WatERP

Water Enhanced Resource Planning

“Where water supply meets demand”

GA number: 318603

WP 1: Water Supply Knowledge Base

D1.5: Report of final taxonomy, functional model and ontology based on the outcome of their application during the project

V1.0 20/09/2015

www.waterp-fp7.eu

Document Information

Project Number	318603	Acronym	WatERP
Full title	Water Enhanced Resource Planning "Where water supply meets demand"		
Project URL	**http://www.waterp-fp7.eu**		
Project officer	Grazyna Wojcieszko		

Deliverable	Number	1.4.5	Title	Extension of taxonomy and ontology to the pilots
Work Package	Number	1	Title	Water Supply Knowledge Base

Date of delivery	Contractual	M36	Actual	M36
Nature	Prototype ☐ Report ☑ Dissemination ☐ Other ☐			
Dissemination Level	Public ☑ Consortium ☐			

Responsible Author	Gabriel Anzaldi Varas	Email	gabriel.anzaldi@eurecat.org
Partner	BDIGITAL (EURECAT)	Phone	+34 93 553 45 40

Abstract (for dissemination)	Description of the final version of the generic ontology that was developed in WatERP project. The novelty of this ontology lies on including human interactions with the natural paths as a mechanism to understand the affectation of different decision-making procedures into the water resources management. The main objective is to match water supply with demand. Hence, these interactions cover multi-disciplinary decisions ranged from water quantity (flows), infrastructural and financial/economic decisions. Thus, the ontology is able to represent most of the decisions taken by the water stakeholders from Water industry (Regulators, Water Service Providers, Bulk Water Suppliers, Water Utilities) to the final end users (consumers). The elaborated knowledge representation is supported by data provenance mechanisms in order to contextualize the information in order to ensure data nature. Moreover, the presented semantic model follows standard hydrologic data representation by the semantic alignment with trustworthy and standard ontologies and vocabularies as the ones published by NASA, CUAHSI, OGC and W3C. .
Key words	Water Knowledge Base, Ontology, Water Ontology, Water Knowledge Representation, Knowledge Base Population, Ontology documentation

Glossary of Terms

(Geo)SPARQL- Geographic SPARQL

A-box- Assertion Box

API- Application Program Interface

CQ- Competency Question

CUAHSI- Consortium of Universities for the Alliance of Hydrologic Science, Inc.

DMS- Demand Management System

DSS- Decision Support System

DX.Y- Deliverable X.Y

HMF- Hydro-Metereological Forecast

HTML- HyperText Markup Language

HTTP- Hypertext Transfer Protocol

IRI- Internationalized Resource Identifier

LODE-Live OWL Documentation Environment

MAS- Multi-Agent System

MMO- Marine Metadata Ontology

NeON- Neon methodology for building Ontology Networks

O&M- Observation and Measurement conceptual model

OGC- Open Geospatial Consortium

OWL-Ontology Web Language

PIM- Pilot Integration Manager

QUDT- Quantity, Units Dimension and Type ontology

RDF- Resource Description Framework

SOA- Service Oriented Architecture

SOA-MAS- Service Oriented Architecture-Multi-Agent System

SOS- Service Observation Sensor

SPARQL- SPARQL Protocol and RDF Query Language

SSN- Semantic Sensor Network

SWEET- Semantic Web for Earth and Environmental Terminology

T-Box- Terminology Box

URI- Unified Resource Identifier

URL- Unified Resource Location

W3C- World Wide Web Consortium

WaterML2.0- Water Markup Language v2.0

WatERP-KB- WatERP- Knowledge Base

WatERP-WDW- Water Data Warehouse

WPx- Work Package X

XML- eXtensible Markup Language

XSLT- eXtensible Stylesheet Language Transformations

Executive Summary

Deliverable 1.5 (D1.5) consists in the report of the final version of the knowledge base elaborated in the framework of the WatERP project. The work presented in this document has been based on explaining all the semantic resources that describes water manager's expertise with the aim of managing water supply and distribution chain. This final version of the knowledge base is the result of the continuous vertical integrations built along the project to align the project modules towards pilot's implementation. Thus, the knowledge base covered in this document is the outcome of continuous modifications performed under deliverables 1.4.x- "*Extension of taxonomy and ontology to the pilot's*" beginning with the generic ontology presented in D1.3- "*Generic ontology for water supply distribution chain*".

The novelty over the WatERP-KB over current developments lies in the representation of the human interactions that are involved in the water supply distribution covering also the interaction of different stakeholders. This aspect permits to evaluate the affection of multi-disciplinary decisions towards enhancing the management of water supply and distribution chain. These interactions have been modelled in the ontology through abstracted entities (semantic classes) focused on grouping the geo-spatial objects (water infrastructure) by behaviour (including decisional-procedures). Thus, the water network (from sources to distribution elements) is modelled as the linkage between mentioned abstracted classes. **The highlighted result is the decoupling of the decisional procedures from geographic location of the elements**. Complementing this abstraction, the information collected from the water infrastructure has been contextualised using **the Observation and Measurement (O&M) conceptual model**. The incorporation of this model into the knowledge base **highlights to ensure data provenance from water systems**.

Focusing on the represented types of decisions, this version of the knowledge base has been expanded towards **interrelating water management informational flows with financial and economic flows**. Then, **water stakeholders are able to understand the enhancement of the water infrastructure through the measurement of key performance indicators that helps to apply several strategies to make more efficient the use of water** (e.g. efficient Irrigation methods, subsides of water technology, establishing efficient water management methods for buildings, etc.). Additionally, the decisions can be applied in certain geo-spatial location through the usage of the geospatial reasoning (using the incorporated Geo-SPARQL ontology).

Transversally to the knowledge base development, the present deliverable also highlights the **elaboration of automatic population method**. This method facilitates the semantic resource maintenance during time. Furthermore, this methodology also permits to maintain updated the information from the water infrastructure, making the semantic resources accessible through a semantic repository (e.g. Sesame).

Therefore, the elaborated knowledge base is capable of: *(i)* abstracting human-decisions from physical context in order to support coordinated decision-making between water industry; *(ii)* transparently updating the knowledge base instances throughout a population process; *(iii)* reasoning over the water resources and geographical information to support the decision-making process in the water supply and distribution chain, *(iv)* representing economic factors aligned with the actors involved in the water supply and distribution chain; *(v)* completely integration with the current standards in water information modelling such as Water Markup Language 2.0 (WaterML2.0) and the Open Geospatial Consortium (OGC)-stack; *(vi)* fully alignment with current standard vocabularies as World Wide Web Consortium (W3C)-Semantic Sensor Network (SSN), Consortium of Universities for the Alliance of Hydrologic Science, Inc. (CUAHSI) and Semantic Web for Earth and Environmental Terminology (SWEET); *(vii)* representing water quality data for further reuse; and *(viii)* providing mechanism to automatically document and understand the semantic model, fostering the ontology sharing and enhancement since the ontology development perspective.

To understand this document the following deliverables have to be read.

Number	Title	Description
D1.1	Generic Taxonomy for water supply distribution	This deliverable summarizes the taxonomy of the domain knowledge and the initial version of the ontology, including the scope, purpose and implementation language to be used.
D1.2	Generic functional model for water supply and demand usage data	Report describing the approach that will be used in the WatERP project to represent the processes required to match supply with demand across the water supply distribution chains. It includes processes and decisions involved in the pilot cases.
D1.3	Generic ontology for water supply distribution chain	Description of the generic ontology that was developed within WatERP project. This deliverable introduce into the incorporation of human-made interactions inside natural water paths in order to better understanding of the decisions to be adopted. Furthermore, data provenance and Linked Open Data Cloud (LOCD) mechanism are also introduced.
D1.4.1	Extension of taxonomy and ontology to the pilot's	This deliverable depicts ontology modifications focused on (i) enhancing observation and measurement procedure, (ii) aligning the WatERP ontology development in parallel with other hydrology standards, and (iii) reinforcing human interaction inside water natural paths in order to improve water management.
D1.4.2	Extension of taxonomy and ontology to the pilot's	This deliverable depicts the enhancements done over WatERP-KB by (i) introducing the "Observation Quality" concept in order to specify the data quality aligned to a certain time series; (ii) implementing alignments according with other hydrology standards (CUASHI, W3C-SSN, NASA-SWEET, HY_FEATURES and WaterML2), and (iii) reinforcing human interaction inside water natural paths (e.g. new feature of interest defined in the ontology that are aligned with the pilots information –e.g. "Gauging station"-;

		and instantiation of the modified logical model for the SWKA case). Furthermore, the present deliverable also describes the population process implementation for a large scenario by applying XSLT transformations.
D1.4.3	Extension of taxonomy and ontology to the pilot's	This deliverable depicts the enhancements done over WatERP-KB by (i) introducing the "Observation Quality" concept in order to specify the data quality aligned to a certain time series; (ii) implementing alignments according with other hydrology standards (CUASHI, W3C-SSN, NASA-SWEET, HY_FEATURES and WaterML2), and (iii) reinforcing human interaction inside water natural paths (e.g. new feature of interest defined in the ontology that are aligned with the pilots information –e.g. "Gauging station"-; and instantiation of the modified logical model for the SWKA case). Furthermore, the present deliverable also describes the population process implementation for a large scenario by applying XSLT transformations.
D1.4.4	Extension of taxonomy and ontology to the pilot's	This deliverable describes the enhancements performed over the ontology derived from the population process and the improvements of the economic and demand models as a result of it representation into the OMP. Moreover, a semantic quality model is also depicted and introduced in this version of the ontology in order to satisfy the latest recommendations provided by the OGC.
D1.4.5	Extension of taxonomy and ontology to the pilot's	This deliverable depicts the improvements performed onto the WatERP general ontology regarding the necessary annotations over the semantic resources in order to document automatically the ontology (HTML). Thus, the generated ontology can be understood and shared across the researches and general public using open standards.
D3.4	WDW Final Prototype	This document provides an overview of the architecture of the water data warehouse and the integration of the pilot sites. It explains the protocols and components that are being used for storing and exchanging both ontological information and observation results between the pilot site, water data warehouse and SOA-MAS. Further, it describes how data mining and geospatial reasoning have been implemented and what considerations have been made about performance, stability and extensibility.
D5.5	Water Demand Management System and relevant documentation	The report presents the terms and general concepts behind water demand management and the WDMS in particular, leading to a system's analysis of management processes along the water supply-distribution chain, by considering knowledge gained from auditing and interacting with WatERP's Pilot sites. This, in turn,

		leads to the specification of user requirements and the design of the WDMS software Toolbox. The models and methods used in WDMS are discussed, with reference to the deliverables D5.1 to D5.4, where a more comprehensive description is provided, and the overall WDMS architecture and functionality is described, followed by a description of the actual software development environment and processes.
D7.1.3	Holistic Auditing	This deliverable consists on the validation of the knowledge base towards adjusting the ontology to the development standards and best practices. Furthermore, the ontology has been compared with the most representative water domain-related ontologies towards semantic expressivity comparison. As a result, some improvements and corrections are suggested to drive future developments and enhancements of the ontology towards the pilots' needs.

Table of contents

List of figures

List of tables

List of Listings

1. Introduction

In the WatERP project, the intelligent architecture supports decision-making process by the integration of a set of WatERP modules (called building blocks) using communication standards. This modular integration is mainly generated by the Multi-Agent System (MAS). The MAS is aimed at performing intelligent orchestration between Decision Support Systems (DSS) and Demand Management System (DMS) and Hydro-Meteorological Forecasts (HMF), using the WatERP-Knowledge Base (WatERP-KB) to standardize the transference of information between modules. Specifically, the WatERP-KB generates multi-level inferences that permit to: *(i)* coordinate the management in each step of the water supply distribution chain; *(ii)* interact among functions involved in each step of the water transport from the hydro-meteorological data to the end-user; *(iii)* interact between currently separated control and optimization systems such as reservoirs or hydroelectric plant, DSSs or water treatment and distribution management tools; and *(iv)* reduce of the energy consumption thanks to improve water distribution.

One remarkable aspect of the WatERP architecture is the confluence between the Service Oriented Architecture Multi-Agent System (SOA-MAS) and the WatERP-KB. The intelligent combination between both modules provides the whole architecture with needed interoperability and orchestration mechanisms to synchronize all defined WatERP modules. On the one hand, the MAS provide the architecture with the orchestration by applying a matchmaking process. The matchmaking process uses yellow and white pages to offer the building blocks needed informational requirements. The MAS is complemented by a Service Oriented Architecture (SOA) aimed at making available the syntactic interoperability between elements of the architecture. Thus, the syntactic interoperability is performed by publishing building blocks functionalities (services) and standard mechanisms to execute their services (see D2.3-"*Open Interface Specification*" on Section 3.2 in page 24). On the other hand, the WatERP knowledge base provides the whole architecture with a common and shared vocabulary. This shared vocabulary permits to move up syntactic to semantic interoperability by facilitating the understanding of the concepts related with the water domain.

