CALLISTO: A Semantics-Based Platform for Sharing FAIR Scientific Data

Thierry Louge[1], Emmanuel Courcelle[1], Michelle Sibilla[2], Nathalie Aussenac-Gilles[2], Christophe Marteau[1], Marianna Braza[3], Jean-Baptiste Tô[3] and Laurent Orgogozo[4]

[1]*UAR CALMIP 3667 Université de Toulouse, CNRS, INPT, INSA, ISAE, UPS, Toulouse, France*
[2]*UMR IRIT 5505 Université de Toulouse, CNRS, Toulouse, France*
[3]*UMR IMFT 5502 Université de Toulouse, CNRS, INPT, Toulouse, France*
[4]*UMR GET 5563, Université de Toulouse, Université Toulouse Paul Sabatier - Toulouse III, France*

Abstract

In this paper, the platform CALLISTO (CALmip Launches an Interface for a Semantic Toolbox Online, https://callisto.calmip.univ-toulouse.fr/) is presented, aiming at open science construction around semantic description of scientific data, software and papers. This work aims tackling the needs for a fine-grained description of multi-disciplinary scientific collaborations, from data creation to cross-fertilization. A framework for hosting comprehensive and semantic metadata for multi-disciplinary scientific teams is proposed. The platform provides a user-friendly online interface for data description, browsing and access. This research has been conducted in collaboration with the multi-disciplinary dataNooS academic alliance. CALLISTO is developed and maintained at the computing centre CALMIP.

Keywords

ontologies, FAIR principles, semantic web, multidisciplinary data sources

1. Introduction

The constant growth of scientific data available online and the growing supply of processing software can paradoxically make the reuse of data and algorithms harder. The lack of widely-used fine-grained description of scientific data, the need for explicit and ergonomic online interfaces for finding data are some of the challenges ahead towards open science. These concerns are at the heart of the work to encourage the widest possible adoption of the FAIR (Findable, Accessible, Interoperable and Reusable) principles. The simultaneous efforts at a national and European level to adopt the FAIR principles, as well as the existing difficulties within scientific projects in managing multidisciplinary data at several reading levels, contribute to raising scientists' awareness of the issues involved in the management and use of their data. Within the University of Toulouse, an academic alliance has been created and named dataNooS[1]. It involves researchers and engineers from different scientific disciplines, whose main role is to study solutions for efficient data handling, interoperability and re-usability generalizable to future projects. As a member of dataNooS, the computing centre CALMIP developed and

SEMANTICS 2022 EU: 18th International Conference on Semantic Systems, September 13-15, 2022, Vienna, Austria
thierry.louge@inp-toulouse.fr (T. Louge)
0000-0002-3709-1976 (T. Louge); 0000-0001-8199-6732 (M. Sibilla); 0000-0003-3653-3223 (N. Aussenac-Gilles)

[1]https://datanoos.univ-toulouse.fr

maintains CALLISTO[2] for "CALmip Launches an Interface for a Semantic Toolbox Online", implementing ideas from dataNooS[3] in a production-level site. At the first step, the goal of CALLISTO consists of sharing data among partners with heterogeneous scientific backgrounds, ensuring their description and easing reuse in a secured calculation environment. At a second step, the objective is to open data, results and analysis processes to the public. CALLISTO is designed with a user-driven development process, through the study of data challenges coming from interdisciplinary projects such as the H2020 project Smart Morphing and Sensing for aeronautical configurations (SMS, http://smartwing.org/SMS/EU/) funded by European Commission (project No. 723402) and HiperBorea[4], funded by the French National Research Agency (ANR) (project ANR-19-CE46-0003-01). These projects illustrate the challenges for scientific projects concerning the sharing and the description of heterogeneous and multidisciplinary data.

In this paper, CALLISTO is presented as an illustration of how the FAIRness of scientific data can be achieved, and supplemented by other aspects such as descriptions of the software and algorithms used to create, analyse or transform the data. This article presents creation of a platform for linking data to software, scientific papers and arguments through the adoption of well-known ontologies such as Micropublications [1] or Software Ontology [5]. An ontology is defined for mapping all these elements in a homogeneous architecture, designed to allow in the medium term the production of automatic data processing and retrieval workflows. We describe this ontology called ARCAS (ARming CAllisto with Semantics), its role in relation to the other elements composing the system.

