Project Title	Expanding FAIR solutions across EOSC
Project Acronym	FAIR-IMPACT
Grant Agreement No.	101057344
Start Date of Project	2022-06-01
Duration of Project	36 months
Project Website	https://fair-impact.eu/

D4.1 - Semantic artefact governance models and disciplinary approaches for inclusion within EOSC

Work Package	WP 4, Metadata and Ontologies
Lead Author (Org)	Parham Ramezani (LifeWatch), Nina Grau (INRAE), Clement Jonquet (INRAE), Nicola Fiore (LifeWatch)
Contributing Author(s) (Org)	Ilaria Rosati (CNR), Enrica Nestola (CNR), Maria Poveda-Villalon (UPM), Lucia Sanchez Gonzalez (UPM)
Due Date	2024-07-31
Date	2024-07-31
Version	v2 - NOT YET APPROVED BY THE EUROPEAN COMMISSION

Dissemination Level	
X	PU: Public
	PP: Restricted to other programme participants (including the Commission)
	RE: Restricted to a group specified by the consortium (including the Commission)
	CO: Confidential, only for members of the consortium (including the Commission)

Funded by the European Union

Versioning and contribution history

Version	Date	Author	Notes
0.1	2024.04.15	Parham Ramezani (Lifewatch)	TOC and V0.1
0.2	2024.06.04	Nina Grau (INRAE)	Sections 1.1.1. Context, 1.1.3. FAIR-IMPACT project and T4.1 goals, 1.2.1. Scope of the D4.1, 1.2.2. Methodology
0.2	2024.05.12	Parham Ramezani (Lifewatch)	Sections 1.1.2. Challenges, 1.2.3. Semantic Artefact governance definition
0.3	2024.05.17	Parham Ramezani (Lifewatch)	Sections 2.1. Engineering methodologies, 2.3. Governance framework
0.4	2024.05.20	Parham Ramezani (Lifewatch)	Section 3. Process followed
0.4	2024.05.30	Nick Juty (UNIMAN)	Review and general comments on the document
0.5	2024.06.12	Nina Grau (INRAE)	Section 2.2. Organisational structure, Appendices
0.6	2024.06.12	Nina Grau (INRAE), Parham Ramezani (Lifewatch)	Appendices formating
0.6	2024.06.17	Nina Grau (INRAE)	Revising of 1.2 to 1.2.3 structure and content; formatting of the table of content
0.7	2024.06.18	Parham Ramezani (Lifewatch)	Section 4. Semantic artefact governance models (Governance Framework and Aspects of SA lifecycle)
0.7	2024.06.18	Nina Grau (INRAE)	Section 4. Semantic artefact governance models (Organisational Structure)
0.8	2024.06.19	Nina Grau (INRAE)	Executive Summary
0.8	2024.06.19	Parham Ramezani (Lifewatch)	Conclusion
0.9	2024.06.24	Clement Jonquet (INRAE)	Editing and shortening the Summary. Completed S1.1.1 with EOSC IF and SRIA requirements. Review and general comments on the document.

0.9	2024.06.24	Ilaria Rosati (CNR), Enrica Nestola (CNR)	Review and general comments on the document
0.9	2024.06.28	Maria Poveda-Villalon (UPM), Lucia Sanchez Gonzalez (UPM)	Review and general comments on the document
1.0	2024.06.28	Nina Grau (INRAE), Parham Ramezani (Lifewatch)	Reply to the comments, provide the necessary modification. Version send for reviewing
1.1	2024.07.01	Richard Dennis (reNEW, University of Copenhagen)	Review and edit the documents. New version of the Summary.
1.2	2024.07.26	Simon Hudson (CODATA)	Review and general comments on the document
1.3	2024.07.29	Nina Grau (INRAE), Parham Ramezani (Lifewatch)	Reply to the comments and provide the necessary modification
1.4	2024.07.30	Nina Grau (INRAE)	Formatage of the document update of table of content, list of figures ,list of tables and Terminology
1.5	2024.07.31	Clement Jonquet (INRAE)	Final review to produce v2. Final list of contributors. Re-merge the 2 versions of the Summary. Fixed Figure sizes and captions.

Disclaimer

FAIR-IMPACT has received funding from the European Commission's Horizon Europe funding programme for research and innovation programme under Grant Agreement no. 101057344. The content of this document does not represent the opinion of the European Commission, and the European Commission is not responsible for any use that might be made of such content.

Table of Contents

List of Tables

List of Figures

Terminology

Acronym	Description
API / APIs	Application Programming Interface
BASF	Badische Anilin- und SodaFabrik
CO or CROP	Crop Ontology Project
CSP	Standing Committee on Science Priorities
DL	Description Logics
EFO	Experimental Factor Ontology
ETSI	European Telecommunications Standards Institute
EM	Engineering Methodologies
EOSC	European Open Science Cloud
FAIR	Findable, Accessible, Interoperable, Reusable principle
FAO	Food and Agriculture Organization
FOAF	Friend Of A Friend
GOMO	Governance Operational Model for Ontologies
HTTP	HyperText Transfer Protocol
IF	Interoperability Framework
IG / IGs	Interest Group
IILG	IUA-IVOA Liaison Committee
IRI	Internationalised Resource Identifier
IVOA	International Virtual Observatory Alliance
LOT	Linked Open Terms
MIRO	Minimum Information for the Reporting of an Ontology
NFDI	french National Research Institute for Agriculture, Food and Environment
PID	Persistent Identifier
PURL	Persistent Uniform Resource Locator
OBO Foundry	Open Biological and Biomedical Ontology Foundry
ODK	Ontology Development Kit
OMO	OBO Metadata Ontology
OWL	Web Ontology Language
RACI	Responsible, Accountable, Consulted and Informed responsibility assignment matrix
RDF	Resource Description Framework
RDF-S	RDF Schema
SA / SAs	Semantic Artefact
SAC /SACs	Semantic Artefact Catalogue
SAREF	Smart Applications REFerence
SCSP	Standing Committee on Standards & Processes
SemVer	Semantic Versioning specification
SKOS	Simple Knowledge Organisation System
SKOSMOS	Open source web-based SKOS vocabulary
SOPs	Standard Operating Procedures
SPARQL	SPARQL Protocol and RDF Query Language
SPOT	Samples, Phenotypes and Ontology Team
SRIA	Strategic Research and Innovation Agenda

Acronym	Description
TCG	Technical Coordination Group
TG / TGs	Task Group
UMLS	Unified Medical Language System
UUID / UUIDs	Universally Unique IDentifier
URI	Uniform Resource Identifier
URL	Uniform Resource Locator
W3C	World Wide Web Consortium
WG /WGs	Working Group
XML	Extensible Markup Language

Executive Summary

Semantic artefacts (SAs - terminologies, taxonomies, thesauri, vocabularies, metadata schemas and standards) are essential for standardising data representation and annotation, encapsulating the highest level of meaningful knowledge within interoperability frameworks. These artefacts are fundamental to sustainable, quality-verified data practices, aligning with the FAIR Data Principles, particularly Principle I.2. Although experts, including those in the EOSC Interoperability Framework, emphasise the need for governance policies for SA, no universal governance model exists that fits all scientific communities or projects. This deliverable presents various governance models to guide and inspire stakeholders in addressing the governance of semantic artefacts. It is the final and primary deliverable (D4.1) of FAIR-IMPACT's T4.1 on "Semantic artefact disciplinary governance," which reviews and analyses community practices to propose strategies tailored to SAs' characteristics.

This document sets out the baseline modality of semantic artefact governance derived from surveyed community-based practices. It provides actionable models for communities or institutions looking to establish or update their semantic artefact governance frameworks. The report is structured into four main sections. The first section introduces the overarching context and the inherent challenges of SA governance, setting the stage for understanding the complexities involved. The second section delves into the three primary components derived from community-based practices: engineering methodologies, organisational structure, and governance framework. Engineering methodologies cover the technical processes and lifecycle management of SAs. The organisational structure addresses the roles and responsibilities within institutions that support effective governance, and the governance framework outlines the high-level principles and standards guiding SA governance. The third section outlines the methodology used to develop the governance models, detailing the steps taken to synthesise community practices into structured, applicable models. The final section presents the baseline governance models, aligning them with specific use cases to demonstrate practical applications. These models address the diverse needs of different scientific communities and projects.

As a result of this comprehensive analysis, FAIR-IMPACT T4.1 has formulated initial requirements and developed three distinct models of semantic artefact governance based on community-based stewardship practices. These models facilitate synchronicity, promoting coordinated efforts across scientific communities and ensuring consistent application of best practices. They aim to ensure long-term sustainability by supporting the enduring use and maintenance of semantic artefacts and aiding EOSC communities in seamlessly integrating semantic artefacts within the European Open Science Cloud, thus advancing the FAIR data agenda. This deliverable empowers stakeholders with the tools and frameworks to establish robust governance structures for semantic artefacts, fostering a more interconnected and sustainable data ecosystem.

Funded by
the European Union

1. Introduction

1.1 Topic presentation

1.1.1 Context

Currently, data sharing is actively promoted by policy making bodies such as the European Commission[1], supported by evidence that increased data sharing generally fosters innovation. Semantic Artefacts (SA) are employed to represent and annotate data in a standardised manner. For this work, we have adopted the definition of SA provided by the FAIRsFAIR project and embraced it FAIR-IMPACT): *"a machine-actionable and -readable formalisation of a conceptualisation enabling sharing and reuse by humans and machines"*[2]. In other words, we use the term SA as a broader term to include ontologies, terminologies, taxonomies, thesauri, vocabularies, metadata schemas, and standards. SAs embody the highest level of meaningful knowledge representation within an interoperability framework, making them particularly sensitive to changes, versioning, sharing, management, etc., all governance-related aspects.

The governance of SA has been identified as a requirement by the European Open Science Cloud (EOSC) Interoperability Framework (IF) and the EOSC Strategic Research and Innovation Agenda (SRIA) (v1)[3]:

- *"SRIA: Develop governance structures for coordinating the work on metadata and ontologies within EOSC, both for specific disciplinary communities and overall coordination.*
- *SRIA: This governance should be built primarily around existing discipline-based communities but needs to be coordinated across these communities within EOSC*
- *SRIA: The work that these governance structures coordinate should include registries that describe metadata schemata in a standardised and machine-actionable way, better researcher-focused tools and services working with these metadata, crosswalks between existing metadata schemata, and training and documentation.*
- *IF: Repositories of semantic artefacts rules with a clear governance framework.*
- *IF: Documents explaining terms and conditions and acceptable use policies for interoperability services are needed. For instance, providing clear descriptions of the service-level agreements of those providing catalogues and registries of semantic artefacts."*

FAIR-IMPACT T4.1 was created to address these requirements when deploying EOSC.

[1] http://data.europa.eu/eli/dir/2019/1024/oj
[2] https://doi.org/10.5281/zenodo.4314320
[3] https://eosc.eu/wp-content/uploads/2023/08/SRIA-1.0.pdf

1.1.2 Challenges

SA has become central to sustainable, quality-verified data recommendations to preserve their integrity, such as the FAIR Principles[4]. However, the FAIR Principles include a notion of “long-term care,” expert groups[5] advocate for establishing governance policies for these dynamic entities. Despite the acknowledged importance of governance for SA, there is no standard definition for SA governance available in the literature[6], underscoring the complexity of this task. Indeed, governance comprises the framework through which an organisation is directed and managed, including the procedures ensuring accountability for the entity and its members. It involves the processes of decision making, establishing rules, and implementing enforcement mechanisms to steer the operations of an organisation or community. Because of its application in different domains, governance is often combined with other terms such as metadata governance, platform governance, infrastructural governance, project governance, software governance, IT governance, etc. Despite the various formats the term encompasses, governance aims to enhance efficiency, objectivity, and transparency in an organisation's structures and decision-making processes.

1.1.3 FAIR-IMPACT project and T4.1 goals

The FAIR-IMPACT project supports harmonising and synchronising the FAIR Principles and practices to realise a FAIR EOSC environment. FAIR-IMPACT recognises the critical role of governance in sustaining FAIR research outputs and services, dedicating a specific task to this subject with a focus on SAs. FAIR-IMPACT’s T4.1, entitled “Semantic artefact disciplinary governance”, seeks to review and analyse community governance practices for SAs and to develop governance models.

By doing so, T4.1 advances expert recommendations by providing models for SA governance. Community-based stewardship practices in SA governance will facilitate synchronicity across scientific communities and ensure long-term sustainability.

1.2. Objectives of the T4.1 deliverable (D4.1)

1.2.1. Scope of the D4.1

This deliverable aims to: (1) identify governance modalities for SAs and (2) provide governance models applicable in various disciplines for making relevant SA-related decisions with the appropriate stakeholders. To achieve this, the approach involves comparing and analysing existing SA governance practices within various communities connected to EOSC.

In Task 4.1, we conducted a comprehensive survey on the prevailing practices of SA governance within scientific communities. During a workshop held in September 2023[7], we invited esteemed speakers from various exemplary communities who contributed to

[4] https://doi.org/10.1038/sdata.2016.18
[5] https://doi.org/10.5281/zenodo.10287011
[6] https://doi.org/10.1016/j.ijinfomgt.2019.07.008
[7] “Semantic artefact workshop”, webpage of the FAIR-IMPACT event. 2023. URL : https://fair-impact.eu/events/fair-impact-events/fair-impact-semantic-artefact-governance-workshop

Funded by the European Union

elucidating SA governance models. We have compiled nine exemplary cases of governance practices about semantic artefacts. A Milestone entitled MS4.1: "Semantic Artefact Governance Models: Examples of Community Practices"[8], has been produced, serving as the baseline for the final deliverable of governance models for semantic artefacts within the EOSC in Deliverable 4.1 (D4.1).