The WatERP-KB performs informational sharing between elements by using standard hydrologic-domain concepts. These standard concepts have been acquired from representative ontologies (W3C-SSN, NASA-SWEET and CUAHSI) and vocabularies (OGC-WaterML2.0 and HY_Features). These aligned vocabularies and models have permitted to represent contextual, spatial and temporal water-related information regarding hydrologic domain. Then, the WatERP-KB is capable of ensuring hydrology data provenance by describing observations and measurement processes. This description has been based on the Observation & Measurement (O&M) model defined by the OGC and the semantic model implemented by the World Wide Web Consortium (W3C) to manage all related information regarding a specific sensor. Additionally, the WatERP ontology contains conceptual elements aligned with the human interactions that are involved in the water supply distribution chain in form of different stakeholders. This last definition is one of the main contributions of the WatERP ontology against current state of the art *(Anzaldi et al., 2014; Corcho & García-Castro, 2010; Malewski, Broring, Maué, & Janowicz, 2013)*.

Hereafter, **the WatERP-KB main objective is to analyse the influence of human interaction and decisions that occurs inside the water supply distribution chain towards the generation of a knowledge representation that gives necessary support to decisional systems of the water resource management** *(Anzaldi et al., 2014)*. The proposed semantic model combines the explicit water manager's knowledge (information from pilot's) and the implicit knowledge (hidden knowledge) inferred by the exploitation of the defined semantic resources. Thus, **new water resource management strategies are derived from the WatERP-KB by making explicit (water manager known) the implicit knowledge (non-direct assertions)**. Hence, the ontology plays the role inside the WatERP platform of *(i)* supporting the water management by exploiting the relations between water resources; *(ii)* providing a mechanism to easy access water information; *(iii)* generating inferences about the impact and dependence between water resources availability, and *(iv)* understanding water resource management and water distribution by the analysis of human-made decisions and their interactions with the water supply and distribution chain.

This deliverable will fully describe the final version of the WatERP-KB. This version of the ontology is the result of the application of an ontology development cycle since the beginning of the WatERP project. The applied ontology development cycle corresponds to the Neon methodology for building Ontology Networks (NeON). This methodology standardizes the process of constructing, testing and maintaining ontological resources in large scenarios *(Gómez-Pérez, Fernández-López, & Corcho, 2007; Suárez-Figueroa, 2010)*. It is divided in three different phases ("*Early Analysis*", "*Analysis and design*", and "*Implementation and Test*") to evolve the ontology towards fulfilling the water stakeholder's requirements/needs and managing the water supply and distribution chain (see Figure 1).

Focusing on the stages that composes the NeON methodology, the "*Early analysis*" (first stage) mainly corresponds to deliverable 1.1 "*Generic taxonomy for water supply distribution chain*" (D1.1) and deliverable 1.2 "*Generic Functional Model for Water Supply Demand and usage Data*" (D1.2). In this stage, the main aim is to establish the bases of constructing the WatERP-KB. That means, this stage was focused on defining the knowledge base purpose, the implementation language, the involved users and the ontological objectives represented in form of Competency Questions (CQs). These aspects were taken into account due to correspond WatERP-KB with informational water manager necessities (detailed on D1.1 on Sections 2, 3 and 4 respectively). In addition, this stage covered a systematic review over the known and representative ontologies/vocabularies regarding sensor and hydrology-domain semantic representation. Thus, semantic models such as Word Wide Web Consortium-Semantic Sensor Network Ontology (W3C-SSN), Semantic Web for Earth and Environmental Terminology (SWEET), Consortium of Universities for the Advancement of Hydrologic Science, Inc ontology (CUAHSI and Marine-Metadata Ontology (MMO) and controlled vocabularies (WaterML, WaterML2.0 and HY_Features) were analysed. Main conclusion of this analysis was reflected on D1.2

in Section 4 (page 44) which highlights that "*most of the developed hydrological ontologies are focused on representing the hydrological cycle from the natural cycle by using an environmental perspective*". Mainly, this perspective is useful for the scientist and biologist in order to represent the hydrosphere events. However, analysed approaches are not able to represent management viewpoint. Moreover, the semantic relations between elements involved in the water supply and distribution chain are not completely defined on the studied ontologies and vocabularies. Therefore, the ontologies can only represent semantically explicit knowledge by generating static semantic resources. Therefore, the novelty of the WatERP ontology is focused on **using the current standards and ontological concepts as a base to construct a knowledge base that also can model the management perspective related to the water cycle**. As a result of the "*Early Analysis*" stage, a controlled vocabulary (taxonomy) for representing the water supply and distribution chain elements and human interactions was defined (see D1.1 in Section 6 on page 39). This vocabulary was derived from pilot's information, functional models that was defined over the pilots, and the analysis performed against the current developments in hydrological ontologies (see D1.2 in Section 3 on page 32).

Next stage of NeON methodology corresponds to the "*Analysis and design*" which is mostly covered under Deliverable 1.3-"*Generic Ontology for water supply distribution chain*" (D1.3). This part of the ontological development was focused on transforming the defined initial acquired vocabulary (taxonomy) towards a knowledge structure. Therefore, the WatERP-KB is able to infer relevant knowledge from the extracted pilots' information using the defined assertions and restrictions. Then, the work performed was based on constructing ontology by defining proper ontological resources. This task was covered by the definition of object properties, data properties, axioms, restrictions and mappings with externals resources (SSN, CUAHSI, SWEET, HY_FEATURES and WaterML2.0). Mainly, the ontology construction was based on transforming the Water Markup Language v2.0 (WaterML2.0) schema and inherited procedures into semantic structure. This transformation permitted the rest of the building blocks to exchange and understand the defined information of WaterML2.0 language by sharing and defining a controlled vocabulary. Additionally, **this conversion also permitted the WatERP-KB to enhance data provenance by the implementation of the "*Observation and Measurement*" conceptual model defined by the OGC in order to preserve data nature**. Furthermore, concepts and relations of the ontology were enriched with semantically annotations to facilitate human readability and comprehension of the ontology towards driving the continuous development of the ontology. As a result of this stage, a first version of the WatERP-KB was published in order to start with the population procedure of the pilot's information.

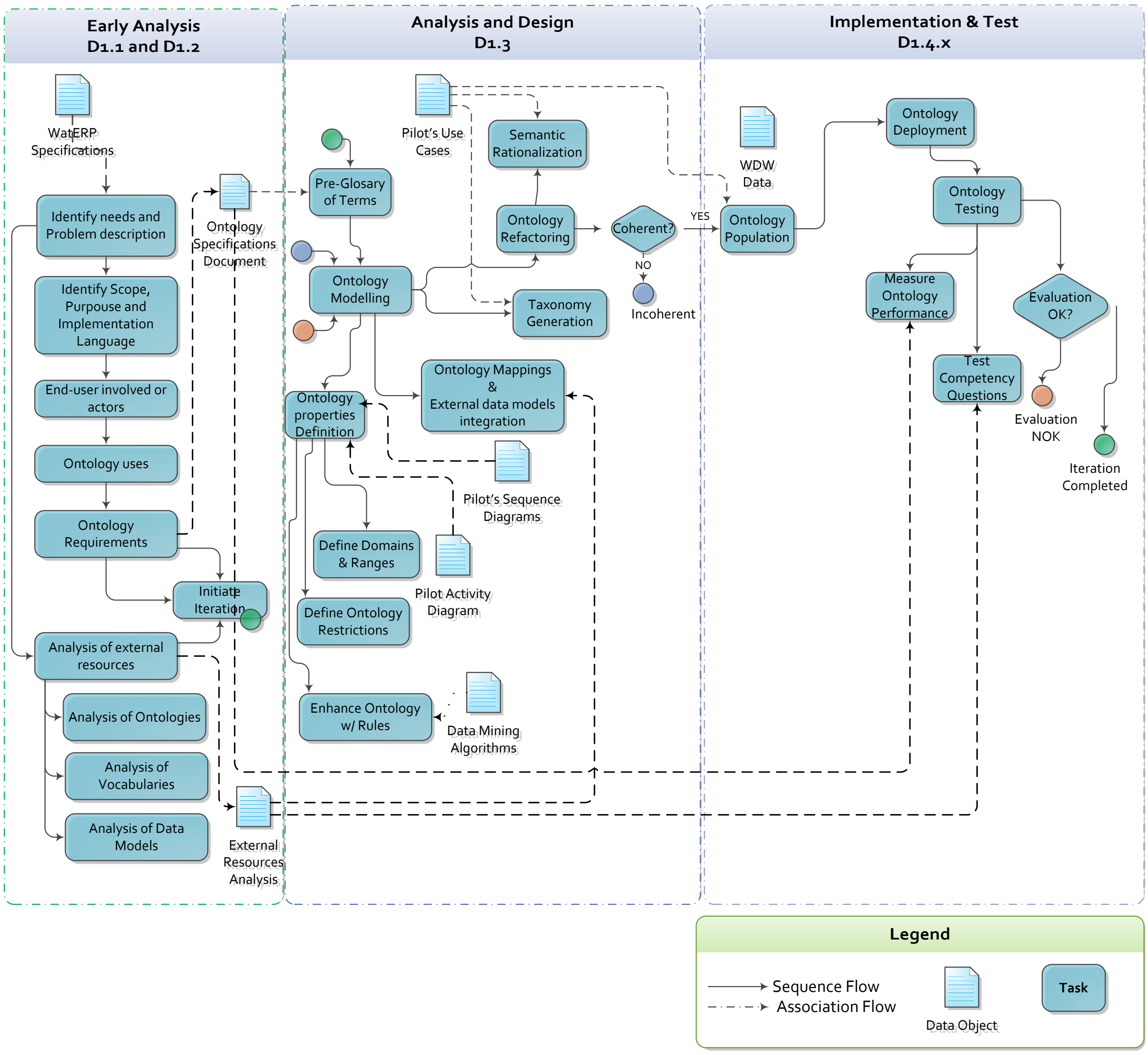

Figure 1 "NeON methodology applied to the WatERP knowledge base development"

Last part of the methodology ("*Implementation and Test*") was based on feeding and testing the WatERP ontology with pilot's information. As depicted in the *Figure 1*, this specific part of the methodology carries out a loop until all evaluations and ontological objectives are satisfied. During this mentioned loop, the process is focused on enhancing the ontology and adapting the semantic concepts according to the pilot's reality. This work is formed throughout the series of deliverables 1.4.x- "*Extension of taxonomy and ontology to the pilots*" (D1.4.x). This series of deliverables are focused on describing the ontological enhancements, population strategy and ontological interaction with rest of the defined building blocks inside the WatERP architecture. Regarding the ontological enhancements, these series of deliverables have permitted to suitable organise the semantic concepts such as "*Phenomena*", "*FeatureOfInterest*" and "*Procedures*" towards an accurate representation of the WaterML2.0 schema. Then, this modification enhances the semantic model by associating the real-object with the observed phenomena and procedure followed to obtain a specific sensor/system measurement (see *D1.4.1 on Section 3.1 in page 27*; and *D1.4.2 on Section 3.1 in page 32*). Furthermore, the WatERP-KB has been enriched with a water data quality model (see D1.4.3 on Section 3.1 in page 28; and D1.4.4 on Section 3.3 in page 32). This water data quality model is based on the latest best practices published by the OGC. This best practices highlights the possibility to link data quality observation with the O&M representation of the information *(Bray & Ramage, 2012; Cox, 2011; OGC, 2011)*. Another highlighted aspect of these enhancements refers to the incorporation of the WatERP-KB geospatial ontology in order to enable semantic geospatial reasoning over the collected information from pilots (see *D1.4.1 on Section 3.1.2 in page 30*; and *D1.4.x on Section 3.1.2 in page 35*).

