D1.5 - Data Models, Taxonomy and Ontology Industrial Water Use

WP1 - Water Efficiency Enhancement Applications Framework and Baseline Assessment

20/01/2023

Authors: Aitor Corchero (EURECAT); Robert Sanfeliu Prat (EURECAT)

The AquaSPICE project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 958396.

Document Information

GRANT AGREEMENT NUMBER	958396		**ACRONYM**		AquaSPICE
FULL TITLE	Advancing Sustainability of Process Industries through Digital and Circular Water Use Innovations				
START DATE	1st December 2020		**DURATION**		48 months
PROJECT URL	www.AquaSPICE.eu				
DELIVERABLE	D1.5– Data Models, Taxonomy and Ontology Industrial Water Use				
WORK PACKAGE	WP1 – Water Efficiency Enhancement Applications Framework and Baseline Assessment				
DATE OF DELIVERY	**CONTRACTUAL**	03/2022		**ACTUAL**	01/2023
NATURE	Report		**DISSEMINATION LEVEL**		Public
LEAD BENEFICIARY	EURECAT				
RESPONSIBLE AUTHOR	Aitor Corchero (EURECAT)				
CONTRIBUTIONS FROM	Robert Sanfeliu Prat (EURECAT) Chuan Jiet Teo (KWR) Johann Poinapen (KWR) George ARAMPATZIS (TUC) Nikos SAVVAKIS (TUC) Kostas Kalaboukas (MAGGIOLI) Babis Magoutas (ICCS)				
ABSTRACT	D1.5 is mainly focused on the elaboration of a semantic model (ontology) devoted to intertwin information to water and industrial domain. The result of this deliverable is the demonstration of the implemented knowledge graph to represent the information acquired through AquaSPICE. To this end, this deliverable describes all the construction process including (i) the analysis of existing data models, taxonomy and ontologies; (ii) selection of semantic models and terms; (iii) construction of the ontology; and (iv) evaluation and validation of the constructed model. Furthermore, this version of the deliverable iterates over the initial version (03/2022) by providing details on the work carried following the next steps identified initially. More precisely, the semantic repository and data exploration tools deployed and the instantiation of the ontology for the different case studies.				

Document History

VERSION	ISSUE DATE	STAGE	DESCRIPTION	CONTRIBUTOR
V0.0	31/03/2021	DRAFT	Outline definition	Aitor Corchero (EURECAT)
V1.0	11/03/2022	VERSION	Initial version of the deliverable	Aitor Corchero (EURECAT)
V2.0	16/01/2023	REVISION	Revision of the deliverable that includes the explanation of how future work describes d on the initial version has been conducted. More precisely a description of the semantic repository, the exploration tool and the instantiation of the ontology has been added. External references revised.	Robert Sanfeliu Prat (EURECAT)

Disclaimer

Copyright message

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ABBREVIATIONS/ACRONYMS

WoT	Web of Things
CQ	Competency Question
EU	European Union
IIoT	Industrial Internet of Things
IoT	Internet Of Things
JSON(-LD)	JavaScript Object Notation (-Linked Data)
OWL	Ontology Web Language
RDF	Resource Description Framework
SPARQL	SPARQL Protocol and RDF Query Language
SPIN	SPARQL Inferencing Notation
SWRL	Semantic Web Rule Language
TTL	Turtle
WP	Work Package

1. Executive summary

This document presents the construction of the AquaSPICE ontology towards the construction of a semantic model (ontology and knowledge graph) that permits to break the existent industrial siloes to empower cross-domain decision-making and enabling industries to share internally and externally information. This paradigm in industries will serve to identify process symbiotic interlinks and support the generation of circular economy strategies through contextualised data exploration. Considering these aspects, this ontology has been constructed following an agile semantic developing methodology called SAMOD to facilitate ontology construction, documentation, and publication. For the construction of the ontology we extended a base ontology created under the H2020 project called ULTIMATE:

https://aolite.github.io/WSISOntology/

Using this base ontology, we have extended it through representative terms specifically for AquaSPICE based on the available example of the AGRICOLA case-study. This case-study, has been published also in the ontology web page as a usage of such ontology:

https://raw.githubusercontent.com/aolite/WSISOntology/main/examples/AQUASPICE-AGRICOLA.ttl

The work performed has been based on existing standards related to SAREF4WATER, WSIS ontology, NGSI-LD v2 and also, Web of Things (WoT) architecture. This interlink with reference architectures and also, with standard data models, has been included to facilitate the information interlink between FIWARE Context broker (Orion Context Broker) and the different case-studies and modules to be elaborated along with the project.

The design and implementation of the AquaSPICE semantic model add significant value to:

- Support the data representation to enable circular economy and process symbiosis strategies.
- Provide metadata and context-based information to interlink water management information with industrial process, material and energy industrial fluxes.
- Generating open linked data related to industrial process and the process symbiosis information.
- Support the construction of data models that permit to harmonize the information exchange and the production of Open APIs to explore such information.
-

2. This document is an iteration of the deliverable presented in 03/2022 adding to that initial version the outcomes of the work performed since that date. As a consequence, the present iteration describes the deployment of a semantic repository (Apache Jena Fuseki) to store the ontology instantiations of each case study and exhibits a GRAPHQL query API that allows other systems to query these ontologies. Moreover, this iteration of the document presents the integration of an user oriented tool (Rhizomer) for the exploration of said ontologies.Introduction

2.1. Scope

This document is mainly focused on the description of the semantic model (ontology) of AquaSPICE. In this regard, the envisioned ontology is envisioned as a part of the Real-Time monitoring and control system to be elaborated under the WP3. Accordingly with this vision, the initial step to facilitate data harmonization and exploration is the development of the ontology to agree in a common terminology and data context for the information. Thus, the AquaSPICE ontology (i) catalogue evidence-based practices and water re-use and recycling strategies to be sharable across platforms and industries; (ii) support system modelling and data analytics to enhance the detection of abnormal behaviors and optimize industrial processes; (iii) understand potential symbiotic links between industries. The elaborated ontology can be accessed through the following URL:

https://raw.githubusercontent.com/aolite/WSISOntology/main/examples/AQUASPICE-AGRICOLA.ttl

The AquaSPICE ontology will be one of the main elements to represent the information and the subsequent context. This ontology will support the decision-making with the interlink between water management, industrial processes and also, the use of other reusable resources in industry. Therefore, one of the main challenges of the ontology is to support the data understanding about circular economy and industrial symbiosis practices. Moreover, the ontology will be able to harmonize information exchange, understanding and exploration among different modules that compose the AquaSPICE architecture. Despite the availability of different ontologies like SIM4NEXUS ontology [1], WatERP ontology [2], OFIS [3] or e-Symbiosis [4] developed during the past years, there is the challenge to **represent wider linked information to establish a more large-scale, representing more operational information about the process to enable the generation of monitoring and control tools**. Also, there is a need to disaggregate the information from site-specific information, making the ontology like a **catalogue of potential process representation, material representation and also a model to detect waste resource events**. Moreover, latest developments on the field for enabling the approaching of industrial symbiosis strategies have been deprecated and not further maintained. This impedes to share knowledge and data using common vocabularies and digital industrial tools. Therefore, this situation has derived to the highlight of industrial siloes and the limitation of the circular economy solutions implementation, scalability and also, transferability.

Recent trends in ontology modelling are confluence to the use of standard data models and ontologies represented by NGSI-LD specification and SAREF ontological ecosystem. At a more industrial level, there are proliferating ontologies related to raw materials and materials modelling. In this regard, there are published standard ontologies like MODA, CHADA, OSMO [5]or EMMO. All these semantic model representations are maintained by relevant institutions like the European Commission, EMMC, or ETSI standardization body. Using this practice, the long-term maintenance of the ontologies is ensured.

Considering these aspects, key innovation aspects of the AquaSPICE ontology is the extension of the WSIS ontology elaborated under H2020 EU funded project towards the **representation of industrial processes and data**. This innovation derives in the informational interlink between resource parameters, process parameters and also process/industries that consumes or produces certain resources. This will definitely contribute to aid decision-making and **establish a set of semantic rules and axioms that facilitate the detection of waste processes or circular strategies hotspots.** Considering these aspects, the present deliverable will:

- Analyse the existing semantic development methodologies to select the most appropriate for the development of AquaSPICE semantic model.
- Describe the requirements collected to build the AquaSPICE semantic model.
- Describe the current version of the ontology and the extension performed over the WSIS ontology.
- Describe the mechanism to test the ontology and the initial dataset loaded into an open shared repository.
- Describe the mechanism to publish the ontology.

The work described under this document corresponds to the efforts performed under WP1 and specifically, under the Task 1.4 entitled as "*Data Models, Taxonomy and Ontology for Water Use in the Process Industry*". In detail, this task aimed at *"convey environmentally, water and industrial informational siloes through a knowledge graph (ontology) that permit systems to (i) catalogue evidence-based practices and water re-use and recycling strategies to be sharable across platforms and industries; (ii) support system* *modelling and data analytics to enhance the detection of abnormal behaviours and optimise industrial processes; (iii) understand potential symbiotic links between industries. For the elaboration of the knowledge graph, the initial action is to analyse semantic models (SAREF4WATR, SSN, SIM4NEXUS Ontology, etc.) and existing datasets and data exchange formats for industrial domain. This analysis will be used to select appropriate terms from existing resources and to align informational requirements with standard representation of the information. A taxonomy of terms will be created based on the initial analysis. This taxonomy will incrementally be enhanced and enriched with properties.*". The result performed under this task will serve as an input for the WP3 (Real-Time Water Efficiency Monitoring Platform for the Process Industry), WP4 (Digital Twin with Smart Analytics and Cognitive Services for Water Efficiency) and WP5 (Water Cyber Physical System: The Industrial Water Efficiency Management System).

2.2. Document Structure

The document is organised as follows:

- Section 3 presents the different semantic development methodologies.
- Section 4 describes the main ontology use-cases and requirements.
- Section 5 focuses on the ontology development and the corresponding parts.
- Section 6 describes the ontology validation and testing.
- Section 7 describes the ontology documentation and publication.
- Section 8 describes the semantic repository deployed to store the ontology instantiation for the different case studies.
- Section 9 presents the UI tool that allows the exploration of these ontologies.
- Section 10 briefly shows the ontologies of the different case studies.
- Section 11 concludes the document and discusses the future work.

3. Semantic Development Methodologies

This section focuses on the selection of a suitable ontology development methodology for the AquaSPICE project. The construction of the semantic model in AquaSPICE will comply with the envisioned requirements and, for that, we need to adopt a methodology to ensure all requirements are considered within the ontology modelling. The methodology must provide an efficient way to construct the ontology through different development cycles that comprises the development, testing and publication of the ontology.

Several methodologies have been proposed for the construction of ontologies in the semantic web domain. Among these methodologies are the predictive ontology development methodologies (e.g., Methontology [6], On-To-Knowledge [7] and NeoN [8]). Newer methodologies comprise the application of adaptive ontology development methodologies to build the ontology, based on agile (e.g., SAMOD [9]). Regarding the predictive methodologies, they are mainly focused on the use of controlled and specific phases to elaborate the ontology. A predictive planning strategy may fail when confronted by significant project specification changes or customer modifications, but it will also be more likely to generate the anticipated result. The development starts with the requirements gathering and finishes with the ontology implementation and publication. Such ontology construction methodologies are similar to iterative software development process methodologies (Figure 1).

