STANDARDS AND OPEN DATA REPORT

SmartH2O

Project FP7-ICT-619172

Deliverable D8.3 WP8

Deliverable
Version 1.1 – 31 July 2015
Document. ref.:
D83.SUPSI.WP8.V1.1

Programme Name:	ICT
Project Number:	619172
Project Title:	SmartH2O
Partners:	Coordinator: SUPSI Contractors: POLIMI, UoM, SETMOB, EIPCM, TWUL, SES, MOONSUB
Document Number:	smarth2o.D82.SUPSI.WP8.V1.1
Work-Package:	WP8
Deliverable Type:	Document
Contractual Date of Delivery:	31 March 2015
Actual Date of Delivery:	31 July 2015
Title of Document:	Standards and Open Data Report
Author(s):	Andrea E. Rizzoli, Bruno Storni, Alessandro Facchini, Piero Fraternali, Andrea Castelletti, Matteo Giuliani, Chiara Pasini, Julien Harou, Jasminko Novak, Isabel Micheel, M. Boeckle, Luigi Caldararu, Sever Calit, Mark Holt.
Approval of this report	Approved by the Project Coordinator
Summary of this report:	A brief overview of relevant standards and open data initiatives in the sector of smart water solutions. Issue of recommendations to promote interoperability and data reuse in the smart water sector.
History:..	see the Document History section
Keyword List:	Smart Water Standards, Interoperability, Open Data
Availability	This report is public

This work is partially funded by the EU under grant ICT-FP7-619172

Document History

Version	Date	Reason	Revised by
0.1	17/11//2014	Initial version, table of content is provided	A.E. Rizzoli
0.2	22/12/2014	Initial review of Smart Metering standards	B. Storni
0.3	1/2/2015	Review of user data and social interaction	C. Pasini, B. Fraternali, M. Boeckle
0.4	3/3/2015	Software interoperability added. Open data review added.	L. Caldararu, S. Calit, M. Boeckle
0.5	25/3/2015	Review of standards for water data	A. Castelletti, M. Giuliani
0.6	25/3/2015	Minor revision	A.E. Rizzoli
0.7	31/3/2015	Final revision	P Fraternali, G Baroffio, J. Novak, M. Boeckle
1.0	31/3/2015	Final approval	A.E. Rizzoli
1.1	31/7/2015	Revision incorporating the reviewers' comments. The executive summary has been rewritten to give more coherence to the document. A new subsection 1.1 on the role of the Internet of Things for Smart Water Systems has been added. A new subsection 3.4 on common practices for serious games development has been added.	A.E. Rizzoli, P. Fratrenali, L. Galli. A. Diniz do Santos

Disclaimer

The research leading to these results has received funding from the European Union Seventh Framework Programme (FP7-ICT-2013-11) under grant agreement n° 619172.

Table of Contents

Executive Summary

This document is an output of Task 8.5 (Open data and standards), which has as main objectives the use of appropriate standards in the design of the SmartH2O platform and also the publication of data generated by the project as Open Data sets.

This deliverable is produced at the end of Year 1, in order to provide useful information to WP6 (Platform Implementation and Integration) for a design of the SmartH2O platform that is aware of the relevant data representation and exchange formats. This awareness will improve the interoperability of SmartH2O and other systems, both commercial water management tools and research water demand management prototypes.

The deliverable is structured as follows:

- The introduction focuses the context of ICT for smart water management for the SmartH2O project labelling it as a system for meeting water demand in cities of the future and it highlights the central role played by data, motivating the need for standards and openness.
- A section titled “the keys to interoperability” outlines the requirements for data and software interoperability and how standards can enforce such requirements, providing the necessary conditions for the future evolution of smart water systems.
- The deliverable the contains a section devoted to a brief survey of the various standards adopted in different sectors, from smart metering to geospatial information, from social data to water management, which are relevant to the SmartH2O project. It provides a brief description of each standard and advises about their adoption in the project.
- A survey of the Open Data initiatives in the area of water management is then introduced. It focuses on initiatives of interest to SmartH2O as a future way for publishing the data sets collected during the experimental activity of the project to the international water management research community. A open data strategy to be followed in the SmartH2O project is also detailed.
- Finally the deliverables proposes a **set of general recommendations** for the adoption of standards and open data for the development of smart water solutions.

1. Introduction

The ITU-T report titled “ICT as an enabler for Smart Water Management” [ITU2010] provides an excellent overview of how ICT can provide a beneficial impact on water management, which is summarised in Figure 1.

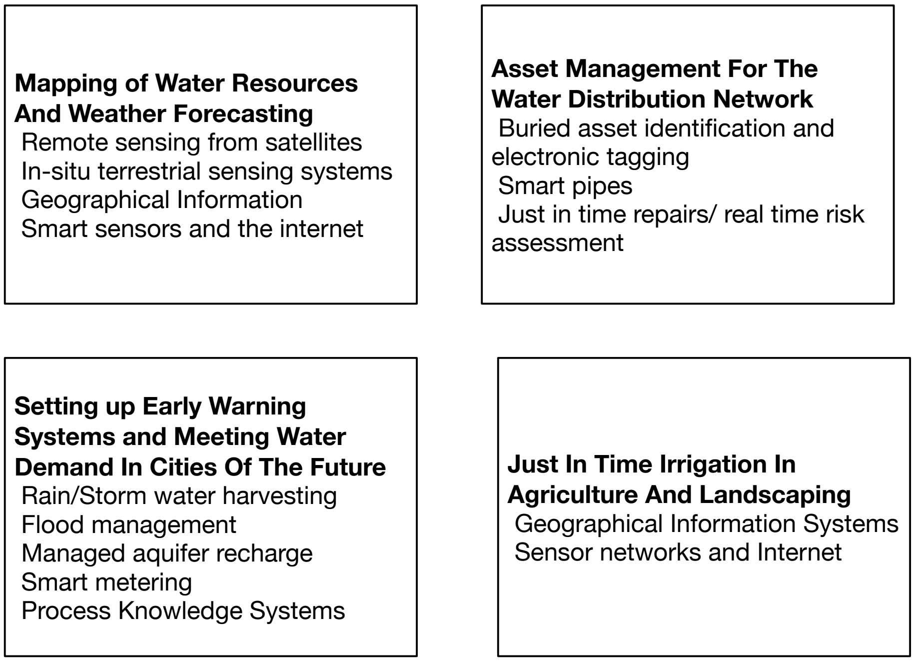


Figure 1. Major areas for ICT in water management. Adapted from [ITU2010].

The SmartH2O project falls in the category of “Systems meeting water demand in cities of the future”. It delivers a software platform to promote water savings thanks to increased awareness, social feedback, and innovative pricing schemes. In order to achieve its aim, the SmartH2O project needs to access, store, and process considerable amounts of data, while at the same time producing the software architecture able to deal with such data, and interfacing with a smart metering infrastructure.

1.1 Smart water systems and the Internet of Things

The components of a smart metering infrastructure can be framed within the Internet of Things initiative (IoT). IoT is a network of physical things endowed with the capability to transmit data over the Internet. The first versions of smart meters were not originally designed to interface with the Internet, as they privileged the direct connection with the utility gathering the data for billing purposes, but in order to broaden the scope of applications of smart meters to provide consumers with access to the data collected by the meters, connectivity over Internet protocols is desirable.

The ITU has launched an “Internet of Things Global Standards Initiative” described in the overview document [ITU2012]. The ITU initiative has to be taken into consideration by developers of smart water solutions who will need to integrate their hardware and software solutions in the IoT context. Also the IEEE is engaged in the development of standards for the IoT.

The integration of Smart Water Systems in the IoT is expected to be beneficial for both the end user and the utilities. In a not so distant future the end user will be able to monitor high resolution individual usage of single fixtures, as the Daiad project[1], an FP7 EU funded project member of the ICT4Water cluster[2], is already showing. The monitoring devices will be IoT enabled, so data can be shared with the utility and/or third parties for monitoring and analytics purposes. Clearly this raises security and privacy issues, which are one of the foremost challenges for the future development of IoT solutions.

1.2 Smart water systems and data

Data are the primary ingredient in the SmartH2O workflow: data about water consumption are gathered by a smart metering infrastructure, and they are stored in our database. Such data are integrated with psychographic data about the user characteristics and are then used to classify users according to their water consumption patterns. Consumers are then advised about potential actions for water saving on the basis of their characteristics and they are engaged in a gamification platform in order to motivate them to continue their progress towards the achievement of their water saving aims.

In the SmartH2O project we have defined a proprietary data model, which we are using to represent the internal information regarding the consumer data model and the user gaming model (see Deliverable D3.1 Databases of User Information). Our choice to develop a proprietary (yet open and public) data model was justified by the requirements of the SmartH2O platform, which has a broad, multi-dimensional spectrum (from water consumption, to users' profiles, to games and social network links) that goes beyond what currently supported the sectorial standards and data exchange formats.

At the same time, we are aware of the need to allow for the interoperability of the different solutions proposed in the Smart Water area, and more in general in the Smart Cities initiatives, and have used the aspect-specific and sectorial standards as blueprints for designing the multi-faceted SmartH2O data model, so to ease the export and import of data from/to the SmartH2O platform.

In the remainder of this document we will analyse how to enhance the interoperability of ICT solutions for urban water management (also defined in the following smart water solutions) thanks to the adoption of standards and the open data initiatives. At the same time, we will highlight which standards are being adopted by, or have influenced the design of, the SmartH2O project.

1 http://www.daiad.eu

2 http://ict4water.eu

2. The keys to interoperability

In this section we analyse how interoperability of smart water solution and open access to data can positively affect the development of research and industry in this sector. For this we also issue a short set of recommendations that we hope they will be adopted not only by our project but also by future projects and initiatives in the area of smart water solutions.

2.1 The data and knowledge supply chain of a smart water solution

Data and software standards are key for the interoperability of “smart solutions” for urban water management. We can assume that a smart solution is based on high quality and high frequency data about the water distribution network, and the core structure of a smart solution is based on the following steps (see Figure 2):

1. data acquisition, using smart meters
2. data transmission, over a communication medium
3. data storage, in a persistent database system
4. data processing and visualisation, by means of data analytics components

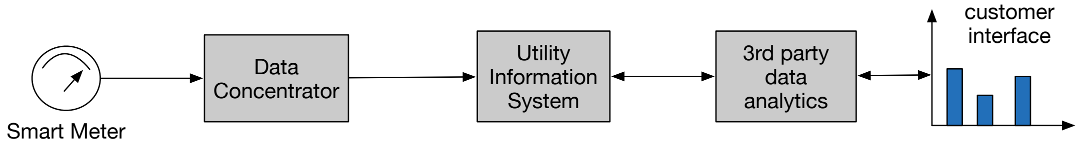


Figure 2. The information flow among the components of a “smart water solution”.

