Project Number: 101057264 Start Date of Project: 01/06/2022 Duration: 36 months

Deliverable D2.1

Compliance Assessment Specification

Dissemination Level	PU
Due Date of Deliverable	(31/05/23), Month (M12)
Actual Submission Date	(31/05/23)
Work Package	WP2 - Compliance Assessment Toolkit
Task	T2.1
Type	Report
Version	V 1.0
Number of Pages	p.1 – p.69

Deliverable Abstract

The report provides a set of requirements and specifications for a Compliance Assessment Toolkit, to be implemented in support of the EOSC PID Policy. The specifications are based on a Conceptual Model for compliance assessment, developed by validation against a number of compliance assessment case studies.

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.

Funded by the European Union

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

DELIVERY SLIP

Contribution	Name	Partner/Activity	Date
Lead author(s)	W Hugo	KNAW-DANS	11-05-2023
Contributor(s)	W Steinhoff	KNAW-DANS	11-05-2023
	D Turner	KNAW-DANS	11-05-2023
	M Buys	DataCite	11-05-2023
	T Zamani	GRNET	11-05-2023
Reviewer(s)	J Hakala	UH (National Library of Finland)	17-05-2023
	R Wilk	AGH/AGH-UST	17-05-2023

DOCUMENT LOG

Issue	Date	Comment	Author/Editor/ Reviewer	ORCID ID
v.0.1	17-04-2023	Inception and first draft	W Hugo	0000-0002-0255-5101
v.0.2	26-04-2023	Elaboration of Sections 1 and 2	W Hugo	0000-0002-0255-5101
v.0.3	11-05-2023	Version for internal review	W Hugo	0000-0002-0255-5101
			M Buys	0000-0001-7234-3684
			D Turner	0000-0003-2216-8181
			T Zamani	0000-0003-1148-1738
			W Steinhoff	0000-0003-1106-8441
v.0.4	17-05-2023	Internal Review	J Hakala	0000-0003-1067-5020
	17-05-2023	Internal Review	R Wilk	n/a
v.0.5	29-05-2023	Address internal review comments	W Hugo	0000-0002-0255-5101
v.1.0	31-05-2023	Editorial review and final version	A Märkälä	n/a

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	A 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	Regulations aimed at protecting the right to privacy of individuals in the EU in a digital context.
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 may 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.

Table of Contents

List of Figures

List of Tables

List of Annexure Figures

List of Annexure Tables

Executive Summary

This report outlines a set of standards, API specifications, and vocabularies that define the nature and capabilities of compliance assessment, encoding, and verification services and infrastructure. It is divided into two sections. The first section provides a review of the conceptual model for the Compliance Assessment Toolkit (CAT). The second section details the requirements and specifications.

A portfolio of case studies identified in the Landscape Assessment Report [2] was selected to validate and extend the Conceptual Model. This ensures that it is generic enough for a variety of compliance assessment applications. The Conceptual Model developed in support of the CAT addresses many of the issues identified in the Landscape Assessment, including unstructured assessment, semantic ambiguity, differences and overlaps, granularity (criteria and entities), multiple evaluation mechanisms, multiple outcomes for the same criteria, and applicability to all research outputs. Solution characteristics identified previously were also reviewed and amended for the CAT implementation and model development.

Three solution components comprising the CAT include an agreed Conceptual Model or ontology for the concepts, entities, and relations between them, identification of vocabularies and registries required to support instances of the model, and a service specification that defines generic methods and payloads required to create, maintain, and add to instances of the Conceptual Model elements. The requirements and specifications define the specific implementation of the CAT for the EOSC PID Policy, while ensuring that it is interoperable with other similar graphs.

Section 1: A review of the conceptual model for the Compliance Assessment Toolkit.

Section 2: Requirements and Specifications, consistent with the FAIRCORE4EOSC template for requirements documentation.

An important result from the work to date is presented in Annexure D.3.1. Here, the EOSC PID Policy is analysed and aligned with the Conceptual Model, and a number of recommendations for refinement of the policy - largely in respect of semantics and organisation - are made. These recommendations are only stated in general terms in the Annexure, since a process of engagement between FAIRCORE4EOSC, the EOSC PID Policy Task Force, and FAIR-IMPACT is in place whereby these recommendations can be discussed and adopted, as applicable.

Section 1: Conceptual Model

1. Context

The FAIRCORE4EOSC Compliance Assessment Toolkit (CAT) is being developed to assist actors in the PID ecosystem[1] with assessment of their compliance with the policy. The reasons for conducting such an assessment can vary, and it is worth elaborating this aspect of assessment at the outset.

Compliance assessment is a widely used term and ranges from informal review processes (that are difficult to reproduce), to formal processes such as audits (that are often conducted in such a way that it is somewhat or fully reproducible). Assessment can be performed by humans or machines (or a combination of the two), and the tests that are employed can vary significantly in respect of formalisation, standardisation, and benchmarking. Many of these considerations are identified, evaluated, and included into the Conceptual Model, which is discussed in Section 1 of this report.

From our analysis of a range of compliance assessment case studies [2], and based on External Advisory Board feedback [3], the following broad considerations are identified[2]:

1. **Consequences**: The consequences of assessment can be significant (financial or reputational loss, opportunity loss, exclusion, and so on), or not (internal identification of risks or performance shortcomings, assessment of maturity to direct future focus).
2. **Reproducibility**: The nature of the criteria included in an assessment, and the way in which these are evaluated (measures or metrics), can vary in terms of reproducibility. In cases where there are significant consequences attached to the outcome of an assessment, reproducibility to the maximum possible extent is important. Qualitative measures are more difficult to reproduce, but techniques are available to assist (benchmarking, peer review, expert consensus, pairwise comparisons, etc.).
3. **Transparency**: The level of transparency required of an assessment process is also linked to the consequences of assessment. If the consequences of a negative or unfavourable assessment is significant, transparency (and possibly recourse and right of appeal, re-assessment, and so on) will be an important aspect.
4. **Privacy**: assessment results can be public or private. If consequences of negative outcomes are significant, it may be required to keep assessment results private. There is, however, ***an important set of exceptions*** to these that broadly cover public claims (claiming trustworthy repository status, GDPR compliance, EOSC Policy compliance, FAIR alignment, and so on). It is clear that public claims must be substantiated by publicly available evidence. It is, however, also reasonable to make only successful assessments public, as is, for example, the practice at CoreTrustSeal.

The CAT is a tool, and does not determine the way in which it will be applied - but it is important to accommodate the elements described above. This includes, *inter alia*, provision for public and private assessment records, the ability to perform self-evaluation prior to a possible formal evaluation, and publication of code, tests, benchmarks, and the like.

[1] These are formally defined in the EOSC PID Policy. Suggestions for refinement are made in Section 2.

[2] The code and infrastructure for the CAT can be reused for other types of assessment, and while these are not the primary aim of the work done in the FAIRCORE4EOSC project, it is nevertheless important to include them here to provide context.

2. Landscape: Compliance Assessment in General

Repositories that curate and preserve research outputs face a complex array of expectations and demands in terms of their performance and quality. These demands cover a wide range of topics, ranging from governance and sustainability, through process and systems characteristics, to the scope and nature of service and infrastructure provision.

The major current focus on repository and service compliance with community expectations, as exemplified by TRUST [21], FAIR [25], and to a lesser extent, CARE [36], has highlighted the need for a uniform approach to the encoding, assessment, recording, and application of compliance assessment and measures [8]. In addition to these broad-based community expectations, compliance is also required in respect of regulations and policies (for example PID policy in EOSC [1], requirements for reproducibility [52], [53], privacy legislation and regulations in the EU [56]), and desirable in respect of the principles and well-established architectural patterns in the research data infrastructure domain specifically, and for digital systems in general.

Funders have also adopted Open Science objectives [9] as a broader (and somewhat aspirational) expectation of the research community's contribution to society, and these aspirations place indirect requirements and expectations on the repository infrastructure community to generate and provide metrics about performance and compliance.

Finally, there is an understanding that participation in regional and global infrastructures will require some form of compliance monitoring, for example membership of the World Data System [10], joining a Data Commons [46] or the EOSC Rules of Participation [11].

The need for unambiguous compliance specification, encoding, measurement, and monitoring is thus required by an intersection of all of the above, with nuances and variations dependent on context. In practice, such an approach will help to avoid duplication and divergence.

The foundational aspects of the Conceptual Model were documented originally in an internal FAIRCORE4EOSC Report [2], and a summary version of this report can be accessed online. Original work on the conceptual model was done within RDA in the RDA/ WDS Certification of Digital Repositories IG [8], and the contributions3 of interest group members are acknowledged.

In the next section a number of specific case studies are identified. These will be analysed to validate and refine the Conceptual Model.

[3] Barbara Sierman, Jonathan Petters, Bob Downs, Dawei Lin, John Westbrook, Wim Hugo.

3. Case Studies and Model Validation

A ***Conceptual Model*** is seen as a cornerstone of the Compliance Assessment Toolkit, since its data model, required vocabularies, and Linked Open Data registries will be based on the model. A number of candidate case studies were identified in the Landscape Assessment [2], and these were further analysed by FAIRCORE4EOSC Task 2.1 to validate and extend the Conceptual Model originally proposed in RDA [8]. This process ensures that the Conceptual Model is both generic enough for extension to a variety of compliance assessment applications, and sufficiently mature (TRL4) to use as a design basis for the Compliance Assessment Toolkit.

The case studies identified in the Landscape Assessment Report [2] represent a variety of **Motivations**[4], and a subset of cases were selected for detailed analysis. This scope is summarised below and details are presented in Annexure D.1, and the implications of the case studies for model organisation is discussed in Annexure D.2.

Table 1 – Scope of Case Studies

Motivation Type	Category	Case Studies	All	Analysis		
				M12	M18	M32
Community Expectations	TRUST	CoreTrustSeal, TRUST Principles, Nestor, ISO 16363	4	2	2	
	FAIR	FAIR Principles, a range of FAIR evaluation tools, FDO	8	6	2	
	CARE/ Ethics	CARE, Data Access Requests, GDPR	3		1	1
	Reproducibility	CORE-2, RO-Crate, CURE-FAIR	3		1	2
	Others	POSI, Publisher's Requirements for Data Repositories	2	1		1
Policy/ Regulatory	Legal	GDPR, Licence Compliance	2		2	
	Policy	EOSC PID Policy	1	1		
Rules of Engagement	Network Membership	World Data System, Data Spaces, FIP, FDP, Data Commons	4	1	1	2
	Infrastructure Inclusion	EOSC Rules of Participation	1	1		
	Calls and Recruitment	Open Call Example	1	1		

Notes:

1. The total number of cases identified per category are shown as ‘All’.
2. This report is published in month 12 of FAIRCORE4EOSC (M12), and the column shows the number of cases evaluated to date. Additional cases will be analysed to be included in documentation updates (M18, M32).
3. Totals may not add up since not all identified cases will be analysed, and some cases apply in more than one category.

[4] Motivations are defined categories in the Conceptual Model, and as such, form part of the vocabulary.

4. Typical Issues

In the Landscape Assessment, several systemic issues were identified [2]. Many of these, in turn, were originally identified in RDA [8]:

- **Unstructured assessment** of important portfolios of principles, such as TRUST, FAIR, and other areas of compliance assessment and monitoring are occurring.
- **Semantic ambiguity** in respect of concepts (principles, criteria, benchmarks, metrics, indicators, best practices, maturity, recommendations, standards, …) and the relationship between these concepts need to be addressed.
- No clear definition of the **differences and overlaps between sets of principles** (e.g. TRUST and FAIR) and semantic alignment between them.
- Not all criteria, benchmarks, and best practices are specified at the same **level of detail** or **granularity**. Moreover, metrics for levels of maturity associated with these criteria can apply at many levels of detail.
- **Multiple evaluation mechanisms** and tools for the same criteria are emerging. There is a diversity of test methods and implementations of tests, often for the same criterion.
- The **same criteria can be duplicated** in more than one assessment approach, sometimes resulting in different outcomes.
- Open Science expectations require the principles of TRUST, FAIR, and similar initiatives to **apply across all research outputs**, not only data. It is likely that many criteria, practices and recommendations, and metrics will be similar or identical, but there are also specific cases where this will not be possible.

5. Elements of A Solution

5.1 Identified Problems

The Conceptual Model developed in support of the Compliance Assessment Toolkit (CAT) implementation will address many, but not all of the issues identified above. In the table below, we summarise this for convenience:

Table 2 – Problems Identified

Issue or Problem	Conceptual Model	CAT Implementation
Unstructured assessment	Addressed directly through modelling of and vocabularies for measures (metrics), tests, benchmarks, and similar.	Available to encode and encourage structured assessment mechanisms for qualitative measures. Populated for EOSC PID Policy.
Semantic ambiguity	Conceptual model defines main concepts, entities, relations, and vocabularies required for precise description and encoding.	Implemented with specific data that defines the vocabularies, code lists, registries, and concepts/ entities associated with EOSC PID Policy.
Differences and Overlaps	All assessment approaches for the same topic can be encoded and described in parallel, assisting with identification of differences and overlaps.	No specific application since only one motivation (EOSC PID Policy) is implemented.
Level of detail, granularity (criteria)	The conceptual model makes provision for recursive relations between criteria, which enables capturing of this complexity.	No specific application since only one topic (EOSC PID Policy) is implemented, and no recursion is foreseen.
Level of detail, granularity (entities)	The model is extended to include Entities, and the Objects and Services they own.	A specific model extension is developed for EOSC PID Policy, and foundational data to support this extension is pre-populated in the model.
Multiple evaluation mechanisms	The model defines Measures, Tests, and Benchmarking of results, which means that differences can be identified and potentially aligned: 1. Different tests for the same measure 2. Different benchmarks for the same test	This scenario will not arise in the CAT implementation - tests are agreed with the EOSC PI Policy Task Force in consultation with experts, and they will define benchmarks for the tests.
Multiple Outcomes, Same Criteria	The model accommodates multiple outcomes for the same criterion, for example in cases of more than one independently developed evaluation mechanism for the same set of criteria[5]. This assists with detection of divergence.	This scenario will not arise in the CAT implementation: each criterion has one metric (measure).

[5] This scenario is especially prevalent in FAIR evaluation - in all, eight case studies were identified for validation, and there may be many more.

Applicable to All Research Outputs	The model is agnostic of the type of object or service being evaluated, and hence can be applied to all or any.	By design, the CAT will assess PID ecosystem elements applicable to all research outputs. In practice, the focus might be largely on data and software.

5.2 Desirable Characteristics of a Solution

The Landscape Assessment also adopted solution characteristics identified earlier by RDA [8]. These have been reviewed and amended for the CAT implementation and model development:

Table 3 – Desirable Characteristics of a Solution

#	Design Element	Description	Reference
1	Conceptual alignment	Reduce the large diversity of opinions and definitions about *[motivations and assessment needs]*, their criteria and implementation practices, how they are measured	[8], [0]
2	Simple, using existing standards	Any solution should be based on existing web and data infrastructure standards and not require any new standards, but rather develop recommendations in respect of the semantics of compliance encoding, recording, and measurement based on existing standards.	[8]
3	Federated in practice, conceptually a single entity	It is highly unlikely that all compliance information for a specific object will ever be recorded, preserved, and published by a single source: in practice, such information will be scattered in many locations and services, and the best possible solution will limit the complexity of the federated information space by standardising its encoding, implementation, and vocabulary. *[The EOSC PID Policy compliance assessments can be federated in theory, in practice they will likely be centrally available via an EOSC service]*	[8], [0]
4	Machine and human readable, machine actionable	Solutions need to consider from the start that navigation and application of the compliance information for an object, service, subject (nodes), or collections of these will be complex, potentially involve many thousands or even millions of records, and may not result in a unanimous assessment of the compliance characteristics of the node in question. With this in mind, machine actionability - both in terms of aggregation and subsetting, as well as analysis and potential ML and AI applications, is a design imperative.	[8]
5	Precision and flexibility	Map and define relations between sets of principles, criteria, metrics, and doing so flexibly - for example, allowing nested criteria with metrics and levels of maturity coupled to any level of detail, as required. Unambiguously define best practice, mandatory or optional recommendations, guidelines, and so on.	[8]
6	Reduced Complexity	Improve understanding about principles, criteria, and metrics, etc. and on how these can be applied in practice.	[8]
8	Parsimony	Minimise the set of applicable criteria and their formulations, standards, metrics, and *[benchmarking]*/ maturity definitions, and/	[8], [0]

		or reduce duplication and complexity.	
9	Relational nature of compliance information	Compliance information should be seen as a property or properties of a relation between a digital object or service and a context - in this case, a compliance measurement event. Compliance information cannot and should not be seen as part of the metadata associated with a digital object in a one-to-one relation, since it can be evaluated by multiple tools, mechanisms and institutions, against several divergent or competing criteria, with varying levels of assertion, and so on.	[8]
10	Universally applicable	Avoid duplication of compliance principles, implementations, criteria, measures, and metrics across different research outputs.	[8]
11	Accommodate humans and machines	Make the certification ecosystem machine-readable and actionable where feasible, recognising that some measurements rely heavily on human assessment - sometimes on site. One should accommodate the likely increased reliance on AI and ML to assist evaluation of complex or qualitative metrics. *[The revised version of the Conceptual Model accommodates automated tests, irrespective of whether these are deterministic, probabilistic, or AI-derived. Test properties and provenance should allow user evaluation of this aspect].*	[8], [0]
12	Standardised measures and reporting	Map institutional/ repository/ object/ service compliance onto a formalised structure and agree on mechanisms for evaluation and disambiguation of multiple and possibly divergent assessments of the same node or object.	[0]

5.3 Solution Components

A solution to the typical issues and requirements expressed above can be obtained by defining and implementing a ***Compliance Assessment Toolkit***, which in turn has three components:

1. An agreed Conceptual Model or ontology for the concepts, entities, and relations between them;
2. The conceptual model identifies vocabularies and registries required to support instances of the model, and will develop initial versions of these to be validated by and transferred to the community for future maintenance;
3. A service specification, which defines generic methods and payloads required to create, maintain, and add to instances of the conceptual model elements.

