

FAIRCORE4EOSC

Developing EOSC-Core components to enable a FAIR EOSC ecosystem

14 | 05 | 2024 by Wim Hugo, DANS, Themis Zamani, GRNET





Funded by the European Union Project 101057264 — FAIRCORE4EOSC



Details on PID Policy

Expectations

Current Collaboration Blueprint

Three threads ...

meosc

The **PID Policy and Implementation Task Force** (PID TF) is dedicated to identify the gaps in the PID ecosystem that has been mentioned in the current SRIA version, especially to highlight mature and recognised PID infrastructures for emerging resource types, to standardise the PID graph, to integrate PIDs into FAIR data management, and to address PIDs and sensitive data (among others). It will provide different kinds of recommendations on PIDs management and **will set up criteria and certification of PIDs**



The **Compliance Assessment Toolkit** will support the EOSC PID policy with services to encode, record, and query compliance with the policy. It delivers a **conceptual model** for assessment together with **vocabularies**, **API services**, and **user interfaces** whereby actors can assess compliance with PID policy applicable to their context.



- Setting up a **coordination mechanism** for EOSC PID service providers
 - Ensuring end user needs are met when the EOSC PID landscape and solutions are developed
- Ensuring **alignment with EOSC policies** and architecture of the PID infrastructures
- Delivering **support instruments for facilitating uptake** (best practices, assessment toolkit, user documentation).





Compliance Assessment Toolkit

EOSC PID Policy compliance for all actors in the ecosystem.

Conceptual Model EOSC PID Policy Mapping to the Model Knowledge Base Development

Entities (Actors and Roles)



The actors in the PID ecosystem are related in specific ways, and this is already partly captured by the EOSC PID Policy. FC4E has made some specific recommendations for amendments and extension of the actor definitions, and this is being refined in F-I WP3.

Introduce the concept of a 'PID Stack':

Scheme	Standards Body	Authority	МРА	Provider	Manager	
Handle	IETF	DONA DOI Foundation DataC		DataCite	Zenodo	
Handle	Handle IETF DONA DOI Fo			DataCite	DANS	
Handle	IETF	DONA	DOI Foundation	IGSN	e.g GeoScience Aus	
Handle	dle IETF CNRI DOI Foundation		CrossRef			
URN	IETF	URN:NBN	N/A	URN:NBN-NL	КВ	

Principles, Criteria, Metrics, and Benchmarks

Criteria

#	Principle or Objective	Suggested Criterion	Description	Metric	Benchmark
C1	Preferred Unambiguous Interoperability	Minimum Operations	Service providers SHOULD provide a common Application Programming Interface to interact with PIDs, supporting a minimum set of operations (create, resolve and modify PID and PID Kernel Information)	ΣT _{1, n}	=0 → 0 >1 → 1
C2	Secure Viable, Trusted	Sensitive Metadata	Sensitive kernel metadata MAY require access control and/or encryption of the Kernel Information.	ΣT _{2, n}	<5 → 0 =5 → 1

Criteria and Actors

Applicability of Criteria to Actors

#	Criterion	Imperative	Scheme	Authority	Service	Manager	Owner
C1	Minimum Operations	SHOULD			1		
C2	Sensitive Metadata	MAY		1	1		
СЗ	Ownership	MUST		1	1		
C4	Maintenance	SHOULD					~
C5	Update Functionality	MUST			1	1	
			8			1	

Tests

Tests

#	Test	Description	Туре	Method	Guidance
T _{1,1}	CREATE	Create a PID and provide kernel information: API exists and evidence (URL) is available	Binary	Yes = 1 No = 0	G1
T _{1,2}	UPDATE	Update kernel information for existing PID: API exists and evidence (URL) is available	Binary	Yes = 1 No = 0	G1
Т _{1,3}	Resolution Service	Resolution API (URL) or URI Pattern exists, evidence is provided	Binary	Yes = 1 No = 0	G1
T _{2,1}	Secure - Encrypted	API services are encrypted (HTTPS)	Binary	Yes = 1 No = 0	G2