In order to promote a healthy economic ecosystem of business solutions, enterprises, consulting agencies and so on, it is expected that each one of the above steps could be performed by a different actor. It should be possible, for a water utility, to easily integrate a new brand of smart meters in a seamless way, without changing anything in the following layers, if standards are respected.

Clearly this goes against the interests of many solutions providers who try to lock the water utility in a vendor specific solution, which only rarely and at high costs can be made interoperable.

2.2 The levels of interoperability

In the previous paragraph we have briefly outlined how interoperability can enable the realisation of smart water solutions that are more efficient, economic, but they are also more resilient: if a given vendor disappears from the market, the current investment of the water utility is preserved as the negative impact can be limited to a component of the solution, and it does not affect its overall functionality.

Aspects of interoperability have been analysed in the work of Tolk et al [Tolk2006] who identified seven levels (six, if we don’t count the “no interoperability” level) of interoperability, when referred to modelling solutions. The same approach, shown in Figure 3, can be applied to smart water solutions, where the “simulation/implementation” level has to be intended as the level of interoperability of the data representations, while the ”modelling/abstraction” level refers to the data analytics solutions used to extract knowledge from data and provide decision support to utilities.

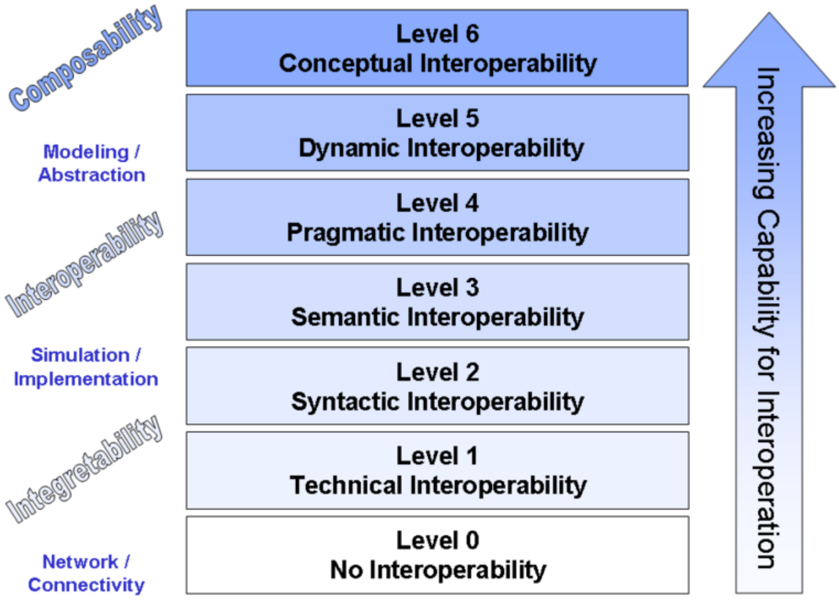


Figure 3. Levels of interoperability in simulation modelling [Tolk2006].

2.3 Openness as a key to interoperability

Another major issue preventing the diffusion of smart solutions for water management is the lack of open accessible data sets. Such data sets are essential to train and develop increasingly better algorithms to profile customer behaviour and habits and provide improved decision support to water utilities.

Tim Berners Lee has proposed a 5 star evaluation scale about the openness of data [TBL2006]. In Table 1 we report the scale as presented in [TBL2006].

Table 1. Tim Berners Lee's 5-star evaluation scale.

★	Available on the web (whatever format) but with an open licence, to be Open Data
★★	Available as machine-readable structured data (e.g. excel instead of image scan of a table)
★★★	as (2) plus non-proprietary format (e.g. CSV instead of excel)
★★★★	All the above plus, Use open standards from W3C (RDF and SPARQL) to identify things, so that people can point at your stuff
★★★★★	All the above, plus: Link your data to other people's data to provide context

Not providing full access to all data should not necessarily have a negative connotation, it is well understood that private enterprises might limit access to proprietary data, but we believe that citizen's data should be in the citizens' hands.

Yet, openness of data can raise privacy and data protection concerns. In Deliverable 8.2 we discuss in Section 4.2 how privacy concerns affect the technological and commercial side of the development of smart meter solutions. Here we highlight how the SmartH2O system adopts the "data protection by design" principles, as no data that can lead to the identification of a specific user are required to be stored on the SmartH2O servers. The upload of sensitive data will be totally under the user's control, as described in Section 4.4 of this document.

In the remainder of this document we will briefly explore which standards are asserting themselves in the smart metering market for urban water and we will also address how open data can partially close the gap between the current availability of historical data and the actual demand of such data for scientific and industrial research.

2.4 Relevant EU Directives and the availability of data

Before reviewing some of the most relevant standards for smart metering, here we mention relevant EU Directives that should be taken into account when developing smart water solutions, especially with respect to the annotation of data and to publishing data related to smart water experiences and studies.

The main directives that should be taken into consideration are **INSPIRE** [EC2015a], which defines the guidelines for interoperability and standards for spatial data, and the Water Framework Directive, which specifies the characteristics of WISE [EC2015b], the water information system for Europe.

The INSPIRE directive aims at creating a spatial data infrastructure for the European Union. Among the main principles on which INSIPRE is based, is the fact that data should be collected only once, and maintained where it is most efficient and effective.

INSPIRE delivers a GeoPortal[3] which allows to browse for data all over the EU member states. Currently searching for “water consumption”, “water demand”, “water use” in this GeoPortal returns 0 results.

In the context of the Water Framework Directive, the Water information system for Europe **WISE** has been created. Some aggregated data on water consumption can be accessed through the Water Data Centre of the European Environmental Agency [EEA2015], but a real geolocated search is not possible. WISE also refers to the Eurostat website [EC2015c], where some datasets on Water use balance are available (only at the country level), but unlike for energy, where Energy consumption statistics at the household level are available, no similar data for water consumption is made available.

Two directives apply to privacy and data protection in the EU: the **Data Protection Directive** (Directive 95/46) and the **e-Privacy Directive** (Directive 2002/58). These Directives will be soon updated in order to also account for issues raised by Smart Metering and one expert group was specifically established to produce “Regulatory Recommendations for Privacy, Data Protection and cyber-security in the Smart Grid Environment”.

[3] http://inspire-geoportal.ec.europa.eu/discovery/

3. A quick overview of relevant standards

In Section 2 we introduced the context and motivated the claim that standards and open data can provide the key to enable interoperability along the data flow chain described in Figure 2, from meter data to human and computer interaction.

In this section we mention some of the relevant standards, which can be employed in the various phases of a smart water solution:

- **Standards for smart metering**, that define how smart meters acquire and exchange data and the communication protocols to be adopted (OMS and OPEN meter). These standards pertain to the data acquisition and data transmission layers.
- **Standards for customer data and water consumption representation**. These standards pertain to the representation of data inboth the water utility information systems and in the third party analytical systems. We have identified the following sub areas:
  - representation of geolocation data;
  - representation of geolocated sensor data and water related data.
  - representation of the user profile data and social interaction connections.
- **Standards and good practices for software interoperability**, useful to interconnect and interoperate software components (web services standards and OpenMI). These standards apply to the interfaces between data analytics components and portals/user interfaces.

Finally we add a subsection that is specific to a key component of the SmartH2O project, which attempts to use gamification techniques to motivate users to a responsible use of water:

- **Common practices in the development of serious games and games with a purpose**. This section does not present standards which are still non-existent, but it points to the most common techniques and systematic analyses of the field of gamification, especially for environmental awareness related applications.

3.1 Standards for smart metering

We briefly describe two major initiatives leading the path towards standardisation for smart metering in Europe: the OMS group [OMS2014] and the OPEN meter project[4]. The former originated mainly from the association of German producers of Smart Meters, while the latter is the result of a EU backed project.

3.1.1 OMS, the open metering system

The basic and most common communication infrastructure in water metering application is Open Metering System OMS [OMS2014]. It has been defined in the frame of a Multi Utility Communication (MUC) concept intended to integrate the different utility metering services, from electricity, to gas, to heat (hot water) and drinking water, in particular to provide a coherent set of ICT hardware and software tools for billing purposes. OMS is based on the European Meter-Bus standard (EN 13757) as well as the Dutch NTA 8130. Meter –Bus (M-Bus) is a European standard (EN 13757-2 physical and link layer, EN 13757-3 application layer) for the remote reading of gas, water, or electricity meters.

An overview of how OMS is supposed to operate is provided in Figure 4.

OMS specifications have been implemented by working groups whose members are issued from Municipal utilities and companies producing metering products. In particular FIGAVA (Bundes Vereinigung der Firmen im Gas und Wasserfach e.V.,Cologne) and ZVEI (Zentralverband Elektrotechnik- und Elektronikindustrie e. V. Frankfurt/Main) who acted as

---

4 http://openmeter.com

chairs.

Communication is organized in 3 levels: primary, secondary and tertiary.

Primary communication defines communication protocols, transmission techniques and media between multi-discipline metering devices (electricity, gas, heat, water meters).

Secondary communication specifies a routing protocol for extended coverage on wired, and wireless communication. It also considers multi hopping.

Tertiary communication is the interface between the Multi Utility Communication (MUC) systems and back office systems for Automated Meter Management (AMM). The protocol defines push and pull data flow procedures. It also specifies how data acquisition, presentation, event handling, configuration, control and clock synchronization should take place.

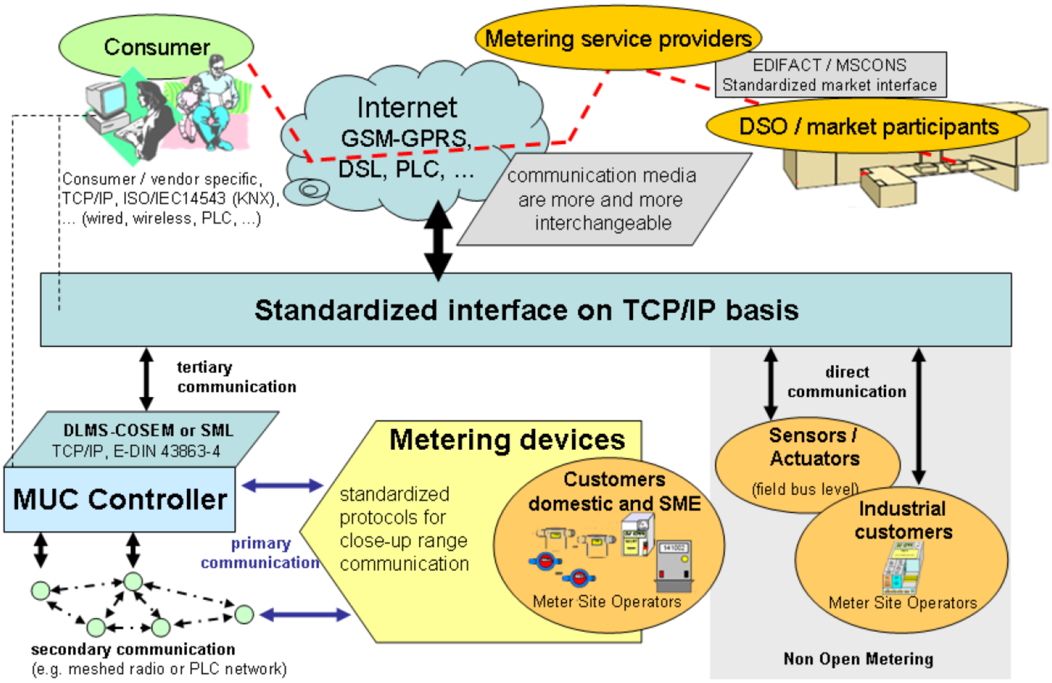


Figure 4. OMS system overview (OMS Specification Vol. 1 General Part).