The remainder of this section deals with these three solution components. A requirements and specifications document, which defines a specific implementation while ensuring that it is interoperable with other similar graphs, is the subject of section 2.

6. Conceptual Model

The Conceptual Model developed initially by RDA [8] identified a number of entities and concepts, categorised in four main groups:

1. Motivations: what are the Origins, Principles, and Objectives that motivates the assessment?
2. Verification: what will signify that a certain principle or objective has been satisfied or achieved? Moreover, who will do this, and what is their Mandate or Authority to do so?
3. Implementation: How will achievement or satisfactory performance be determined, assessed, or evaluated? What Tests, Metrics, Benchmarks, and Maturity Levels are involved, and how are these applied?
4. Elaboration: on what basis are criteria identified, and what does this mean in practice? What is the role of Best Practices, Guidance, and Recommendations in all of this?

Extensions and refinements were identified by FAIRCORE4EOSC after validation of the RDA model [2] against case studies:

1. Entity details and relations, and a need to define the Objects and Services that are being assessed;
2. The same entity, object, or service may be assessed for more than one Use Case, implying a change in applicable criteria[6].
3. The need for an Assessment Method definition: how will a set of metrics be assessed to determine compliance (or not)?;
4. An additional relation between Metrics and Tests. The RDA version of the model [2] only indicates that metrics are derived from tests, the extension explicitly defines an Algorithm that formalises this relationship.

> Before the model is discussed in detail, it should be noted that there is an important element to many of the typologies and vocabularies proposed here: the model identifies that such a vocabulary is required, but the content of the vocabularies or typologies are proposals that serve as a starting point for model implementation. In many cases, the vocabulary will be refined with community input during the course of the project, and it will most likely be adjusted over time.

6.1 Motivation

The model starts with standardising the motivation for assessment, and the scope of the assessment.

There can be many motivations for (or origins of) the need for assessment, and these have been categorised in the model using a **Motivations Typology[7].** For convenience, these are also included in the diagram, and they are explained below:

1. **Policies and Regulations**: compliance is often mandatory and is intended to ensure legality or a minimum level of performance. Examples include the EOSC PID Policy, or GDPR.
2. **Rules of Engagement:** These ensure a minimum level of performance, readiness, or maturity as a prerequisite for participation in a network, an infrastructure, or a consortium.
3. **Community Expectations**: These are broad community expectations of the performance of the ecosystem, and some of these are formalised in sets of principles (FAIR, TRUST, CARE) and in some cases, realised as sets of criteria and/ or elaborated as expected behaviour or levels of performance (guidance, best practices).
4. **Good Digital Systems**: these motivations include risk aversion (for example by requiring two-factor

[6] For example - DOIs may be used to reference digital objects, specimens, and instruments, amongst others - in which case the criteria for governance or the duties of PID managers could be different.
[7] An example of a proposed vocabulary.

authentication, open-source code, and the like)[8].

5. **Architectural Patterns and Frameworks**: for systems engineering, these will be important considerations **-** elements such as interoperability, modularity, topology, and scalability are included in this set of motivations[6].

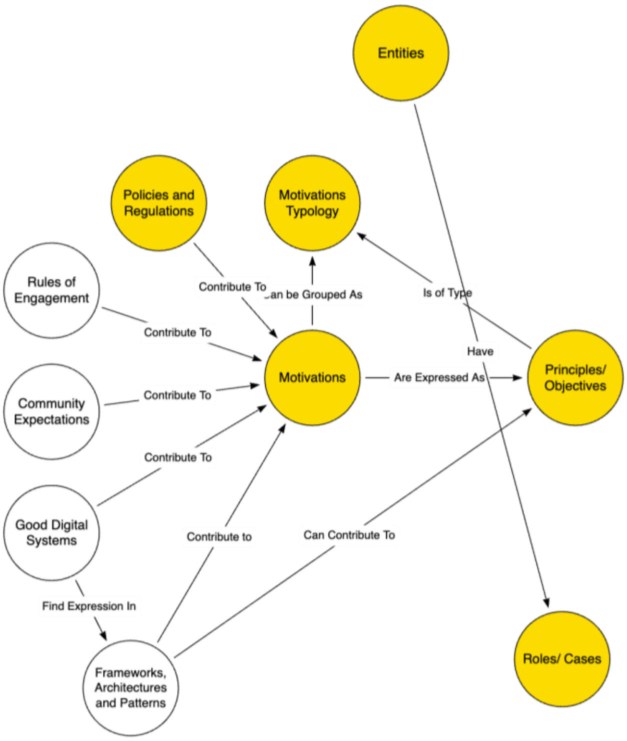


Figure 1 – Motivations for Compliance Assessment

The scope of the motivation is also important. If we develop a set of **Principles** or **Objectives** to support a policy, for example, the **Entities** to which it is applied, and the different **Use Cases** and **Roles** involved is an important differentiator: not all criteria apply to all use cases and/ or to all entity roles, as we will see. In other words, These entities have Roles (and/ or implement Use Cases), as defined in the policy.

The entities may be responsible for or own **Objects** or **Services** that require persistent identification, and this aspect is discussed later.

Motivations can be arranged in a hierarchy if needed, and a motivation can be linked to more than one typology.

In this specific case (FAIRCORE4EOSC CAT), the **Motivation** is to develop and apply **Policy** to manage the quality and use of PIDs in EOSC, and it applies to **Entities** (organisations, individuals) participating in the ecosystem. Not all of these entities participate in the same **Use Cases**. These nodes are shown in yellow in the diagram.

6.2 Verification

The desirable properties or outcome of a specific motivation is usually expressed as a set of **Criteria** or **Provisions**. We will provide detail in later sections on the origin of these criteria and provisions. These criteria and provisions serve to ***verify*** that principles are adhered to or objectives are met. Meeting criteria or provisions is considered to signal compliance with a principle or alignment with/ support of the objective. Note that multiple **Principles** (or Objectives) can reference the same criterion.

Criteria can have **Typology** that are not necessarily based on or aligned with the principles they are supporting.

[8] We did not investigate or analyse this aspect in great detail in the current document but may do so in future.

The use cases provide examples of these, where, for instance, the EOSC PID Policy [1] groups provisions and criteria into sections. The CoreTrustSeal criteria [22], in another example, are also grouped into a 'focus', some dealing with governance and the organisation, some with technology, etc.

It may be possible and desirable to generalise these into a typology, but we have made no attempt to do so authoritatively in the current model version. There are some suggestions in the typology in respect of model policy elements based on work published for data policies [60].

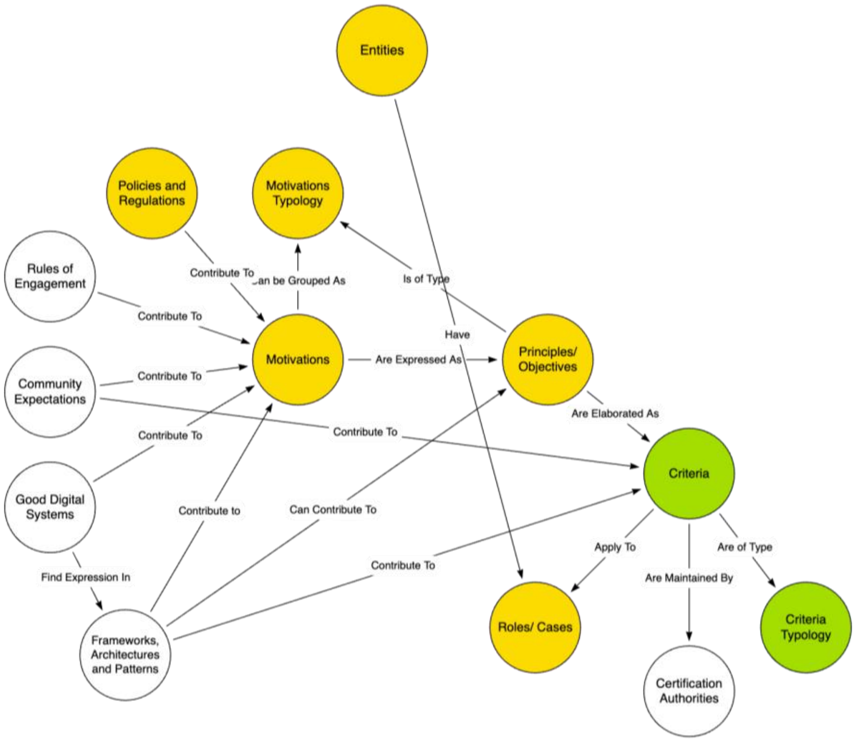


Figure 2 – Criteria for Compliance Assessment

6.3 Implementation: Metrics and Tests

To determine compliance requires measurement of the actual performance against a goal (Standard or Benchmark) - in short, a Metric.

Good metrics are reproducibly measurable by anyone, and preferably quantitatively, but in some cases, one cannot avoid qualitative measures. Moreover, a criterion is not much use if it cannot be measured, because compliance cannot be evaluated.

A metric is measured using one or more Tests. These are then compared to Benchmarks (Standards) to contextualise performance. Each test has a Typology that is not fully depicted in the diagram, but deals with aspects such as Precision, Reproducibility, and classification of the Result Type. In later sections, we define vocabularies for classification of tests. In addition, the standards or benchmarks can be normative or informative, and may involve a formal definition of this aspect [14].

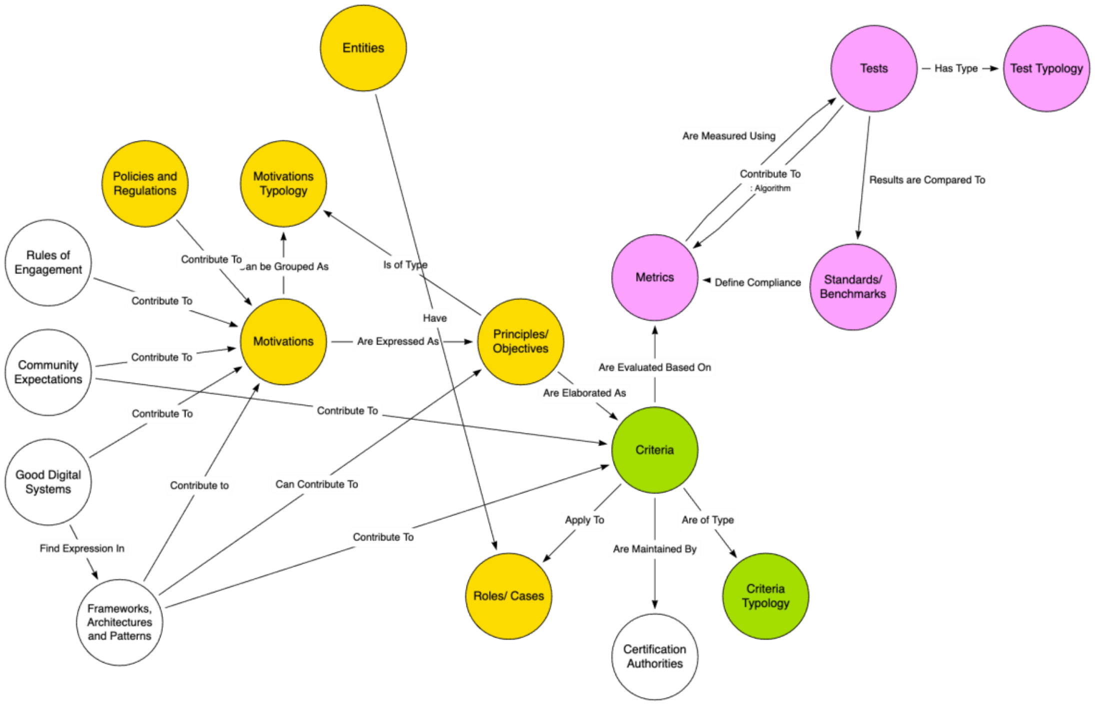


Figure 3 – Metrics, Tests, and Benchmarks

Finally, a metric is often composed of multiple test results, combined by way of an Algorithm. These algorithms can vary quite considerably in terms of approach and complexity, and this is dependent on the test typology. It is, for example, easier to combine quantitative results (binary, real or integer numbers) into an algorithm than it is for completely qualitative assessments that are expressed by way of narrative or text (for example - comments by a reviewer). In some cases, the algorithm belongs to an algorithm family (for example multi criteria analysis, or analytic hierarchy processes).

6.4 Implementation: Assessment

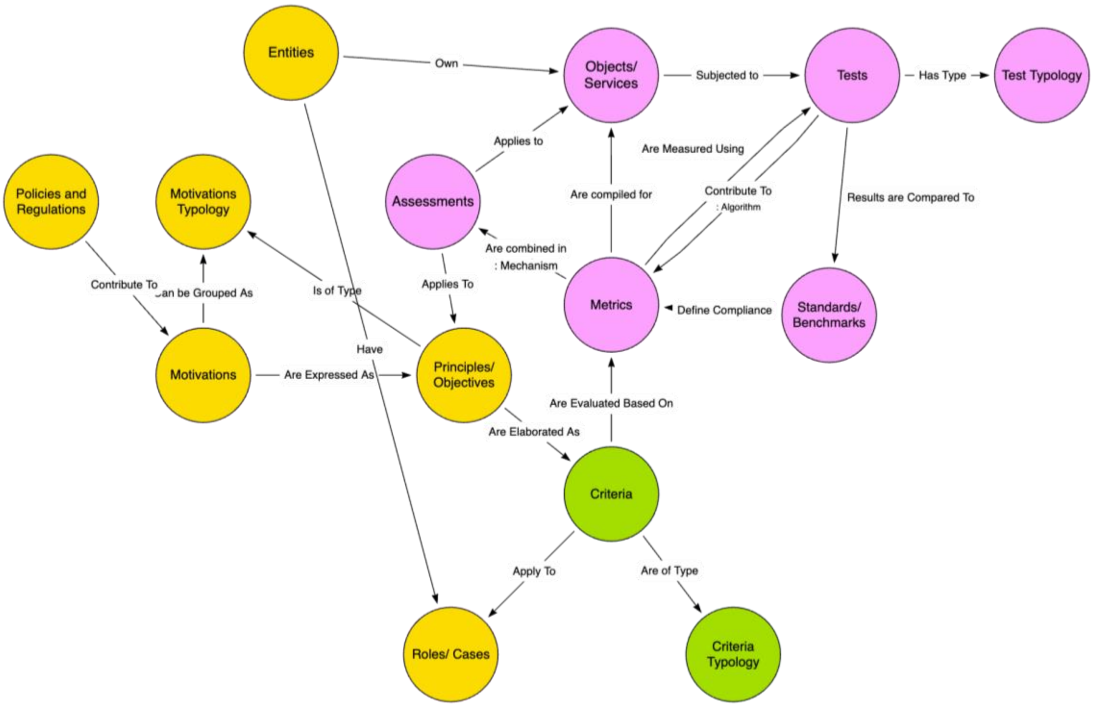


Figure 4 – Assessment

The tests that are applied, apply to a specific Object or Service owned or offered by an entity. Once all metrics have been computed, using an algorithm applied to test results, it is possible to compile an Assessment.

To do so, one needs to evaluate the results for each criterion or provision, based on a Mechanism or Methodology for assessment.

A common scenario entails classifying criteria by principle or objective, in which case an assessment method and outcome is available for each principle or objective. Mechanisms for assessment are often characterised by some type of weighting and ranking. In these cases, criteria are assigned a weighting, or a combination of mandatory and optional criteria: all mandatory criteria need to be complied with, whereas optional criteria could be combined into some form of ranking. These mechanisms can, again, also make use of complex methodology such as multi-criteria analysis.

The model makes provision for storing the elements required for assessment, but the data values required for the EOSC PID Policy Compliance Assessments must still be determined in consultation with the EOSC PID Policy Task Force. Example data sets and vocabularies have been developed.

6.5 Elaboration

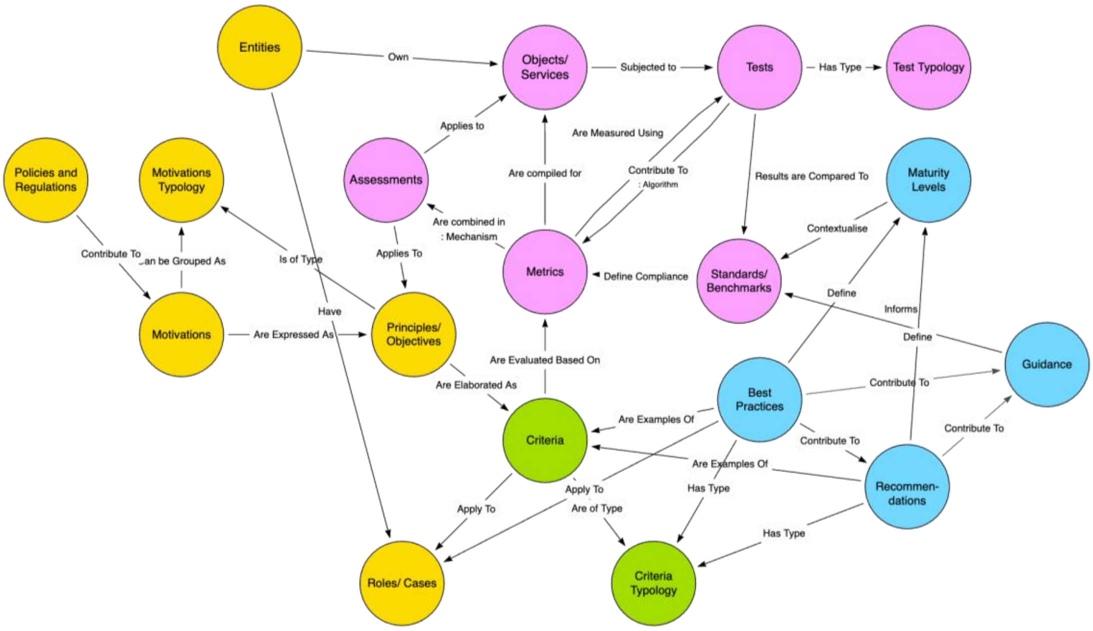


Figure 5 – Additional Context for Compliance Assessment

These model elements deal with the context associated with assessment and are capable of encoding and linking several community-driven resources.

