This paper was published in Kolbe, Bill, Donaubauer (eds.): Geoinformationssysteme 2017 - Beiträge zur 4. Münchner GI-Runde, Wichmann Verlag, Heidelberg, ISBN 978-3-87907-626-0

Integrating and Managing the Information for Smart Sustainable Districts - The Smart District Data Infrastructure (SDDI)

Mandana MOSHREFZADEH[1], Kanishk CHATURVEDI[1], Ihab HIJAZI[1,2], Andreas DONAUBAUER[1] and Thomas H. KOLBE[1]
[1]Chair of Geoinformatics · Technical University of Munich · Arcisstraße 21 · 80333 München
[2]An-Najah National University · Department of Urban Planning Engineering · Nablus, Palestine
E-Mail: mandana.moshrefzadeh@tum.de
E-Mail: kanishk.chaturvedi@tum.de
E-Mail: ihab.hijazi@tum.de
E-Mail: andreas.donaubauer@tum.de
E-Mail: thomas.kolbe@tum.de

Abstract

The goal of making cities smart and sustainable leads to an urgent need for the development of a stable information architecture which is interoperable, functional, extensible, secure and transferable. A main part of this architecture is the data infrastructure, which covers the services supporting the dynamic data collected by various sensors and also a virtual district model, representing the physical district's objects, which can be enriched with semantic information, i.e. thematic information from different application domains. Open Geospatial Consortium (OGC) standards such as CityGML and Sensor Web Enablement (SWE) are crucial in the establishment of this model. In this paper, the process and the concept of the smart data infrastructure developed for Smart Districts are described in detail from five viewpoints according to the standard ISO 10746 "Information technology – Open Distributed Processing – Reference model". The paper concludes with an example explaining the configuration of the so-called SDDI (Smart District Data Infrastructure).

1 Introduction and background

According to the Department of Economic and Social Affairs of the United Nations Secretariat, 53% of the world's population resided in urban areas in 2014 and the world continues to urbanize rapidly (UNITED NATIONS 2014). The process of urbanization is accompanied with challenges regarding economic, social and ecologic aspects. For example, even today, 60-80 per cent of energy is consumed in cities, and cities account for 75 per cent of carbon emissions. Consequently, "to make cities inclusive, safe, resilient and sustainable" is one of the sustainable development goals which were developed as an outcome of the United Nations Conference on Sustainable Development (Rio+20) and came into force in 2016 (UNITED NATIONS 2016).

Figures like those above are quite often used by authors in the field of "smart cities" in order to motivate their work. Companies like SIEMENS[1], IBM[2], Microsoft[3] and CISCO[4] cite those figures in their advertising material for their smart grid, smart homes, smart traffic and the Internet of Things (IoT) solutions, which try to improve living in rapidly growing megacities.

A closer look at the situation of cities reveals more diversity than the figures above suggest. First, the definition of the terms "city" and "urban area" varies from country to country. For example, in Norway, a locality with more than 200 inhabitants is called an urban area whereas Japan requires a city to be called urban area if it has at least 50000 inhabitants[5]. Second, the challenges and problems of cities are quite diverse in different countries of the world and they can be heterogeneous even within a country and even within a city itself. Challenges in the fields of governance, economy, mobility, environment, people and living as observed for example by (MONZON 2015) can therefore range from urban sprawl, traffic congestions, and an urban ecosystem under pressure to shrinking cities, unemployment, and oversized infrastructure.

There is no doubt that information and communication technologies have the potential to assist citizens and government in facing these challenges, but it also becomes obvious that monolithic IT systems or sectoral smart city solutions might not be suited well for taking into account all the diverse challenges.

Furthermore, the systems and technologies as they are offered by the big players from the IT domain are quite often top down proprietary solutions whereas distributed bottom up solutions might be more suitable taking into account the variety of stakeholders involved in smart city projects: these are owners (e.g. municipality, housing companies, citizens), administrations (e.g. municipality), service owners and providers (e.g. energy providers, transportation and mobility companies) and citizens, all having different interests, goals and tasks and last but not least already existing IT systems helping them to fulfill their specific tasks. The solutions proposed by the big players are mostly related to IoT and Big (unstructured) Data, often not taking into account the variety of structured data sources – such as cadastral data, utility networks and so on – which are already available in diverse IT systems used by city administrations.

Moreover, although nearly 100% of the data analyzed in the context of smart cities is related to artificial and natural physical objects in the city, geospatial information often does not play a major role in smart city projects.

The Smart District Data Infrastructure (SDDI) concept described in this paper takes into account these insights and proposes an open, distributed, spatially-enabled system architecture capable of integrating structured as well as unstructured data as well as a bottom up processes for addressing the specific information needs of specific city challenges. The SDDI is currently being implemented in four big European cities, and first experiences are reported for

[1] https://www.siemens.de/digitalisierung/smart-city.html

[2] http://www.ibm.com/smarterplanet/us/en/smarter_cities/overview/

[3] https://enterprise.microsoft.com/de-de/industries/citynext/

[4] http://www.cisco.com/c/dam/en_us/solutions/industries/docs/gov/everything-for-cities.pdf

[5] http://www.prb.org/Publications/Articles/2009/urbanization.aspx

a district in East London, i.e. Queen Elizabeth Olympic Park – the former area of the Olympic Summer Games in 2012, which is now subject to a large urban transition process.

2 Smart and Sustainable Districts (SSD)

2.1 District as a governing body

According to the Oxford dictionary, district means, "An area of a country or city, especially one characterized by a particular feature or activity" or "A region defined for an administrative purpose" (OXFORD DICTIONARY 2016). According to this definition, it is hard to define exactly the border of a district. Indeed, the district has different definitions varying from country to country. From the governing and administrative point of view even in central Europe, the way a district is defined varies considerably. Thus, dealing with districts means working with a vast variety of features, challenges and opportunities of a part of a city.

One of the reasons that the focus from cities has been shifted to one lower scale, i.e. district, is the fact that in the district scale, it is more likely that overlaps of different sections and fields can be discovered. It is also more realistic to bring different stakeholders to the table on this scale and to discuss with them and find out what and where the barriers are. If smart and sustainable actions in one district are successfully applied, other neighbouring districts will no doubt be willing to look at this district and try to adapt the solutions in their areas.

There are many activities in the context of the smart cities projects in different domains for instance in energy, mobility, water, last mile logistic, business, etc. Looking at all of them from above, it can easily be seen that these hot topics all have some overlaps with each other. Offered solutions in the framework of Smart City projects are mainly focused on new technologies such as "Humble lamppost", "parking sensors", etc. for various domains such as energy, mobility, crowd management. Each of these technologies is well adapted to the present needs of the cities and performs precisely in such a way that all the requirements for the specific challenges of the cities are met. Nonetheless, these are top down solutions (MOSHREFZADEH & KOLBE 2016). In fact, the lack of adequate bottom-up approaches to a dominance of top-down and supply-focused solutions resulted in indirectly ignoring the sustainable integrated solutions (SÁNCHEZ et al. 2013). It is at this point that the idea of focusing on the district scale and working to offer bottom-up solutions comes to mind.