3.1.2 The OPEN meter project

The OPEN meter project was an FP7 funded project which had the objective to specify a comprehensive set of open and public standards for the Advanced Metering Infrastructure, supporting electricity, gas, water and heat metering, based on the agreement of all the relevant stakeholders in this area, and taking into account the real conditions of the utility networks so as to allow for full implementation. The project ended in 2011, and it also produced a specification of Open Meter OSI layers and multi-metering networking interfaces[5].

OPEN meter draft standards have already been accepted by major European utilities and are being fed into the European and international standardization process.

Some OPEN meter partners are also members of the OMS group, so an even stronger convergence of the efforts is expected in the future.

[5] http://www.openmeter.com/files/deliverables/D3.2_v2.0.pdf

3.2 Standards for customer data and water use

3.2.1 Data geolocation and GIS standards

Water use and water demand data must be geolocated to make sense. Note that a precise geolocation, at the level of the individual household, could raise privacy concerns; water and energy data are highly sensitive, as they could reveal a lot about the habits of the consumer, and be used for malicious intents, such as burglars wanting to be sure about the absence of the landlords. In the SmartH2O project we will use data geolocated at the postcode level, or at the level of a district metering area, thus making it impossible to associate the consumption trace with a specific household.

Despite this premise, the adoption of a standard to represent spatial data is recommended. The principal organisation providing standards for spatial data is the Open Geospatial Consortium (OGC). Among relevant standards issued by the OGC we mention: GML the Geography Markup Language, the Geographic Metadata XML encoding (GMD).

In Figure 5 we show the role of OGC standards as a bridge between more abstract ISO standards, which define the terms of discourse (content, vocabulary), and the W3C and IETF standards, which focus on the implementation infrastructure to deliver information on the Web.

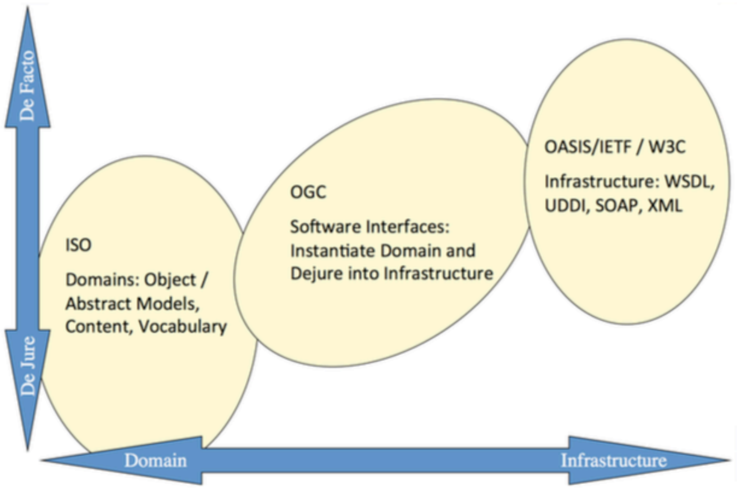


Figure 5. Where the OGC fits in the standards world [OGC2015].

The OGC is organised in Domain Working Groups, which focus on specific sectors and aspects of standardisation of location data. Of particular relevance for the SmartH2O project is the Energy and Utility Domain Working Group (E&U DWG[6]), which was charted in 2012 with the task of developing standards for the Smart Grid. The OGC has then released the OGC Smart Cities Spatial Information Framework[7], which is not yet a standard, but a white paper. This document approaches the comprehensive problem of spatial and location standards from the smart city, defined as an entity that "… provides effective integration of physical, digital and human systems in the built environment to deliver a sustainable, prosperous and inclusive future for its citizens".

In the Energy & Utility Domain Working Group the OGC is actively integrating different standards that can be jointly used to provide an integrated solution to standardisation in Smart Cities. In particular the OGC's Sensor Web Enablement (SWE) standards enable to

---

[6] http://www.opengeospatial.org/projects/groups/energyutilities

[7] https://portal.opengeospatial.org/files/?artifact_id=61188

make sensor data repositories discoverable, accessible and useable via the Web. The SWE is composed of:

- Observations & Measurements (O&M) –The general models and XML encodings for observations and measurements.
- Sensor Model Language (SensorML) – Standard models and XML Schema for describing the processes within sensor and observation processing systems.
- Sensor Observation Service (SOS) – Open interface for a web service to obtain observations from one or more sensors.
- Sensor Planning Service (SPS) – An open interface for a web service by which a client can 1) determine the feasibility of collecting data from one or more sensors or models and 2) submit collection requests.
- PUCK Protocol Standard – Defines a protocol to retrieve a SensorML description, sensor "driver" code, and other information from the device itself, thus enabling automatic sensor installation, configuration and operation.
- SWE Common Data Model – Defines low-level data models for exchanging sensor related data between nodes of the OGC® Sensor Web Enablement (SWE) framework.
- SWE Service Model – Defines data types for common use across OGC Sensor Web Enablement (SWE) services.

Thanks to the SWE stack of standards, access to sensor data in Smart Cities, including data on smart water sensors, will be made accessible in a transparent and uniform manner.

It should be finally remarked that both INSIPIRE and GEOSS (Global Earth Observation System of Systems) offer their data according to the Geospatial Portal Reference Architecture offered by the OGC.

3.2.2 WaterML

While the OGC is developing standards for Smart Cities, it has to be noted the most attention is being dedicated to Energy data, while Water data, at least in the urban context, is somehow lagging behind. The above mentioned ITU report “ICT as an enabler for Smart Water management” mentions the OGC Sensor Web, but it fails to mention the emerging standard for Water Observation Data of the OGC WaterML [Taylor2013] which was not yet released when that report was issued in 2010. The ITU report nevertheless mentions the Water Information Research and Development Alliance, lead by the Bureau of Meteorology of the Australian Government, which was one of the leading organizations driving the new WaterML standard.

A fundamental document which paved the way to the creation of the WaterML standard was the OGC discussion paper on the harmonization of water observation data [OGC2010]. The main achievement of that document was the integration of a number of national standards developed by various agencies (the above mentioned Australian BoM, the US EPA), an initial version developed by CUHASI, a consortium of US universities, and de facto standards into a unique framework. The result was WaterML, which is a technical standard and information model used to represent hydrological time series structures.

WaterML is grounded on existing OGC standards, especially from Observations & Measurements (O&M) and the Geography and Markup Language. WaterML allows to define hydrological time series by denoting:

1 the observed phenomenon;

2 the spatial and temporal context;

3 which procedure was used in generating the time series (e.g. raw data from a sensor);

4 result-specific metadata, such as time series qualifiers, interpolation types, comments, quality codes etc.;

5 information on Monitoring points;

While the WaterML standard could be applied also to urban water contexts, most applications are for natural water bodies, such as hydrological measurements from monitoring stations. It is expected that, as the Smart Cities standards evolve, the use of WaterMl will become more widespread also in the urban context.

3.2.3 Semantics for water data

Semantics allow machine to interpret the meaning of information. The enrichment of data with their semantic information is typically made by using ontologies, which are explicit representation of knowledge, typically expressed using an ontology language. Among the multiple languages that are available, OWL is widespread (Ontology Web Language) being a W3C standard and also being usable on the web.

While ontologies provide formal naming and definition of the types, properties, and relationships of the entities in a domain, simple representations such as thesauri, controlled vocabularies are also used such as DCAT[8] (Data Catalog Dictionary).

There is not much done in the area of semantics for water data. The Water Data Center of CUHASI (Universities Allied for Water Research) maintains two semantic systems for the CUAHSI Hydrologic Information Center. The first is a number of controlled vocabularies[9] that are used to populate fields in the Observations Data Model. The second semantics system is an ontology[10] for properties measured that facilitates keyword search.

The W3C has recently launched the semantic water interoperability model community group[11] but no reports have been published yet, as it was recently founded in March 2014.

3.2.4 User data and social interaction

SmartH2O does not only represent data about smart grids, location and water consumption, but also about the users that consume water and participate to the gamified applications.

Also in the sector of user modelling and data exchange, the need arises for more interoperable user models; this need has been already identified in the recent work [Carm2011, which focuses on context adaptation of applications, that is to the adaptation of the content and functionality of the interface to the current profile, preferences, and even activity of the user.

This broad survey identifies characteristic dimensions, such as the users' profile, task features, social relations, and the context.

Another survey [Sosn10] overviews the ontological technologies for user modelling for the purpose to aggregate and align many RDF/OWL user models in a decentralized approach.

A great amount of available user and social representation formats have been proposed lately, which calls for a conceptualization effort aimed at providing a unified model.

Figure 6 adapted from [Kaps11] shows the current panorama of social networks most frequently used, with their essential social interaction features, and the most notable efforts in the definition of formats and models for the representation and exchange of people social data.