Firstly, community efforts can identify sets of Best Practices. These best practices are applicable to one or more use cases or roles, and some of them can be formulated as Recommendations to the community. The best practices and recommendations, in turn, can be included into Guidance.

It is useful and prudent if criteria are based on community-developed best practice and recommendations, but this is not always the case. The model should allow such relations, ensuring that it is possible to encode links where they exist.

Finally, one can use guidance, best practice, and recommendations to derive benchmarks, and use them to define Maturity Levels. The maturity levels, in turn, can contextualise performance and benchmarks. Maturity levels are often arbitrary, and we map some of the examples we identified in use cases against the original Capability Maturity Model (CMM) [59].

6.6 Tools and Certification Authorities

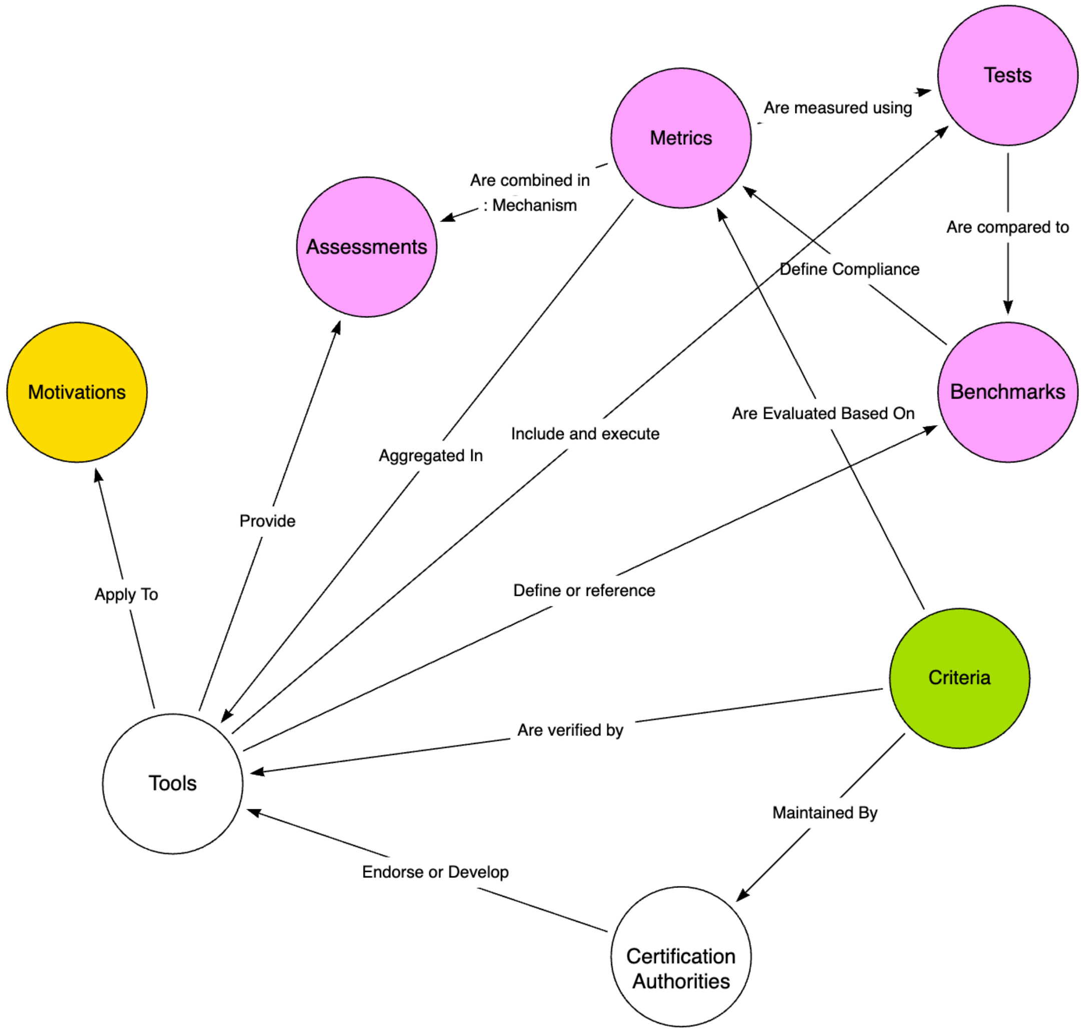


Figure 6 – Tools and Certification

In some cases, Tools are developed to include and execute the Tests needed to measure compliance, and it optionally does so by contextualising the test results against Benchmarks or Standards. Benchmarks can also define compliance categories for a specific Metric. Tools could optionally provide a method of Assessment and are applicable to a Motivation.

Moreover, there may be one or more Certification Authorities involved in assessment of compliance, and these authorities may develop and provide, or endorse Tools. Note that even informal or unsophisticated assessment approaches are nevertheless regarded as tools and can be encoded as such if needed.

6.7 Entities

The entities and roles of those entities (“Actors”) referenced in the EOSC PID Policy [41] has an implied model that defines relations between them. Each motivation may have its own specific model. The model for the EOSC PID Policy is shown below, based on the actors and stakeholders defined in the policy. The actors and roles are discussed in more detail in Section 2.

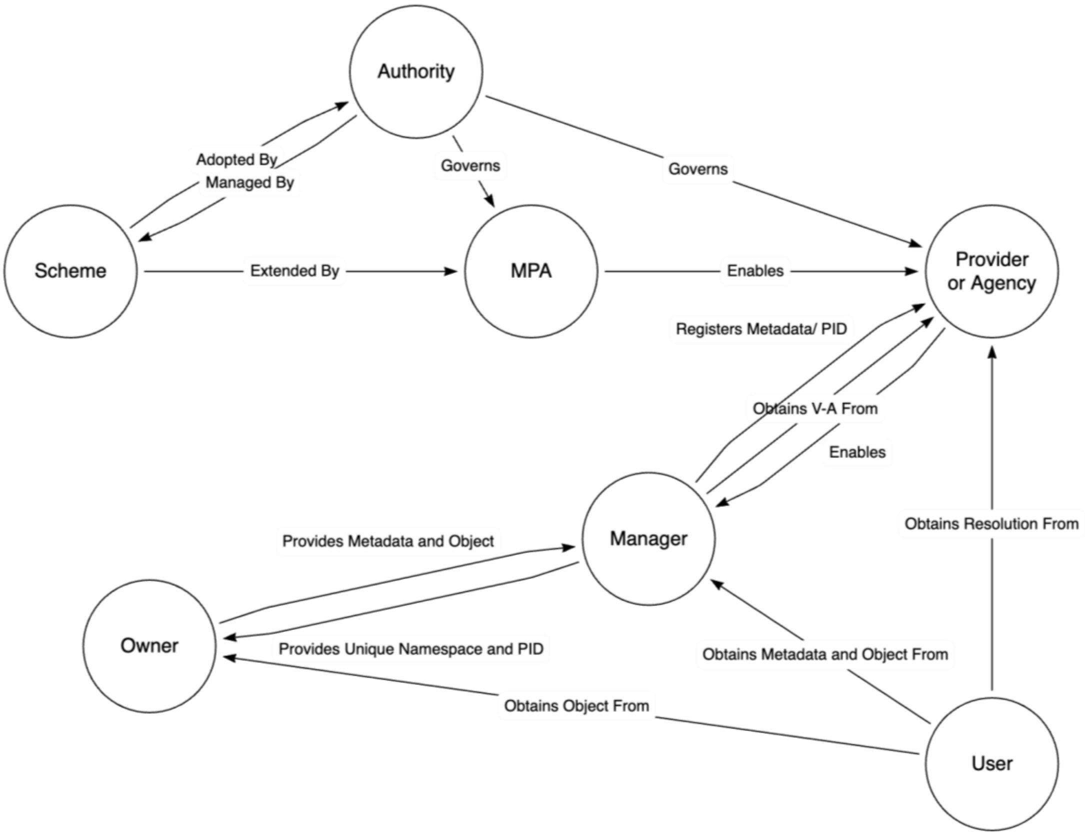


Figure 7 – Ecosystem Entities and Relations

7. Vocabularies and Registries

Implementation of the conceptual model to support any specific implementation requires a number of vocabularies and registries to be available. These can be grouped as follows:

1. Assessment Dependence: some vocabularies and registries depend on the nature of the assessment being made, while some are foundational to the conceptual model.
2. Ownership: some vocabularies and registries are maintained by the community or by external service providers, while some are owned by the Compliance Assessment Toolkit. In some cases, proposals are seeded in the vocabulary or registry, but the community can add to the resource ("Seeded").

The table below summarises the list of vocabularies and registries required by the Conceptual Model, and categorises them in terms of the above.

Table 4 – Vocabularies and Registries identified in the Model

Vocabulary Requirement	Type	Owner	Dependency	Typical Values
Concepts, Entities, and Relations. Relations are shown below.	Ontology	Model	Foundational	The ontology of the compliance graph. Concepts are typically captured in vocabularies, and entities in registries, but not exclusively so.
Domain - Use Case	Relation	Model	Foundational	These relations are fixed, and form part of a model relation set.
Principles - Motivations	Relation	Model	Foundational	
Principles - Criteria	Relation	Model	Foundational	
Criteria Use - Cases	Relation	Model	Foundational	
Guidelines - Criteria	Relation	Model	Foundational	
Guidelines - Benchmarks	Relation	Model	Foundational	
Criteria - Certifier	Relation	Model	Foundational	
Assessments - Objects/Services	Relation	Model	Foundational	
Metrics - Assessments	Relation	Model	Foundational	
Objects/Services - Metrics	Relation	Model	Foundational	
Metrics - Criteria	Relation	Model	Foundational	
Metrics - Tests	Relation	Model	Foundational	
Benchmarks - Tests	Relation	Model	Foundational	
Entities - Entities	Relation	Community	Foundational	These are sets of relations (multiple relations possible between the same nodes), and may

Entities - Use Case	Relation	Community	Foundational	depend on community modification.
Entities - Objects/Services	Relation	Community	Foundational	
Guidelines - Guidelines	Relation	Community	Foundational	
PID Use Case Ontology[9]	Ontology	Seeded	Assessment	A classification of use cases in terms of features and performance requirements (governance, cost, volume, …)
Entities	Registry	Seeded	Assessment	Registry of entities (institutions) and object types in the ecosystem
Users	Registry	Community	Application	Users registered and optionally validated
Use Cases	Registry	Seeded	Assessment	Use cases for which criteria apply
Domain	Registry	Community	Application	The domains for which use cases apply
Certifier	Registry	Seeded	Assessment	The certification authorities mandated to assess compliance
Objects/Services	Registry	Seeded	Assessment	The objects and/ or services being assessed
Assessments	Registry	Seeded	Assessment	An assessment for a service or object
Standards (Benchmarks)	Registry	Community	Assessment	Registry of standards and benchmarks - links to Information on how to interpret a test result
Recommendations	Registry	Community	Assessment	A registry of recommendations as annotations from published work, linked to criteria
Guidance	Registry	Community	Assessment	A registry of guidance as annotations from published work, linked to criteria
Best Practices	Registry	Community	Assessment	A registry of best practices as annotations from published work, linked to criteria
Principles	Registry	Seeded	Assessment	{TRUST, FAIR, CARE, PID Policy, …}
Cases	Registry	Seeded	Assessment	A registry of use cases, typified using the PID Use Case Typologies
Criteria	Registry	Seeded	Assessment	FAIR criteria, CTS criteria, PID policy provisions, …
Metrics	Registry	Seeded	Assessment	An inventory of registered metrics

[9] This aspect is a separate initiative in FAIR-IMPACT and can be used to tag (classify) use cases.

Tests (Methods)	Registry	Seeded	Assessment	Registry/ inventory of test methods
User Roles	Vocabulary	Seeded	Application	User roles defined by application(s)
User Identifier Typology	Vocabulary	Community	Application	Types of user identities, defined elsewhere.
Entity Typology	Vocabulary	Model	Foundational	Supports a registry of entities
Principles Typology	Vocabulary	Model	Foundational	A classification of principles and objectives.
Objects/ Services Typology	Vocabulary	Model	Foundational	Classification of objects and services
Metrics Typology	Vocabulary	Model	Foundational	Types of metrics (qualitative, quantitative, and elaboration of these, how repeatable they are)
Test Typology	Vocabulary	Model	Foundational	Types of tests (automated, manual, machine actionable, …), and possibly whether tests are standardised or not
Roles	Vocabulary	Seeded	Assessment	{PID Scheme (Component), PID Authority (Role), PID Service Provider (Role), PID Service (Component), …}
Maturity/ Level of Compliance	Vocabulary	Seeded	Assessment	{Ad-hoc, aware, implementation phase, managed and measured, optimised}
Motivation Typology	Vocabulary	Seeded	Foundational	{Community expectation, mandatory, rules of engagement, digital systems architecture, …}
Criterion Typology	Vocabulary	Seeded	Foundational	{Recommendation, Best Practice, Guideline, …}

8. Interface Methods and Payloads

The Compliance Assessment Toolkit requires a set of methods and payloads, available via API, to ensure that compliance assessments can be encoded, recorded, and retrieved. These methods form the backbone of the toolkit.

The APIs should follow the specifications developed in Section 2 [13] for general API design, and Section 2 provides a detailed inventory of API calls.

The following table summarises the scope of the API calls.

Table 5 – Summary of API Calls

API Scope	GET	POST	PUT	PATCH	DELETE
Application Administration (Users, User Roles, ...)	Admin	Admin	Admin	Admin	Admin
Foundational Model Aspects (Model typologies, ontologies)	All	Admin	Admin	Admin	Admin
Assessment-Related Typologies	All	Admin	Admin	Admin	Admin
Model Foundational Registries	All	Admin	Admin	Admin	Admin
Assessment-Related Registries - Entities, Objects, and Services, as well as Assessments	All	Validated	Validated	Validated	Validated
Motivation-Related Registries - Principles, Criteria, Tests, Metrics	All	Validated	Validated	Validated	Validated
Guidance-Related Registries (Benchmarks, Best Practice, ...)	All	Validated	Validated	Validated	Validated
Private Self-Assessment	Identified	Identified	Identified	Identified	Identified

Section 2: Requirements and Specifications

1. Releases

The following releases are foreseen during the life of the project:

- M18 (Beta Release, TRL6): this release coincides with the general beta release of the services and components being developed by FAIRCORE4EOSC.
- M32 (TRL 8 #1): This is the first production release of the Compliance Assessment Toolkit. It will include improvements identified during the beta testing phase[10], and incorporate any refinements to vocabularies and the Conceptual Model.
- M36 (TRL 8 #2): Ensuring integration with EOSC, improvements identified during production use. This release is optional.

Requirements and specifications in the annexures are defined in terms of technology readiness level in assigned releases, and ranked in terms of priority. This information should be interpreted as follows:

Table 6 – Releases and Expected Readiness Levels

Release	High	Medium	Low
M18 (TRL 4-6)	TRL6 mandatory	TRL6 optional	TRL 4 optional
M32 (TRL 6-8)	TRL8 mandatory	TRL8 optional TRL6 mandatory	TRL6 optional
M36 (TRL 8)	TRL8 mandatory	TRL8 mandatory	TRL8 optional

Note: Interpret the table as follows based on an example: all functionality marked at TRL6 in M18 is mandatory for that release if the priority is high, else it is optional. At M36, all issues and bugs identified in the M32 release must be addressed, and optional features are those that are assessed to be ‘nice to have’.

[10] FAIRCORE4EOSC is discussing a collaboration with FAIR-IMPACT and the EOSC PID Policy Task Force to include the beta release of the CAT into community engagements to test and evaluate usability.

Funded by the European Union

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

2. Actors and Stakeholders

A list of Actors and Stakeholders is presented below. There are two specific groups: the actors involved in or directly referenced by the EOSC PID policy [1], and referred to as 'Entities' in the model, and the actors or stakeholders that may want to make use of the CAT services [0].

