Water Analytics and Management with Real-time Linked Dataspaces

Umair ul Hassan, Souleiman Hasan, Wassim Derguech, Louise Hannon, Eoghan Clifford, Christos Kouroupertroglou, Sander Smit, Edward Curry

Abstract Due to predictions of water scarcity in the future, governments and public administrations are increasingly looking for innovative solutions to improve water governance and conservation. The problem is exasperated due to low levels of awareness about water consumption among the general public. This calls for a holistic approach to effectively manage resources during all stages of water usage. Implementation of such an approach heavily relies on advanced analytics technologies that combine data from different sources to enable decision support and public engagement. The next-generation of water information management systems must overcome significant technical challenges including integration of heterogeneous and real-time data, creation of analytical models for diverse users, and exploitation of ubiquitous devices to disseminate actionable information. This chapter presents a new approach for water analytics in public spaces that is built upon the fundamental concepts of Linked Data technologies. The chapter also presents a concrete realization of the Linked Data approach through the development of water analytics applications for buildings in public educational institutions.

Umair ul Hassan, Souleiman Hasan, Wassim Derguech, Edward Curry
Insight Centre for Data Analytics, National University of Ireland Galway, Ireland

Louise Hannon, Eoghan Clifford
College of Engineering & Informatics, National University of Ireland Galway, Ireland

Christos Kouroupertroglou
Ultra4, Greece

Sander Smit
BMChange, The Netherlands

1 Introduction

One of the sustainable development goal set out by the United Nations, as part of its agenda for 2030, is to ensure availability and sustainable management of water and sanitation for all [12]. Furthermore, recent projections by the Organization for Economic Cooperation and Development estimate that more than 40% of the world's population will be living in areas under severe water stress by 2050 [25]. This problem is expected to worsen due to a high global demand for water from manufacturing, thermal electricity generation and domestic use. Commercial uses of water are depleting the world's freshwater supply in both quantity and quality. A key factor contributing towards scarcity of water is the historical belief that water is not a vital resource that needs to be managed. Nonetheless, a recent study has highlighted the effects of water scarcity on economic growth [17]. The same study also recommends conserving water through increased efficiency in existing uses. This underlines a significant opportunity for research and development of ICT tools to raise awareness, improve management, and increase conservation of water [26].

In order to manage water holistically, it is important to use decision support tools that present meaningful and contextual information about usage, pricing, and availability of water in an intuitive and interactive way. Different users have different information requirements to manage water, from home users managing their personal water usage, business users managing the water consumption of their commercial activities, to municipalities managing regional distribution and consumption at the city level. In order to develop water information services for such diverse users, it is necessary to leverage knowledge from across a number of different domains, including metering, collection and catchment management, environmental, water quality, energy usage, utility information, end-user feedback, occupancy patterns, meteorological data, etc. However, many barriers exist to interoperability across domains and there is little interaction between these islands of information. The design of next-generation water information management systems poses significant technical challenges in terms of information management, integration of heterogeneous data, and real-time processing of dynamic data.

Linked Data technology leverages open protocols and W3C standards for sharing structured data on the Web. In this chapter, we discuss the use of Linked Data as an enabling technology for water data services. The objective of this approach is to create an integrated well-connected *Real-time Linked Dataspace* [10, 16] of information relevant to managing water in public spaces. Representing water usage data within the Linked Data format makes it open; thus, allowing it to be easily combined with data from other relevant domain silos. This chapter describes the fundamentals of the Linked Data approach for water data services [5]; in addition, it details a concrete implementation of this approach for water analytics in public spaces. Section 2 motivates the need for contextual water information management. Section 3 introduces the main concepts of the Linked Data approach. Section 4 details the architecture developed for enabling this approach, in the context of Waternomics project. Section 5 describes the pilots used for testing and validation of proposed approach. Section 6 details the water management applications designed a university building and a school. Section 7 discusses related literature and Section 8 provides a brief summary of this chapter.

2 Motivation

Sustainability requires information on the use, flows, and destinies of energy, water, and materials including waste, along with monetary information on environmental costs, earnings, and savings. This type of information is essential if we are to understand the causal relationships between the various actions that can be taken, and their impact on sustainability. However, the problem is broad in scope, and the necessary information may not be available, or difficult to collect. Within the context of water management, improving the sustainability of water consumption, especially through changing the way a household, organization, or city operates [6]. This requires a number of practical steps that will include the need for a systematic approach for information-gathering and analysis.

2.1 Contextual Water Management

One of the key problems of modern water management systems is the lack of data management and decision support tools that present meaningful and personalized information about usage, pricing, and availability of water in an intuitive and interactive way to end-users. This introduces limitations in the efforts to manage water as a resource, including:

- **User Awareness**: End-users do not have access to water information (i.e. availability, consumption, and pricing) at the moment water consumption decisions are being taken.
- **User Incentives**: Due to billing, pricing, awareness, or metering aspects, end-users may not have an incentive to change their behavior.
- **Integrated Information Provision & Analysis**: Decision makers do not have access to information platforms to make organizational changes. Personalized water information can only be created by combining publicly available water data with private water usage data that is only available to water service providers.
- **Benchmarking**: End-users do not know if their individual water consumption pattern is high or low compared to others.

2.2 Water Footprint & Water Information Ecosystems

The demand for business transparency is driving multinational companies towards more holistic assessments of their water footprint and associated impact. By understanding all the freshwater sources and uses related to a business or a product, decision-makers can identify environmentally conscious and programmatic changes to reduce their freshwater impact or footprint. Water footprint assessments are emerging concepts that require obtaining water data from many participants within an organization's supply chain. Numerous data sources can be used for this purpose, including weather data, geo-location data, historical records, product usage data, user behavior habits, etc. There is no single source to provide such data and a considerable number of different data sources must be integrated to collect the information necessary to generate an accurate water footprint. In short, successful management of water data requires consideration of all sources of water consumption, including indirect ones, augmented with water network distribution information.