1.2.2 Methodology

SA governance is a new concept that we aimed to define and elaborate in this deliverable. The need for a holistic approach to SA governance aims at avoiding ad-hoc practices, project by project or even ontology per ontology. A first bottom-up approach enables us to capture the actual practices employed by different communities in dealing with their SAs. Through the “FAIR-IMPACT Semantic Artefact governance workshop” from Sept. 2023, we were able to explore various governance strategies and habitual practices across multiple communities by examining the decision domains of data and IT governance, as outlined by Khatri’s[9] 2010 reference “data governance” paper. The milestone entitled “MS4.1. Semantic artefact governance models: example of community practices”[10] reports describe the communities surveyed and the multiple approaches to community-driven governance of semantic artefacts examined during this workshop. Secondly, we employed a top-down approach to define SA governance by adapting existing definitions of data governance from the literature to the critical components of SAs. This dual approach integrates the functions and concepts of SA governance, allowing us to extract SA governance components and develop models closely aligned with observed standard practices. To achieve this, we decided to extract the effective use of SA governance and integrate it into the SA governance components derived from the top-down approach.

This document considers SAs specifically utilising semantic web technologies and standards within their development processes. Consequently, the deliverable's output, which comprises potential SA governance models, is aligned with Semantic Web practices.

1.2.3 Semantic Artefact governance definition

To conceptualise the definition of SA governance, we based our approach on a data governance definition derived from a structured literature review and presented by Rene Abraham et al. 2019:

- “*Data governance specifies a cross-functional framework for managing data as a strategic asset.*
- *data governance specifies decision rights and accountabilities for an organisation’s decision making about its data,*
- *data governance formalises data policies, standards, and procedures and monitors compliance*.”

---

[8] “MS4.1. Semantic artefact models: examples of community practices”. 2023 T4.1 milestone report DOI: 10.5281/zenodo.10287010
[9] https://doi.org/10.1145/1629175.1629210
[10] “M4.1. Semantic artefact models: examples of community practices”. 2023 T4.1 milestone report DOI: 10.5281/zenodo.10287010

To align our definition of SA governance with data governance, SA must be regarded as a strategic asset within an organisation, necessitating development, publication, and maintenance throughout its lifecycle. Another characteristic emphasises an organisation's decision making about its data; in our case, we narrow the scope and consider only decisions that need to be made about SA management, how these decisions are made, and who in the organisation has the right to make these decisions. Finally, in formalising policies, standards, procedures, and monitoring, compliance must be limited and aligned to SA development practices and technologies instead of considering generic aspects of data management.

Based on this, we conceptualise the SA governance definition as:

A principled approach for standardising different aspects of semantic artefacts development throughout the semantic artefact lifecycle, from acquisition to use to disposal. It specifies decision rights and accountabilities for an organisation's decision making about semantic artefact.

Thus, we broke down SA governance into: 1) engineering methodology, 2) organisational structure and 3) governance framework, which will be scrutinised for all the surveyed organisations and communities. This separation allowed us to categorise practices for SA development and management that align with the features of each component. These components are then modular to form comprehensive SA governance models for several target groups.

2. Semantic artefact governance components

The SA governance definition allowed us to identify three main components:

(1) The first component is SA **Engineering Methodologies (EM)**, encompassing the development process, lifecycle, and methods for building SA. Despite various classifications of EMs, they share the commonality of structuring activities, tasks, and operations into systematic and formal phases necessary for SA development. In FAIR-IMPACT, T4.2 and T5.3 investigate the development of a FAIR-by-design methodology.

(2) The second component is **Organisational Structure**, which includes engagement of stakeholders and roles and responsibilities. We will explore this topic by addressing the question: who does what, and how?.

(3) The third component is the **Governance Framework**, which defines a set of high-level fundamental rules and agreed-upon specifications. This framework facilitates interoperability, the development and use of standard tools, and the application of best practices, and it can help address skills gaps.

2.1. Engineering Methodologies

The EM acts as a backbone for implementing governance methods. SA EM, or with a narrower focus, like Ontology Engineering methodologies, primarily fall under SA

Funded by the European Union

management due to their emphasis on practical implementation and operational aspects of creating, maintaining, and evolving SAs. However, to integrate the governance and management of SAs, a methodological approach to SA engineering has been included as an SA governance component. We illustrate the diversity of EM by showcasing initiatives' decisions to either adopt existing strategies and tools or develop their methods, along with specific lifecycles and related tools. Further details are referenced in FAIR-IMPACT's T4.2, which produced a revised Linked Open Terms (LOT) methodology ("Processes & tools to engineer FAIR semantic artefacts" T4.2 milestone report)[11]. The EM design varies depending on the type of SAs under development, the SAs domain, and the maturity level of infrastructure, tools, and processes in SA management. Initiatives may design methodology and workflows internally, customise all steps based on their requirements and objectives, or adapt existing EM to adhere to best practices and increase interoperability in the development processes.

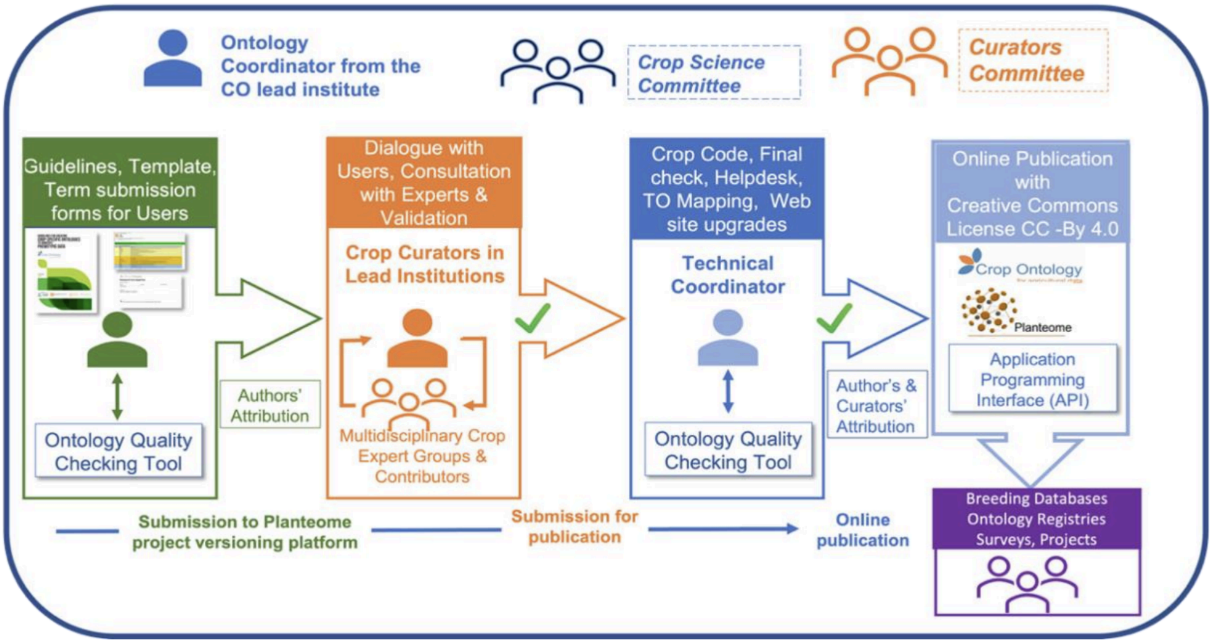

Figure 1. Schema of the workflow between users, curators and technical coordinators within the Crop Ontology project.[12]

The **Crop Ontology project** is one of the initiatives that has designed its methodology for creating domain/crop-specific trait (phenotype) ontologies based on a predefined semantic model. They have established workflows to annotate structured data (CSV and Excel files) using the Crop Ontology generic semantic model and publish them on a searchable library[13]. Quality and validation checker tools and submission processes are embedded in the workflows to semi-automate the development process. Fig. 1 represents the roles and workflow for creating a crop-specific ontology and submitting terms.

[11] https://doi.org/10.5281/zenodo.10551054
[12] https://hdl.handle.net/10568/118001
[13] https://cropontology.org/

Funded by the European Union

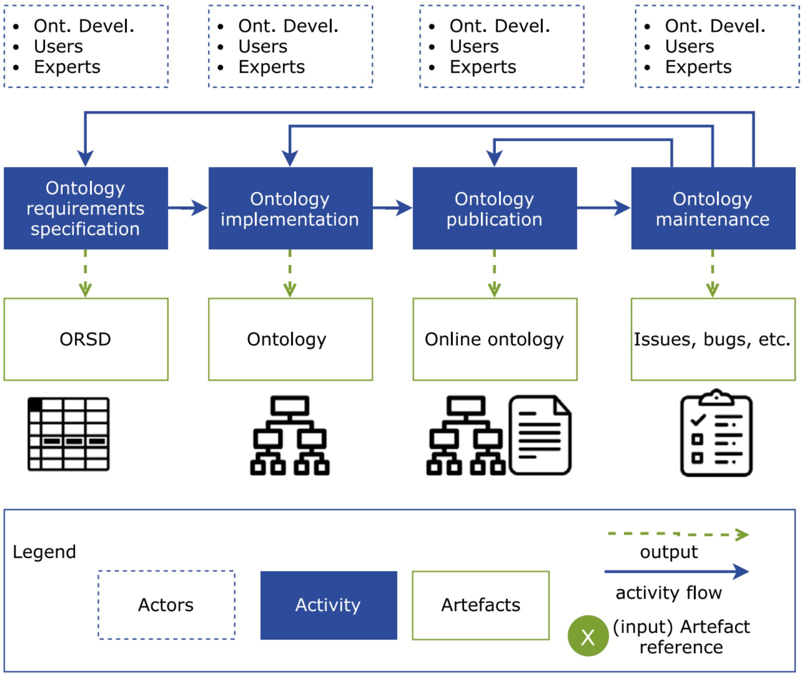

Figure 2. Linked Open Terms (LOT) Methodology.[14]

In contrast, the **BASF** organisation has adopted the LOT methodology[15] (Fig. 2), a lightweight methodology for developing ontologies and vocabularies. The LOT methodology defines iterations over a basic workflow composed of the following activities: ontological requirements specification, ontology implementation, ontology publication, and Ontology maintenance. This methodology emphasises compatibility with software development techniques, increases reusability, and focuses on publication and online ontology.

Another example of an initiative using its strategies to address EM is the Smart Applications REFerence (**SAREF**) set of ontologies developed by ETSI. They have defined a project-oriented workflow (Fig. 3 & 4) to create new versions of the SAREF ontology and its extensions. Instead of focusing on specific sequential phases of ontology development, they have designed a workflow articulated around the use of issues in the corresponding SAREF project issue tracker on the ETSI public forge[13], as it *"enables not only to have a single point of interaction for development but also to keep track of the development activity and discussions"* [13].

[14] https://lot.linkeddata.es/
[15] https://doi.org/10.1016/j.engappai.2022.104755

Funded by the European Union

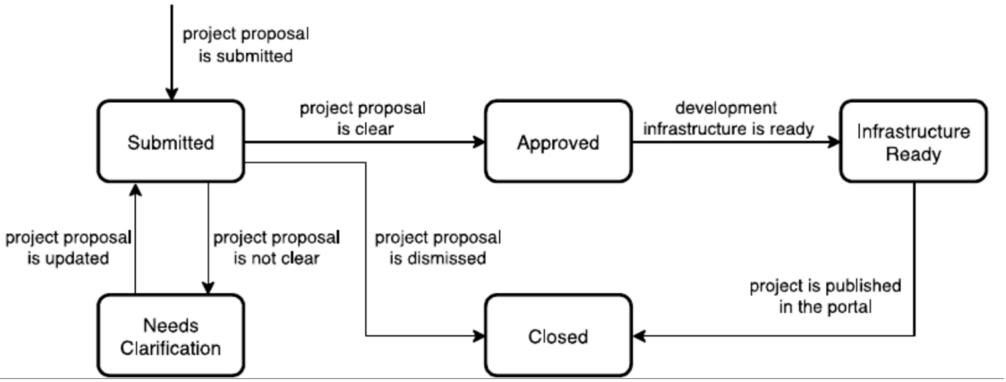

Figure 3. Different states of a new SAREF project version proposal and the transitions among them.[16]

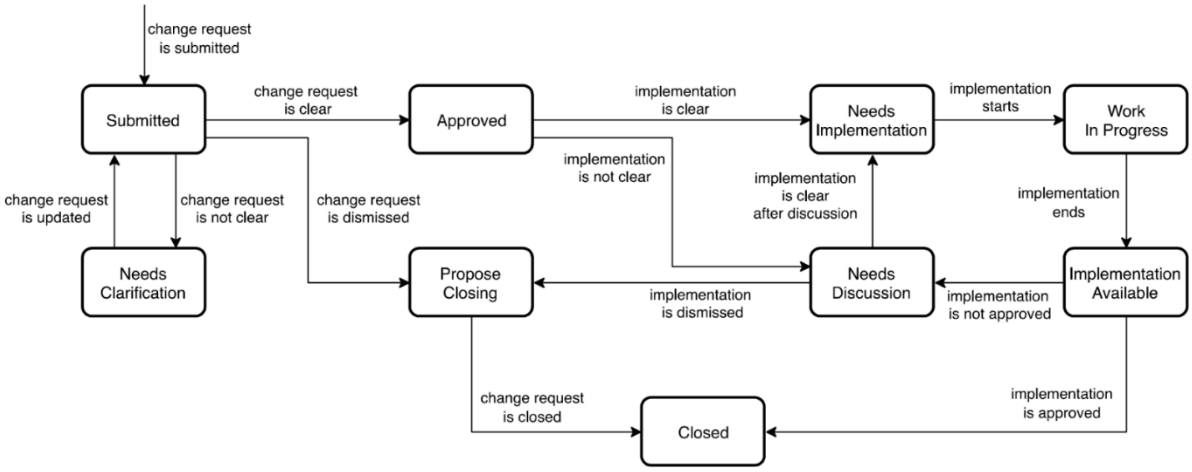

Figure 4 Different states of SAREF project version development and the transitions among them. [13]

In the biological and biomedical domain, one of the most prominent initiatives, **OBO Foundry**, promotes a practical approach by offering centralised infrastructure through the Ontology Development Kit[17] for managing the ontology lifecycle. It comprises two architectural components:

- A toolbox containing all the necessary tools for ontology development, from Unix command-line tools like rsync and git to specialised ontology pipeline programs like ROBOT. The ROBOT source code consists of 'robot-core' and 'robot-command'.
    - 'robot-core' is a library supporting everyday ontology development tasks, referred to as "operations" (such as Convert, Reasoning, Extract, Query, Report, etc.).
    - 'robot-command' provides a command-line interface divided into "commands", each of which wraps a 'robot-core' operation.
- A set of executable ontology-engineering workflows delivered as a directory of scripts, build rules, and source files. Examples of these workflows include initialisation and

[16] https://www.etsi.org/deliver/etsi_ts/103600_103699/103673/01.01.01_60/ts_103673v010101p.pdf
[17] https://doi.org/10.1093/database/baac087

update workflows, editors' workflows, quality control and continuous integration workflows, release workflows, and workflows for importing and reusing existing ontologies.