Guidance

Guidance

		1				
#	Guidance					
G1	One may extend the tests to recognise typical and popular standards for API implementation, such as REST, SmartAPI, and the like.					
G2	A series of 5 tests are proposed, all of which need to be satisfied to enable encryption and access control for sensitive metadata. Some of these tests may not apply to all use cases - a topic for future refinement.					
G3	In practice, this may require multiple tests, since in some cases, the ownership is encoded in metadata at the Manager (entity metadata), and in some cases with the Provider or the Authority. In addition, it will require definition of the path or retrieval mechanism for the information, which may be different for each					
	Profile. Suggestion: store the retrieval instruction 3.3, available end May 2024	T 1				
C 1	In practice, evoluction is yerry difficult, due to two factors:	Г				

Compliance Assessment Toolkit

EOSC PID Policy compliance for all actors in the ecosystem.

Workflows, Use Cases, Best Practices, Entities, ...

(René van Horik, Wim Hugo, DANS)

Use Cases, PID Stacks, and Desirable Features

Analysis of the landscape to generate a structured body of knowledge

FAIRCORE4EOSC

meosc

Ca	se study	Ŧ	Function/ Entities		-	PID	=	= l	Use Case Elements	Ŧ	Comments	-	= R	tesolvable	Ŧ	Reference =
Fin	land		Researchers, persor	IS		ORCID			0		ORCID cannot be	created posthumous	ly			[48]
Fin	land		Researchers, persor	IS		ISNI			0							[48]
Fin	land		Organisations			VAT Number			6				N	lot Yet		[48]
Fin	land		Organisations			RoR			6							[48]
Fin	land		Organisations			ISNI			6							[48]
Fin	land	# =	Country =		Elomo		=	DII	D Stack =	Sta			Pof	arence =		[48]
Fin	land	# -	country .	Use case	e Lieme			F IL	D Stack	512			Ren			[48]
Fin	land	1	Netherlands	Identific	ation of	researcher(s) for funding and reporting	g	OF	RCID	Re	commended	https://hyp.is/dr5		[44]		[48]
Fin	land	2	Netherlands	Identific	ation of	research output for reporting		DC	וכ	Re	commended	https://hyp.is/dr5		[44]		[48]
		3	Netherlands	Identific	ation of	research output for reporting		UR	N Recommended <u>https://hyp.is/dr5</u>			[44]				
		4	Netherlands	Identific	ation of	research output for reporting		ISS	SN	Re	commended	https://hyp.is/dr5		[44]		
		5	Netherlands	Identific	ation of	research output for reporting		На	andle	Re	commended	https://hyp.is/dr5		[44]		
# =	Use Case	or Wo	kflow	Ŧ	Descrip	tion $\overline{\mp}$	DOI	ŀ	Benefits	-			Ŧ	Reference	÷ A	Annotation Link 😑
1	Grant app	licatio	n processes											[47]	h	https://hyp.is/jsNzLIIV
1	Grant app	licatio	n processes		Researc	hers and institutions pass PIDs for previous	DOI		1) Less manual data	a en	try			[52]	h	https://hyp.is/1enHZF
					grants, o	outputs, organisations, people, and projects	ORCID		2) More accurate d	ata	ling contactual infor	mation				
					to grant	application systems	RAID		5) More time for in	ciuu	ing contextual mon	mation				
1	Grant app	licatio	n processes											[44]	h	https://hyp.is/TtWjdF
2	Content s	ubmiss	ion processes											[47]	h	https://hyp.is/jsNzLllV
2	Content s	ubmiss	ion processes		Researc	hers share their ORCID iD when submitting	DOI		1) More accurate d	ata a	about co-authors, af	filiations, and funding		[52]	h	https://hyp.is/1enHZF
					new out	tputs and connect to ROR IDs for	ORCID		2) Time/effort savir	ngs i	n review and metad	ata creation				
					Instituti	ional affiliations and grant DOIs for funding	KOK		3) Easier sign-in	mor	aant					
									5) Easier reviewer s	selec	ction and recognition	n				