---

8 http://www.w3.org/TR/vocab-dcat/

9 http://his.cuahsi.org/mastercvreg/cv11.aspx

10 https://www.cuahsi.org/Ontology

11 https://www.w3.org/community/swim/

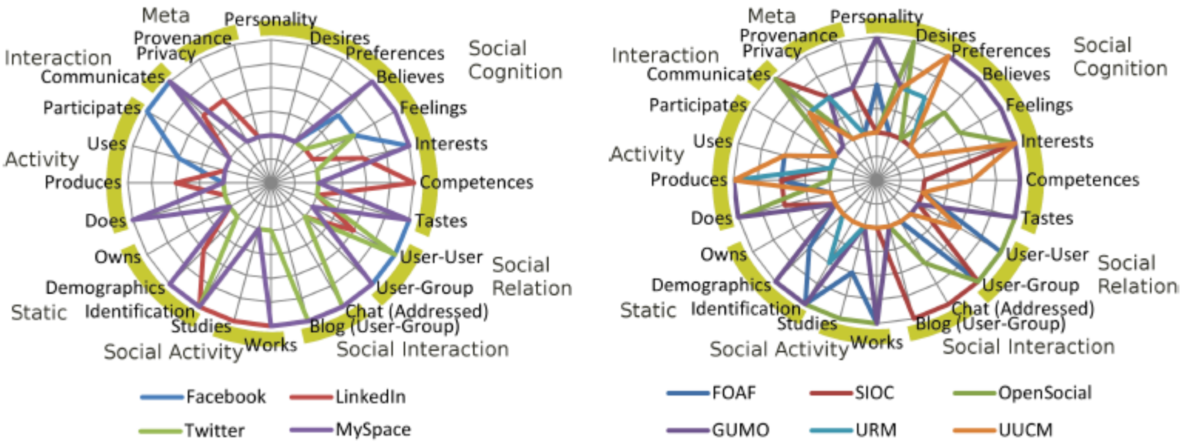


Figure 6: Comparison of social networks (left) and user models (right) (from [Kaps11])

In the following, we survey the most notable proposals for the representation of data about users, their interactions and social context. These proposals have been considered during the design of the user profile and social link representation part of the SmartH2O data model, documented in D3.1 Databases for user information.

The **General User Model Ontology** (GUMO) [Heck05] [Heck07] is a uniform interpretation of distributed user profiles models in intelligent environments based on OWL. However, it does not support the modelling of social relationships.

The **Cultural User Modeling Ontology** (CUMO) [Rein07] is used to allocate culture dimension to a certain user behaviour. Building user models from cross-system platforms is another tentative to aggregate the distributed profiles for the same user in case of obsolete or scattered data.

The **Unified User Context Model** (UUCM) [Meht05] advocates the reuse of user profiles in different systems as context passport (e.g., social networks platforms, on line communities, etc.) and minimize the sparse or missing information.

User Role Model (URM) [Zhan06] is another example for modelling users and their access roles to support cross-system personalization. While this model covers social relationships of users, it does not capture profile information. However, the great amount of available user and social representation formats calls for an additional conceptualization effort aimed at providing a unified model that integrates also social features [Carm07]

The **Friend of a Friend** (FOAF[12]) is a project that aims to describe users, their links and activities within Web pages using RDF/XML. However, it lacks the expressive power to represent social interactions (commenting, etc.), social cognition and meta information.

The **OpenSocial API**[13] is another recent effort with the goal of providing a common set of APIs to access major social net-working applications (such as Myspace, Orkut, Yahoo!, etc.). OpenSocial provides a very abstract and generic data model, so to be able to fit information coming from several social networks.

Semantically-Interlinked Online Communities (SIOC[14]), is an RDF/OWL representation to link and describe online community sites (e.g., message boards, wikis, weblogs, etc.). SIOC documents may use other existing ontologies to enrich the information described such as Dublin Core metadata, FOAF, RSS1.0, etc. The literature about recommender systems also proposed user modelling on the social web as a tool for recommendation

12 http://www.foaf-project.org

13 http://www.opensocial.org

14 http://rdfs.org/sioc/spec

strategies based on interests and topical affinity.

The **ImREAL** project[15] or **Grapple**[16] are such examples. This latter [Aroy10] appears very suited to the representation of user profiles and can be freely extended (RDF-based), but does not cover social relationships and social actions.

TheHiddenU[17] project is an attempt in the direction of social user model unification, as it compares several models in terms of their coverage of different user and social concepts. For semantic enrichment and mining of users profiles from the social Web, some frameworks are relevant, such as **USem** [Abel11].

The **U2M**[18] project is an example of a user model and context ontology integrating GUMO with UserML for future Web 2.0 extensions.

TweetUM[19] is a Twitter-based User Modeling Framework which allows developers to acquire Twitter-based profiles in RDF format. It creates semantically enriched user pro- files inferred from users' tweets, retrieving all topics that the user is interested in, the entities information, the top hash tags or top entities cloud.

GeniUS[20] is another project that can be used to enrich the semantics of social data: given a stream of messages from the users' status, it can generate the topic and some information about the user.

The **Social Web User Model** (SWUM) [Cena11] is a generic model that focuses on the social web domain; its goal is to support data sharing and aggregation of social networks user models by aligning concepts, attributes (e.g., personal characteristics, interests and preferences, needs and goals, knowledge and background), user behaviours (e.g., previous search behaviour using Google dashboard), location, and time. User's category or community can be inferred but there is no tracking of the user's actions.

The **Cubrik** data model [Kara13] extends the user data representation models with information specific for user's activities in gamified applications, and is the main base of the SmartH2O representation of water users' data. It extends previous models with a notion of user's action, which reflects any user's activity that the system is aware of, and of reward, which is a set of possible engagement measures that the system computes in order to acknowledge the actions of the user, both at the individual and at the social level.

Actions are employed to record both gaming actions and crowd tasks, the former being related to the participation to the playing of serious, persuasive and educative games, the latter describing the execution of data collection, enrichment, or validation actions by the user.

Table 2 summarizes and compares the reviewed social and user models in the light of their main characteristics.

---

15 http://www.imreal-project.eu

16 http://www.grapple-project.org

17 http://social-nexus.net

18 http://www.u2m.org

19 http://www.wis.ewi.tudelft.nl/tweetum/

20 http://www.wis.ewi.tudelft.nl/genius/

Table 2. Comparison of the features of the most relevant user data representation formats.

User Model Dimension	GUMO	URM	UUCM	FOAF	OpenSocial	SIOC	GUMF	HiddenU	SWUM	CUbRIK
Demographics	Yes	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Preferences	Yes	No	Yes	No	Yes	No	Yes	Yes	Yes	No
Context-awareness	Yes	No	Yes	No	No	No	Yes	Yes	Yes	No
Skills, Interests	Yes	Yes	Yes	Yes	Yes	No	Yes	Yes	Yes	No
Social Interaction	No	Yes	No	Yes	Yes	Yes	No	Yes	Yes	Yes
Activity, Task History	Yes	Yes	Yes	No	Yes	Yes	No	Yes	Yes	Yes
Game History	No	No	No	Yes	No	No	No	Yes	Yes	Yes
User Behaviour	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
User Network Role	No	Yes	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Resources Capabilities	Yes	No	No	No	Yes	Yes	Yes	Yes	Yes	Yes
Other Cognitive Patterns	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Topic Affinity	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
GeoCultural Affinity	Yes	No	Yes	No	Yes	No	Yes	No	Yes	Yes
User Category	No	Yes	No	No	No	No	No	No	Yes	Yes
User Action	No	No	No	No	Yes	No	No	No	No	Yes
Conflicts	No	No	No	No	No	No	No	No	No	Yes
Multiple Platforms	No	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

The Models and Technologies overviewed in Table 2 comprise both academic proposals and representation formats used in industrial projects. They have been assessed to form the basis of the SmartH2O models. SmartH2O has the need to model users, the context in which they operate, their actions, both in business scenarios and gaming ones, and thus its data model is a principled fusion of features from multiple existing models, which have addressed, individually, these domains.

3.3 Standards for software interoperability

The SmartH2O platform is composed by an assembly of heterogeneous applications and tools as they address different types of processes and data:

- Raw metered consumption.
- Psychographic user data.
- Gamified actions and rewards.
- User modelling and Agent Based Modelling systems.
- Data Analytics applications.
- External user portals.

Due to different nature of its components, interoperability between components and interoperability with external systems were selected to be guiding design rules for SmartH2O platform.

To ensure maximum of interoperability between platform's components, open standards for software interoperability were used or are evaluated for upcoming developments:

- Service Oriented Architecture was fully adopted by all platform components.

- REST (Representational State Transfer) and WSDL (Web Services Description Language) based interfaces were implemented by all platform components to communicate with other components.
- The OpenMI standard is evaluated as a possible interoperability standard for modelling components.

3.3.1 Web services

A web service is the interface of a software function, which is made available on the web.

The service oriented architecture (SOA) is a design pattern that guides the development of complex applications based on the integrated interaction of web services.

The SOA proposes to organise services across five layers:

1 Consumer Interface Layer: GUI for end users or apps accessing apps/service interfaces.

2 Business Process Layer: services representing business use-cases in terms of applications.

3 Services: they are consolidated together for whole enterprise in service inventory.

4 Service Components: they are used to build the services, like functional and technical libraries, technological interfaces etc.

5 Operational Systems: the data models, enterprise data repository, technological platforms etc.

The Web Services Description Language (WSDL) is an XML-based interface definition language that is used for describing the functionality offered by a web service. It is similar to the concept of signature of a procedure or a function in a programming language.

The Representation State Transfer (REST) is a software architecture that delivers guidelines and best practices for building web applications based on web services that display the properties of being scalable, performing and simpler to maintain. REST provides a simpler alterative to WSDL as it exploits the Hypertext Transfer protocol (HTTP) using the same commands (GET, POST, PUT, etc.). Clearly this imposes some limitations, and thus REST and WSDL can often coexist.

In SmartH2O the SOA guidelines are being adopted in the design and implementation of the software architecture. Both WSDL and REST web services are being employed.

3.3.2 OpenMI

OpenMI[21] (Open Modelling Interface) is an open standard that defines an interface that allows models from different vendors to exchange data at run-time.

OpenMI is the result of Water Framework Directive, co-funding of European Commission and support from the leaders of attempts enabled by FP5 project HarmonIT led by Centre for Ecology and Hydrology. This first version was at most a research outcome that was further developed into a standard for operational practice under a second project under the “Sustainable management of ground water and surface water management" of the European Commission’s LIFE Environment programme.

The OpenMI Environment facilitates a set of tools that can help make new and existing model code Open-MI compliant and help combine Open-MI compliant components into integrated modeling system. The interaction of different models may produce better models for each actor of the interaction.

The OpenMI interface has three functions:

- *Model definition*: To allow other linkable components to find out what items this model can exchange in terms of quantities simulated and the locations at which the quantities are simulated.

---

21 http://www.openmi.org

- *Configuration*: To define what will be exchanged when two models have been linked for a specific purpose.
- *Run-time operation*: To enable the model to consume or provide data at run time.

The general interoperability scenario of OpenMI compliant applications is depicted in Figure 7

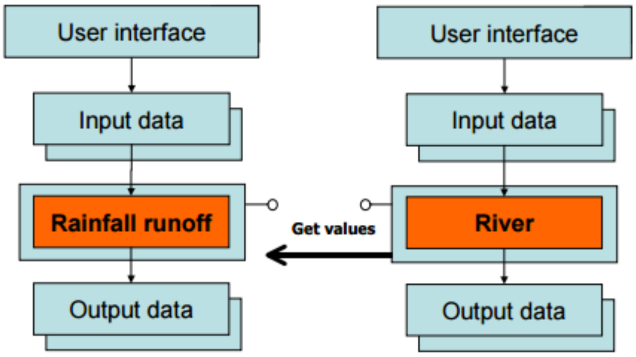


Figure 7. OpenMI compliant modeling applications interaction (from Moore2010).