However, in many communities of practice, initiatives need to follow engineering methodologies explicitly. Within our reviewed communities, for instance, the **Food and Agriculture Organization (FAO)**, the owner of **AGROVOC** and **INRAE**, the owner of **INRAE Thesaurus,** approach the SA development and curation by leveraging common editorial collaboration platforms (e.g., VocBench[18] and Protégé[19]) to enable decentralised resource management. However, they do not report relying on a specific established EM. Editorial groups usually schedule regular updates and releases continuously. The SAs result from these processes is generally deployed in systems like SKOSMOS or published in semantic artefact catalogues (e.g., AgroPortal[20], EcoPortal[21], etc.) to ensure the accessibility and long-term preservation of resources. In other cases, in addition to the previous strategy by FAO and INRAE, initiatives develop extra tools in a workflow manner to address a particular part of the SA lifecycle. For example, **EMBL-EBI,** as an owner of **the Experimental Factor Ontology (EFO)** owner, created workflows for receiving new term requests and implementing term changes. Another example is **NFDI4Biodiv,** which designs standard pipelines to transform taxonomies into OWL products.

Organisations like the **International Virtual Observatory Alliance (IVOA)**, which deals with weak semantic types (e.g., controlled lists and schemas), employ document-oriented processes to build consensus within the community around virtual observatory technology. IVOA working drafts become recommendations by following this process. Fig. 5 shows the labels that describe increasing levels of maturity and consensus in the standards process.

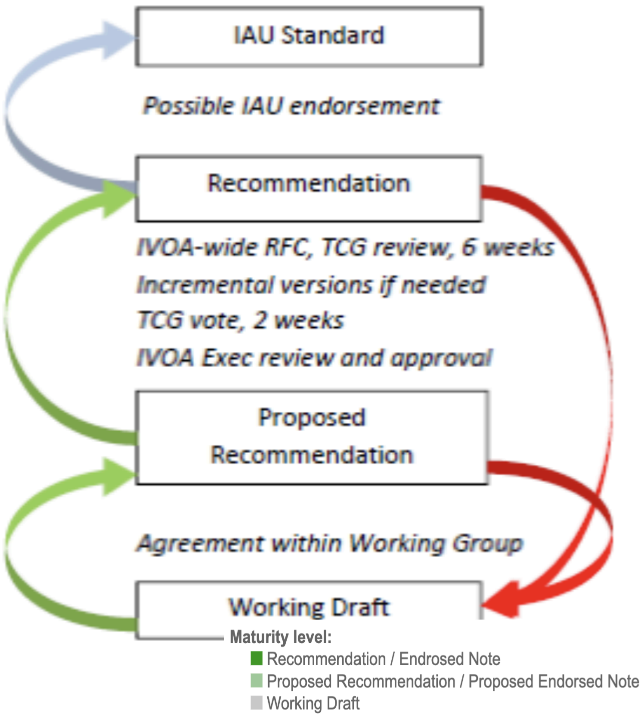

Figure 5. IVOA document promotion process. [22]

[18] https://vocbench.uniroma2.it/
[19] https://protege.stanford.edu/
[20] https://agroportal.lirmm.fr/
[21] https://ecoportal.lifewatch.eu/
[22] https://www.ivoa.net/documents/DocStd/20170517/

Funded by the European Union

2.2. Organisational Structure

This section focuses on characterising the Organisational Structure component of SA governance. We will explore and identify (1) the stakeholders involved in the decision-making process, (2) the framework of decision types and accountabilities associated with SA governance, and (3) the operational mechanisms for disseminating decisions among various stakeholders.

The first part provides information on the composition and structure of groups involved in different governance decision domains. The second part outlines the types of decisions made in areas of responsibility, and the third part examines the operational procedures that link decision makers with task supervisors. By processing this information, we can get a comprehensive overview of the crucial decisions and accountabilities involved in the SA governance and how information is disseminated among stakeholders to ensure effective decision implementation.

To refine our analysis of the roles and responsibilities of each actor involved in SA governance among the surveyed communities, we utilised additional resources referenced during the FAIR-IMPACT Semantic Artefact governance workshop and T4.1's milestone report produced after this workshop[23] or other reports/documents available online and referred by the survey communities. All the additional resources collected and used for the component's analysis are listed in the "Appendices" section. For convenience, these have been split into three appendices:

- Appendix A for the Crop Ontology project (CROP), OBO Foundry and BASF,
- Appendix B for INRAE Vocabularies, IVOA and NFDI4Biodiv,
- Appendix C for EMBL-EBI, SAREF and AGROVOC communities.

Although additional resources are available online, only the descriptions collected for BASF (Appendix B) and EMBL-EBI (Appendix C) are limited to the actors' names assigned to their activities. This means that most of the communities surveyed (7 out of 9) provide substantial information on their stakeholders' roles and accountabilities, allowing for a representative analysis.

The first analysis of the subcomponent of the Organisational Structure of the SA governance lies in defining the distribution of roles and tasks among stakeholders. To achieve this, we distribute the stakeholders of each community, collected in Appendixes A, B, and C, according to their affiliation, either to a group or to a single individual of actors associated with a specific activity (Table 1).

Table 1. Distribution of stakeholders of each community based on their affiliation to a specific group or an individual activity. Italic is used to represent the stakeholders' names assigned by each community, and underlines represent each group name.

[23] https://doi.org/10.5281/zenodo.10287011

Funded by the European Union

Community	Groups	Individuals
CROP	-Committees: *Curators Committee; Strategy Committee; Scientific Advisory Committee* -Team: *CO team (Technical and Ontology coordinators)*	*CO management institute*
OBO Foundry	-Committees: *Operations Committee (3 working groups); Code of Conduct Committee* -Working groups *(WG): Editorial WG; Technical WG; Outreach WG*	-
BASF	-Team: *Data semantic team*	*Ontology developer; Ontology owner; Ontology curator; Domain Expert; End-user*
INRAE Voc	-Committees: *Editorial Committee* -Groups: *Tasks groups (TGs)*	*SA authors, External experts*
IVOA	-Committees: *Executive Committee; Standing Committee on Science Priorities (CSP); Standing Committee on Standards & Processes (SCSP); IUA-IVOA Liaison Committee (IILG)* -Groups: *Technical Coordination Group (TCG); Working groups (WGs); Interest group (IGs); Media Group*	-
NFDI4Biodiv	-Committee: *Taxonomies Editorial Committee* -Board: *Taxonomies Editorial Board* -Groups: *Tasks groups*	*Ontology managers, Experts in the biodiversity domain*
EMBL-EBI	-Team: *Samples, phenotypes and ontology (SPOT) team:* *-Ontology Application team* *-Semantic data integration service team*	*Project leaders; Managers; Coordinators; Editors; Developers; Curators; Data engineers*
SAREF	-Committee: *ETSI SmartM2M Technical Committee* -Board: *Steering Board members; Technical Board members*	*Project leader; Ontology developer; Contributor; Ontology user*
AGROVOC	-Board: *Core Board* -Team: *AGROVOC team*	*Manager; Curator; Technical lead; Technical support; Communication support; Editors, Development, Curators; Experts*

OBO Foundry and **IVOA** are the only two surveyed communities that have entirely centralised their stakeholder activities within formal groups. A significant proportion of the communities (5 out of 9) have adopted a mixed stakeholder organisation, consisting of formal groups of authorities and individuals assigned specific tasks. This is the case for **CROP**, **INRAE Voc**, **NFDI4Biodiv** and **AGROVOC**. For **BASF** and **EMBL-EBI** communities, we have yet to find a formal organisation of the stakeholders involved in the SA governance. **BASF** and **EMBL-EBI** are considered task-centred communities. Indeed, while a "*Data semantic team*" exists within **BASF**, stakeholders' roles are distributed according to individual responsibilities, such as the ontology owner, developers, ontology curator, domain expert, and end-user. Although the EMBL-EBI institute has a team called "Samples, Phenotypes and Ontology (SPOT)", which is further divided into two sub-teams, the activities depend on the service provided; these activities do not appear to be related to specific stakeholder responsibilities.

Thus, we identified six types of formal groups: Committees, Boards, Teams, Working groups (WGs), Interest groups (IGs), and Task Groups (TGs). Additionally, we categorised eight individual task activities assigned to specific actions, including domain expertise, management, coordination, project investigation, ontology development, curation, use, and contribution. While the names assigned either to the stakeholder groups or to the individual task are highly heterogeneous depending on the communities, we can deduce that the establishment of a structured framework for a cohort of stakeholders engaged in SA governance appears essential, given that nearly all communities have organised their stakeholders within at least one formal group, and supplement them with specific task assignments which mainly involve specific SA activities and coordination.

The second subcomponent of the Organisational Structure in SA governance outlines the broad spectrum of accountabilities assigned to stakeholders.

Understanding this subcomponent necessitates examining the responsibilities defined within each formal group or individual stakeholder. Indeed, Table 1 demonstrates that diverse designations exist within the same formal group that do not clarify specific accountabilities. This underscores the importance of delineating decisions and responsibilities assigned to each structured group and individual task. To address this issue, we identified and compiled stakeholders' responsibilities by domain of activity (Table 2).

A domain of activity is defined when the same task or responsibility is observed in at least two communities. Thus, we identified eight domains of activity in SA governance:

- Editorial: draft policies and guidelines and provide recommendations;
- Expertise: advice on domain-specific accuracy;
- Outreach: maintain a relationship between the community and the end user or SA owner;
- Coordination: ensure governance policy consistency by maintaining consistent standards and technical compliance;
- Technical: maintain the infrastructure, website, and SA tools;
- Project management: ensure the disposal and following the implementation of recommendations within the project;
- SA curation: ensure the SA's semantic and technical compliance and manage versioning;
- SA development: involve the SA development cycle and release.

Funded by
the European Union

*Table 2. Distribution of stakeholder names assigned by each surveyed community according to their domain of activity. Italics represent the name of the stakeholder group or individual assigned by each community, and * identify stakeholders who do not have descriptions of their accountabilities.*

Domain of activity / **Communities surveyed**	**Editorial**	**Expertise**	**Outreach**	**Coordination**	**Technical**	**Project management**	**SA curation**	**SA development**
CROP	*Strategy committee; Scientific committee*	*Scientific Committee*	*Ontology coordinator*	*CO team (Technical and ontology coordinator)*	*Technical coordinator*	*CO team; CO management institute*	*Curator committee; Ontology coordinator*	*CO team*
OBO Foundry	*Editorial Committee*	-	*Outreach WG*	*Technical WG; Editorial WG*	*Technical WG*	-	*Technical WG*	*Technical WG*
BASF	-	*Domain experts**	-	-	*Developers**	*Ontology owner**	*Ontology curators**	*Developers**
INRAE Voc	*Editorial Committee*	*External experts*	-	-	-	*SA authors**	*Editorial committee; experts*	*Editorial Committee*
IVOA	*WGs; Executive committee, IGs*	*CSP committee*	*CSP committee; Media Group*	*Executive committee; TCG; SCSP committee*	-	-	-	*TCG*
NFDI4Biodiv	*Tasks groups: Taxonomies editorial board*	*Task groups; Experts*	-	*Taxonomies editorial committees*	-	-	*Experts*	*Ontology manager*
EMBL-EBI	*Ontology editorial lead**	*Scientific curator*; scientific programme manager**	-	*Coordinators* (Catalogue*; ontology project coordinator)**	*Software developers*; ontology tools developer*; ontology developer*; web developer**	*Project leaders**	*Curators**	*Ontology developers*; semantic data engineer**
SAREF	*ETSI SmartM2M Technical Committee*	*Expert**	-	*Ontology developer; Steering board members*	*Technical board*	*Project leaders, Contributors*	*Steering board members*	*Project lead; ontology developer*
AGROVOC	*Core board*	*External experts, ontology editors*	*AGROVOC team (Communication support)*	*Core board*	*Technical Board members*	*Editors*	*Agrovoc team (curator); experts*	*Editors: AGROVOC team (Technical support and manager)*

Funded by the European Union

Almost all communities (8 out of 9), except for **BASF**, provide detailed descriptions of stakeholders' responsibilities in editorial activities. While this activity is always organised among stakeholders, only **EMBL-EBI** assigns individual stakeholders to this role. Expertise activity oversight is provided either by external actors, as in the cases of **INRAE Voc** and **AGROVOC**, or by dedicated groups within the community, such as in **CROP**, **IVOA**, **NFDI4Biodiv**, and **EMBL-EBI**. For **BASF** and **SAREF**, the existence of this stakeholder role is confirmed, but no further details are available. For outreach, **OBO Foundry**, **IVOA**, and **AGROVOC** have dedicated groups for this task. In contrast, **CROP** manages this responsibility by a non-specific group already assigned to other activity domains. Coordination activity is broadly shared among the surveyed communities (7 out of 9). While **CROP**, **IVOA**, and **NFDI4Biodiv** have formally involved specific groups of stakeholders in this activity, only **EMBL-EBI** assigns this role to individual actors. Coordination can also be managed by stakeholders involved in other activities, as seen in the **OBO Foundry**, **IVOA**, **SAREF**, and **AGROVOC** communities. For technical activities, the **CROP**, **OBO Foundry, SAREF**, and **AGROVOC** communities involve groups of stakeholders. Conversely, **BASF** and **EMBL-EBI** engage individual stakeholders or multiple actors in these tasks. Project management is present in 6 out of 9 communities (**CROP**, **BASF**, **INRAE Voc**, **EMBL-EBI**, **SAREF**, **AGROVOC**). It is primarily fulfilled by SA authors or owners who are part of the community's stakeholders. Finally, the last two domains of activity are SA curation and SA development: Curation is explicitly specified by four of the nine communities (**CROP**, **BASF**, **EMBL-EBI**, **AGROVOC**) and usually requires expert opinions for **INRAE Voc**, **NFDI4Biodiv**, **SAREF**, and **AGROVOC** communities. Development activities can be assigned to specific individual actors, such as **BASF**, **EMBL-EBI**, **SAREF,** and **NFDI4Biodiv**. However, in **CROP**, **OBO Foundry**, **INRAE Voc**, **IVOA**, and **AGROVOC**, this activity is mainly assigned to stakeholders already involved in other domains of activity.