2.2 Climate-KIC SSD flagship project

Climate-KIC, launched in 2010, is one of the six original Knowledge and Innovation Communities (KIC) set up by the European Institute of Innovation and Technology (EIT). Their mission is to deliver innovative solutions to climate change via a dynamic alliance of European partners drawn from academia, industry and the public sector. In Climate-KIC, the activities are driven by four main themes in different scales. According to the scale of the project and the required support from Climate-KIC, the projects are categorized in three different types in which the "Flagship Project" has the highest importance in terms of scale and scope. "Smart Sustainable District (SSD)" is a Climate-KIC flagship project, started in 2014, in the theme "Urban Transitions" that aims at collaborating with the most ambitious district level developments in the cities and regions represented in the Climate-KIC. It will demonstrate how new thinking, coupled with effective tools, technologies and policies, can lead to 'factor

4 improvements' in city district performance across a range of sustainability measures. 'Factor 4 improvement' is a concept of sustainable management with the goal of having twice the environmental impact for half the cost (VON WEIZSÄCKER et al. 1998).

2.3 SSD structure, partners and districts

The SSD structure consists of seven work packages covering all its activities, one of which concerns data and digitization. The other work packages focus on other aspects such as building physical elements, modelling tools, socio-technical aspects, integration of solutions and process and project management. In this project, about 16 European partners from different organizations including industries, academia, and research institutes are involved in diverse branches of expertise and different activities.

SSD is currently focusing on different parts of Europe to bring in more districts with different structures. At the moment, the involved districts in the project are from Rotterdam (Stadshaven Harbour), Utrecht (The new centre), London (Queen Elizabeth Olympic Park) Paris (Les Docks de Saint Ouen), Berlin (Moabit West), Gothenburg (Johanneberg), Malmö (Southeast), Helsinki (Kalasatama) and Copenhagen (Energy block). The activities in these districts are divided into two types called 'deep dive' and 'non-deep dive' activities. This categorization is based on the level of involvement and investment of Climate-KIC in these different districts. The so-called deep dive districts (DDDs) are those under the main focus. From the above-mentioned districts, Utrecht CS, London QEOP, Berlin Moabit West and Paris Les Docks de Saint Ouen have been chosen as DDDs. For the non-DDDs, the process is defined on a general level, enriched by the outcome and experiences gained from DDDs activities that are shared with other districts. Of course, this has been designed in such a way that all districts benefit.

In general, districts undertake a multi-stage process (CLIMATE-KIC SSD 2016):

- Priorities, strategies and opportunities must first be determined.
- Then tangible 'factor four' outcomes from cross-sector synergy are identified, either through demonstrating unconsidered benefits, or through bringing in new data and modelling scenarios.
- SSD makes the business case for sustainability, achieving environmental, social and economic outcomes.
- The last phase of the process concerns managing, evaluating and refining the proposition. This is done by sharing best practice, integrating and layering project data, understanding interactions, and by applying new techniques.

2.4 District challenges and their related data issues

According to the original structure mentioned in subsection 2.3, one of the work packages is completely devoted to data and its related issues as it was found that "data" is a common challenges of all the districts studied in SSD. We as the leader of this work package have been working on developing an approach that takes into account the requirements of the existing and possible future challenges of the districts.

The first step was the preparation of several surveys, spread to the district networks in order to have a better understanding of the current situation. Additionally, it was also important to get to know about other existing technologies and solutions in the framework of smart cities. Therefore, many research articles as well as engineering reports and Best Practice reports were studied to find out which technologies are available and what gaps need to be worked on. Through multiple discussions and workshops with the districts partners, the main challenges and barriers with respect to the data were studied and the opportunities were discovered.

A list of the barriers, requirements, focus points and opportunities was collected in order to understand what must be offered to the districts to cover their requirements. For example one of the difficulties mentioned was that in addition to the group who facilitates the use of the data, different stakeholders and those who provide the data are highly influenced by the way the data can be accessed. On the other hand, there is more than one group that is interested in using the data. This again shows the necessity of considering interoperability. Therefore, a service oriented architecture which can address issues such as data communication and transformation among different functional units, ensure data security and extensibility of data structure and be independent of changes in data providers and technologies, is key to the solution.

2.5 State of data in districts and their requirements

The outcome of the workshop highlighted an urgent need for a platform that provides links to various data sources, which must be analysed for decision makers, citizens and general users. Indeed, in the district we are not only faced with one system but rather a system of systems, which makes the process very complex. Moreover, new technologies and smart solutions such as Internet of Things (IOT) offer and require open systems.

During interactions with all these districts, it was observed that most of the time the problem is not a lack of data but how to access it and how to know which data are useful for different use cases in the district. Finding the right persons or organizations who have the required information, is on its own a challenging task but convincing them to offer their information in such a project is a big barrier in the transition process. The system and data providers (who are mostly private bodies) cannot be forced to share their own data and services with others but should rather be encouraged to give access to others to work with their data. To do so, they should know that they can link their systems to the platform without a huge effort and that they can define the level of access to others in a controlled manner. In this way, the data security and privacy concerns can be well handled and the stakeholders can be encouraged to integrate their system and data into such a platform. Such a data platform cannot be a monolithic system that can be installed on one machine but consists of multiple individual components some used by everybody and some specific to individual stakeholders.

On the other hand, the way this distributed systems should be managed is not yet clearly stated in any of the given solutions in the “Smart Cities” initiatives. How these distributed systems are related to each other is often overlooked. This is very essential especially when it comes to evaluate the effects of one solution on other solutions and aspects and elements of the system. Therefore, the lack of a common relator may result in misinterpretation of the consequences of the decisions.

The abovementioned facts have inspired us to think about a smart data infrastructure that can be used by any district or city for managing their data infrastructure in a standard manner. The result of this, led to the development of "Smart District Data Infrastructure" (SDDI).

3 Smart District Data Infrastructure (SDDI)

3.1 Introduction and motivation

Existing challenges of the cities have proven that in order to smartly manage the physical infrastructure, a well-designed communication system between different actors, parties and organizations as well as services to the citizens are required (DEGBELO et al. 2015). This also applies to the district scale.