In the context of SmartH2O project, The Agent Based Modelling component can make use of Open-MI to integrate different user behavioural models like:

- Models of user consumption based on previous consumption.
- Models of user consumption based on social stimuli.
- Models of user consumption based on pricing.

3.4 Common practices for serious games and games with a purpose

It is not possible to talk about "standards" when talking of Serious Games or Games with a Purpose (GWAP), but for sure there are common elements, architectural decisions, patterns, which are slowly emerging to provide guidance and reference points to developers of such solutions. In this section we briefly review the main distinctive features of gamification approaches as describe in [Galli2015]. In particular we describe the game components, we define serious games, games with a purpose and gamification techniques. Finally, common gamification mechanics solutions and methods for the impact evaluation are presented. This field is too novel to have assessed standards, but we expect that some of the concepts presented in this section will evolve towards a more formal representation, following what has happened for the field of software design.

3.4.1 The game components

The main components of a game are players, objectives, rules, conflicts, resources and outcomes.

Players are voluntary participants that accept the rules and constraint of a game and strive to employ worse rather than better means to reach a particular objective..

Objectives are the specific goals that its players have to reach when participating in the experience. In games, the objective is a key element without which the experience loses much of its structure, and our need to work toward the objective is a measure of our involvement in the game.

Rules limit player behaviour and proscribe reactive events. Rules are fundamental pieces of any gameful experience that define allowable actions by the users and consequential reaction from the system; they may be used also to define game objects and concepts that could be used to reach a particular objective in the game.

The rules and the admissible actions that can be performed in a game tend to deter players from accomplishing goals directly. This particular challenge for the players is one of the distinctive element of games: **conflict**, which the players work to resolve in their own favor.

Resources are particular objects or elements that hold a high value because they can help the players in reaching their objectives, but they are usually scarce in order to pose an additional challenge.

The last element that most of the games have in common is that for all their rules and constraints, the **outcome** of both experiences is uncertain, though there is the certainty of a measurable and unequal outcome of some kind (e.g. a winner, a loser).

3.4.2 Serious games and games with a purpose

The massive amount of time that people spend in online gaming is being more and more exploited by developing games that trascend pure entertainment purposes. According to the Serious Games initiative[22] which started in 2002, games that have these features are considered to be part of a particular category called Serious Games. Serious Games refer to "applications of interactive technology that extend far beyond the traditional videogame market, including: training, policy exploration, analytics, visualization, simulation, education and health and therapy".

The application of gaming technology, process, and design to the solution of tasks that are relatively easy to complete by humans but computationally rather infeasible to solve has given birth to a special subgenre of Serious Games, called Games with a Purpose (GWAP) [vonAhn2008]. These games tyypically rely on the following patterns:

- **Intuitive decisions** tasks are related to combinatorial optimization tasks like packing problem that are known to be NP-hard. It has been proven [Corney2010] that humans are able to solve by intuition even complex tasks belonging to this category of problems, thus human computation can be employed as a mean to use mental abilities of the participants to find solutions or algorithms able to solve combinatorial problems disguised as puzzles.
- **Aesthetic Judgment**: The design and implementation of computational systems capable of having human-level perception and understanding of aesthetics, like the quality of an image, a piece of music or the proportions in a picture is still an unsolved
- **Contextual Reasoning**: Most cognitive processes are contextual in the sense that they depend on the environment, or context, inside which they are carried on. Human Reasoning tasks that are not feasible for the machines often involve semantic understand- ing and abstraction capabilities typical of human beings. Examples for the application domain of contextual reasoning regards tasks such as image and audio annotations.

3.4.3 Gamification

There are numerous opinions as to what Gamification exactly is. One of the most general and used definition is:

"**Gamification** is the use of **game elements** and **game-design techniques** in **non-gaming contexts**." [Werbach2012]]

Gamified applications use **elements of games** that do not spawn entire games. Of course, the boundary between games and applications with game elements can be very blurry, because often this boundary is personal, subjective and social. Self representation with avatars, three-dimensional environments, narrative context, feedback, reputations, ranks, levels, marketplaces and economies, competition under rules that are explicit and enforced,

22 http://www.seriousgamessummit.com

team parallel communication systems that can be easily configured, and time pres- sure are all game elements.

It is not so simple to decide which game elements to put, where and how to make a successful gamified experience and where the game-design techniques should be included. Thus **game design** is a mixture of science and art and a lot of analysis of successful past experiences is required to accomplish satisfactory results.

There are three particular **non-game contexts**: internal, external and behavior change.

- **Internal game context**: companies use gamification to improve productivity within the organization in order to foster innovation, enhance camaraderie or to otherwise derive positive business results through their own employees. Internal gamification is something called “enterprise gamification” and in this case there are two distinguishing scenarios. In the first one, the players are already part of a defined community, the company knows them and how they interact with each other on a regular basis. The scenario derives from the first but motivational dynamics of gamification must interact with the firm’s existing management and reward structures. Internal gamification can work for core job requirements, but it is not always applicable, thus there must be some novel motivations.
- **External gamification** involves customers or prospective customers and such applications are generally driven by marketing objectives. Gamification in this case, attempts to improve the relationships between businesses and customer, producing increased engagement, identification with the product, stronger loyalty and ultimately higher revenues.
- **Behavior-change gamification** aims at creating new beneficial habits among the population; this may range from encouraging people to live in a healthier way, to study more, to maintain a sustainable and eco-friendly behavior and so on. Behavior change programs are often run or sponsored by nonprofits and governments, but they can also produce results that are beneficial even for private institutions.

3.4.4 Common gamification elements and gamification mechanics solutions

When creating a Gamified system, the objective is the starting point, it can be employee engagement, it may be increasing sales of a product, or a reduction in water or energy consumption. The goal is to meet that objective.

The next step is to start to design the gamified system. First and foremost in many systems will be the metrics to be applied to evaluate the objective. The metrics are what will allow you to know if you are on target to meet the objective or not. In the case of a water consumption awareness system such as SmartH2O it could be the litres of water per day.

Then the designer must consider what **gamification elements and mechanics** will best help you achieve the goal and start to put them into your system. In this section we focus on the most common gamification elements and mechanics that are encountered in Gamified applications.

Points.

Points or Player scores are numerical values that represent a measure of the skill of a player. We often see points used to encourage people to do things by collecting them. The assumption is that people will buy more widgets or will work harder in exchange for points. This is a simple approach that occasionally works to motivate those people who like collecting things

Points are used in gamification for a number of reasons:

1. Keeping score
2. Determine the winning state when it is necessary to define a winning condition for the players, points can be an easy mean to achieve the result.
3. Connecting progression and rewards.
4. Provide clear feedback

5. Provide quantifiable data

Bear in mind that points are very limited. They are uniform, abstract, interchangeable, and well, pointlike. To put it another way, a point is a point. Each additional point simply indicates a greater magnitude, and nothing more. This is one reason why badges are often found in conjunction with points systems.

Achievement and badges.

An Achievement is a set of tasks, defined by a designer, for the player to fulfill so to achieve a milestone and progress in the game. A Badge is an artifact associated to the completion of an achievement and given to a player after completing the achievement, or, in gaming terms, after “unlocking the achievement”.

Badges are a chunkier version of points. A badge is a visual representation of some achievement within the gamified process. Some badges simply demar- cate a certain level of points. Fitbit is a gamified system that allows people to use a wireless pedometer to track the number of miles they walk or run. The system displays a badge when the user exceeds certain point thresholds, such as 50 miles in a week or 10,000 steps in a day.

Other badges signify different kinds of activities. Foursquare, a service that engages users with local businesses by encouraging them to check in to a location with their cellphone, has numerous badges for all manner of achievements. Users unlock the Adventurer badge as soon as they check into ten places registered with the Foursquare system, and they receive the Crunked badge for checking into four bars in one night. (No one said that badges need to be socially responsible.)

Achievements are now so popular in the gaming culture that the reasons for which they have been introduced are often overlooked; however to make a reward system effective, it is necessary to keep in mind the purpose for which they have been developed.

Leaderboards

A Leaderboard is an ordered list of players, with the scores they have obtained in a specific game, It can be considered as the early ancestor of the achievement concept. Leaderboards are problematic gamification elements: On one hand, players often want to know where they stand relative to their peers.

A leaderboard gives context to progression in a way the points or badges cant. If performance in the game matters, the leaderboard makes that per-formance public for all to see. In the right situation, leaderboards can be powerful motivators. Knowing that its just a few more points to move up a slot or even to emerge on top can be a strong push for users. On the other hand, leaderboards can be powerfully demotivating. If you see exactly how far you are behind the top players, it can cause you to check out and stop trying.

Leaderboards can also reduce the richness of a game to a zero-sum struggle for supremacy, which inherently turns off some people and makes others behave in less desirable ways. Several studies have shown that introducing a leaderboard alone in a business environment will usually reduce performance rather than enhance it. There are various ways to make leaderboards work for your gamified system. A leaderboard need not be a static scoreboard, and it need not only track one attribute.

3.4.5 Techniques for evaluation and comparison

After the definition of the game mechanics that can be introduced in a gami- fied application, another fundamental aspect is the definition of what are the metrics by means of which we can say if the Gamification approach contributed positively or not.

While the term game metric has become something of a buzzword in game development in recent years, metrics have arguably been around for as long as digital games have been made even though he application of game telemetry and game metrics to drive data-driven design and development has expanded and matured rapidly just in the past few years across the industry[140]. Game metrics start with raw telemetry data, which can be stored in various database formats, ordered in such a way that it is possible to transform the data into various interpretable measures, e.g. average completion time as a function of individual game levels or revenue per day.

The game metrics used more often are:

- ARPU: Average revenue per user, the total revenue divided by the number of subscribers
- Churn: the turnover rate (or attrition rate) of social games active players
- Retention: It can be considered as the opposite of churn. Retention is how well you maintain your user base.
- DAU: Daily Active Users and MAU Monthly Active Users
- Cohort: a cohort is a group of subjects who have shared a particular event together during a particular time span, In social gaming metrics, cohorts are used for analyzing retention
- Engagement: The term engagement, in a business sense, indicates the connection between a consumer and a product or service
- Re-Engagement: Gamers stop playing eventually. Re-engagement measures how many you get back
- Entry Event: An entry event is the first action a user performs when they enter the game
- Exit Event: The opposite of entry events
- Viral rate/K Factor: Measured by K-Factor, the Viral Rate shows how much your users are promoting, evangelizing and spreading the application/game
- Lifetime Network Value: it is the value a user provides to your network over the course of their entire lifetime on the network. For instance, is the user contributing to viral effects? Evangelizing the game? Contributing positively to ARPU?

