AGGREGATION OF SEMANTIC SENSOR DATA

Graduation proposal

by

Ivo de Liefde

November 23, 2015

CONTENTS

ACRONYMS

API application program interface . . . 3
IoT internet of things . . . 1
JSON javascript object notation . . . 3
μSDI micro spatial data infrastructure . . . 2
OGC open geospatial consortium . . . 1
O&M observations and measurements . . . 4
OSM open streetmap . . . 3
OWL web ontology language . . . 7
RDF resource description framework . . . 3
REST representational state transfer . . . 2
SensorML sensor modelling language . . . 3
SEL semantic enablement layer . . . 4
SOS sensor observation service . . . 2
SSN semantic sensor network . . . 1
SSW semantic sensor web . . . 1
SWE sensor web enablement . . . 1
W3C world wide web consortium . . . 1
XML extensible markup language . . . 3

1 INTRODUCTION

This document should include:

- motivation / problem field /relevance
- position in the academic and professional debate
- problem statement, objectives, research questions
- approach, theoretical framework, methodology
- references
- preliminary project set up and results

an introduction in which the relevance of the project and its place in the context of geomatics is described, along with a clearly-defined problem statement

1.1 BACKGROUND

Egenhofer (2002) argues ”there would exist a much higher potential for exploiting the Web if tools were available that better match human reasoning” (Egenhofer, 2002, p. 1). Sheth et al. (2008) proposed adding semantics to the sensor web which has resulted in the so-called semantic sensor web (SSW). The sensor web enablement (SWE) standards by the open geospatial consortium (OGC) have been combined with ontologies in the SSW to add meaning to sensor data (Pschorr, 2013; Henson et al., 2009). The world wide web consortium (W3C) has contributed to this development by creating the semantic sensor network (SSN) ontology (Compton et al., 2012).

Smart cities can be defined as ”enhanced city systems which use data and technology to achieve integrated management and interoperability” (Moir et al., 2014, p. 18). Research on smart cities has identified a number of potential uses for semantic sensor data. Often this is presented in the context of the internet of things (IoT) (Zanella et al., 2014; Wang et al., 2015a). Fell (2014) describes the IoT as ”the result of technological progress in many parallel and often overlapping fields, including those of embedded systems, ubiquitous and pervasive computing, mobile telephony, telemetry and machine-to-machine communication, wireless sensor networks, mobile computing, and computer networking” (Fell, 2014, p. 11). Wang et al. (2015a) argues that sensor data fusion plays an important role in the IoT. He defines sensor data fusion as ”a data processing technique that associates, combines, aggregates, and integrates data from different sources” (Wang et al., 2015a, p. 2).

To keep up with the developments of the IoT and the smart city OGC has identified the role of their standards in the ’smart cities spatial information framework’ (Percivall, 2015). This is in line with the smart city vision of

Townsend (2013) for creating ”an organically evolved set of open standards and software that anyone can build on” (Townsend, 2013, p. 290).

1.2 PROBLEM STATEMENT

The vision of the IoT and the smart city is very appealing. However, there are still a number of missing links considering the role of the SSW. First of all, current SSW research focusses mainly on sensor data from a single source. It is therefore difficult to discover sensor data on the internet that can be queried using the SWE standards. Second of all, data from different sources cannot easily be shared, integrated, combined and aggregated for data fusion (Wang et al., 2015b; Ji et al., 2014; Corcho and Garcia-Castro, 2010).

To overcome the issues of sensor data discovery and integration Janowicz et al. (2013) suggest a linked data model and a representational state transfer (REST)ful proxy for OGC's sensor observation service (SOS). They mention future work should focus on the 'Sensor Plug&Play infrastructure' in which ”sensors can by automatically registered” (Janowicz et al., 2013, p. 21). Also, it is argued that these developments could lead to a micro spatial data infrastructure (μSDI) and will enable an ubiquitous geoweb.