Table 7 – Actors and Stakeholders, User Roles

Actor	Description	Reference
PID Standards Body	A PID Standardisation organisation (IETF, IANA, The DONA Foundation, ISO) which appoints the PID Authority and is responsible for the PID Scheme.	[0]
PID Scheme (Component)	A set of rules and standards defining the nature of a PID. This would include a set of lexical formatting rules for PIDs within a namespace. It could also define for example: associated PID Type; definition of associated metadata; quality assurance conditions; usage rights, terms and conditions, and algorithmic methods for generating PID names and enforcing PID properties.	[1]
PID Authority (Role)	A controller responsible for maintaining the rules for defining the integrity of PIDs within a PID Scheme. These rules may include setting standards for lexical formats, algorithms, and protocols to ensure global uniqueness, together with setting quality of service conditions to enforce compliance to the rules. PID Authorities may be organisations (e.g., DOI.org), which enforce control over a PID infrastructure. A Persistent Identifier (PID) policy for the European Open Science Cloud (EOSC) 8 may also be Authorities which do not have a central control (for example Software Heritage persistent identifiers1 and W3C's Decentralised Identifiers), but provide a community standardisation mechanism that specifies the conformance of PIDs to a PID Scheme.	[1]
PID Service Provider (Role)	An organisation which provides PID services in conformance to a PID Scheme, subject to its PID Authority. PID Service Providers have responsibility for the provision, integrity, reliability, and scalability of PID Services, in particular the issuing and resolution of PIDs, but also lookup and search services, and interoperability with a generic resolution system.	[1]
Multi-Primary Administrator (MPA)(Role)	"Each credentialed MPA operates its own GHR Services in accordance with the DONA Foundation Policies & Procedures for the GHR and coordinates its GHR Services with other MPAs and DONA in the distributed operation of the GHR on a multi-primary basis." (DOI and Handle Only)	[7]
PID Service (Component)	Basic services are those that create, manage, and resolve PIDs and their associated kernel information which conforms to a PID Scheme. Advanced, value-added services may also be provided, for example attribute search or metrics.	[1]

PID Manager (Role)	PID Managers have responsibilities to maintain the integrity of the relationship between entities and their PIDs, in conformance to a PID Scheme defined by a PID Authority. A PID Manager will typically subscribe to PID services to offer functionality to PID Owners within the PID Manager's services. One example is a Service Provider which uses PID Services as part of its own service delivery. For example, PID Managers may include a provider of a data repository, a data catalogue, or a research workflow system.	[1]
PID Owner (Role)	An actor (an organisation or individual) who has the authority to create a PID, assign PID to an entity, provide and maintain accurate Kernel Information for the PID. A new PID Owner must be identified, and these responsibilities transferred, if the current PID Owner is no longer able to carry them out.	[1]
End User (Role)	The end user of PID Services, for example researchers, or software, or services produced to support researchers.	[1]
Compliance Monitoring (Role)	On completion, the work will support an additional role and associated component for the EOSC PID Policy, as follows: Compliance Monitoring (Role) - One or more organisations that provide services to monitor and/ or enforce compliance (with PID Policy), resulting in interoperable and aggregable compliance metrics for the roles and components foreseen in the policy.	[0], [1]
Casual User/ General Public	Users that engage with the system to search for and find existing evaluation records, statistics, and guidance in respect of compliance.	[0], [4]
Identified User	Users may want to specify preferences and save context, and to do so, users need to be identified unambiguously. No other information is needed except some globally unique identifier.	[0], [4]
Validated User	Users that contributed external evaluations and assessments and may want to contribute a self-assessment. These users will often want to obtain additional guidance and best practices for the elements where their level of performance is lower than the benchmarks. These users can have any of the detailed roles identified in 01 t0 08 above.	[0], [4]
Admin User	Administrative users that manage vocabularies, users, and databases associated with the CAT.	[0], [4]

3. Detailed Requirements and Specifications

These are intended as stand-alone documents in practice, and are collected in the Annexures. The scope is as follows:

Table 8 – Scope of Documentation

Aspect	Annexure	Description
Component Information	Annexure A	Required by FAIRCORE4EOSC as administrative metadata.
User Stories	Annexure B.1	Examples of how the service or component will be applied by end users and systems.
User Requirements	Annexure B.2	User requirements - describing the main use cases to be accommodate by the system

References

No	Description/Link
R0	Developed in this document
R1	European Commission, Directorate-General for Research and Innovation, Hellström, M., Heughebaert, A., Kotarski, R., et al., A Persistent Identifier (PID) policy for the European Open Science Cloud (EOSC), Publications Office, 2020, https://data.europa.eu/doi/10.2777/926037
R2	FC4E Compliance Toolkit - Landscape Assessment, Internal Milestone Report, December 2022, https://docs.google.com/document/d/1XxiTpqkD6vMOipylwH5PScaXskaYTRzsRP3uw0J1f3E/edit?usp=sharing
R3	Comments and panel discussion during FC4E Technical Meeting, Kajaani, 18 April 2023, https://docs.google.com/document/d/19NmhdTEvqedQWx299AipmKLyv8G-77xF/edit#bookmark=id.3m9px47smzq7
R4	European Commission, Directorate-General Research and Innovation, Wierenga, K., Johansson, L., Kanellopoulos, C., et al., EOSC Authentication and Authorization Infrastructure (AAI) : report from the EOSC Executive Board Working Group (WG) Architecture AAI Task Force (TF), Publications Office, 2021, https://data.europa.eu/doi/10.2777/8702
R5	Collaboration proposal, FAIRCORE4EOSC, FAIR-IMPACT, and EOSC PID Policy TF, https://docs.google.com/presentation/d/1tNEb47Xq-M9Fs87u1oYSRuSb9c6bCXzSlM_6jghdyv4/edit#slide=id.g22cc66b4c73_0_103
R6	Roberto Di Cosmo, Morane Gruenpeter, Stefano Zacchiroli. Identifiers for Digital Objects: the Case of Software Source Code Preservation. iPRES 2018 - 15th International Conference on Digital Preservation, Sep 2018, Boston, United States. pp.1-9, ff10.17605/OSF.IO/KDE56ff Ffhal-01865790v4f
R7	The DONA Foundation, https://www.dona.net/handle-system/
R8	RDA/ WDS Certification of Digital Repositories IG, https://www.rd-alliance.org/groups/rdawds-certification-digital-repositories-ig.html
R9	CODATA Coordinated Expert Group, Berkman, Paul Arthur, Brase, Jan, Hartshorn, Richard, Hodson, Simon, Hugo, Wim, Leonelli, Sabina, Mons, Barend, Pergl, Hana, & Pfeiffenberger, Hans. (2020). Open Science for a Global Transformation: CODATA coordinated submission to the UNESCO Open Science Consultation. Zenodo. https://doi.org/10.5281/zenodo.3935461
R10	World Data System Membership Categories, https://worlddatasystem.org/members/
R11	European Commission, Directorate-General for Research and Innovation, EOSC rules of participation, Publications Office, 2021, https://data.europa.eu/doi/10.2777/30541
R12	Conceptual Model Validation case studies, FAIRCORE4EOSC Working Documentation, https://docs.google.com/spreadsheets/d/12BV6jL0tpq1UQlWiG5dnTsTeYvINIwI5r2GIhT-X_48/edit#gid=987836026&range=A1
R13	Internal Specification, API Design Practices, https://docs.google.com/document/d/1QA6xun4R5tOJTueVpBtJMznHrwc3h1TWUer1aRtWEnk/edit?usp=sharing
R14	IETF, FRC 2119: Key words for use in RFCs to Indicate Requirement Levels, https://www.rfc-editor.org/rfc/rfc2119.html
R15	Herbert Van de Sompel, Michael L. Nelson (2015). Reminiscing About 15 Years of Interoperability Efforts, DOI: 10.1045/november2015-vandesompel
R16	EU, General Data Protection Regulation (GDPR), https://gdpr.eu/tag/gdpr/

R17	Oscar Corcho, Fajar J. Ekaputra, Ivan Heibi, Clement Jonquet, Andras Micsik, Silvio Peroni and Emanuele Storti, 2023, A maturity model for catalogues of semantic artefacts, https://doi.org/10.48550/arXiv.2305.06746
R18	de Castro, Pablo, Herb, Ulrich, Rothfritz, Laura, & Schöpfel, Joachim. (2023). Building the plane as we fly it: the promise of Persistent Identifiers. Zenodo. https://doi.org/10.5281/zenodo.7258286
R19	de Castro, Pablo, Herb, Ulrich, Rothfritz, Laura, & Schöpfel, Joachim. (2023). Failed PIDs and unreliable PID implementations, Zenodo. https://doi.org/10.5281/zenodo.7330527
R20	Sansone, Susanna-Assunta, McQuilton, Peter, Cousijn, Helena, Cannon, Matthew, Chan, Wei Mun, Callaghan, Sarah, Carnevale, Ilaria, Cranston, Imogen, Edmunds, Scott, Everitt, Nicholas, Ganley, Emma, Graf, Chris, Hrynaszkiewicz, Iain, Khodiyar, Varsha, Leary, Adam, Lemberger, Thomas, MacCallum, Catriona, McNeice, Kiera, Murray, Hollydawn, … Threlfall, Jonathan. (2020). Data Repository Selection: Criteria That Matter. Zenodo. https://doi.org/10.5281/zenodo.4084763
R21	Lin, D., Crabtree, J., Dillo, I. et al. The TRUST Principles for digital repositories. Sci Data 7, 144 (2020). https://doi.org/10.1038/s41597-020-0486-7
R22	CoreTrustSeal Standards and Certification Board. (2022). CoreTrustSeal Requirements 2023-2025 (V01.00). Zenodo. https://doi.org/10.5281/zenodo.7051012
R23	DIN 31644, Information und Dokumentation - Kriterien für vertrauenswürdige digitale Langzeitarchive.
R24	ISO 16363, Which repositories are TRUSTWORTHY? https://www.iso.org/standard/56510.html
R25	Wilkinson, M., Dumontier, M., Aalbersberg, I. et al. The FAIR Guiding Principles for scientific data management and stewardship. Sci Data 3, 160018 (2016). https://doi.org/10.1038/sdata.2016.18
R26	Wikipedia, FAIR Data.
R27	Signposting the Scholarly Web, https://signposting.org/
R28	FAIR Digital Objects Framework, https://www.go-fair.org/today/fair-digital-framework/
R29	Schwardmann, U., 2020. Digital Objects – FAIR Digital Objects: Which Services Are Required?. Data Science Journal, 19(1), p.15. DOI: http://doi.org/10.5334/dsj-2020-015
R30	F-UJI Tool, https://www.f-uji.net/
R31	Anusuriya Devaraju, & Robert Huber. (2020). F-UJI - An Automated FAIR Data Assessment Tool (v1.0.0). Zenodo. https://doi.org/10.5281/zenodo.4063720
R32	Hugo, Wim, Le Franc, Yann, Coen, Gerard, Parland-von Essen, Jessica, & Bonino, Luiz. (2020). D2.5 FAIR Semantics Recommendations Second Iteration (1.0). Zenodo. https://doi.org/10.5281/zenodo.5362010
R33	FAIR Data Maturity Model Working Group. (2020). FAIR Data Maturity Model. Specification and Guidelines (1.0). https://doi.org/10.15497/rda00050
R34	FAIR EVA, https://www.faireva.org/
R35	Your first step towards your FAIR data(set), https://fairaware.dans.knaw.nl/
R36	CARE Principles for Indigenous Data Governance, https://www.gida-global.org/care
R37	De Vos, M., Kirrane, S., Padget, J., Satoh, K. (2019). ODRL Policy Modelling and Compliance Checking. In: Fodor, P., Montali, M., Calvanese, D., Roman, D. (eds) Rules and Reasoning. RuleML+RR 2019. Lecture Notes in Computer Science(), vol 11784. Springer, Cham. https://doi.org/10.1007/978-3-030-31095-0_3
R38	Esteves, Beatriz, Pandit, Harshvardhan J., & Rodríguez-Doncel, Victor. (2021, July 17). ODRL Profile for Expressing Consent through Granular Access Control Policies in Solid. International Workshop on Consent Management in Online Services, Networks and Things (COnSeNT 2021), Virtual. https://doi.org/10.5281/zenodo.5111540
R39	The SOLID Project, https://solidproject.org/

R40	EOSC Task Force on Rules of Participation, https://www.eosc.eu/advisory-groups/rules-participation-compliance-monitoring
R41	EOSC PID Policy and Implementation Task Force, https://www.eosc.eu/advisory-groups/pid-policy-implementation
R42	European Commission, Directorate-General for Research and Innovation, Hellström, M., Heughebaert, A., Kotarski, R., et al., A Persistent Identifier (PID) policy for the European Open Science Cloud (EOSC), Publications Office, 2020, https://data.europa.eu/doi/10.2777/926037
R43	FAIR Data Spaces in NFDI, https://www.nfdi.de/fair-data-spaces/?lang=en
R44	FAIR Data Spaces Specifications, https://websites.fraunhofer.de/fair-ds/existing-specifications/
R45	FAIR Data Spaces - EOSC Pillar, https://www.eosc-pillar.eu/federated-fair-data-space-f2ds
R46	Data Commons - https://docs.datacommons.org/
R47	FAIR Implementation Profiles, https://www.go-fair.org/how-to-go-fair/fair-implementation-profile/
R48	World Data System Membership Application, https://worlddatasystem.org/members/application_membership/
R49	[13] M. Bar-Sinai, L. Sweeney and M. Crosas, "DataTags, Data Handling Policy Spaces and the Tags Language," 2016 IEEE Security and Privacy Workshops (SPW), 2016, pp. 1-8, doi: 10.1109/SPW.2016.11.
R50	FAIR Data Point, https://www.fairdatapoint.org/
R51	Hugo, Wim. (2013). A Maturity Model for Digital Data Centers. Data Science Journal. 12. WDS189-WDS192. https://doi.org/10.2481/dsj.WDS-032.
R52	Soiland-Reyes, Stian & Sefton, Peter & Crosas, Merce & Castro, Leyla & Coppens, Frederik & Fernández, José & Garijo, Daniel & Grüning, Björn & Rosa, Marco & Leo, Simone & Carragáin, Eoghan & Portier, Marc & Trisovic, Ana & Community, RO-Crate & Groth, Paul & Goble, Carole. (2022). Packaging research artefacts with RO-Crate. Data Science. 5. 1-42. 10.3233/DS-210053
R53	Research Objects, https://www.researchobject.org/
R54	RDA Webinar: 10 Things for Curating FAIR and Reproducible Research
R55	Peer, Limor, Arguillas, Florio, Honeyman, Tom, Miljković, Nadica, Peters-von Gehlen, Karsten, & CURE-FAIR subgroup 3. (2021). Challenges of Curating for Reproducible and FAIR Research Output (2.1). https://doi.org/10.15497/RDA00063
R56	GDPR: https://www.consilium.europa.eu/en/policies/data-protection/data-protection-regulation/
R57	Principles of Open Scholarly Infrastructure, https://openscholarlyinfrastructure.org/
R58	RDA TIGER Open Call, https://www.rd-alliance.org/rda-tiger-launches-open-call-rda-working-group-facilitation-support-0

R59	Paulk, Mark; Curtis, William; Chrissis, Mary Beth; & Weber, Charles. Capability Maturity Model for Software (Version 1.1). CMU/SEI-93-TR-024. Software Engineering Institute, Carnegie Mellon University. 1993. http://resources.sei.cmu.edu/library/asset-view.cfm?AssetID=11955
R60	Hugo, W and Beck, J. Deliverable number: 5.2 - Guidance on Data Policy and Supporting Open Licences - Data Policy Considerations, Project: 730995 - Supporting EU-African Cooperation on Research Infrastructures for Food Security and Greenhouse Gas Observations (SEACRIFOG), 2019, http://dx.doi.org/10.13140/RG.2.2.11940.04485
R61	FAIRCORE4EOSC WP5: Internal PID System Review, https://docs.google.com/document/d/1oD75aCkKp7LB-9qD7WZ_Ib83lVdOwNYTH2ClDdVfBOY/edit?usp=sharing
R62	W Hugo, M Buys: Compliance Case Study Analysis: POSI, 2023, http://doi.org/10.5281/zenodo.7933879
R63	Sanderson, R., Phillips, M., van de Sompel, H. Analysing the Persistence of Referenced Web Resources with Memento, Open Repositories 2011 Conference, 2011, https://doi.org/10.48550/arXiv.1105.3459
R64	Klump, J. and Huber, R., 2017. 20 Years of Persistent Identifiers – Which Systems are Here to Stay?. Data Science Journal, 16, p.9., http://doi.org/10.5334/dsj-2017-009

Annexure A: Component Information

Component	CAT - Compliance Assessment Toolkit
Category	Operational EOSC Services
Contact person	Wim Hugo (DANS) Themis Zamani (GRNET)
Email address	wim.hugo@dans.knaw.nl
Contributors	W Hugo, W Steinhoff (DANS), T Zamani (GRNET), M Buys (DataCite), S Bingert (GWDG), Maarten Hoogerwerf (SURF)
Version	0.1
Date	31-05-2023

Overview
The FAIRCORE4EOSC Compliance Assessment Toolkit (CAT) will assist actors in the PID ecosystem with assessment of their compliance with the policy. The toolkit is by design capable of accommodating a wider variety of compliance assessment use cases, as detailed in Section 1, but this section deals with PID Policy Compliance Assessment only.

Objectives	
#	Short description
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.

Out of Scope/ Important Dependencies	
#	Short description
1	The conceptual model vocabularies used for development of the CAT will be, in many cases, proposals based on analysis of the use cases, but should and will be revised by the community (primarily FAIR-IMPACT and the EOSC PID Policy Task Force) with a view to finalisation by M32 of the project.
2	The implementation described below allows self-assessment to be conducted by identified users, and by design allows end users to import assessment results from external tools, but the FC4E CAT project will not be developing such tools.
3	The FAIRCORE4EOSC project will collaborate with FAIR-IMPACT and the EOSC PID Policy Task Force to develop and encode guidance and best practices for use in the CAT, but the completeness and scope of such guidance depends on activities outside the control of FAIRCORE4EOSC.
4	Any operational product or service requires more than technical implementation: contracts and legal documentation, privacy statements, marketing and guidance materials, governance and community engagement mechanisms, and so on. These aspects will be addressed in Task 2.3.

Annexure B: Requirements

Red italics: links to or has dependencies on other FAIRCORE4EOSC components that have been identified but for which details are not yet known or published.

Priority: H=High, M=Medium, L=Low, interpreted as explained in the section on Releases.