Observing districts that are passing through the smart transition of their services and structures highlights the complexity of these systems. District is a system of systems, which are deeply interconnected with each other. In most cases, changes in one system or service will affect the others. Therefore, it is essential to break down the complexity of districts which are indeed complex distributed systems. On the one hand, the system is open in that it should be extensible. This means that different partners can be part of the system in different ways. On the other hand, the system is called distributed because a number of different stakeholders (e.g. owners, operators, solution providers, citizens, and visitors), agents, communities and various data layers including sensors, analysis tools, etc. are present in it (MOSHREFZADEH & KOLBE 2016). This stresses the debate around a centralized and monolithic approach and its disadvantages (VAN ALSTYNE et al. 1995). Although the *centralized approach* allows pumping of all the various information from different sources into a single repository, the limitations with this approach such as the unwillingness of different source providers for releasing their data into a central repository, difficulty in management of semantics of various data, etc. makes the centralized approach impractical and will rule it out from the discussion.

Therefore, the strategy of SDDI is to adapt the concept of a *distributed system*, consisting of heterogeneous components, which are connected by standardized interfaces. Naturally working in such a complex, distributed system means coping with different aspects of heterogeneity, i.e. different types of data (structured and unstructured data), data models, data formats, applications, stakeholders, software systems and so on. Thus, to manage such a complex distributed system, it is necessary to define the smart transition process using an architectural approach organized as a set of viewpoints.

3.2 The SDDI process

To define the process of SDDI we needed to structure our viewpoints according to the idea of having an open and distributed system. In the "OGC Smart Cities Spatial Information Frameworks" white paper it is recommended to follow the standard ISO 10746 "Information technology — Open Distributed Processing – Reference model (ODP-RM)" (PERCIVALL et al. 2015).

It is necessary to look at the entire task from different views. The standard (ISO/IEC 10746-2:2009) provides a coordinating framework for the standardization of open distributed processing (ODP). This supports distribution, interworking, portability, and platform and technology independence (ISO/IEC 10746-3:2009) and defines the following viewpoints which each focus on different aspects of the system: a) enterprise viewpoint, b) information viewpoint, c) computational viewpoint, d) engineering viewpoint and e) technology viewpoint.

Based on this standard, we have developed a planning process by which the open and complex distributed system can be developed and managed. This design process is very helpful in defining the challenges of a district in the project and distinguishing the responsibilities and tasks on different levels. The overall structure of this process is depicted in Fig. 1.

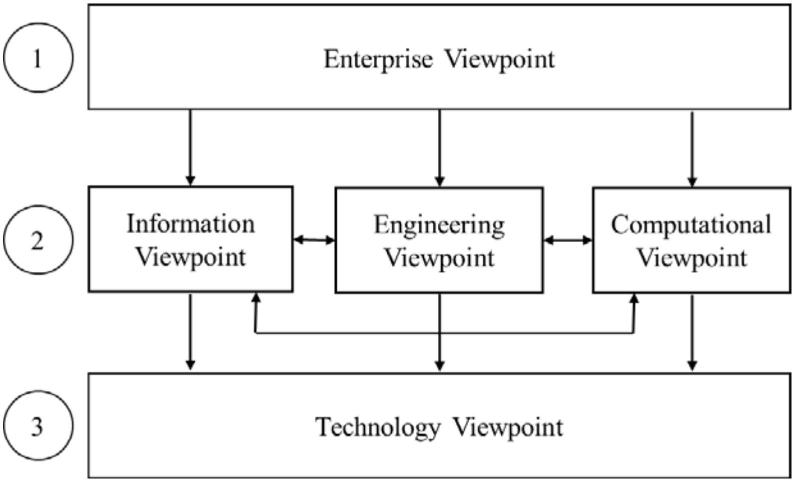

Fig. 1: Planning process for an open and complex distributed system in the context of smart districts.

As a first step, it is necessary to understand the involved parties and stakeholders and to appreciate their roles and interests. The purpose, scope, and policies for the system are what the "Enterprise View" focuses on. It describes the business requirements and ways to meet them. The enterprise view plays a crucial role as it defines the direction of the other steps.

In a second step, the system is structured from three different viewpoints: The "Information View" focuses on data modeling and the semantics of the information to be managed. It defines ontologies specified by data models and their semantic definitions. For example, in the SDDI we employ specific data models for sensor descriptions, sensor observations, and the 3D district model. The "Engineering View" looks how to build the information infrastructure from modular, distributed elements. In our case this includes the decision to employ a service oriented architecture (SOA) using web service interface and protocol definitions from the Open Geospatial Consortium (OGC). The "Computational View" describes the processes as well as the data and control flows. Here, it is specified how the user tasks (and, thus, the purpose of the entire system as defined in the enterprise view) can be realized using the entities and concepts as specified in the information and engineering views.

In the third and last step, the system is being seen from a technical and implementation perspective, i.e. the "Technology View". Here, appropriate software and hardware products are chosen and their configuration and the roll-out of the system are described.

This three step approach also makes clear that choosing specific hardware and software products should only be done, when the other four views have been clearly examined and defined. Nevertheless, because we suggest to use standardized components, the list of available implementations can be investigated and potential costs can be assessed early.

A key aspect of the SDDI is interoperability, i.e. the capability of different components, systems, actors, and datasets to mutually work together. Semantic interoperability is achieved by choosing standardized data models / ontologies in the information view where possible. Syntactic interoperability is required in order to directly be able to connect different components to each other. Again, this is ensured by employing OGC interface and encoding standards in the engineering viewpoint. This also gives freedom to choose from the large variety of software products in the technical viewpoint which comply to these standards avoiding any vendor lock-in.

3.3. The SDDI reference model

Following the steps defined in the SDDI process (subsection 3.2), we start the work by identifying and understanding the involved stakeholders in the project. As already explained, the main activities within the SSD project were focused on four different districts. However, the intermediate outcomes were adjusted in discussions with all other involved districts in the SSD districts network. The enterprise view depicts interests, conflicts of interests, requirements and the most important use cases and their clients. By holding discussions in different districts and with different stakeholders, local partners and citizens, common challenges were examined. Interestingly, it was discovered that almost all the districts involved in SSD are facing very similar barriers. Although, these barriers are specific and different from one place to another in terms of national regulations, they are in general very similar especially with respect to the technical side.

On the data side, looking at existing challenges proves that there is a huge need for a standard solution that can address the existing issues and be capable of covering future requirements. Despite the differences in the various use cases chosen in SSD districts and the necessary techniques for carrying out the computation and analysis, more than half of the data required for the calculations are common amongst use cases. This applies both in domain and in cross-domains applications that show the necessity of considering these different use cases with regard to each other. This of course can be done by providing a comprehensive data infrastructure. (PERCIVALL et al. 2015) have called it in their OGC white paper "a framework of trusted/authoritative data"; for example, core reference data in 2D and 3D (i.e. topography), identifiers and addressing, smart infrastructure (BIM, smart grid), sensor feeds, etc. to build the backbone of the Smart City framework. A Smart city also needs to be open to different "data types, such as volunteered, unstructured and linked data. Such a framework needs a robust data integration platform" (PERCIVALL et al. 2015).