1.3 USE CASE

Identifying urban heat islands in cities. For this temperature data is required together with statistics, based on addresses.

1.4 RESEARCH QUESTION

This thesis aims to gain knowledge on how to move the SSW closer to an ubiquitous geoweb by answering the following research question:

How can the semantic sensor web improve the discovery, integration and aggregation of distributed sensor data?

2 RELATED WORK

a related work section in which the relevant literature is presented and linked to the project;

2.1 SENSOR WEB

OGC SWE standards such as SOS and sensor modelling language (SensorML) (Botts et al., 2007), (Botts et al., 2008)

Different data formats: extensible markup language (XML) (SWE), EXI (W3C) and javascript object notation (JSON) SensorThings application program interface (API) (Zanella et al., 2014)

OGC has been working on standards for geodata discovery that allow users to find datasets based on geographic and temporal properties (Goncalves, 2014). It is available as an extension for existing formats like RSS and Atom. However, it does not mention support for sensor data discovery.

Sensor metadata visualisation has been studied by Yoo (2014) in order to improve the discovery of sensor data. How to integrate or combine different datasets is not addressed in this approach.

2.2 LINKED DATA

Linked Data (Berners-Lee et al., 2001)

Web of Data Bizer et al. (2009)

Research by Auer et al. (2009) under the name of 'linked geodata' shows how to convert open streetmap (OSM) data into resource description framework (RDF) triplets and publish it in the semantic web. The linked OSM data used to be accessible online at `http://linkedgeodata.org/OnlineAccess`. However, this is no longer the case. They do offer a manual and open source software for publishing OSM data as linked data at `https://github.com/GeoKnow/LinkedGeoData`. Anyone interested in using linked OSM data can refer to this for their own implementation.

Publishing geodata as RDF and mapping on-the-fly (Missier, 2015)

2.3 SEMANTIC SENSOR WEB

SSW (Sheth et al., 2008), (de Mel et al., 2011), (Bakillah et al., 2013)

Huang and Javed (2008) add semantics to sensor data, but do not use the SWE standards by OGC.

Extending RDF with the ability to represent spatial and temporal data (Koubarakis and Kyzirakos, 2010)

Janowicz et al. (2010) argue that the current SWE standards are not sufficient for dealing with sensor data on the semantic web. A transparent semantic enablement layer (SEL) for OGC services is proposed that replaces the observations and measurements (O&M) and SensorML standards. This research formed the basis for the development of the W3C SSN ontology (Compton et al., 2012).

Henson et al. (2009) and Pschorr (2013) present the SemSOS architecture which adds semantics to sensor data and publishes the semantically enriched sensor data via SOS and the O&M encoding. However, in SemSOS the semantics are not being used to integrate or combine data from different sources, but rather to enrich data from a single source.

Applications using OGC SOS and W3C SSN (Keßler and Janowicz, 2010), (Barnaghi et al., 2010)

Three layer model: the sensor data source layer, the data integration layer and the application layer Wang et al. (2015b).

2.4 INTERNET OF THINGS

More and more devices connected to the internet. Also a growing amount of research on using sensors of smart devices. (Waher, 2015), (Žarko et al., 2015), (Calbimonte et al., 2011)

Research on connecting smart devices to SSW (de Vera et al., 2014)

OpenIoT platform (Calbimonte et al., 2014)

2.5 SMART CITIES

The role of sensors in smart cities (Zanella et al., 2014)

The role of OGC standards in smart cities (Percivall, 2015)

3 RESEARCH OBJECTIVES

the research objectives and/or research questions are clearly defined, along with the scope (ie what you will not be doing);

3.1 RESEARCH QUESTION

How can the semantic sensor web improve the discovery, integration and aggregation of distributed sensor data?

3.2 OBJECTIVE

Develop a method to semantically link sensor metadata to real world objects for spatial, semantic and temporal data aggregation.