In reference to the ontological interaction with the rest of WatERP building blocks, information related to the demand management systems developed under WP5 (see D5.1- "*Demand Forecasting Models*" on *Section 6 in page 95*) were included as remarked on *D1.4.2 on Section 3.2 in page 36*. Hereafter, the demand management system information is encapsulated into the decisional process of the water management and new inference and information reinforce can be done. Moreover, the ontology has been also enhanced with the main information regarding the "*economic instruments for water demand management*" (see D5.3-"*Tools for assessing economic instruments for demand management and relevant guidelines & documentation*"). These enhancements have permitted to merge into the management perspective the economic aspects related to hydrologic-domain (see *D1.4.3 on Section 3.2 in page 30* and *D1.4.4 on Section 3.3 in page 34*). As a result of the test performed on D7.1.2 and D7.1.3 both entitled as "*Holistic Auditing*", more data properties and naming refactor were performed in order to accomplish the data needs, ontological development standards and enrich the ontology in concordance with the kind of data that is required to be modelled.

Regarding the ontology population, an automatic process has been developed and tested in order to collect all the real pilot's information (ACA and SWKA) uplifting the data into knowledge. The data is converted into knowledge by contextualising the collected information using the WatERP-KB. Hence, this contextualization is based on using eXtensible Stylesheet Language Transformations (XSLT) in order to transform WaterML2.0 pilot data into semantic resources (see *D1.4.1 on Section 2 in page 13*; *D1.4.2 on Section 2 in page 15*; *D1.4.3 on Section 2 in page 16*; *D.1.4.4 on Section 2 in page 15*; and *D1.4.5 on Section 2 in page 18*).

As an outcome of applying all iterations over the ontology, a final complete version of the ontology capable of automatically publishing pilot's data information to the rest of building blocks are performed. Therefore, the present deliverable is sought of depicting the final version of the ontology including all the sub-parts and the mappings with external resources (see *Section 2*). Complementing the WatERP-KB description, the automatic population procedure is also depicted in *Section 2.1*. In *Section 4*, the WatERP-KB has been documented (see also *D1.4.5 on Section 3.1.5 in page 34*) and published. Finally, main conclusions are presented in *Section 5*.

2. WatERP Ontology

This section is aimed at describing the last stable version of the WatERP-KB. Main purpose to develop the ontology can be found on the *D1.1 on Section 2.2 in page 26*: "***The main purpose of the ontology is to represent the water supply distribution chain with the perspective of resource management. Managing the resources in each part of the water supply distribution chain means monitoring relevant variables aligned with natural cycle of water in conjunction with the man-made (or "human-altered") water infrastructure systems (water levels, water flows, storage, release or treatment volumes, etc.)***". This main purpose of the ontology is aligned with the decision-making process (see *Table 1*) subdivided into different strategies depending on the time horizon (short, mid or long term) and the spatial scale (river basin or end use location). These different management strategies (Operational, Management and Planning) are attached to different water stakeholders (water operators, water utilities, water authorities). Due to this correspondence between stakeholders, the decision-making applied in current water supply and distribution chain is performed in isolation for each stakeholder without considering the other types of decisions. Similarly, the ICT systems involved in operational, management and planning decision-making procedures work as an isolated islands due to stakeholders data needs. However, the interconnection between the decisions at spatial scale (upper and lower parts) and temporal scale are plausible making the holistic water management a promising aspect.

Decision Level		**Spatial scale (Δx)**	
		Upper (River Basin)	**Lower (end use)**
Temporal scale (Δt)	**Operational** (short term)	Time step: day Time frame: week *(rivers, channels)*	Time step: hour Time frame: day *(pipe)*
	Management (seasonal-annual)	Time step: month Time frame: 3-6 months to a year *(reservoir)*	Time step: week-month Time frame: 3-6 months to a year *(tank)*
	Planning (long-term)	Time step: year Time frame: 10-30 years *(regulating infrastructure)*	Time step: year Time frame: 10-30 years *(distribution infrastructure)*

Table 1 "Water supply management decision levels according to spatial and temporal scale"

Focusing on the holistic management perspective, the WatERP-KB covers an abstract representation of all decisional elements involved in the management of the water supply and distribution chain. This abstraction covers economic, financial and water management decisional flows towards representing the different stakeholder's decision-making perspectives aligned with the real measurements collected from the water supply and distribution chain. Hence, these abstractions are represented in the created semantic model (see *Figure 2*) by differentiating the following parts:

- **Matching between supply and Demand (detailed on Section 2.1)**. The elaborated semantic model is capable of representing the matching between water supply and demand by the representation of a "*water resources management*" network/s that indeed, are composed by "*water resources*". These water resources corresponds to different water decisional behaviours such as: *(i)* the management of the source of water; *(ii)* the management of water storage infrastructure (tanks, aquifers, reservoirs); *(iii)* water transportation between two different places; *(iii)* water transformation procedures; *(iv)* distribution of water throughout a certain place (e.g. city). The water resources are aligned with real-elements (device, system, informational source, etc.) defined as "*Feature of Interest*" that at same time collects several measurements according with certain phenomena (named as "*Observation*"). Hence, **the correspondence between water resources (abstractions) and the real-objects (features of interest) permit to separate water management procedures from geometric/geographic perspective**.
- **Observation & Measurement (detailed on Section 2.2).** The Observation & Measurement part of the WatERP-KB is aimed at representing the measurements performed by a real water infrastructure, including natural and human-made infrastructure. These real water infrastructure measurements are represented in form of "*Observations*". The "*Observations*" are capable of depict a "Phenomena" measurement in form of "*ObservationResult*" (e.g. time series, specific value, etc.) through certain measurement procedure ("*Procedure*"). **Using the O&M model to represent the hydrological measurements, the collected hydrologic data is contextualised meanwhile data provenance of the information is ensured.**
- **Decision procedures (detailed on Section 2.3).** This part of the WatERP-KB is focused on representing the alerts and recommendations (actions) of the DSS over the water management structure. Hence, this specific part provides the water managers with proper knowledge generated by the DSS about the behaviour of the water resource management. This representation of the alerts and recommendations is performed by the "*State*" entity to establish current status of the water supply and distribution network and their subsequent water resources. The "*State*" entity could generate different "*Alerts*" that are solved by applying several "*Actions*" over the water resources or the entire network.
- **Water Management Flows (detailed on Section 2.4)**. The water management flows are aligned with the information required for the Demand Management System (DMS) in order to calculate the forecasted demands based on the consumers behaviour. Thus, the managed information is related with the daily habits of the consumers ("*Activities*") and the hydro-meteorological predictions generated by the HMF system. Therefore, this part of the constructed ontology is aimed at **representing the water quantity (flows) transferred between the water industry (Bulk Water Suppliers, Water Utilities and Water Authorities) towards satisfying the end-users**

(consumers) needs. For this purpose, this model also categorises the consumers habits in form of activities (washing clothes, water consumed by appliance) towards better understanding the consumer behaviour.

- **Economic Management Flows (detailed on Section 2.5)**. The economic management flows refers to the water prices that are agreed between the different water stakeholders's to pay the amount of consumed water. Moreover, the semantic representation of the economic water flows are aligned with the key performance indicators for water infrastructure subsides (named as "*Indicators*") and the corresponding mechanisms to enhance the water infrastructure (named as "*Instruments*"). Therefore, the economic management flow representation permits to **manage the water supply and distribution chain since the financial point of view with also planning perspective to enhance the water infrastructure**.

As a conclusion, the combination of these differentiated parts into a global semantic model has permitted to merge different hydrological perspectives (water operational management, financial management, planning management) separated from the physical (real) and geographic dispersion of the elements that compose the water supply and distribution chain (see *Figure 2*). The subsequent sections will describe in depth all of these sub models in order to facilitate the comprehension of the constructed WatERP-KB.

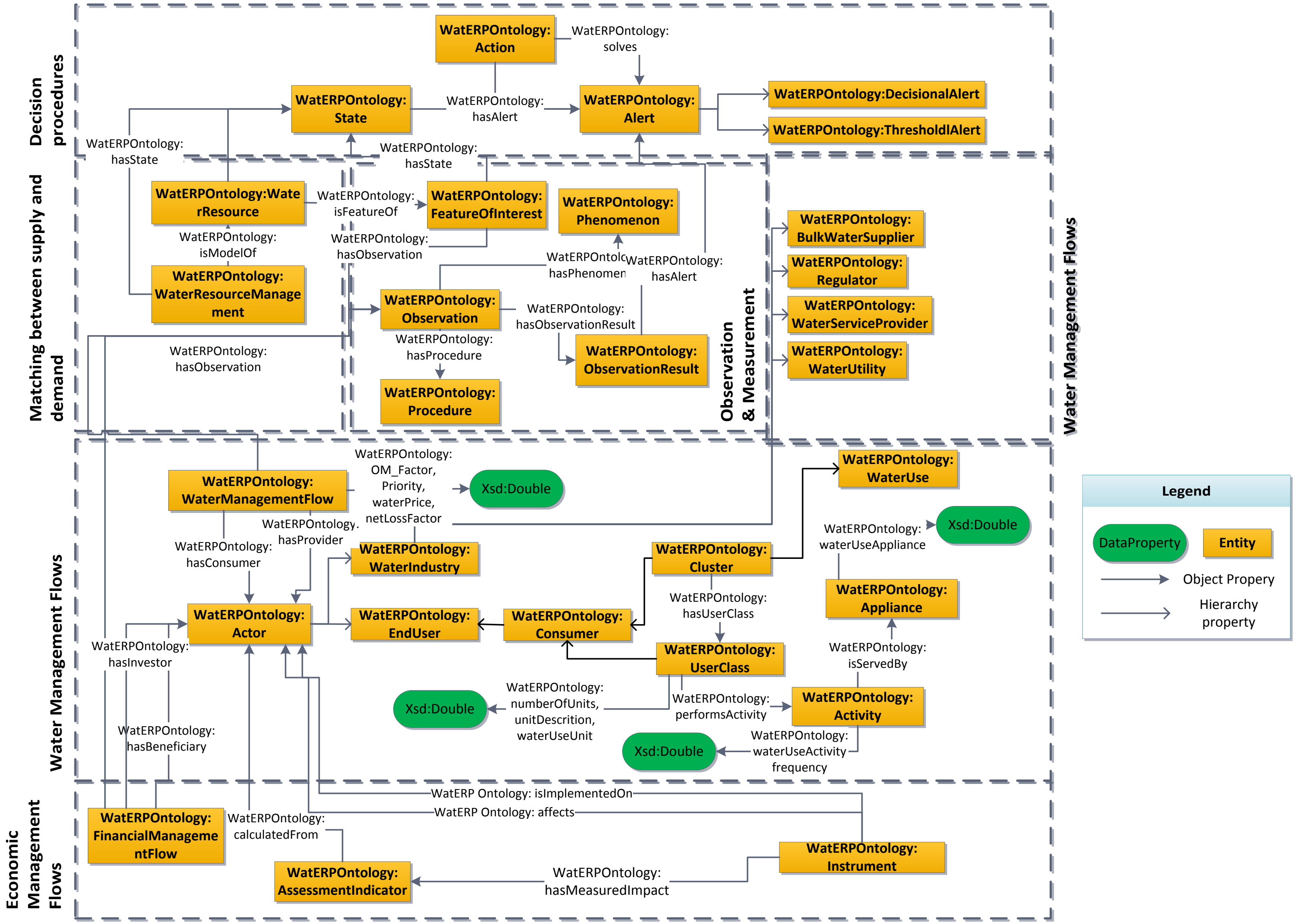

Figure 2"General schema of the WatERP-KB"

2.1 Matching between water supply and demand

The matching between water supply and demand part (see Figure 3) is represented in the WatERP-KB through the "*WaterResourceManagement*" entity. The "WaterResourceManagement" concept corresponds to the procedure to allocate water on an equitable basis to satisfy all demand and uses. Then, the "WaterResourceManagement" corresponds to a network where different water management strategies convey to satisfy the final demand or water use. These different water management strategies are named "*WaterResources*" due to the different water management characteristics that can be found in the water supply and distribution chain. Hence, the "Water Resources" can be (detailed on *D1.1 in Section 6.3 on page 41* and *D1.2 in Section 3 on page 32*):

- **Source**. A water resource defined as an element that provides water resource to the entire water supply and distribution system.
- **Transport** .A water resource capable of moving water (resources) from one place to another in the water chain.
- **Storage.** A water resource that corresponds to retain of water resources for a period for later introducing them back into the system when needed.
- **Transformation.** A water resource capable of modifying the properties of a water resource by applying several procedure/s (e.g. desalination plant that transforms seawater into drinking water).
- **Sink.** A water resource capable of subtracting water from the water supply system (e.g. water consumption from users). Then, this water resource element depends on the water use ("*WaterUse*") or activities ("*Activities*"). The different water uses are categorized in the semantic model into agricultural purposes (e.g. irrigation), environmental use (water left into the natural course), industrial use in the production process, recreational use and urban use (e.g. drinking water consumption).