Entities

Currently being integrated with <u>FABIO</u>

# =	Entity =	Motivation	Reference =
1	Datasets	Data set Identifier for data set is needed to locate and identify the source of the (reused) data.	[56]
1	Datasets	PIDs for research data	[59]
2	Person	A person identifier is needed to identify to whom the (reused) data belongs to, and possibly how to a	[56]
2	Person	PIDs for researchers and contributors	[59]
2	Person	A core principle of Linked data is to use URIs to identify important entities such as people. The followi	[63]
2	Person	A RO-Crate SHOULD have contact information, using a contextual entity of type ContactPoint.	[63]
3	Organisation	An organi a belongs to, and ho	[56]
3	Organisation	PIDs for o	[59]
3	Organisation	An Organ cholarlyArticle or af	[63]
4	Software	Research	[56]
4	Software	PIDs for s	[59]
		Printed Music Instruments Books Series	

Researchers

Individuals

Stakeholders

# =	Category =	Entity -	Reference =	Annotation Link =
	Institutions	Public research performing organisations (RPO)	[47]	https://hyp.is/WwdROFITEe6
	Institutions	Higher Education Institutions	[47]	https://hyp.is/WwdROFITEe6
	Institutions	Research organisations	[47]	https://hyp.is/WwdROFITEe6
	Institutions	Research funding organizations (RFOs)	[47]	https://hyp.is/WwdROFITEe6
	Institutions	Policymakers	[47]	https://hyp.is/WwdROFITEe6
	Institutions	Libraries	[47]	https://hyp.is/WwdROFITEe6
	Institutions	(Science) Publishers	[47]	https://hyp.is/WwdROFITEe6
	Institutions	Service providers	[47]	https://hyp.is/WwdROFITEe6
	Institutions	Research Infrastructures	[47]	https://hyp.is/WwdROFITEe6
	Groups	Researchers	[47]	https://hyp.is/WwdROFITEe6
	Groups	Librarians	[47]	https://hyp.is/WwdROFITEe6
	Groups	Open Science/Open Access managers/coordinators	[47]	https://hyp.is/WwdROFITEe6
	Groups	CRIS system managers	[47]	https://hyp.is/WwdROFITEe6
	Groups	Repository managers	[47]	https://hyp.is/WwdROFITEe6
	Groups	Data stewards	[47]	https://hyp.is/WwdROFITEe6
	Groups	Data curators	[47]	https://hyp.is/WwdROFITEe6
	Groups	Researchers	[61]	https://hyp.is/kKYQnF4UEe6
	Institutions	Research organisations	[61]	https://hyp.is/7qA6VI4UEe6y
	Institutions	Research funding organizations (RFOs)	[61]	https://hyp.is/7qA6VI4UEe6y

Compliance Assessment Toolkit

Best Practices for Managers

Sources of Best Practice

- EOSC PID Policy
- FAIR-IMPACT
 - Review of Data and PID Policies (National, Institutional)
 - RDA Outputs and recommendations
 - Review of PID Stack documentation (recommended practices from providers, ...)
 - Published use of PIDs in workflows and specific use cases

Sources of Best Practice

Elements, Characteristics and Attributes of PID Stacks

This inventory and classification is in development by the FAIR-IMPACT project. It consolidates the expectations of the community, assertions and features advertised by PID Stacks, and the content of the EOSC PID Policy into a multi-level hierarchy.

Live version

Compliance Assessment Toolkit

Design

Main Components and Services

Compliance Assessment Toolkit

Demonstration

FAIRCORE4EOSC Compliance Assessment Toolkit

An easy to use Toolkit to measure compliance (ex. with PID Policy - 1st version) based on the results of Conceptual Model (definitions, relationships)

set of definitions for the concepts involved and the relationships between them

landscape of compliance assessment

CAT at a glance

the EOSC PID policy with services to encode, record, and query compliance with the policy. To do so, a wide range of compliance requirements (TRUST, FAIR, PID Policy, Reproducibility, GDPR, Licences) have been evaluated as use cases for definition of a conceptual model.

The Fundamentals

Users: The user that is using the Service

Actors: Different actors that perform specific actions on behalf of the Organization they belong to.