I would like to bridge the efforts by Henson et al. (2009), Pschorr (2013), of adding semantics to SOS, and the efforts by Auer et al. (2009) of adding semantics to OSM data using the SSN ontology proposed by Compton et al. (2012) in order to improve the discovery, integration and aggregation of sensor data from different sources.

The thesis research should result in a prototype implementation which consists of two parts:

1. Firstly, an application that takes locations (HTTP addresses) of SOS servers as input and automatically links them to the OSM data. It results in a(n extended) mapping of the sensor web that will be used by the aggregation queries in the second part of the implementation. However, these mappings between sensors and objects are also freely accessible on the (semantic) web which improves the discovery of sensors for other applications.

2. Secondly, an application that allows users to query aggregated sensor data from different sources. This takes an OSM feature and a time interval as input, optionally with other spatial/temporal parameters (like a value for a buffer operation or a time interval to aggregate on). It returns a set of aggregated sensor data.

3.3 SCOPE

focus on OGC's SWE standards / SOS and W3C's SSN ontology. Not going into evaluation of different standards. Not specifically focussing on smart devices, but on SWE enabled sensors.

4 METHODS

overview of the methodology to be used;

4.1 STANDARDS

4.1.1 Sensor Web Enablement

SOS and SSN

4.1.2 Semantic Web

- Store OSM data: create RDF on-the-fly to prevent double storage.
- Retrieve sensor metadata: create RDF from O&M which is returned by SOS getCapabilities/describeSensor requests
- Use SSN as ontology for semantic sensor metadata
- Use web ontology language (OWL) as language for storing semantic RDF triples

Query metadata: SPARQL, geo-SPARQL or stSPARQL? and which query engine?

4.2 ONTOLOGIES

The O&M ontology is retrieved from sensors
Use SSN ontology ontology to store metadata from sensors.

4.3 MIDDLEWARE

Creating own middleware to link sensors semantically and retrieve aggregated data