The "*WaterResources*" involves different real systems named as "*FeatureOfInterest*". These real-systems are able to monitor certain hydrologic-related phenomena (see Section 2.2). The "WaterResourceManagement", "WaterResource", "FeatureOfInterest", and the "WaterUse" entities cover specific spatial locations where different decisions are applied. In order to represent these spatial regions, the Geo-SPARQL ontology developed for the OGC has been imported. The main benefit that the geo-SPARQL offers is based on the geo-spatial reasoning. This kind of reasoning automatically infers the impact of the decision based on the geographic location.

Moreover, the defined concepts have been aligned with external and trustworthy ontologies and vocabularies to homogenize as much as possible all the hydrologic vocabulary. Thus, the "*WaterResourceManagement*" and the "WaterResource" concept has been aligned with "HY_*HydrometricNetwork*" and "HY_*HydrometricFeature*" of the HY_Features ontology *(Atkinson & Dornblut, 2011; Dornblut & Atkinson, 2013)*. This alignment will permit to understand and enrich the WatERP modules with information encapsulated with the HY_Features vocabulary. Additionally, the

"*FeatureOfInterest*" concept has been aligned with the W3C-SSN ontology (Compton et al., 2012) in order to link the WatERP-KB real-objects with the elements categorized under the "*FeaturesOfInterest*" in this ontology and vice versa.

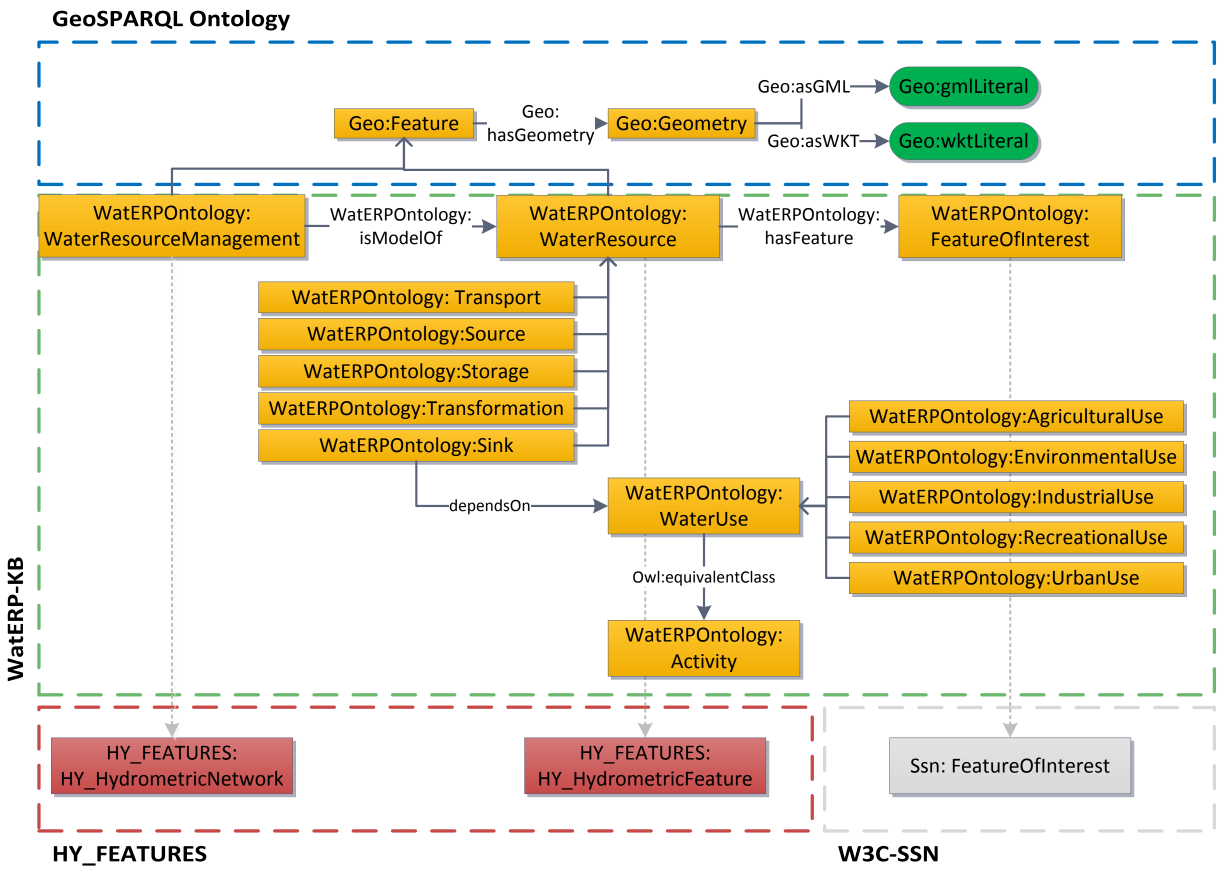

Figure 3 "Matchmaking between water supply and demand representation in the WatERP-KB"

As a conclusion, the incorporation of the "*Matching supply with demand*" semantic model in the WatERP ontology provides **needed abstractions and ontological resources to represent the decisional paths and interrelations that involve the process of the water resource management in the water supply and distribution chain**.

2.2 Observation & Measurement

The Observation & Measurement model *(Kuhn, 2009)* is a standard conceptual model to represent the different measurements produced by the real-objects ("*FeaturesOfInterest*"). This model permits to contextualize the measurements sensed by different sensors installed along the water supply and distribution chain. The main element of this contextualization refers to the "*Observation*" (see Figure 4). The "*Observation*" is the type of measurement performed by a "*FeatureOfInterest*" (using the "*hasFeature*" relation). Each "*Observation*" is defined by the "*Phenomenon*" or observation property that measures (using "*hasPhenomenon*" property); the followed "*Procedure*" to perform the measurement;

and the "*ObservationResult*" referring to the specific measurements of the observation (using the "*hasObservationResult*" properties). This depicted model is semantically aligned with the W3C-SSN ontology ("*ssn:Observation*", "*ssn:ObservationProperty*" and "*ssn:Phenomenon*"). Moreover, the features of interest are aligned with the HY_Features semantic model at the entities "*HY_BodyOfWater*", "*HY_AtmosphericHydroFeature*", "*HY_SubSurfaceHydrofeature*" and "*HY_HydroFeature*".

In detail, the "*SampleFeatures*" and "*SamplingFeatures*" are the types of "*FeatureOfInterest*" entity to represent the real world (as defined in the WaterML2.0 schema). The "*SampleFeature*" refers to the natural water infrastructure ("*BodyOfWater*)" and the human-made hydrological infrastructure ("*Infrastructure*"). Thus, the "*BodyOfWater*" is subdivided into the same categories has defined in the hydrosphere or water cycle ("*AtmosphericHydrology*", "*GroundWaterHydrology*" and "*SurfaceHydrology*"). Hence, the "*AtmosphericHydrology*" corresponds to any hydrology element above the land surface (e.g. "*Wind*"). The "*GroundwaterHydrology*" refers to the subsurface types of water as "*Aquifers*", "*Infiltration*" and "*Reservoirs*" among others. The "*SurfaceHydrology*" categorizes the concepts for the surface (over land) water bodies as "*Basin*", "*Channel*", "*Dam*", "*Sea*", "*Land*", etc. On contrary, the "*SamplingFeature*" defines the features of interest by their geometric structure categorizing the real-objects into "*SF_SamplingCurve*", "*SF_SamplingPoint*", "*SF_SamplingSolid*", etc; according to the WaterML2.0 documentation at *(Hydrology Domain Working Group, 2014; OGC, 2011)*.

The "*Phenomenon*" entity is composed by all the properties that are observed by a certain feature of interest. Most of specific categorized phenomenon are semantically aligned with the SWEET ontology (e.g. "*Medium*", "*Momentum*", "*Rainfall*" among others) *(DiGiuseppe, Pouchard, & Noy, 2014; Raskin & Pan, 2005)*. The "*Phenomenon*" entity is subdivided in the different natural properties that can be found in the environment, categorized partially following the CUAHSI ontology *(Maidment, 2005; Open Geospatial Consortium Inc. & Inc., 2007)*. These properties refer to the "*ChemicalPhenomena*", the "EconomicPhenomena", the "*HydrologicPhenomena*" and the "*PhysicalPhenomena*". The "*ChemicalPhenomena*" represents the composition, conformation and other chemical properties of water. Then, "*ChemicalPhenomena*" covers the water "*Concentration*", "*Humidity*", "*HumicSubstance*" and "*Material*" among others. The "*HydrologicPhenomena*" as the name indicates is focused on categorizing the hydrological domain properties that can be measured (e.g. "*LandCover*", "*PanEvaporation*", "*RunOff*", etc.). The "*PhysicalPhenomena*" refers to physico-natural properties such as "*Evapotranspiration*", "*Temperature*", "*Flows*", etc.

The "*Procedure*" entity is able to represent the different used process to measure the observation and indeed, the phenomenon. Therefore, the different identified procedure refers to *(i)* "*Algorithms*" for those observations that uses a combination of different specific variables inside an equation/algorithm; *(ii)* "*ManualMethod*" for those observation that are manually measured; *(iii)* "*Sensor*" categorizing those observation generated by sensors; *(iv)* "*Simulation*" for contextualizing the observations that are result of using simulation tools; *(v)* "*Unknown*" for those process that are not explicit defined. All of these procedure methods are directly aligned with the procedures defined in the WaterML2.0 schema.

Closing the observation contextualization, the "*ObservationResults*" or the resultant specific measurement for an observation is formed by "*SpecificValues*", "*TimeValuePair*" or "*TimeSeriesObservation*". The "SpecificValues" refers to values that are not dependent on time. This data type reflects only the measured value. The "*TimeValuePair*" is related with a pair composed by the date value of the measurement and the specific value for the defined time stamp. The "*TimeSeriesObservation*" is aligned with a set of "*TimeValuePairs*" in order to represent a time series of values ("*isSetOf*" relation). For all of the "*ObservationResults*" the units are directly aligned with the "*Quantity, Units Dimension and Type*" (QUDT) ontology *(Hodgson, Keller, Hodges, & Spivak, 2014)*. More precisely, the "*ObservationResult*" is aligned with the "*qudt:Unit*" entity through the "*qudt:unit*" property. Then, each observation result is aligned with a standard semantic model to represent the collected measurements.

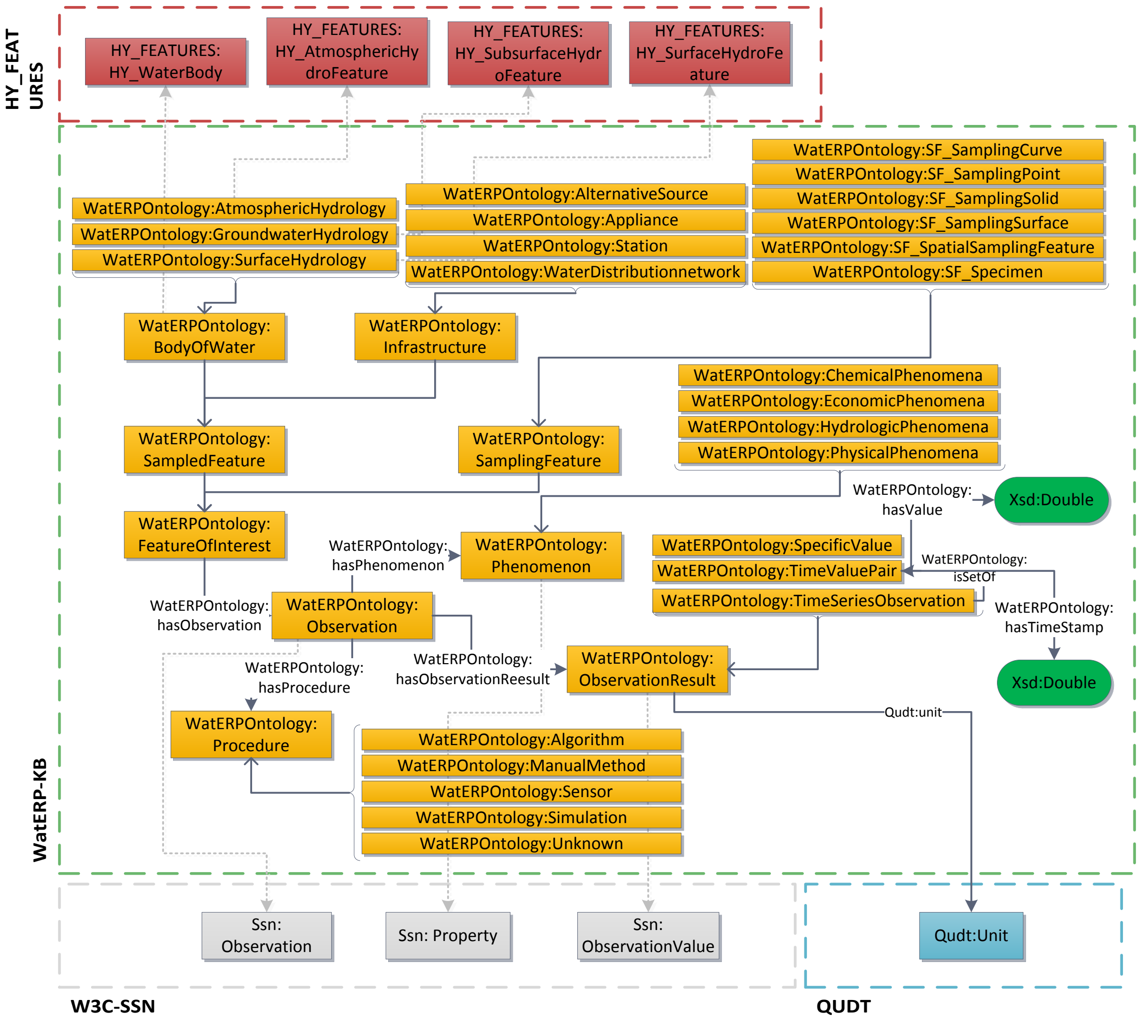

Figure 4 "Observation & Measurement model applied in the WatERP-KB"