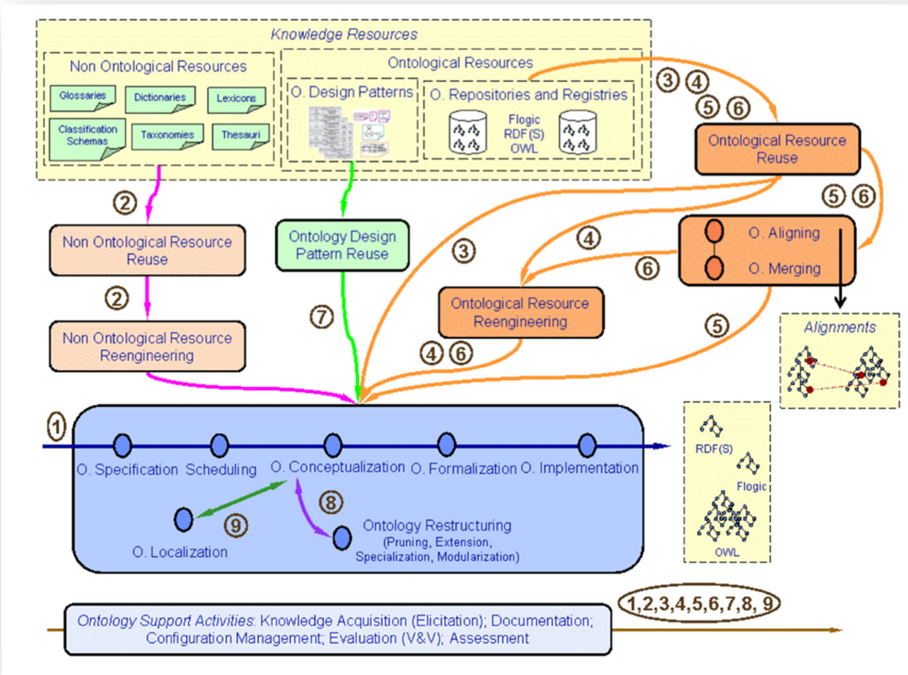

Figure 1. NeoN terative methodology process [8]

On the contrary, adaptive software methodologies are mainly focused on the high-uncertainty in ontology development, giving short-development cycles to adapt to future customer needs. In the end, these methodologies are more susceptible to the ontology development evolution. These kinds of methodologies mainly provide for the elaboration of different cycles to build the ontology. Each cycle comprises the definition of a scenario and then their corresponding modelling based on the different user stories (requirements) generated. At the end of each cycle, a newer version of the ontology is produced, enabling an agile-based extension of the semantic model (Figure 2).

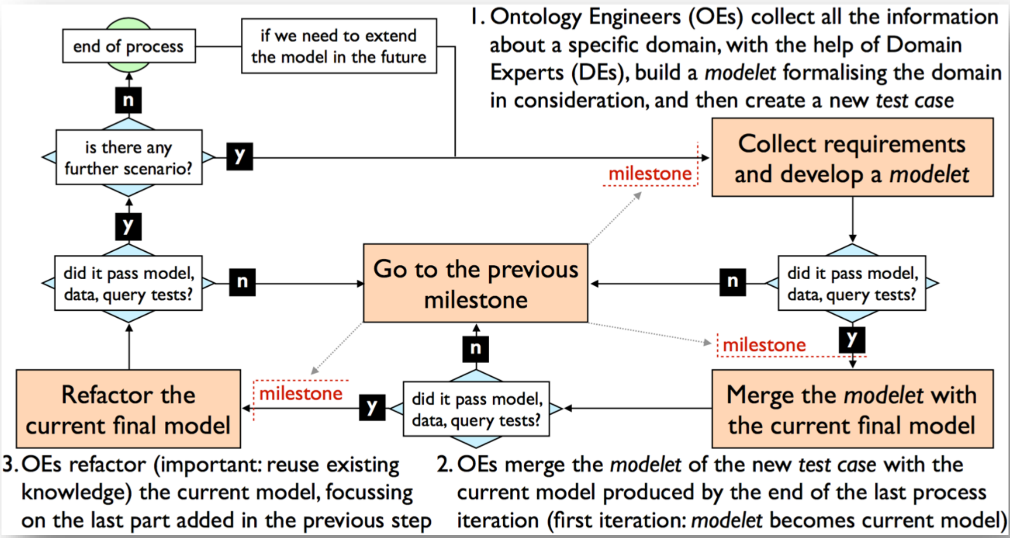

Figure 2. Predictive ontology development methodology

3.1. Adopted Methodology in AquaSPICE

Considering these types of methodologies, we selected the SAMOD methodology for the elaboration of the AquaSPICE semantic model. This ontology development methodology offers AquaSPICE ontology the ability to perform enhancements of the ontology in parallel with the developments of the various modules and tools. Hence, this methodology will allow evolving the data model and the knowledge representation model in parallel to the elaboration of the real-time monitoring and control tool (WP3) and the rest of AI and cloud models developed under WP4 and WP5. Therefore, the AquaSPICE ontology will contribute to the elaboration of the data models for the FIWARE compatible architecture to be developed under the project. Moreover, the SAMOD ontology development technique facilitates the elaboration of large-scale ontologies, providing for the collaboration between different teams (in this case, knowledge engineers and domain experts), reusability and the possibility of re-engineering knowledge resources, with the ability to reuse and incorporate other semantic models (e.g., SAREF4WATR, WATERP, GEOSPARQL, etc.) and other data models and schemas (e.g., Smart Water Data models). A summary of the comparison of the chosen methodology against others considered is presented in Table 1, along with certain attributes.

AQUASPICE Needs	Methontology	DILIGENT [10]	On-To Knowledge	NeoN	SAMOD
Collaboration between different teams	Partially	YES	NO	YES	YES
Inclusion of existent ontologies (or standard ontologies)	YES	NO	NO	YES	YES
Re-engineering of knowledge resources	NO	NO	NO	YES	YES
Continuous enhancement and evolution of the ontology	Linear	Based on user's role	Linear	Modules	Cycles
Validation methodology for the ontology	Evaluation based on technical elements, documentation, and environment (Not Detailed)	NO	Competency Questions and Application environment	Competency Questions, Documentation Validation and Technical elements of Ontology	Competency Questions and user-story requirements based on scenarios
Aligned with large ontology construction	YES	YES	YES	YES	YES
Support guided ontology construction (Requirement	YES	Based on user's Role	YES (based on processes)	YES	YES, based on cycles.

s, Design, Implementation)					

Table 1 – Ontology development methodologies comparison

Based on the analysis presented in Table 1, the adopted methodology for the semantic model development is presented in Figure 3. The process of each cycle starts with the collection, identification, and selection of the requirements (user stories) to be implemented. This process involves the ontology developers, the different case studies, and technology providers to identify data modelling needs for the elaboration of the FIWARE compatible architecture in AquaSPICE. These aspects are modelled in the form of requirements and competency questions to enable continuous integration and publication of the semantic model (ontology). Once the main framework is selected, the second phase of the cycle relies on the implementation of the ontological part. In the case of the initial model, this part corresponds to all the initial structures envisioned. This part generates the proper ontology file in OWL/RDF/N3 notation. When this is ready, it will be merged with the rest of the semantic model. This merging is performed under a GitHub repository to facilitate conflict resolution of the code. When completed, the ontology is tested using OOPS! And FOOPS! Technology and is documented and published and documented using the LODE technology.

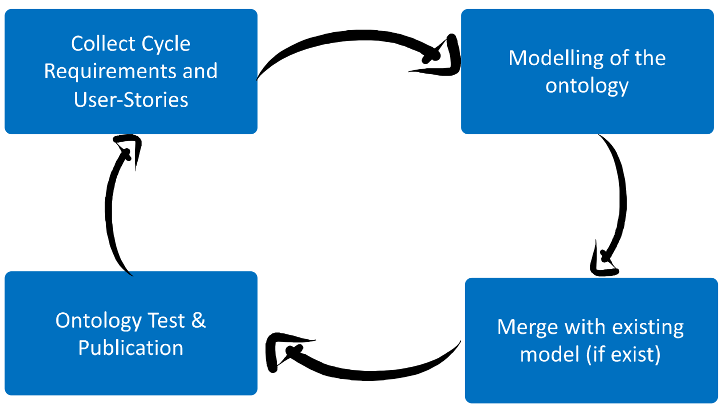

Figure 3. AquaSPICE ontology development methodology

The described cycle corresponds to one cycle of the methodology. We envision to execute three different cycles in the framework of the ontology development. The initial cycle (the one that is currently described in this deliverable) corresponds to the initial version of the ontology that includes the main structural aspects of the semantic model aligned with the requirements exposed in the D1.2. "User Analysis, Use Cases Requirements and Quality Criteria". This initial cycle will comprise the following main features: (i) catalogue evidence-based practices and water re-use and recycling strategies to be sharable across platforms and industries; (ii) support system modelling and data analytics to enhance the detection of abnormal behaviours and optimise industrial processes; (iii) understand potential symbiotic links between industries; (iv) compatibility with other representative data models as NGSI-LD and Web of Things. This initial model of the AquaSPICE ontology was developed during the period of time between M8 and M16 of the project. From M16-M24 the model was used to define an initial instantiation of the ontology for all the case studies. Furthermore, the semantic data exploration tool to navigate through the information was incorporated. Continuing from that stage, last cycle will be focused on the adaptation of the instantiation of the ontology for each case study to adapt it to possible changes, expanding the model if necessary.

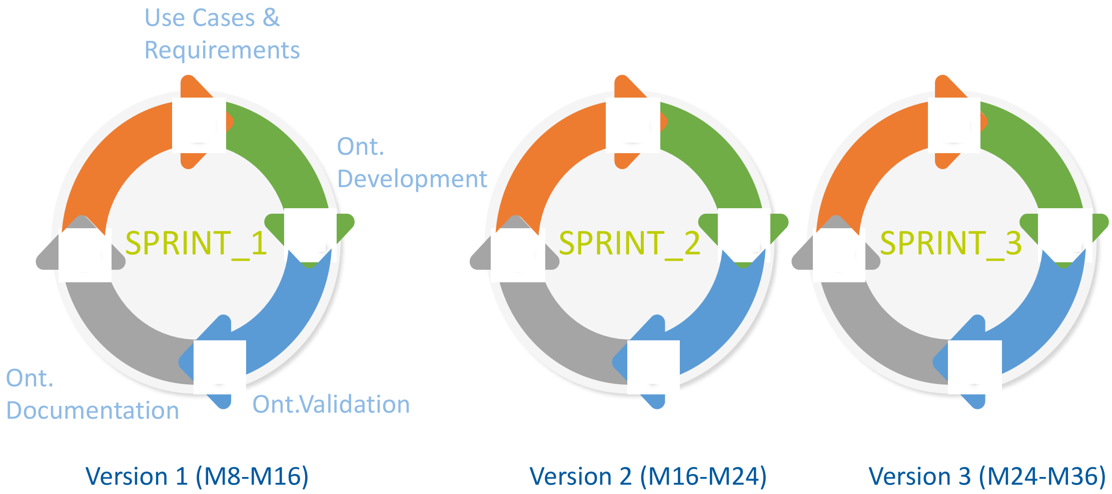

Figure 4. AquaSPICE Ontology and Semantic Repository Sprints

4. AquaSPICE Ontology Requirements and use-cases

The main purpose of the AquaSPICE semantic model is to interlink industrial processes and industries to better understand symbiotic pathways and resource reuse strategies. For that, we have to integrate contextual information like process information, industry location and also type of industries involved. Complementary, we need to integrate operational information about the processes covering aspects like databases, sensors, Internet of Things (IoT) devices, monitoring and control tools, etc. The proposed ontology should be able to harmonize different type of information as same time as reasoning about processes to detect complex operational events that derive in resource waste and inefficiencies. The ontology will provide a common data understanding of this information based on the context (their use). Moreover, semantic enriched data representation should be aligned with current standards from the integrated domains.

Considering these aspects, the main purpose of the AquaSPICE ontology is to provide a complete semantic model about all knowledge and decision-making procedures related to the interlink between the industries, process and operational information to contextualize and support the elaboration of circular and industrial symbiosis strategies.

4.1. Scope

The main scope of the AquaSPICE ontology is to represent both types of information (i) Context-based information about the processes and industries; and (ii) represent the information from multiple systems that are observing a heterogeneous number of industrial systems, water systems and energy systems that takes part of the AquaSPICE case-studies. These variables and information will conform the semantic modelling and contextualization of operational industrial processes linked with industrial symbiosis strategies and technologies. The type of the variables in most of the cases represent the devices and industrial systems related to process monitoring in a broad of domains (industry, farming, textile, etc.). So, one of the main aspects of the ontology is to make understandable the information collected from the different devices and harmonize such information under a common data model compatible with FIWARE. FIWARE is a curated framework of open-source platform components to accelerate the development of smart solutions [11]. This compatibility is required due to the adoption of FIWARE as a reference architecture for data ingestion, collection, knowledge discovery and visualization. Based on this, the ontology needs to follow standard definitions and terms adoptions from representative organizations (World Meteorological Organization -WMO-, OGC, INSPIRE, W3C, FIWARE, etc.).

4.2. Implementation Language

The selected implementation language for the ontology is OWL, the Ontology Web Language. The serialization format will be performed with the use of the Turtle (TTL) language and JSON-LD serialization to expose the corresponding information and variables acquired and stored in the different data sources as the semantic repository and non-structured databases. In case required, we will explore the possibility of using semantic rules to perform ontology alignments (using SPARQL Inferencing notation -SPIN- [12], Semantic Web Rule Language -SWRL- [13]), and also model water management, energy management and material management events that occur at the operational level inside the industries (e.g. alerting about water, energy and material wastes at operational level).

4.3. Uses of the AquaSPICE ontology

The potential end-users that will consume and explore the information modelled under the AquaSPICE ontology are represented in the Table 2.