3 Linked Data for Water Information Integration

Information integration projects typically focus on one-off point-to-point integration solutions between two or more systems in a customized but inflexible and ultimately non-reusable manner. The fundamental concept of Linked Data is that information is created with the mindset of sharing and reuse. Linked Data leverages open protocols and W3C standards, emerging from research into the Semantic Web, for sharing structured data on the Web. It proposes an approach for information interoperability based on the creation of a global information space. Linked Data has the following advantages:

- Separate systems that are designed independently can be later linked at the edges.
- Interoperability is added incrementally when needed and where it is cost-effective.
- Data is expressed in a mixture of vocabularies.

Linked Data is facilitating the publishing of large amounts of structured data on the web. The resulting interlined data can be considered as a Web scale dataspace supported by the Semantic Web technologies. The *Linked Open Data*[1] represents a large number of interlinked datasets that are being actively used by industry, government, and scientific communities. Linked Data promotes four basic principles for exposing, sharing and connecting data. The first principle encourages the use of *Uniform Resource Identifiers* (URIs[2]) for naming things. The second principle recommends the use of *Hyper Text Transfer Protocol* (HTTP[3]) for URIs, so that data can be retrieve from names using standard protocols. The third principle promotes the use of standard web formats, such as the *Resource Description*

[1] http://lod-cloud.net/

[2] https://tools.ietf.org/html/rfc3986

[3] https://tools.ietf.org/html/rfc2616

Framework (RDF[4]) or the *JavaScript Object Notation* (JSON-LD[5]), for making data available through URIs. The fourth principle encourages contextualization of data by providing links to other related URIs, thus creating a data network. Within the context of water analytics, following these principles enables standardized access and supports interoperability for applications that aim to exploit water information.

4 Linked Real-time Dataspace for the Waternomics Project

The goal of the *Waternomics*[6] project is to provide personalized and actionable information about water consumption and water availability to households, companies, and cities in an intuitive and effective manner at a time-scale that is relevant for effective decision making [5]. Access to such information will increase end-user awareness and improve the quality of the decisions regarding water management and governance. Waternomics accomplishes this by combining water usage related information from various sources and domains to offer water information services to end-users. The Waternomics platform enables sharing of water information services across different groups of users by providing a convergence layer on top of existing water infrastructures with minimal disruption. The objective is to expose the data within existing systems, but only linking the data when it needs to be shared. Representing water usage data within the Linked Data format makes it open; thus, allowing it to be easily combined with data from other relevant domains.

4.1 Architecture

The main components of the envisioned architecture, as illustrated in Figure 1, are the data sources of water usage on existing metering systems, a dataspace consisting of Linked Data, a set of support services, and the resulting applications for water management.

- **Water Metering**: At the bottom of the architecture are the operational and legacy information systems. Adapters perform the “RDFization” process, which transforms multiple formats and legacy data to lifts it to the dataspace. Linked Data principles play a crucial part here since they enable interoperability and cross-linking of water information across different sensors. Furthermore, this RDFization enables contextualization of local water information with the openly available Linked Data such as geographical and meteorological information.
- **Data Integration**: The Linked Data integrates at the information-level (data), instead of at the infrastructure-level (system), by focusing more on the conceptual

[4] https://www.w3.org/RDF/
[5] http://json-ld.org/
[6] http://www.waternomics.eu

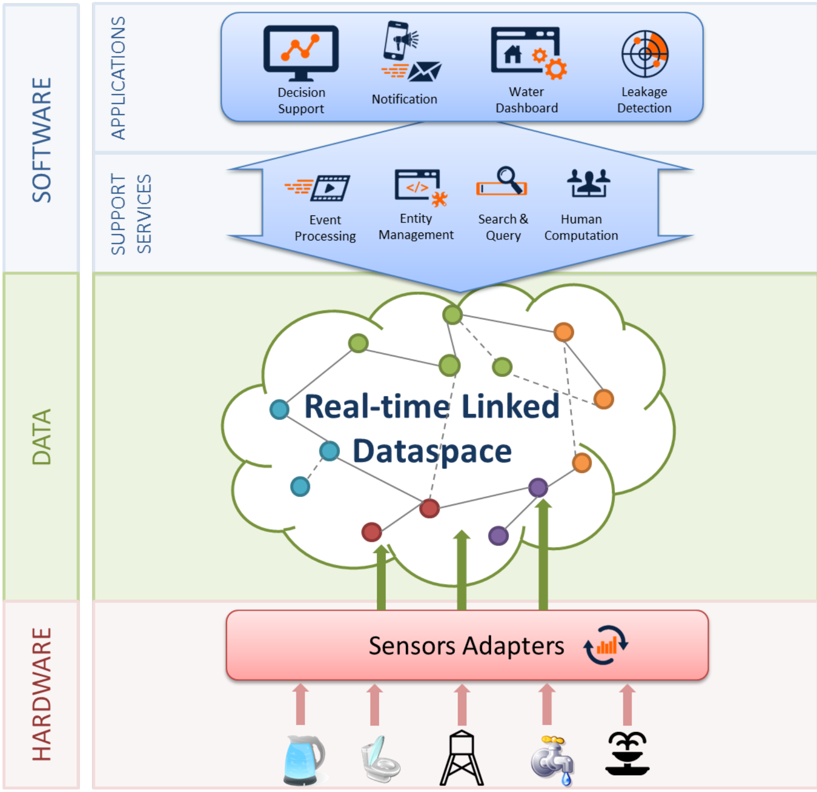

Fig. 1 The Waternomics Platform

similarities (shared understanding) between information. The resulting Real-time Linked Dataspace is rich with knowledge and semantics about water usage performance indicators and forms the basis for real-time water usage analytics and other applications with the help of support services. A key aspect of integration based on Linked Data principles is the re-use or mapping of existing vocabularies and ontologies for describing water data, thus facilitating semantic linkage within and beyond the dataspace.