We refer to [Galli2015] for a detailed description of these game metrics. Here we simply recall that the main purpose is to verify and assess the imp[act of the game on the collective achievement of the Gamified application objectives.

4. Open Data

In this section we introduce the Open Data concept of SmartH2O. We have already highlighted in Section 2 how data on water consumption in major urban cities across Europe is inconsistently accessible. High resolution data sets, which also detail the water end use per fixture type, are even harder to obtain and often they can only be acquired for a fee.

With the advent of smart metering systems we hope that this situation is going to change, but for this it is required that data owners, that is the citizens represented by their governments, will adhere to the principles of open data.

4.1 What are open data

Open data are defined by the Open definition[23]:

"Open means anyone can freely access, use, modify, and share for any purpose (subject, at most, to requirements that preserve provenance and openness)."

The full definition contains more details, but the highlights are:

1. **Availability and Access:** the data must be available as a whole and at no more than a reasonable reproduction cost, preferably by downloading over the internet. The data must also be available in a convenient and modifiable form.
2. **Reuse and Redistribution:** the data must be provided under terms that permit reuse and redistribution including the intermixing with other datasets.
3. **Universal Participation:** everyone must be able to use, reuse and redistribute - there should be no discrimination against fields of endeavour or against persons or groups. For example, 'non-commercial' restrictions that would prevent 'commercial' use, or restrictions of use for certain purposes (e.g. only in education), are not allowed.

The term "Open Data" then describes the process of how scientific data may be re-used and published without price and permission barriers [Rust2008], where "availability and access", "reuse and redistribution" and "universal participation" are playing a dominant role. This new open way of serving data, integrates the public that they are becoming part of it and may enrich data, process it, or combine it with different sources [Janssen2012]. The traditional boundaries of organizations are vanishing and actually everybody has access to those data and can use it for self-purposes. If the government or other organizations open this more or less closed model, the consequences can be twofold. On the one hand side, the control of the data will be lost and public managers fear different scenarios: being confronted with a variety of stakeholders which may help to improve the data [Janssen2012], but also seen as a threat, if specific goals are not reached, which is also connected to public responsibility and accountability. Another cause of transparent data is the expectancy of the society to intervene, which can lead to common dissatisfaction. On the other hand the opening of data creates benefits like social innovation, economic growth and trust [Janssen2012]. One of the most promising and tremendous resource in the open data domain is open government data, which is largely untapped. A distinct number of areas, which are already adding value are following: "Transparency and Control", "Participation", "Self – Empowerment", "Improved or new private products and services", "Innovation", "Improved efficiency of government services", "impact measurement of policies" and "New knowledge from combined data sources and patterns in large data volumes". While in those areas open data is already creating social and economical value, there are a lot of new unknown areas where the combination of data can create new knowledge and give new insights, which has not yet been discovered.

---

23 http://opendefinition.org

4.2 Open Data for urban water management

In this section the focus lies on the available open water data resources in general as well on tools and methods, which support the water management and decision making process. A blog post on the *futureearth* website[24] mentions that collecting and analyzing data on water can help to connect researchers and governments with individual water users and will raise the awareness of water management challenges. This is just possible if the available data are in the form that supports their further processing.

4.2.1 Open water data Europe

The European Union (EU) has developed a platform[25] where various datasets are available. The records are categorized by subject (e.g. economics, trade, finance, environment and so on) and can be located through an integrated search function. The search term "urban water" delivers ten search results including "urban waste water treatment directive" as the most relevant one. Another resource[26] from the European environment agency provides few datasets about urban waste water. From the UK different datasets are available through an interactive map search[27], where the region of interest can be marked and submitted. The datasets focus more on resource availability, quality and water frameworks directives. The available resources of open data are often dependent on the city and their government support of open data initiatives.

Table 3 lists available open data water resources by description, application area, and available format with a focus on Europe.

Table 3. Examples of open data water resources

Description	URL	Available Format
Status of bathing water. Popular bathing places are monitored for indicators of microbiological pollution Further search results with the search term "water data": - water productivity - waterbase, water quantity - water exploitation index - water use balance - water made available by use	http://open-data.europa.eu/en/data/dataset/e83MrtBw7tGcDVOxVStPSg http://open-data.europa.eu/en/data/dataset?q=water+data&op=&ext_boolean=all&page=2	XLS
Water use in the manufacturing industry by activity and supply category / Europe	http://appsso.eurostat.ec.europa.eu/nui/show.do?dataset=env_wat_ind&lang=en	XLS, CSV, HTML, PC Axis, SPSS, TSV, PDF
Annual freshwater abstraction by source and sector / Europe	http://appsso.eurostat.ec.europa.eu/nui/submitViewTableAction.do	XLS, CSV, HTML, PC Axis, SPSS,

24 http://www.futureearth.org/blog/2014-aug-27/virtual-flood-information-open-data-sustainable-water-management

25 http://open-data.europa.eu/en/data/

26 http://www.eea.europa.eu/data-and-maps/data/waterbase-uwwtd-urban-waste-water-treatment-directive-4

27 http://data.gov.uk/data/search?q=water&ext_bbox=-0.53,51.66,-0.36,51.73

		TSV, PDF
Water use by supply category and economical sector / Europe	http://appsso.eurostat.ec.europa.eu/nui/show.do?dataset=env_wat_cat&lang=en	XLS, CSV, HTML, PC Axis, SPSS, TSV, PDF
Generation and discharge of wastewater in volume / Europe	http://appsso.eurostat.ec.europa.eu/nui/show.do?dataset=env_ww_genv&lang=en	XLS, CSV, HTML, PC Axis, SPSS, TSV, PDF
Renewable freshwater resources / Europe	http://appsso.eurostat.ec.europa.eu/nui/show.do?dataset=env_wat_res&lang=en	XLS, CSV, HTML, PC Axis, SPSS, TSV, PDF
Population connected to wastewater treatment plants Europe / Europe	http://appsso.eurostat.ec.europa.eu/nui/show.do?dataset=env_ww_con&lang=en	XLS, CSV, HTML, PC Axis, SPSS, TSV, PDF
Complete Statistical Overview / Europe	http://ec.europa.eu/eurostat/web/environment/water/main-tables	Zip Package
Urban Waste Water Treatment Directive / Europe	http://open-data.europa.eu/en/data/dataset/ZKWId3G3YMye9AzEUxwmog	TXT, XLS, CSV, Microsoft Access
Bathing Water Data England	http://data.gov.uk/dataset/bathing-water	CSV, HTML, JSON, RDF, SPARQL, RDF
Bristol River Quality / England	http://data.gov.uk/dataset/bristol-river-water-quality	CSV
DFID Live data page for energy and water consumption / England	http://data.gov.uk/dataset/dfid-energy-and-water-consumption	CVS, HTML

4.2.2 Open water data management tools

The European research project “Urban Water”[28] develops an innovative ICT-based platform for efficient and integrated water management of urban water resources. The platform makes use of weather prediction, water reserves (e.g. reservoirs) data, household consumption data, water distribution data and statistics from other sources. The online tools listed in Table 5 focus on the region Europe.

Table 4. Examples of open data water management tools.

Description	URL
Tool to visualize the values inside of the drinking water for predefined places / Germany	http://blog.opendatalab.de/opendata/2014/03/25/trinkwasser/
Through an interactive map,	http://data.gov.uk/data/map-based-search

28 http://urbanwater-ict.eu

environmental data (including water data) can be explored / England	
The leaf water management tool helps farmers to manage their water resources, assess risks and find new ways to improve efficiency / England	http://www.leafuk.org/leaf/farmers/watermanagementtool.eb
The GWT provides a free publicly resource to identify cooperate water risks	http://www.wbcsd.org/work-program/sector-projects/water/global-water-tool.aspx

Recently the non-profit social enterprise "Open Water Foundation" (OWF) (http://openwaterfoundation.org/) came up with a collection of tools, which provide data management to a certain level and focus on the region North America.

The OWF tools support the users in questions about water supply and quality for the environment, cities, agriculture and industry. One application case is the use of TSTools[29], which access time series, tables and spatial data in databases, files and web services. OWF is working on a decision support system to understand, analyze and model complicated water systems. The model involves water laws, hydrologic and geographic variability, surface and groundwater, operations and a lot more other factors. The state of Colorado provides databases, software tools and baseline data sets. OWF is trying to build a platform of tools for water resources and extends the current CDSS (Colorado's decision support system), which can be found online[30,31].

CDSS is a water management system with online tools, which provide water data from various categories.

The list of tools reported in Table 5 illustrates examples of available water management tools and their functionalities.

Table 5: examples of water data management tools

Name	Description
Groundwater (Water Levels) / Colorado	Groundwater (Water Levels) provides detailed well information and related groundwater parameter (Surface water evaluation, Location data, Depth, Permit dates) http://cdss.state.co.us/onlineTools/Documents/Data_Dictionary_GroundwaterLevels.pdf
Streamflow Stations / Colorado	Provides historical streamflow data for specific gages, either monthly volume or daily flow rate http://cdss.state.co.us/onlineTools/Documents/Data_Dictionary_Stations.pdf
Water Rights / Colorado	This tool supports a detailed water rights information for water structures such as ditch diversions, reservoirs and wells http://cdss.state.co.us/onlineTools/Documents/Data_Dictionary_WaterRights.pdf
Aquifer Determination Tools / Colorado	The aquifer tools can be used to research aquifer characteristics for the Denver or Dakota/Cheyenne aquifers to determine the volume of water at a specific location

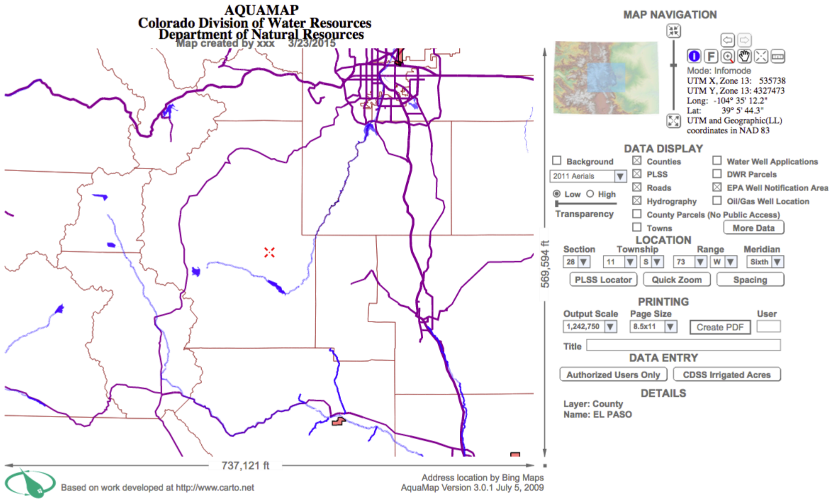
29 http://openwaterfoundation.org/software-tools/tstool