2. Running example: SMS project

SMS for "Smart Morphing and Sensing" is the European project mentioned in the introduction. This project is a collaboration between experts with different scientific backgrounds (aeronautics, fluid mechanics, structural mechanics, control system...). Different teams from different countries worldwide make their own experiments in their competency areas, and then need to agree on common concepts in order to work with one another. The software libraries used by the various teams may be shared, and projects papers shall be understood by all the project's stakeholders. A major requirement is then to have a common metadata set that all stakeholders understand, and all future users will also understand. Meanwhile, metadata shall also be able to express very narrow elements for specialists.

The first element to ease data sharing among project members is to provide a shared space with storage space and metadata browsing. This is provided by Dataverse, which is used as a metadata catalog and data registry in CALLISTO. An important element for experiments reuse and reproducibility, is a good understanding of the analysis processes the data went through [2], the scientific claims they support and of the overall context of the experiments. ARCAS and ARCADIE (ARCAS Domain Implementation, ARCAS enriched and populated for a specific project)

[2]http://callisto.calmip.univ-toulouse.fr
[3]This research has been partially funded by RTRA: http://www.fondation-stae.net/
[4]http://hiperborea.omp.eu
[5]http://theswo.sourceforge.net/

handle these elements. The bibliography for a project is described using Micropublications ontology. Software libraries use SWO for non-grounding elements (e.g. algorithms provided, languages...) and GEOS[3] (Generic Ontology for Services, based on OWL-S) for grounding elements (e.g. Inputs and outputs combinations, URLS, parameters...). The prospective issue of data analysis from OBDAs raised in [4] is addressed by the inclusion of SWO and the GEOS (GEneric Ontology for Services) module extracted from ASON [5] already used for automatic composition of semantic services, the ultimate goal of CALLISTO, in the astrophysics domain. GEOS and SWO allow accurate software description and automation whenever possible. The following Table 1 summarizes the requirements from SMS project stakeholders regarding "FAIR-ification" of their data following FAIR)Findable, Accessible, Interoperable, Reusable) elements classification.

Table 1: Requirements for FAIR data sharing

Reference	Requirement	FAIR element
1	Provide a comprehensive metadata set understandable by all the stakeholders	I
2	Allow the extension of metadata through unambiguous references	I
3	Specify the analysis process the data went through	R
4	Link data to papers, claims and arguments	R
5	Access through user-friendly interfaces with no technical prerequisite	A

3. Related works

3.1. Semantic description of data and metadata through ontologies

Ontologies, following the widely-used Gruber definition are to be considered as "explicit specification of a conceptualization"[6]. They are at the heart of the Semantic Web (SW) for several reasons, as underlined by Tim Berners-Lee in his grounding paper [7] and its update in 2006 [8]. Ontologies are by nature intended to be browsed and interpreted by algorithms as well by humans. For this reason, they benefit from several advanced tools allowing their edition by graphical means (Protégé [9], inline browsers[6]), and several query languages (SPARQL is one of the most well-known). Among data-related ontologies, DCAT-AP[7] is a DCAT-based European specification for datasets, intended to standardize their description in order to improve cross-portal searches. Ensuring interoperability of data and services through the use of ontologies is not a new idea. It has been described in particular in a paper by David Chen [10] as "federated" interoperability. Still, it remains little used in practice despite its possibilities, as building a usable and efficient ontology-based system requires ontological expertise, in addition to the strict scientific skills related to a particular field. Nevertheless, high level abstraction

[6]http://vowl.visualdataweb.org/webvowl.html
[7]https://github.com/SEMICeu/DCAT-AP

ontologies, and ready for use ontologies exist, making it easier to find relevant ontologies for a specific use-case. The Micropublications [1] ontology defines concepts and relations that aim at modeling argumentation and counter-argumentation. While describing datasets is not in the scope of Micropublications, its modeling of methods and data as scientific evidences supporting scientific claims makes it particularly adapted to the use cases that CALLISTO proposes to address. The connection between metadata and ontologies concerning data semantics can be studied from a number of perspectives, whether controlling the data entered in metadata fields [11], improving search capabilities in data warehouses [12], or classifying and enriching metadata formats according to user preferences [13]. In this paper, as discussed in sections 4.3 and 4.4, ontologies are used as metadata model (section 4.3) as well as metamodel for data (section 4.4, concerning services composition).