B.1 User Stories

User stories		
#	Description of the user story	Reference
US-1	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).	[1], [5]
US-2	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.	[1]
US-3	Institutional users of persistent identifiers want to obtain guidance on best practices and recommendations in terms of implementation. The guidance applies in two categories: ● how to use or provide actionable persistent identifiers best, depending on the use case, ● and how to develop policy in respect of identifiers.	[5]
US-4	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 approach proposed there to accommodate PID Policy compliance assessment.	[2]

B.2 User Requirements

User Requirements				
#	Short description	Priority	Feasibility	Reference
UR-1	A user registers with the intent of contributing information about its own role(s) in the PID Ecosystem. Such requests are first identified using AAI, and then validated and confirmed by an administrator. The user selects the actor type from a list of options.	H (M18)	4	US-2
UR-2	A user wants to register with the intent of preserving preferences in the system. Such registrations need only identification via AAI, no validation required.	H (M18)	5	US-1 US-2 US-3
UR-3	Any registered user can select and specify their own preferences and save contextual information.	M (M 32)	4	US-1 US-2 US-3
UR-4	All users, irrespective of registration, have context preserved in cookies. Permission for cookie use must be confirmed by the user.	M (M 18) M (M 32)	5 5	US-1 US-2 US-3
UR-5	Administrators are enabled to record self-assessments, or import assessment test results, on behalf of actors identified by the project as important and foundational.	H (M 18) H (M 32)	5 5	US-2
UR-6	Specific actors (service providers, PID managers) are able to register and be validated by admin users. Once validated, these users can record self-assessments (Type 1). Validated users can also provide links to external test results as a basis for assessment (Type2).	H (M 18) H (M 32)	5 3	US-2
UR-7	All users have access to dashboard views that summarise, in a non-identifiable way, the scope of assessments and the current state of PID compliance across the ecosystem.	H (M 18)	5	US-1 US-2 US-3
UR-8	Dashboard views serve as representations of search and discovery facets, and can be used to filter assessments for review by an end user.	M (M 18) H (M 32)	5	US-1 US-2 US-3

UR-9	End users contributing assessments (of either type) can elect to keep the assessment results private. In principle, actors that claim compliance with EOSC PID policy publicly, will be required to publish an assessment as evidence.	H (M 18)	5	US-2
UR-10	In all dashboard views, private information can be aggregated into depersonalised results provided it is not possible to identify the actor from the result. If it is possible, such data should be excluded.	M (M 18) H (M32)	4 4	US-2
UR-11	All users have access to a search and discovery interface that lists assessment summaries on the basis of facets. The facets include the actor type, allowing a view of the subset of specific schemes, authorities, services, managers, and so on. Only public and user-owned assessments are listed.	H (M 18)	5	US-1 US-2 US-3 US-4
UR-12	Summary lists of assessments link to detailed assessment views for the following cases: 1. Public assessments 2. Assessments owned by the user The detailed assessment views also provide contextual information on guidance and best practice applicable to the assessment criteria.	H (M 18)	5	US-1 US-2 US-3 US-4
UR-13	Guidance and best practice information is available to the system on the basis of shared criteria and use case keys. Any user can access the knowledge base of guidance and best practice, and when viewing a specific assessment, applicable guidance and best practices are available to the user.	M (M 18) H (M32)	4	US-1 US-2 US-3 US-4
UR-14	Administrative users are able to register and maintain vocabularies and registries, or to maintain references to external vocabularies and registries.	M (M18) H (M32)	4	US-4
UR-15	Publication of a manually captured assessment results in a citable version of the assessment, which is recorded in a suitable repository and provided with a PID. Automated assessments that are linked to the CAT are only published if a PID does not yet exist.	L (M18) H (M32)	5	US-2

B.3 Functional Requirements

Functional Requirements			
#	Short description	Priority	Reference
1	Use cases that enable the creation, management, and referencing of the vocabularies and registries required by the Compliance Assessment Toolkit *Shared with MSCR vocabulary service*	See below	See details below
1.1	Create or add a reference to a vocabulary or registry, manage or deprecate the entry at a future time.	M (M18)	JIRA UR-14
	Map a vocabulary or registry to an externally maintained semantic service. *PID Resolver may list some sources - to be confirmed.*	M (M23)	JIRA UR-14
1.2	Support for the listing and querying of available vocabularies are registries, and the properties of these. List the available vocabularies, elements of a vocabulary, and the details of a specific vocabulary element.	H (M18)	JIRA UR-14
1.3	Create and edit/ update vocabulary items For a given vocabulary, add an element (label or term).	M (M18)	JIRA UR-14
	For a given vocabulary, or add/ update relations between vocabulary elements, or register a new relation type.	L (M32)	JIRA UR-14
2	Add and manage compliance assessments via API	See below	See details below
2.1	Add compliance assessments via bulk population of the graph database as exemplars 1. Selected Schemes 2. Selected Authorities 3. Selected MPAs 4. Selected Service Providers *Link to PID MetaResolver: Add to and supplement data already available from the PID MetaResolver via API.*	M (M18) H (M32)	JIRA UR-6 UR-9 UR-5
2.2	Add a compliance assessment - manually - for the following actors: 1. Manager	H (M18)	JIRA UR-6 UR-9

2.3	Add compliance assessment manually for all of the above, as well as 1. PID Owners	M (M32)	JIRA UR-6 UR-9
2.4	Ingest a remotely performed assessment for any of the above	L (M32)	JIRA UR-6 UR-9
3	Administration and Maintenance API: Register, characterise, and manage users and their specific roles in the CAT. *Re-use any available framework API at GRNET to manage user types and roles.*		JIRA
3.1	Placeholder user accounts for each defined role. Add, manage, and define user roles	H (M18)	JIRA UR-1
3.2	User registers as an 'Identified' user -role is assigned automatically, no further action required.	H (M23)	JIRA UR-1 UR-2
3.3	User registers as a 'validated' user - admin needs to validate and assign a role.	H (M23)	JIRA UR-1 UR-2
4	Query the CAT database in respect of compliance status of actors.	See below	C.1-01
4.1	Issue SPARQL Query to the graph database and receive a return. Query returns are handled differently depending on licence: 4.1.1 Private assessments: aggregate data only 4.1.2 Public assessments: assessment detail is visible.	H (M18)	JIRA UR-7 C.2-02
4.2	Format and transform SPARQL to HTML based on React Template	H (M32)	JIRA UR-7 C.2-03
5	Administrator UI Cases	See Below	JIRA
5.1	Administrator uses the CSC interface to manage vocabularies Administrator opens a dashboard-like inventory of vocabularies, and can list, add, remove, and test vocabulary links from the interface	H (M18) L (M32)	JIRA UR-14 JIRA UR-14
5.2	Administrator opens a dashboard-like view of users, able to edit, add, and remove users.	H (M32)	JIRA UR-1 UR-2
5.3	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	H (M32)	JIRA UR-1 UR-2

6	Validated User - Assessment Use Cases	See below	JIRA
6.1	Add, edit, and delete registry information for the user - provide 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. (see (A validated user may be the owner of multiple services or objects. This situation is handled in case 7 below.)	M (M18) H (M32)	JIRA UR-3
6.2	A validated user requests or registers a new manual external assessment. This assessment has the following steps and sub-components: 1. Provide or confirm information about the submitter (=logged-in user account) 2. Select the type of assessment (Scheme, Authority, Provider, Manager, …) 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.	H (M32)	JIRA UR-6 UR-9
6.3	A validated user requests or registers a new manual assessment. This assessment has the following steps and sub-components: 1. Provide or confirm information about the submitter (=logged-in user account) 2. Select the type of assessment (Scheme, Authority, Provider, Manager, …) 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 by filling in the evaluation form. Support is provided by making benchmarking, guidelines, and best practices information available.	H (M18) M (M32)	JIRA UR-6 UR-9
7	Use Cases - Validated User Dashboard. This is used to add, edit, and delete registry elements for specific actors. User navigates to a dashboard list associated with the account - at a minimum from ‘My Profile’ or similar.	See below	JIRA UR-3
7.1	Default view or user selects ‘Assessment Status’. Displays summary charts in respect of each actor category associated with the account. *Note: uses the same query as the home-page summary dashboard, filtered for account.*	H (M18)	JIRA UR-3

7.2	User selects “Account Information”. Displays information captured at registration, and allows modification, including password reset. Issue a token for API access. 1. Validated account: R/W 2. Identified account: Read-Only	M (M32) L (M32)	JIRA UR-3
7.3	User selects “My Preferences”. User is able to modify 1. preferences (style, default views, etc. if any), 2. and saved searches.	M (M32)	JIRA UR-3
7.4	User selects ‘My Objects and Services’. The dashboard lists all services or objects associated with the account with assessment status. Subcases include: 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)	H (M18)	JIRA UR-6 UR-9
8	User selects the Assessment Status Dashboard. This is also the default home page view.	H (M18)	
8.1	Summary View. Clicking a facet in the summary opens a filtered listing (8.2).	H (M18)	JIRA UR-7 C.1-06
8.2	Listing/ Cards - all assessments matching the filter specification. Clicking a list entry opens a detailed view.	H (M18)	JIRA UR-7 C.1-07
8.3	Detailed View (equivalent to 9.3)	M (M18)	JIRA
9	Search and Discovery Dashboard - 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	H (M18)	UR-7 UR-8 UR-10 UR-11 UR-12 C.1-07
9.1	Dashboard: similar to 8.1		JIRA
9.2	Facets and Listing per Actor: the index is set to facets that can filter and characterise actors, and actor-related information is displayed when a detailed record is selected.		JIRA UR-7 UR-8 UR-10 UR-11 UR-12 C.1-09
9.3	Detailed Object or Service View: the index is set to facets and properties applicable to services or objects, and service-		JIRA UR-7 UR-8

	related information is displayed when a detailed record is selected.		UR-10 UR-11 UR-12 C.1-09
10	Guidance dashboard - List, query, and explore guidance.		JIRA
10.1	Select a portfolio of use cases, domains, and ecosystem roles from a checklist (facets) Open a list of applicable guidance		JIRA UR-13 C.1-08
10.2	List guidance elements meeting filter specifications, and show details when a record is selected		JIRA UR-13 C.1-08

B.4 Non-Functional Requirements

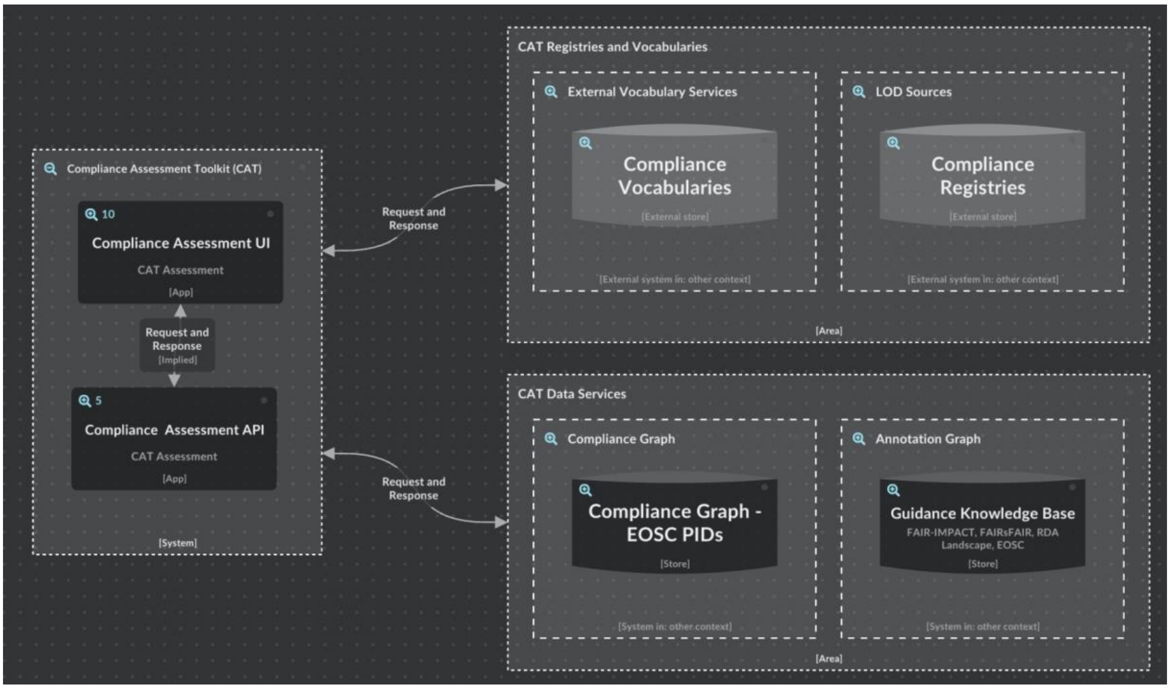
Non-functional requirements			
#	Short description	Priority	Reference
NF-1	The CAT SHOULD use existing APIs and services for registration of vocabularies and graph data whenever possible, and if existing instances are not available, it MUST base internal instances on a generally accepted specification for such services in published examples. *Example: Graph database compliant with OpenAIRE work on the FC4E Research Graph or DataCite work on the PID Graph.* *Example: Vocabulary API compatible with or re-using CSC Vocabulary Server* *Example: Type definitions compatible with the FC4E DTR*	H (M18)	[5]
NF-2	The CAT SHOULD defer all user communication to a ticket system, and integrate required information to the CAT UI via ticket system API. The ticket system should provide the minimum functionality specified in C.1.	M (M18) H (M32)	[0]
NF-3	Assessment Event: each recording of or change to an assessment introduces an assessment event, and these events need to be versioned and be persistently identifiable. A specification is provided for creation of the PID and the data model underlying the PIDs. See NF-7 below.	M (M18) H (M32)	[0] NF-7 C.2-07
NF-4	The CAT MUST implement the minimum status code specification for HTTP requests for all APIs.	L (M18) H (M32)	[13]

NF-5	The CAT MUST write significant events to a log and write critical events to the ticket system.	L (M18) H (M32)	[0] C.2-08
NF-6	The CAT SHOULD write and update a cookie on significant events and user choices, preserving state	L (M18) H (M32)	UR-4 C.3-01
NF-7	A self-assessment is published formally, is citable, and requires a PID. 1. New assessment of a service or object: a. Compile a metadata record of the assessment b. Publish to a suitable repository, Zenodo is proposed c. Register PID in the CAT 2. Updated assessment of the same service or object: a. Update the metadata record of the assessment b. Add previous PID to metadata c. Publish to a suitable repository, Zenodo is proposed d. Register PID in the CAT e. Link to previous assessment	L (M18) H (M32) M (M32) H (M36)	UR-15 C.2-07

Annexure C: Specifications

Architectural design

C2 (C: Applications and Services

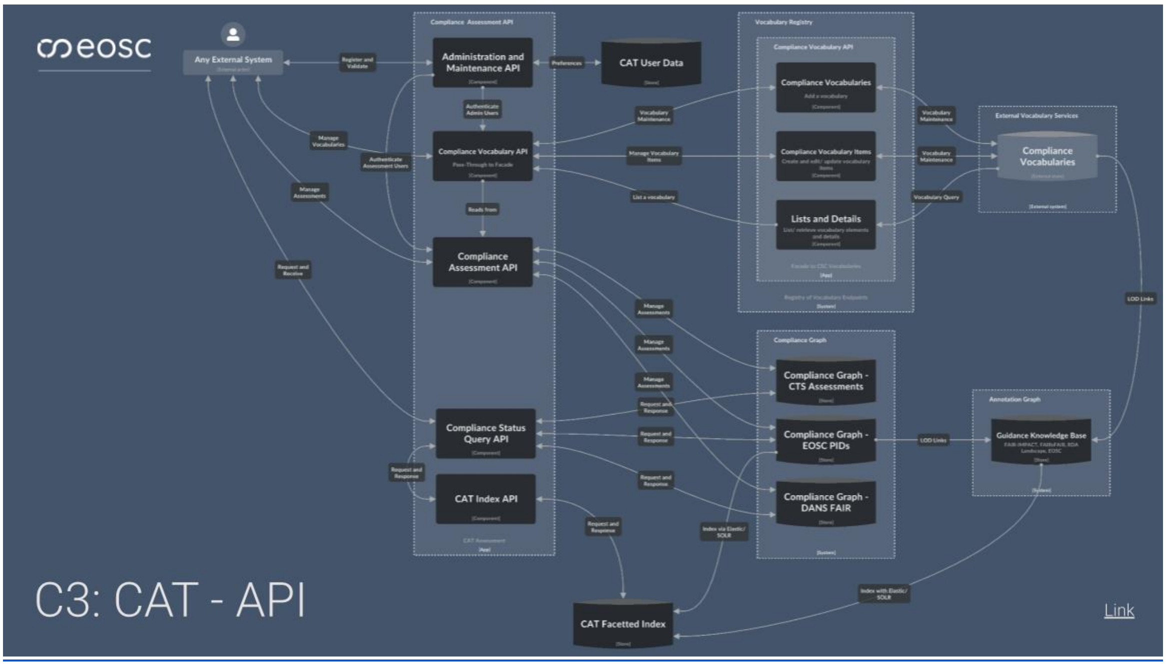


The Compliance Assessment Toolkit consists of four interdependent applications and services:

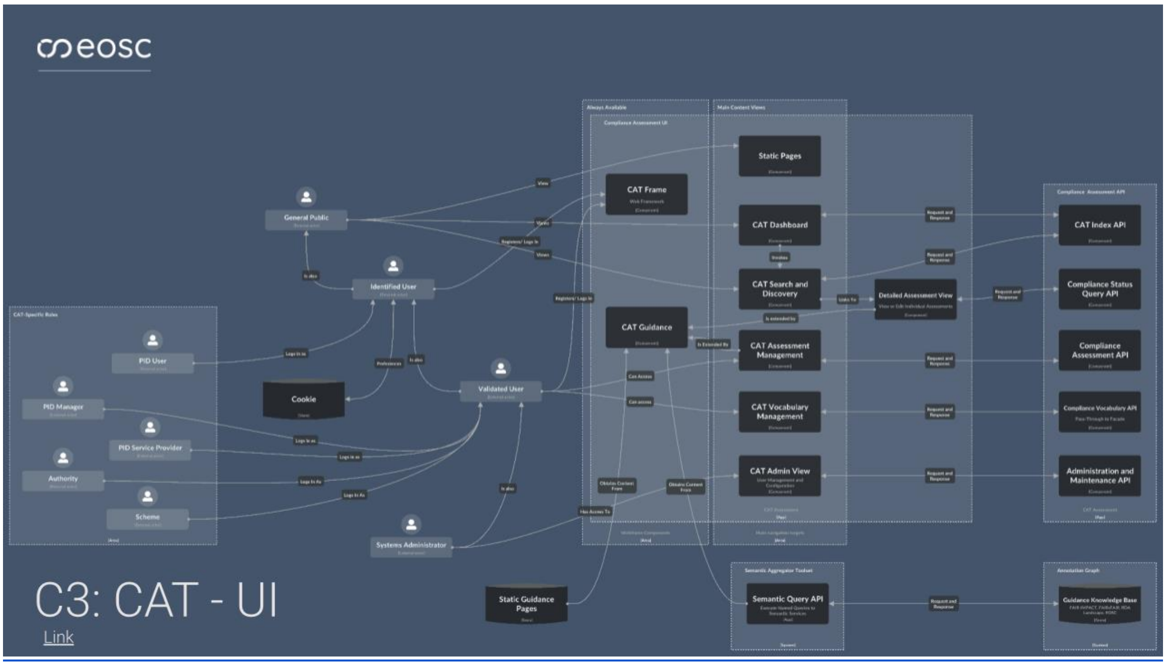
1. **A vocabulary and registry service**. This is a facade for a number of contributing sources of vocabulary and registry entries that are distributed in the ecosystem. Some vocabulary is owned by the CAT and will be maintained in a vocabulary server nominated by FAIRCORE4EOSC.
2. **Data services**. These services are hosted on infrastructure provided for the CAT. It requires two logically distinct, but interrelated datasets:
    a. Compliance assessment data, stored in a graph database that uses vocabulary and registry PIDs as relational nodes for the data, and
    b. Guidance data, linked to the vocabulary, that assists with the assessment and evaluation process and to improve maturity.
3. **Compliance Assessment API**. This API provides a single endpoint for all data, registry, and vocabulary services required by the CAT.
4. **Compliance Assessment UI**. This provides a variety of user-focused functionality and features based on the data sources and semantic services.

