

Generic Evaluation Reference Model

Introduction	2
Overview	3
Status of an Entity	3
A Governed Standard and Process	4
Governance Policy	4
Standard Requirements	5
Evaluation Process	5
Assessment Method	6
Outcome	6
Validity Terms	6
Artefacts, Actors, and Agents	7
Acknowledgements	8

Introduction

As domains of practice evolve and mature they tend to develop agreed best practices. These approaches may be formalised into standards to support consistency and provide the foundation for further improvement.

Ruusalepp et al define a standard as the “specification of precise criteria to be used consistently and appropriately”, noting that they may be either *normative*—setting requirements for quality and actions, or *informative*—describing and guiding the use of methods¹.

Standardisation may be applied to a wide variety of entities (people, processes, organisations, objects, etc.). For normative standards, a variety of evaluation approaches may be applied. Evaluating the status of an entity requires that some defined attributes (skills, quality levels, capabilities, characteristics) are compared to standardised requirements. Evidence must be available to support the evaluation.

Evaluation processes may be delivered at varying levels of formality. Assessment methods may be entirely machine-actionable (e.g. conformity of an XML file to a schema) or mediated by human actors (e.g. an audit and certification process). Aligning with standard requirements and creating and managing evidence require resources. An organisation or individual may be subject to more than one type of evaluation.

A number of proposals and projects around the European Open Science Cloud (EOSC²) and FAIR³ data envision the development or expansion of evaluation, audit or certification processes. The purpose of these is to demonstrate good practice or to set out rules of engagement between the people, processes and technologies that comprise data infrastructures. Targets for evaluation include data repositories, metadata registries, software, services and digital objects.

This paper describes a generic model in terms of the actors, concepts and processes necessary for the development, comparison, and implementation of evaluation frameworks.

¹ Raivo Ruusalepp, Christopher A. Lee, Bram van der Werf and Matthew Woollard, ‘Standards Alignment’ in Nancy Y. McGovern, ed., *Aligning National Approaches to Digital Preservation* (Atlanta, GA: Educopia Institute, 2012), pp. 115-166.

² <https://www.eosc-portal.eu/>

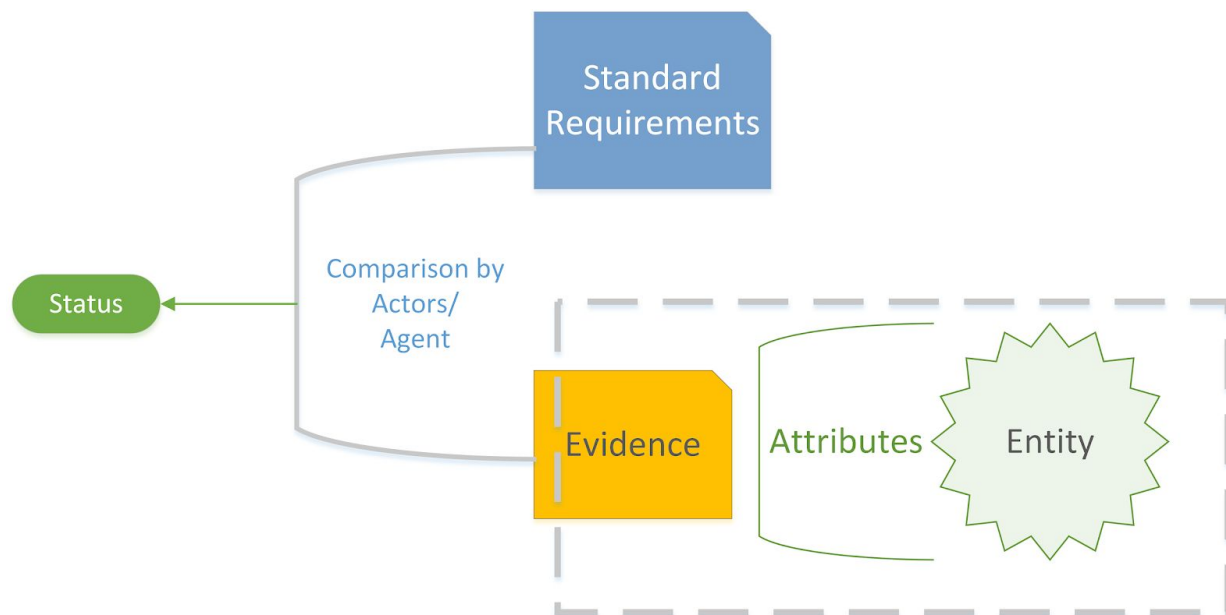
³ <https://www.force11.org/group/fairgroup/fairprinciples>

Overview

An **evaluation process** defines the **status** of the **attributes** of an **entity** by comparing **evidence** to **standard requirements**.