User Group (UG)	Name	Description
UG-1	Industrial Managers	Group focused on (i) managing processes inside industries; (ii) control efficiency in the process; and (iii) monitor resources waste
UG-2	Industrial Operators	Group focused on (i) controlling the different processes; (ii) defining local parameters of the industrial processes; (iii) determining industrial process thresholds; and (iv) monitoring industrial infrastructure.
UG-3	Water Authorities	This group will be focused on: (i) elaborating water management and water quality policies; (ii) determining hazardous contaminants and spills.
UG-4	Water Managers	This group will be focused on: (i) determining water management procedures; (ii) analysing water related indicators; (iii) identifying water management alerts/anomalies
UG-5	Water Operators	This group will be focused on: (i) defining water parameters for the operations; (ii) determining water management thresholds; (iii) monitoring water infrastructure inside industries too.
UG-6	Risk managers	This group is devoted to (i) elaborate industrial risk plan and specially in industries with harmful materials, (ii) interlink vulnerabilities with contextual industrial information.

User Group (UG)	Name	Description
UG-7	Farmers	This group is devoted to (i) monitoring and control farming production; (ii) establish thresholds for ensuring the production; (iii) reuse water that permits to create an efficient production and reduce costs.
UG-8	Society/Citizens	This group is aimed at (i) exploring circular economy strategies; (ii) visualise the benefits of the industrial symbiosis and circular economy to the region.

Table 2 – User groups represented in the AquaSPICE ontology

4.4. AquaSPICE Use-Cases

Following the SAMOD methodology, the use-cases (or test-cases) identified have been collected from the work performed under WP1 entitles as "Water Efficiency Enhancement Applications Framework and Baseline Assessment". Specifically, the use cases have been aligned with the work performed under the D1.2. "User Analysis, Use Cases Requirements and Quality Criteria" and D1.3. "Use Cases Definition with Baseline Assessment". Considering these sources of information, the use-cases identified for the elaboration of the ontology are represented in the following tables:

Name	Water Management Information Exchange
Description	In AquaSPICE, capturing heterogeneous information coming from water management infrastructure to analyse current scenarios of water management in urban areas and industries. This will comprise the information collection of the state of the lakes, tanks, pumps, etc. It should comprise the information collection of the supply and distribution value chain of the water infrastructure
Example-1	There is a need to represent the (waste) water tank levels
Example-2	There is a need to determine current state of the pumps

Table 3 – Use-Case of Water Management

Name	Process interlink and management
Description	In AquaSPICE, there is a need to contextualise the industries and also, to characterise the resource providers and resource consumers. In this regard, the consumption and supply of resources should be represented at both, industrial and process level.
Example-1	Determine the type of resource inputs and outputs in the process
Example-2	Determine the quantity of resources needed

Table 4 – Use-Case of process interlink

Name	Energy Management Information exchange
Description	Consideration of the energy fluxes in the processes and industries. In this regards, there is a need to consider energy consumption and supply for the processes
Example-1	Consideration of energy supply and consumption by each of the processes.

Table 5 – Use-Case of Energy Management

Name	Material management in the industries

Description	Consideration of the material fluxes in the processes. In this regards, there is a need for the determination of the material supply and generated in the industrial processes
Example-1	Consideration of material supply and consumption by each of the processes.
Example-2	Categorization of the materials at industrial level

Table 6 – Use-Case of Material management

4.5. AquaSPICE Uses

Based on the different use-cases identified, we consider the following ontology uses that will serve to conform the actions to be represented inside the different types of requirements. Thus, the following uses are envisioned for the AquaSPICE ontology:

Uses	**Concept**	**Description**
Use-1	Integrate heterogeneous information (systems)	Pull and Push information from different data sources (Sensor-based information, BBDDs, decision-making tools, thematic/general models, hydrometeorological models, wearables, etc.) coming from water infrastructure, energy infrastructure or industrial processes.
Use-2	Standard-based serializations	Provide the information to the systems and visualization engine using standard-based serializations coming from water, energy, materials and industrial level.
Use-3	Navigate through the information	Permit to navigate throughout the information in order to support semantic facet navigation
Use-4	Navigate through geospatial information	Provide navigation through different maps and types of layers.
Use-5	Determine input/outputs process resources	Provide information about the inputs and outputs of processes and industrial resources.
Use-6	Determine symbiotic states	Determine symbiotic states between processes and industries.

Table 7 – AquaSPICE ontology uses

4.6. AquaSPICE Requirements

This part of the ontology development aims at describing the functional and non-functional requirements that the ontology must fulfil. The identified functional requirements are represented in the form of Competency Questions (CQs). These questions need to satisfy the user requirements and serve as ontology validations about the informational uses of the representative users. The non-functional requirements of the ontology describe the extra features that the semantic models need to satisfy (e.g., multi-language representation). As a result, the informational requirements of the user are identified and planned to be included in the ontology to be elaborated. This will establish the base of the entire platform and informational navigation of the semantic repository.

4.6.1. Functional Requirements

For the elaboration of the functional requirements, we adopted a bottom-up approach for a better understanding of the domain. Specifically, we have involved the different users and information throughout the AquaSPICE case-studies. and technology providers of the different modules and FIWARE-compatible tools (Real Time Monitoring Tool from WP3 based on the context broker). This involvement has been produced coming from the different studies and questionnaires performed under WP1, offline information exchange and plenary meetings. This interest has been translated into competency questions that include for each of the presented questions, the feature (question) and an example of the corresponding answer from the system. The identified competency questions are:

Id	Competency Questions	Answer
ACQ-1	Which assets are involved in the water/energy/industrial infrastructure?	Water pipe, energy storage system, DAF, etc.
ACQ-2	What types of sensors are involved in the infrastructures?	Energy meter, water meter, water flows, tank level
ACQ-3	What type of actuators are involved in infrastructures?	Pumps, valves, etc
ACQ-4	What type of processes are involved in the demo-case X?	Farming Operation, Oil refining, etc
ACQ-5	What type of technology have been implemented in the industry X?	Screening, Mixing Tank, etc
ACQ-6	What is the volume of water exiting process X?	123L
ACQ-7	What is the energy flux in process X between 01-01-2022 and 01-03-2022?	01-01-2022 200Kwh, 02-01-2022 203Kwh, 03-01-2022 190Kwh, …
ACQ-8	What are the physical parameters of measurement X?	Energy flux, water flux, etc

Table 8 – CQs related to industrial symbiosis

Id	Competency Questions	Answer
ACQ-9	What is the region of the industry X?	POLYGON (10,10,10,10)
ACQ-10	What are the industries related to the industrial region X?	Resource provider X, resource consumer Y
ACQ-11	What kind of technology takes part of industry X?	Technology A, technology B, etc
ACQ-12	What kind of resources are needed in process X?	Water, energy, etc
ACQ-13	How much resource are required in process X?	200kwh of energy
ACQ-14	What kind of resources are generated by process X?	Water, energy, etc
ACQ-15	How much resource are generated by process X?	100 m3 of water

Table 9 – CQs related to process interlink

4.6.2. Non-Functional Requirements

The proposed non-functional requirements are focused on determining transversal functionalities that the ontology should fulfil. In this regard, the following table represents the non-functional requirements for the presented ontology:

Id	Competency Questions
NFR-1	The ontology needs to be compatible with water standards for information exchange as WaterML2, NGSI-LD, WoT or INSPIRE.
NFR-2	The ontological terms should have their specific definition
NFR-3	The ontology must support a multilingual information representation in the following languages: English, and case studies local languages

Table 10 – Non-functional requirements of the ontology

5. AquaSPICE Ontology Development

This part of the AquaSPICE ontology development comprises the construction of the ontology. In this regard, this specific section will initially cover other relevant initiatives considered for the ontology development (Section 5.1) including relevant associations, H2020 projects and standards. These initiatives have been used to select a more relevant glossary of terms to be exposed under Section 5.2. represents the different parts of the ontology to cover the different requirements exposed within Section 5.3.

5.1. Relevant Initiatives

This section mainly represents the relevant initiatives used for the elaboration of the glossary of terms and relevant definition of the AquaSPICE ontology. For that, we revised relevant associations, H2020 projects and standards that contain data specifications, data models and ontologies.

5.1.1. Standardization activities

Standardization Initiatives	Description
ETSI STF 534	This group launched by ETSI is devoted to the creation and maintenance of the SAREF extensions to the domains of smart cities, smart industry and manufacturing, smart agri-food, smart water and environment. All of these are under the umbrella of SAREF to integrate semantically information about vertical domains. The terms of this initiative will be considered for the construction of AquaSPICE, specially SAREF4WATR and SAREF4ENV.
ETSI-GS-CIM-001	This standard refers to the Context Information Management specification under the so-called NGSI-LD specification. This specification is mainly devoted to allow users to provide, consume and subscribe to context information in multiple scenarios and involving multiple stakeholders. It enables close to real-time access to information coming from many different sources (not only IoT data sources).
Open Geospatial Consortium (OGC)	The Open Geospatial Consortium has defined WaterML 2.0, an information model for the representation of water observations data, with the intent of allowing the exchange of such data sets across information systems (https://www.opengeospatial.org/standards/waterml). Moreover, the latest version of WaterML 2.0 (part 3) includes the informational model called HY_FEATURES that is focused on representing hydro-science and water network topology. The latest developments

Standardization Initiatives	Description
	comprise the elaboration of Sensor Things API (https://www.ogc.org/standards/sensorthings). Complementing this information, the OGC also offers CityGML ADE, for the representation of utility networks in 3D city models (http://www.citygmlwiki.org/index.php/CityGML_UtilityNetworkADE)
OneM2M	IoT standard devoted to the elaboration of IoT reference architecture. This initiative contains ontological models for the elaboration of syntactic and semantic interoperability of IoT Systems. Additionally, it also includes XML schemas.
ISO/TC 184	ISO standard focused on automation systems and their integration for design, sourcing, manufacturing, production and delivery, support, maintenance and disposal of products and their associated services
ISO/TC 184/SC 4	ISO standard focused on industrial data standards.

Table 11 – Standardization Activities

5.1.2. Associations initiatives

Associations Initiatives	Description
AIOTI	The Alliance for Internet of Things Innovation (AIOTI) is a multi-stakeholder platform for inspiring IoT Innovation in Europe that combines large and small companies, academia, start-ups and scale-ups, end-users, policymakers and representatives of society in an end-to-end approach (https://aioti.eu/). Two of the AIOTI working groups are relevant for three work: Smart Water Management, Manufacturing and IoT Standardization (semantic interoperability)
Water Europe	The Water Europe (http://watereurope.eu/) mission is to foster collaborative, innovative and integrated European research and technologies development, to ensure the European growth and competitiveness of the water sector, to provide global answers to global challenges for the next generations, and to address the challenges of integrated and sustainable management of water resources.
ICT4WATER cluster	The ICT4Water cluster is a hub for EU-funded research and innovation projects developing digital innovations for the water sector (https://www.ict4water.eu/). It brings projects together supporting them to exchange information and best practices, disseminate and

Associations Initiatives	Description
	exploit project outputs, contribute to defining digital water strategies, and contribute to policy development in the digital and water domain.
IWA	The International Water Association (IWA) is an open platform in which both innovators and adopters of new technologies and approaches can generate creative friction; it is a place for diffusion, benchmarking and evidence (https://iwa-network.org/). The IWA develops research and projects focused on solutions for water and wastewater management, organizing events that bring the latest science, technology and best practice to the water sector at large, and working to place water on the global political agenda and to influence best practice in regulation and policymaking.
Corss-Synergies	The cross-synergy working group is the alliance of the latest H2020-SC5-11 projects to elaborate common actions on ontologies and data models, sensors, AI data-driven models, business and communication and dissemination activities. Concerning the AquaSPICE ontology, interests are on the Smart Data Models (https://www.fiware.org/developers/smart-data-models/) alignment with the reference architecture and the ontology.
EFFRA	The Made In Europe partnership will be the voice and driver for sustainable manufacturing in Europe. It will boost European manufacturing ecosystems towards global leadership in technology, circular industries and flexibility. The Partnership will contribute to a competitive, green, digital, resilient and human-centric manufacturing industry. It will be at the centre of a twin ecological and digital transition, both a driver and subject to these changes.
ONTOCOMMONS	OntoCommons is an H2020 CSA project dedicated to the standardisation of data documentation across all domains related to materials and manufacturing. OntoCommons lays the foundation for interoperable, harmonised and standardised data documentation through ontologies, facilitating data sharing and pushing data-driven innovation, to bring out a truly Digital Single Market and new business models for European industry, exploit the opportunities of digitalisation and address sustainability challenges.
BRIDGE	BRIDGE wants to provide field experience, feedback and lessons learned from the participating projects to help overcome the barriers to effective innovation. It aims at gathering coordinated, balanced and coherent recommendations to strengthen the messages and maximize their impacts towards policy makers in view of removing barriers to innovation deployment.