To summarise, we identified eight domains of activity involved in SA governance, which enabled us to design the SA governance framework. The significance of each domain can be assessed by the number of communities that have assigned stakeholders to manage each activity (see Box 1 below). SA development seems to be the most significant domain, with all communities assigning stakeholder accountability to this area. Editorial, curation, and expertise are the second most essential domains, with 8 out of 9 communities assigning stakeholder accountability to these areas. The third most significant domain involves the coordination between decision-makers and task-responsible actors. The fourth most important domain, tied in importance, involves technical and project management aspects.

Box 1. Domains of activity involved in SA governance for each community.

1-SA development (9) (CROP; OBO Foundry; BASF; INRAE Voc, IVOA, NFDI4Biodiv; EMBL-EBI, SAREF, AGROVOC)
2-SA curation (8) (CROP, OBO Foundry; BASF; INRAE Voc; EMBL-EBI; SAREF; AGROVOC)
3-Editorial (8) (CROP, OBO Foundry, INRAE Voc, IVOA, NFDI4Biodiv, EMBL-EBI, AGROVOC)
4-Expertise (8) (CROP; BASF; INRAE Voc, IVOA; NFDI4Biodiv; EMBL-EBI; SAREF; AGROVOC)
5-Coordination (7) (CROP; OBO Foundry; IVOA; NFDI4Biodiv; EMBL-EBI; SAREF; AGROVOC)
6-Project management (6) (CROP; BASF; SA authors; EMBl-EBI; SAREF; AGROVOC)

7-Technical (6) (CROP, OBO Foundry, BASF; EMBL-EBI; SAREF; AGROVOC)
8-Outreach (4) (CROP; OBO Foundry; BASF; IVOA)

Finally, outreach is the domain with the least involvement, as only 3 out of 9 communities have assigned stakeholders to this area, reflecting its relatively lower importance within SA governance. While the SA development domain of activity is present in SA governance for all the communities, this responsibility is mainly assigned to polyvalent or individual stakeholders. However, the editorial and SA curation activities are allocated to specific groups of actors, highlighting the importance of collectively managing those.

To finalise our analysis of the Organisational Structure of stakeholders involved in SA governance, we examine how information is disseminated among different actors to ensure effective decision implementation. To achieve this, we categorise the actors by their domains of responsibility (Table 3):

- Decision making, defined as the responsibility being of to establishing best practices to ensure policy accuracy ;
- Executive processes, defined as the responsibility to carry out the tasks and inputs for the effective management and implementation of the SA;
- Orchestration defined as the responsibility to ensure the effective compliance between the policy rules and their implementation. Typically, orchestration comes into the picture to verify the correctness of the execution processes.

Table 3. Distribution of stakeholder names assigned by each community according to their roles in decision making, executive processes and orchestration.

Domain of responsibility / Communities surveyed	Decision making	Executive processes		Orchestration	
	Editorial	SA curation	SA development	Coordination	Procedures
CROP	*Strategy committee; Scientific committee*	*Curator committee; Ontology coordinator*	*CO team*	*CO team (Technical and ontology coordinator)*	-
OBO Foundry	*Editorial Committee*	*Technical WG*	*Technical WG*	*Technical WG; Editorial WG*	*Standard Operating Procedures (SOPs)*
BASF*	-	Ontology curators*	Developers*	-	RACI
INRAE Voc	*Editorial Committee*	*Editorial committee; experts*	*Editorial Committee*	-	-
IVOA	*WGs; IGs; Executive committee*	-	*TCG*	*Executive committee; TCG; SCSP committee*	*Technical specifications*
NFDI4Biodiv	*Tasks groups: Taxonomies editorial board*	*Experts*	*Ontology manager*	*Taxonomies editorial committees*	-
EMBL-EBI*	Ontology editorial lead*	Curators*	Ontology developers; semantic data engineer*	Coordinators (Catalog; ontology project coordinator)*	-

Funded by the European Union

SAREF	*ETSI SmartM2M Technical Committee*	*Steering board members*	*Project lead; ontology developer*	*Ontology developer; Steering board members*	*Technical specifications*
AGROVOC	*Core board*	*Agrovoc team (curator); experts*	*Editors: AGROVOC team (Technical support and manager)*	*Core board*	*Editorial guidelines; Semantic data interoperability*

For **NFDI4Biodiv** and **EMBL-EBI**, stakeholders are divided into three distinct roles, with specific actors or groups involved in decision making, executive processes, or orchestration. In **CROP** and **SAREF**, a particular group of stakeholders handles decision making, while a shared group manages executive processes and orchestration roles. For **OBO Foundry** and **IVOA**, actors involved in decision making also participate in orchestration activities, while those responsible for orchestration are as well involved in executive processes. **AGROVOC** has a similar organisation, except that executive responsibilities are specifically assigned. **INRAE Voc** is the only community where stakeholders are involved in both decision making and executive processes. Finally, we found that 4 out of 9 communities have formal procedures that precisely describe the execution process, as seen in **OBO Foundry**, **IVOA**, **SAREF**, and **AGROVOC**. **BASF** has only defined roles and responsibilities of the stakeholders using the RACI (Responsible, Accountable, Consulted, and Informed) responsibility assignment matrix.

To summarise this third component of the Organisational structure of SA governance, we found that **CROP, NFDI4Biodiv, EMBL-EBI**, and **SAREF** have decentralised roles and responsibilities, with stakeholders involved only in decision making. In contrast, **INRAE Voc, OBO Foundry, IVOA,** and **AGROVOC** centralised decision makers who also hold other responsibilities, facilitating the orchestration and implementation of policies and guidelines, but are also usually supplemented by formal procedure documents such as SOPs or technical procedures.

To conclude this section on the analysis of the organisational structure of SA governance, we could define the stakeholders' roles related to the development, implementation, enforcement, and facilitation of rules of engagement and the governance framework concerning activities related to SA. Based on our analysis, we identified three types of actors involved in SA governance to varying degrees of involvement (Table 4):

- (1) Governing stakeholders, who are responsible for editorial tasks, coordination, and SA activities such as development and curation;
- (2) Intermediate participants, who closely interact with the governing stakeholders and are involved in domain expertise, service enablement, and community support;
- (3) Participatory actors initiate the SA transaction by providing initial content. They can be project managers who are primarily considered SA holders or SA owners, authors, and contributors.

Table 4. Synthesised roles (type of actors) involved in the SA governance and relation to the domain of activity.

Funded by the European Union

Types of actors	Roles	Descriptions	Domain of activity
Governing stakeholders	**Decision makers**	Stakeholder(s) that is(are) accountable for the governance of a particular governance framework. e.g. draft policy and ensure their agreement with the principles	**Editorial**
	Executors	Stakeholder(s) that provide(s) a service enabling or facilitating trustworthy SA transactions for the participants. e.g. allow connections with other SA and provide technical support	**SA development & curation**
	Orchestrators	Stakeholder(s) that orchestrate(s) the SA ecosystem and ensure(s) its functioning and that participants abide by the agreed standard rules and principles. e.g. manage continuity of SA implementation and take care of onboarding new participants.	**Coordination**
Intermediate participants	**Specialists (semantic or domain)**	Actor(s) that advise(s) on technology and semantic standards in a specific domain to ensure accuracy and pertinence of SA to the relevant discipline.	**Expertise**
	Infrastructures providers	The party that maintains technical infrastructure and the services they provide through connection of multiple services, e.g. maintaining website and SA tools	**Technical**
	Outreach coordinators	The party that animates and provides support to their community to onboard and increase awareness in the ecosystem, e.g. manage communication channels and organise events	**Outreach**
Participatory actors	**Managers**	Party that, in the context of a specific SA transaction, can have a right or duty to ensure SA quality and compliance technically	**Project management**
	Contributors	The participant that is identified as a legal person who is involved in multiple levels of mechanisms of giving consents, e.g. SA owers or SA authors	**SA development & curation**

2.3. Governance Framework

The **Governance Framework** is the third component within our interpretation of SA governance; it provides mechanisms for management by establishing high-level rules and standards. In the following paragraphs, we analyse the practices of our survey communities.

As one of the leading organisations in governance and management of ontologies in the biological and biomedical domain, OBO Foundry provides principles for integrating and harmonising open and FAIR ontologies. Fifteen principles have been defined, outlining recommendations, best practices, and standards for developing and reusing ontologies. "*Originally, these principles were not precisely formulated, and interpretation was subjective. Consequently, they have been formally encoded as operational rules to enable the establishment of automated validation checks*".[24] The OBO Dashboard[25] (Fig. 6) has been developed to utilise these operational rules to provide a set of automated tests that guarantee a minimum level of compliance with OBO Principles and best practices. Through this approach, OBO Foundry has been structured and assessed based on criteria that help to improve overall quality and interoperability, which is crucial for making data FAIR. However, a limitation of OBO Foundry is the lack of a well-structured framework to identify the main components of their governing strategy explicitly.

[24] https://doi.org/10.1093/database/baab069
[25] https://dashboard.obofoundry.org/

Funded by the European Union

Furthermore, *only some of the OBO principles are expressed at the same level of maturity or granularity*[26], making it challenging for other initiatives attempting to implement similar principles. In the context of ontology development standards, the URI/Identifier Space principle (P3) exemplifies high maturity with its well-established requirement for each ontology to possess a unique IRI (Permanent URL). This principle is widely adopted and ensures clarity in ontology identification. However, the implementation details, such as best practices for managing IRIs across various systems, may vary. This variability in technical specifics introduces granularity differences in how the principle is operationalised and applied across different ontology projects.

Ontology (click for details)	Open	Format	URIs	Versioning	Scope	Definitions	Relations	Documented	Users	Authority	Naming	Maintained	Responsivene	ROBOT Report
apollo_sv	✔	✔	✔	✔	✔	⚠	i	✔	✔	✔	✔	✔	✔	⚠
cl	✔	✔	✔	✔	✔	⚠	i	✔	✔	✔	✔	✔	✔	⚠
disdriv	✔	✔	✔	✔	✔	⚠	✔	✔	✔	✔	✔	✔	✔	⚠
doid	✔	✔	✔	✔	✔	⚠	i	✔	✔	✔	✔	✔	✔	⚠
dpo	✔	✔	✔	✔	✔	⚠	i	✔	✔	✔	✔	✔	✔	⚠
eco	✔	✔	✔	✔	✔	✔	i	✔	✔	✔	✔	✔	✔	⚠
envo	✔	✔	✔	✔	✔	⚠	i	✔	✔	✔	✔	✔	✔	⚠
fbcv	✔	✔	✔	✔	✔	⚠	i	✔	✔	✔	✔	✔	✔	⚠
fbdv	✔	✔	✔	✔	✔	✔	i	✔	✔	✔	✔	✔	✔	✔
go	✔	✔	✔	✔	✔	⚠	i	✔	✔	✔	✔	✔	✔	⚠
hancestro	✔	✔	✔	✔	✔	⚠	i	✔	✔	✔	✔	✔	✔	⚠

Figure 6. A fundamental analysis provided by the OBO Foundry dashboard, the left column lists the ontologies under the test, and each row represents the result of the test for each ontology concerning OBO principles. [20]

The need for a holistic approach to ontology governance has led **BASF company** to create a Governance Operational Model (GOMO)[27] for BASF ontologies, which offers a framework to coordinate and control the development, management, and curation of ontologies throughout the stages of the ontology lifecycle. GOMO consists of four main components: principles, standards and associate quality assurance method, along with best practices, training and outreach (Fig. 7). It results from a collaborative effort between industry and academia in the semantic web field, integrating concepts from data governance, FAIR principles, schema.org, and the OBO and IOF Foundries.