C3 (Components, Stores)

API Specifications



User Interface Specifications



C.1 Functional Specifications

Functional specifications			
#	Short description	Priority	Reference
C.1-01	Specification: User roles and behaviours mapped to use cases	H (M12)	JIRA K.22.1
C.1-02	Dashboard and management - user account	M (M12)	
C.1-03	Dashboard and Management - administration of user accounts	M (M12)	
C.1-04	Dashboard and Management - vocabularies and registries	L (M12) H (M18)	
C.1-05	Dashboard and Management - Information about a validated user 1. Assessment Summary 2. Account Information 3. Preferences 4. My Objects and Services	H (M12)	
C.1-05	Manual Assessment 1. User interface specification 2. Workflows 3. Configuration	H (M12)	
C.1-06	Dashboard - Assessments 1. Layout specification 2. Configuration specification	H (M12)	
C.1-07	Dashboard - Search and Discovery 1. Listings of assessments 2. Facets and indexing of assessments 3. Detailed view of an assessment 4. Configuration	H (M12)	

C.1-08	Dashboard - Guidelines (Guidance, Recommendations, Best Practices) 1. Listings of guidelines 2. Facets and indexing of guidelines 3. Detailed view of a guideline 4. Configuration	H (M12)	
C.1.09	Dashboard - Actors - Summarises what is known of higher-level, publicly available actors 1. Listing of actors 2. Facets and indexing - actors 3. Detailed view of an Actor (C.1-05 but limited to public data) (*PID MetaResolver Information*) 4. Configuration	M (M12) H (M32)	PID Kernel MetaResolver

C.2 Service Specifications

Service Specifications			
#	Short description	Priority	Reference
C.2-01	Specification: Vocabulary API *Defer to CSC Vocabulary Service Specification*	H (M12)	
C.2-02	Specification: Compliance Status Query API - Request and Response API *Extend FC4E Research Data Graph/ PID Graph Specification*	H (M12)	
C.2-03	Specification: Compliance Status Query API - Response Formatting	H (M12)	YasGUI
C.2-04	Specification: HTTP Status Codes and API Design Principles	H (M12)	[13]
C.2-05	API Calls: these need to be standardised to align with the specifications for the FC4E RDGraph component. *Extend FC4E Research Data Graph Specification*	H (M12)	CAT API RDGraph[11]

[11] Restricted until publication by the European Commission

C.2-06	Assessment Exchange Format: a JSON-LD format for exchange of an assessment between tools and instances of the Compliance Assessment Toolkit	M (M12) H (M32)	
C.2-07	Register a compliance assessment in Zenodo and obtain a PID 1. Metadata specification 2. API - Request and response specification	M (M12) H (M32)	Zenodo API
C.2-08	Registering significant events and errors in the log and ticket system.	L (M12) H (M36)	

C.3 Operational Specifications

Operational specifications			
#	Short description	Priority	Reference
C.3-01	State information to be stored in a cookie on the user’s machine. This specification has an impact on the privacy statement required for operational deployment.	L (M18) H (M32)	

C.4 Integration with EOSC Core Components

Integration with EOSC Core components			
#	Short description	Priority	Reference
1	Authentication	H (M32)	
2	Help Desk	H (M32)	
3	Monitoring	M (M32)	

Annexure D: Supplementary Information

D.1 Assessment of Case Studies - Detail

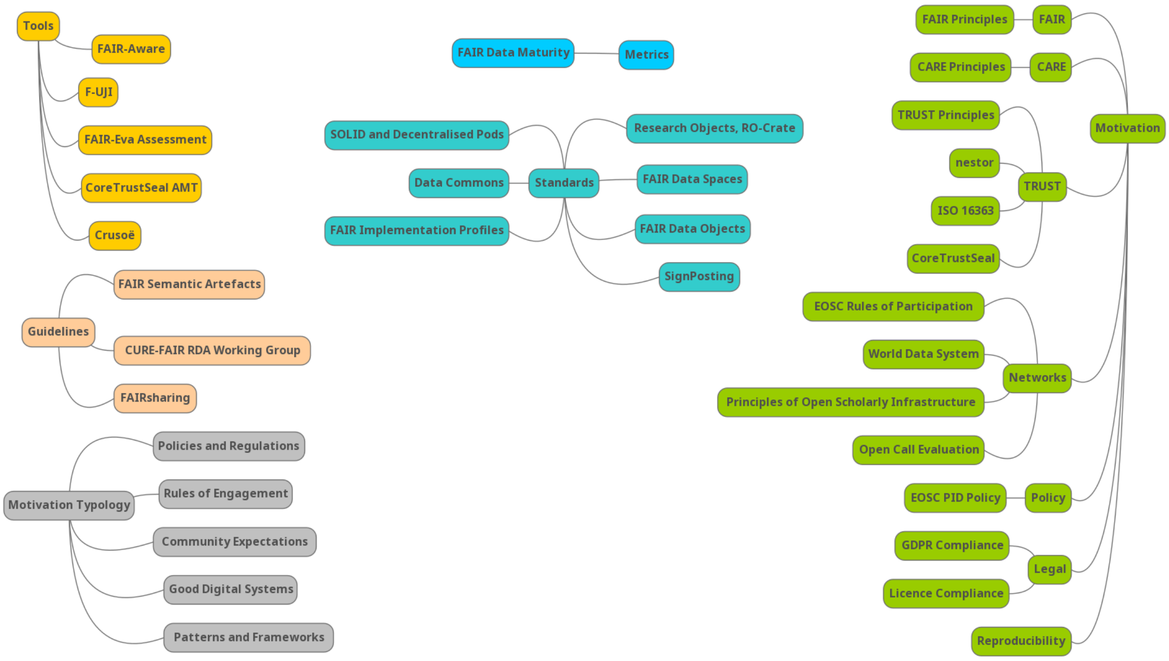
Motivation Type	Motivation (Use Case)	Description	Reference
Community Expectation	nestor	German National Standard DIN 31644, used for trustworthy long-term repositories	[23]
Community Expectation	ISO 16363	ISO 16363:2012 defines a recommended practice for assessing the trustworthiness of digital repositories. It is applicable to the entire range of digital repositories. ISO 16363:2012 can be used as a basis for certification.	[24]
Community Expectation	FAIR Principles	“FAIR data are data which meet principles of findability, accessibility, interoperability, and reusability. The acronym and principles were defined in a March 2016 paper in the journal Scientific Data by a consortium of scientists and organisations. The FAIR principles emphasise machine-actionability.”	[25], [26]
Community Expectation	FAIR Data Objects	FAIR Digital Objects (FDO) provide a conceptual and implementation framework to develop scalable cross-disciplinary capabilities, deal with the increasing data volumes and their inherent complexity, build tools that help to increase trust in data, create mechanisms to efficiently operate in the domain of scientific assertions, and promote data interoperability.	[28], [29]
Community Expectation	F-UJI	“F-UJI is a web service to programmatically assess FAIRness of research data objects at the dataset level based on the FAIRsFAIR Data Object Assessment Metrics”.	[30], [31]
Community Expectation	FAIR Semantic Artefacts	“17 preliminary recommendations related to one or more of the FAIR principles, and 10 best practice recommendations on semantic artefacts were documented.”	[32]
Community Expectation	SignPosting	“Signposting is an approach to make the scholarly web more friendly to machines. It uses Typed Links as a means to clarify patterns that occur repeatedly in scholarly portals. For resources of any media type, these typed links are provided in HTTP Link headers. For HTML resources, they may additionally be provided in HTML link elements.”	[15], [27]
Community Expectation	FAIR Data Maturity	“To remedy the proliferation of FAIRness measurements based on different interpretations of the principles, an RDA Working Group … established a set of indicators and maturity levels for those indicators.”	[33]

Community Expectation	FAIR-Eva Assessment	“Fair EVA is an open source project that is gathering resources and building tools to help researchers and developers, technology activists and voice technology users evaluate and audit bias and discrimination in voice technologies.”	[34]
Community Expectation	FAIR-Aware	“The tool is discipline-agnostic, making it relevant to any scientific field. … The self-assessment consists of 10 questions with additional guidance texts to help you become more aware of what you can do to make your data(set) as FAIR as possible. The assessment will take between 10-30 minutes, after which you will receive an overview of your awareness level and additional tips on how you can further improve your FAIR skills.”	[35]
Community Expectation	CARE	“The current movement toward open data and open science does not fully engage with Indigenous Peoples rights and interests. Existing principles within the open data movement (e.g. FAIR: findable, accessible, interoperable, reusable) primarily focus on characteristics of data that will facilitate increased data sharing among entities while ignoring power differentials and historical contexts. The emphasis on greater data sharing alone creates a tension for Indigenous Peoples who are also asserting greater control over the application and use of Indigenous data and Indigenous Knowledge for collective benefit.”	[36]
Mandatory and Regulatory	Licence Compliance	Various approaches are being developed, using ORDL-compliant encodings of conditions, to be evaluated as automatically as possible against the performance of the requestor.	[37], [38], [49]
Rules of Participation	SOLID and Decentralised Pods	“Solid is a specification that lets people store their data securely in decentralised data stores called Pods. Pods are like secure personal web servers for data. When data is stored in someone's Pod, they control which people and applications can access it.” Evaluating these access requests provides a standardised basis for compliance assessment.	[39]
Rules of Participation	Service Registration	The EOSC Rules of Participation is being refined and developed by the Task Force on Rules of Participation.	[40]
Mandatory and Regulatory	EOSC PID Policy	A policy developed for PID ecosystem in EOSC, supplemented by refinements and extensions from the EOSC Task Force on PID Policy and evaluated against community expectations.	[18], [19], [41], [42]
Rules of Participation	FAIR Data Spaces	“The Federated FAIR Data Space (F2DS) provides tools for both data producers and data consumers contributing to enhance the overall FAIRness of datasets natively dispersed across heterogeneous	[43], [44], [45]

		repositories by realising services for datasets homogenisation, enrichment and onboarding and services for seamless discovery and access.”	
Rules of Participation	Data Commons	“The Data Commons Graph aggregates data from many different data sources into a single database. Data Commons is based on the data model used by schema.org; for more information, see the guide to the data model. The Data Commons API allows developers to programmatically access the data in Data Commons.” In enabling such seamless access, Data Commons are defining at least some compliance criteria for ensuring the semantic interoperability of federated data services.	[46]
Rules of Participation	FAIR Implementation Profiles	“The FAIR Implementation Profiles representing the implementation strategies of various communities can be used as the basis to optimise the reuse of existing FAIR-enabling resources and interoperation within and between domains. Ready-made and well-tested FAIR Implementation Profiles created by trusted communities can find widespread reuse among other communities, and vastly accelerate convergence onto well-informed FAIR implementations.” As such, the FIPs represent an encoding and implicit vocabulary for identification of FAIR criteria and measures.	[47]
Rules of Participation	World Data System	The World Data System provides an interesting perspective: in addition to requiring CTS certification, repositories are required to satisfy a set of additional criteria of membership.	[48]
Rules of Participation	FAIR Data Points	“A FAIR Data Point (sometimes abbreviated to FDP) is the realisation of the vision of a group of authors of the original paper on FAIR on how (meta)data could be presented on the web using existing standards, and without the need of APIs.”	[50]
Community Expectation	Research Objects, RO-Crate	“An RO-Crate is a structured archive of all the items that contributed to a research outcome, including their identifiers, provenance, relations and annotations … RO-Crate simplifies the process of making research outputs FAIR while also enhancing research reproducibility.”	[52], [53]
Community Expectation	CURE-FAIR RDA Working Group	The working group has published a number of resources to assist with reproducibility.	[54], [55]
Mandatory and Regulatory	GDPR Compliance	Across the EU, this is a major consideration, and while exact requirements are slightly different between countries, compliance assessment can be standardised to a large degree. Complex array of	[16], [56]

		compliance elements.	
Rules of Participation	Principles of Open Scholarly Infrastructure	POSI offers a set of guidelines by which open scholarly infrastructure organisations and initiatives that support the research community can be run and sustained. Others have since built on the foundation of POSI to discuss and propose how all those that support scholarly communications can use these principles to hold each other accountable.	[57]
Rules of Participation	Open Call Evaluation	An anonymised open call aimed at funding applications meeting a set of defined criteria.	[58]
Rules of Participation	Publisher's requirements for repository selection	Representatives from journals, journal publishers and scholarly communication organisations have come together in the FAIRsharing Community to propose a set of criteria for the identification and selection of those data repositories that accept research data submissions.	[16]
Community Expectation	TRUST Principles	"The TRUST Principles provide a common framework to facilitate discussion and implementation of best practice in digital preservation by all stakeholders."	[21]
Community Expectation	CoreTrustSeal	"The CoreTrustSeal Trustworthy Data Repositories Requirements reflect the characteristics of trustworthy repositories. As such, all Requirements are mandatory and are equally weighted, standalone items. Although some overlap is unavoidable, duplication of evidence sought among Requirements has been kept to a minimum where possible."	[22]

D.2 Assessment of Case Studies - Focus Areas



Annexure Figure 1 – Classification of Case Studies in Terms of Primary Focus Area

From assessment of case studies, we could identify six broad model elements that should be accommodated. Colours correspond to the colours in the figure. We discuss these model elements below with some examples from the case studies:

1. ● **Motivations**: these represent the rationale for development of a set of principles, policy provisions, or objectives. The motivations can be thought of as a hierarchy, and it may well be that some motivations represent more than one higher-level node[12]. There are some noteworthy special cases in the hierarchy:
    a. The TRUST principles were developed *after* a number of approaches to certification of trustworthy repositories had been developed, and serves as an attempt to harmonise the objectives of all of them. To our knowledge, the TRUST principles have never been formally mapped[13] to nestor, ISO 16363, or CoreTrustSeal. As a result, these three certification mechanisms are not shown as children of the TRUST principles but as siblings - this may change if a complete mapping proves possible.
2. ● **Standards:** Several of the case studies could be identified as largely contributing to standardisation of implementation. FAIR Data Objects, Signposting, RO-Crate, and so on provide a set of implementation standards that will contribute to or guarantee alignment with the criteria and objectives expressed in the Motivations. There are some cases (Data Commons, RO-Crate, for example) where a mixture of motivations and objectives or principles and standards are present. Benchmarks and tests are somewhat synonymous with standards..
3. ● **Metrics**: One of the case studies (FAIR Data [33] or Semantics [17] Maturity) is the best representative of a collection of metrics (or indicators), each metric identifies benchmarks or standards whereby the measured value will be contextualised. Metrics point to tests that determine, as

[12] In graph database terms, this is not a major concern.

[13] It may be possible to do so in later phases of this project.

reproducibly as possible, a value to compare to the benchmark or standard. Note that tests can also be standardised - but these standards are distinct from ‘performance’ standards discussed above.

4. ● **Tools**: Several of the case studies represent tools that can be used to determine level of compliance or performance against a benchmark: F-UJI, FAIR-Eva, and FAIR-Aware are examples of such tools for FAIR Principles, and the CoreTrustSeal AMT and Crusoë are tools for CoreTrustSeal compliance assessment. Some tools include benchmarks as well as standardised tests in the toolkit.
5. ● **Guidelines**: Guidelines, recommendations, and best practices are provided by several of the case studies. These include FAIR Semantic Artefacts, CURE-FAIR RDA Working Group, and FAIRsharing - but again, the classification is not absolute: these resources also include objectives and principles, and benchmarks and standards in some cases.
6. ● **Typology**: The motivations can be classified broadly in respect of intent, and some motivations fall into multiple categories. As an example - RO-Crate can also be viewed as a motivation with a set of objectives, in which case it will support ‘Community Expectations’ as well as ‘Patterns and Frameworks’ in the typology.