SDDI is designed considering the requirements of the districts as well as aiming at providing a framework which offers a bottom-up approach for the effective integration of solutions. The key characteristics of the SDDI framework are as follows:

- Redundancy avoidance: In many cases there are datasets describing or related to a specific object. This object is often defined differently in different sources or by various providers. This leads to ambiguity and redundancy of the data which need to be interpreted later. For example, applications such as energy simulation, pedestrian flow simulation, applications involving real-time sensor observations in buildings all require to work with information about the districts' buildings which might be respresented redundantly within each of the applications. In order to avoid data redundancy, standards play a crucial role. SDDI is designed based on standards from OGC and ISO. For example

CityGML can be used to represent buildings just once for all of the applications mentioned above.

- Well-specified data semantics: The challenging point here is that the data are often interpreted differently. This leads to the misuse of the data over time by different users. It is crucial to use data models which present meaningful information understandable by everyone, and therefore a well-specified data semantic is needed. Standards from ISO or OGC are good examples for this characteristic and are considered in the SDDI model.

- Virtual District Model: The two aspects "redundancy avoidance" and "well-specified data semantics" are addressed by introducing a virtual district model (VDM). The VDM contains objects such as buildings, roads, city furniture, water bodies, etc. in addition to networks such as water utility, smart grid or transportation networks. (PERCIVALL et al. 2015) argue that space is a principle method to organize the Smart City. From our point of view space (coordinates, geometry) is not the only method but semantic objects (with spatial properties) – as they are provided by the VDM – should be used as a common denominator for representing and organizing the information pieces from the various application domains of the Smart Districts. Our detailed analyses of the SSD deep dive districts clearly show that nearly all thematic and sensor information are directly related to the objects of the VDM. Some sensors are even measuring properties of the real world objects (e.g. Smart Meters are measuring the power consumption of buildings). Hence, linking the sensors with the respective building objects and properties implicitly specifies the semantics of the sensor observations.

  The VDM is a response to what is mostly missing in the data management of today's other Smart City initiatives. This is the management of the data through a common digital model of the physical urban environment as the information hub. This can be seen, for example, in IoT and Big Data analytics centered Smart City concepts where obviously the concept of linking the devices to a common data hub is lacking. Based on the experiences gained in the SSD project and through the work with various districts, we can conclude that for almost all cases the districts need to work with or refer to district objects in one way or another. These objects are defined regarding their locations and their physical characteristics in the real world. Hence, it is necessary to have a virtual model of these physical elements of the area – whether it is just for a district or the entire city. Above, the VDM is also key to diverse types of simulations (e.g. energy, traffic, and environmental simulations) and to the estimation of the impacts of planned changes to the district.

- Interoperability: According to ISO 2382-1 (c.f. ISO/IEC 10746-2:2009), the term interoperability is defined as "the capability to communicate, execute programs, or transfer data among various functional units in a manner that requires the user to have little or no knowledge of the unique characteristics of those units". Interoperability is one of the most important characteristics of this model which covers the semantic and syntactic interoperability. This in fact is the key role of the SDDI which overcomes obstacles such as institutional barriers and avoids vendor lock-in, thus providing openness for extensions, and leading to the sharing of information.

- Extensibility: The realization of the SDDI as a modular, open, and interoperable set of distributed functional components ensures the easy extensibility by new stakeholders, users, sensors, thematic information, and analysis tools. Furthermore, the model should not be stopped at the current development, as technologies are rapidly developing. The structure of SDDI is designed in a way which can be extended in order to meet the future needs and cases.

- Functionality: A standard solution ensures the functionality of the approach and model apart from the use cases. This means that the model is designed such that it can be used for different use cases.
- Transferability: What makes SDDI powerful is that this platform is not developed only to be implemented for one use case or one district but to be implemented in different places in similar ways. This characteristic of SDDI again is due to the extensive use of standards in this infrastructure. For example, there are many cities in the world, which have already developed the 3D model of their cities following the OGC CityGML standard.

3.4 SDDI components

The proposed "SDDI" has five main tiers, each of which is shown in Fig. 2 and briefly explained here:

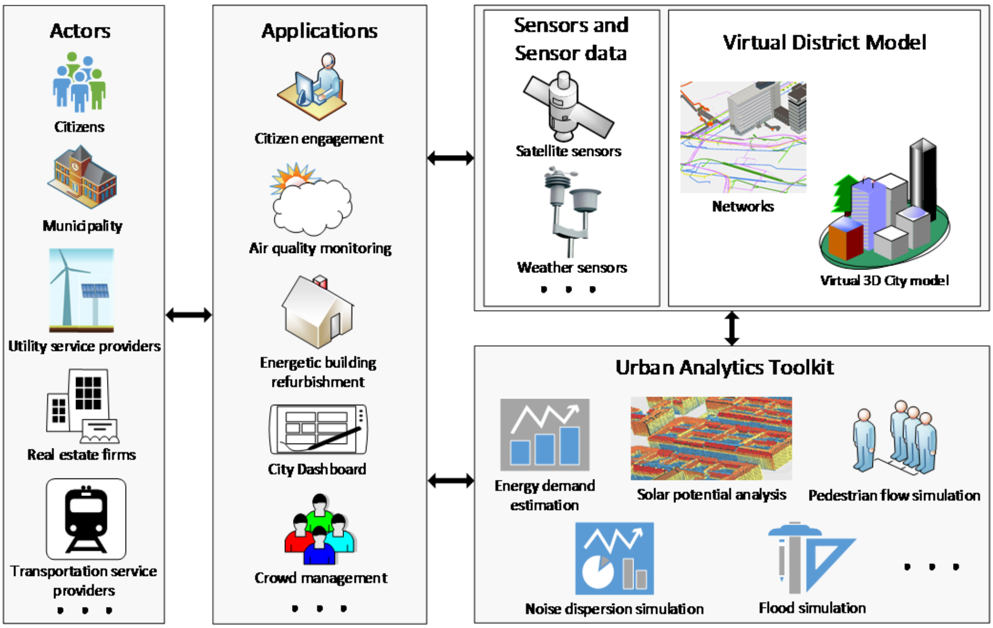

Fig. 2: Smart District Data Infrastructure (SDDI) concept.

i. **Actors**

"Actors" are all end users such as citizens, municipalities, utility and transportation service providers, real estate firms etc. This is the group of people and organizations, i.e. the stakeholders, for which the SDDI is offering services.