- ❑ A **governing body** manages **standard requirements** and an **evaluation process** according to a **governance policy**.
- ❑ An **applicant** represents the **entity** in the evaluation process.
- ❑ **Evidence** about the **attributes** of the entity is compared to the standard requirements
- ❑ The **status** is the level of alignment between the evidence and the standard requirements.
- ❑ In the evaluation process, an **assessment method** measures the status, leading to an **outcome** which has **validity terms**.

Status of an Entity



HLH-GenericEvaluationReferenceModel-A-_v02_00.vsd

Diagram A: Entity, Attributes, Evidence, Standard Requirements and Status

An entity can be any “thing”: person, organisation, object etc.

- ❑ The entity must be clearly and unambiguously identified.
- ❑ The attributes of an entity that are relevant to the standard requirements must be defined

Example types of attributes of entities.

- Research skills of a person
- Logical Integrity of a digital object
- Completeness of a metadata record
- Long term preservation capability of a repository
- Information security level of a data environment
- Functionality of a piece of software
- Conformance of an XML file
- Quality of a product or service

☐ Evidence is provided about the attributes of the entity

Evidence could include:

- Policies, procedures or other documentation
- Information published to the internet
- Metadata about a digital object
- Direct inspection by a reviewer
- A written response to an examination paper
- Content within an internal wiki

Applying an evaluation process:

- ☐ Evidence about the attributes of the entity is compared to the standard requirements
- ☐ The status is the level of alignment between the evidence and the standard requirements.

Note: the statement “the **status** of the **entity** is how it compares with the **requirements**” is valid, but not sufficient. To align with the model an evaluation framework must acknowledge that:

- Only some aspects of the entity are being evaluated: the defined **attributes**.
- It may not be possible to directly inspect every attribute, so rely on **evidence**.

Confidence in the evaluated status of the entity depends on standard requirements and evaluation processes which are clear and well-governed.

A Governed Standard and Process

A **governing body** manages **standard requirements** and an **evaluation process** according to a **governance policy**.

Governance Policy

The governance policy:

- ❑ Defines the authority of the governing body
- ❑ Describes the management (including change management) of the policy, standard requirements, and evaluation process

Standard Requirements

The standard requirements define the criteria which the attributes of the entity will be evaluated against:

- ❑ The standard can be any set of criteria
- ❑ The standard must define which entities it applies to

Examples of standards.

- Technical specification
- Certification standard
- Rights statements
- Policies
- Legislation

Evaluation Process

An **evaluation process** defines the **status** of the **attributes** of an **entity** by comparing **evidence** to **standard requirements**.

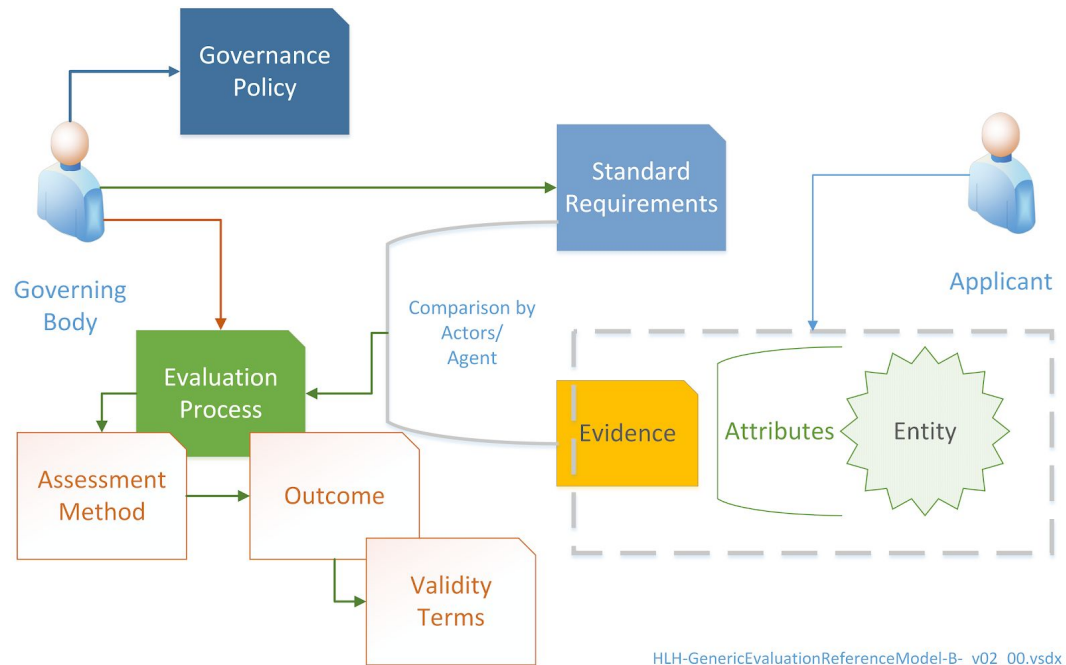


Diagram B: Generic Evaluation Reference Model

- ❑ The evaluation process must be a defined set of actions

Assessment Method

The assessment method defines the level of alignment between the standard requirements and the evidence about the attributes of the entity (the status).