- **Dataspace Support Services**: The support services are designed to simplify the consumption of the Real-time Linked Dataspace by encapsulating common services for reuse (e.g. search and query, entity management, event processing, etc.). These support services exploit Linked Data technologies and provide additional tools for aggregation, analysis, and improvement of basic data gathered through water metering. Furthermore, these services enrich the aggregated data for complex analytical queries. The primary purpose of these services is to provide Application Programmable Interfaces (APIs) over the dataspace that can be re-used by application developers.

- **Water Analytics Applications**: At the top of the architecture are the water usage and management applications that consume the resulting data and events from the Real-time Linked Dataspace. These applications not only consume the information from the dataspace but also generate user-friendly views over the underlying data.

The support services play a crucial role in realization and exploitation of the Real-time Liked Dataspace. These services include but are not limited to:

- **Search & Query Service**: The query service concerns the technical aspects of enabling access to the data in the Real-time Linked Dataspace through structured queries or RESTful API calls. The query service also enables low latency data analysis. The search service provides keyword-based lookup queries over underlying data sources and their descriptions. The objective of such a service is to help developers and applications in a situation when their queries are not well-defined.
- **Entity Management Service**: This service provides a catalog that serves as the central registry of entities, datasets, and data sources. Within the catalog, all water related datasets, entities, and other sources of information are declared along with their descriptions. This includes a) the list of entities such as sensors or locations that are important for understanding water data and b) open data sources that are relevant to water management such as weather observation stations or forecast services. Besides the APIs and query endpoints provided by the individual data sources, the catalog also provides queries services over the descriptions of entities and datasets.
- **Event Processing Service**: The event processing service allows automatic matching of events similar to users defined rules based on a semantic model for water management. Thus, it simplifies the task of water sensor management. It allows the system to go up early, while administrators can add more meta-data for sensor management in a pay-as-you-go manner [13, 15, 8].
- **Human Computation Service**: The support services, as described above, are primarily focused on providing management tools and programmable access to the constituent information of the Real-time Linked Dataspace. These services are further complemented with a human computation service that is concerned with the collaborative aspect of data management [27, 30]. Essentially, it allows small tasks for data management to be distributed among people who are willing to participate in the dataspace management and improvement process [29]. The same service is further utilized for spatial tasks of data management in public spaces [28].

5 Water Management in Public Spaces

One of the distinguishing aspects of the Waternomics project is its wide variety of end-users. Waternomics has four pilot sites to test and validate its research activities, data platform, and applications. The pilot sits represent use cases of water management in public spaces, as summarized in Table 1.

Table 1 Summary of pilots for the Waternomics project

Pilot	Usage	User Groups
Linate Airport	Corporate	Corporate Staff, Travelers, Shop Owners
Municipality of Thermi	Domestic	Families
NUIG Engineering Building	Public	Building Managers, University Staff, and Students
Coliste na Coiribe School	Public	Building Managers, School Staff, and Students

5.1 Linate Airport

The Linate Airport pilot targets corporate users that are staff members of the airport including building managers, technicians, and engineers. These are adult users that have an advanced level of education and skills to work in such environment. Besides staff members of the airport, target users also include passengers that range from a wide variety of casual to business travelers from different age groups from kids to adults.

Linate Airport is deeply embedded in the urban belt of the city of Milan in Italy. It has a total area of approximately 350 hectares. The airport clientele is predominantly passengers on particular national and international particular routes. In 2012, the Linate airport has operated for 6.3% of the passengers, and 2.2 % of the goods in transit through Italian airports. The airport has two runways for landing and take-off. The first runway (2,442 meters long) is intended for commercial aviation and the second runway (601 meters long) is intended for general aviation. The airport aprons, ramps, and parking stands allow for the simultaneous parking of 46 aircraft. The passenger terminal extends over five levels with a total area of about 75000 meters square (of which about 33000 meters square are open to the public). The terminal is equipped with 71 check-in counters and 24 gates, five of which serve as a loading bridge. Approximately 21% of the area open to the public is dedicated to commercial activities (shops, restaurants, bars, car rentals, banking services, post offices, branches of public services) and 7.5% to the services provided by airlines (check-in, ticketing).

Given the complexity of an airport, a key aspect of this pilot site has been the cooperation with the company that operates the airport. In particular, information on commercial activities and information on key water consumers within the airport, as well as the water and wastewater infrastructure, have been readily made available.

5.2 Municipality of Thermi

The pilot concerning the Municipality of Thermi in Greece targets domestic users. Families are the primary users including children, young adults, and adults. The Municipality of Thermi is situated in the eastern area of the prefecture of Thessaloniki, at a distance of 15 kilometers from the metropolitan center of Thessaloniki. The Municipality of Thermi consists of 14 communities with Thermi being the seat of the Municipality, covering an area of 38.34 hectares. The total population of the Municipality of Thermi is 53,070 according to the census in 2011; however, the actual population is now estimated at 70,000.

The main land use in the area is agriculture; however, land use is changing with more land being dedicated to various types of buildings and infrastructure. Thermi has a strong developmental relationship with an urban area located in close proximity: the Thessaloniki Urban Agglomeration (TUA). It is a rapidly growing and economically viable zone, which is developed as a residential expansion of the TUA, but also as a pole for the location of industrial plants, tertiary sector activities, and highly specialized services, maintaining, at the same time, the characteristics of a developed suburban agricultural economy. At the southeast part of the settlement, there is a planned area of soft manufacturing activities. Finally, there are some large land property areas, such as the military installations, the airport, the American Farm School and the buildings of Aristotle University of Thessaloniki (AUTH), which cover a significant amount of land in the area.