3.2. Ontologies for software description

The goal pursued in CALLISTO by the means of software description is quite different from software registries found in astrophysics[8], geosciences[9] or domain-agnostic registries such as Software Heritage [14] (a repository for long-term software preservation). The software in those registries are not connected to the data they help to produce or analyse. With CALLISTO, the goal is not that much capturing software metadata, but rather to allow service composition with the fine-grained description of elements involved in this composition. To this end, ARCAS ontology concentrates in a single point a description of the scientific domain, a directory of services with their grounding elements for accessing the data and linking them with processing. Software Ontology (SWO) [15] is an ontology for software description. Although originally designed for use in the field of bioinformatics, SWO offers a large number of concepts and relations to describe software elements, regardless of their application domain. There are dozens of data types, algorithms and formats already defined, using concepts and relations from the OBO foundry vocabulary [16]. Relying on this vocabulary ensures maintainability and increased interoperability with all ontologies using Basic Formal Ontology (BFO) [17], a high-level ontology specifically designed for interoperability and reusability. The software description ontology (SD)[10] shares a similar scope with SWO, and the two seem very complementary. SD focuses more on data specification, transformation and structure than SWO wich is more oriented towards algorithms specifications and software functionalities. SWO offers a more fine-grained description of software than its counterpart used in Dockerpedia[11], that is more oriented towards the description of Docker images and the dependencies that they contain. The `vocab` ontology underlying Dockerpedia[12] is not directed towards the same goals as SWO or SD, and is far from offering the same level of details concerning the insights of the software elements (e.g. inputs, algorithms). The same applies to the `Codemeta` vocabulary[13], extending Schema.org and describing software elements at a much lower level of detail than SWO. The Research Objects

[8]https://ascl.net/
[9]https://csdms.colorado.edu/wiki/Model_download_portal
[10]https://w3id.org/okn/o/sd
[11]https://cbuil.github.io/dockerpedia-ontology/release/0.1.0/ontology.xml
[12]http://dockerpedia.inf.utfsm.cl/vocab
[13]https://codemeta.github.io/terms/

community released ro-crate[14], an effort to capture software data and bibliography. Ro-crate uses a combination of JSON-LD and RDF langages to describe data, software, workflows and other elements. While ro-crate seems very efficient in its field, it does not apply on CALLISTO where the focus is not on metadata describing the workflows, but rather on the use of elements allowing a dynamic services composition leading to a workflow. In the case of the joint use of several repositories from different scientific projects inside CALLISTO, we will eventually have a network of ontologies but all elements will be based on ARCAS. It is of additional interest to ensure the maximal interoperability among the ontologies within this network thanks to high-level ontologies sharing concepts with a maximum of other ontologies and the use of SWO. SD may seem an interesting addition, but we will favour SWO concepts and relationships because of this concern for interoperability in OBO. Nevertheless, a close study of the possibilities offered by SD should be conducted in the next steps of CALLISTO developments and will be mentioned in the future works.

By combining SWO in an ontology and data access description with domain knowledge, it is possible to go as deep as the languages used by the software, the libraries, and the different algorithms with their functionalities. This is very useful especially in science, where many algorithms in a single software may accomplish the same functionality (e.g., solving an equation) with different implementations. In CALLISTO, these formal descriptions rely on the ARCAS ontology that serves as a single register for these different elements. By reasoning on these description the system builds processing chains (workflows) adapted to the users' requests when using the platform.

4. The CALLISTO platform

CALLISTO is the platform for collaborative science proposed by CALMIP. CALLISTO is available for public use[15] and the underlying ARCAS ontology is downloadable[16]. The SMS example of using ARCAS in a scientific project is also available online[17]. Public use of CALLISTO is restricted to items (data, software, ontologies...) made public by CALMIP users.