ii. Applications

All actors interface with the SDDI by applications, which implement the application logics according to specific tasks; they also provide the user interface to the actors and they make use of 1) sensors, 2) the VDM, 3) the processing and computation tools provided by the Urban Analytics Toolkit.

iii. Urban analytics toolkit

This includes all modelling, analysis and simulation tools which are interfaced to the SDDI using standardized service interfaces. This can cover a huge warehouse of existing and developing tools which is independent of where each tool will be deployed. This empowers the transferability and replicability of the whole SDDI as well as bringing a competition between various developers to improve the quality and performance of their analytical tools.

iv. Virtual District Model

The VDM consists of three parts:

- 3D spatio-semantic model: It represents both spatial characteristics of the district's physical objects and semantic information. The latter defines functions, thematic properties and characteristics of the district's objects and the interrelationship between them. This model represents each real world object of the districts by a unique virtual object (a so-called 'digital twin') and each of these modelled objects has a unique and stable identifier used for unambiguously referring to the object. This 3D model may include buildings, roads, water bodies, etc. Each of these objects can be further used as a reference object to which other information and devices in the district such as sensors can be linked (MOSHREFZADEH & KOLBE 2016). It is proposed to base this spatio-semantic model on the Open Geospatial Consortium (OGC) CityGML standard (GRÖGER et al. 2012).
- The network model: It defines functional behaviours, and resources and their flows, e.g., transportation, energy, water, or communication networks. Such network elements correspond to the 3D spatio-semantic model. An option for modelling the network elements is to use the utility network ADE for CityGML (KUTZNER & KOLBE 2016). However, it has to be mentioned that this model is not yet fully complete and needs to be further developed.
- The visualisation models: In order to illustrate the 3D spatio-semantic model of the district for rendering the district objects or 2D maps. In general, the visualisation model is useful for many purposes such as presentation, communication, interaction, etc. The visualisation model should be automatically derived from the semantic 3D model or (at least) be closely linked to it.

v. Sensors

Nowadays, sensors and IoT devices play a very important role in developing smart city applications. These devices may be air quality sensors, weather stations (e.g. monitoring temperature, humidity etc.) or even smart meters (measuring real-time electricity consumption). However, these sensors in general belong to different stakeholders with different rights and interests. They may also belong to different platforms, which can be either open or proprietary. In order to integrate diverse sensors and IoT devices with city information models within one operational framework, interoperability plays an important role in ensuring different components from different vendors can work together. The OGC already provides the Sensor Web Enablement (SWE) standards suite (OGC SWE 2015) for realizing interoperable sensor web infrastructures. SDDI focuses on providing interoperability of sensors using OGC SWE. Within the OGC SWE standards suite, sensor descriptions are encoded in OGC SensorML format and sensor observations in OGC O&M format. The web services such as the Sensor Observation Service (OGC SOS 2012) and the SensorThings API (LIANG et al. 2016) allow retrieval of sensor descriptions and observations using different requests.

4 Case study

This section presents a case study for one of the districts within the SSD project in which the very first implementation of SDDI was performed. The following sub-sections provide details of the overall process of implementing the SDDI starting from gathering data requirements to the running implementations in Queen Elizabeth Olympic Park (QEOP) in London.

4.1 General description

Queen Elizabeth Olympic Park (QEOP) is a sporting complex built for the 2012 Summer Olympics and is now one of the key locations in East London. The park is spread over 200 ha. QEOP was selected by Climate-KIC as one of the first districts to start working with their European-wide SSD programme to offer the opportunity to co-develop and demonstrate how new thinking, coupled with effective tools, technologies and policies, can lead to factor-4 improvement (twice the environmental impact for half the cost) in city district performance across a range of sustainability measures. The developed solutions are intended to be replicated city-wide and also in other districts.

4.2 Analysis of business needs

Within the scope of the SSD programme, four key themes were identified:

- Resource Efficient Buildings – Focusing initially on the iconic London Aquatics Centre and Copper Box Arena, this workstream is intended to create tools and approaches to enable low cost, low energy, low environmental impact management and maintenance of future ready buildings.
- Energy Systems – The energy systems workstream aims to create an efficient, smart, low carbon, resilient energy ecosystem, with specific focal points including optimization of district energy systems, community engagement and benefits and increased renewable energy generation.
- Smart Park / Future Living – Implementing user-facing digital and data solutions that deliver financial and CO_2 efficiencies, and prioritize quality of life improvements for those who live, work, and visit the Park.
- These are all underpinned by the fourth workstream: Data Architecture and Management – Implementing efficient and robust data management solutions that support the identification and trialing of innovative solutions and provide the foundation for improved

park operations, user experience and approaches that can be replicated by others, including the London Data Store.

4.3 Extracting and grouping data into SDDI components

Based on the identified workstreams, several workshops were conducted in London with the partners. For different use cases, the required data were identified and accordingly, it was mapped whether the data was available or not. During the listing of data items, it was found that many data items belong to the description of the physical reality in the QEOP district. It was also observed that many data items were required for more than one opportunity. Hence, the data was listed into following categories:

i. Data about the physical environment, such as GIS data, maps, or 3D models.

ii. Networks with stocks and flows, for example, transportation (multi modal), energy, water, or waste network nodes and connections.

iii. Sensors, including data about sensors (sensor description, location etc.) and data generated from sensors (observations).

iv. Thematic data on different domains, such as costs and performances, key performance indicators (KPIs), demands, consumptions, productions and generations.

v. Further data, such as weather, events and agents (such as people/visitors moving through the park, or simulated/virtual people).

The identification and listing of data items helped to classify them into major groups: (i) virtual district model (VDM), describing the data about the physical environment (ii) sensors, and (iii) other data, such as environment, weather, agents, events etc. (see Fig. 3).