Table 12 – Association Activities

5.1.3. European Projects initiatives

EU project activities	Description
H2020-SIM4NEXUS	SIM4NEXUS **¡Error! Referencia de hipervínculo no válida.**aims to predict society-wide impacts of resource use and relevant policies on sectors such as agriculture, water, biodiversity and ecosystem services through a model-based analysis. This project aims to adapt existing knowledge and develop new expertise on the Nexus; to reduce uncertainty, and to show the implementation by a network of regional and national cases.
H2020-ULTIMATE	ULTIMATE (https://ultimatewater.eu) aims to operating as a catalyst for Water Smart Industrial Symbiosis (WSIS), in which water/wastewater plays a key role within a dynamic socio-economic and business oriented industrial ecosystem. ULTIMATE will demonstrate the multiple uses of municipal and industrial wastewater through nine high-level demonstrations in Europe and the south-eastern Mediterranean from the agro-food processing, beverage, heavy chemical/petrochemical and biotech industries.
H2020-B-WATER-SMART	B-WATER-SMART (https://b-watersmart.eu) aims to speed up the transition to water-smart economies and societies in coastal Europe and beyond. To achieve this, it will adopt a large-scale systemic innovation approach to select, connect and demonstrate tailored solutions for multiple users and sectors. It will further create new business models based on circular economy and water-smartness
H2020-FIWARE4WATER	This project began in June 2019 and intends to link the water sector to FIWARE© by demonstrating its capabilities and the potential of its interoperable and standardized interfaces for both water sector end-users (cities, water utilities, water authorities, citizens and consumers), and solution providers (private utilities, SMEs, developers).
LIFE-eSymbiosis	The project aimed to develop a knowledge-based service that will promote, demonstrate and advance Industrial Symbiosis (IS) in Europe.
H2020-SMARTFAN	SMARTFAN (https://www.smartfan-project.eu) aims at the micro and Nano components, which will be used due to their special physico-chemical properties, in order to develop smart (bulk) materials for final application on intelligent structures. CFs for reinforcement and conductivity variance, CNTs and CNFs for sensing, Micro-containers for self-healing, Electro-Magnetic nanoparticles for fields detection and shielding, colouring agents for marking cracks and defects, piezoelectric materials can be the base for manufacturing new smart materials.

EU project activities	Description
H2020-EMMC-CSA	EMMA-CSA aimed at establishing current and forward looking complementary activities necessary to bring the field of materials modelling closer to the demands of manufacturers (both small and large enterprises) in Europe. The ultimate goal is that materials modelling and simulation will become an integral part of product life cycle management in European industry, thereby making a strong contribution to enhance innovation and competitiveness on a global level

Table 13 – AquaSPICE related projects

5.1.4. Relevant Ontologies

Ontologies Initiatives	Description
WAWO+ ontology	Waste Water Ontology (WaWO) (https://upcommons.upc.edu/handle/2117/106298) has been developed to model information about wastewater treatment tasks through the definition of the basic terms and relations comprising the vocabulary of the wastewater treatment area. It is the demonstration of a shared understanding of the wastewater domain that is agreed among several agents: mainly, experts in chemical, environmental and engineering. The main aim of the WaWO ontology is to design a model that: presents a vocabulary for the wastewater domain that all agents can jointly use and understand by describing the meaning of each term in an as unambiguous and precise manner as possible.
SAREF4WATR	This ontology is mainly focused on representing water infrastructure and the corresponding measurements. Despite this focus, the ontology considers water quality monitoring and also, the alignment of different water quality events with different types of waters.
SAREF4ENVI	SAREF4ENVI has two main aims: on the one hand, to be the basis for enabling the use of SAREF in the environment domain and, on the other hand, to exemplify how to enable interoperability between environmental devices in cooperation.
SAREF4INMA	SAREF4INMA focuses on extending SAREF for the industry and manufacturing domain to solve the lack of interoperability between various types of production equipment that produce items in a factory and, once outside the factory, between different organizations in the value chain to uniquely track back the produced items to the corresponding production equipment, batches, material and precise time in which they were manufactured
SAREF4ENER	SAREF4ENER focused on extending SAREF to the representation of energy management in buildings and other infrastructures. It was the

Ontologies Initiatives	Description
	initial ontology in which part of the ontology have derived on the ontology today we know as SAREF.
WSIS	WSIS ontology making reference to model industrial symbiosis and also, related indicators to the different regionalities.
OFIS	The ontological framework for Industrial Symbiosis (IS) exploits semantic knowledge modelling and enables structural data transformation for identification of potential synergies between various industries and hence formation of one to one and complex symbiotic networks.
e-Symbiosis	A new ontology framework for Industrial Symbiosis by pioneering the use of ontology engineering in the field. Semantics are used to model Industrial Symbiosis flows, to model enabling technologies and to systematise the development of a matching service.
MODA	A web portal devote to the categorization of models equations and simulations in relation to industrial symbiosis and process modelling.
CHADA	A data management portal and framework devoted to material characterization based on datasets that covers (i) sample, (ii) method, (iii) raw data and (iv) data analysis as the main component of the metadata associated to any characterization experiment.

Table 14 – AquaSPICE related ontologies

5.1.5. Other Initiatives

Ontologies Initiatives	Type	Description
INSPIRE	Directive	The INSPIRE Directive aims to create a European Union spatial data infrastructure for EU environmental policies or activities which can have an impact on the environment (http://inspire.ec.europa.eu/). This European Spatial Data Infrastructure will enable the contribution of environmental spatial information among public sector organizations, providing public access to spatial information across Europe and assisting in policy-making across boundaries.
EU WFD	Directive	The EU Water Framework Directive (WFD), adopted in 2000, requires effective water management and helps Member States prepare for extreme weather events which, due to climate change, are becoming more frequent and cause tremendous damages (https://environment.ec.europa.eu/topics/water/water-framework-directive_en).

Ontologies Initiatives	Type	Description
WISE	Data Repository	The Water Information System for Europe (WISE) is a partnership between the European Environment Agency and European Commission (DG Environment, Joint Research Centre and Eurostat) (https://water.europa.eu/). It is a collective database created around the subject of water management in Europe that holds all the significant information on this matter, hence composing a new, comprehensive, shared EU data and information system for water, with river basins. It presents water-related information and data organized in four broad areas: projects, policy, links, themes, and data.
GEOSS	Data Repository	The Global Earth Observation System of Systems (GEOSS) is a collection of independent, coordinated, Earth observation, information and processing systems that relate and provide access to heterogeneous information for a huge range of users in both private and public sectors (http://www.geoportal.org/).
Copernicus	Data Repository	Copernicus is a European system for controlling the Earth and is coordinated and maintained by the European Commission (https://www.copernicus.eu/). The development of the observation infrastructure is processed under the aegis of the European Space Agency for the space segments and by the European Environment Agency and EU countries for the in-situ component. It is a collection of complex systems which gather data from different sources: earth observation satellites and in situ sensors such as airborne sensors, sea-borne sensors, and ground stations.

Table 15 – AquaSPICE data repositories to be interlinked

5.2. Glossary of Terms

After the compilation of the information from the mentioned initiatives, this section describes main glossary of terms and their alignment with existing ontologies, data models, vocabularies, and standards. Hence, the following table compiles the primary study about the vocabulary used during the definition of the competency questions. This used vocabulary has also been used in the design and modelling of the AquaSPICE ontology:

Concept	Terms	Ontology/Data Model
A reference to time appears	The word 'period' appears	Time Ontology
	Some of the following expressions appear: 'how many times', 'how often'	
	Some of the following nouns appear: 'date', 'hour', 'minute', 'second', etc.	
A reference to digital systems appears and the corresponding measurements	The word 'sensor', 'device' or 'system' appears	SAREF4WATR/SAREF4ENV/OneM2M
	Some of the following expressions appear: 'indicators value', 'observations', 'measurements', 'indicators', etc	
	Some of the following nouns appear: 'observes', 'observedProperty', etc	
Industrial Symbiosis	Some of the following expressions appear: "Resource Provider", "resource consumer", "Industry", "Process"	OFIS/eSymbiosis/WSIS
	The word "Industry" appears	
Units and Geographical Information	Some of the following expression appears: 'measures located in', 'events in indicators/scenarios'	QUDT, W3CGeo, GeoSPARQL
	Some of the following aspects are considered: 'units attached measures'	

Table 16 – AquaSPICE Glosary of terms

5.3. AquaSPICE Ontology

This section is mainly devoted to the presentation of the AquaSPICE ontology. The work comprises the development of the ontology focuses on the design of the ontology and the subsequent implementation in OWL format.

The construction of the ontology considered the requirements (competency questions) covered within Section 4 concerning the industrial symbiosis and also process contextualization to interlink industrial heterogeneous information. As mentioned, the main advancement of the ontology relies on the catalogue of potential process representation, material representation and also a model to detect waste resource events. Additionally, the presented ontology is also devoted to interlink smart data models (of FIWARE) with the corresponding context-broker and other modules implemented across WP3, WP4 and WP5.

Based on these aspects, the core ontology is presented in the Figure 5. This ontology represents main digital and physical assets related to the representation of industrial processes and also, the resources interlink with each of the process. Therefore, the presented ontology makes a representation about the industrial digital systems with interlink with materials management and specification, water infrastructure and also, energy infrastructure. For enabling this representation, it is indicated the provenance ontologies (e.g., WISIS ontology -wsis- or SAREF4WATR -s4watr-). The advancements of the AquaSPICE ontology have been included inside the WSIS ontology (wsis) as enabling to a common framework for industrial symbiosis around EU.

PREFIX	Namespaces
dcterms	http://purl.org/dc/terms/>
Owl	http://www.w3.org/2002/07/owl#
rdf	http://www.w3.org/1999/02/22-rdf-syntax-ns#
xsd	http://www.w3.org/2001/XMLSchema#
rdfs	http://www.w3.org/2000/01/rdf-schema#
saref	https://saref.etsi.org/core/
s4city	https://saref.etsi.org/saref4city/
vann	http://purl.org/vocab/vann/
s4syst	https://saref.etsi.org/saref4syst/
time	http://www.w3.org/2006/time#
geosp	http://www.opengis.net/ont/geosparql#
sf	http://www.opengis.net/ont/sf#
schema	http://schema.org/
s4watr	https://saref.etsi.org/saref4watr/
om	http://www.ontology-of-units-of-measure.org/resource/om-2/
wsis	https://w3id.org/def/wsis/
AquaSPICE	https://w3id.org/def/wsis/examples/AquaSPICE/agricola/

Table 17 – AquaSPICE ontological prefix

For the description of the ontology figures and shapes, we used arrows to represent the relationships between classes (object properties). These relationships represent RDF, RDFS, OWL constructs and object properties. Precisely:

- *Plain arrows* with a white triangle represent the rdfs:subClassOf relation between two classes. The origin of the arrow is the class to be declared as a subclass at the destination of the arrow.

- *Dashed arrows* between classes indicate a local restriction in the origin class. More frequent restrictions refer to object properties that can be instantiated between the classes in the origin and the destination of the arrow. The identifier of the object property is indicated within the arrow.
- *Dashed arrows* with no identifier are used to represent the rdf:type relation, indicating that the element in the origin of the arrow is an instance of the class in the destination arrow.

In the diagrams, datatype properties are denoted by attributes (rectangles) adhered to the classes in a UML oriented notation. Dashed boxes represent local restrictions in a class (e.g., datatype properties that can be applied to the class it adheres to).

Individuals (or ontology instances) are denoted by rectangles in which the identifier is underlined.

5.3.1. Digital Measurements & Variables from IIoT

As presented in Figure 6, industrial representation of the IIoT systems and their connection with observations and measurements is performed using the ontological patterns represented in SAREF and SAREF4WATR ontology. To avoid duplication in the documentation of the standards SAREF4WATR, specific details can be shown inside the documentation. In an overview, SAREF and SAREF4WATR follow up the Observation and Measurement pattern for the representation of temporal measurements.

Under this pattern, specific measurements (*saref:Measurements*) are interlinked with specific Devices (*saref:Device*) to catalogue the specific digital system that performs the observation procedure. Both devices and measurements refer to a specific property (*saref:Property*). This property refers to the specific types of measurement (e.g., water flow, temperature, energy flows, etc). These properties and subsequent measurements are observed in real elements or feature of interests (*saref:FeatureOfInterest*). These real objects normally are related to a geospatial representation (*geosp:Feature*) that covers points or specific areas (e.g., polygons, multi-polygons, etc).