4.1. ARCAS and ontologies

The primary intent of CALLISTO is to provide efficient data sharing, allowing for fine-grained description of data and the scientific context they belong to, and that is realized by using project-scale ontologies. All the afore mentioned ontologies are based on a common core, the generic, domain-independant ontology called ARCAS (ARming CAllisto with Semantics). Modeling the treatments associated with the data, whether for their creation or for future analysis, is important. It provides information on the technical and scientific provenance of the data, and allows future processing to be considered based on the precise description of algorithms, functions and data formats acceptable to the software. ARCAS uses to this end SWO [15] with elements from edamontology [18]. The purpose of the data is to produce scientific

[14]https://www.researchobject.org/ro-crate/1.1/
[15]https://callisto.calmip.univ-toulouse.fr/
[16]https://callisto.calmip.univ-toulouse.fr/ARCAS.owl
[17]https://callisto.calmip.univ-toulouse.fr/SMS.owl

arguments. It is interesting to be able to relate the arguments, the data, their provenance and the provenance of the articles (which is not necessarily the same as the provenance of the data, for example when reusing data for new scientific work). ARCAS uses MicroPublications [1] for modeling articles and arguments, and PROV-O [19] for data and article provenance. CALLISTO is organized around a specialization of ARCAS by project. These ontologies are generically referred to as "ARCADIE" (ARCAS Domain Implementation). In ARCADIE, it is possible to link scientific arguments and the data concerned, in order to obtain a complete overview of the data, from their business meaning to the surrounding software tools. Those descriptions emerge from a practical community which is consistent with the development of the semantic Web [8].

4.2. Datasets repository

Dataverse[18] is used by CALLISTO as a data repository. AllegroGraph and Gruff [20] allows for the visualization, query and update of the associated ontologies. To ensure homogeneity and complementarity between the repository and the ontologies, CALLISTO performs a bidirectional exchange between the ontologies and Dataverse: selected ontology elements can be passed to the Dataverse metadata sets and reversely, the descriptions of the datasets automatically populate the associated ontology. Thus, the user can populate metadata fields coming from the ontology to Dataverse, and supplement this metadata with concepts from the ontology. Figure 1 shows the functional overview of CALLISTO.

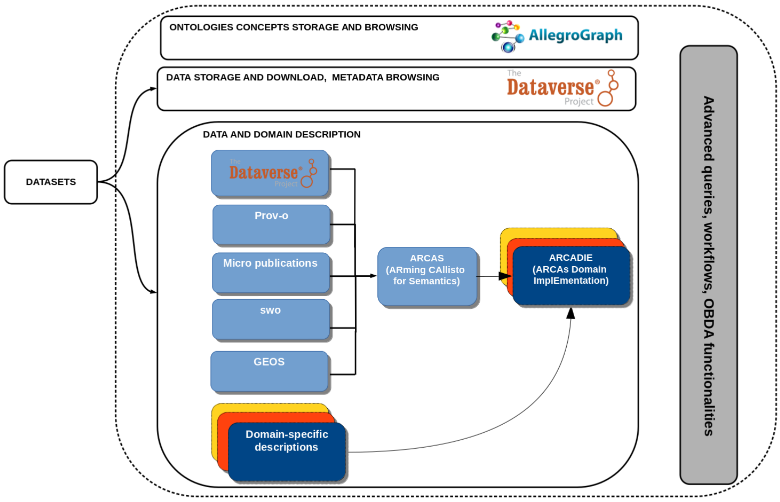

Figure 1: Functional overview of CALLISTO

Thanks to the GEOS module, CALLISTO aims at providing to its users advanced functionalities such as automatic or semi-automatic services composition. These functionalities rely on Allegrograph as the triple-store storing CALLISTO ontologies.

[18]https://Dataverse.org/

4.3. Expressing ontology content into metadata elements

The role of metadata is to describe the strongest common block of description agreed upon by the majority of project participants. Therefore, the minimum common concepts can be understood by all regardless of the scientific field. In our approach, metadata refers to ontology items, allowing the fields to be filled in with a degree of flexibility while referring unambiguously to domain entities. This is an advantage over a ”classic” metadata field that can sometimes be filled in differently depending on the user and whose exact meaning can become ambiguous. The conclusion of the above, is that ontologies and metadata need to be related to each other. Nevertheless, being able to reliably match the content of metadata with the ontological representation of a knowledge domain is a delicate problem, difficult to generalize and often implemented in an ”ad hoc” way. In CALLISTO this is resolved by specifying the metadata themselves in the domain ontology so that they become a part of the domain knowledge representation. ARCAS provides a simple way of expressing that part of its content should be part of a metadata scheme. It uses three elements to this end: the annotation property ”metadataLevel” and the two relations ”hasMetadata” and ”hasMandatoryMetadata”. The property ”metadataLevel” indicates that the element it annotates is to be considered as an element of the metadata model for the ontology content. Its values may be 1 or 2, a value of 1 indicating that the element is an optional metadata field, and 2 indicating that the element is a mandatory metadata field. Figure