Similarly as contextualising the collected measurements from the real-objects, the O&M conceptual model has been also used to measure water data quality. This kind of measurements is representative in order to determine the data usefulness to re-use it in other hydrologic systems or simulation models. In order to be compatible with a coherent water data quality contextualization, the WatERP-KB (see Figure 5) has extended the previous O&M model with: *(i)* definition of a "*WQ_Observation*" class as a child of "*Observation*", *(ii)* creation of "*WQ_Measurement*" class as a child of measurement to represent quality procedure; *(iii)* incorporation of "*Quality*" term from QUDT and making it equivalent to "*WatERPOntology:Quality*" term that indeed is a child of "*Phenomenon*" class; *(iv)* creation of "*WQ_MeasurementTimeSeriesTVPObservation*" and "*WQ_MeasurementTimeSeriesTVP*" class as a child of "*TimeSeriesObservation*" and "*TimeValuePair*" respectively to measure the time series observation and the specific values pairs that takes part of certain time series. Furthermore, the "*ObservationResult*" (time series) has been enhanced by the incorporation of a property ("*qudt:unit*") and a "*Unit*" class from the QUDT ontology to describe the unit of the time series instead of the "*Unit*" data property defined in previous version of the WatERP ontology.

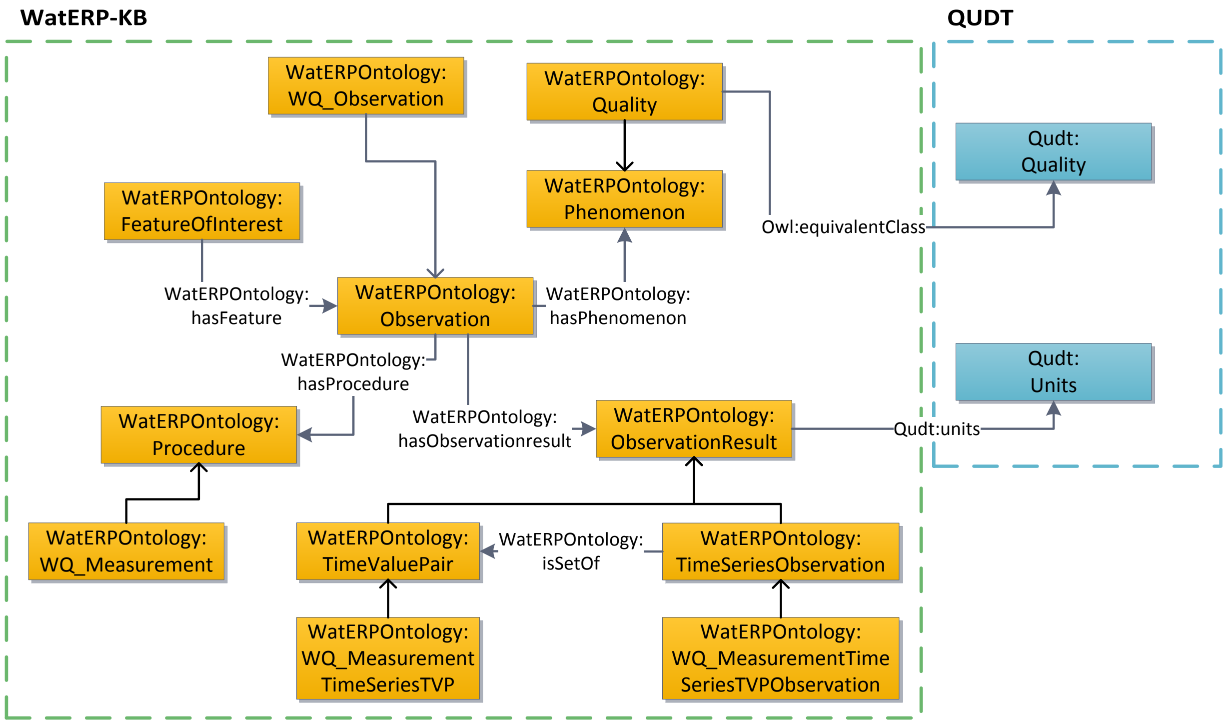

Figure 5 "Water Data quality modelling in the WatERP-KB"

As a conclusion, the depicted part of the WatERP-KB ensures a suitable **contextualization of the measurements collected from the hydrologic systems**. Furthermore, this semantic model also permits to **establish the water data quality of the information for further reuse of the information**. This last aspect also differentiates the WatERP ontology from current semantic models.

2.3 Decision procedures

The decision procedure represents the different alerts, recommendations, states and actions that are involved in the water resource management of the water supply and distribution chain (see Figure 6). This model is strongly aligned with the information covered by the decisional tools (DSS) developed in the WatERP-KB.

In detail, the "*WaterResourceManagement*", "*WaterResource*" and "*FeatureOfInterest*" entities have attached certain state ("*hasState*" semantic object property) as a result of the application of the Rule-Based reasoning (see *D4.4 on Section 2.2 in page 31*). This "*State*" refers to the description of the current situation regarding the environmental variables over these elements ("*Abundance*", "*Normality*" or "*Scarcity*"). Depending of the state, some "Alerts" can be generated ("*hasAlerts*" relation). The "*Alert*" serves to notify of an approaching abnormality, danger, emergency, or opportunity regarding the management of water resources. Hence, the "*Alert*" is subdivided into "*DecisionalAlert*" and "*ThresholdAlert*". The "DecisionalAlert" states those abnormalities generated from applying some decisions over the water resources. On contrary, "*ThresholdAlert*" are caused by anomalous values in certain observation results. Independently of the type of alert, some "*Actions*" are performed to solve it (using "solve" relation). With the matching done between alerts and actions, water managers can know which actions have been applied to solve specific alerts and the ontology can provide them further similar cases.

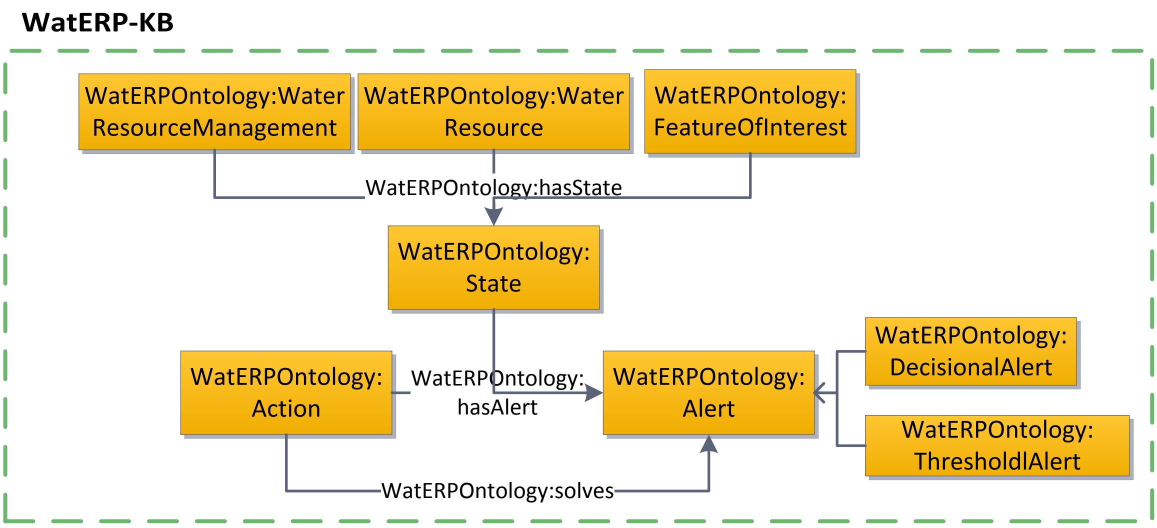

Figure 6 "Decision representation in the WatERP-KB"

As a conclusion, the management of decisions, alerts and actions are beneficial in order to know the **current state of the water resources and the whole water supply and distribution network**. Furthermore, this model also has served to make interoperable the information provided by the DSS of the WatERP platform.

2.4 Water Management Flows

The water management flow is aimed at representing different water uses for the involved water stakeholders. Moreover, this part of the ontology is also focused on depicting the transferred amount of water between water stakeholders. This kind of information is highly valuable for the Demand Management System developed in the WatERP (see *D5.3 on Section 8 page 72*).

The proposed model (see *Figure 7*) categorizes the "*Actors*" into "*EndUsers*" and "*WaterIndustry*". On the one hand, "*EndUsers*" actors denote the water consumers that indeed are subdivided into different "*Clusters*" (groups of consumers grouped by similar behavior) and "*UserClasses*" (family, households, etc.). On the other hand, "*WaterIndustry*" actors depicts the water industry subdivided into "*WaterServiceProvider*", "*BulkWaterSupplier*", "*Regulator*" and "*Water utility*".

The "*Cluster*" depends on several "*WaterUses*" and takes part on several "UserClass" ("*hasUserClass*" relation). Different user classes (family, households, etc.) perform several "*Activities*" during the day ("*performsActivity*" relation). The performance of several activities during a day contributes to the water consumption by the usage of different "*Appliance*" ("*isServedBy*" relation). Therefore, this representation permits to a better understanding the water consumption from the consumer's point of view.

Since the perspective of the water industry, the amount of water transferred until achieving consumers' needs is represented through "*WaterManagementFlow*". The "WaterManagementFlow" is a N-ary relation pattern *(Hayes et al., 2006)* in charge of semantically modelling the water flows between actors, activities and appliances. Specifically, the defined entity permits to specify: *(i)* the water provider ("*hasProvider*") as the actor in charge of providing water to other actor ("*WaterIndustry*" entity); *(ii)* the water consumer defined as the "*Actor*" that consumes the water ("*EndUser*" or "*BulkWaterSuppliers*" or "*WaterServiceProvider*"); and *(iii)* the water consumption performed in form of "*Observation*". Moreover, attached to the water management flow information can be included: *(i)* a "*priority*" of the water flow when an actor receives water from multiple providers; *(ii)* an "*OM_Factor*" to represent the annual operation and management cost per unit of water flow; *(iii)* "*netLossFactor*" that measures the water loss factor for the distribution of water through the water network; and *(iv)* "*waterPrice*" to associate to a "*WaterManagement*" for a water price.

As a conclusion, this representation permits to understand the water transference between water stakeholders in the water supply and distribution chain. Thus, **the water manager is able to know how the water is consumed and how much water is provided from regulators and service providers, enabling to adjust the water consumption with the economic cost of water** (see Section 2.5).