For the purpose of water management pilot, a selection of 10 households was made. These households were selected so that they represent a wide variety of types of houses and families in order to examine the effects of different types of domestic users.

5.3 NUIG Engineering Building

The pilot in National University of Ireland Galway (NUIG) targets staff members (including managers, technicians, and researchers) and students (including undergraduates and postgraduates). While staff members are interested in understanding water usage behaviors and detecting saving opportunities, students are interested in visualizing the building consumption and water consumption data in their projects and research works. The age groups of this pilot site range from young adults to adults. NUIG is one of the Ireland's national universities, founded in 1845, it is ranked among the top 2% of universities in the world. Located in the city of Galway (population 70,000 approximately) on the west coast of Ireland, NUIG has more than 17,000 students and 2,500 staff. The Engineering Building at NUIG is a state of the art educational facility designed to be a “living laboratory” where the building itself is an interactive teaching tool. The Engineering Building opened in 2011, it is the largest engineering building in Ireland and includes lecture halls, classrooms, offices, laboratory facilities, a café, showers, and bathrooms. The building accommodates approximately 1,100 students and 100 staff on four floors (in 14,000 meters square of floor space). The majority of students are undergraduates aged 18-24 years.

5.4 Coláiste na Coiribe School

Similar to the university, a secondary school in Galway has both staff members and students as target users. The main difference is the age groups of users which range from kids to adults on this site. Coláiste na Coiribe (CnaC) is an Irish language secondary school with approximately 350 students and 25 teaching and administrative staff. The existing school is housed at a small location in the city center. To facilitate the demand for places at the school and to address space pressures, a new school (7,400 meters square) was under constructed at a suburban location in Galway.

This new school building serves as a pilot for Waternomics. The new school accommodates up to 720 students (aged 12-18) and includes classrooms, offices, sports halls and associated toilet and shower facilities. As the school was identified as a suitable pilot site at the early stages of construction it provided an opportunity for the Waternomics project team to engage with the designers and contractors in the deciding on the provision of water metering and water information infrastructure for the building. In addition, it provided a unique opportunity to monitor this new building from the beginning of its occupation.

The new school building opened in October 2015, it facilitated engagement with students at an early age regarding water consumption behavior. Furthermore, these students tested and gave feedback to the project on how the platform functions in communicating complex water-related data to a wider audience. The collaboration between the school and the Waternomics project resulted in students actively providing inputs to the project (e.g. user interface design, applications etc.). The school management faces key budgetary and conservation targets; to date reporting on water and associated energy consumption has been very limited. The pilot informs future design of similar buildings with particular focus on water conservation measures and rainwater harvesting systems.

6 Realization of Waternomics Platform

In this section, we provide a concrete realization of the Real-time Linked Dataspace using the tools and techniques discussed in previous sections. We have implemented the dataspace, for the Waternomics project, as a realization of the Lambda architecture. The Lambda architecture was introduced with the aim of allowing seamless ingestion and processing of streaming events data [31]. It consists of three layers: the batch layer deals with processing of large quantities of historical data, the speed layer processes real-time data to minimize latency, and the serving layer provides combined query access to data from other two layers. Our implementation departs from the original Lambda architecture due to the central role of catalog service in the implementation of the batch, speed, and serving layers. The support services in the dataspace are mainly implemented through customization of following open source software: Druid, Apache Spark, MySQL, Apache Kafka, and Apache Cassandra.

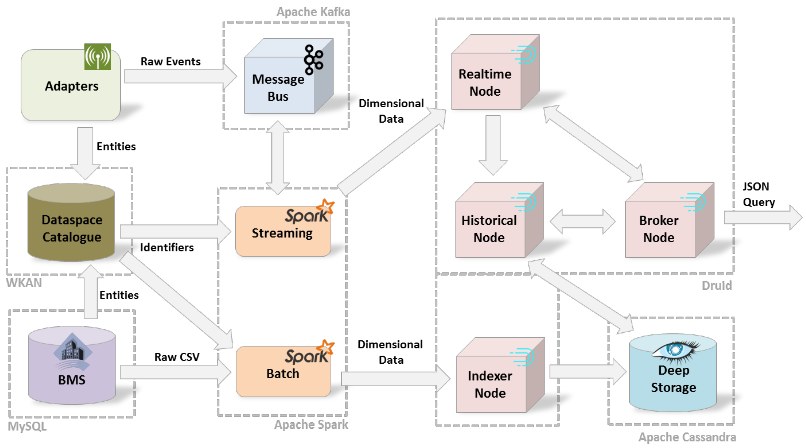

Fig. 2 Lambda architecture realization of the Real-time Linked Dataspace for Waternomics project

Figure 2 shows the data from a building management system (BMS) and water sensors in the Galway pilot being processed in the dataspace. All data sources and entities are defined in the catalog (WKAN). The batch layer is implemented using Spark SQL when historical data from BMS is fed into the indexer node of Druid. Real-time data from sensors is fed into the Kafka message broker, which provides a high availability integration point for speed layer data from the different pilots. Real-time data from Kafka is processed through Spark Streaming to a real-time node of Druid. The combined code from Spark Streaming and Spark SQL provides a standardized way of generating dimensional data that is served using the Druid cluster. The Druid nodes use Cassandra as deep storage for historical data. The batch data is made available through the historical node and streaming data is made available through the real-time node. Periodically, the streaming data is pushed to the historical node as new data arrives. The broker node of Druid seamlessly exposes batch data and real-time data, without the need for writing queries for real-time and batch data separately.

6.1 Data Sources & Open Data

The pilots in the Waternomics project aim to collect both real-time and historical data for water management. For instance, the NUIG and CnaC pilots include following data sources for large buildings.