- ❑ The assessment method can be the application of any metric. A metric can range from a defined objective measure to a subjective opinion, or any combination of these approaches.

Examples:

- A range from A to F
- Score from a total 84/100
- Strongly Agree to Strongly Disagree
- Checksums match (or not)

- ❑ If different metrics interact to deliver a final status, the interaction should be defined.

Example:

- Each of the criteria is measured separately on a scale from 0-5, the final status is a cumulative score without weighting.

An assessment method leads to an **outcome** which has **validity terms**.

Outcome

Assessment methods are separate from outcomes e.g. a “PASS” outcome could be derived from a score of 60/100 or of 80/100 score.

- ☐ All of the possible outcomes must be defined
- ☐ Outcomes may be positive, negative or neutral

Example outcomes:

- The entity is out of scope for the standard (neutral)
 - The evidence is insufficient to reach an outcome (neutral)
 - A pass (positive)
 - A fail (negative)
-
- ☐ All outcomes must have associated validity terms

Validity Terms

The validity terms are the rules related to the outcomes.

- ☐ Validity terms for each possible outcome must be defined

For example:

- An outcome could be valid for a defined time period
 - The entity is certified for 5 years
- An outcome could be dependent on fulfilling a corrective action plan
- There may be a right to appeal
 - The applicant may request an alternate reviewer
- There may be terms associated with repeating the evaluation
 - The applicant cannot re-apply for 6 months
 - The entity may only be re-evaluated if the evidence is revised
- Further evaluations may be required to maintain the outcome
 - The applicant must undertake ongoing internal audits to maintain certification

Artefacts, Actors, and Agents

- ☐ Artefacts must be defined

Entity/attributes definitions, standard requirements, governance policy and an evaluation process (including assessment method, outcome and validity terms) are the minimum artefacts

to be defined. This generic model does not define how these different artefacts should be grouped or presented.

- ❑ All roles must be defined

Governing body and applicant are the minimum roles to be defined. A particular evaluation framework may include additional roles. For example, a framework may include a reviewer to compare the evidence to the standard, but this role may also be taken by a (machine) agent.

- ❑ How roles interact with artefacts must be defined

For example, the governance policy, standard requirements and evaluation process must all be associated with a governing body. The governing body is not necessarily the original creator of the artefact. The model presents a single governing body for simplicity, but different artefacts may have different governing bodies.

- ❑ If there is more than one governing body, interactions between them must be defined

The applicant has a defined relationship to the entity and the evidence. The applicant is responsible for any interactions with the governing body or other identified actors.

- ❑ An applicant acts as the representative of the entity

When a particular evaluation framework is implemented the governing body, applicant and any other roles must be clearly and unambiguously associated with actors (people, groups or organisations).

- ❑ Which actors play each role must be clearly identified

Conclusion

The generic evaluation reference model seeks to define the minimum set of terms and concepts necessary to support a common approach to evaluation. The model supports the design, evaluation and comparison of evaluation frameworks.

It is increasingly likely that applicants will need to comply with more than one evaluation framework. Comparison of evaluation frameworks using the generic model permits consideration of the resource impact on applicants. The model will also support applicants in developing information management systems that can efficiently support evidence for multiple evaluation frameworks

There are a number of characteristics of evaluation approaches which the authors consider important, but out of scope for the model. For example, it is assumed that for an evaluation approach to be successful there must be a perceived beneficial outcome to the applicant, and a degree of confidence in the authority of the governing body. Another broadly desirable characteristic is transparency. The validity, adoption and confidence in any evaluation process are likely to be increased by maximising the transparency of entity definitions, governance, standards and processes (including assessment methods, outcomes and validity terms).

The next steps are to open the model up for further discussion and to develop exemplars based on evaluation frameworks which exist or are in development. Exemplars will be used to validate the model. Subsequent versions will amend, clarify, correct, reduce or expand the model as necessary.

Acknowledgements

The authors would like to thank Professor Matthew Woollard for his initial feedback on the model including elaboration on the interactions between entities, attributes and evidence.