[26] https://doi.org/10.1186/s13326-023-00286-8
[27] https://doi.org/10.5281/zenodo.7007495

Funded by
the European Union

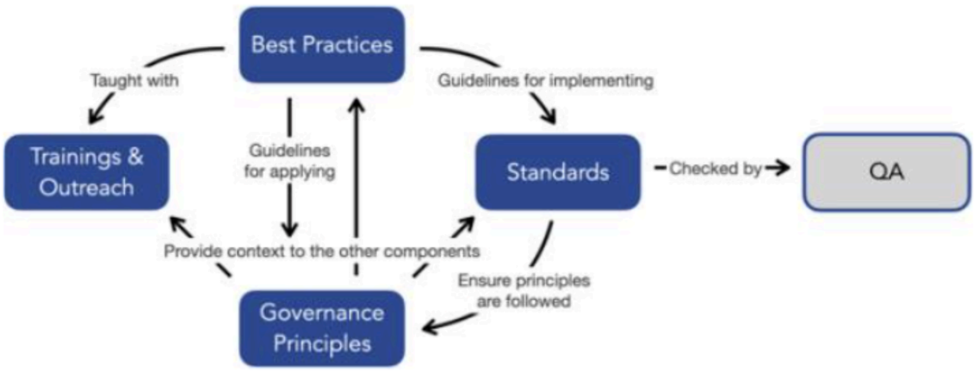

Figure 7. GOMO components and interlinked interactions [25]

The following definitions of components have been extracted from the paper presenting GOMO[25]:

- *“**Governance principles** define a set of high-level fundamental rules on how the organisation should develop, publish, maintain, and consume ontologies.*

- ***Standards** are a set of agreed-upon specifications to be followed during the ontology lifecycle. Each standard is associated with a **quality assurance** method that allows human—or software-based evaluation of its correct implementation.*

- ***Best Practices** Comprise is a set of recommendations and guidelines that explain and illustrate how to follow the principles and implement the standards while providing the user with the background knowledge required to perform specific activities.*

- ***Training and outreach** Comprise a series of interactive, expert-led workshops to explain the concepts and guidelines from best practices and standards with tailored examples for the target audience.”*

BASF currently has defined seven principles for ontology governance, including 1. Availability, 2. Access Rights and Security Policy, 3. persistency, 4. documentation, 5. fairness, 6. modularity, and 7. community.

SAREF also provided four principles: 1. reuse and alignment, 2. modularity, 3. extensibility, and 4. maintainability. These principles are dedicated to the design of the SAREF core ontology and its different extensions.

However, many initiatives have yet to define principles explicitly. Organisations like INRAE, EMBL-EBI, and NFDI4Biodiv have aligned their governance approaches with part of existing principles or methodologies (OBO, FAIR, and LOT) where they are applicable. Others, like IVOA, must have overarching principles but develop extensive specifications (standards) within their domain.

In all cases, these initiatives prioritise the quality of their work by adopting good practices and recommendations. Automated or semi-automated tools (e.g., the trait dictionary quality

Funded by the European Union

checking tool in the Crop Ontology project or O'FAIRe methodology and tool[28] for assessing the level of FAIRness in ontologies) are typically employed in their workflows to validate the quality of the results.

Analysing communities of practice has enabled us to identify critical components of the governance framework: principles, standards, and quality. We have conceptualised these elements in the following model (Fig. 8) to provide a harmonised structure for governance that will increase interoperability among initiatives that develop and maintain SAs.

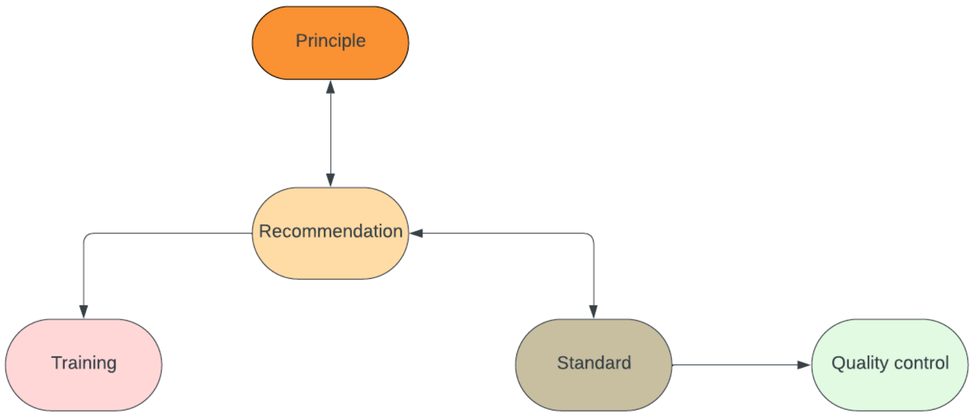

Figure 8. General governance model of the establishment of principles and recommendations within an organisation.

Principles can be simple statements or concrete policies that cover various aspects of the semantic artefact lifecycle (e.g., metadata, versioning, contribution). For example, "*The process of ontology versioning should be encoded by reusing multiple relevant metadata properties from existing vocabularies*" is a fundamental principle (rule) covering the versioning aspect of the SA lifecycle. Each principle should lead to the reuse or creation of specific **Recommendations** that guide the methodological application of the principle. "*About versioning ontologies or any digital objects with clear semantics*"[29] (a research article produced in FAIR-IMPACT T4.2) discusses versioning for SAs by describing methods for encoding versioning and other relevant information in metadata properties. This article provides valuable recommendations regarding the versioning principle we have defined.

Training is a component connected to recommendations. It highlights the need for initiatives to train different stakeholders by explaining the concepts, guidelines, and usage of these recommendations to apply the associated principles effectively. Another component is **Standards**, which are established norms or requirements about processes or practices. They should be extracted from recommendations and developed by consensus, serving as definitions to ensure quality, interoperability, and efficiency. For example, in the context of SA versioning, providing a separate Internationalised Resource Identifier (IRI) for identifying

[28] https://doi.org/10.1504/IJMSO.2022.131133
[25] https://doi.org/10.5281/zenodo.7007495
[29] doi: 10.4126/FRL01-006444994

Funded by the European Union

each version of an ontology (version IRI) could be an internal standard, along with the semantic versioning specification (SemVer), which is a system for version numbers. Web ontology language (OWL) could be another standard for providing properties for establishing the versioning process (e.g. *owl:versionInfo* or *owl:verionIRI*). Establishing these standards enables human or software-based evaluation of their correct implementation through a **Quality control** process. Continuing with the example, an automated test could be defined to query the semantic artefact to retrieve the version IRI. If found, this IRI is compared to a regex pattern to determine if it adheres to SemVer. For instance, if an ontology is annotated with the IRI “https://w3id.org/example/1.0.0” using the *owl:versionIRI* OWL property, it will successfully pass the test and validate its quality based on versioning standards.

3. Process followed

Once we analysed the three main components of SA governance and outlined how communities manage them, we developed a methodology to formalise SA governance models. In the first step, we aggregated practices associated with different aspects of the SA lifecycle (e.g., versioning, documentation, metadata, etc.). In the second step, we identified potential target groups for the model. We then developed SA governance models for these groups by integrating practices related to Engineering Methodology and SA lifecycle (Section 2.1 and result of step 1: Table 6), Organisational Structure (Section 2.2), and Governance Framework (Section 2.3). Table 7 presents these models, structured according to the approaches reviewed by the communities under our scrutiny.

Step 1) Our strategy is to examine the principles of each initiative to pinpoint the general aspects of the SA lifecycle that each principle aims to cover. The OBO Foundry principles are our primary source for this step due to their clear documentation and well-established reputation in the academic discipline , making them highly reliable and comprehensive. Their extensive application across various ontologies provide a solid foundation for our analysis.

Principles from BASF serve as a secondary resource, particularly the fifth principle dedicated to developing and maintaining FAIR ontologies. These principles are valued for their operational approach to ontology governance and their innovative contributions to the field. FAIR ontologies are essential for ensuring that data and information are Findable, Accessible, Interoperable, and Reusable, which enhances their utility and integration across diverse systems.

Additionally, we consider other resources such as SAREF design principles, methods from LOT, and ontology governance frameworks to address potential gaps in the primary sources. These supplementary materials help ensure a comprehensive evaluation by incorporating diverse perspectives.

Step 1. a) Align the BASF and OBO principles to identify gaps in each organisation’s policy for managing SAs. These gaps, shown in grey boxes in Appendix D, indicate areas of the SA lifecycle not covered by the principles of one of these initiatives.

Funded by the European Union

Step 1. b) We merged these principles into a single list to address the gaps by combining overlapping principles and retaining the rest from both organisations. We then normalised them to identify the general aspects of the SA lifecycle that each principle aims to cover. The result is shown in Table 5, where each element is tagged with the principle from which it was initially derived.

Table 5. Aspects of SA life cycle result of step 1.b.

Availability (licencing) **BASF(P1)-OBO(P1)**	Access rights and security policy **BASF(P2)-OBO(P1)**	Scope **OBO(P5)**	Community **BASF(P7)**	Documentation **BASF(4)-OBO(8)**	Persistency (Versioning and Deprecation) **BASF(P2)-OBO(P4)**
Unique identification for (meta)data **BASF(P5)-OBO(P3)**	Relevant attributes, metadata and provenance **BASF(P5)-OBO(P6,P9)**	(Meta)data access and semantic repository **BASF(P5)**	information model **BASF(P5)-OBO(P2)**	(Meta)data reuse **BASF(P5)-OBO(P7)**	SA visualisation **Other sources**
Language **Other source**	Naming Conventions **OBO(P12)**	Commitment to collaboration and contribution **OBO(P10,P11,P16)**	Channels for communications **OBO(P13,P20)**	Modularity and extensibility **BASF(6)-SAREF(P2, P3)**	

Step 1. c) We collected various approaches from the SA governance workshop and related resources to address different aspects of the SA lifecycle (Table 6). Each element may be further divided into several categories to provide more detail.

Table 6. Aggregation of all the practices from the SA governance workshop (Sept. 2023) for different domains within the SA lifecycle.

Governance aspects	Implementations observed (workshop)
Availability (licencing) ● open licence ● limited licence ● closed **Community** ● SA is not assigned to a particular community ● SA receive support from a particular community ● SA has been derived from a specific community ● SA is a cross-community/domain	**Access rights and security policy** ● role-based access control (e.g. An SA editor role may have permissions to modify and update the ontology, while a viewer role may only have read-only access.) ● user attribute-based and resource attribute-based access control (access to an SA might be restricted based on user attributes like department or job title.) ● ad hoc discretionary access control (e.g. a researcher with appropriate permissions on the system who developed an ontology may grant access to specific colleagues or collaborators while restricting access to others.)
Persistency (Versioning and Deprecation) ● Range: ○ not defined policy or commitment ○ versioning and depreciation conventions are defined only for the SA ○ versioning and depreciation conventions are also defined for other documents (e.g., requirements, conceptualisation, tests, etc.). ● Version and deprecation granularity: ○ SA level ○ SA element (term) level ○ both ○ neither ● Mechanism for versioning: ○ manual versioning (e.g. defining file naming	**Documentation** ● Data Management Plan: ○ not stated in the documentation ○ not required ○ optional/recommended for projects in the early stages ○ mandatory ● Granularity: ○ SA and SA element-level documentation ○ project level documentation ● Topic: ○ requirements documentation ○ design documentation ○ versioning documentation ○ usage documentation ○ development documentation

<table>
<tr>
<td>

conventions like ontology_v1.0_2024-05-19.owl)
automate versioning (e.g. implementation of version control system like git, relying on repositories versioning system)

Version numbers convention:

defining internal convention
adopting semantic versioning specification (SemVer)

</td>
<td>

maintenance and support documentation
testing documentation
compliance and licensing documentation
metadata and provenance documentation
collaboration and contribution documentation
deployment documentation

Audience:

human-readable documentation
machine-readable documentation

</td>
</tr>
<tr>
<td rowspan="3">
Relevant attributes, metadata and provenance

Metadata standards:

no standard
has its metadata standards (e.g., OBO Metadata Ontology (OMO))
adopted existing control vocabularies and ontologies

Mandatory/recommended and optional metadata (on ontology level and ontology element level)

not defined
defined internally
adopted existing rules and recommendations on minimum metadata (e.g., Minimum Information for the Reporting of an Ontology (MIRO))

Type:

structural and descriptive attributes (class, property, hierarchy, instance, annotation property, labels, text descriptions, definitions, synonyms, comments)
descriptive metadata (abstract, keywords, references, authors and contribution)
administrative metadata (ontology name, title, version information, publisher, project/resource owner, date, licence, documentation, funding information, funding)
structural metadata (analysis, methods, sampling procedure and size, categories and variables)
provenance (general provenance, creation date, authors and contributors, tools and methodologies, source data, change history, modification date, change reasons, latest changes, usage and application context)
access and distribution metadata (publication date, distribution channels like platforms or repositories landing page, access restrictions, downloadable files, API and SPARQL endpoints)

</td>
<td>
Unique identification for (meta)data

Type:

Ontology identifier:

(local) unique identifiers (e.g., UUIDs)
global unique and persistent identifiers (e.g., DOI)

Ontology elements identifier:

internal identifier (e.g., stable URIs)
external identifier

versioned identifier (e.g., version IRI)
specific metadata identifier
data and metadata share the same identifier (metadata embedded with the data)

Management:

centralised management

registry
authority control system

decentralised management (e.g., federated systems)

Audience:

human-readable identifiers (e.g., descriptive strings)
machine-readable identifiers (e.g., UUIDs, hash codes)

Resolvability:

unresolvable identifier
resolvable identifier

</td>
</tr>
<tr>
<td>
SA visualisation

Diagrams:

class hierarchies (e.g., tree diagrams)
entity-relationship diagrams

UML Diagrams

Graph-based visualisation:

node-link diagrams

SPARQL query results visualisation
tabular visualisation:

Tables and spreadsheets

</td>
</tr>
<tr>
<td>
Language

Monolingual
Multilingual

</td>
</tr>
</table>

Funded by
the European Union

<table>
<tr>
<td>
information model

Knowledge representation data model
Resource Description Framework (RDF)
Knowledge representation formalism
Description Logics (DL)
Knowledge representation language

RDF schema (RDFS)
Simple Knowledge Organisation System (SKOS)
web ontology language (OWL)
Unified Medical Language System (UMLS)
OBO language

Formality level

Controlled vocabulary and schema
Terminologies
Taxonomies
Thesaurus
Ontology

Syntax

only serialised in RDF/XML
several serialisations are available (turtle, N-Triples, JSON-LD, N3, etc.)