- Historical and batch data from building management system
- Real-time data from ultrasonic water sensors

A set of relevant open datasets for both pilot sites are included in the Waternomics catalog:

- Open data from weather prediction and observation services
- Public calendar data used by analytics services for distinguishing between water consumption in working days and holidays.
- Drought data in Ireland
- Updates from Irish Water services

All of the above-mentioned data sources joined the real-time dataspace for NUIG and CnaC pilots through definition in the WKAN catalog. Figure 3 shows a list of datasets for the NUIG pilot. It shows summary meta-data for each dataset in the form of tags and description. Users can select a data source to reveal further meta-data which includes the location of data. As a convention, all datasets for historical and real-time data from sensors of pilots are tagged as private. This way there associated meta-data is only visible to authorized users. By comparison, open data sets are defined as public datasets which can be used by everyone.

6.2 Applications

The applications that may be built on top of Waternomics dataspace are diverse; they include water awareness dashboards, decision support for the different targeted users (i.e. domestic users, organizations, cities), and water availability/forecasting, dynamic pricing, and water footprints.

- **Water awareness**: Low comprehension of water flows by users and over usage is one of the biggest causes of water wastage. A lack of awareness on the amount of water consumed leads to the lack of incentives to monitor and affect the situation. Water awareness requires different information for household, company, and city

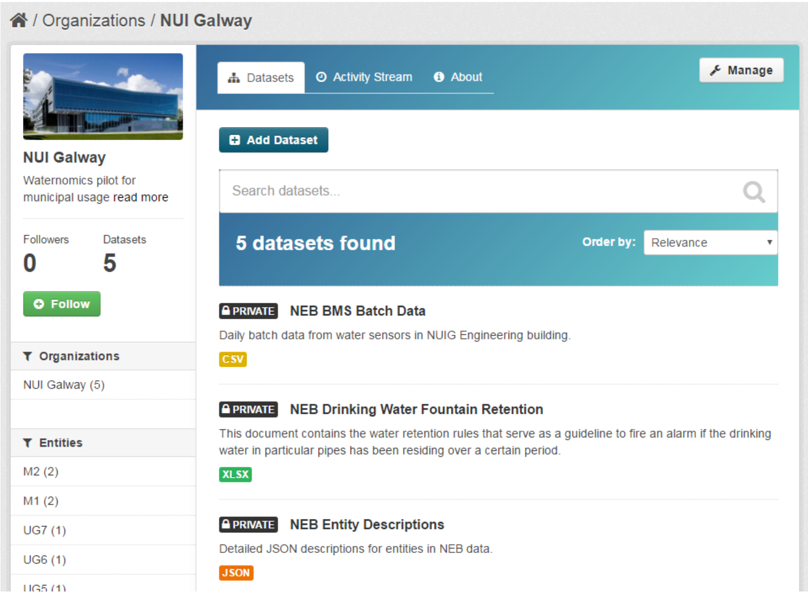

Fig. 3 Datasets and data sources in the WKAN catalog for the NUIG pilot

level, and where different decisions are taken to manage water on these levels. Therefore, water awareness dashboards need to be tailored to different needs of different water usage levels. The data collected by smart water meters is enriched with contextually Linked Data and processed in real-time; hence, allowing for deeper data analysis and faster reactions.

- **Water consumption**: Hydro-meteorological forecasts predict natural demand and supply of water and can be used to prepare and adjust water supply. Forecasting systems can achieve different goals depending on the level of the system deployment. At the household level, forecasts include analysis of occupants behavior and water consumption based on similar historical water usage. These forecasts can be incorporated into dashboards and used as the drivers for water-saving goal. Forecasting models can further leverage Linked Open Data at the neighborhood or city-level. At the company-level forecasts similar to those of the household level are also augmented by models or simulations of the water needs of subsystems within the organization. Linked Data can be used to perform benchmarking between similar organizations to identify areas of potential water optimization.
- **Water education**: Understanding the impact of a product or service requires an analysis of all potential water consumption associated with its entire life-cycle. For instance, a water footprint of a product would provide a quantitative cradle-to-grave analysis of the product/services global water costs (i.e., water used in raw materials extraction, through materials processing, manufacture,

distribution, use, repair and maintenance, electricity generation, and disposal or recycling). Building a water footprint requires the gathering of water data from many participants within the supply chain. Linked Open Data can be a key enabler for the development of a global information ecosystem of water footprint inventory data on products, services, and organizations.

In the following, we present a set of applications developed for the NUIG Engineering Building and the CnaC School. Since both pilot sites have many commonalities, they share two main applications: the Public Display and the Manager Dashboard. Table 2 presents an overview of the applications developed in terms of their objective and target user groups.

Table 2 Mapping of applications against objectives and user groups

	Public Display	Manager Dashboard	Water Retention Time Observer	Observatories Control Panel	Wearable Info-centre	Goal-oriented Accessing Water
Objectives of Applications						
Increase Water Awareness	✓	✓	✓	✓	✓	✓
Reduce Water Consumption				✓		✓
Promote Water Education	✓					
User Groups in University						
President		✓	✓			
Building Services Manager		✓	✓	✓		
Chief Technical Officer		✓	✓	✓	✓	
Consultants / Contractors		✓				
Technicians		✓	✓	✓	✓	
Staff / Lecturers	✓		✓			✓
Researchers (PG / PD)	✓		✓			✓
Students	✓		✓			✓
User Groups in School						
School Principal		✓		✓		
Building Contractor		✓		✓		
Teachers	✓					
Students	✓					

6.2.1 Public Display

A key objective for both pilot sites in Galway was to increase water usage awareness in public spaces. Towards this objective, setup of a kiosk with an interactive dashboard can help attract people and engage them with discovering water usage details of their building. A public dashboard is a web application that shows generic information of the site's water consumption as compared with social norms; in addition, it displays information related to consumption per student, toilet flushes per day, etc.