The measurements are specified by a temporal specification (*time:TemporalEntity*) and a unit of measure (*saref:UnitOfMeasure*). There exists a special measurement related to the feature of interest to measure specific general indicators over a region.

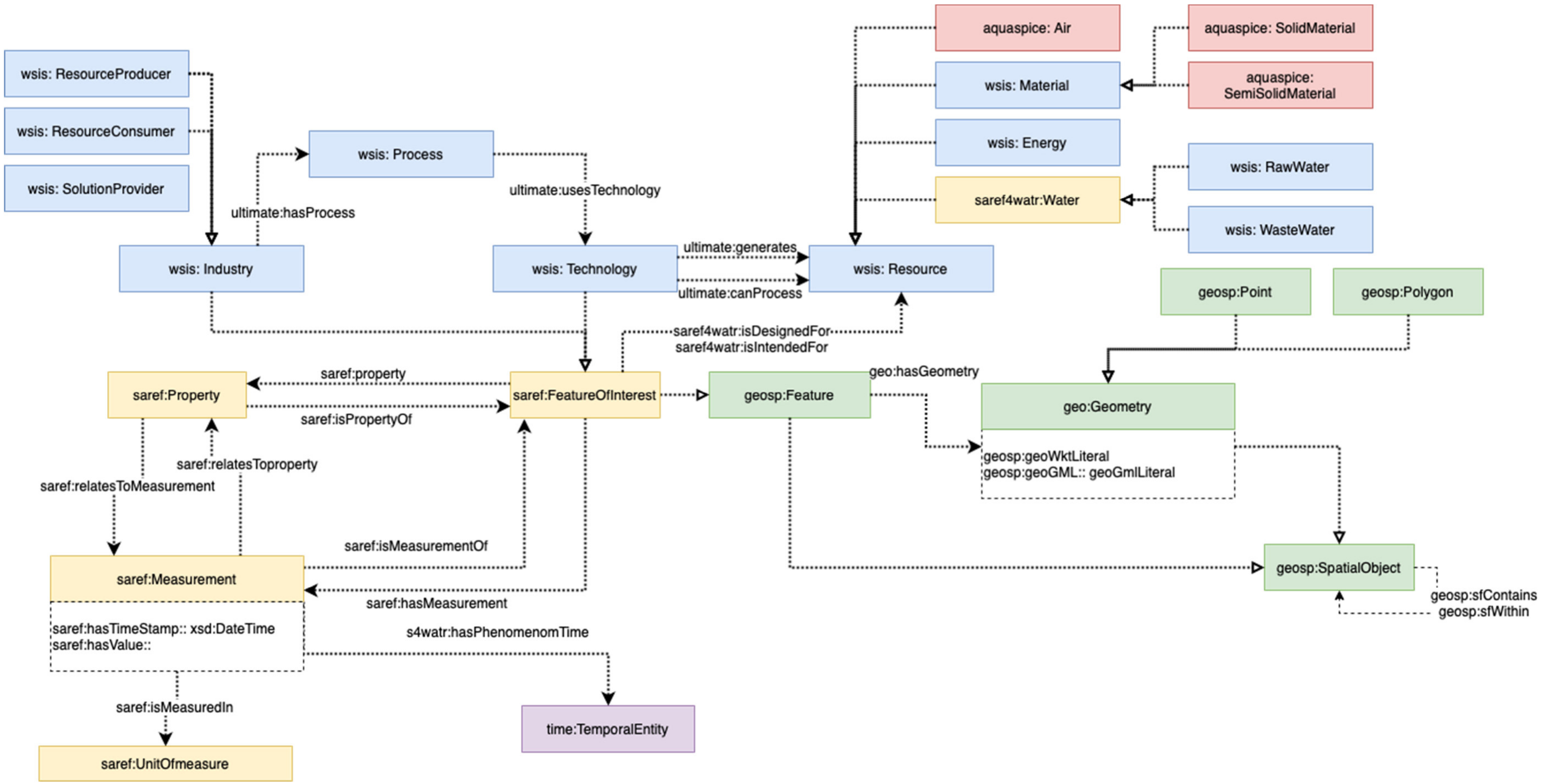

Figure 5. AquaSPICE Core Model

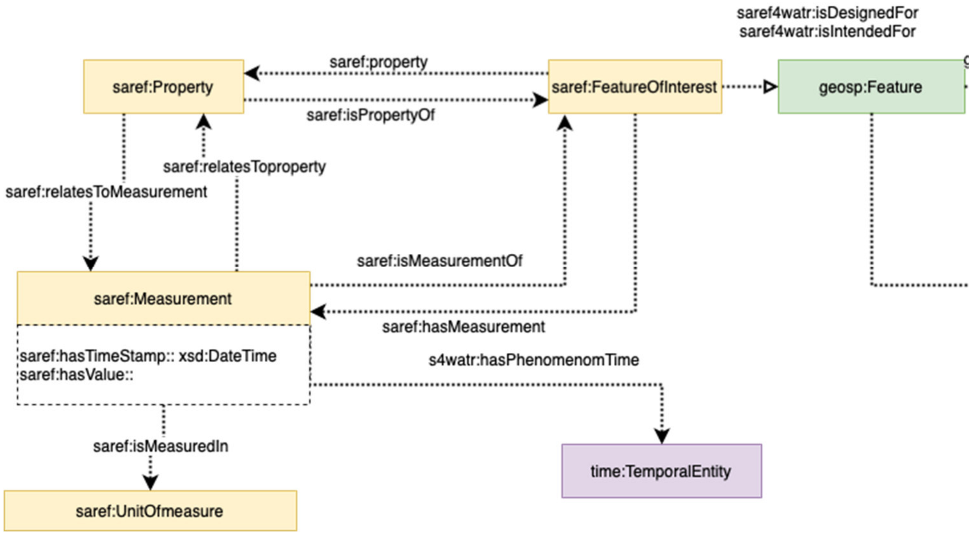

Figure 6. Observation & Measurement pattern

5.3.2. Resource Management

The resource management representation (Figure 7) mainly relies on the interrelation between specific technology (*wsis:Technology*) and the generation (*wsis:generates*) and processing (*wsis:canProcess*) of specific resources (*wsis:Resource*). The resource are categorized into water (*s4watr:Water*), energy (*wsis:Energy*), material (*wsis:Material*) and air (*AquaSPICE:Air*). Considering materials, they are categorized into solid materials (*AquaSPICE:SolidMaterial*) and also, semi-solid materials (*AquaSPICE:SemiSolidMaterial*). An important aspect is the alignment of the measurements and physical aspects with the corresponding resources to associate the different flows, volumes with each associated resource.

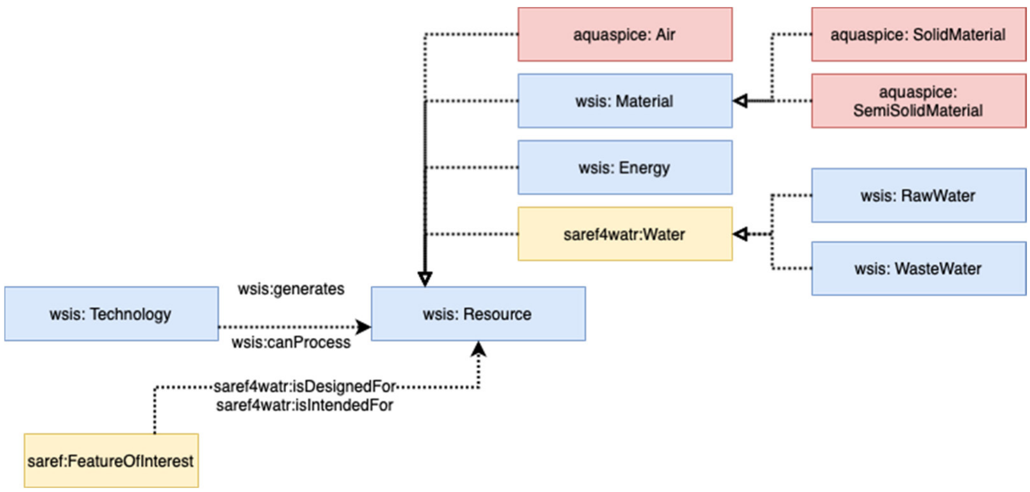

Figure 7. Resource Management Representation

5.3.3. Industrial Symbiosys

This part of the ontology (Figure 8) is mainly devoted to represent industries and the process and technology involved. Thus, the industry (*wsis:Industry*) can be resource producer (*wsis:ResourceProducer*), resource consumer (*wsis:ResourceConsumer*) and also, solution provider (*wsis:SolutionProvider*). This types of industries contains (wsis:hasProcess) several processes (*wsis:Process*) to the generation of products. This processes uses (*wsis:usesTechnology*) one or several technologies (*wsis:Technologies*) to process the resources and also generate new resources. Finally, the technology is also located in specific region and location and also, perform specific generation and processing of resources (measurement). For that, technologies are also a feature of interest inside the ontology (*saref:FeatureOfInterest*).

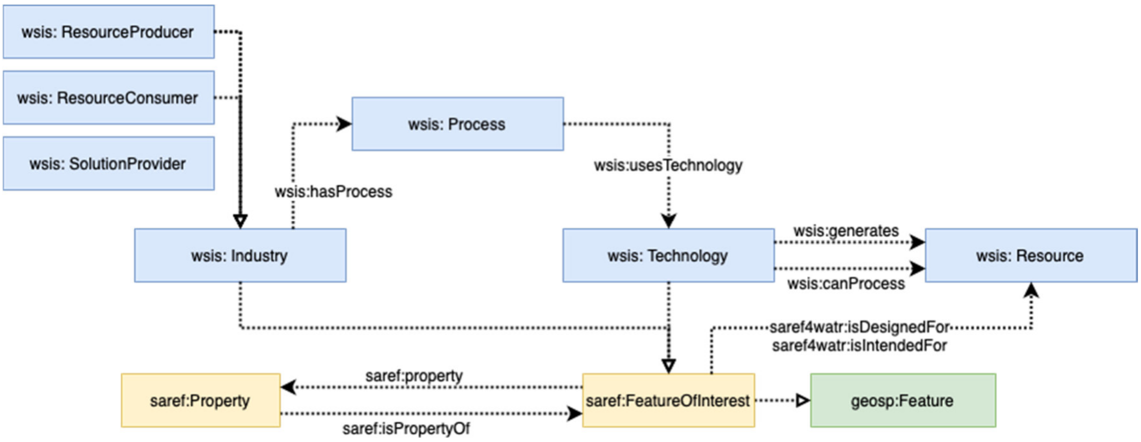

Figure 8. Industrial symbiosis representation

5.3.4. NGSI-LD Alignment

This ontology requires to be aligned with the structural data model and standard of NGSI-LD (Figure 9) in order to link information with FIWARE compatible modules and information exchange. Considering this meta-model, the AquaSPICE ontology connects device representation (*saref:Device*) with the representation of IoT systems inside NGSI-LD (*ngsild:Entity*). This entity contains properties (*ngsild:Property*) that have similar representation as properties defined in SAREF (*saref:Property*). Finally, the different properties in NGSILD are interrelated to include specific measurements and other information. Thus, it is aligned through relationship resources (*ngsild:Relationship*).

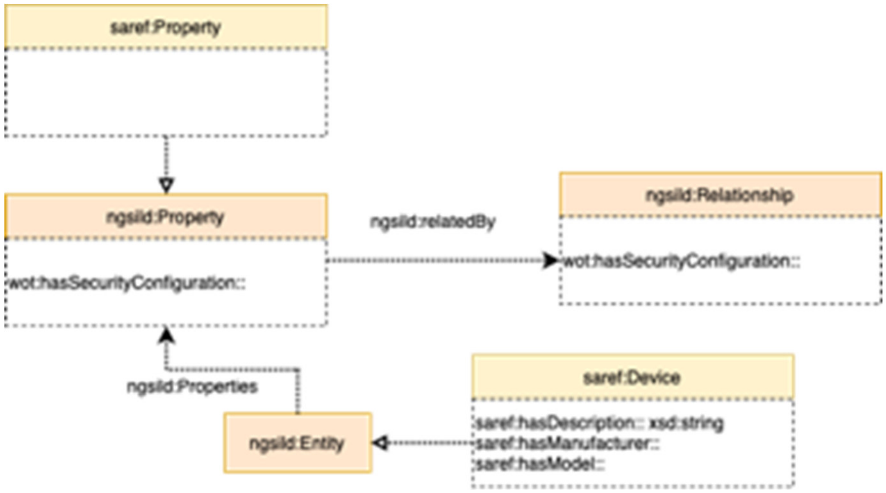

Figure 9. Alignment with NGSI-LD meta-model

6. AquaSPICE Ontology Validation

This section is mainly devoted to the testing and validation of the ontology. For that, this part of the document initially validates the core ontology to ensure correctness in the ontological schema and representation (Section 6.1). With this validation, we followed up with the testing of the elaborated ontology according to the use case of Agricola (Section 6.2). Thus, we performed unit test over the ontology to ensure the alignment with the Competency Questions (Section 4).