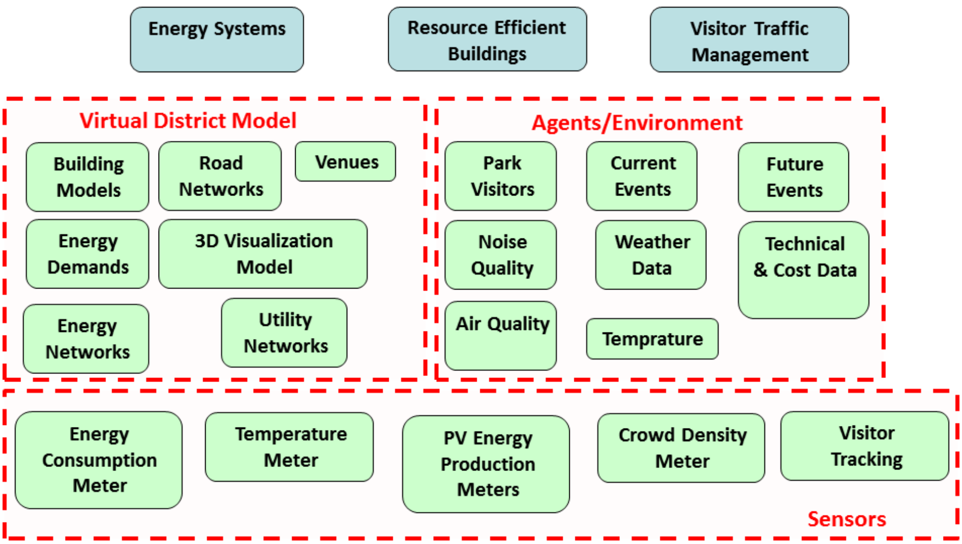

Fig. 3: Grouping of QEOP data items into categories, which can be related to SDDI components. Please note, that this figure just shows a simplified sketch. The complete documentation covers multiple large spreadsheets.

4.4 Linking the VDM with distributed data/sensors using web services

Fig. 4 illustrates a specific configuration of SDDI components, allowing the integration of distributed data items in a standardized way. The use case is the simulation of pedestrian crowd movement within the park.

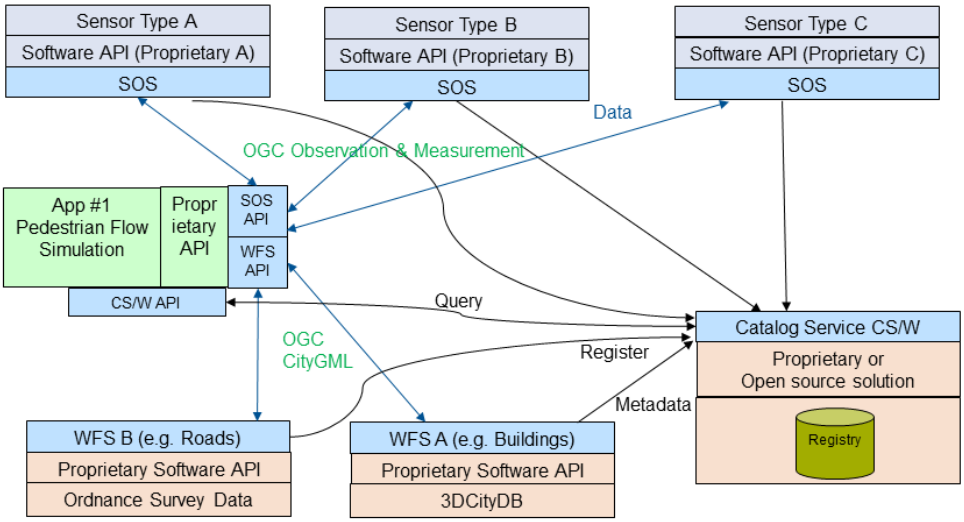

Fig. 4: Example configuration of the SDDI distributed architecture based on OGC web services. The use case is the simulation of pedestrian crowd movement within the park.

The architecture shows setting up following OGC web service interfaces on top of distributed data repositories, sensor devices, and the application program.

(i) Sensor Observation Service (OGC SOS 2012) – SOS facades are setup for proprietary sensor devices. It allows exporting sensor data in OGC O&M standard. This standard defines the semantics as well as the exchange format of the data.

(ii) Web Feature Service (OGC WFS 2014) – this OGC standard specifies the discovery, retrieval, and querying of object based data. The WFS uses GML application schemas to define the specific properties of each type of feature. In the SDDI the CityGML application schema is being used to define the semantics and exchange format of the objects, properties, and interrelationships of the VDM. Hence, objects like buildings, roads, etc. may be accessed utilizing the WFS. Similarly, Web Coverage Service (OGC WCS 2012) facades can also be set up. WCS is an OGC standard, which defines retrieval of field-based data, e.g. weather information.

(iii) Catalogue Service for the Web (OGC CS 2007) – The Catalogue Service for the Web interface (CS/W) facilitates registration of all datasets, services, and applications. It is responsible for managing and querying of all the metadata. It also allows automated harvesting of metadata from the registered services.

The user application is augmented by interoperable service interfaces to access data from different sources in a standardized way. It may include different interfaces for multiple services such as WFS, WCS, SOS, and CS/W, allowing to search / retrieve the data items.

4.5 Implementations

a. Virtual District Model

The VDM includes the topography information model of the QEOP, which currently consists of 3D building and road models. The semantic 3D building and road models of QEOP are represented in level of detail 2 (LoD2) according to the CityGML Building and Transportation modules respectively. In order to provide the CityGML LoD2 data, the following datasets were used:

- Digital Surface Model (DSM), with a grid resolution of 50 cm
- Digital Terrain Model (DTM) with a resolution of 50 cm
- 2D Building footprints from Ordnance Survey (OS) Master Map
- Building and Road address data from OS AddressBase Premium
- Imagery data from OS MasterMap Imagery

The QEOP data contains 3708 objects and includes the original (and stable) object ID values from Ordnance Survey called 'TOID'.

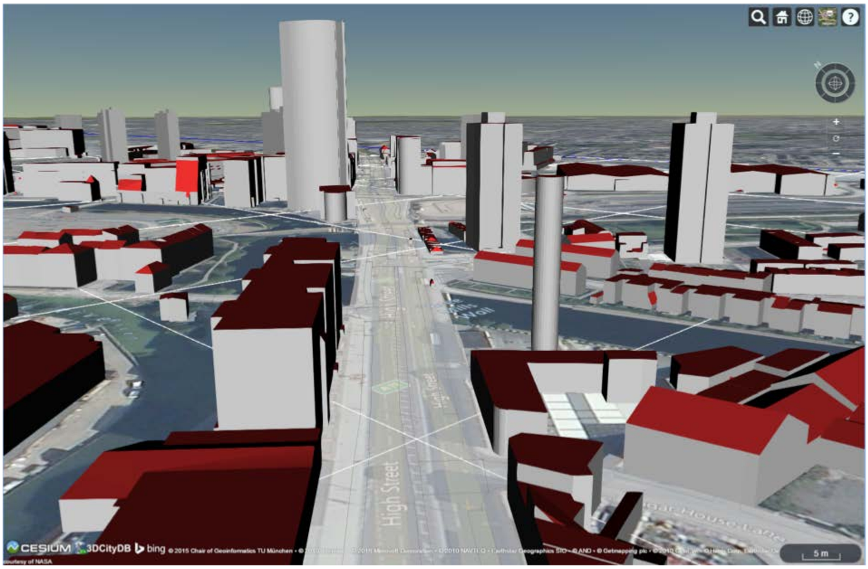

Fig. 5: Virtual District Model of QEOP (Screenshot taken from 3DCityDB Web Client developed by TUM (CHATURVEDI et al. 2015))

b. OGC Web Feature Service

The topography information model for the QEOP, which currently includes 3D buildings and roads, is based on the OGC CityGML standard. The company 'virtualcitySYSTEMS GmbH' offers a Web Feature Service implementation (virtualcityWFS[6]) allowing web-based access to the 3D city objects stored in the 3D city database (3DCITYDB 2016). WFS clients can directly connect to this interface and retrieve 3D content for a wide variety of purposes with the help of operations such as *GetCapabilities*, *DescribeFeature*, and *GetFeature*.