30 http://openwaterfoundation.org/software-tools/colorados-decision-support-systems

31 https://sites.google.com/site/cdssstaging/

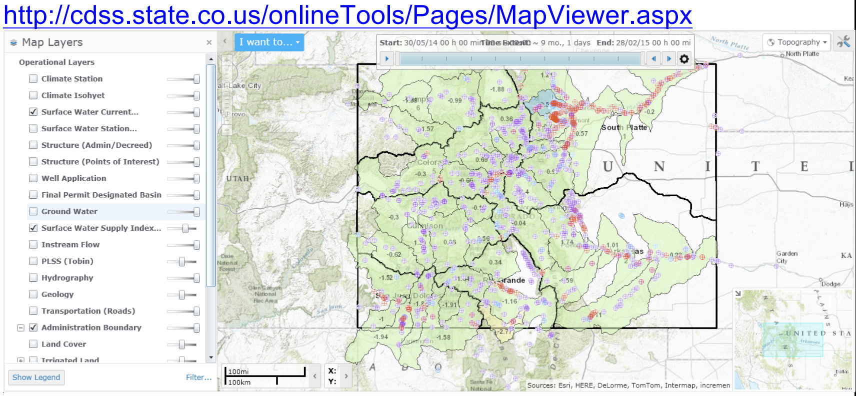
<table>
<tr><td></td><td>http://cdss.state.co.us/onlineTools/Pages/AquiferDeterminationTools.aspx</td></tr>
<tr><td>AquaMap / Colorado</td><td>Is a web-based interactive map for viewing stream, structure, aquifer and well data

http://water.state.co.us/DataMaps/GISandMaps/AquaMap/Pages/default.aspx



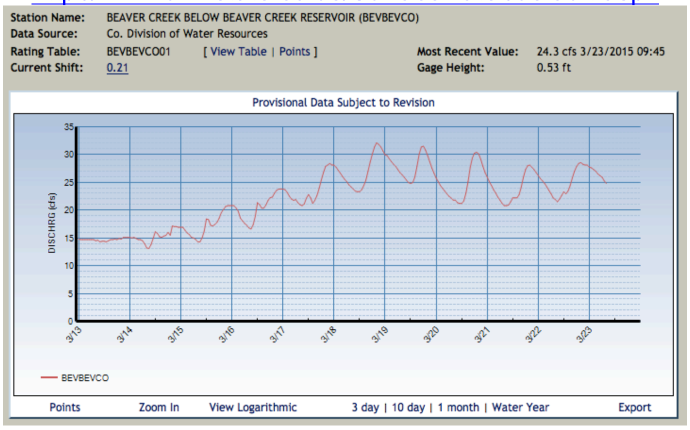
Aqua Map

Different data can be displayed: Counties, PLSS, Roads, Hydorgraphy, County Parcels, Towns, Water Well Applications, DWR Parcels, EPA Well Notification Area and Oil/Gas Well Location</td></tr>
<tr><td>MapViewer</td><td>The MapViewer is a web-based GIS application that provides water feature information in the form of an interactive map and offers various functionalities: annotation functionality, import and export of shapefiles and a limited analysis can be done

http://cdss.state.co.us/onlineTools/Pages/MapViewer.aspx


MapViewer</td></tr>
<tr><td>Real-Time Surface Water Conditions</td><td>The real-time water conditions provides real-time satellite monitoring data from gaging stations maintained by DWR and US Geological Survey. The common use is to find current streamflow for a specific stream or ditch, download provisional flow data for the current water year or find station description information for specific gages. Below an example of visualization of real time surface water conditions of the “Beaver Creek Reservoir” station.</td></tr>
</table>

	http://www.dwr.state.co.us/Surfacewater/default.aspx



Real-Time Surface Water Conditions

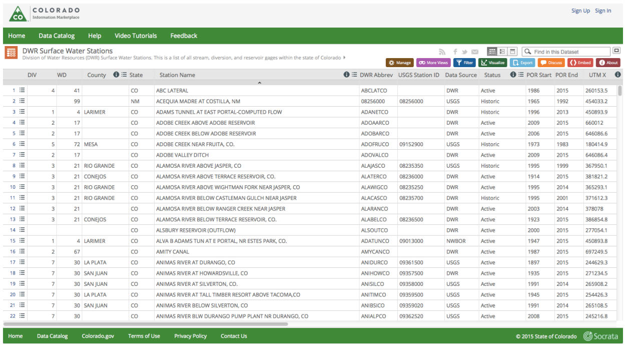
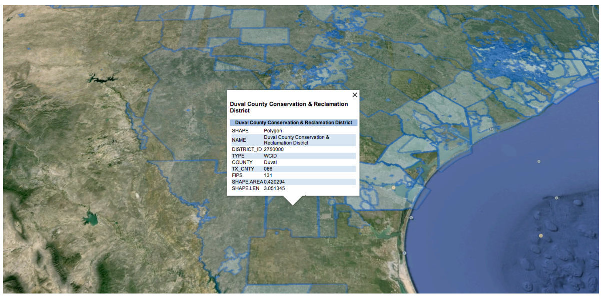
4.2.3 Open water data North America / Canada / Australia

A quite large number of States and / or cities in America, Canada and Australia are offering an open data platform, respectively a catalogue where open data can be accessed.

Table 6 reviews the main efforts in those countries.

Table 6: examples of open water data management and publication efforts in US, Canada and Australia

Description	URL	Available Format
United States federal agencies & national portals; 38,130 datasets were found for the keyword “water” and 366 for “water usage” which includes residential water usage (e.g. Los Angeles)	http://catalog.data.gov/	Mostly CSV, RDF, JSON, XML, HTML
The open data portal of the state of California (http://www.waterboards.ca.gov/) offers a wide range of information about their water resources: • There are maps and websites about their regional water quality • Water boards and data maps (e.g. drinking water supply service area lookup tool, “is my property near a nitrate-impacted water well?”, water quality control board and so on) • Drought information • Through the California environmental data exchange network (http://www.ceden.org/) open water data from the categories “water quality, toxicity, tissue, benthic, habitat” can be accessed • This portal provide maps and focus on monitoring the water quality (e.g. “is our water safe to drink?”, “is it safe to swim in	http://www.waterboards.ca.gov/waterboards_map.shtml http://www.waterboards.ca.gov/data_maps/index.shtml http://www.waterboards.ca.gov/waterrights/water_issues/programs/drought/index.shtml http://ceden.waterboards.ca.gov/AdvancedQueryTool	

our waters?", "is it safe to eat fish and shellfish from our waters?"	http://www.mywaterquality.ca.gov/	
The state of Colorado provides an information marketplace where water datasets like "DWR water level", "DWR Surface water stations" can be accessed. There are approximately 650 water data sets available. Data from different water stations have been collected. The following image shows the Colorado Open Data Catalogue dashboard where the data can be managed (more views, filter, visualize, export, discuss, embed)  **Figure 4 Open Data Dashboard**	https://data.colorado.gov/browse?category=Water	CVS, PDF, XML, JSON, XLS
On the open data portal[32] of Oklahoma more than 2550 datasets with the keyword "water" have been found (e.g. water contaminant level, initial groundwater assessment and so on). The first few pages in the search results are representing the most relevant data for water usage.	https://data.ok.gov/browse?sortBy=relevance&utf8=✓&q=water	CSV, JSON, PDF, RDF, RSS, XLS, XLSX, XML
The open data portal of Texas also offers a wide range of water data from the categories (drinking water, water use and availability and water quality)  **Figure 5 KML File**	http://www.tceq.state.tx.us/agency/water_main.html	SHP, GDP, KML, Meta
The world bank (e.g. water productivity) offers free and open access to data about development	http://data.worldbank.org/indicator/ER.GDP.F	EXCEL, XML, CSV

32 https://data.ok.gov/

countries in and round the globe (e.g. water productivity, improved water source, annual freshwater withdrawals)	WTL.M3.KD	
The open data catalogue from Utah delivers 167 search results with the search term “water” delivers a lot of information on “water well drilled”	https://opendata.utah.gov/browse	CSV, JSON, PDF, RDF, RSS, XLS, XLSX, XML
The Australian Government provides an open data catalogue where 8 search results have been found with the search term “water” (e.g. Melbourne Water Use by Postcode)	http://data.gov.au	CSV, XLSX, docX
Also the City of Palo Alto offers data of the annual water consumption on their website, which can be accessed via an API and XLS, CVS export	http://data.cityofpaloalto.org/datastreams/77374/palo-alto-utilities-annual-water-consumption/	API, XLS, CVS

The open data catalogue of Toronto[33] offers just one dataset, which analyzes the water billing by ward but also includes the annual residential water usage (in cubic meters).

Further Canadian datasets can be found here on the Open Data Canada website[34] with a focus on flood extent and natural disaster. The city of Ottawa represents itsown open data website[35] where the ward data from national households or drinking water summaries can be accessed. A water quality data management system has been developed by Google[36], that allows users to manage water quality and monitoring locations, samples, results and projects inside a web based environment.

In general there are more open data catalogues available for the region North America and England than for the rest of Europe.

In those countries civil society organizations, media and the senior management dedicate themselves for a new data policy[37].

4.3 Open data in scientific publishing

The European Commission has been taking steps forward to move from open access (publications) in scientific publishing to “open science” (publications and data). In 2012, it encouraged member states to make publicly funded research results available to all in order to guarantee responsible “research and innovation” [EC2012].

This principle concerning scientific research and publications is illustrated in Horizon 2020, a EU Research and Innovation programme with nearly €80 billion of funding available over 7 years. This principle is implemented by compelling entities, whose projects receive funds from Horizon 2020, to insure that any peer reviewed journal article they publish is openly accessible. This principle has been extended also to research data, which must be shared with the scientific community. In order to achieve this aim, it is required that clear policies for dissemination are defined [EC2012], which shall address the following: concrete objectives

33 http://www1.toronto.ca/wps/portal/contentonly?vgnextoid=289bd103cd8b1310VgnVCM1000003dd60f89RCRD&vgnextchannel=8517e03bb8d1e310VgnVCM10000071d60f89RCRD

34 http://open.canada.ca/data/en/dataset?q=water+usage&sort=metadata_modified+desc

35 http://data.ottawa.ca/dataset/2006-and-2011-ward-data-from-the-census-and-2011-national-household-survey-nhs

36 https://code.google.com/p/open-waters/

37 http://www.bpb.de/gesellschaft/medien/opendata/64063/globale-entwicklung

and indicators to measure progress, implementation plans, financial planning. The effective implementation of these policies would result in accessible research data that can be re-used through digital e-infrastructures. When making research data accessible, a major critical aspect should concern “privacy, trade secrets, national security, legitimate commercial interests and intellectual property rights”. Furthermore, accessible datasets must be easily identifiable and information must be provided to let everybody use it and connect to previous studies. Publicly funded academic institutions must support the process of sharing research data, also through the development and training of new professional profiles endowed with technical competencies consistent with the aim of the project.