6.1. Validation of the AquaSPICE Ontology

The evaluation of the ontology is a process to ensure good practices in the representation of ontology resources. For that, we initially used OOPS! Tool for ensuring the correctness of the following aspects among other relevant ones:

- The domain or range of a relationship is defined as the intersection of two or more classes. This warning could avoid reasoning problems in case those classes could not share instances.
- No naming convention is used in the identifiers of the ontology elements. In this case the maintainability, the accessibility and the clarity of the ontology could be improved.
- A cycle between two classes in the hierarchy is included in the ontology. Detecting this situation could avoid modelling and reasoning problems.

Thus, the initial step has been to ensure the correctness of the core ontological models using OOPS! Tool (Figure 10). After executing the tool, we received two recommendations inside the evaluation report. Summarising the evaluation results, we received the following comments:

Evaluation Result	Problem & Solution
CRITICAL	PROBLEM- ONTOLOGY NOT AVAILABLE ON THE WEB. his pitfall occurs when the ontology code (OWL encoding) or its documentation (HTML document) is missing when looking up its URI. This pitfall deals with the first point from the Linked Data star system that states "On the web" ([10] and [11]). Guidelines in [12] also recommends to "Publish your vocabulary on the Web at a stable URI". This pitfall is also related to the problems listed in [8] and [5]. SOLUTION. WE published the ontology (Section 7) at following URL: https://aolite.github.io/WSISOntology/
CRITICAL	PROBLEM-NAMESPACE HIJACKING. It refers to reusing or referring to terms from another namespace that are not defined in such namespace. This is an undesirable situation as no information can be

Evaluation Result	Problem & Solution
	retrieved when looking up those undefined terms. This pitfall is related to the Linked Data publishing guidelines provided in [11]: "Only define new terms in a namespace that you control" and to the guidelines provided in [5]. This pitfall appears in the following elements: › http://purl.org/vocab/vann/preferredNamespacePrefix › http://purl.org/vocab/vann/preferredNamespaceUri › http://www.w3.org/1999/02/22-rdf-syntax-ns#comment For detecting this pitfall we rely on TripleChecker. See more results at TripleChecker website. Up to now this pitfall is only available for the "Scanner by URI" option. SOLUTION. We put it same URI inside the ontology for their consistency.

Table 18 – OOPS Report and solutions provided by AquaSPICE

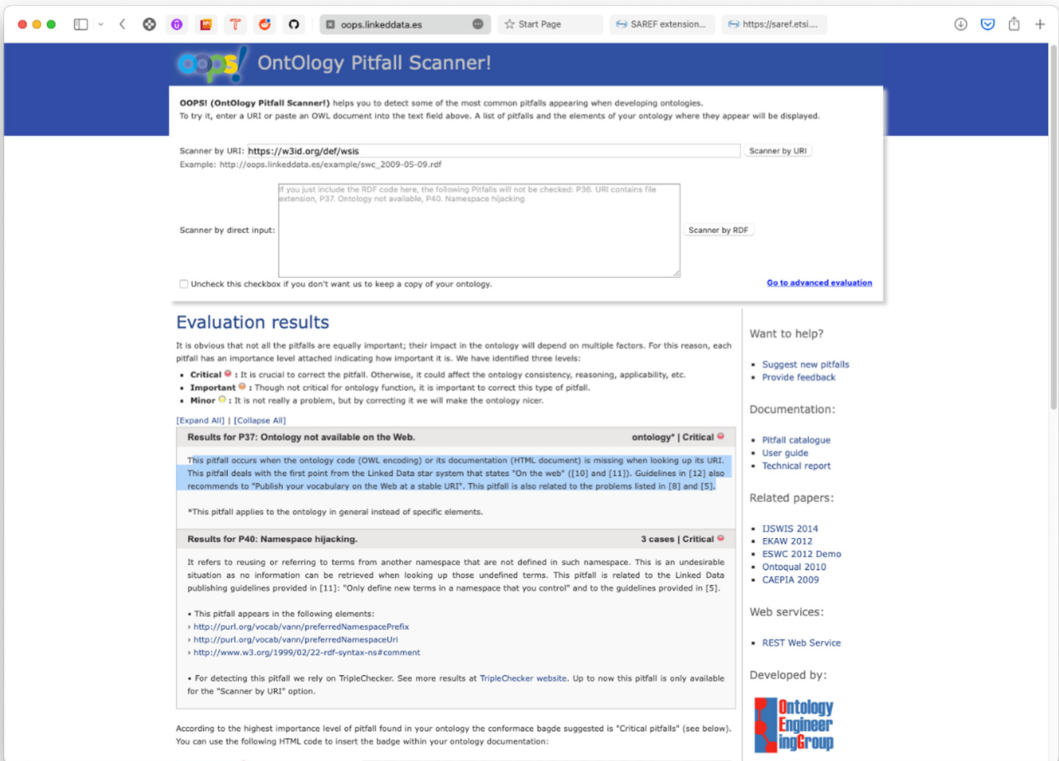

Figure 10. Ontology Evaluation using OOPS!

Once performed the solutions to the problems found by OOPS!, the evaluation report gives some minor aspects that can be solved during next versions of the ontology in relation to the definition of ranges and annotations (Figure 11).

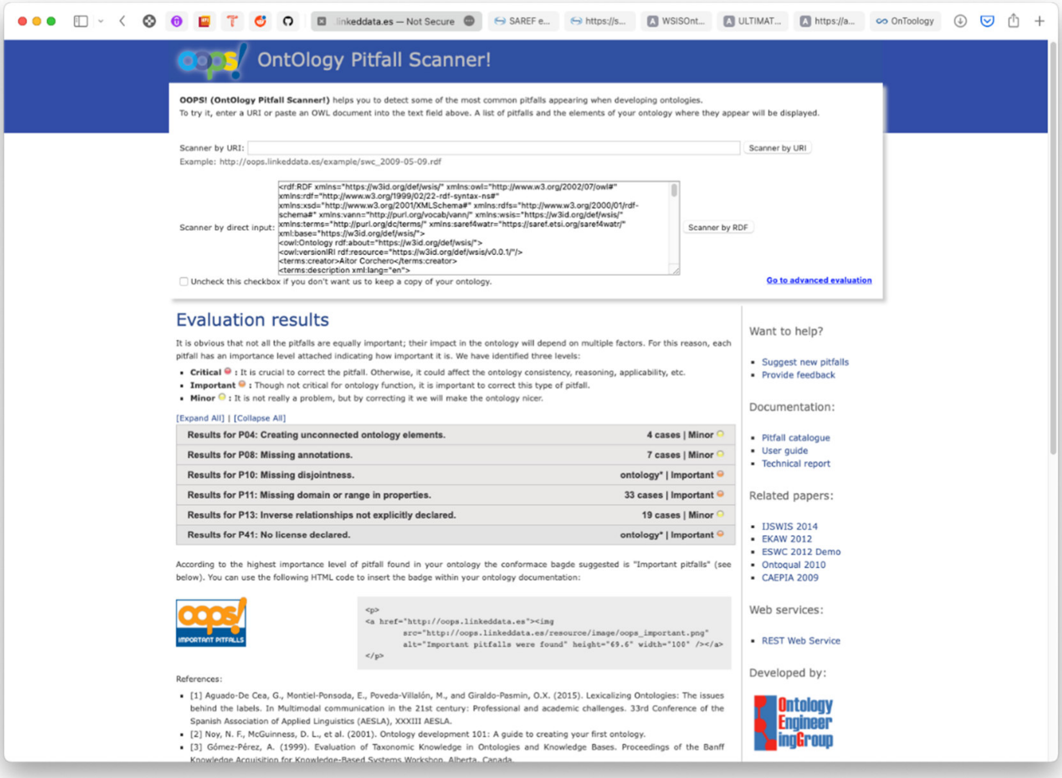

Figure 11. Second validation using OOPS!

After the validation of the ontology structure, the second validation performed relates to the FAIR principles in relation to ensure the reusability principles of the ontology. For that purpose, we used [FOOPS!](#) Tool. This tools permit to provide the means for researchers to assess whether a vocabulary (OWL or SKOS) conforms or not to the best practices for publishing ontologies on the Web.

The execution of the FOOPS! Tools revealed (Figure 12) that the ontology is almost aligned with the principles in a (72%). Indeed, the ontology needs to revise during the next version, the following aspects:

1. Check the IRI version in some of the description.
2. Make public the prefix of the ontology.
3. Check some ontology metadata to ensure the annotations and comments for better understanding of the terms and relations.

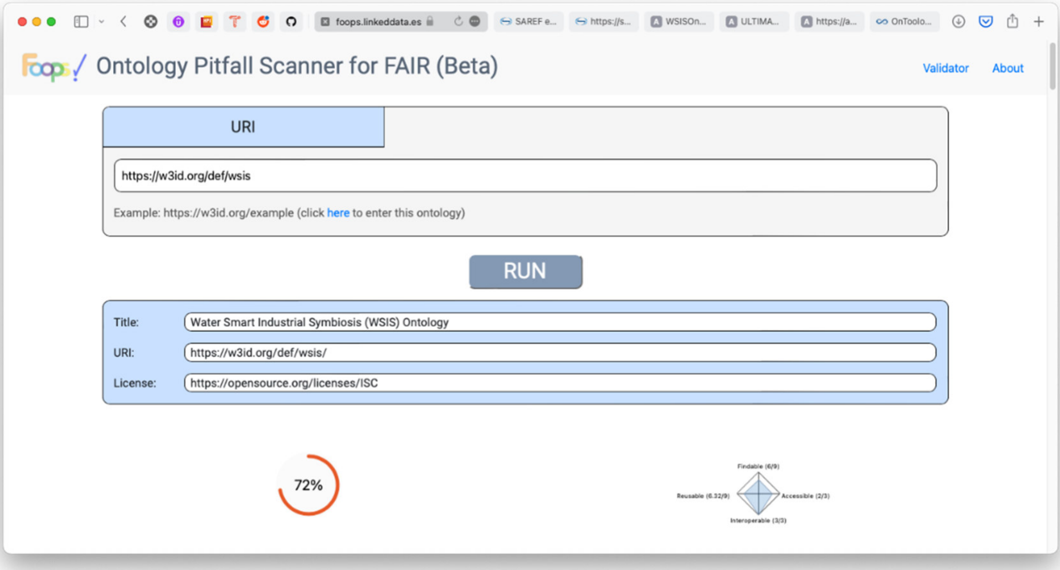

Figure 12. AquaSPICE validation with FOOPS!

6.2. Testing of the AquaSPICE ontology

For performing a unit test of the ontology, we initially performed an ontology specification and instantiation of one of the AquaSPICE case-studies. In this regard, we used the AGRICOLA case-study for that purpose. The Agricola case-study (Figure 13) is located at Agricola International Slaughterhouse (Agricola slaughterhouse) in Bacău. The main aim of the case-study to use a smart solution for sustainable water reuse, with the objective of improving water efficiency monitoring and process optimisation. At the moment, an upgrade of the production plant is taking place so that the maximum production capacity of the processing plant to 550,000 chickens/week (i.e. 850 tonnes per week) to be achieved. This upgrade will consequently raise water needs of the plant and the treatment needs that will be developed and implemented through AquaSPICE project. The indication is that the 15,000 chickens/day more that the plant will be able to process will require 200 m3 more fresh water. In general, this case study is not affected by seasonality; there is a certainty that chicken production is higher in the summer following the market needs, but it does not seem to be monitored and just follows the market needs.