The web application developed for Galway pilots, as shown in Figure 4, serves as a communication medium to display the amount of water being consumed in various parts of the building. This application shows volumetric values of water usage in other dimensions such as cost, metaphors, and footprints. The image on left side of Figure 4 shows the amount of water in terms of the number of standard size of water cooler bottles. This application visualizes water quantities in circles using colors to indicate if the water usage is high or low. This application also uses social media to inform users about the technology used within the Waternomics project, and its updates. Users can interact further with the application to explore the water usage data over a full month; furthermore, they can get more details about technologies used within the Waternomics project.

Fig. 4 Public Display showing water data from the NUIG Engineering Building

The public dashboard application uses the Water Analytics Support Service for querying the data from the dataspace. This service has been extended to serve as an extension of the public dashboard and allow users to explore further the water data by scanning QR codes near the sensors. As shown in Figure 5, users are able to visualize the entire month's consumption. This extension aims to support students

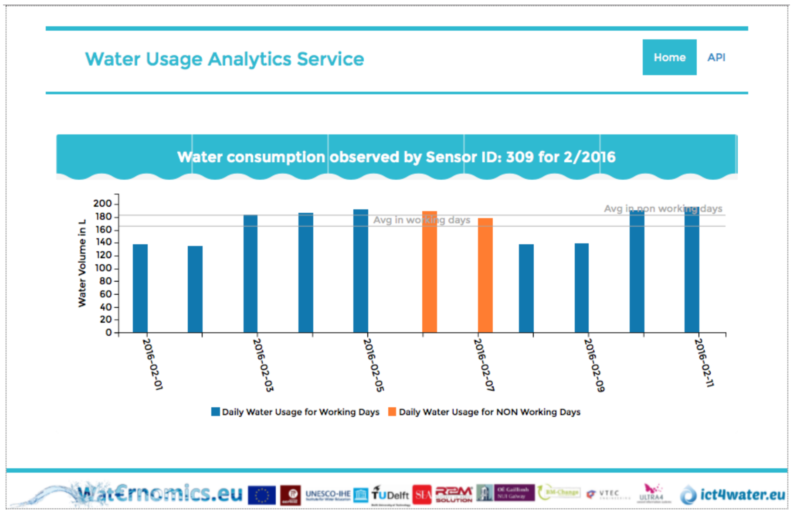

Fig. 5 Visualizing Water Analytics for the sensor 309 in February 2016

and researchers who can retrieve the data from this service via its API to use in their research projects.

6.2.2 Manager Dashboard

Managers in the NUIG Engineering Building and CnaC school are interested in watching the consumption at different points of the water network. In both pilots, dashboards can be considered rather as a family of applications targeting the specific needs of managers than as a single application aiming to solve all problems for all users. One of the key elements used in the Manager dashboards are historical graphs showing the consumption in various points or groups of interests (see Figure 6). Goal setting and tracking is also an important aspect for managers in the Galway pilots so comparison graphs are an important part of their dashboards.

6.2.3 Water Retention Time Observer

Making drinking water available becomes a major concern in public spaces. This can be guaranteed through a carefully selected location for drinking water fountains in order to make sure that water is always flowing in the pipes. However, in spaces such as a university building, drinking water fountains can remain unused during long holidays and weekends. Consequently, drinking water can reside in the pipes for long periods. Building managers want to make sure that residual water is still safe to drink.

Fig. 6 Manager Dashboard for NUIG Engineering Building

In this context, the water retention time observer application can assist managers to guarantee that they receive timely notifications regarding water that has been residing for a long period in drinking water pipes. This is done by allowing them to setup a set of rules for tracking periods of inactivity in specific measurement points and automatically send notifications through the system to selected user-groups. Figure 7 shows the list of active alarms detected by the application. The application is well aligned with one of the objectives of Waternomics project, i.e. giving actionable information to water users and managers.

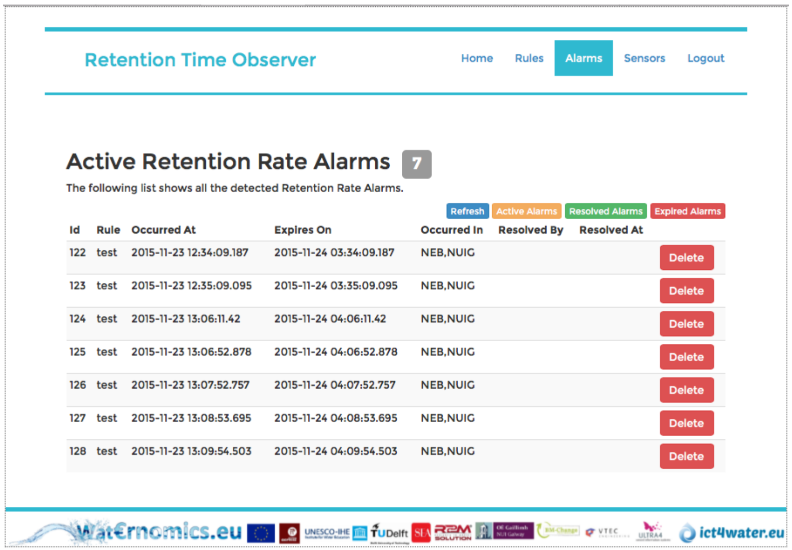

Fig. 7 Retention Time Observer and Active Alarms

6.2.4 Observations Control Panel

Both pilot sites in Galway also aim to improve water network management by assisting staff in coordinating and making better-informed decisions. An additional aspect in this context is the ability to communicate messages to specific user-groups related to their consumption in order to require actions or encourage behavior change. The observations control panel is an application that gives an overview of the status of all notifications within a timeframe. It provides an interface for managing notifications that can originate from any application that uses data from the dataspace. Based on the activity logging on different notifications, the user can see how much time it takes from the time of creation of a notification to the time of action or expiry. The application also allows to filter notifications based on the group they were targeting, criticality level and the source application. The control panel allows managers to not only show but generate custom notifications themselves to facilitate this communication with specific user groups.