In order to implement these policies it is necessary to have a sound e-infrastructure for the dissemination of scientific data. This requires an effort from the academic institutions and funding entities to make the process of sharing and using datasets as efficient as possible. This can be achieved by defining all information specific to the dataset along its life cycle. Moreover, it is necessary to train young researchers in the field of data-analysis and also to develop new profiles of technicians and data managers. A further concern refers to the development of analysis tools and new software necessary to fully exploit the datasets once they are shared.

Currently, OpenAire acts as an e-infrastructure compliant with the principles stated by Horizon 2020, since it enables sharing both scientific papers as well as datasets, free of charge.

4.4 The Open Data Strategy and Policy of SmartH2O

According to the Open definition[38] there are four fundamental steps to be taken to open up the data:

1 ***Select the dataset(s) to open.***

2 ***Apply an open license.***

   a. Determine what intellectual property rights exist in the data.

   b. Apply a suitable ‘open’ license that licenses all of these rights and supports the definition of openness discussed in the section above on ‘What Open Data’

3 ***Make the data available*** - in bulk and in a useful format. You may also wish to consider alternative ways of making it available such as via an API.

4 ***Make it discoverable*** - post on the web and perhaps organize a central catalogue to list your open datasets.

In the SmartH2O project we commit to this procedure to release anonymised datasets of water consumption at the individual household level. We will make available as many datasets as we will be able to, as this depends eventually on the user agreement. The users taking part in the SmartH2O project will be asked to select one of the following options regarding to the level of data openness:

1. No consumption data shared with the SmartH2O project. This is an opt-out only option. By default, anonymised consumption data is made available to the SmartH2O project by the water utility. The information will be nevertheless used to provide aggregated water consumption. No individual water consumption data will be published as open data.
2. Consumption data shared only with the project but no household information revealed to anyone, not even the project partners. If the user selects this option s/he will be able only to access his/her own raw consumption data on the non-gamified SmartH2O platform. The consumption data are being anonymously shared by the

[38] http://opendefinition.org

water utility with the SmartH2O project. No individual water consumption data will be published as open data.

3. Consumption data and household information shared only with the SmartH2O project. The user fully participates to the project and uses the Gamified platform, but no individual water consumption data will be published as open data.
4. Consumption data anonymously shared, household data is kept private. Only individual household water consumption is made available as Open Data. No information on household features is revealed.
5. Full anonymous disclosure: water consumption and household features are being offered for release as Open Data.

All aggregated data will be made available by the SmartH2O project provided prior consent of the water utility.

5. Conclusions

In this document we have briefly reviewed the most relevant standards that can be applied for smart water solutions. While the development of standards supporting communication between the smart water meter and the utility information systems has already reached a satisfactorily level of maturity, especially thanks to the push of smart meter manufacturers, standards for the representation of structured and high level data related to urban water use are still evolving. Here the role of the Open Geospatial Consortium and its Smart Cities initiatives is seen as highly relevant.

On the basis of our analysis of standards for smart water solutions we have identified five recommendations as key enablers to make possible a real interoperability of data and solutions for the smart water sector.

- **Recommendation #1**. *We recommend that the whole supply chain of a smart solution should be vendor and technology neutral*.
- **Recommendation #2**. *We recommend that all smart water solutions be ranked against the interoperability scale of Tolk et al. [Tolk2006], in order to assess their suitability to integration, interoperability and composability.*
- **Recommendation #3.** *We recommend that all smart water solutions be evaluated with respect to the level of access provided to the data they produce following Berners-Lee scale [TBL2006], in order to assess their positive impact on the society and the research community*.
- **Recommendation #4**. *We recommend that data on water consumption from municipal level up to country level is collected and published in INSPIRE, so that policy makers and citizens could understand the past, present and future water needs of their regions.*
- **Recommendation #5.** *It is recommended to make all efforts to prevent the malicious use of data produced by the smart meters. The adoption of the "Data Protection by Design" principles is expected in all smart water solutions, especially by avoiding to store sensitive customer data which are not essential to the aims of the smart water solution*.

Once data are standardised, such data should be made available, where possible, for future studies and research efforts. We have therefore also made a tour d'horizont on Open Data initiatives in the water sector and we have finally outlined our strategy to publish the data that are being generated by the SmartH2O project. Our final recommendation is therefore an extension to Recommendation #4:

- **Recommendation #6**. Whereas possible, water consumption data should be made available following the principles and the protocols of Open Data initiatives.

6. References

[Abel11] Fabian Abel, Ilknur Celik, Claudia Hauf, Laura Hollink, and Geert-Jan Houben. U-sem: Semantic enrichment, user modeling and mining of usage data on the social web. *CoRR*, abs/1104.0126, 2011.

[Aroy10] Lora Aroyo and Geert-Jan Houben. User modeling and adaptive semantic web. *Semantic Web*, 1(1-2):105-110, 2010.

[Carm07] Francesca Carmagnola, Federica Cena, and Cristina Gena. User modeling in the social web. pages 745-752, 2007.

[Carm11] Francesca Carmagnola, Federica Cena, and Cristina Gena. User model interoperability: a survey. User Model. User-Adapt. Interact., 21(3):285{331, 2011.

[Cena11] Federica Cena, Antonina Dattolo, Ernesto William De Luca, Pasquale Lops, Till Plumbaum, and Julita Vassileva. Semantic adaptive social web. pages 176-180, 2011.

[Corney2010] J. R. Corney, I. Kowalska, et al. "Crowdsourcing human problem solving strategy". In: CrowdConf10 Proceedings of the 1st International Conference on Crowdsourcing. San Francisco, CA, USA. 2010.

[EC2012] COMMISSION RECOMMENDATION of 17.7.2012 on access to and preservation of scientific information.

[EC2015a] European Commission, INSPIRE: infrastructure for spatial information in the European Community. http://inspire.ec.europa.eu. Last accessed: March 2015.

[EC2015b] European Commission and European Environment Agency, The Water Information System for Europe. http://water.europa.eu. Last accessed: March 2015.

[EC2015c] European Commission, Eurostat, http://ec.europa.eu/eurostat/web/main. Last accessed: March 2015.

[EEA2015] European Environment Agency, The Water Data Centre, http://www.eea.europa.eu/themes/water/dc. Last accessed: March 2015.

[Heck05] Dominik Heckmann, Tim Schwartz, Boris Brandherm, Michael Schmitz, Margeritta von Wilamowitz-Moellendorff. Gumo – The General User Model Ontology. User Modeling 2005. Lecture Notes in Computer Science Volume 3538, 2005, pp 428-432.

[Heck07] Dominik Heckmann, Eric Schwarzkopf, Junichiro Mori, Dietmar Dengler, and Alexander Kröner. The user model and context ontology gumo revisited for future web 2.0 extensions. In *Proc. of the Int. Workshop on Contexts and Ontologies: Representation and Reasoning (C&O:RR)*, 2007.

[IEEE2015]IEEE Standards Association, Internet of Things - IEEE standards enabling products with real-world applications, http://standards.ieee.org/innovate/iot/

[ITU2010] ITU-T, ICT as an enabler for Smart Water Management, ITU-T Technology Watch Report October 2010.

[ITU2012] ITU-T Study Group 13, Overview of the Internet of things, http://handle.itu.int/11.1002/1000/11559

[Galli2015] Galli, L. (2015) Methodologies and Tools for Games with a Purpose design and Gamified Applications. PhD Thesis, Politecnico di Milano.

[Janssen2012] Janssen, M. , Charalabidis, Y., Zuiderwijk A. (2012) Benefits, Adoption Barriers and Myths of Open Data and Open Government, Information Systems Management, 29:4, 258-268, DOI: 10.1080/10580530.2012.716740

[Kara13] Roula Karam, Piero Fraternali, Alessandro Bozzon, Luca Galli: Modeling End-Users as Contributors in Human Computation Applications. MEDI 2012: 3-15

[Meht05] Bhaskar Mehta, Claudia Nieder´ Avare Stewart, Marco Degemmis, Pasquale ee, Lops, and Giovanni Semeraro. Ontologically-enriched unified user modeling for cross-system personalization. pages 119-123, 2005.

[Moore2010] Roger Moore (Editor), OpenMI Document Series: Scope for the OpenMI

(Version 2.0). The OpenMI Association.

[OGC2010] Open Geospatial Consortium. 2010. Harmonising Standards for Water Observation Data - Discussion Paper. OGC 09-124r2

[OGC2015] Open Geospatial Consortium. 2015. OGC Smart Cities Spatial Information Framework. http://www.opengis.net/doc/WP/smart-cities-sif

[OMS2014] OMS Open Metering System Specification, http://www.oms-group.org/. Lasta cessed: March 2015.

[Rein07] Katharina Reinecke, Gerald Reif, and Abraham Bernstein. Cultural user modeling with cumo: An approach to overcome the personalization bootstrapping problem. 2007.

[Rust2008] Rust, P. 2008: Open Data in Science, Serials Review, 34:1, pp. 52-64.
[Sosn10] Sergey A. Sosnovsky and Darina Dicheva. Ontological technologies for user modelling. IJMSO, 5(1):32{71, 2010.

[Taylor2013] Peter Taylor, Simon Cox, Gavin Walker, David Valentine and Paul Sheahan. 2013. WaterML2.0: development of an open standard for hydrological time-series data exchange. Journal of Hydroinformatics.

[TBL2006] Berners Lee, Tim. 2006. Linked Data.
http://www.w3.org/DesignIssues/LinkedData.html

[Tolk2006] Tolk A, Diallo SY, Turnitsa CD, Winters LS (2006) "Composable M&S Web Services for Net-centric Applications," Journal for Defense Modeling & Simulation (JDMS), Volume 3 Number 1, pp. 27-44, January 2006.

[vonAhn2008] L. Von Ahn and L. Dabbish. “Designing games with a purpose”. In: Commun. ACM 51 (2008), pp. 58–67.

[Werbach2012] K. Werbach and D. Hunter. For the Win: How Game Thinking Can Revolutionize Your Business. Wharton Digital Press, 2012. url: http://books.google.it/books?id=abg0SnK3XdMC.

[Zhan06] Fuzhi Zhang, Zhizheng Song, and He Zhang. Web service based architecture and ontology based user model for cross-system personalization. pages 849-852, 2006.