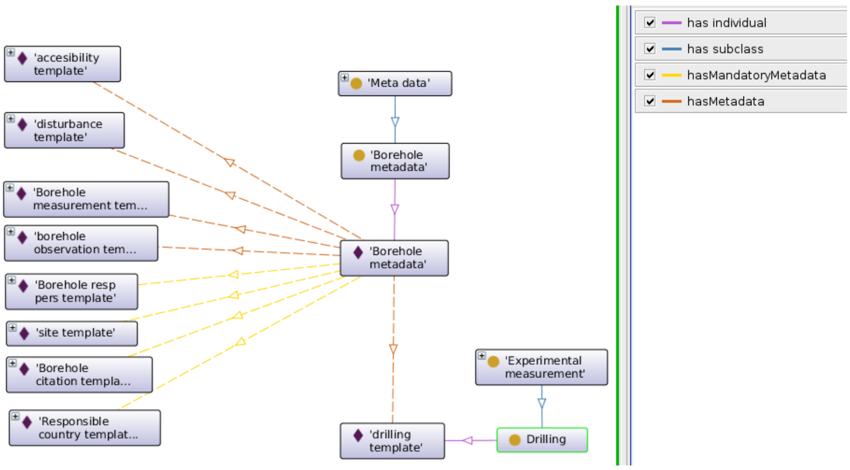

Figure 2: Modeling of Borehole metadata elements in HiperBorea instance of ARCAS

2 shows an example of metadata expression, for the HiperBorea instance of ARCAS. HiperBorea aims quantifying the impact of climate change on boreal permafrost, and drilling is used for experimental measurements. The HiperBorea scientific leaders chose to follow GTN-P metadata modeling of boreholes. As a consequence, elements in HiperBorea ontology have been named to reflect this schema. A template for drilling individuals is available in the dedicated HiperBorea ARCADIE version, providing properties and relations for relevant drilling description. These elements are part of the metadata associated to the boreholes. Appropriate relations and properties have been created, using the ARCAS ”MetaDataLevel” properties, the ”hasMetadata” and ”hasMandatoryMetadata” relationships, thus allowing automated Dataverse-compliant

metadata generation

4.4. Interoperability, request mapping and services composition

The description of the software, together with the description of the scientific claims related to the data and the papers, are part of the "V.R.E." (Virtual Research Environment) offered by CALLISTO, exposed in Figure 3. This V.R.E serves two main purposes:

- Services composition including data access and data analysis in several steps (workflows). This service composition, based on the algorithm developed for astrophysics [5] uses reasoning on the ontology. Elements of such a composition are shown in Figure 4.
- Localization of datasets for accessing paper-related data.