6.2.5 Wearable Info-centre

Managers at the Galway pilots are very mobile and they require instant notifications of important aspects of the water consumption in their building. In this case, however, users are more technology friendly and expressed their willingness to use more advanced mechanisms for receiving notifications through smart devices. The wearable info-center application was developed for mobile notifications.

Fig. 8 Wearable info-centre, receiving notification from the water retention time observer

The wearable info-centre is an application that the user installs on a smartwatch to display notifications as they are received on the mobile phone. This way users dont have to check on their phones every time they receive notifications from the platform. Instead, they can check their smartwatches which is less obtrusive while communicating the information at any time. Figure 8 shows an example of using the wearable info-centre. The application provides an interface for displaying notifications that can originate from any application which uses data from the real-time dataspace. So, the application is indirectly uses all kinds of data provided by the real-time dataspace.

6.2.6 Goal-oriented Accessing Water

One of the ideas explore during the user tests was the concept of allowing users to track their own personal consumption patterns. The patterns are based on the applications that are activated and connected with specific micro-sessions of a user consuming water such as preparing coffee, drinking water, washing hands etc. One of the key outcomes in user tests was that mobile and wearable devices can offer a great opportunity in personalized tracking but this is hard to do if it requires an additional action to already existing routine (e.g. if it requires you to get your phone out of your pocket and scan a QR code). So, in the goal-oriented accessing water application we experimented with the idea of replacing an action in user's routine while in parallel providing some short pieces of information.

This concept challenges the centuries old mechanism of operating a faucet, which in fact is a valve of various designs. The new system transforms water usage into goal-oriented activity such that accessing water is no longer just about enabling a valve. By setting up a touch enabled sink display next to a faucet (without its

Fig. 9 Goal-oriented Accessing water application

original turning knob), users were able to choose certain water activity such as “one cup of tea” or “one bottle of water”. This message will be sent through wireless to a solenoid valve connected to a water pipe or faucet that provides a certain amount of water. In this manner, users were always aware of their water usage thus lowering the chance of wastage. A social network system was also implemented into the system such that users could report issues to each other or even to the building manager so that urgent problems can be solved more rapidly to prevent waste of water in any case.

7 Related Work

In general, data management is seen as a challenge for smart infrastructures [7, 24, 4]. As recent surveys show, a number of policies and standards for smart metering have been adopted in different countries, but most standards still contain a fragmented set of solutions with little support for adding contextual data [22]. Most policies and standards appear in the smart grid area and are adopted by other areas [9]. Hydro-meteorological information is mainly described by drought indicators [1] such as Standardized Precipitation Index [3] and Temperature Condition Index [19]. Mostly these indices describe the present state of the system [2].

It has been shown that water consumption awareness and the strength of motivation greatly affect the potential for water saving. For example, in [18] the deployment of an experimental system that provided detailed water usage information in the shower showed the resulting decrease in water consumption. It also showed the division of users into two groups: those who continued to pursue conscious water behavior even after the experiment was over, and those who returned to previous water habits after the removal of informational displays. An overview of pro-environmental behavior models and key human-computer interaction techniques to promote and motivate such behavior are presented in [11].

In [23], a display to present gas, electricity, and water consumption in an artistic way is described. In [21], a persuasive application to promote a responsible attitude towards natural resources, food, and water during family interactions is described. The comparison between lightweight ambient and numeric displays is performed in [20]. Results showed that an abstract ambient display with color-coded visualization of water usage causes bigger water-saving behavior changes comparing to a numeric display. In [14], group-based feedback is used to reduce the consumption of paper within an office environment.

All of these techniques are complementary to real-time dataspaces for water analytics. The approach we propose here aims to make it easier to implement such applications by reducing the cost of gathering the necessary data to drive the applications.

8 Summary

This chapter motivates the need for efficient water information management in public spaces and presents a Real-time Linked Dataspace approach for water data services. A high-level architecture, for the Real-time Linked Dataspace, realizes this approach in the context of the Waternomics project. The Waternomics project established the utility of this approach with the help of four pilot sites that represent different scenarios of public spaces. This chapter describes a concrete instantiation of the Real-time Linked Dataspace approach for two educational institutions, along with applications supported by the water data services.

Acknowledgements The research leading to these results has received funding under the European Commission's Seventh Framework Programme from ICT grant agreement WATERNOMICS no. 619660. It is supported in part by Science Foundation Ireland (SFI) under Grant Number SFI/12/RC/2289.

References

1. Barua, S., Ng, a.W.M., Perera, B.J.C.: Comparative Evaluation of Drought Indexes: Case Study on the Yarra River Catchment in Australia. Journal of Water Resources Planning and Management **137**(2), 215–226 (2011)
2. Boken, V.K.: Improving a drought early warning model for an arid region using a soil-moisture index. Applied Geography **29**(3), 402–408 (2009)
3. Cancelliere, A., Mauro, G.D., Bonaccorso, B., Rossi, G.: Drought forecasting using the standardized precipitation index. Water Resources Management **21**(5), 801–819 (2007)
4. Cavanillas, J.M., Curry, E., Wahlster, W.: New Horizons for a Data-Driven Economy: A Roadmap for Usage and Exploitation of Big Data in Europe. Springer Publishing Company, Incorporated (2016)
5. Curry, E., Degeler, V., Clifford, E., Coakley, D., Costa, A., van Andel, S.J., van de Giesen, N., Kouroupetroglou, C., Messervey, T., Mink, J., Smit, S.: Linked Water Data for Water