c. OGC Sensor Observation Service

Within the project, the real-time sensor observations from different weather stations set up and operated by the 'Intel Labs London' are being used. The weather stations are named as ICRI_0001, ICRI_0002, and ICRI_0003 and measure in total 15 properties in the park including temperature, humidity, wind speed etc. The observations are recorded for every minute. They are accessed from Intel's platform via a Hypercat registry and encoded using the SenML format. Hypercat[7] and SenML are industry IoT standards that are independently being developed from the OGC SWE standards. In order to integrate the sensors into the SDDI, SOS facades were developed for the sensors for retrieving sensor observations and sensor descriptions in a unified interoperable way. For setting up the SOS, the open-source implementation from 52° North (52°NORTH SWE 2016) was used. We are currently developing also SOS facades for Smart Meters of selected buildings, which use a proprietary application programming interface (API).

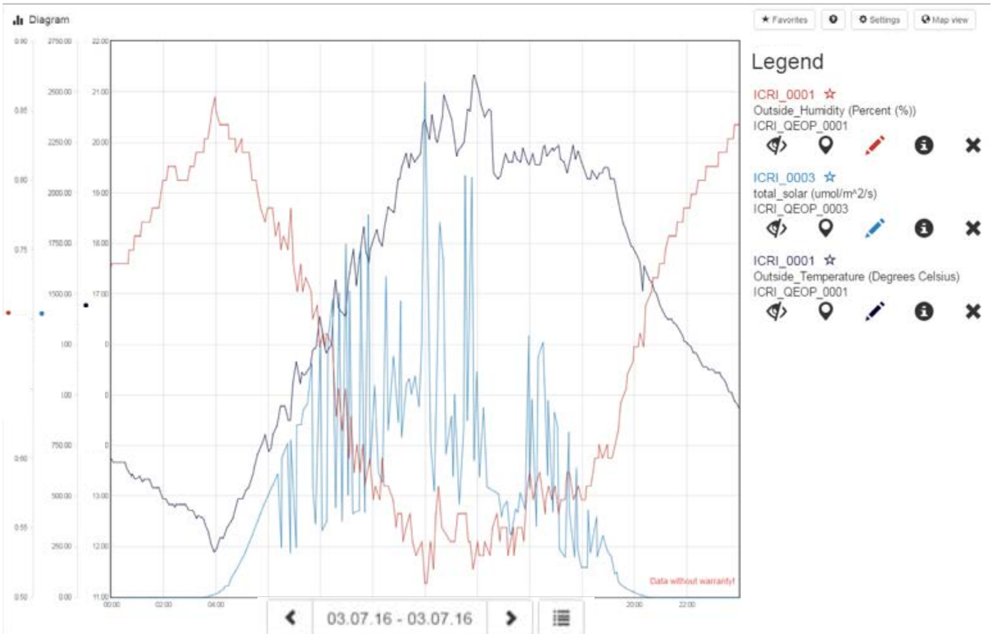

Fig. 6: Timeseries graph visualization of real-time sensor observations from London QEOP (screenshot taken from 52° North SOS Client). The graph shows the outside temperature, humidity, and solar irradiation on the 3rd of July 2016 of two stations overlayed within the same view.

d. OGC Catalogue Service for the Web

CS/W defines common interfaces to discover, browse, and query metadata about data, applications, services, and other potential resources within the SDDI. This allows the registration of different web services such as the Web Feature Service (WFS) that is managing the Virtual District Model in CityGML and the Sensor Observation Service (SOS) that is providing access to Intel's sensor devices and (in the future) other sensors in the park. Once the WFS/SOS is registered, the CS/W can be configured for automatic regular harvesting of the information contents of the WFS and SOS in order to create the proper metadata for the catalogue. Hence, if new data is added to the WFS, it will automatically become visible in the CS/W later. Similarly, if new sensors will be added to the SOS or new sensor data is available it will automatically become visible in the CS/W later. The Catalogue Service for the Web (CS/W) has been set up using the GeoNetwork[8] open-source project.

5 Conclusion and outlook

SDDI is becoming a major resource for access to geospatial data and services partnerships between the different stakeholders and partners of Deep Dive Districts in the Smart Sustainable Districts project. The SDDI is contributing to sound decision making and better services. It is necessary for energy planning and simulation, urban planning, and environmental protection. In order for the SDDI to be open, vendor neutral, and extensible, it has to be interoperable. Interoperability can only be achieved through consistent and structured usage of interface and encoding standards. For this purpose, SDDI utilizes a large suite of OGC standards. Moreover, the SDDI is a process-based initiative emphasizing partnerships and multi sectoral collaboration among different district partners. The SDDI is composed in a bottom-up approach which reflects the different aspirations of various stakeholders. One key aspect of the SDDI is the use of the Virtual District Model (VDM) as an information hub for all the district data related to the physical objects of the district. This distinguishes the SDDI from a general Spatial Data Infrastructure (SDI). The VDM is also used to visualize and analyze the current situation as well as planned changes within the city district.

In order to meet today’s information needs, the SDDI is conceptually laid out as a service oriented architecture (SOA). In the development of the SDDI framework we recognize and acknowledge the diversity and heterogeneity of the various stakeholders. We tackle this challenge by achieving a high degree of standardization and uniformity while being flexible and open for extensions. The SDDI is not just an information architecture, but also comprises the process to determine the requirements of the different districts using the reference model defined by ISO 10746 “Information technology – Open Distributed Processing – Reference model (ODP-RM)”.

The SDDI development strategy was driven and examined through real use cases (called "districts opportunities" in the SSD context) from four districts in Europe. Examining data sharing strategy and spatial-data requirements for different districts in different countries provides a proof of concept for the applicability of the SDDI development strategy, particularly from an organizational perspective for several reasons. Firstly, the requirements are defined based on the information needs of high priority tasks of the districts. This allowed us to learn

[8] http://geonetwork-opensource.org

from the state of the art of the practice of different stakeholders who are responsible for managing the different districts and who best know their tasks, workflows and develop generic tools that meet their needs. Secondly, the implementation of the SDDI in different districts allows the asking of the "how" and "why" research questions and investigation of the nature and complexity of spatial data sharing partnerships.