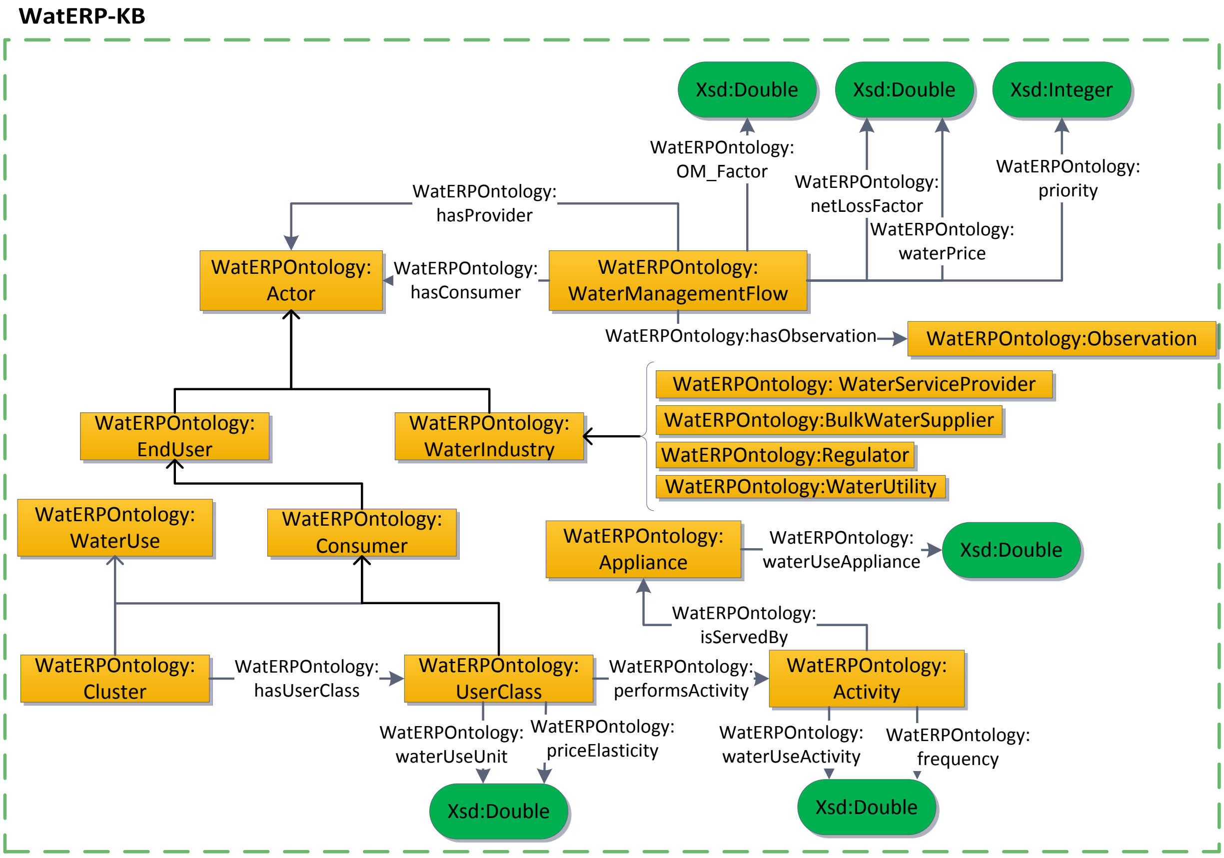

Figure 7 "Water management Flows represented in the WatERP-KB"

2.5 Economic Management Flows

The economic management flows model aims to represent the financial aspects that are involved in the improvement of the water resource management chain. Thus, this section is focused on representing the investments and economic water prices agreed between the water stakeholders. Moreover, this part of the ontology is also aimed at representing and contextualizing the improvement lines of the water supply and distribution chain through the definition of an investment model based on key performance indicators of the network. The represented financial information has been gathered from the informational model implemented in the WatERP modules named as Economic Instruments for enhancing the Demand Management System (see *D5.3 on Section 3 in page 23*).

Thus, the implemented model (see Figure 8) is mainly focused on the interaction between "*Actors*" that are subdivided into the "*EndUsers*" and "*WaterIndustry*". Both, "*WaterIndustry*" and "*EndUsers*" categorization is widely depicted on Section 2.4. Consequently, "*BulkWaterSupplier*", "*Regulator*", "*WaterServiceProvider*" (water industry) and "*Consumers*" (end-users) negotiates between them the water prices and its payment depending on the amount of consumed water. This payment/investment is represented through the "*FinanceManagementFlow*". This entity follows a N-ary pattern to represent: *(i)* the "*Actor*" who pays or performs the investment ("*hasInvestor*"); *(ii)* the "*Actor*" who is beneficiary of the payment or the investment ("*hasBeneficiary*"); and *(iii)* the economic "Observation" (restricted only to those observations with economic phenomena) that includes, for example, the investments, water costs, revenues, water tariffs, abstraction charges and other types of payment.

Additionally, the "*Actors*" are affected from water infrastructure obsolescence and needs to plan it enhancement. Therefore, the "*Actors*" are affected from improvement strategies or "*Instruments*" that are currently applied or are tentative to be applied into the water infrastructure ("*affects*" and "*implementedOn*" relationships respectively). These instruments correspond to mechanism for intensify the improvement of water infrastructure depending on the water use. Then, the identified instruments corresponds to: *(i)* efficient irrigation methods ("*EfficientIrrigationMethodInstrument*"); *(ii)* improvement of water recycling and reuse ("*WaterRecyclingAndResuseInstrument*"); *(iii)* reducing network loses in the water transportation and distribution ("*ReductionNetworkLoosesInstrument*"); *(iv)* subsidizes the adoption of water technology to enhance the infrastructure ("Subsidizes *TechnologyAdoptionInstruments*"); *(v)* adoption of standard technology or smart water technology based on standards ("*TechnologyStandardsInstrument*"); *(vi)* improve water consumption and use in buildings and neighbourhoods ("*WaterEfficiencyForBuildingsInstrument*"); and *(vii)* establishing water prices ("*WaterPricingIntrument*").

The decision for promoting the adoption of one instrument from other is performed by the measurement of key performance indicators, named as "*AssessmentIndicators*" (through "*hasMeasuredImpact*" relation). The "*AssessmentIndicators*" are measurements that analyses specific aspects of the water infrastructure. These specific aspects correspond to the cost recovery ("*CostRecovery*"), affordability ("Affordability") and water saving potential ("*WaterSavingPotential*"). These indicators are calculated from the information that each actor can provide depending of their decisional duties ("*calculatedFrom*" relation).

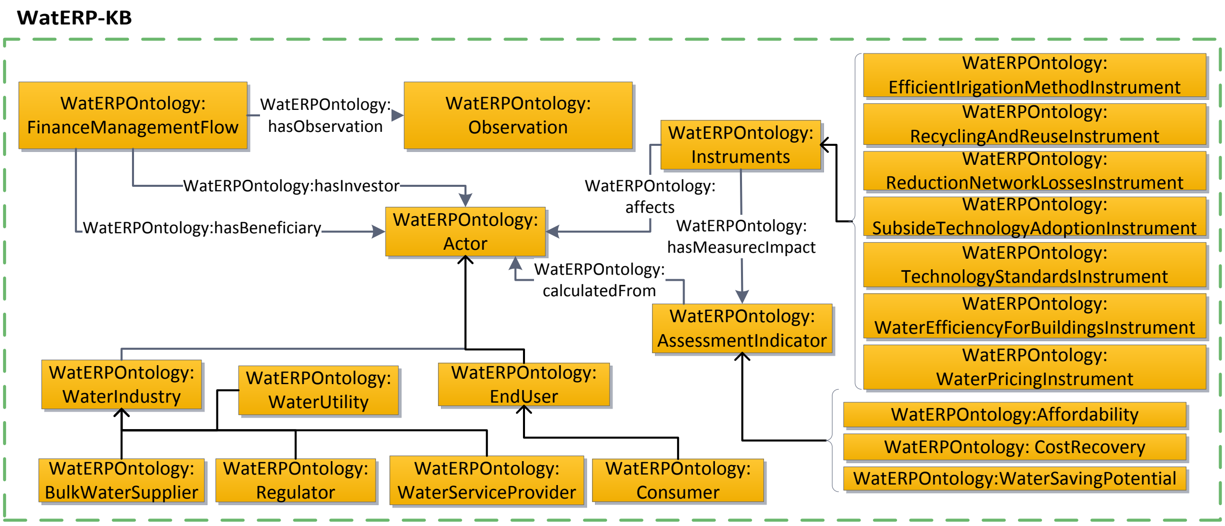

Figure 8 "Financial Flows represented in the WatERP-KB"

As a conclusion, this financial/economic model **describes the economic aspects to transfer water towards satisfying user demands**. Moreover, this model also facilitates the comprehension of the **financial objectives of the water infrastructure by providing needed mechanism to measure key performance indicators. This key performance indicators measurement derives in the implementation or subsidise of new mechanism to enhance the water supply and distribution network**.

3. Ontology Population

This section provides an in depth description of how the ontology is populated with the information from the pilots (ACA and SWKA). This process was designed with the aim of providing to the rest of building blocks with understandable required information (see Figure 9).

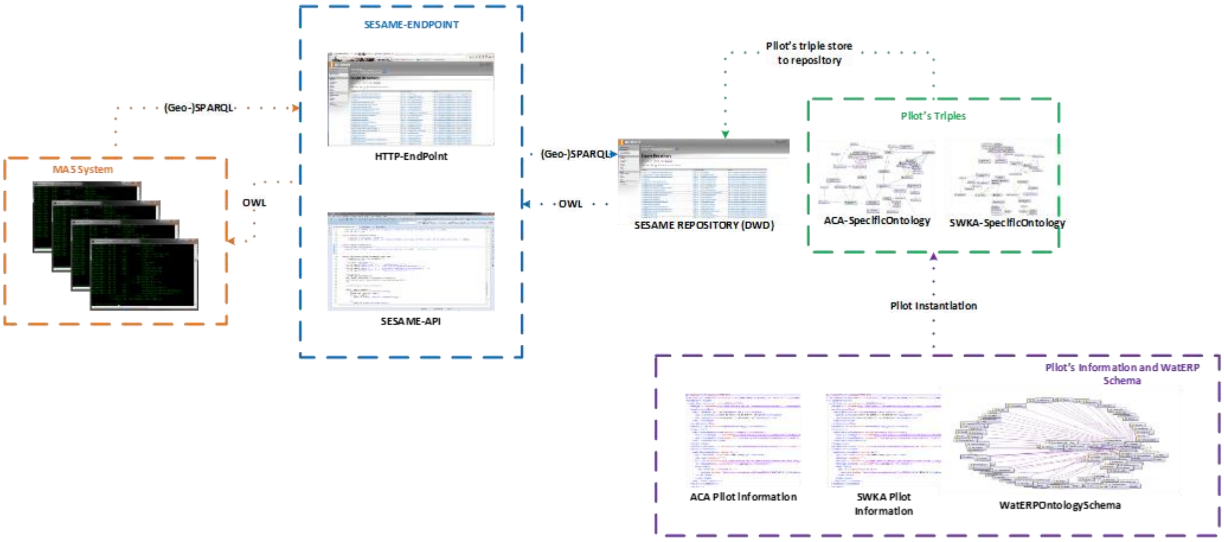

Figure 9 "General population procedure applied in the WatERP project"

This process starts by gathering pilot information serialized in WaterML2.0 notation in order to uplift sensor data into contextual knowledge (specific ontologies). This WaterML2.0 is collected from the Sensor Observation Service (SOS) through the Pilot Integration Manager (PIM) developed under WP3 (see D3.3-"WDW 2nd Prototype"). Hence, this process uses the "*GetObservation*" SOS operation to collect all observations available for each demo-site. This operation is encapsulated in a XML-SOAP request to the SOS OGC server as depicted in Listing 1. The XML-SOAP request permits to ask the SOS-PIM server for all the observations as it is defined under the "sos:GetObservation" tag due to no restriction/filter over the information has been defined. Hence, the response to this operation corresponds to a WaterML2.0 file format that follows the definition given under the "*reponseFormat*" tag.

```
<?xml version="1.0" encoding="UTF-8"?><env:Envelope
xmlns:env="http://www.w3.org/2003/05/soap-envelope"
xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
xsi:schemaLocation="http://www.w3.org/2003/05/soap-envelope http://www.w3.org/2003/05/soap-
envelope">

<env:Body>"

<sos:GetObservation xmlns:sos="http://www.opengis.net/sos/2.0"

xmlns:fes="http://www.opengis.net/fes/2.0"

xmlns:gml=\"http://www.opengis.net/gml/3.2"
```

```
xmlns:swe=\"http://www.opengis.net/swe/2.0"

xmlns:swes=\"http://www.opengis.net/swes/2.0"

xmlns:xlink=\"http://www.w3.org/1999/xlink"

xmlns:xsi=\"http://www.w3.org/2001/XMLSchema-instance"

xsi:schemaLocation="http://www.opengis.net/sos/2.0
http://schemas.opengis.net/sos/2.0/sos.xsd\"" service="SOS" version="2.0.0">

<sos:responseFormat>http://www.opengis.net/waterml/2.0</sos:responseFormat>

</sos:GetObservation></env:Body></env:Envelope>
```

Listing 1 "OGC-SOS request for gathering pilot's information"

The specific ontologies (green colored on Figure 9) are derived from the gathered WaterML2.0 pilot's information (for ACA and SWKA) and the WatERP schema (purple colored on Figure 9) following a transformation process. This transformation process is dependent from the source format. In case of the WatERP project, the input file format is an eXtensible Markup Language (XML) file that follows the - WaterML2.0 schema. Hence, the transformation process is based on reading a WaterML2.0 that contains the observations associated with pilot's information; use the mentioned file to gather the defined specific information, and convert this information into ontological resources by using the definitions of the WatERP knowledge base schema (concepts, object properties, data properties, etc.). To perform this transformation (A-box), a XSLT that indicates how to transform the eXtensible Markup Language (XML) encapsulated information into the WatERP-KB vocabulary, was applied (see Figure 10).