Knowledge base (graph) compatible product (appropriate modelling with the vision that the SA will finally use to create knowledge graphs)

</td>
<td>
(Meta)data access and semantic repository

Network communication protocols:

only through HTTP(S)
HTTP(S) and SPARQL

Methods to access the ontology:

HTTP(S)
RDF/OWL Dump files
APIs
SPARQL endpoint

Repository type:

Internal to project or organisation.
external

Repository domain:

domain specific (e.g., EcoPortal, BioPortal)
Domain generic (e.g., fairsharing.org)

Repository maturity level:

indexed
no searchable list
simple searchable list
ontology library/catalogue (e.g., Ontology Lookup Service)
SA catalog (e.g., AgroPortal, EcoPortal)

Repository technology:

triple store
property graph (e.g., Neo4j)
relational database

</td>
</tr>
<tr>
<td>
(Meta)data reuse

SA selection:

Strategy:

SA selection by standardisation (e.g., reuse of ontologies like PROV-O or Time Ontology, which are the W3C standard)
SA selection by popularity (reuse of FOAF ontologies)
SA selection by performing analysis

Implementation:

use of SA repository services (e.g., browse or terms search)
use of dedicated tools (e.g., Ontofox)

Policies for reuse:

direct reuse

soft reuse (referring to external SA elements URIs)
hard reuse (import the entire SA)

Indirect reuse:

terms from external SA are reused as a template in the new ontology

hybrid reuse

Ontology integration methods:

Modular composition (allow separation and recombination of different parts of the SA depending on specific needs)
Orthogonality and Merging

Ontology networking:

Cross-product definitions (reuse terms from other SA in cross-product definitions where appropriate)
Conflict Resolution (e.g., addressing overlapping and competing ontologies in the same domain)

</td>
<td>
Scope

SA lacks a statement of the domain or subject matter it intends to cover
The scope of the SA is defined by the repository where the SA is hosted

The domain is selected from predefined categories provided by the repository during the SA submission process

The domain is clearly defined by giving an explicit scope statement

The documentation clearly outlines the boundaries and content of the SA in a dedicated section of the project’s website

Practices:

focused scope (narrow scope to prevent duplication of terms)
handling out-of-scope terms (Import and reuse terms from existing ontologies when required terms fall outside the defined scope)
separate modules for out-of-scope terms (Place out-of-scope terms in separate SA that can be imported or exported as needed)
extensibility to allow further growth of the SA (different stakeholders specialise the SA concepts according to their needs and points of view, add more specific concepts, relationships and hypotheses to refine the general (standard) semantics expressed in the reference SA)
document and respond to community feedback regarding the scope and content of the SA

ensure that generic terms meet the needs of the broader community

</td>
</tr>
</table>

Naming Conventions	Channels for communications

• no establishment of naming conventions
• align naming conventions with existing standards and guidelines (e.g., OBO naming convention) where applicable
• establishment of concrete naming conventions
• Type:
  ◦ domain-specific
  ◦ generic or cross-domain (extendible)
• Target format:
  ◦ OBO
  ◦ OWL
	• None
	• Direct communication:
	  ◦ personal email
	  ◦ project-specific email
	  ◦ Messaging platforms:
	    ▪ instant messaging (e.g., slack, Microsoft Teams)
	    ▪ group chats
	• Public discussion forums:
	  ◦ project/community mailing list
	  ◦ announcement mailing list (for pre-announcing changes and updates)
	  ◦ online forums
	    ▪ project-specific forum
	    ▪ public SA related forum (e.g., BioPortal Forum)
	• Collaborating platform:
	  ◦ Version control repositories:
	    ▪ GitHub/GitLab issues
	    ▪ pull requests and code reviews
	    ▪ GitHub pre-releases
	  ◦ project/community wiki pages
	• Community contribution:
	  ◦ feedback forms
	  ◦ community surveys
	• Specialised communication channels
	  ◦ dedicated support channels (e.g., helpdesk)
	  ◦ Advisory panels (reflect changes in scientific consensus to keep the SA accurate over time):
	    ▪ scientific advisory boards
	    ▪ user advisory panels

Commitment to collaboration and contribution
• Community engagement
  ◦ open forums and discussion boards
  ◦ workshops and meetings
  ◦ contribution guidelines (transparent contribution processes and agreements)
  ◦ training and mentorship programs
• Tools for supporting collaborative ontology development
  ◦ online document editors (e.g., Google Docs)
  ◦ collaborative editing platforms (e.g., web protege)
  ◦ issue tracker
• Contact authority:
  ◦ primary contact person
    ▪ optional
    ▪ mandatory
  ◦ alternative contacts
  ◦ to date, contact information
    ▪ included in metadata embedded with the SA file)
    ▪ submitted in the registry (which will be available in a separate metadata file)	

Step 2) In practice, SA governance is inherently situational, reflecting each organisation's unique characteristics and needs. The approach to management is influenced by several factors, including organisational goals and structure, technological capabilities and infrastructure, regulatory and compliance requirements, data complexity and volume, stakeholders' involvement, and cultural factors. However, despite these differences, it is essential to establish consistent, interoperable principles that can harmonise SA governance across various domains. Such principles provide a common framework that facilitates alignment and integration, ensuring that governance practices can effectively support cross-domain interoperability while accommodating diverse organisational contexts.

Step 2. a) To develop effective models that accommodate a wide variety of initiatives and communities, we compared several factors mentioned in the previous paragraph across nine communities from the SA governance workshop. This comparative analysis led us to identify distinct models, each representing a set of common features and governance practices observed across these communities. These models provide a framework to which different components of SA governance can be applied in a manner that reflects the unique characteristics and needs of each initiative. As a result, we defined the following general models that capture these shared characteristics and governance approaches:

- **1) SA Project-based initiatives** managing single-core SA, such as application ontologies (e.g., **EFO** and **SAREF**) or Linked Open Data sets (e.g., **AGROVOC**). Within

Funded by
the European Union

this category, SA is developed to address the need for standard vocabulary in a target domain. These SA enable semantic interoperability and support the creation of knowledge graphs by providing domain-specific concepts and properties, enhancing data integration between systems that deploy them.

- **2) Distributed SA development and support-based initiatives** support communities in developing and maintaining (FAIR and open) SA. Individuals or groups of researchers independently build and maintain artefacts in a decentralised manner. Organisations such as **INRAE Voc** and **NFDI4Biodiv** operate within this category.
- **3) SA Harmonisation initiatives** that provide principles and guidelines, along with infrastructure, to harmonise the management of SA across communities or organisations. Ontology developers can use the upper-level ontologies, modularisation methods, and other standard workflows and practices established by these initiatives to ensure they develop interoperable SA. **The OBO Foundry** is an example of this category, serving as a pivotal point in using semantic web technologies in biological and biomedical sciences. **BASF** also falls into this category but restricts its developed ontologies to internal organisational use**.**

Step 2. b) Providing concise statements that can be considered good practices and recommendations for SA governance tailored for the target groups. These statements are specified to cover all subcomponents of SA governance that we have defined.

Following this structured methodology, we comprehensively analysed and synthesised the practices, principles, and resources related to SA governance. This approach enabled us to develop tailored SA governance models that address several target groups' unique needs and characteristics. In the next section, we present these models in detail, showcasing the practical outcomes of our process and offering a framework for effective SA governance.

4. Semantic artefact governance models

The governance of SA extends beyond decision making. While the interaction between actors and their accountability for different domains of the SA life cycle is crucial, it must be guided by well-defined rules aligned with the organisation's or community's goals. A holistic view is essential, encompassing all requirements for SA management, from the maturity level of infrastructure and processes to the capabilities of human resources. This comprehensive approach ensures all elements work cohesively to achieve the desired outcomes.

Organisations and communities establish different methods and processes to develop and manage SA per their organisational structure and capabilities. As long as they achieve their goals, their strategy is functional. Thus, many approaches fit all. We aim to address this diversity by proposing SA governance models structured to align with the approaches reviewed by several communities, focusing on three distinct types of initiatives: project-based initiatives, distributed SA development and support initiatives, and SA harmonisation initiatives.

Table 7. FAIR-IMPACT proposed Semantic Artefact governance models for pre-defined target groups.

SA Governance	components	Model #1: SA project-based Initiatives	Model #2: Distributed SA development and support-based Initiatives	Model #3: SA harmonisation Initiatives
Governance Framework	**Principles**	Where applicable, follow existing principles (e.g., OBO, FAIR, LOT, 5 Star Data[30]).	Inform the community about existing principles and provide guidelines for establishing them.	Define clear principles and provide infrastructure to support the community or organisation toward its goals and objectives for managing SA.
	Recommendations	Consider good practice and guidelines.	Guide the community to existing recommendations and guidelines.	Create comprehensive recommendations, implementation guidelines and examples for different aspects of the SA life cycle.
	Standards	Adopt common standards and specifications to ensure a high level of quality.	Set the baseline standards across communities to facilitate the quality evaluation of SA.	Adhere to expected standards and specifications, contribute to them and define new ones as needed.
	Quality control	Adopt software engineering practices (e.g., implementing version control systems) and use ontology validators and evaluators.	Guide the community in using software engineering practices and quality evaluator tools throughout the different stages of the SA lifecycle. Develop pipelines and tests tailored to the specific needs of the community.	Integrate the infrastructure with quality evaluator tools. Develop (openly) available quality control processes, tests, and tools for validating the implementation of principles and standards defined across the community or organisation.
	Training	Relying on training materials available, maybe set up training sessions for primary users (e.g., systems that implemented the SA).	Organise training sessions for the community to educate them about the principles and recommendations followed and how to ensure the quality of SA.	Hold workshops to elaborate on each principle and recommendation within the community or organisation. Engage in long-term training and support activities.
Aspects of SA lifecycle	**Availability (licencing)**	Make an SA available under an open licence.	Inform the community about the benefits of choosing an open licence to maintain their copyright while enabling the public to use and remix their SA.	We are developing guidelines and workflows to support the licensing process. Associate the licensing to a principle.
	Access rights and security policy	Manage the permissions granted to different individuals within the SA lifecycle.	Provide guidelines on methods and tools for implementing a secure access policy.	Establish an access rights management system. Create a document that outlines policies and strategies for a community or organisation to maintain the security of its SA.

[30] https://5stardata.info/en/

Funded by the European Union

Versioning	Identify each version of the SA using a unique IRI.	Guide the community towards relevant practices and encourage them to establish a plan (workflows) for the periodic release of new versions of SA.	Provide a formal versioning process to track different iterations of the SA (and associated documentation), ensuring transparency and effective change management. Provide infrastructure and tools to support these aspects.
Deprecation	Establish a deprecation process at the SA element (term) level and notify primary users of updated or alternative terms upon deprecation.	Guide the community towards relevant practices for deprecation.	Establish a deprecation process at the SA and SA element (term) levels. Provide formal mechanisms to encode and support these aspects.
Documentation	Create a human-readable description of the SA (HTML pages) to enhance understanding of its structure and content. Where relevant, provide documentation on critical aspects of the SA lifecycle (e.g., versioning, testing, data access, etc.).	Motivate the community to provide documentation for their SA development and curation processes and potential SA usage. Encourage the use of selected semantic artefact catalogues.	Provide documentation on the potential use of infrastructure services. Develop or integrate services into the infrastructure to automate the generation of HTML pages based on SA structure and content. Work along with selected semantic artefact catalogues.
Unique identification for (meta)data	Assign a globally unique and persistent identifier to the SA, SA metadata, and SA elements. Resolve each of them to their different respective representations as requested by the client, using redirection practice and content negotiation mechanisms.	Define the namespace naming conventions. Inform the community about redirection, resolution, and PID provider services. To support this, maybe associate them with selected semantic artefact catalogues.	Develop the policy for creating harmonised and interoperable namespaces within the community or organisation. Integrate redirection and PID provider services into the ontology development workflows and systems to automate the seamless assignment of identifiers to SA, SA metadata, and SA elements and facilitate the management of related representations. Work with selected semantic artefact catalogues to reflect these aspects.
Relevant attributes, metadata and provenance	Use common metadata standards to provide rich context to ontology and ontology terms in different aspects such as description, administration, and provenance.	Establish minimum standards metadata for reporting SA within the community. Encourage the use of selected semantic artefact catalogues dealing with metadata and provenance.	Establish internal metadata schemas alongside existing metadata standards to provide customised classes and properties for reporting SA. Implement these schemas and standards within semantic artefact catalogues.
(Meta)data access and semantic	Host the SA (metadata) on web (application) servers or tools, such	Introduce the community to domain-specific semantic artefact catalogues where they can publish	Develop semantic artefact catalogues or components, such as web applications, repositories, dedicated APIs, SPARQL