Validations: Request to be a specific type of actor on behalf of your Organization

Template of assessment: A way to create different types of assessments based on specific policies (using principles, criteria, tests), on specific actors and on compinatioon.

Assessment : Assess and document the current state of compliance of your object based on rules extracted from the policy you want (ex. the PID Policy)

Subject : The subject that the User wants to access (service,)infrastructure , url etc)

The Connections

Users: The user that is using the Service

Roles: The different types of Roles the user can have in the Service (Identified, Validated)

Actors: Different actors that perform specific actions on behalf of the Organization they belong to.

Validations: Request to be a specific type of actor on behalf of vour Organization

Template of assessment: A way to create different types of assessments based on specific policies (using principles, criteria, tests), on specific actors and on compination.

Assessment : Assess and document the current state of compliance of your object based on rules extracted from the policy you want (ex. the PID Policy)

Subject : The subject that the User wants to access (service, infrastructure , url etc)

Actors

- Different types of actors based on PID Policy
- A user can request different types actors **Requirement for a validation**

Validations

- Everyone can log in (identified)
- Multiple validation requests
 - different actors Ο
 - different organizations 0

Requirement for an assessment

Assessments

- Follows the assessment methodology (principles, criteria, tests)
- Json schema to support the relationships of the principles, criteria, tests
- Library of different types of tests
 - binary, value 0
 - working on automated tests (fetching results from external sources)
- Benchmarks, metrics supporting the calculation of required results.

ldentified

gets the role

automatically

Profile

id:c0f15711b0f26749734 003ca3c....194f5819@einf ra.grnet.gr Type: Identified Registered on: 2023-07-02T14:48:28.000 +03:00

id:c0f15711b0f26749734 003ca3c....194f5819@einf ra.grnet.gr Type: Identified Registered on: 2023-07-02T14:48:28.000 +03:00

The User Experience

PID Scheme

The User Experience / In Action requests a validation A user automatically Admin logins gets the role via AAI **L**Identified profile create new validation request id: c0f15711b0f26749734003ca3cbc30fbd1f6baffb8f3c315a771ee57194f5819@einfra.grnet.gr Account type: Identified registered on: 2024-01-23T13:44:08.000+02:00 Organization Name (*) Organization Source (*) Organization Website (*) ~ ROR Personal Details Name: Select... Surname: validation is moderated by Email: Organization Role (*) Actor (*) Update Details the administrators Select Actor \sim Validation Requests A You should update your personal details in order to be able to create validation requests User Details (*) Name: Themis Surname: Zamani Email: themis@admin.grnet.gr Assesments Submit \bigcirc 0 Subjects **Personal Details** Name: Themis \bigcirc Surname: Zamani Email: themis@admin.grnet.gr 0 Update Details 0 \bigcirc

The Assessment Experience

The Assessment Experience

create assessment	i General Info						
Step 1. Actor Step 2. Submission Step 3. Assessment	Name (*) GRNET-Service-Provider						
	Type eosc pid policy						
i General Info		\sim					
L Submitter							
Subject of assessment (Object, Entity or Service)	Subject of assessment (Object, Entity or Service)						
Rights, Licencing or Re-use	A Rights, Licencing or Re-use						
If you are not yet ready to share an assessment result or it is being done for int are visible to others it is private by deafult	If you are not yet ready to share an assessment result or it is being done for internal purposes only, keep it set to 'private'. Only the results of 'public' assessments are visible to others it is private by deafult						
 The Assessment is Public and Licensed with CC 4.0 BY The Assessment is Private 							
$\leftarrow \operatorname{Prev} \qquad \qquad \qquad Next \rightarrow \qquad $	Create	Close					

create assessment	L Submitter						
Step 1. Actor Step 2. Submission Step 3. Assessment	Name Themis	Surname Zamani					
	Affiliation National Infrastructures for Research and Technology - GR	ORCID ID					
i General Info		~					
Le Submitter		\checkmark					
Subject of assessment (Object, Entity or Service)	Subject of assessment (Object, Entity or Service)						
Rights, Licencing or Re-use		^					
If you are not yet ready to share an assessment result or it is being done f are visible to others it is private by deafult	for internal purposes only, keep it set to 'private'. Only the resu	Ilts of 'public' assessments					
The Assessment is Public and Licensed with CC 4.0 BY	The Assessment is Public and Licensed with CC 4.0 BY						
The Assessment is Private							
$\leftarrow \operatorname{Prev} \qquad \qquad \qquad Next \rightarrow \qquad $		Create Close					