RESTful service

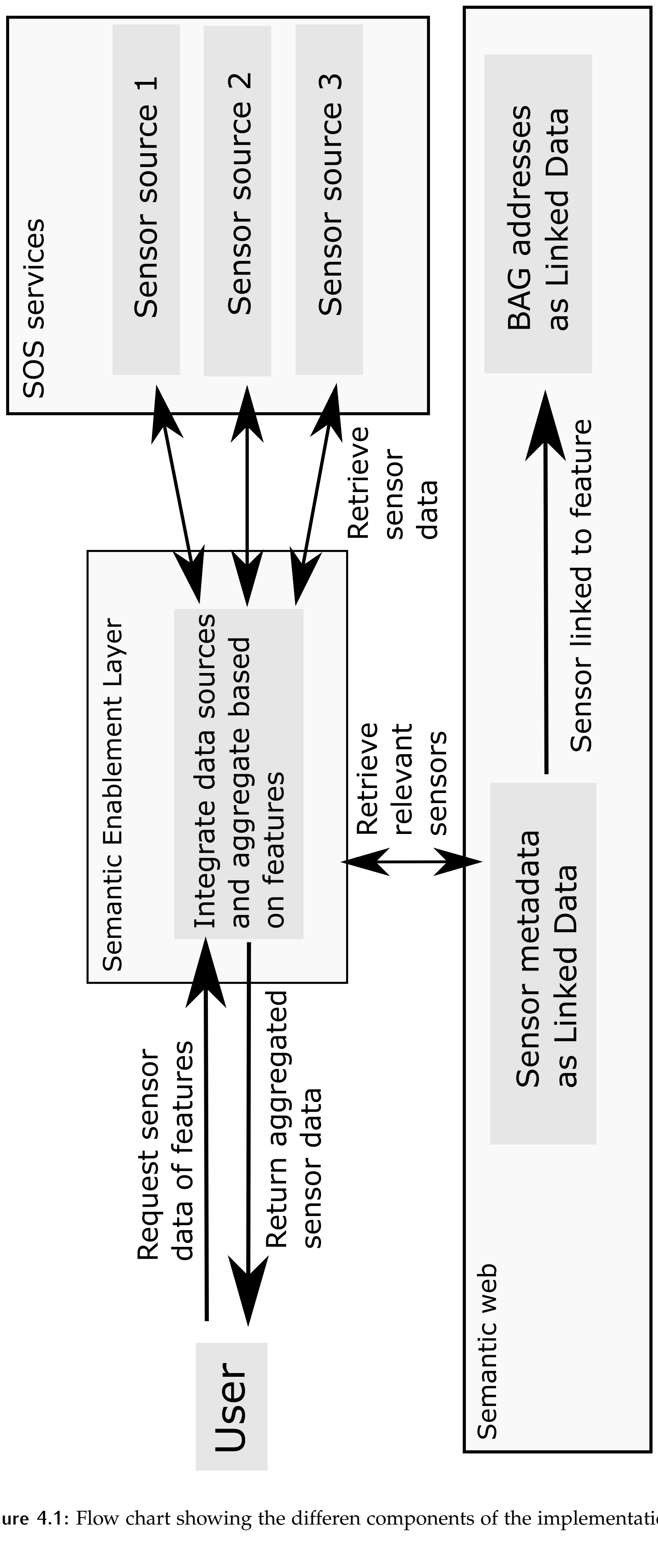

Figure 4.1: Flow chart showing the differen components of the implementation

5 PLANNING

time planning—having a Gantt chart is probably a better idea then just a list;