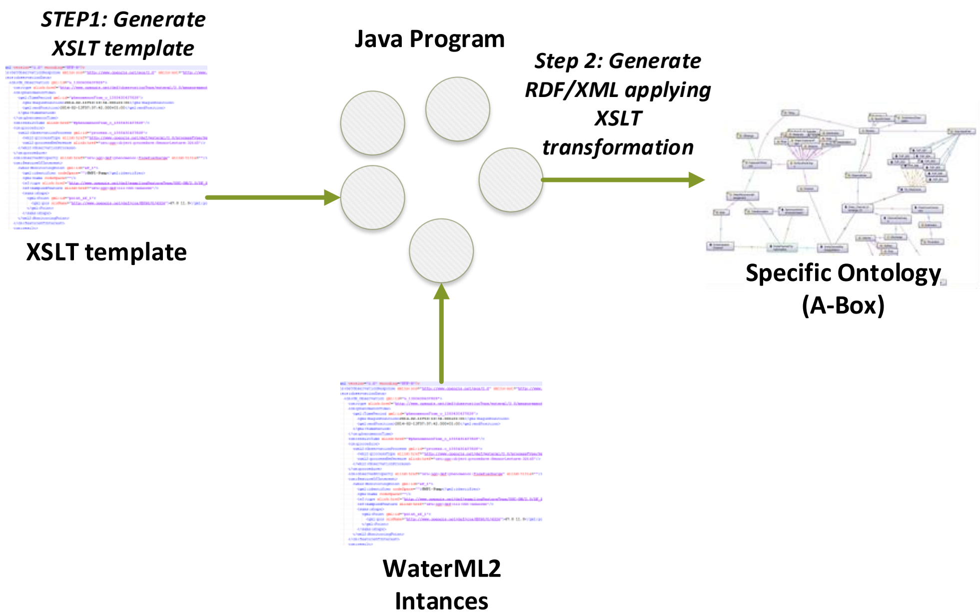

Figure 10 "Applied XSLT transformation procedure"

In detail, the process begins with the generation of a XSLT template (step 1) by following the recommendations proposed by the W3C (Kay, 2007; Pope & Gilman, 2001). The XSLT template structure (Listing 2) is a combination of commands (e.g. “xsl:if”) and XML tags (e.g. “*WatERPOntology:hasProcedure*”) that permit to navigate throughout the XML file (e.g. WaterML2 instance) with the aim of selecting desirable information and transform it to other format (e.g. RDF/XML document).

```
<xsl:if test="om:procedure/wml2:ObservationProcess">
  <WatERPOntology:hasProcedure>
    <rdf:Description                                   rdf:about="http://www.waterp-fp7.eu/WatERPOntology.owl#{om:procedure/wml2:ObservationProcess/@gml:id}">
      <WatERPOntology:id
rdf:datatype="http://www.w3.org/2001/XMLSchema#string"><xsl:value-of
select="om:procedure/wml2:ObservationProcess/@gml:id"/></WatERPOntology:id>
      <WatERPOntology:name
rdf:datatype="http://www.w3.org/2001/XMLSchema#string"><xsl:value-of
select="om:procedure/wml2:ObservationProcess/wml2:processReference/@xlink:href"/></WatERPOntology:name>
      <xsl:if
test="om:procedure/wml2:ObservationProcess/wml2:processType/@xlink:href">
      <rdf:type
rdf:resource="{om:procedure/wml2:ObservationProcess/wml2:processType/@xlink:href}"/
>
      </xsl:if>
    </rdf:Description>
  </WatERPOntology:hasProcedure>
</xsl:if>
```

Listing 2 " XSLT template for the 'has Procedure' ontological resource"

Once the XSLT template has been created, a Java program (step 2) is in charge of collecting WaterML2.0 files and transforming into Resource Description Framework (RDF)/XML by using a specific XSLT template (e.g. “*WatERPXSLT.xsl*”). Specifically, the Java program (Listing 3) reads in an input stream a XSLT template and a WaterML2.0 file. By using these two read streams, the transformation is done by using the “*transform*” function that is aimed of merging both input streams and applying the defined XSLT rules (codification). As a result, an output stream (e.g. “*outResult*”) is created. Similarly, as performed in the previous section, the generated output stream that corresponds with a RDF/XML file is validated and saved into another file by using Jena library *(Jena, 2007; McBride, 2002)* . If the validation is performed correctly, then the resultant RDF/XML (A-Box) is a valid file that can be introduced into the Sesame server.

```
transformer = factory.newTransformer(xslSource);
transformer.transform(xmlSource, outResult);
transformOutputStream.close();
ByteArrayInputStream             modelInputStream             =             new
ByteArrayInputStream(transformOutputStream.toByteArray());
Model rdfModel = ModelFactory.createDefaultModel();
rdfModel.read(modelInputStream, null, outputformat.getName());
rdfModel.write(outputStream, outputformat.getName());
outputStream.flush();
return outputStream;
```

Listing 3 " JAVA code to perform XSLT transformation"

The Sesame repository is located inside the WatERP-Water Data Warehouse (WatERP-WDW) (see D3.2–"*WDW 1st prototype*"). The process to feed the Sesame server with the generated A-Box files is performed by using the Sesame library (Broekstra, Kampman, & Harmelen, 2002; Schenk, 2008). Hereafter, this process is initiated by establishing the Sesame server connection (see Listing 4). The sesame connection is defined by the Unified Resource Location (URL) of the sesame server (e.g. "*http://172.20.10.196:8083/openrdf-sesame*") and the repository (e.g. "*useekm-owlim-1*") where the information is going to be stored.

```
private void  init() {
          Repository repo = new HTTPRepository(url, repository);
          try {
            repo.initialize();
          } catch (RepositoryException e) {
            e.printStackTrace();
          }
          try {
            con = repo.getConnection();
          } catch (OpenRDFException e) {
            // handle exception
            System.out.println(e.getMessage());
          }
    }
```

Listing 4 "Code for establishing Sesame Java Connection"

Once the connection has been established, any procedure in reference with Sesame can be done. In this case, the next procedure is focused on uploading (adding) into the repository the RDF that contains the A-box ontological information from the pilots (result of the population process). To perform this task, the file name that contains the A-Box and the base Unified Resource Identifier (URI) of the ontological model must be provided (see Listing 5).

```
public void populateRepository(String fileName, String baseURI)  {

                 File file= new File (fileName);
                 URIImpl uri= new URIImpl(fileName);
                 try {
                         con.add(file, baseURI, RDFFormat.forFileName(fileName),uri);
                         con.commit();

                 } catch (RDFParseException e) {
                         e.printStackTrace();

                 } catch (RepositoryException e) {
                         try {
                                 con.rollback();
                         } catch (RepositoryException e1) {e1.printStackTrace();}
                         e.printStackTrace();
                 } catch (IOException e) {
                         e.printStackTrace();
                 }
        }
```

Listing 5 "Code to introduce A-Box file into Sesame"

With the pilot's information uploaded into Sesame, the repository also offers HyperText Transfer Protocol (HTTP) and Sesame Application Program Interface (API) end-points in order to exploit the ontological information stored in a non-structured data base (internal to Sesame repository). Therefore, the MAS system is able to execute (geo) SPARQL queries by using the defined end-points. In fact, the end-points receive the queries formulated by MAS and execute the petitions in the Sesame repository. When results were available, Sesame returns the proper response by using the end-point towards the MAS.

As a conclusion, the depicted population process permit to maintain automatically the collected semantic resources from pilots. Furthermore, the depicted procedure also promoted the interaction with the rest of WatERP building blocks using the SOA-MAS as a bridge. The SOA-MAS uses the SESAME connection to perform the needed queries and return needed information/knowledge to the specific building block.

4. Ontology Publication

This section is focused on describing the followed methodology to make the WatERP-KB human understandable and readable for publishing and sharing it with the scientific community. In order to convert the WatERP knowledge base into a readable document (HyperText Markup Language –HTML-), the generated ontology (including all the imported models) has been converted in HTML using Live OWL Documentation Environment (LODE) tool (Peroni, Shotton, & Vitali, 2012). To perform this transformation correctly, the knowledge base has been enhanced with the necessary annotations to document the semantic model (see Table 2). These annotations are being introduced into the WatERP-KB along their development process started with the D1.3-"*Generic Ontology for water supply distribution chain*" (see *D1.3 on Section 2.5 in page 44*).

Annotation	Description
dc:title	Establish the title of the ontology.
dc:date	Date definition of the latest version of the ontology.
owl:versionInfo	Definition of the ontology version.
dc:creator	Definition of the main author.
dc:contributor	Tags to define main contributors to the ontology.
owl:imports	Definition of the imported ontologies.
rdfs:comment	Semantic annotation used to describe semantic resources (classes, data properties or object properties)
dc:description	If the contained information is text, represents the introduction of the ontology. If image (PNG), the image is uploaded into the documentation
rdfs:label	Definition of the name of the semantic resource (classes, data properties or object properties)

Table 2 "Required annotation to perform the ontology publication"

Once the annotations have been defined, the ontology has been exported into RDF/XML syntax to initiate the conversion. The conversion has been performed using the OWLAPI (Horridge & Bechhofer, 2011) as stated in the LODE web-tool (see Figure 11). Mainly, the process is focused on importing the ontology into the OWLAPI. With the model loaded in memory, the LODE offers the possibility to perform directly the documentation over the model or the inferred model. In case of publishing the inferred model, Pellet reasoner (Sirin, Parsia, Grau, Kalyanpur, & Katz, 2007) is used. The conversion between OWLAPI serializations (RDF/XML, Turtle, OWL, etc.) and HTML is performed using XSLT similarly as the depicted in the population section (see Section 3).

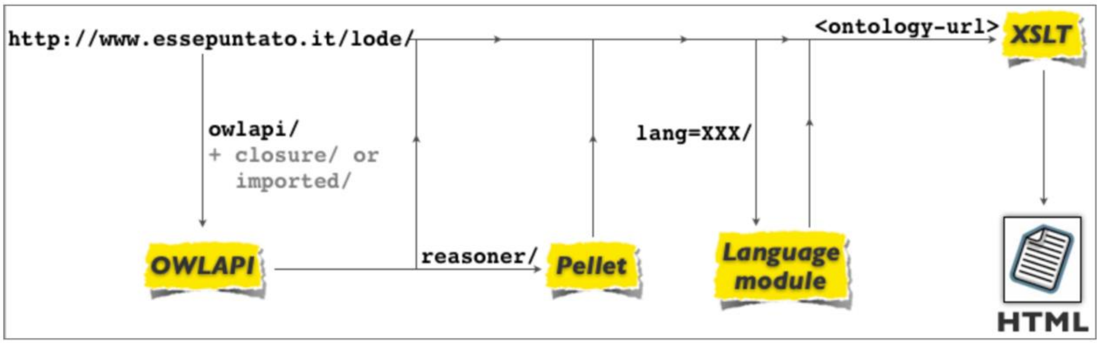

Figure 11 "LODE documentation engine schema (source: (Peroni, Shotton, & Vitali, 2012))"

The result of exploiting these annotations using LODE process is the fully description of the ontology resources in readable and natural language (see Figure 12). As depicted in such figure, the ontology title, authors, ontology version and code is presented in the first part of the document. Following this initial part, the next section depicts the summary (introduction) of the ontology including a description and a graph (image) that defines a conceptual linkage between the main resources described in the ontology.