Funded by the European Union

repository	as an ad hoc SKOSMOS installation or a triple store (e.g., Virtuoso), to make it accessible via HTTP(S).	their SA, making them accessible through services for term lookup, ontology explorations and dedicated APIs.	endpoints, and related services, to facilitate SA (metadata) management and access.
Information model	Utilise RDF(s), OWL, and SKOS (W3C standards) as core vocabularies for encoding and structuring SA. Offer one or more serialisations of the SA in acceptable syntax formats.	Advocate for using W3C standards when selecting knowledge representation languages for SA development within the community. Recommend using specific representation languages and syntax to ease interoperability.	Enforce certain (W3C) standards for knowledge representation languages in SA development within communities or organisations. Develop standard pipelines for transformation between different SA formality levels. Additionally, it provides services to produce SA source files and convert them between various SA formats (serialisations).
(Meta)data reuse	It involves reusing specific terms and structures defined in existing SA rather than creating new ones from scratch.	Provide a primary list of selected SA for reuse or direct the community to tools that help locate their desired SA for reuse.	Enforce some practices, including specific properties, in terms of reuse to help create interoperable SA networks within a community or organisation. Reuse upper-level ontology to support broad semantic interoperability. Provide methodologies for modular development and design pattern reuse.
Naming Conventions	Choose standard label properties (RDFS, SKOS) to provide a primary label or synonym for every SA term.	Encourage the community to follow established naming conventions based on W3C standards.	Provide detailed documents for establishing naming conventions within the organisation. Contribute and complete naming convention standards when necessary for a harmonised use.
Scope	Prepare a brief text document that clearly states the extent of the domain the SA intends to cover.	Ensure that SA developed within the community has a well-defined scope and content consistent with that scope.	Urge the organisation to clearly state what is out of scope while also identifying and linking to the specific users and use cases the SA aims to address.
Commitment to collaboration and contribution	Utilise tools that facilitate collaborative SA development, such as collaborative editing platforms.	Ensure all SA is assigned a contact person (with detailed information) to facilitate communication between the community and the SA developers. Suggest tools (e.g. forums on semantic artefact catalogues or issue trackers) to ease collaboration.	Organise collaborative workshops to ensure the orthogonality of distinct ontologies. Address topics such as reusing content from other ontologies, determining domain divisions between ontologies, evaluating the potential merger of ontologies into a single artefact, etc. Set up and enforce required collaboration practices and tools.
Channels for communications	Create responsive channels to collect feedback and term requests from users and domain experts and keep them informed about updates	Host public forums to promote knowledge sharing among various stakeholders in the community. Suggest tools (e.g. forums on semantic artefact catalogues or issue trackers) to ease feedback and	Establish scientific and user advisory boards within the organisation. Develop curation and maintenance workflows (e.g., term requests and issue resolution) and integrate them into the SA lifecycle.

Funded by the European Union

		to the SA.	communications.	
	SA visualisation	Provide visual representations for the SA.	Direct the community to tools for visualising SA.	Integrate visualisation tools into the systems to provide hierarchical, graphical, and tabular representations of the SA at different stages of its lifecycle.
Organisational structure	**Governing roles**	Decision makers and executors play a crucial role, with particular attention to technical support from individuals involved in the SA's development and curation. However, the orchestrators who oversee the effective implementation of these decisions may be engaged in other activities or dedicated solely to this task.	Decision making and orchestration are activities that are primarily conducted by the same individuals who facilitate the coordination of procedures. As for the executors, they are mainly the specialists involved in the curation process.	Complete decentralisation among the governing stakeholders, with particular emphasis on the orchestrators activity, which is explicitly described and has formal procedures in place.
	Intermediate roles	Specialists, external or internal, and service-enabling actors provide the most essential intermediary roles. The outreach coordinators are optional.	Key roles are provided by the specialists, who are usually part of the community, and the outreach coordinators play an optional role.	Infrastructure providers play a critical intermediate role by maintaining the technical infrastructure and harmonising connections with multiple services. While the outreach coordinators appear optional, its role is essential in providing additional support to users.
	Participatory roles	The managers and contributors directly provide uploads of new SA.	The managers and contributors primarily undertake this part, but it can also be managed by an additional project investigator associated with the community.	The managers and contributors directly provide uploads of new SA.

Funded by
the European Union

5. Conclusions

The exploration of SA governance within the EOSC context (and beyond) highlights its multifaceted and critical role in digital data management and interoperability. As the EOSC strives to establish a unified, open environment for data and services, the governance of semantic artefacts becomes indispensable for seamless integration and interoperability across varied scientific domains.

Our analysis has underscored the complexity of SA governance, revealed through the diversity of methodologies and practices across different scientific communities. This diversity points to a lack of a universally accepted definition of SA governance; however, it broadly aligns with the principles and frameworks that standardise and manage the lifecycle of semantic artefacts. These include their creation, development, utilisation, and eventual disposal, incorporating decision-making structures, stakeholder roles, and operational mechanisms to maintain the integrity, interoperability, and longevity of SAs.

We identified three foundational pillars of SA governance: *Engineering Methodologies*, *Organisational Structure*, and *Governance Frameworks*, each contributing uniquely to the governance landscape. Each pillar shapes how SA is managed and governed within different communities and organisations. These pillars facilitate systematic approaches to SA development and management, define roles and responsibilities within the governance process, and set overarching rules and standards that enhance adherence to FAIR principles and ensure data quality and metadata management.

SA governance model implementation varies significantly across initiatives, reflecting scientific communities' diverse needs, contexts, and objectives. While some adopt centralised governance structures with formalised decision-making bodies, others favour decentralised, flexible, community-driven approaches. This diversity highlights the adaptable nature of SA governance and underscores the necessity for models that can evolve with changing technologies, scientific practices, and community dynamics.

Moving forward, the development of robust, reusable SA governance models presents both challenges and opportunities. It requires a profound synthesis of practices and principles from successful initiatives like the OBO Foundry, Crop Ontology project, and GOMO-BASF. **Our proposed models:**

- **Model #1: Project-based Initiatives** (i.e., initiatives managing single SA),
- **Model #2: Distributed SA Development and support based initiatives** (i.e., initiatives supporting decentralised SA development efforts),
- **Model #3: SA Harmonisation Initiatives** (i.e., initiatives harmonising SA management practices across a community),

are tailored to meet the distinct needs of various target groups. They are designed to enhance data sharing, collaboration, and innovation within scientific communities, thereby advancing the objectives and vision of the EOSC.

The journey towards effective SA governance in the context of EOSC is ongoing. Stakeholders are encouraged to continue refining and harmonising practices, leveraging lessons learned from a broad spectrum of initiatives to foster innovation, accelerate research outcomes, and ensure the sustainable impact of semantic artefacts in advancing scientific knowledge and discovery. This collaborative approach will enhance the interoperability and reusability of semantic artefacts and pave the way for more efficient and transparent data management practices in support of FAIR principles. By addressing challenges and embracing collaborative governance models, stakeholders can foster innovation, accelerate research outcomes, and ensure the long-term sustainability and impact of semantic artefacts in advancing scientific knowledge and discovery, thereby significantly contributing to the objectives and vision of the EOSC.

6. Appendices

Appendix A. List of CROP, OBO Foundry and BASF stakeholders' structuration and description of their roles and responsibilities

Community / Categories	CROP	OBO Foundry	BASF
Findability	Yes	Yes	Yes
Sources	-Governance document[31] -Actors Roles webpage[32] -Actors members webpage[33]	-Webpage[34]	-Governance document[35]
Formal groups	Yes	Yes	No
Name of the groups (Stakeholders)	-Committees: -Curators Committee -Strategy Committee -Scientific Advisory Committee -Team: -CO team (Technical and Ontology coordinators)	-Committees: -Operations Committee (Editorial, Technical and Outreach working groups) -Code of Conduct Committee -Working groups (WG): -Editorial WG -Technical WG -Outreach WG	-Team -Data semantic team
Description of stakeholders' roles and responsibilities	Yes	Yes	No
Raw description (Stakeholders group Responsibilities	CO management institute *-Accept responsibility to **manage the overall CO project*** *-Appoint the Ontology Coordinator and Technical Coordinator*	Operations committee: *Members of the OBO Foundry Operations Committee aim to*	-ontology -developer -ontology owner -ontology curator

[31] CROP Ontology "Governance and Stewardship Framework" document: https://cgspace.cgiar.org/server/api/core/bitstreams/548034d0-445b-4de3-9050-708941f0a790/content
[32] CROP Ontology stakeholder's roles webpage : https://cropontology.org/page/Advisoryroles
[33] CROP Ontology stakeholder's members webpage :https://cropontology.org/page/MembersAC
[34] OBO Foundry stakeholder's organisation webpage : http://obofoundry.org/about-OBO-Foundry.html
[35] BASF "Governance operational model of ontologies" document : https://doi.org/10.5281/zenodo.7007495

Funded by the European Union

Individual)	*-Provide necessary resources to maintain the website and develop necessary new features* CO team: Technical and Ontology Coordinators -Ontology coordinator *-**Lead the CO team** at the global level on behalf of the CO Management Institute* *-Manage the product, its content, its **website** and the **quality assurance process*** *-**Manage the regular update of the Guidelines and Template, quality checking tools, validation and that curation roles** are in place.* *-**Maintain a practical governance framework*** *- **Regularly consult the community of curators to get feedback and needs*** *-Interact with the Strategy Committee to get timely recommendations on global CO strategy* *-**Validation of new crops after checking if the governing roles** are in place* *-Allocate unique crop codes from the CO list* *-Guide users* *-Initiate fundraising and engage resources* *-**Lead the outreach activities** with the Ontologies Community of Practice* *-**Engage community members** in decision making and communication* -Technical coordinator *-Member of the CO team* *-**Update the CO content** by providing guidance and expertise to curators* *-Promote the use of best practices concerning Crop Ontology development* *- Contribute to the discussion around content and CO domain boundaries* *-Guide the use of other ontologies functional in agriculture* *-**Contribute to the improvement and maintenance of the Template and Guidelines*** *-**Maintain & upgrade, in collaboration with the developer, the CO website backend & frontend** according to users' requirements* *-Maintain the domain name of the CO website* *-**Check crop ontologies for quality and inconsistencies*** *-**Operate ontology helpdesk*** *-Promotion of the CO* *-Create the term mappings with Planteome* *-**Technical interactions with ontology registries and the breeding databases*** *-Training of curators*	*improve the **flow of operations** and make things happen within the OBO Foundry. Such operations include, but are not limited to, establishment of policies, review of resources, outreach and education.* -Editorial WG *is responsible for:* *-Principles: drafting text and creating the workflow and guidelines for the principles development process* *-Ontologies: reviewing and creating the workflow and guidelines for the ontology review process* *-Policies: crafting the policies for the above* -Technical WG *is involved in **maintaining the technical infrastructure** for the OBO Foundry. This includes establishing **policies** to be **implemented** in standard **tools**, **website** maintenance, etc.* -Outreach WG *is involved in **public relations** for the OBO Foundry. This includes monitoring and following up discussions on mailing lists, preparing documentation and educational materials, and presenting OBO Foundry activities at workshops, conferences, or other venues.*	-domain expert -end-user

	Curators committee *-Group of Curators representing all crop ontology curators* *-Selected by the members of the curator community for a term of up to 2 years* *-Acts as a standing body to facilitate communication between the entire curator community and the CO Team* Strategy Committee *-Committee with representatives from crop domains, CO global project domain and external domain (including representatives of Breeding databases and ontology registries)* *-**Provide recommendations** to the CO Ontology Coordinator on **strategic project issues*** Scientific Advisory Committee *-Committee with representatives from crop domains and external (scientific) domain* *-**Provide advice to the CO Team on scientific issues** relevant to the CO* *-**Provide guidance and recommendations** for the evolution of the Crop Ontology content and t**echnology, identification of best practices** regarding the use of CO or other ontologies* *-**Provide expertise and guidance for the homogenisation** of entities and attribute terms across species – e.g. a tissue name* *-V**alidate the quality assurance tools** proposed by the Community that can be promoted through CO* *-Validate for some species a group of experts who could evaluate the submissions (homogeneity). The group would make critical assessments (redundancy checking, homogenisation of terms, etc.).* *-Validate the modifications of the Trait Dictionary Template and accompanying Guidelines in consultation with the Ontology and Technical Coordinators*	Code of Conduct committee *is responsible for ensuring the OBO Code of Conduct is upheld.*	

Appendix B. List of INRAE Voc, IVOA and NFDI4Biodiv stakeholders' structuration and description of their roles and responsibilities

Community / Categories	INRAE Voc	IVOA	NFDI4Biodiv
Findability			
Sources	No	Yes	No
Formal groups (Stakeholders)	-	-Actors webpage[36] - IVOA wikipage[37] - Decision process document[38]	
Formal groups (Stakeholders)	Yes	Yes	Yes
Name of the groups (Stakeholders)	-Committees: -Editorial Committee	-Committees: -Executive Committee -Standing Committee on Science Priorities (CSP) -Standing Committee on Standards & Processes (SCSP) -IUA-IVOA Liaison Committee (IILG) Groups: -Technical Coordination Group (TCG) (WG chair and vice-chair; IG chair and vice-chair; -TCG chair and vice-chair; -IVOA chair and vice-chair; Chair of SCSP) -Working groups (WGs) (Applications, Data Access Layer, Data Model, Grid and Web Services, Registry and Semantics) -Interest group (IGs) (Data Curation & Preservation, Education, Knowledge Discovery in Databases, Operations, Solar System, Theory, Time Domain, Radio, High Energy) -Media Group	-Committees: -Taxonomies Editorial Committee -Board: -Taxonomies Editorial Board -Groups: -Tasks groups
Description of stakeholders' roles and responsibilities	Yes	Yes	Yes
Raw description (Stakeholders group	Editorial Committee Provide editorial guidelines in SA	Executive committee **Review and approve WG recommendations**	Taxonomies Editorial Committees *oversee taxonomy-related*