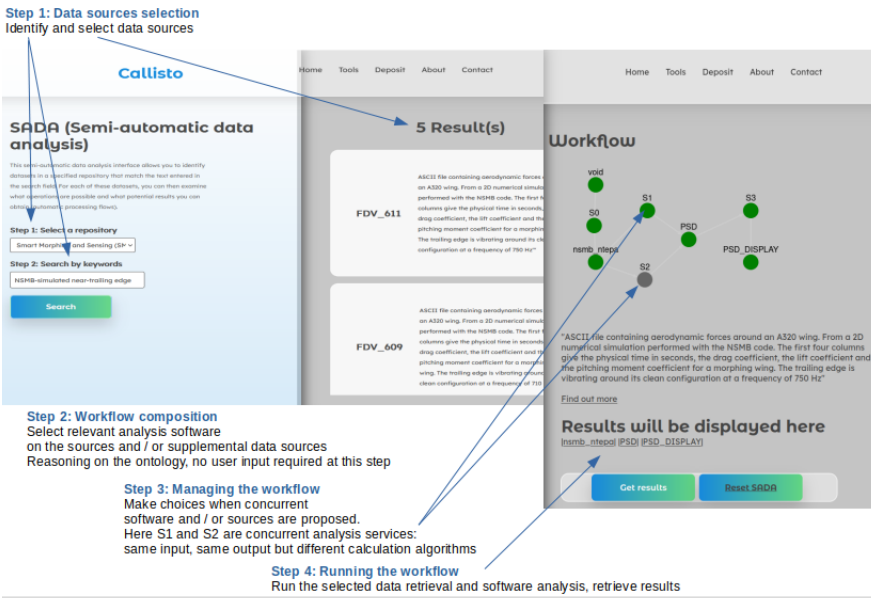

Figure 3: Service composition in CALLISTO

Figure 3 presents the four steps for a service composition in CALLISTO. From a similarity matching between the user request and ontology concepts, datasets are found. Reasoning on the ontology content elects some services composition related to the datasets, user selects the more relevant services for his needs when multiple elements may render the same service (e.g., in the example S1 and S2). The workflow is then run, using the GEOS elements that describe the actual call to the services. Figure 4, presents important elements for services composition with bibliographic references.

In this Figure, we see a claim in the center, lower part. This claim "is qualified by" two elements, NSMB[21] and Power spectral density (PSD). It also "supports" a scientific paper. "is qualified by" and "supports" are two relations from Micropublications ontology. These two elements come from the domain description for SMS project. The PSD is then an output of a data analysis service, in this example calculated by the periodogram method. PSD and the service are linked through the relation ”isCombinedToParam”, combining the PSD and an aggregate defining the PSD together with its format and unit. This service outputs PSD, and has an input that is ”NSMB-simulated near trailing edge activation” which one can find in some of the datasets in CALLISTO Dataverse instance (e.g. ”FDV 601” dataset in this example). This dataset ”has query software” (a relation in the ontology coming from GEOS module) some software that automates its access. Therefore, it is possible to go from a claim in a scientific paper to ontological elements and from these elements to data analysis pipelines.

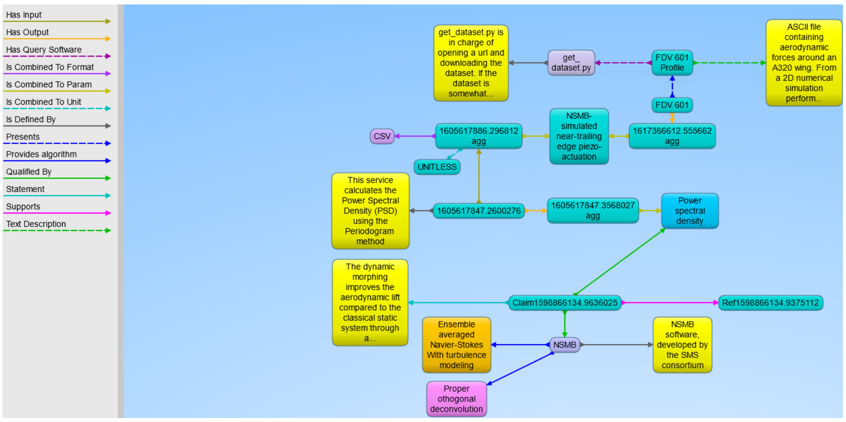

Figure 4: Elements of service composition in the ontology through browsing in Gruff, provided by Allegro

Interoperability between data and analysis software is ensured by the description of the elements in the ontology. Aggregates in ARCAS follow the specification of GEOS module of ASON [3] services ontology structure. GEOS being based upon OWL-S[22], allows describing Web and non-Web services down to the details of their implementation, the parameters to be provided and the URLs to be contacted. This structure distinguishes three key concepts to ensure interoperability of data and processing software: the unit in which a quantity is expressed, the format adopted to communicate it, and the unit itself as a concept in the ontology.