D.3 Case Study Analyses - Selected Examples

Case study analyses will be published separately in Zenodo as 'Working Papers' and referenced in the deliverable D2.1. Some examples are placed here for convenience.

D.3.1 EOSC PID Policy

D.3.1.1 Principles and Objectives

Annexure Table 1 – EOSC PID Policy Principles and Objectives

#	Principle or Objective	Description
P1	Application	PID application depends on unambiguous ownership, proper maintenance, and unambiguous identification of the entity being referenced.
P2	Secure	PID services for EOSC need to address (a wide variety of) applications (including those) that require secure mechanisms built into the PID Infrastructure.
P3	Ecosystem	An ecosystem of PID Infrastructures is needed to support the wide variety of scientific applications and offers sufficient flexibility (service providers, scheme, attribute set) and capacity.
P4	Levels of Granularity	The PID ecosystem ideally supports multiple levels of granularity and encourages/ fosters links between them.
P5	Type Support	Classes of digital objects may need different attribute sets a PID is resolved to. It is the responsibility of a community of practice to define and document these attribute sets (PID Kernel Information Profiles).
P6	Diversity	PIDs can identify many different entities. These can be born digital (e.g. documents, data, software, services - otherwise known as digital objects - and collections made of them), physical (e.g., people, instruments, artefacts, samples), or conceptual (e.g., organisations, projects, vocabularies).
P7	Services	Services are mature, managed with high availability and uptime, and are capable of integration into research and data infrastructures.
P8	Integrated	Services need to integrate well with European Research Infrastructures, but not at the exclusion of the broader research community.
P9	Resolution	There is a need for a generic, global PID resolution system across all PID systems and service providers.
P10	Governance	PID Service Providers should apply appropriate community governance to ensure that their PID Services and Systems adhere to these policies and are agile and responsive to the needs of research, Open Science and EOSC.

D.3.1.2 Criteria

Annexure Table 2 – EOSC PID Policy: Criteria, Metrics, and Benchmarks

#	Principle or Objective	Suggested Label	Description	Metric	Benchmark
C1	Application	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)	$\Sigma T_{1,n}$	=0 → 0 =1 → 1
C2	Secure	Sensitive Metadata	Sensitive kernel metadata **MAY** require access control and/or encryption of the Kernel Information.	$\Sigma T_{2,n}$	<5 → 0 =5 → 1
C3	Application	Ownership	PID ownership **MUST** be visible to other actors in the ecosystem.	T_3	=0 → 0 =1 → 1
C4	Application	Maintenance	The PID owner **SHOULD** maintain PID attributes.	T_4	=0 → 0 =1 → 1
C5	Application	Update Functionality	The PID manager **MUST** provide the functionality required to maintain PID attributes.	T_5	=0 → 0 =1 → 1
C6	Application	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.	T_6	=0 → 0 =1 → 1
C7	Application	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.	T_7+T_{35}	<2 → 0 =2 → 1
C8	Levels of Granularity	Guidance	The PID Service **SHOULD** publish guidance on the use cases, levels of granularity, and community best practices that are satisfied by their PID services.	T_8	=0 → 0 =1 → 1
C9	Ecosystem	Community Engagement	The PID Service **SHOULD** engage the end user community to determine changes in needs and practices and adjust their services and guidance accordingly.	$\Sigma T_{9,n}$	=0 → 0 >1 → 1
C10	Application	Versioning - Schema	PID services **SHOULD** support versioning.	T_{10}	=0 → 0 =1 → 1
C11	Application	Versioning - Procedure	PID services and PID Managers **SHOULD** have clear versioning policies.	T_{11}	=0 → 0 =1 → 1

C12	Persistence	Persistence - Authority	PID Authority **MUST** ensure that the PID cannot be deleted.	T_{12}	=0 → 0 =1 → 1
C13	Persistence	Persistence - Service	PID Service **MUST** ensure that the PID issued by the PID Authority cannot be deleted in its records.	$\Sigma\ T_{13,\ n}$	=0 → 0 >0 → 1
C14	Persistence	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.	T_{14}	=0 → 0 =1 → 1
C15	Application	Type Information	The PID Authority **SHOULD** provide information on the referenced object's fundamental type and management policy in a machine-actionable way.	$\Sigma\ T_{15,\ n}$	<2 → 0 =2→ 1
C16	Diversity	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.	T_{16}	=0 → 0 =1 → 1
C17	Diversity	Kernel Information Profiles	PID Services **MUST** engage the community to develop one or more Kernel Information Profiles appropriate to the use cases addressed by their services.	T_{17}	=0 → 0 =1 → 1
C18	Machine-Readability	Automation	The PID Service **SHOULD** maintain entity metadata as part of its PID Kernel information, but this source is not authoritative. Its main purpose is automation.	T_{18}	=0 → 0 =1 → 1
C19	Ecosystem	Accurate Entity Metadata	The PID Service **MUST** maintain entity metadata as accurately as possible in collaboration with the PID Owner. This copy is the authoritative version.	T_{19}	=0 → 0 =1 → 1
C20	Integrated	Openly Available	Services **MUST** be available to all researchers in the EU.	T_{20}	=0 → 0 =1 → 1
C21	Integrated	RI Integration	Services **SHOULD** allow integration with European Research Infrastructures.	$\Sigma\ T_{21,\ n}$	=0 → 0 >0 → 1
C22	Services	No End User Cost	The basic services of PID registration and resolution **SHALL** have no cost to end users.	T_{22}	=0 → 0 =1 → 1
C23	Services	Basic Service Maturity	A PID Service infrastructure **MUST** be at a minimum technology readiness level of 8. This applies to basic services (registration, resolution).	T_{23}	<8 → 0 =8 → 1 =9 → 1

C24	Services	Maturity - Value Added Services	Added value services **MAY** be offered at technology readiness levels lower than 8. ***OR*** Added value services **SHOULD** be offered at technology readiness level 8.	max (T_{23};$T_{24,1}$;$T_{24,2}$)	? → 0 ≥6 → 1
C25	Services	Availability - Measure	PID Services **MUST** meet 999 availability and uptime.	max ($T_{25,1}$;$T_{25,2}$)	? → 0 <8.77 → 0 ≥8.77 → 1
C26	Services	Availability - Procedure	PID Service Providers **SHOULD** document a summary of their maintenance and availability provisions publicly.	T_{26}	=0 → 0 =1 → 1
C27	Sustainable	Continuity	PID Service Providers **MUST** have a clear sustainability and succession plan with an exit strategy that guarantees the continuity of the resolution of its PIDs registered with the service.	$\Sigma\ T_{27,\ n}$	<3 → 0 =3 → 1
C28	Governed	Certification	PID Authorities and Services **MUST** agree to be certified with a mutually agreed frequency in respect of policy compliance.	T_{28}	=0 → 0 =1 → 1
C29	Services	Agreed Responsibilities	PID Services **SHOULD** agree with PID Managers the responsibilities for Kernel Information maintenance, preferably via contract.	T_{29}	=0 → 0 =1 → 1
C30	Resolution	Global Resolution	PID Service Providers **MUST** ensure their system supports the necessary API for global resolution services.	T_{30}	=0 → 0 =1 → 1
C31	Governed	Community Inclusion	PID Services **MUST** include representatives of the EU research community.	$\Sigma\ T_{31,\ n}$	<2 → 0 =2 → 1
C32	Governed	Justifiable Cost	PID Services **SHOULD** be provided at justifiable cost to PID Owners and PID Managers within EOSC.	T_{32}	=0 → 0 =1 → 1
C33	Governed	Global Governance	PID Service governance structures **SHOULD** align or be embedded in global governance structures	T_{33}	=0 → 0 =1 → 1

D.3.1.3 Additional Suggested Criteria

Annexure Table 3 – EOSC PID Policy: Additional Criteria, Metrics, and Benchmarks

C34	Persistence Median	Persistence	PID Services **SHOULD** aim for a persistence median that is acceptable to and aligns with community and dependency expectations.	T_{34}	=0 → 0 =1 → 1

C35	Resolution Percentage	Resolution	PID Service **SHOULD** resolve at least p percent of PIDs in a randomised sample, where p is determined by community and dependency expectations.	T_{35}	=0 → 0 =1 → 1

D.1.3.4 Applicability of Criteria to Actors

Annexure Table 4 – EOSC PID Policy: Mapping Criteria to Actors

#	Criterion	Imperative	Scheme	Authority	Service	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			✓	✓	
C12	Persistence - Authority	MUST		✓			
C13	Persistence - Service	MUST			✓		
C14	Resolution Authenticity	MUST				✓	
C15	Type Information	SHOULD		✓			
C16	Digital Representation	MUST				✓	
C17	Kernel Information Profiles	MUST			✓		
C18	Automation	SHOULD			✓		
C19	Accurate Entity Metadata	MUST				✓	
C20	Openly Available	MUST			✓		

C21	RI Integration	SHOULD			✓		
C22	No End User Cost	SHALL				✓	
C23	Basic Service Maturity	MUST			✓		
C24	Maturity - Value Added Services	SHOULD			✓		
C25	Availability - Measure	MUST			✓		
C26	Availability - Procedure	SHOULD			✓		
C27	Continuity	MUST			✓		
C28	Certification	MUST	(✓)	✓	✓	(✓)	
C29	Agreed Responsibilities	SHOULD			✓	✓	
C30	Global Resolution	MUST			✓		
C31	Community Inclusion	MUST			✓		
C32	Justifiable Cost	SHOULD			✓		
C33	Global Governance	SHOULD			✓		
C34	Persistence Median	SHOULD	✓	✓	✓	✓	
C35	Resolution Percentage	SHOULD	✓	✓	✓	✓	

D.3.1.5 Tests

Annexure Table 5 – EOSC PID Policy: Proposed Tests

#	Test	Description	Type	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
$T_{1,3}$	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
$T_{2,2}$	Sensitive - Indication	Sensitive PID Kernel Metadata can be defined - evidence is provided.	Binary	Yes = 1 No = 0	G2
$T_{2,3}$	Secure -	Sensitive PID Kernel metadata can be	Binary	Yes = 1	G2

	Encrypted Kernel Metadata	encrypted - evidence is provided.		No = 0	
$T_{2,4}$	Secure - Access	Sensitive PID Kernel Metadata requires access to be granted - evidence is provided.	Binary	Yes = 1 No = 0	G2
$T_{2,5}$	Secure - Authentication	Sensitive PID Kernel Metadata requires users to be authenticated - evidence is provided.	Binary	Yes = 1 No = 0	G2
T_3	Ownership is visible	A test determines if an ownership attribute is available for the PID. Evidence is provided of a mechanism to retrieve this information.	Binary	Yes = 1 No = 0	G3
T_4	Maintenance	A test to determine if the entity (PID) attributes are being maintained.	Binary	Yes = 1 No = 0	G4
T_5	UPDATE	Test is the same as test $T_{2,1}$	Binary	Yes = 1 No = 0	G5
T_6	Ownership Transfer	Public evidence of a contract or procedure that specifies ownership transfer provisions.	Binary	Yes = 1 No = 0	G6
T_7	Conformance Test	Testing that the relation between PID and entity, maintained by a manager, is conformat with Authority requirements. Existence of public evidence (declaration) is required.	Binary	Yes = 1 No = 0	G7
T_8	Use case guidance	Public evidence is available of community guidance on appropriate granularity and application in one or more use cases.	Binary	Yes = 1 No = 0	G8
$T_{9,1}$	Community Engagement - User Forum	Public evidence is provided of a periodic user forum.	Binary	Yes = 1 No = 0	G9
$T_{9,2}$	Community Engagement - User Forum	Public evidence is provided of periodic member or subscriber assemblies.	Binary	Yes = 1 No = 0	G9
T_{10}	Versioning support	Public evidence of versioning support in Kernel Information Profile or in user guidance.	Binary	Yes = 1 No = 0	G10
T_{11}	Versioning Policy	Public evidence of versioning policy.	Binary	Yes = 1 No = 0	G11
T_{12}	PID cannot be deleted	Public evidence is provided of the fact that the PID will never be deleted.	Value	Yes = 1 No = 0	G12
$T_{13,1}$	PID Persistence - Service - Evidence	Public evidence is provided by the Provider (Service) that PIDs cannot be deleted.	Binary	Yes = 1 No = 0	G13
$T_{13,2}$	PID Persistence - Service -	An inventory of PIDs issued by the Authority on behalf of the Service *a* is compared to	Value	$s/a<b = 0$ $s/a\geq b = 1$	G13

	Evidence	the inventory of PIDs published by the service *s* and the ratio is larger than a benchmark *b* determined by the community.			
T_{14}	Resolution Efficiency/ Integrity	This test is equivalent to T_{35}, but the result applies to a Manager.	Binary	Yes = 1 No = 0	G14
$T_{15,1}$	Machine-actionable type information	A pathway or published API call to verify the type is available.	Binary	Yes = 1 No = 0	G15
$T_{15,2}$	Machine-actionable management policy	A pathway or published API call to obtain the management policy is available.	Binary	Yes = 1 No = 0	G15
T_{16}	Digital Representation Exists	The test involves determining the percentage *f* of resolved PIDs that result in a viable entity, compared to a community expectation *p*.	Value	$f<p$ = 0 $f \geqq p$ = 1	G16
T_{17}	Community Involvement - Kernel Information Profiles	Public evidence of community involvement exists.	Binary	Yes = 1 No = 0	G17
T_{18}	Metadata is machine readable	Metadata is available in machine-readable format from the resolution target. This is publicly available by providing a URL pattern, API, or code example.	Binary	Yes = 1 No = 0	G18
T_{19}	Assuring accurate entity metadata	Public evidence of procedures or policies at Managers.	Binary	Yes = 1 No = 0	G19
T_{20}	Services are Open	Services (Providers) need to supply public evidence of open availability of services.	Binary	Yes = 1 No = 0	G20
T_{21}	Integration with EU RIs	Test is one or more evidentiary URLs to demonstrate use of the Service in a RI. Each instance is counted.	Binary	Yes = 1 No = 0	G21
T_{22}	No end user cost	Public evidence of cost structure or free services offered.	Binary	Yes = 1 No = 0	G22
T_{23}	Service Version Age - TRL	Number of months of operational availability of the current and previous version of PID registration service (*m*), compared to a benchmark *b*.	TRL Value	$m<p$ = 8 $m \geq p$ = 9	G23
$T_{24,1}$	Statement - TRL - beta	Public statement of new service in beta testing is provided	TRL Value	No = ? Yes = 7	G24

$T_{24,2}$	Statement - TRL - labs or experimental	Public statement of new service available as test, experiment, or labs version.	TRL Value	No = ? Yes = 6	G24
$T_{25,1}$	Availability	Public assertion of availability expressed as average annual downtime *d,* less than 8.77 hours per year.	Value	No = ? Yes = *d*	G25
$T_{25,2}$	Availability	Heartbeat monitoring of a service endpoint designated by the Service, expressed as average annual downtime *d,* less than 8.77 hours per year.	Value	No = ? Yes = *d*	G25
T_{26}	Maintenance and Availability Provisions	Publish public evidence of relevant provisions	Binary	Yes = 1 No = 0	G26
$T_{27,1}$	Continuity Provisions - plan	Publish public evidence of a continuity plan	Binary	Yes = 1 No = 0	G27
$T_{27,2}$	Continuity Provisions - exit strategy	Public declaration that an exit strategy is presented in the plan	Binary	Yes = 1 No = 0	G27
$T_{27,3}$	Continuity Provisions - exit strategy	Public declaration that continued resolution is addressed in the plan	Binary	Yes = 1 No = 0	G27
T_{28}	Certification	Public declaration of willingness to be certified	Binary	Yes = 1 No = 0	G28
T_{29}	Contract - Services and Managers	Evidence of a contract between Services and Managers exists - URL is available.	Binary	Yes = 1 No = 0	G29
T_{30}	Global Resolution Possible -	Public declaration of the countries, if any, where services are not available (count=c). If the number exceeds b, the provision is not satisfied.	Value	$c>b$ = 0 $c\leq b$ = 1	G30
$T_{31,1}$	Representative Governance - EU Researchers	Public declaration of representation on governance structure by member(s) of EU research community	Binary	Yes = 1 No = 0	G31
$T_{31,2}$	Representative Governance - Evidence	Public evidence is available of composition of governance structures	Binary	Yes = 1 No = 0	G31
T_{32}	Justifiable Cost	Publicly confirm that time-limited funds are used only for time-limited activities and that operational services are funded from membership and subscription fees. An appropriate test is formulated for POSI [62]: Ratio of structural income vs operational	Value	$\geq 1=1$ $<1=0$	G32

		expenditure is computed.			
T_{33}	Global Governance	Publicly confirm global governance participation	Binary	Yes = 1 No = 0	G33
T_{34}	Persistence Mean	The test involves a random statistically significant sample for a provider, and determining a distribution of resolvable PIDs as a function of time since creation. From this, a mean m can be evaluated against a norm n.	Value	$m<n$ = 0 $m \geqq n$ = 1	G34
T_{35}	Resolution Percentage	The test involves determining the percentage *f* of resolved PIDs that result in a viable entity, compared to a community expectation *p*.	Value	$f<p$ = 0 $F \geqq p$ = 1	G35

D.3.1.6 Guidance

Annexure Table 6 – EOSC PID Policy: Proposed Guidance

#	**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 scheme, authority, service, or manager. It could also be different depending on the Kernel Information Profile. Suggestion: store the retrieval instruction/ path as an attribute of a Service.
G4	In practice, evaluation is very difficult, due to two factors: ● It requires that a sample of millions of PID owners be evaluated across all services, and ● Some entities may never have to be maintained and are, despite years of non-maintenance, up to date. A measure of the mean update frequency of PIDs across a specific service, and monitoring its change over time against a benchmark, may be the only realistic assessment mechanism.
G5	This test is the same as $T_{2,1}$ - in cases where the Provider updates relevant attribute changes on behalf of the owner with the Authority. FOr now, we assume that this will always be the case.
G6	Public evidence is available of contractual or procedural provisions for ownership transfer. This will be a self-assessment for the foreseeable future.
G7	This test is proposed initially to be based on evidence provided by the Manager by way of a public declaration (e.g. on their website). Measuring conformance with authority resolution requirements will be much more complex if it is automated - depends on authority and possibly on Manager

	implementations.
G8	Test is initially proposed to be the provision of public evidence.
G9	A minimum level of community involvement requires user forums and/ or member and subscriber assemblies.
G10	We assume that in most instances, it will be possible for Services to point to public evidence (e.g in Kernel Information Profile schema or in user guidance) of versioning support.
G11	Managers can indicate public evidence of versioning policies or procedures.
G12	Authorities will usually state this fact prominently on their websites, or it may be contained in the published specification for the schema.
G13	Initially, it may be simpler to base the test on publicly available evidence, but an automated test could also be possible in future. The number of PIDs recorded by the Authority and those recorded by the Service will be different in practice, since it may include tests, for example. These known differences should be accommodated in an automated test.
G14	A randomised, statistically significant sample needs to be evaluated on a periodic basis (annually?) to gauge the efficiency of resolution of services. See also G35. For a Manager, it will be difficult to determine the percentage of tombstones in the total resolvable PIDs.
G15	The mechanisms whereby the information can be obtained are unlikely to be standardised or interoperable between authorities. A published API call or code example will be acceptable.
G16	This test can be the same as the Resolution Percentage test (T35), but the scope of computation is by Manager.
G17	Community involvement can include working groups with community representation, or a community consultation process prior to release of schema.
G18	There will be significant variation between Services, but publicly available instructions (URL pattern, API call, code example) will be adequate.
G19	It is not practically feasible to assess compliance from basic principles - hence public evidence from Managers that policies or processes exist will be the most suitable alternative.
G20	This is best achieved by pointing to a publicly available statement, or endorsement of a set of principles such as POSI [57].
G21	This evidence can be based on example URLs that illustrate the use of the Service in a European Research Infrastructure. It may require a validated list of such Research Infrastructures. If the measure and tests are implemented in this way, it can also be applied to gauge penetration of a Service in RIs.
G22	Evidence of the absence of costs for basic services to end users (PID requisition and registration, resolution).
G23	Services that have been in operation for some months can generally be regarded as production-tested (TRL9). Objective assessment of technology readiness level is somewhat more complex and is not being considered.