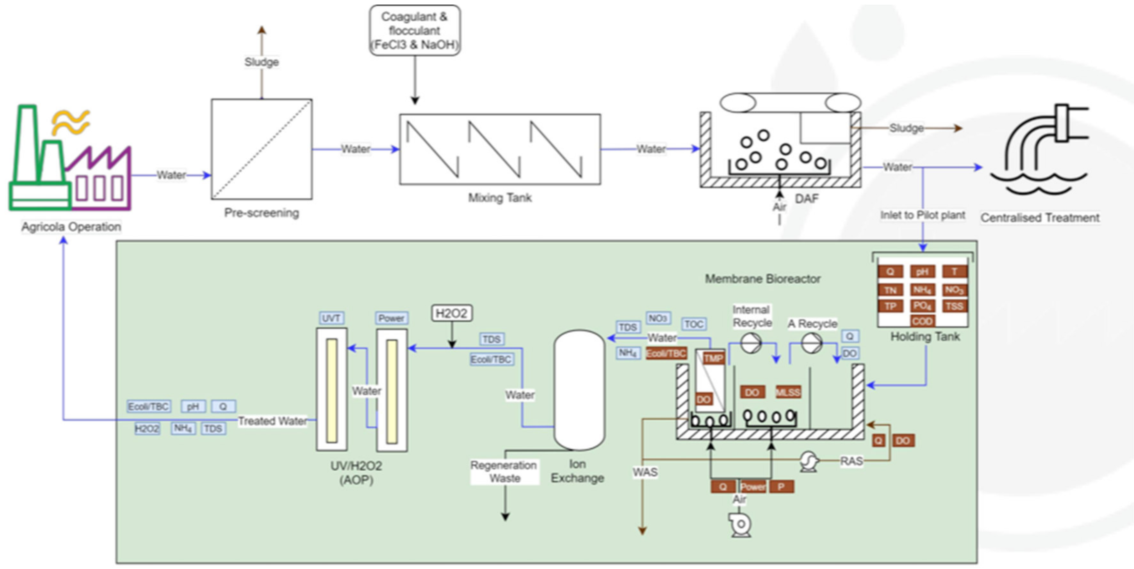

Figure 13. Sensor map of the Agricola case-study

Considering this process and the subsequent improvement to be performed within AquaSPICE, the work has been focused on translating this diagram into an ontology instance based on the core model described (Figure 14). Considering this meta-model, we elaborated specific data sets of the information to be introduced publicly in Triply. Triply is a web-based semantic store that permit and facilitate static data sharing. This triple store enables a SPARQL endpoint over your information to query the data. Thus, it will permit us to test the ontology considering a triple store and SPARQL end point.

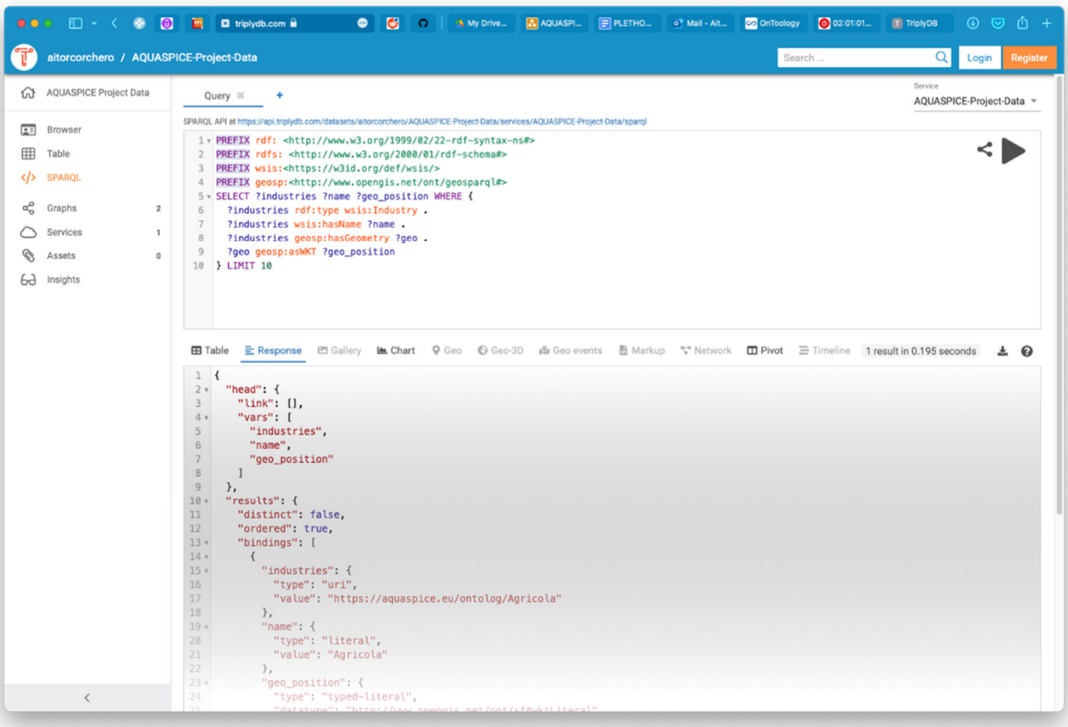

Figure 14. Example of the SPARQL endpoint for AquaSPICE dataset

Once the data is publicly available, the next step has been to elaborate a program in Node.js to perform the unit test. For that purpose, the program mainly uses the following libraries:

- Jest. For testing and elaborating unit-test over the remote triple-store repository.
- SPARQL ENDPOINT. SPARQL library to query and endpoint at manipulate the corresponding information.

Based in these libraries, we performed the implementation of the 13 Competency Questions described in the Table 8 and Table 9. This implementation mainly focuses on the elaboration of the corresponding SPARQL query and the execution of the SPARL wrapper for getting the result. Once the result obtained, it has been checked to ensure their correctness:

```
const sparqlQuery =`PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
    PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>
    PREFIX wsis:<https://w3id.org/def/wsis/>
    PREFIX geosp:<http://www.opengis.net/ont/geosparql#>
    SELECT ?industries ?name ?geo_position WHERE {
      ?industries rdf:type wsis:Industry .
      ?industries wsis:hasName ?name .
      ?industries geosp:hasGeometry ?geo .
      ?geo geosp:asWKT ?geo_position
    } LIMIT 10`;

    const bindingsStream = await myFetcher.fetchBindings(
        'https://api.triplydb.com/datasets/aitorcorchero/AQUASPICE-Project-
Data/services/AQUASPICE-Project-Data/sparql',
        sparqlQuery
    );

    sparqlPromise = await new Promise ((resolve, reject) => {
        let chunks = [];
        bindingsStream.on('data', (triple) => chunks.push(triple));
        bindingsStream.on("end", () => resolve(chunks));
        bindingsStream.on("error", error => reject(error));
      });

    return expect (sparqlPromise[0].name.value).toBe('Agricola');
```

Therefore, after the execution of the testing, we obtained the following results (Figure 15). As can be appreciated, 3 of the tests failed. The main reason of the failure is because the triple store does not permit the reasoning, thus inverse relations are not resolved. This will be solved once we update the information into a triple store with reasoning capabilities higher than RDF.

```
✓ ACQ-1. Which assets are involved in the water/energy/industrial infrastructure? (507 ms)
✓ ACQ-2. What types of sensors are involved in the infrastructures? (244 ms)
✓ ACQ-3. What type of actuators are involved in infrastructures? (197 ms)
× ACQ-4. What type of processes are involved in the demo-case X? (235 ms)
✓ ACQ-5. What type of technology have been implemented in the industry X? (196 ms)
✓ ACQ-6. What is the volume of water exiting process X? (213 ms)
× ACQ-7. What is the energy flux in process X between 01-01-2022 and 01-03-2022? (205 ms)
✓ ACQ-8. What are the physical parameters of measurement X? (218 ms)
✓ ACQ-9. What is the region of the industry X? (204 ms)
× ACQ-10. What are the industries related to the industrial region X? (220 ms)
✓ ACQ-11. What types of sensors are involved in the infrastructures? (201 ms)
✓ ACQ-12. What kind of resources are needed in process X? (227 ms)
✓ ACQ-13. How much resource are required in process X? (212 ms)
✓ ACQ-14. What kind of resources are generated by process X?
✓ ACQ-15. How much resource are generated by process X?
```

Figure 15. Results of the unit-tests

7. AquaSPICE Ontology Documentation and Publication

The AquaSPICE ontology has been published inside the WSIS ontology publication to facilitate its reuse by the community. The corresponding documentation and version code of the ontology can be found the following link:

https://aolite.github.io/WSISOntology/

The documentation of the ontology has been performed by using an external tool called WIDOCO. This tool permits to automatically generate the documentation of the ontology-based on the labels defined in the semantic model. Despite directly using this tool, we preferred to use a continuous integration methodology for ontology documentation and publication. Hence, this procedure will involve once the ontology model is published on GitHub, tests are passed and the ontology is published using WIDOCO tool. For that purpose, we used Ontoology web tool.

Considering this aspect, the main annotations used in the ontology development for documentation purposes are:

Annotation	Description
Dc:contributor	Definition of the contributors of the ontology
Dc: creator	Definition of the creators of the ontology
Dc: description	Description of the ontology (indicating main purpose)
Dc:source	Tags of the ontology
Dc:title	Title of the ontology
Dc:created	Date of creation
Dcterms:license	Licence determined for the ontology
Dcterms:modified	Date of the last modification
versionInfo	Version of the ontology
Comment	Description of the main entities and classes
label	Human readable label for the classes and the properties (optional)

Table 19 – AquaSPICE ontology prefix

After the end of each of the continuous integration processes, we have the ontology published in a custom landing page that could be increased with other ontologies if we work on the same repository. To ensure accessibility, we used a public link and URL under the W3C domain (Figure 16).

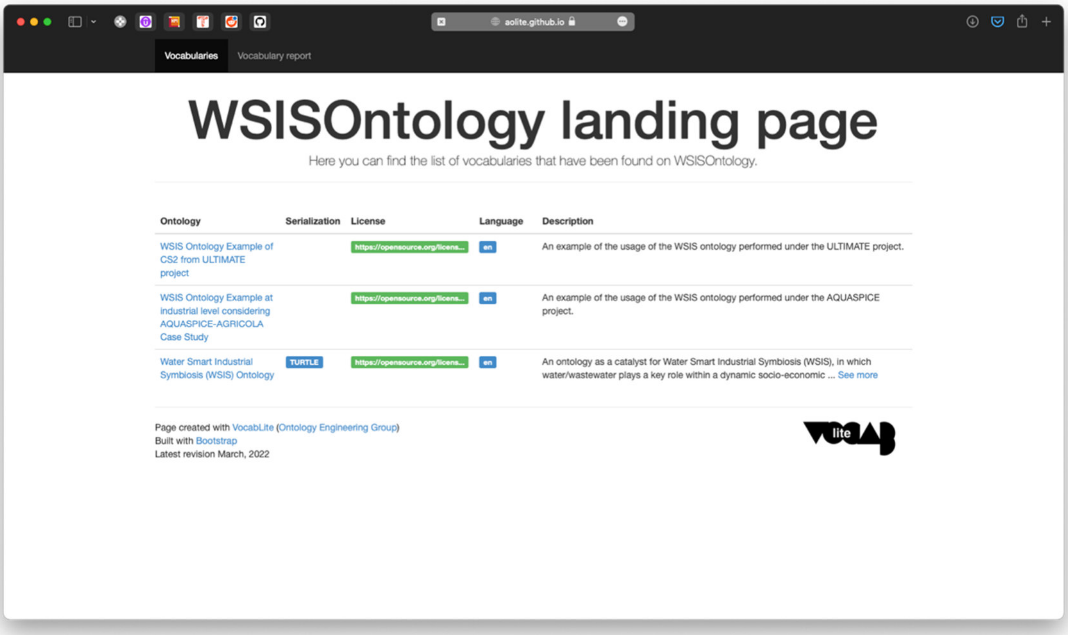

Figure 16. AquaSPICE ontology publication

8. Semantic Repository

In order to allow other systems to interact with the CS ontology population, it is necessary to offer it as an actionable service. To this extent, Apache Jena Fuseki has been deployed. This software package, offered under Apache License 2.0, deploys a SPARQL server with a programmatic API to interact with the existing ontologies. This API, conformant with the SPARQL 1.1 Graph Store HTTP Protocol, allows users to send GRAPHQL Queryes and obtain the results using expressed in RDF.

Using a Dockerized distribution of the software, the SPARQL server has been deployed and necessary configuration files have been edited.

9. Data exploration tool

The already described semantic repository offers the capability to other systems to interact with the ontology that describes the different case studies. However, the SPARQL 1.1 Graph Store HTTP Protocol used by the repository is not suited for direct human interaction. To enhance the human exploration of the CS's ontologies and allow users to effortlessly navigate the knowledge contained in them, the Rhizomer application has been integrated. Rhizomer is a web application that facilitates publishing and exploring Semantic Web data (and particularly Linked Data). Published under GPL-3.0 license the service allows users to visually interact with ontologies. Rhizomer is divided in two components:

- RhizomerEye: a web front-end developed in Angular that allows users to navigate ontologies by means of a web browser.
- RhizomerAPI: a Java Spring backend that transforms the requests received from the RhizomerEye to SPARQL queries and sends them to the Jena Fuseki server.