5. Validation and positioning

An extensive view of the benefits of using CALLISTO in a scientific project is given in a datapaper coming from SMS project[23]. This paper covers the functionalities: Data sharing, workflows description and overall enhancements of data sharing among project stakeholders. This illustrates some use cases benefiting from the CALLISTO’s features. As an example, the contributions of this approach concerning the services composition is the following: During the elaboration of the first processing workflow, it was necessary to define, in the ontology, relevant metadata for the definition of data access services and data analysis. This metadata was then transmitted from the ontology to the Dataverse data repository interfaces. When registering new data, the correct description of the processing and the provision of the right metadata at the time of deposit allowed the automatic inclusion, when appropriate, of the datasets in the workflow. The following Table 2 summarizes the requirements outlined in section 3 and the way CALLISTO meets them. An extensive discussion about the different platforms for scientific FAIR data sharing would go out the scope of this paper, nevertheless we indicate on Table 2 some elements of comparison with the European research infrastructure ELIXIR (www.elixir-europe.org). In the colum ”ELIXIR”, a ”A” indicates that ELIXIR adresses the same requirement (with different technical solutions) where a ”NA” indicates that, to the best of our knowledge ELIXIR does not.

Table 2: Requirements and solutions

Requirement	Solution	ELIXIR
1	Custom metadata sets subset of project-scale ontologies	A
2	Extension of metadata by referencing elements in the ontology	NA
3	Using SWO and GEOS to specify software and automation	NA
4	Using Micropublications and referencing Dataverse datasets as data supporting claims	NA
5	Using Dataverse and AllegroGraph for user-friendly frontends	A

6. Conclusion and future works

In this paper, the platform CALLISTO has been presented, which aims to support Open Science with the added value of incorporating Dataverse as datasets repository, Allegro as a Triple Store and scientific papers description by the means of Micropublications. Furthermore, CALLISTO has been improved by the adoption of a multi-disciplinary point of view, through the production of project-specific ontologies (ARCADIE) all inheriting from the same ARCAS core ontology. CALLISTO is a semantic platform for improving FAIR and multi-disciplinary scientific data sharing that relies, when possible, on proven solutions (Dataverse, well-known ontologies, Allegro) and proposes to unify these elements thanks to the ARCAS and ARCADIE ontologies. The aim of this work is to enable a better understanding of shared data, by including metadata definition elements directly in the ontology and a mechanism for transforming these ontological elements into a format compatible with data repository software, namely Dataverse. Through the use in particular of SWO and MicroPublications, CALLISTO allows describing not only the data, but also the ecosystem in which they are inserted (processing software, scientific publications...). The second important feature of CALLISTO evolution is the development of an online workspace, supported by reasoning using ontologies at the center of CALLISTO. Virtual research environment functionalities like a semi-automatic workflow generation, are available online. A reuse of `ro-crate` ids planned for the description of workflows created inside CALLISTO by the users to enhance their reusability. Finally, an extensive study of the possibilities offered by SD will be conducted.

References

[1] T. Clark, P. N. Ciccarese, C. A. Goble, Micropublications: a semantic model for claims, evidence, arguments and annotations in biomedical communications, Journal of Biomedical Semantics 5 (2014) 28. URL: https://doi.org/10.1186/2041-1480-5-28. doi:10.1186/2041-1480-5-28.

[2] P. Alliez, R. D. Cosmo, B. Guedj, A. Girault, M.-S. Hacid, A. Legrand, N. Rougier, Attributing and referencing (research) software: Best practices and outlook from inria, Computing in Science Engineering 22 (2020) 39–52. URL: https://www.softwareheritage.org/wp-content/uploads/2020/01/2020GtCitation.pdfhttps://hal.archives-ouvertes.fr/hal-02135891. doi:10.1109/MCSE.2019.2949413.

[3] T. Louge, M.-H. Karray, B. Archimède, J. Knödlseder, Ason: An owl-s based ontology for astrophysical services, Astronomy and Computing 24 (2018) 1–16.

[4] G. Xiao, D. Calvanese, R. Kontchakov, D. Lembo, A. Poggi, R. Rosati, M. Zakharyaschev, Ontology-based data access: A survey, in: Proc. of the Twenty-Seventh Int. Joint Conference on Artificial Intelligence, IJCAI-18, International Joint Conferences on Artificial Intelligence Organization, 2018, pp. 5511–5519. URL: https://doi.org/10.24963/ijcai.2018/777. doi:10.24963/ijcai.2018/777.

[5] T. Louge, M. H. Karray, B. Archimède, Z. Maamar, M. Mrissa, Semantic web services composition in the astrophysics domain: Issues and solutions, Future Generation Computer Systems 90 (2019) 185–197.