create assessment	Subject of assessment (Object, Entity or Service)					
Step 1. Actor Step 2. Submission Step 3. Assessment	Select an existing subject from previous assessments or define a new one Subject (*) Select Object Clear Selection • Subject ID (*) • 01213					
i General Info	Subject Name (*) GRNET-PIDs Subject Type (*) Service	~				
Submitter		\sim				
Subject of assessment (Object, Entity or Service)		\sim				
Rights, Licencing or Re-use		^				
If you are not yet ready to share an assessment result or it is being done for internal purposes only, keep it set to 'private'. Only the results of 'public' assessments are visible to others it is private by deafult The Assessment is Public and Licensed with CC 4.0 BY The Assessment is Private						
← Prev Next →	Create	Close				

The Core: Assessment In Action / Guidance

The Core: Assessment In Action

edit assessment id: 79bad374-7dc	-3-485c-a81b-3dd0a66cf9a2	
Step 1. Actor Step 2. Submission Step 3.	Assessment	
Compliance: UNKNOWN 🗠 Ranking:	Mandatory: 0 / 5 Optional: 0 / 7	Here is the main bar of information
C5 - Update Functionality The PID manager MUST provide the functionality required to maintain PID attributes.	Progress C6: Ownership Transfer Required The PID manager SHOLIL D provide policies and contractual arrangements for transfer of ownership should	From here you can see the final result of the assessment
C6 - Ownership Transfer (required) 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	the owner no longer be able to assume responsibilities in compliance with the policy. part of principle P1: Application 1	Compliance: UNKNOWN The actual Ranking
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.	Are the specifics of the ownership transfer provisions publicly available through contract or procedure? Yes No	Ranking:
C11 - Versioning required PID services and PID Managers SHOULD have clear versioning policies.	URL: Please enter and add a valid url to support your claim Add TEST	And will you are typing you get informed about the number of mandatory and optional remaining criteria
C14 - Resolution Authenticity or Efficiency 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.		Mandatory: 0 / 13 Optional: 0 / 13
C34 - Persistence Mean (required)		Mandatory: 11 / 13 Optional: 12 / 13

Technologies Used

Test and send us your feedback

Test the service

Send Feedback

https://cat.argo.grnet.gr

https://forms.gle/gmCr1xhv7VbK4GAF9

New Features / Ideas - Next Steps

Next Steps

- Give the ability to users to create their own Policy and associated templates
 - Support templates of policies
 - Either copy an existing one
 - Create a new one based on a library of principles, metrics, tests
- More Guidance
 - General Guidance
 - Help topics menu
- Use of a Vocabulary service
- Support a Providers Registry
- Monitor the PID ecosystem \rightarrow coming next :)

Compliance Assessment Toolkit

Automated Monitoring of the PID Ecosystem

Automated Monitoring: Link Rot Checker Specification

Term / Abbreviation	Description
Reference rot	A reference that is experiencing either link rot or content drift.
Link rot	A resource identified by a URI that has ceased to exist and hence a URI reference to that resource will no longer provide access to referenced content.
Content drift	A resource identified by a URI of which the content has changed to such an extent that it ceases to be representative of the content that was originally referenced.

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

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

faircore4eosc.eu Twitter: @FAIRCORE4EOSC LinkedIn: company/faircore4eosc Youtube: FAIRCORE4EOSC

Enquiries: Anu Märkälä (CSC) <<u>anu.markala@csc.fi</u>> (Programme Office)

Testing CAT: Themis Zamani (GRNET) < themis@admin.grnet.gr >

Using CAT, Policy Alignment, Model, Survey: Wim Hugo (DANS) <<u>wim.hugo@dans.knaw.nl</u>>

Funded by the European Union