10. Case study population

Continuing the ontological modelling work described in the first iteration of this deliverable, efforts have been put on modelling the different case studies using the ontology described. As a result, an OWL files for the different project case studies have been created and deployed in the Jena Fiseki server. Moreover, Rhizomer has been configured to allow user to interact with these ontologies. With that, a user can now navigate to Rhizomer Eye and list all the available ontologies (Figure 17: Sample of the CS ontologies linked to Rhizomer Eye). Moreover, the user can see a graph visualization of the key concepts of these ontologies (Figure 18: Graph visualization of the ontology for CS1.A) or browse them in detail (Figure 19: Query result visualization).

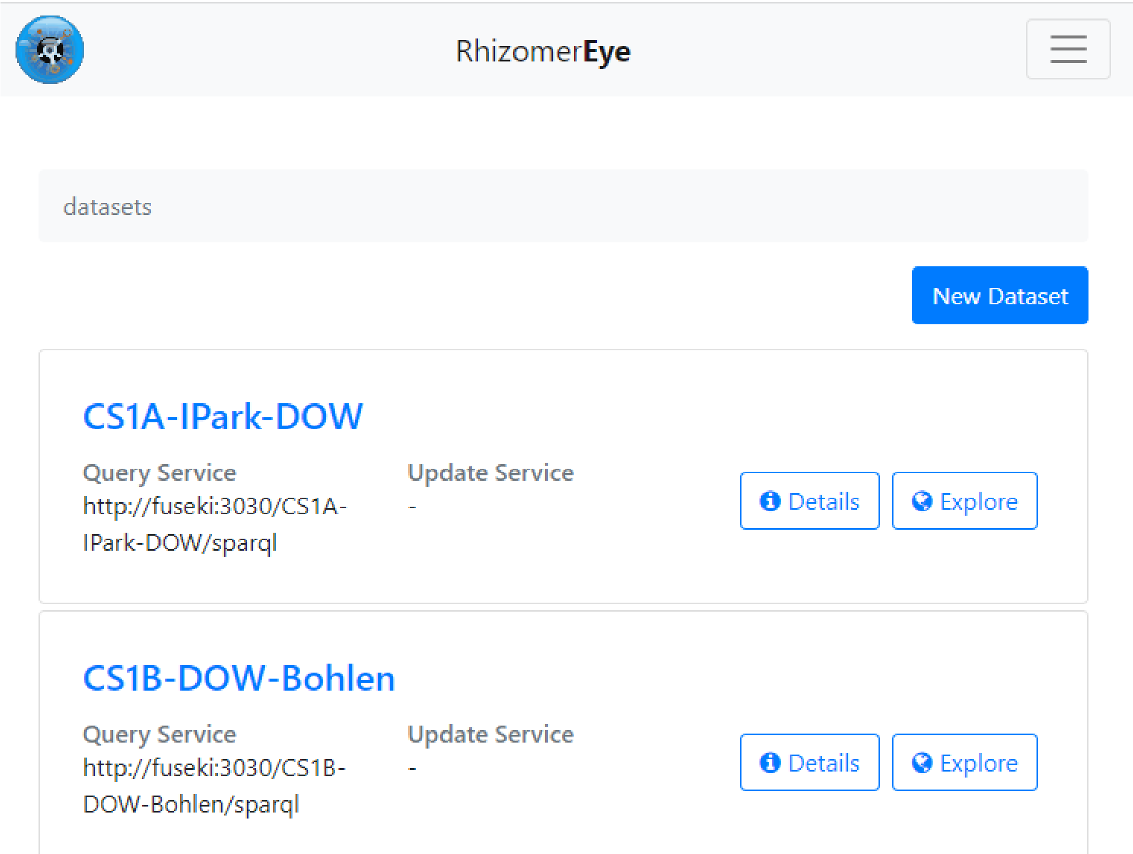

Figure 17: Sample of the CS ontologies linked to Rhizomer Eye

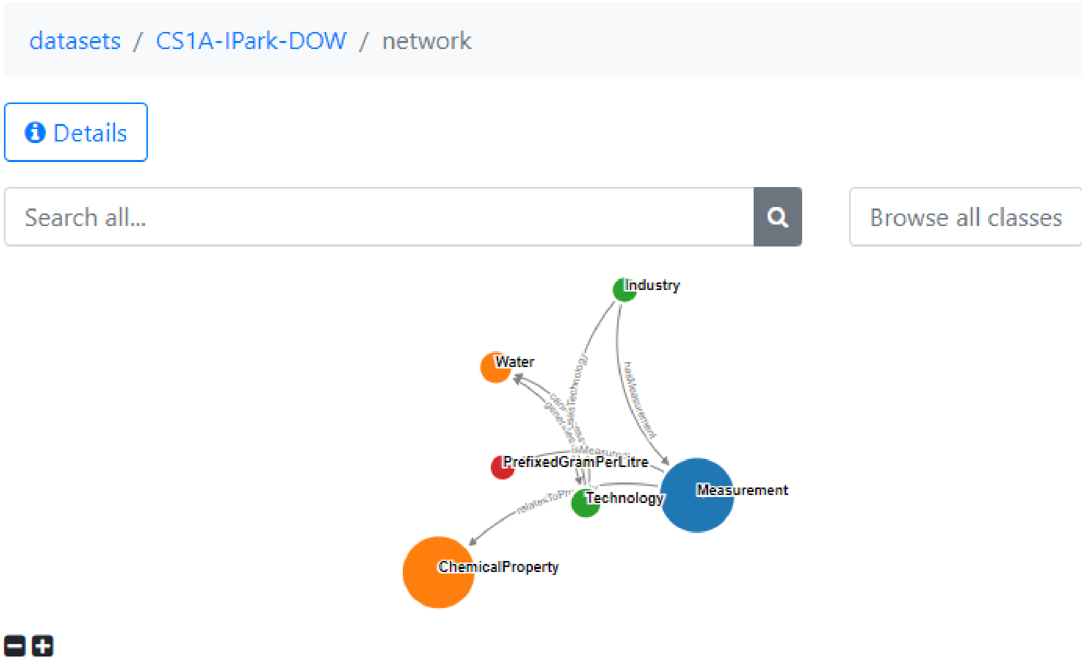

Figure 18: Graph visualization of the ontology for CS1.A

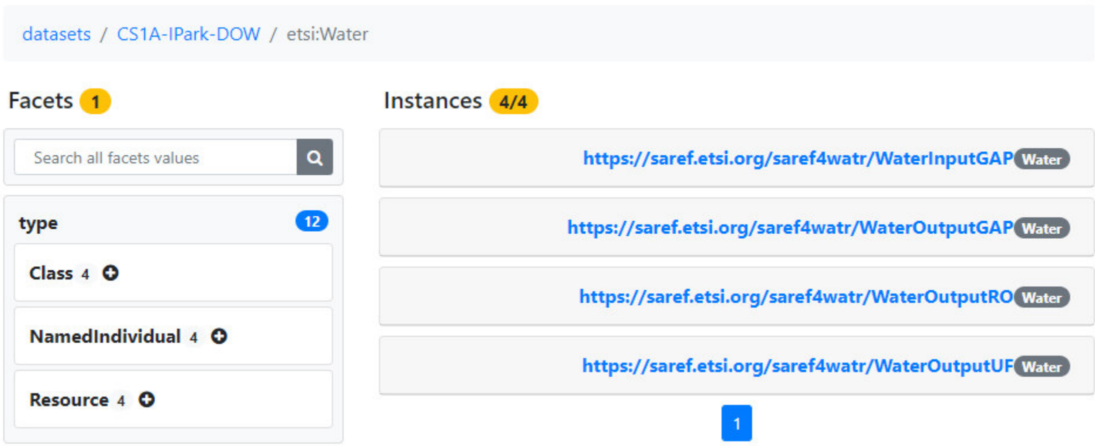

Figure 19: Query result visualization

Using the expressiveness of the AquaSPICE ontology, we are also able to link the instantiation of the ontology with the RTM platform containing the actual time series data stored for that case study. As an example, for CS3.A Antwerp, the different instances of “MeasurementStation” link to the RTM query API (Figure 20: Link from a measurement station instance of CS3.A to the RTM platform API). Although not meant for human usage, this link can serve other software elements navigate the knowledge generated in AquaSPICE.

Figure 20: Link from a measurement station instance of CS3.A to the RTM platform API

11. Conclusions & Future Work

This last section of the deliverable is mainly devoted to the summary of the conclusions and results (Section 11.1) obtained after the development of the Task 1.4 entitled as "*Data Models, Taxonomy and Ontology for Water Use in the Process Industry*". This task was finalized with the present deliverable. However, the ontology will evolve until the end of the project. Hence, subsequent changes will be reported accordingly within the deliverables of WP3.

11.1. Conclusions and Results

The present deliverable has depicted the advancements performed in AquaSPICE in the elaboration of the semantic model (ontology) to: (i) catalogue evidence-based practices and water re-use and recycling strategies to be sharable across platforms and industries; (ii) support system modelling and data analytics to enhance the detection of abnormal behaviors and optimize industrial processes; (iii) understand potential symbiotic links between industries.

For that, the AquaSPICE ontology has been elaborated considering case-study requirements and also, following main EU initiatives, standards and directives for the sector. Moreover, the development of the ontology has been also created considering previous experiences and existing advancements on the semantic modelling for industrial symbiosis, circular economy and also, materials characterization around EU projects. To ensure that the ontology meets use-cases requirements and best practices, we performed a procedure for guaranteeing their accomplishment thanks to pitfalls and also unit testing.

Furthermore, a semantic repository (Apache Jena Fuseki) has been deployed to store the instantiation of the AquaSPICE ontology for the different case studies. Additionally, a data exploration tool (Rhizomer) has also been installed to allow users to visually explore these ontologies.

As a conclusion, task 1.4 establish the bases for a common data model and knowledge exchange in AquaSPICE. It will serve to generate and semantically link data models inside the AquaSPICE architecture.

11.2. Future Work

As mentioned, Task 1.4 has finalized with the present deliverable. Despite this, we envision as a future task the following actions:

Action	Description
Case-Studies population	Continue with the population of the case-studies inside the demonstration strategy considered in AquaSPICE to reflect the evolution in the case-studies

Table 20 – AquaSPICE future work to be accomplished

12. Bibliography

[1] A. Corchero, L. Echeverria and E. Westerhof, “D4.4: SEMANTIC REPOSITORY,” 2020.

[2] WatERP, “D1.3. Generic Ontology for water supply distribution chain,” Barcelona, 2013.

[3] N. Trokanas, T. Raafat, F. Cecelja and A. Kokossis, “OFIS – Ontological Framework for Industrial Symbiosis,” *Computer Aided Chemical Engineering,* vol. 32, pp. 523-528, 2013.

[4] F. Cecelja, N. Raafat, N. Trokanas, S. Innes, M. Smith, A. Yang, Y. Zorgios, A. Korkofygas and A. kokossis, “e-Symbiosis: technology-enabled support for Industrial Symbiosis targeting Small and Medium Enterprises and innovation,” *Journal of Cleaner Production,* vol. 98, pp. 336-352, 2015.

[5] M. T. Horsch, D. Toti, S. Chicchiera, M. A. Seaton and I. T. Todorov, “OSMO: Ontology for Simulation, Modelling, and Optimization,” 2021.

[6] M. Fernandez-Lopez, A. Gomez-Perez, A. Pazos and J. Pazos, “Building a chemical ontology using methontology and tthe ontology design environment,” *IEEE Intelligent Systems & their applications,* vol. 4, no. 1, pp. 37-46, 1999.

[7] S. Staab, H. Schnurr, R. Studer and Y. Sure, “Knowledge Processes and Ontologies,” *IEEE Intelligent Systems,* vol. 16, no. 1, pp. 26-34, 2001.

[8] M. Suarez-Figueroa, “NeOn Methodology for Building Ontology networks specification, scheduling and Reuse,” UPM, 2011.

[9] S. Proni, “SAMOD: An agile methodology for the development of ontologies,” 2016. [Online]. Available: https://essepuntato.it/samod/.

[10] H. S. Pinto, C. Tempich and S. Staab, “DILIGENT: Towards a fine-grained methodology for Distributed, Loosely Controlled and evolving Engineering of <Ontologies,” in *16th European Conference on Artificial Intelligence (ECAI 2004)*, 2004.

[11] FIWARE FOUNDATION, “FIWARE,” [Online]. Available: https://www.fiware.org.

[12] W3C, “SPIN-Overview and Motivation,” 2011.

[13] W3C, “A Semantic Web Rule Language Combining OWL and RuleML,” 2004.