[6] T. R. Gruber, A translation approach to portable ontology specifications, Knowledge acquisition 5 (1993) 199–220.

[7] T. Berners-Lee, J. Hendler, O. Lassila, The semantic web, Scientific american 284 (2001) 34–43.

[8] N. Shadbolt, T. Berners-Lee, W. Hall, The semantic web revisited, IEEE intelligent systems 21 (2006) 96–101.

[9] M. A. Musen, The protégé project: a look back and a look forward, AI matters 1 (2015) 4–12.

[10] D. Chen, Enterprise interoperability framework., in: EMOI-INTEROP, 2006.

[11] S. A. C. Bukhari, M. Martínez-Romero, M. J. O' Connor, A. L. Egyedi, D. Willrett, J. Graybeal, M. A. Musen, K.-H. Cheung, S. H. Kleinstein, Cedar ondemand: a browser extension to generate ontology-based scientific metadata, BMC Bioinformatics 19 (2018) 268. URL: https://doi.org/10.1186/s12859-018-2247-6. doi:10.1186/s12859-018-2247-6.

[12] C. Xiao, N. Chen, X. Wang, Z. Chen, A semantic registry method using sensor metadata ontology to manage heterogeneous sensor information in the geospatial sensor web, ISPRS International Journal of Geo-Information 5 (2016) 63.

[13] R. Brisebois, A. Abran, A. Nadembega, A semantic metadata enrichment software ecosystem based on metadata and affinity models, International Journal of Information Technology and Computer Science 9 (2017) 1–13.

[14] R. Di Cosmo, S. Zacchiroli, Software heritage: Why and how to preserve software source code, 2017.

[15] J. Malone, A. Brown, A. L. Lister, J. Ison, D. Hull, H. Parkinson, R. Stevens, The software ontology (swo): a resource for reproducibility in biomedical data analysis, curation and

digital preservation, Journal of biomedical semantics 5 (2014) 25.

[16] B. Smith, Ashburner, C. M., Rosse, et al., The obo foundry: coordinated evolution of ontologies to support biomedical data integration 25 (2007) 1251–1255. doi:https://doi.org/10.1038/nbt1346.

[17] B. Smith, A. Kumar, T. Bittner, Basic Formal Ontology for Bioinformatics, IFOMIS Reports, 2005.

[18] J. Ison, M. Kalaš, I. Jonassen, D. Bolser, M. Uludag, H. McWilliam, J. Malone, R. Lopez, S. Pettifer, P. Rice, Edam: an ontology of bioinformatics operations, types of data and identifiers, topics and formats, Bioinformatics 29 (2013) 1325–1332. URL: https://doi.org/10.1093/bioinformatics/btt113. doi:10.1093/bioinformatics/btt113. arXiv:https://academic.oup.com/bioinformatics/article-pdf/29/10/1325/710075/btt113.pdf.

[19] T. Lebo, S. Sahoo, D. McGuinness, K. Belhajjame, J. Cheney, D. Corsar, D. Garijo, S. Soiland-Reyes, S. Zednik, J. Zhao, Prov-o: The prov ontology, W3C recommendation 30 (2013).

[20] Franzinc,semanticwebtechnologies http://www.franz.com/agraph/allegrograph/, ????

[21] Y. Hoarau, D. Pena, J. B. Vos, D. Charbonier, A. Gehri, M. Braza, T. Deloze, E. Laurendeau, Recent developments of the navier stokes multi block (nsmb) cfd solver, in: 54th AIAA Aerospace Sciences Meeting, 2016, p. 2056.

[22] D. Martin, M. Burstein, D. McDermott, S. McIlraith, M. Paolucci, K. Sycara, D. L. McGuinness, E. Sirin, N. Srinivasan, Bringing semantics to web services with owl-s, World Wide Web 10 (2007) 243–277.

[23] J.-B. Tô, T. Louge, C. Rouaix, C. Jimenez Navarro, A. Marouf, M. Braza, ”aerodynamic performance in transonic regime around an a320 airfoil by means of electroactive morphing through vibration and slight deformation of the near-trailing edge region at high reynolds number”, datapaper preprint (????). doi:https://dataverse.callisto.calmip.univ-toulouse.fr/dataset.xhtml?persistentId=doi:10.48531/JBRU.CALMIP/19XGRY.