5.1 GANTT CHART

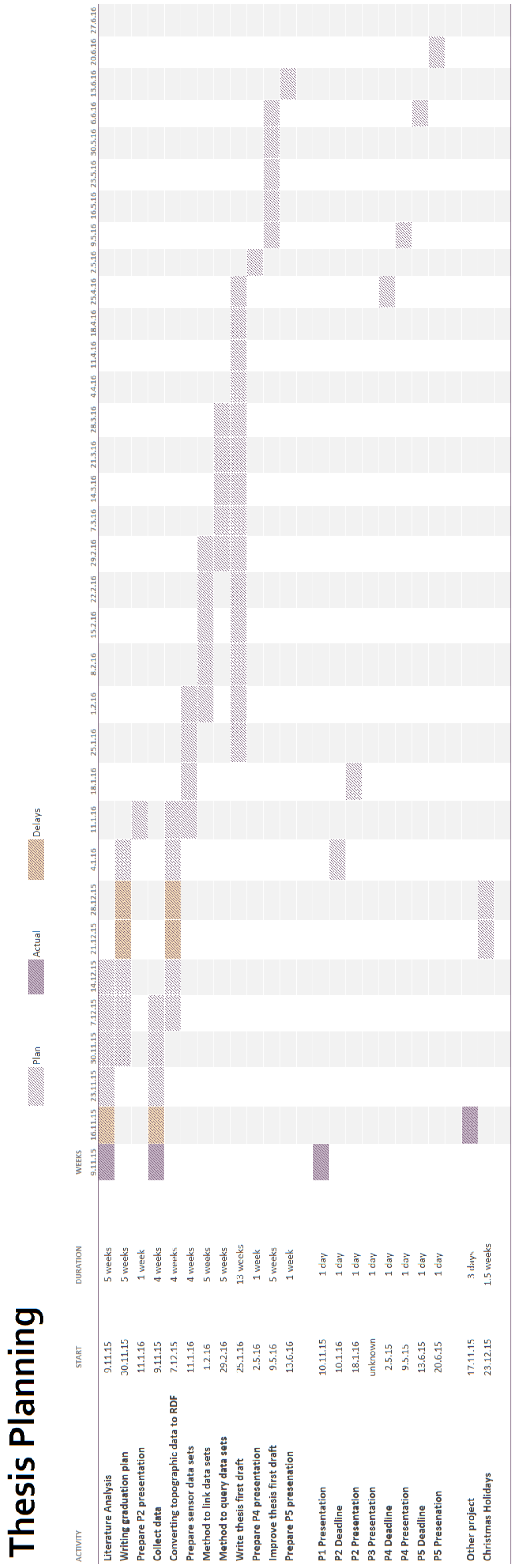

Figure 5.1: GANTT chart showing the planning of the thesis

6 TOOLS AND DATA

since specific data and tools have to be used, it's good to present these concretely, so that the mentors know that you have a grasp of all aspects of the project;

6.1 DATA

Topographic data
Sensor data (different sources)

6.2 DATABASE

Oracle?

6.3 SERVER

Or localhost?

6.4 PROTOTYPE

Use Python programming language and Flask for server sided scripts
Perhaps with Python's Request library for making POST or GET requests.

BIBLIOGRAPHY

Auer, S., Lehmann, J., and Hellmann, S. (2009). *Linkedgeodata: Adding a spatial dimension to the web of data*. Springer.

Bakillah, M., Liang, S. H., Zipf, A., and Mostafavi, M. A. (2013). A dynamic and context-aware semantic mediation service for discovering and fusion of heterogeneous sensor data. *Journal of Spatial Information Science*, 6:155–185.

Barnaghi, P., Presser, M., and Moessner, K. (2010). Publishing linked sensor data. In *CEUR Workshop Proceedings: Proceedings of the 3rd International Workshop on Semantic Sensor Networks (SSN), Organised in conjunction with the International Semantic Web Conference*, volume 668.

Berners-Lee, T., Hendler, J., Lassila, O., et al. (2001). The semantic web. *Scientific american*, 284(5):28–37.

Bizer, C., Heath, T., and Berners-Lee, T. (2009). Linked data-the story so far. *Semantic Services, Interoperability and Web Applications: Emerging Concepts*, pages 205–227.

Botts, M., Percivall, G., Reed, C., and Davidson, J. (2007). Ogc sensor web enablement: Overview and high level architecture. OGC document 06-021r1.

Botts, M., Percivall, G., Reed, C., and Davidson, J. (2008). Ogc sensor web enablement: Overview and high level architecture. In *GeoSensor networks*, pages 175–190. Springer.

Calbimonte, J.-P., Jeung, H., Corcho, O., and Aberer, K. (2011). Semantic sensor data search in a large-scale federated sensor network.

Calbimonte, J.-P., Sarni, S., Eberle, J., and Aberer, K. (2014). Xgsn: An open-source semantic sensing middleware for the web of things. In *7th International Workshop on Semantic Sensor Networks*, number EPFL-CONF-200926.

Compton, M., Barnaghi, P., Bermudez, L., GarcíA-Castro, R., Corcho, O., Cox, S., Graybeal, J., Hauswirth, M., Henson, C., Herzog, A., et al. (2012). The ssn ontology of the w3c semantic sensor network incubator group. *Web Semantics: Science, Services and Agents on the World Wide Web*, 17:25–32.

Corcho, O. and Garcia-Castro, R. (2010). Five challenges for the semantic sensor web. *Semantic Web-Interoperability, Usability, Applicability*, 1.1(2):121–125.

de Mel, G., Pham, T., Damarla, T., Vasconcelos, W., and Norman, T. (2011). Semantically enriched data for effective sensor data fusion. In *SPIE Defense, Security, and Sensing*, pages 80470L–80470L. International Society for Optics and Photonics.

de Vera, D. D. P., Izquierdo, l. S., Vercher, J. B., and Gomez, L. A. H. (2014). A ubiquitous sensor network platform for integrating smart devices into the semantic sensor web. *Sensors*, 14(6):10725–10752.

