocabularies

R28	Klein M, Van de Sompel H, Sanderson R, Shankar H, Balakireva L, et al. (2014) Scholarly Context Not Found: One in Five Articles Suffers from Reference Rot. PLoS ONE 9(12): e115253. https://doi.org/10.1371/journal.pone.0115253
R29	Eve, M. (2024) 'Digital Scholarly Journals Are Poorly Preserved: A Study of 7 Million Articles', Journal of Librarianship and Scholarly Communication. 12(1) doi: 10.31274/jlsc.16288
R30	https://drive.google.com/file/d/1c5I2o6_2OBOFHh7PVnIX2Jg5ZOGuNrur/view?usp=sharing
R31	https://drive.google.com/file/d/1c8CIKGO8cZLNNsL7GpZUjYgS_dxsTNbX/view?usp=sharing
R32	https://drive.google.com/file/d/1c8JRM-b3hrqL6O0S-eOHl5TuQ5Im4vvP/view?usp=drive_link
R33	https://drive.google.com/file/d/1bPkNqZVArVDT_RKAgZ-1HX_KxWNxX8qo/view?usp=drive_link

Annexure A: Supplementary Materials

Resource	Description
Specifications: PRM	A specification for the monitoring service envisaged in support of CST - resolution and persistence monitoring service (to be integrated with PIDMR if possible) https://drive.google.com/file/d/1c8CIKGO8cZLNNsL7GpZUjYgS_dxstNbX/view?usp=drive_link
Specifications: PPT	A specification for computing statistically significant tests for persistence, resolvability, and content drift based on the data collected by the PRM https://drive.google.com/file/d/1c8JRm-b3hrqL6O0S-eOHl5TuQ5lm4vVP/view?usp=drive_link
CAT Test Cases	A set of test cases (automated) for technical validation of the CAT beta release https://drive.google.com/file/d/1bsObiC1yvxHNY2nLOOYLekpzxASETfyQ/view?usp=drive_link
PIDMR Integration	Specifications and a discussion document for integration between PIDMR and CAT https://drive.google.com/file/d/1cEI8rmNMX_PVFIL7TbMORSGrI37NfxX2/view?usp=drive_link
Example Files	Examples of integration data structures/ request responses for PIDMR and CAT integration https://drive.google.com/file/d/1cHwjQ4idzG-P00KFy-wlUH2MbhOeNzZS/view?usp=drive_link
Data Files PID Ecosystem and Service Features	Attribute and relationship data data for the PID Ecosystem, being created in collaboration with FAIR-IMPACT. Work in Progress. Ecosystem: https://docs.google.com/spreadsheets/d/1DHd2COyHDID8iPjka2sms-rkegA3dE87tm79XFxJPZo/edit?usp=drive_link Attributes and Features: https://docs.google.com/spreadsheets/d/1x4YURMaDvxlUw1nu8B2V0glyRBoa_lhHAFKF1v61-WQ/edit?usp=sharing
Data Model Overview	Intended as an exchange of information with the OSTRails project. https://docs.google.com/presentation/d/1RiDsImztkip30DOURXfuXbBIZwprNjFDnT-gYvD2li0/edit?usp=drive_link

<p>EOSC PID Policy Case Study</p>	<p>One of several case studies used to define the conceptual model for CAT - and shared with the EOSC PID Policy Task Force as feedback and suggestions for improvement.</p> <p>https://docs.google.com/document/d/1MabseuZKjew-ExAlpoQhHFBqcyLxjgZzCytGao5BVHc/edit?usp=drive_link</p>
<p>Milestone Report: Beta Release</p>	<p>A milestone report describing the CAT beta release</p> <p>https://tt.eduuni.fi/sites/csc-rdi-fileshare/FAIRCORE4EOSC/WP1%20Project%20coordination%20and%20strategic%20alignment/Submitted%20FAIRCORE4EOSC%20Milestones/FAIRCORE4EOSC_MS_08.pdf</p>
<p>Sample Data</p>	<p>Sample data developed for demonstrators and development, to protect identity of ecosystem actors until agreements/ governance is in place.</p> <p>https://docs.google.com/spreadsheets/d/154dQJsRqiT3jWuSCwCX15ruJknnxLkOCVNRXo4vodEc/edit?usp=sharing</p>
<p>User Interface Improvements</p>	<p>A series of improvements agreed for the CAT UI</p> <p>https://docs.google.com/presentation/d/1pNr2PjpVE-xJzxTUPBnd1yYPPj7vwaSyN80gyIbVxsl/edit?usp=sharing</p>

Annexure B: Presentations and Community Engagements

Host	Event Title	Date	Location	Collaborators	Outputs
EOSC Symposium 2023	Persistent Identifiers (PIDs) Breakout	September 21st, 2023	Madrid, Spain	EOSC-A PID Task Force, RAISE project	https://symposium2023.eoscfuture.eu/symposium/pids/
FAIR-IMPACT All Hands Meeting in the Hague		October 4th-6th, 2023	The Hague, Netherlands	FAIR-IMPACT	F43.2.4.1 CAT Overview.pptx
N/A	Meeting on PID Policy Alignment and Support	November 7th, 2023	virtual	EOSC-A PID Task Force, FAIR-IMPACT	20231107 Meeting on PID Policy Alignment and Support
N/A	EOSC Compliant PID Implementations - Practical Guidelines for Implementing Best Practices	November 21st, 2023	virtual	FAIR-IMPACT	https://fair-impact.eu/events/fair-impact-events/eosc-compliant-pid-implementations-practical-guidelines-implementing-best
Synchronisation Force Workshop	Synchronisation Force 2nd Workshop - Persistent identifiers	November 30th, 2023	virtual	FAIR-IMPACT	https://fair-impact.eu/events/synchronisation-force-events/synchronisation-force-2nd-workshop-november-2023
EOSC Winter School 2024	N/A	January 29th - February 1st, 2024	Thessaloniki, Greece	FAIR-IMPACT, RAISE project, EOSC-A PID Task Force	EOSC Winter School 2024 Session Report_OA1_PIDs.docx
International Digital Curation Conference	Defining the criteria for assessing PID Policies and Services	February 19th, 2024	Edinburgh, Scotland	FAIR-IMPACT	https://zenodo.org/records/10791006