Powered by

Water Management Ontology

IRI:
http://qudt.org/1.1/schema/qudt

Version IRI:
http://www.waterp-fp7.eu/WatERPOntology.owl

Date:
29-04-2015

Current version:
1.9

Previous version:
http://www.waterp-fp7.eu/WatERPOntology.owl (visualise it with LODE)

Authors:
Gabriel Anzaldi (Barcelona Digital Technology Centre)

Contributors:
Aitor Corchero (Barcelona Digital Technology Centre)
Edgar Rubion (Barcelona Digital Technology Centre)

Imported Ontologies:
http://qudt.org/1.1/schema/qudt (visualise it with LODE)
http://qudt.org/1.1/vocab/OVG_units-qudt-(v1.1).ttl (visualise it with LODE)
http://www.opengis.net/ont/geosparql (visualise it with LODE)

Other visualisation:
Ontology source

Abstract

The Water Management Ontology (WMO) is focused on representing human interactions over the water natural paths as a mechanism to understand the affectation of human decisions over the water resource management. The latter objective is to match water supply with demand by making understandable both infrastructure and management decisions. The knowledge representation is supported by standard semantic models that permit to ensure the data provenance and nature. Hence, the representation of the water data is supported by the W3C-SSN, CUAHSI, NASA-SWEET and OGC standard models. In detail, the W3C-SSN is used to reinforce the representation of water data according to the Observation and Measurement (O&M) concept proposed by the OGC; the CUAHSI ontology has been used to categorise the concepts according to a wide-known taxonomy of concepts; and the NASA-SWEET ontology has been used to reinforce the defined concepts with a relevant third-party model. As a conclusion, the water management ontology is the result of a combination of standard ontologies and models with several management, financial and data quality assurance models towards providing the different water managers with the proper knwoledge related to the water management and distribution.

Table of Content

1. Introduction
2. Classes
3. Object Properties
4. Data Properties
5. Namespace Declarations

Figure 12 "Main page of applying LODE documentation engine to the WatERP-KB"

The third part of the documentation refers to the description of the semantic resources of the WatERP-KB such as classes, object properties and data properties. For each semantic resource (see Figure 13), the Internationalized Resource Identifier (IRI), description, ranges, domains and restrictions are described.

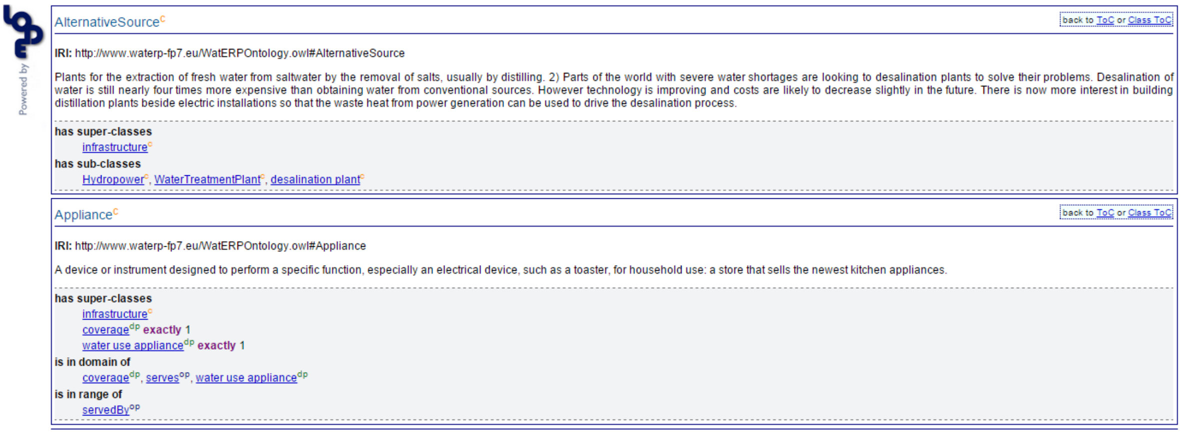

Figure 13 " Documentation of WatERP-KB semantic resources using LODE documentation engine "

As a conclusion of this section, the WatERP-KB has been successfully documented using a semantic documentation tool as a LODE. Hence, the WatERP ontology can be easily published and understood using HTML visualization.

5. Conclusions

The WatERP ontology's main objective is to manage the knowledge of the water supply distribution chain by the perspective of water resource management. Therefore, one of the key aspects of the WatERP-KB is to model the human-made decisions in the water resource management chain towards the semantically structuration and knowledge management of the water resource decisions during different time frames.

Accordingly to this objective, the WatERP-KB has been developed considering water stakeholders decision making process and existent information inside representative pilot's sites as ACA and SWKA. In this line, the WatERP-KB has abstracted the geographic water infrastructure into decisional models. This abstraction has permitted to represent and manage the water supply and distribution chain holistically, considering upper and lower parts as a whole system. Focusing on the decisions to be represented, the WatERP-KB semantically contextualizes and combines water management and financial decisions towards making understandable water payments and planning decisions. This latter aspect permits to enhance the water infrastructure with for example, smart water technology. Transversally, the WatERP-KB uses the standard conceptual model O&M towards contextualizing the different measurements available in the water supply and distribution chain. Furthermore, the ontology is able to perform the geo-spatial reasoning by the incorporation of the Geo-SPARQL ontology. Generally, the WatERP-KB is semantically aligned with standards and representative vocabularies such as CUAHSI, SWEET, WaterML2.0, W3C-SSN and HY_Features to enable the knowledge base to understand the information serialized by these models and then, contribute to semantic interoperability between systems.

Considering all of above aspects, the main capabilities of the resultant knowledge base relies on: *(i)* abstracting human-decisions from physical context in order to support coordinated decision-making between water industry; *(ii)* automatic population process that permit to transparently update the knowledge base information towards offering needed knowledge to the water manage; *(iii)* reasoning over the water resources and geographical information to support the decision-making process in the water supply and distribution chain, *(iv)* represent economic factors aligned with the actors involved in the water supply and distribution chain; *(v)* completely integration with the current standards in water information modelling such as WaterML2 and the OGC-stack; (vi) fully alignment with current standard vocabularies as W3C-SSN, CUAHSI and SWEET; *(vii)* representing water quality data for further reuse; and *(viii)* providing mechanism to automatically document and understand the semantic model, fostering the ontology sharing and enhancement since the ontology development perspective.

6. References

Anzaldi, G., Wu, W., Abecker, A., Rubión, E., Corchero, A., Hussain, A., & Quenzer, M. (2014). Integration of Water Supply Distribution Systems by using Interoperable Standards to Make Effective Decisions. In *11th International Conference on HydroInformatics* (pp. 1–8).

Atkinson, R., & Dornblut, I. (2011). *HY_Features: A common Hydrologic Feature Model Discussion Paper*. *Measurement* (Vol. 2007). doi:http://www.opengeospatial.org/

Bray, A., & Ramage, S. (2012). Open Geospatial Consortium (OGC) OGC Market Report Open Standards and INSPIRE, (April), 1–32.

Broekstra, J., Kampman, A., & Harmelen, F. Van. (2002). Sesame: A Generic Architecture for Storing and Querying RDF and RDF Schema. *International Semantic Web Conference ISWC*, *1*, 54–68. doi:10.1007/3-540-48005-6_7

Compton, M., Barnaghi, P., Bermudez, L., García-Castro, R., Corcho, O., Cox, S., … Taylor, K. (2012). The SSN ontology of the W3C semantic sensor network incubator group. *Web Semantics: Science, Services and Agents on the World Wide Web*, *17*, 25–32. doi:10.1016/j.websem.2012.05.003

Corcho, O., & García-Castro, R. (2010). Five challenges for the semantic sensor web. *Semantic Web*, 2–5. Retrieved from http://iospress.metapress.com/index/F52644V462H28KX4.pdf

Cox, S. (2011). *Observations and Measuremets-XML Implementation*. *Measurement*. doi:http://www.opengeospatial.org/

DiGiuseppe, N., Pouchard, L., & Noy, N. (2014). SWEET ontology coverage for earth system sciences. *Earth Science Informatics*, *7*(4), 249–264. doi:10.1007/s12145-013-0143-1

Dornblut, I., & Atkinson, R. (2013). *HY _ Features : a geographic information model for the hydrology domain*. doi:10.5675/GRDC

Gómez-Pérez, A., Fernández-López, M., & Corcho, O. (2007). Ontological Engineering: with examples from the areas of Knowledge Management, e-Commerce and the Semantic Web. (Advanced Information and Knowledge Processing). Retrieved from http://dl.acm.org/citation.cfm?id=1199560

Hayes, P., Carroll, J., Welty, C., Uschold, M., Vatant, B., Manola, F., … Lawrence, J. (2006). Defining N-ary Relations on the Semantic Web. *W3C Working Group Note*. Retrieved from http://www.w3.org/TR/swbp-n-aryRelations/

Hodgson, R., Keller, P. J., Hodges, J., & Spivak, J. (2014). QUDT - Quantities, Units, Dimensions and Data Types Ontologies. *W3C*. Retrieved from http://qudt.org/

Horridge, M., & Bechhofer, S. (2011). The OWL API: A Java API for OWL ontologies. *Semantic Web*, *2*(1), 11–21. doi:10.3233/SW-2011-0025

Hydrology Domain Working Group. (2014). Standardization of Water Data Exchange: WaterML2.0 and Beyond.

Jena. (2007). Jena Semantic Web Framework. *IEEE Internet Computing*. Retrieved from http://jena.sourceforge.net/

Kay, M. (2007). XSL Transformations (XSLT) Version 2.0. *W3C*. doi:http://www.w3.org/TR/xslt

Kuhn, W. (2009). A functional ontology of observation and measurement. *GeoSpatial Semantics*, 26–43. Retrieved from http://link.springer.com/chapter/10.1007/978-3-642-10436-7_3

Maidment, D. R. (2005). *CUAHSI Hydrologic Information Systems Symposium*. *A Data Model for Hydrologic Observations*.

Malewski, C., Broring, A., Maué, P., & Janowicz, K. (2013). Semantic matchmaking & mediation for sensors on the sensor web, *7*(3), 929–934. Retrieved from http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=6556967

McBride, B. (2002). Jena: A semantic web toolkit. *IEEE Internet Computing*, *6*(6), 55–58. doi:10.1109/MIC.2002.1067737

OGC, C. (2011). *OGC WaterML 2.0: Part 1- Timeseries*. *Measurement*. doi:http://www.opengeospatial.org/

Open Geospatial Consortium Inc., & Inc., O. G. C. (2007). *CUAHSI WaterML*. *OGC 07-041r1*.

Peroni, S., Shotton, D., & Vitali, F. (2012). The live OWL documentation environment: A tool for the automatic generation of ontology documentation. In *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)* (Vol. 7603 LNAI, pp. 398–412). doi:10.1007/978-3-642-33876-2_35

Pope, A., & Gilman, C. P. (2001). XSL: Extensible Stylesheet Language Transformations (XSLT). *Language*, 319–352.

Raskin, R. G., & Pan, M. J. (2005). Knowledge representation in the semantic web for Earth and environmental terminology (SWEET). *Computers and Geosciences*, *31*(9), 1119–1125. doi:10.1016/j.cageo.2004.12.004

Schenk, S. (2008). Sesame RDF Repository Extensions for Remote Sesame RDF Repository Extensions Querying for Remote Querying. *ZNALOSTI Conf*, 375–378. Retrieved from http://znalosti2008.fiit.stuba.sk/download/articles/znalosti2008-Schenk.pdf

Sirin, E., Parsia, B., Grau, B. C., Kalyanpur, A., & Katz, Y. (2007). Pellet: A practical OWL-DL reasoner. *Web Semantics*, *5*(2), 51–53. doi:10.1016/j.websem.2007.03.004

Suárez-Figueroa, M. C. (2010, June 25). NeOn Methodology for Building Ontology Networks: Specification, Scheduling and Reuse. Facultad de Informática (UPM). Retrieved from http://oa.upm.es/3879/2/MARIA_DEL-_CARMEN_SUAREZ_DE_FIGUEROA_BAONZA.pdf