G24	Public statements about services in beta testing, labs, or experimental use are feasible mechanisms for determining other technology readiness levels.
G25	Uptime and availability can be measured by declaration (publicly available value) or by monitoring (this could be expensive and time-consuming).
G26	Public evidence of maintenance, uptime, and availability provisions is the most feasible option.
G27	Public evidence of a continuity plan needs to be provided, as well as declarations that such planning makes provision for an exit strategy and continued resolution.
G28	Authorities and Services need to publicly agree to be certified, which initially involves allowing a public record of self-assessment in respect of policy compliance.
G29	Evidence of contracting between Services and Managers need not involve individual contracts, since these may be confidential - but could point to the standard text of such a contract. The assumption is that Managers cannot operate without entering into a contract based on the standard text.
G30	This provision is somewhat problematic to evaluate objectively. In principle, any HTTP-based service should be available wherever the internet is accessible. In some countries, DNS may not resolve, inter alia due to network configuration or internet censorship policies. As a first step, Services are asked to indicate if they are aware of any countries where their service is not available. This count (or list) has to be compared to a community goal or benchmark.
G31	Provision can be adequately assessed by asserting publicly that EU researchers are included into governance structures, and by providing public evidence of composition.
G32	Determining whether costs to EU or EOSC users is justifiable is not simple: Providers may not be in a position to explain their cost structures, new services may have higher unit costs, and it may be difficult to obtain a representative assessment from all Managers that are in scope should one consider asking them via survey. POSI [10] provides two criteria that can be publicly attested as an alternative: ● “Time-limited funds are used only for time-limited activities. The day to day operations of the infrastructure should be supported by day to day sustainable revenue sources.” This confirms that Manager costs/ fees are not subsidising non-service activities. ● “Mission-consistent revenue generation. The infrastructure revenue should be consistent with the mission.” This confirms that Manager costs/ fees are not subsidising non-PID activities. There are two cases to consider in the tests: 1. Publicly available financial statements are reviewed. 2. A statement is made if financial records are not public.
G33	A public statement about global governance participation will be the simplest option for validation.
G34	Published literature [63], [64] suggest that the mean time that PIDs remain resolvable (a measure of persistence) varies quite significantly with the service in question. A randomised, statistically significant sample needs to be evaluated on a periodic basis (annually?) to gauge the persistence of services.
G35	Published literature [63], [64] indicates that even if a PID is resolvable, the content it resolves to does not always remain meaningful. A randomised, statistically significant sample needs to be evaluated on a periodic basis (annually?) to gauge the efficiency of resolution of services.

D.3.1.7 Assessment/ Evaluation

Assessment depends on the role of the actor in the ecosystem. Mandatory criteria (SHALL, MUST) are regarded as Go/ No Go criteria. These must be met in full to be compliant with the policy. Desirable criteria (SHOULD) and optional criteria (MAY) are grouped together to determine a ranking.

Annexure Table 7 – EOSC PID Policy: Proposed Assessment per Actor

#	Aspect	Compliance	Score (S)
E1	Schemes		
E1.1	Go/ No Go	S=1: Compliant	C24
E1.2	Ranking	Ranking S/2	C34+C35
E2	Authorities		
E2.1	Go/ No Go	S=2: Compliant	C12+C28
E2.2	Ranking	Ranking S/5	C2+C3+C15+C34+C35
E3	Providers/ Services		
E3.1	Go/ No Go	S=11: Compliant	C3+C5+C13+C17+C20+C23+C25+C27+C28+C31+C32
E3.2	Ranking	Ranking S/13	C1+C8+C9+C10+C11+C18+C21+C24+C26+C29+C33+C34+C35
E4	Managers		
E4.1	Go/ No Go	S=6: Compliant	C5+C7+C14+C16+C19+C22
E4.2	Ranking	Ranking S/6	C6+C11+C28+C29+C34+C35
E5	Owners		
E5.1	Go/ No Go	Not Applicable	
E5.2	Ranking	Ranking	C4

D.3.2 POSI

The Principles of Open Scholarly Infrastructure [57] was one of the case studies evaluated and is included here since many of the criteria and tests could overlap with those of the EOSC PID Policy.

D.3.2.1 Principles and Objectives

Annexure Table 8 – POSI: Principles and Objectives

#	Principle or Objective
P1	Governance
P2	Sustainability
P3	Insurance

D.3.2.2 Criteria

Annexure Table 9 – POSI: Criteria, Metrics, and Benchmarks

#	Principle/ Objective	Criterion	Metric	Benchmark
C1	P1	Coverage across the research enterprise. The infrastructure SHOULD have coverage across the research enterprise. It must transcend disciplines, geography, institutions and stakeholders.	$\Sigma\ T_{1,n}$	=0 → 0 =1 → 1
C2	P1	Stakeholder governed. The infrastructure SHOULD be board-governed, drawn from the stakeholder community	$\Sigma\ T_{2,n}$	>2 → 0 =2 → 1
C3	P1	Non-discriminatory membership. The infrastructure SHOULD provide opt-in membership or participation where any stakeholder may express an interest. The process of representation in day-to-day governance MUST also be inclusive with governance that reflects the demographics of the membership.	$\Sigma\ T_{3,n}$	>2 → 0 =2 → 1
C4	P1	Transparent operations. The infrastructure SHOULD make all processes and operations transparent (within the constraints of privacy laws)	T_4	=0 → 0 =1 → 1
C5	P1	Cannot lobby. The infrastructure SHALL NOT independently or collectively lobby to drive regulatory change.	T_5	=0 → 0 =1 → 1

C6	P1	Living will. The infrastructure SHALL have a plan to address the condition under which the organisation would be wound down.	T_6	=0 → 0 =1 → 1
C7	P1	Formal incentives to fulfil mission & wind-down. The infrastructure (and staff) SHOULD have direct incentives to deliver on the mission and wind down.	T_7	=0 → 0 =1 → 1
C8	P2	Time-limited funds are used only for time-limited activities. The day-to-day operations of the infrastructure SHOULD be supported by day-to-day sustainable revenue sources.	$\Sigma T_{8,n}$	=0 → 0 =1 → 1
C9	P2	Goal to generate surplus. The infrastructure should, within constraints provided by their legal status, SHOULD generate a surplus beyond their immediate operating costs.	$\Sigma T_{9,n}$	=0 → 0 =1 → 1
C10	P2	Goal to create a contingency fund to support operations for 12 months. The infrastructure SHOULD prioritise generating a contingency fund that can support a complete, orderly wind down (12 months in most cases).	$\Sigma T_{10,n}$	=0 → 0 =1 → 1
C11	P2	Mission-consistent revenue generation. The infrastructure revenue SHOULD be consistent with the mission.	T_{11}	=0 → 0 =1 → 1
C12	P2	Revenue based on services, not data. The infrastructure SHOULD ensure that data is openly available and a community property.	T_{12}	=0 → 0 =1 → 1
C13	P3	Open source. The infrastructure SHOULD make all software required to run the infrastructure available under an open-source licence.	$\Sigma T_{12,n}$	=0 → 0 ≥1 → 1
C14	P3	Open data (within constraints of privacy laws). The infrastructure, within constraints of privacy laws, SHALL make all data openly available.	T_{14}	=0 → 0 =1 → 1
C15	P3	Available data (within constraints of privacy laws). The infrastructure SHOULD, in addition to providing open data, make the data easily available.	T_{15}	=0 → 0 =1 → 1
C16	P3	Patent non-assertion. The infrastructure SHOULD commit to a patent non-assertion covenant.	T_{16}	=0 → 0 =1 → 1

D.3.2.3 Tests

Annexure Table 10 – POSI: Proposed Tests

#	Test	Description	Type	Method	Guidance
$T_{1,1}$	Coverage Institutional	No current tests exist. Proposal: Public evidence of appropriate	Guided Review	Yes = 1 No = 0	G1

		scope			
$T_{1,2}$	Coverage Geographic	No current tests exist. Proposal: Public evidence of appropriate scope	Guided Review	Yes = 1 No = 0	G1
$T_{1,3}$	Coverage Domain	No current tests exist. Proposal: Public evidence of appropriate scope	Guided Review	Yes = 1 No = 0	G1
$T_{1,4}$	Coverage Output Type	No current tests exist. Proposal: Public evidence of appropriate scope	Guided Review	Yes = 1 No = 0	G1
$T_{2,1}$	Independent Board	No current tests exist. Proposal: Public evidence of board existence	Guided Review	Yes = 1 No = 0	G2
$T_{2,2}$	Stakeholder membership	No current tests exist. Proposal: Public evidence of board composition	Guided Review	Yes = 1 No = 0	G2
$T_{3,1}$	Open Membership	No current tests exist. Proposal: Public evidence of membership process	Binary	Yes = 1 No = 0	G3
$T_{3,2}$	Member Governance	No current tests exist. Proposal: Public evidence of member involvement in governance processes	Guided Review	Yes = 1 No = 0	G3
T_4	Transparent	No current tests exist. Proposal: A problematic criterion, since ‘all processes and operations’ is undefinable in practice. Test by reviewing scope of publicly available process and service documentation and making a subjective assessment of its completeness.	Guided Review	Yes = 1 No = 0	G4
T_5	Cannot Lobby	No current tests exist. Proposal: Evidence of a public declaration by the infrastructure.	Binary	Yes = 1 No = 0	G5
T_6	Living Will	No current tests exist. Proposal: Public evidence of a continuity or transition plan[14].	Binary	Yes = 1 No = 0	G6
T_7	Formal Incentives	No current tests exist. Proposal: Very difficult to test. If public evidence of this exists in the continuity plan, it can be referenced.	Guided Review	Yes = 1 No = 0	G7
$T_{8,1}$	Time-Limited Funds - Financial	No current tests exist. Proposal: Publicly available financial statements are reviewed. Ratio of	Value	$R_1 \geq 1=1$ $R_1<1=0$	G8

[14] Very similar criteria and tests exist for CoreTrustSeal, their guidance can possibly be adopted and amended.

	Statement	structural income vs operational expenditure is computed.			
$T_{8,2}$	Time-Limited Funds - No Financial Statement	No current tests exist. Proposal: A public statement about the ratio between structural funds and operational expenditure is available.	Value	$R_2 \geq 1=1$ $R_2<1=0$	G8
$T_{9,1}$	Surplus: Public	No current tests exist. Proposal - Public information: review published statements	Guided Review	Yes = 1 No = 0	G9
$T_{9,2}$	Surplus: Private	No current tests exist. Proposal- Private information: evidence of public declaration	Binary	Yes = 1 No = 0	G9
$T_{10,1}$	Contingency: Public	No current tests exist. Proposal- Determine months of contingency cover from financial statements	Value	≥12 = 1 <12 = 0	G10
$T_{10,2}$	Contingency: Public	No current tests exist. Proposal- Private information: public declaration of months of cover	Value	≥12 = 1 <12 = 0	G10
T_{11}	Mission-Consistent Revenue	No current tests exist. Proposal- a declaration or public evidence that revenue is supportive of the mission.	Binary	Yes = 1 No = 0	G11
$T_{12,1}$	Sustainable Operational Revenue - Public	No current tests exist. Proposal- Public financial data - ratio of service or membership revenue to operational expenses.	Value	≥1 = 1 <1 = 0	G12
$T_{12,1}$	Sustainable Operational Revenue - Private	No current tests exist. Proposal- Public financial data - ratio of service or membership revenue to operational expenses confirmed publicly as being >=1.	Value	≥1 = 1 <1 = 0	G12
$T_{13, n}$	Open Source	Proposal: one or more (n) verifiable references to a software repository.	Binary	0 = 0 1 = 1	G13
$T_{14,n}$	Open Data	Proposal: At least one API or publicly accessible web page is available for data access.	Binary	0 = 0 1 = 1	G14
T_{15}	Easily Accessible	Proposal: Subjective assessment	Guided Review	Yes = 1 No = 0	G15
T_{16}	Patent non-assertion	Proposal: public declaration available	Binary	0 = 0 1 = 1	G16

D.3.2.4 Guidance

Annexure Table 11 – POSI:Proposed Guidance

#	Guidance
G1	Proposal: guidance on how to annotate and reference public evidence of coverage, and how to categorise based on domain and geographic coverage. Infrastructure should address a diversity in at least 1 category (geographic, output type, institutional, domain).
G2	Proposal: guidance on how to annotate and reference public evidence of the board's existence and composition, and how to verify that the board represents stakeholders.
G3	Proposal: Public evidence of membership process by referencing documentation on membership application process. Public evidence of member involvement in governance processes - e.g. via a member forum or organisation, annual member assembly, etc.
G4	Proposal: Inventory of public evidence of processes and operations. Subjective evaluation of the completeness of the inventory compared to the infrastructures stated products and services.
G5	Proposal: Absence of lobbying activity is best asserted by the infrastructure itself, by way of e.g. a public statement on their website. It implies a mechanism to alter the test result if credible evidence of the contrary emerges.
G6	Proposal: Public evidence of a continuity plan that addresses one or more desirable elements or has one or more desirable attributes.
G7	Proposal: Guidance can be provided to reviewers to determine if this criterion can be or will be satisfied by the infrastructure. A subjective assessment.
G8	Proposal: Either one of two ratios can be computed from published financial statements, or publicly attested if financial statements are private: 1. Ratio of structural income to operational expenditure 2. Ratio of grant income to capital or project expenditure If either of these values is equal to or exceeds 1 for a suitable reporting period (previous 3 years, for example), there is no reason to be alarmed and it is clear that operational expenses do not depend on grant income.
G9	Proposal: Financial information is not always publicly available, and that affects the test. For public information, guidance will be required to interpret and assess financial statements. For private information, a public declaration of adequate surplus and its intended use may be required.
G10	Proposal: Financial information is not always publicly available, and that affects the test. A ratio of contingency funds (liquid assets) to monthly operational expenditure can be calculated from financial statements, if not available, a public assertion of months of contingency cover will be adequate.
G11	This is not easy to measure directly, and the simplest is to ask for a public declaration in this respect.
G12	Proposal: Financial information is not always publicly available, and that affects the test. A ratio of sustainable income (membership fees, service subscriptions) to operational expenditure can be calculated from financial statements, if not available, a public assertion of the ratio being 1 or more will be adequate.

G13	Proposal: one or more references to publicly available software repositories (Github, Bitbucket, …) verified as deployable. Long-term repositories are better (Software Heritage, …).
G14	Proposal: References to API, harvesting and data discovery/ download resources and endpoints are all valid.
G15	Proposal: difficult to assess objectively - guidance includes 1. Formal published user ratings mostly positive - assessment = yes 2. Positive testimonials on site - assessment =yes 3. User forums, feedback mechanisms, ticket systems in place - assessment = yes 4. Evaluator experience is positive - assessment = yes
G16	Proposal: not easy to determine automatically - a world-wide patent search is possible via PATENTSCOPE but this the entity being searched for may be incorrect. Rely on a public declaration.

D.3.2.5 Assessment/ Evaluation

Annexure Table 12 – POSI: Proposed Assessment

#	**Aspect**	**Weight**	**Score**
E1	Mandatory Criteria (MUST)	Go/ No Go	C3+C5+C6+C14=4
E2	Desirable Criteria (SHOULD)	Rank	C1+C2+C4+C7+C8+C9+C10+C11+C12+C13+C15+C16
E3	Optional Criteria (COULD, MAY)	Secondary Rank	