One important lesson learned through the SDDI development process is that SDDI does not give a 'turn-key software installation' that can simply be installed in exactly the same way for all districts. It is a framework that considers the different districts' needs in the development and configuration process. SDDI is a methodology that aims to provide techniques to determine requirements, analyze these requirements and identify the relevant data (spatial and non-spatial) and analysis functionalities that are needed to meet the determined requirements. Moreover, the SDDI implementations within the four deep dive districts are used as best practice for other districts showing how to setup the different IT components and connect applications and tools to the different web services.

A crucial issue for the SDDI is to develop a methodology to deal with the major dilemma of how to provide the needed stability and sustainability in the development, adaption, and not ignore potentially unstable and conflicting environment conditions. SDDI in its current status needs further investigation to institutionalize data maintenance issues and ICT sustainability that SDDI are based on, though. It is also a current challenge to ensure sustainable, long-term operation of the implemented SDDI prototypes, e.g. by operating different components or even the entire infrastructure by commercial providers. In this context, moving SDDI implementations to cloud computing platforms and considering the diverse data security aspects are major points for future research as well as the development of an appropriate business model to bring the SDDI method to the market.

Literature

3DCITYDB (2016), 3D City Database (3DCityDB) Homepage. http://www.3dcitydb.org/ (06.12.2016)

52°NORTH SWE (2016), 52N Sensor Web Community – Sensor Observation Service. http://52north.org/communities/sensorweb/sos/ (06.12.2016)

CHATURVEDI, K., YAO, Z. & KOLBE T. H. (2015), Web-based Exploration of and Interaction with Large and Deeply Structured Semantic 3D City Models using HTML5 and WebGL. In: KERSTEN, T. P. (ED.), Tagungsband der 35.Wissenschaftlich-Technischen Jahrestagung der DGPF, ISSN 0942-2870

CLIMATE-KIC SSD (2016), Smart Sustainable Districts Project Homepage, http://www.climate-kic.org/programmes/smart-sustainable-districts/ (09.11.2016)

DEGBELO, A. GRANELL, C., TRILLES, S., BHATTACHARYA, D., CASTELEYN, S., & KRAY, C. (2015), Opening up Smart Cities: Citizen-Centric Challenges and Opportunities from GIScience. ISPRS International Journal of Geo-Information, 5(2)

GRÖGER, G., KOLBE, T. H., NAGEL, C. & HÄFELE, K.-H. (2012), OpenGIS City Geography Markup Language (CityGML) Encoding Standard – Version 2.0.0. Open Geospatial Consortium, OGC Doc. No. 12-019

ISO/IEC 10746-2 (2009), Information Technology – Open Distributed Processing – Reference Model: Foundations – Part 2.

ISO/IEC 10746-3 (2009), Information Technology – Open distributed processing – Reference model: Architecture – Part 3.

ISO/IEC JTC 1 (2014), Smart Cities – Preliminary Report 2014, Information technology, http://www.iso.org/iso/smart_cities_report-jtc1.pdf

KUTZNER, T. & KOLBE, T. H. (2016), Extending Semantic 3D City Models by Supply and Disposal Networks for Analysing the Urban Supply Situation. In: KERSTEN, T. P. (Ed.), Tagungsband zur Dreiländertagung der SGPF, DGPF und OVG in Bern, Publikationen der DGPF, Band 25, ISSN 0942-2870

LIANG, S., HUANG, C.-Y. & KHALAFBEIGI, T. (2016), OGC SensorThings API Part 1: Sensing, Version 1.0, Open Geospatial Consortium, OGC. Doc. No. 15-078r6, http://www.opengeospatial.org/standards/sensorthings

MONZON, A. (2015), Smart Cities Concept and Challenges: Bases for the Assessment of Smart City Projects. In: Helfert, M., Krempels, K., Klein, C., Donnellan, B., Gusikhin, O. (eds.), International Conference on Smart Cities and Green ICT Systems, Springer

MOSHREFZADEH, M. & KOLBE, T. H. (2016), Smart Data Infrastructure for Smart and Sustainable Cities. In: Proc. of the 3th Int. Conference on Design & Decision Support Systems in Architecture and Urban Planning, Eindhoven, The Netherlands, pp. 328 – 343

OGC CS (2007), OpenGIS Catalogue Service Specification, Version 2.0.2, Open Geospatial Consortium, OGC Doc. No. 07-006r1, http://www.opengeospatial.org/standards/cat

OGC SOS (2012), OGC Sensor Observation Service 2.0 Interface Standard, Open Geospatial Consortium, OGC Doc. No. 12-006, http://www.opengeospatial.org/standards/sos

OGC SWE (2015), OGC Sensor Web Enablement Standards, Open Geospatial Consortium, http://www.opengeospatial.org/ogc/markets-technologies/swe

OGC WCS (2012), Web Coverage Service Interface Standard - Core. Open Geospatial Consortium, OGC Doc. No. 09-110r4, http://www.opengeospatial.org/standards/wcs

OGC WFS (2014), OGC Web Feature Service 2.0 Interface Standard, Open Geospatial Consortium, OGC Doc. No. 09-025r2. http://www.opengeospatial.org/standards/wfs

OXFORD DICTIONARY (2016), Definition of “District”. https://en.oxforddictionaries.com/definition/district. (14.11.2016)

PERCIVALL, G., RÖNSDORF, C., LIANG, S., MCKENZIE, D. & MCKEE, L. (2015), OGC Smart Cities Spatial Information Framework, Open Geospatial Consortium, OGC Doc. No. 14-115

SÁNCHEZ, L., ELICEGUI, I., CUESTA, J., MUÑOZ, L. & LANZA, J. (2013), Integration of Utilities Infrastructures in a Future Internet Enabled Smart City Framework, In: Sensors, 13(11), pp. 14438-14465.

UNITED NATIONS (2014), World Urbanization Prospects. https://esa.un.org/unpd/wup/

UNITED NATIONS (2016), Sustainable Cities, why they matter. http://www.un.org/sustainabledevelopment/wp-content/uploads/2016/08/16-00055K_Why-it-Matters_Goal-11_Cities_2p.pdf

VAN ALSTYNE, M., BRYNJOLFSSON, E. & MADNICK, S. (1995), Why not one big database? Principles for data ownership. In: Decision Support Systems, Vol. 15, pp. 267 – 284.

VON WEIZSÄCKER, E. U., LOVINS, A. B., LOVINS & L. H. (1998), Factor four: doubling wealth-halving resource use: the new report to the Club of Rome. Earthscan