Egenhofer, M. J. (2002). Toward the semantic geospatial web. In *Proceedings of the 10th ACM international symposium on Advances in geographic information systems*, pages 1–4. ACM.

Fell, M. (2014). *Roadmap for the Emerging Internet of Things - Its Impact, Architecture and Future Governance*. Carre & Strauss, United Kingdom.

Goncalves, P. (2014). Ogc opensearch geo and time extensions. OGC Implementation Standard.

Henson, C., Pschorr, J. K., Sheth, A. P., Thirunarayan, K., et al. (2009). Semsos: Semantic sensor observation service. In *Collaborative Technologies and Systems, 2009. CTS'09. International Symposium on*, pages 44–53. IEEE.

Huang, V. and Javed, M. K. (2008). Semantic sensor information description and processing. In *Sensor Technologies and Applications, 2008. SENSORCOMM'08. Second International Conference on*, pages 456–461. IEEE.

Janowicz, K., Broring, A., Stasch, C., Schad, S., Everding, T., and Llaves, A. (2013). A restful proxy and data model for linked sensor data. *International Journal of Digital Earth*, 6(3):233–254.

Janowicz, K., Schade, S., Bröring, A., Keßler, C., Maué, P., and Stasch, C. (2010). Semantic enablement for spatial data infrastructures. *Transactions in GIS*, 14(2):111–129.

Ji, C., Liu, J., and Wang, X. (2014). A review for semantic sensor web research and applications. *Advanced Science and Technology Letters*, 48:31–36.

Keßler, C. and Janowicz, K. (2010). Linking sensor data-why, to what, and how? In *SSN*.

Koubarakis, M. and Kyzirakos, K. (2010). Modeling and querying metadata in the semantic sensor web: The model strdf and the query language stsparql. In *The semantic web: research and applications*, pages 425–439. Springer.

Missier, G. A. (2015). Towards a web application for viewing spatial linked open data of rotterdam. Master's thesis, Delft University of Technology.

Moir, E., Moonen, T., and Clark, G. (2014). What are future cities: Origins, meanings and uses.

Percivall, G. (2015). Ogc smart cities spatial information framework. OGC Internal reference number: 14-115.

Pschorr, J. K. (2013). Semsos: an architecture for query, insertion, and discovery for semantic sensor networks. Master's thesis, Wright State University.

Sheth, A., Henson, C., and Sahoo, S. S. (2008). Semantic sensor web. *IEEE Internet Computing*, 12(4):78–83.

Townsend, A. M. (2013). *Smart cities: big data, civic hackers, and the quest for a new utopia*. WW Norton & Company.

Waher, P. (2015). *Learning Internet of Things*. Packt Publishing Ltd.

Wang, M., Perera, C., Jayaraman, P. P., Zhang, M., Strazdins, P., and Ranjan, R. (2015a). City data fusion: Sensor data fusion in the internet of things.

Wang, X., Zhang, X., and Li, M. (2015b). A review of studies on semantic sensor web. *Advanced Science and Technology Letters*, 83:94–97.

Yoo, B. (2014). Visualization and level-of-detail of metadata for interactive exploration of sensor web. *International Journal of Digital Earth*, 7(11):847–869.

Zanella, A., Bui, N., Castellani, A., Vangelista, L., and Zorzi, M. (2014). Internet of things for smart cities. *Internet of Things Journal, IEEE*, 1(1):22–32.

Žarko, I. P., Hromic, H., Phuoc, D. L., Serrano, M., Antonic, A., Hayes, C., and Decker, S. (2015). Real time analysis of sensor data for the internet of things by means of clustering and event processing. In *Proc. of the IEEE International Conference of Communications (ICC2015)*. Hrvatska znanstvena bibliografija i MZOS-Svibor.

COLOPHON

This document was typeset using LaTeX. The document layout was generated using the `arsclassica` package by Lorenzo Pantieri, which is an adaption of the original `classicthesis` package from André Miede.