[36] IVOA stakeholder's members webpage : https://www.ivoa.net/members/
[37] IVOA wikipage : https://wiki.ivoa.net/twiki/bin/view/IVOA/WebHome
[38] IVOA "Document Standards" : https://www.ivoa.net/documents/DocStd/20170517/

Funded by the European Union

Responsibilities Individual)	quality and maintenance SA authors Decision of SA development and maintenance External experts Ensure scientific quality	CSP *The primary goal of the CSP is to **support the Executive Committee** in its goal of **sustaining** the VO's impact, as follows:* *-**Advise the Executive Committee** on strategies for engaging the astronomical community as a participant in the Virtual Observatory.* *-Demonstrate to the community the benefits of participation in the IVOA, such as the growth in the scientific return of data, the capability to discover and fuse multiple data sets, and the application of the VO in planning new observations and observing strategies.* ***-Recommend scientific priorities and requirements** that will drive the development of new services, protocols, and tools. The IVOA will develop these and coordinate with the TCG. These priorities should be driven by scientific use cases developed in cooperation with the scientific community.* *-Support the TCG in the development of protocols to **ensure** they meet **the actual scientific requirements*** *-Support VO members in **developing tutorials, workshops and scientific training materials**.* SCSP *The Standing Committee on Standards and Processes (SCSP) reviews and updates IVOA processes in response to the needs of the IVOA community. When such updates necessitate the revision of an IVOA standard, the same processes apply to the **review and promotion of that standard** as when the revisions originate with a Working Group, with the chair of the Committee playing the same role as the chair of a Working Group.* IILG *role is to support the IAU's goals in astronomy and promote the value of the Virtual Observatory, especially its value in promoting Open Science and FAIR principles, to the IAU's international community* TCG *-is to **ensure proper technical coordination amongst the various IVOA WGs and IGs and** a **liaison role between them and the IVOA Executive Committee**.* *-the vocabulary is updated, and the IVOA Technical Coordination Group (TCG) endorses this update, potentially prompting additional discussions at the IVOA*	*decisions and quality control.* Taxonomies Editorial Board *ensure the quality and accuracy of taxonomic information by making informed decisions about taxonomy updates and changes.* Task groups *- work collaboratively to establish consensus on data and metadata standards.* *- drive decisions on standards and best practices for biodiversity data through consensus-building processes* Ontology Manager *responsible for ontology structure, content, and versioning decisions.* Experts in biodiversity domains *contribute domain-specific expertise and validate artefacts.*

Funded by the European Union

		level. WG ***Produces*** *specifications, guidelines etc..[...] that may process to* ***recommendation*** IG Share best practices and engage IVOA members' projects in the topic. Media Group *aims to* ***disseminate the latest news and information*** *about Virtual Observatory developments,* ***applications, standards, workshops, meetings,*** *etc., to astronomers via various channels, including social media. Messages will be catered to IVOA members and primarily to the general astronomy community to show the benefits of the VO for their science. To achieve this, the Media Group's central areas are Social Media, the IVOA Newsletter, the IVOA Web page and Outreach.*	

Appendix C. List of EMBL-EBI, SAREF and AGROVOC stakeholders' structuration and description of their roles and responsibilities

Community / Categories	EMBL-EBI	SAREF	AGROVOC
Findability	**EMBL-EBI**	**SAREF**	**AGROVOC**
Sources	Not Found	Yes	Yes
Formal groups (Stakeholders)	-Actors webpage[39] -Team webpage[40]	- Pipeline document[41] - Technical Specification document[42] -Specialist Task Force (STF): 641 webpage[43] and STF 653 webpage[44]	-Webpage[45] -Editorial guidelines document[46] -Semantic data interoperability document[47]
Formal groups (Stakeholders)	Yes	Yes	Yes
Name of the groups (Stakeholders)	-Team -Samples, phenotypes and ontology (SPOT) team -Ontology Application team -Semantic data integration service team		-Team: -AGROVOC team -Board -Core Board
Description of stakeholders' roles and responsibilities	Not Found	Yes	Yes
Raw description (Stakeholders group Responsibilities	SPOT team: *Actively develops ontologies, including the Cell Type Ontology and Experimental Factor Ontology.*	Steering actors: -Steering Board members *belongs to the persons* ***in charge of steering the***	Core board *-comprises the* ***FAO and KTBL team****, and* ***decision-making*** *follows* ***established guidelines,***

[39] EMBL-EBI stakeholder's members webpage : https://www.ebi.ac.uk/about/teams/samples-phenotypes-ontologies/members/
[40] EMBL-EBI team organisation webpage: https://www.ebi.ac.uk/about/teams/samples-phenotypes-ontologies/about/
[41] " SAREF Pipeline and Portal-An Ontology Verification Framework" document: https://hal-emse.ccsd.cnrs.fr/emse-04277942v1/document
[42] SAREF "Development Framework and Workflow, Streamlining the Developement of SAREF and its Extensions" document: https://www.etsi.org/deliver/etsi_ts/103600_103699/103673/01.01.01_60/ts_103673v010101p.pdf
[43] SAREF STF 641 SAREF Digital Twins webpage: https://portal.etsi.org/xtfs/#/xTF/641/what-we-do
[44] SAREF STF 653 SAREF Patterns webpage: https://portal.etsi.org/xtfs/#/xTF/653/how-we-do
[45] AGROVOC webpage : https://www.fao.org/agrovoc/about
[46] AGROVOC "Editorial Guidelines" document : https://doi.org/10.4060/cb8640en
[47] AGROVOC "Semantic data interoperability on food and agriculture" document : https://doi.org/10.4060/cb2838en

Funded by the European Union

Individual)	-Semantic data integration -Ontology application team -Semantic data integration service team *deliver Lookup Service* External collaborations *Projects range in scope from data analysis and generation projects to projects delivering infrastructure.* Project leader: -Technical project leader -Project lead -Ontology project lead Manager: -Scientific programme manager -Research management office lead Coordinator: -Catalogue coordinator - Ontology project coordinator Editors: -Ontology editorial lead - Ontology editor Developers: -Software Developer -Software engineer -Software developer-ontology tools - Ontology developer -Semantic web developer -Web developer Curators:	***SAREF development,*** *including SAREF core and SAREF extensions. The community involvement and the underlying infrastructure**.** [...] then **review these change requests**.* *-is it composed of the SmartM2M Chairman, Vice-Chairman, Technical Officer, and **experts** nominated by SmartM2M?* (Workflow 1: new SAREF project version and Workflow 3: project release) -Technical Board members *belong to the persons in charge of **maintaining the SAREF public forge and the SAREF public portal.*** *- is composed of ETSI Secretariat and **experts** nominated by SmartM2M.* (Workflow 1: new SAREF project version, Workflow 2: project version development, Workflow 3: project release) Development actors: -Project leader *is the person in charge of the SAREF project who carries out the **project management tasks**.[...] may have at l**east the role of Maintainer in the ETSI public forge.*** *The project leader **ensures that SmartM2M approves the change requests and that the implementations satisfy the requested change.*** *(Workflow 1: new SAREF project version, Workflow 2: project version development, Workflow 3: project release)* -Ontology developer *is a member of the ontology development **team** who has high knowledge about ontology development and rights **to modify the ontology and interact in the development cycle** [...] create*	*seeking consensus through annual meetings of the AGROVOC Editorial community.* *-**maintains consistency and coherence** per AGROVOC Editorial Guidelines, continually **improving curation workflows**.* Coordination and technical support (FAO) Development *-key stakeholders include FAO, The Kuratorium für Technik und Bauwesen in der Landwirtschaft (KTBL), and Tor Vergata University (Italy)* *-FAO and Tor Vergata University (Italy): the management of specialised concept schemes is possible within AGROVOC* Maintenance: (FAO) *FAO keeps AGROVOC current, with several institutions and individual domain experts as focal points for specific languages or topics.* *-FAO carries mainly the responsibility for the six FAO languages (English, French, Spanish, Arabic, Chinese and Russian),* *-facilitates the **technical maintenance** of AGROVOC, including its publication as a Linked Open Data resource, and coordinates all editorial activities* **Update**: (editors & AGROVOC team) *Our **editors** and **our team** update it continuously. Updated AGROVOC content is released once a month.* **Editors** *maintained by a vast community of experts and institutions,* AGROVOC is edited using Vocbench and validated by the AGROVOC team

	-Scientific curator -Senior curator -Catalogue curator -Bioinformatic developer -Bioinformatician Data Engineer: -Data Engineer -Semantic data engineer	*and modify the different development artefacts,* ***provide new requirements*** *to the ontology and validate whether they are satisfied or not when implemented, and have decision rights about what contributions can be included in the ontology [...] also review change requests, propose, and review implementations of accepted change requests. (Workflow 2: project version development-* Experts *may be nominated by SmartM2M to become a* ***development actor*** *for some SAREF projects.* Community actors: -Contributor *is knowledgeable about the ontology domain and* ***proposes contributions*** *[...] have an account on the SAREF public forge. The role of the contributor* ***is not assigned beforehand; it is obtained when submitting some contribution.*** (Workflow 1: new SAREF project version, Workflow 2: project version development, Workflow 3: project release) -Ontology user *may start contributing to some SAREF projects and become a Contributor [...] if interested in any of the SAREF projects or in proposing a new project [...] do not have an account on the SAREF public forge [...] include potential end users of the ontology. These software developers, industry stakeholders, researchers, domain experts, etc., will use the ontology within their applications.* The SAREF public forge *allows users to be defined in the following roles: Guest, Reporter, Developer, Maintainer, and Owner,*	Editor Community AGROVOC Dgroup **Curators**(FAO and expert external) -***expert communities*** *can now curate a topic within AGROVOC, enriching AGROVOC with specialist knowledge, with modern infrastructure to share this as part of the AGROVOC Linked Open data* *-curatorial responsibilities* ***involve FAO and a team of 40 experts*** *from 34 organisations across 24 countries.* **Social media** *-[AGROVOC team] organises periodical in-person and virtual meetings with the editorial community and community of experts to maintain and coordinate all contributions to AGROVOC.* *-endeavours to publish and document the work done in the context of standards, technology and good practices in the agricultural domain.* **AGROVOC team (FAO)** -manager -Curator -Technical lead -Technical support -Communication support **External experts** -***40 experts*** *from 34 organisations across 24 countries.*

Funded by the European Union

		each with its permissions. European Commission and the ETSI SmartM2M Technical Committee. *Overseen SAREF governance*	

Appendix D. Alignment of BASF and OBO Principles

BASF principles	OBO Foundry principles
P1) Availability - Ontologies must be available across the organisation. **P2) Access rights and security policy** - Ontologies Must be assigned with their own access rights and security policy.	**P1) Open** - The ontology MUST be openly available to be used by all without any constraint other than (a) its origin must be acknowledged and (b) it is not to be altered and subsequently redistributed in altered form under the original name or with the same identifiers.
P3) Persistency - versioning and deprecation process.	**P4) Versioning** - The ontology provider has documented procedures for versioning the ontology and different versions of ontology are marked, stored, and officially released.
P4) Documentation	**P8) Documentation** - The ontology owners should strive to provide as much documentation as possible.
P5) Findability (F1) - (meta)data are assigned a globally unique and persistent identifier.	**P3) URI/Identifier Space** - Each ontology MUST have a unique IRI in the form of an OBO Foundry permanent URL (PURL).
P5) Findability (F2) - data are described with rich ontology metadata.	
P5) Findability (F3) - (meta)data clearly and explicitly include the identifier of the data they describe.	
P5) Findability (F4) - are registered or indexed in a searchable resource, typically a repository.	
P5) Accessibility (A1) - (meta)data are retrievable by their identifier using a standardised communications protocol	
P5) Accessibility (A2) - metadata are accessible, even when the SA are no longer available	
P5) Interoperability (I1) - (meta)data uses a formal, accessible, shared, and broadly applicable language for knowledge representation.	**P2) Common Format** - The ontology is available in a common formal language in an accepted concrete syntax.
5) Interoperability (I2) - (meta)data use vocabularies that follow FAIR principles	
5) Interoperability (I3) - (meta)data qualified references to other (meta)data	**P7) Relations** - Relations should be reused from the Relations Ontology (RO).
P5) Reusability (R1) - (meta)data are richly described with a plurality of accurate and relevant attributes	**P6) Textual Definitions** - The ontology has textual definitions for most classes and particular top-level terms. **P9) Documented Plurality of Users** - The ontology developers should document that multiple independent people or organisations use the ontology.

Funded by
the European Union

P6) Modularity	
P7) Community - Ontologies MUST be driven by a community	
	P12) Naming Conventions - The names (primary labels) for elements (classes, properties, etc.) in an ontology must be intelligible to scientists and amenable to natural language processing. Primary labels should be unique among OBO Library ontologies.
	P5) Scope - The scope of an ontology is the extent of the domain or subject matter it intends to cover. The ontology must have a specified scope and content that adheres to that scope.
	P10) Commitment To Collaboration—OBO Foundry ontology development should be carried out collaboratively, as with many other standards-oriented scientific activities.
	P11) Locus of Authority - There should be a person who is responsible for communications between the community and the ontology developers, for communicating with the Foundry on all Foundry-related matters, for mediating discussions involving maintenance in the light of scientific advance, and for ensuring that all user feedback is addressed.
	P13) Notification of Changes - Ontologies SHOULD announce significant changes to relevant stakeholders and collaborators before release.
	P16) Maintenance - The ontology needs to reflect changes in scientific consensus to remain accurate over time.
	P20) Responsiveness - Ontology developers MUST offer channels for community participation and SHOULD be responsive to requests.

Funded by
the European Union