Information Management. In: 11th International Conference on Hydroinformatics (HIC) (2014)
6. Curry, E., Donnellan, B.: Sustainable Information Systems and Green Metrics. In: Harnessing Green It: Principles and Practices, pp. 167–198 (2012)
7. Curry, E., Dustdar, S., Sheng, Q.Z., Sheth, A.: Smart citiesenabling services and applications. Journal of Internet Services and Applications **7**(1), 6 (2016)
8. Derguech, W., Bhiri, S., Hasan, S., Curry, E.: Using formal concept analysis for organizing and discovering sensor capabilities. Computer Journal **58**(3), 356–367 (2015)
9. Fang-Yuan Xu, Long Zhou, Yi Lin Wu, Yingnan Ma: Standards, policies and case studies in smart metering. In: IEEE PES General Meeting, pp. 1–5 (2010)
10. Franklin, M., Halevy, A., Maier, D.: From databases to dataspaces. ACM SIGMOD Record **34**(4), 27–33 (2005). DOI 10.1145/1107499.1107502
11. Froehlich, J., Findlater, L., Landay, J.: The design of eco-feedback technology. Proceedings of the 28th international conference on Human factors in computing systems pp. 1999–2008 (2010)
12. General Assembly, United Nations: Transforming our world: The 2030 agenda for sustainable development. Tech. Rep. 1 (2015)
13. Hasan, S., Curry, E.: Approximate Semantic Matching of Events for the Internet of Things. ACM Transactions on Internet Technology **14**(1), 1–23 (2014)
14. Hasan, S., Medland, R., Foth, M., Curry, E.: Curbing Resource Consumption Using Team-Based Feedback. In: Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), vol. 7822 LNCS, pp. 75–86 (2013)
15. Hasan, S., O'Riain, S., Curry, E.: Towards unified and native enrichment in event processing systems. In: Proceedings of the 7th ACM international conference on Distributed event-based systems - DEBS '13, p. 171. ACM Press, New York (2013). DOI 10.1145/2488222.2488347
16. Heath, T., Bizer, C.: Linked Data: Evolving the Web into a Global Data Space. Synthesis Lectures on the Semantic Web: Theory and Technology (2011)
17. Hertel, T.W., Liu, J.: Implications of water scarcity for economic growth (2016)
18. Kappel, K., Grechenig, T.: "Show-Me": Water Consumption At a Glance To Promote Water Conservation in the Shower. Persuasive 09 Proceedings of the 4th International Conference on Persuasive Technology pp. 1–6 (2009)
19. Kogan, F.N.: Application of vegetation index and brightness temperature for drought detection. Advances in Space Research **15**(11), 91–100 (1995)
20. Kuznetsov, S., Paulos, E.: UpStream: motivating water conservation with low-cost water flow sensing and persuasive displays. In: Proceedings of the 28th international conference on Human factors in computing systems - CHI '10, p. 1851. ACM Press, New York, New York, USA (2010)
21. Lepe-Salazar, F., Yamabe, T., Alexandrova, T., Liu, Y., Nakajima, T.: Family interaction for responsible natural resource consumption. Proceedings of the 2012 ACM Annual Conference on Human Factors in Computing Systems (CHI '12) pp. 2105–2110 (2012)
22. Liotta, A., Geelen, D., van Kempen, G., van Hoogstraten, F.: A survey on networks for smart-metering systems. International Journal of Pervasive Computing and Communications **8**(1), 23–52 (2012)
23. Makonin, S., Pasquier, P., Bartram, L.: Elements of consumption: An abstract visualization of household consumption. In: Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), vol. 6815 LNCS, pp. 194–198 (2011)
24. Nam, T., Aldama, F.A., Chourabi, H., Mellouli, S., Pardo, T.A., Gil-Garcia, J.R., Scholl, H.J., Ojo, A., Estevez, E., Zheng, L.: Smart cities and service integration pp. 333–334 (2011)
25. OECD: OECD Environmental Outlook to 2050. OECD Environmental Outlook. OECD Publishing (2012)
26. Pereira, Â.G., Rinaudo, J.D., Jeffrey, P., Blasques, J., Quintana, S.C., Courtois, N., Funtowicz, S., Petit, V.: ICT Tools to Support Public Participation in Water Resources Governance & Planning: Experiences from the Design and Testing of a Multi-Media Platform. Journal of Environmental Assessment Policy and Management **05**(03), 395–420 (2003)

27. Ul Hassan, U., Bassora, M., Vahid, A.H., O'Riain, S., Curry, E.: A collaborative approach for metadata management for Internet of Things: Linking micro tasks with physical objects. In: Proceedings of the 9th International Conference on Collaborative Computing: Networking, Applications and Worksharing, pp. 593–598. IEEE (2013)
28. Ul Hassan, U., Curry, E.: Efficient task assignment for spatial crowdsourcing: A combinatorial fractional optimization approach with semi-bandit learning. Expert Systems with Applications **58**, 36–56 (2016)
29. Ul Hassan, U., O'Riain, S., Curry, E.: Towards Expertise Modelling for Routing Data Cleaning Tasks within a Community of Knowledge Workers. In: Proceedings of the 17th International Conference on Information Quality. Paris, France (2012)
30. Ul Hassan, U., Zaveri, A., Marx, E., Curry, E., Lehmann, J.: ACRyLIQ: Leveraging DBpedia for Adaptive Crowdsourcing in Linked Data Quality Assessment. In: E. Blomqvist, P. Ciancarini, F. Poggi, F. Vitali (eds.) Knowledge Engineering and Knowledge Management, pp. 681–696. Springer International Publishing (2016)
31. Yang, F., Tschetter, E., Léauté, X., Ray, N., Merlino, G., Ganguli, D.: Druid: a real-time analytical data store. In: Proceedings of the 2014 ACM SIGMOD international conference on Management of data - SIGMOD '14, pp. 157–168. ACM Press, New York, New York, USA (2014)