# A Lightweight Approach for the Sensor Observation Service to Share Environmental Data across Europe

Simon Jirka[1], Arne Bröring[1,4,5], Peter Christian Kjeld[2], Jon Maidens[3], Andreas Wytzisk[1]

[1]52°North Initiative for Geospatial Open Source Software GmbH, Germany

[2]European Environment Agency, Copenhagen, Danmark

[3]Atkins Danmark A/S, Copenhagen, Danmark

[4]Institute for Geoinformatics, University of Muenster, Germany

[5]Faculty ITC, University of Twente, Enschede, The Netherlands

## Abstract

The importance of near real-time access to environmental data has steadily increased over the last years. In this article, the European Environment Agency (EEA), which receives environmental data from a large number of providers, is in focus. The heterogeneous data formats and data transfer mechanisms make the data collection and integration a difficult task for the EEA. An approach is needed for facilitating the interoperable exchange of environmental data on a large scale. A core element of this approach is the Sensor Web Enablement (SWE) technology of the Open Geospatial Consortium (OGC) which allows the standardised, interoperable, vendor and domain independent exchange of sensor data. The main contribution of this article is a lightweight profile for the OGC Sensor Observation Service that ensures the necessary interoperability for seamlessly integrating the environmental data provided by the EEA's member states and thus forms the foundation for the developed data exchange mechanisms. This is complemented by information about the resulting Sensor Web architecture and the integration into the EEA's existing IT infrastructure. In summary, this paper shows a practical scenario in which SWE technology enables the exchange of near real-time environmental data on a large scale.

## 1 Introduction

Over the last years, the importance of near real-time access to environmental data has steadily increased. For organisations such as the European Environment Agency[1] (EEA), which receives environmental data from a large number of providers, an efficient exchange of data is challenging. In case of the EEA, huge amounts of environmental data are collected from various agencies of its member states. The large numbers of data formats (e.g., binary, text based, or different XML variants) and data access interfaces (e.g., FTP or proprietary web services) are causing huge integration efforts (Kjeld et al. 2011). As shown in Figure 1, to access a new data source, the EEA needs to *understand* the interface. Hence, for each new type of data source, the existing data brokering system of the EEA needs to be manually adapted. This is the key challenge underlying this research, to answer the question, of how to build a system that enables environmental data steams on a large scale across multi-organizational environments. Thereby, a data stream represents the transfer of environmental observations through a web based infrastructure from the source, here member state agencies, to the aggregating organization, here the EEA, as well as the forwarding to end users of environmental data.

[1] http://www.eea.europa.eu/

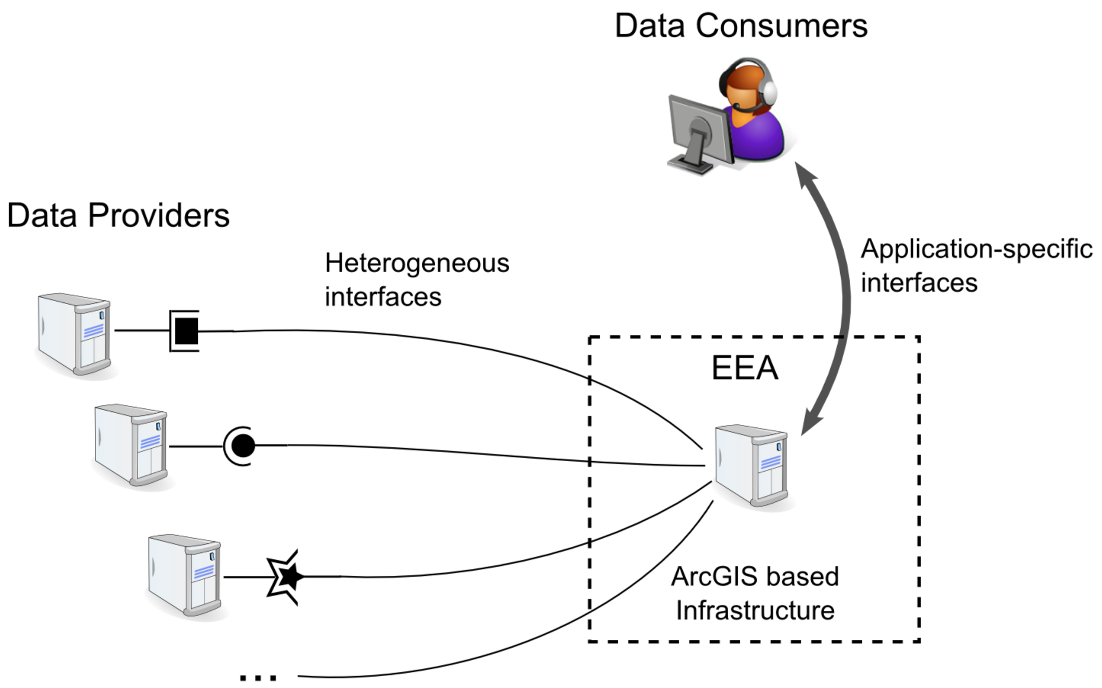


Figure 1 - Current Situation - Data Sets from Member States are provided through Heterogeneous Mechanisms

In the past, different standardization efforts and legal frameworks have been initiated to address this challenge for establishing interoperability in context of environmental data exchange. In Europe, the INSPIRE (European Commission 2007) initiative and the leveraged Open Geospatial Consortium (OGC) standards are of most importance. Hence, also this research builds up on those building blocks. The key element to facilitate the integration of multiple data providers is the Sensor Web Enablement (SWE) framework of standards (Bröring et al. 2011) governed by the OGC. These are used here, since they provide (1) openness and interoperability, (2) vendor independency, and (3) domain independency.

To answer the research question and establish SWE as a standard for a homogeneous exchange of near real-time environmental data across Europe, several challenges need to be addressed. A first step to enable interoperable flows of environmental data is the definition of profiles that reflect a common agreement how to apply the SWE standards. As the SWE standards are defined in a rather generic manner to support a very broad range of use cases and scenarios, a pragmatic approach that reduces their complexity is needed. On the one hand, such focused SWE profiles increase interoperability by avoiding ambiguous interpretations. On the other hand they reduce the necessary efforts for implementing a SWE infrastructure at the EEA as well as the member states. This resulted in a lightweight SOS profile which constitutes the main contribution of this paper.

In addition, two further aspects will be introduced. First, it is shown how SWE standards are integrated into legacy workflows, here, by looking at the use case of the EEA. Therefore, the existing ArcGIS Server (Bader 2005) environment needs to be adapted. Second, by providing re-usable software and documentation, the member state agencies receive support in implementing the SWE technology. These two aspects complement the lightweight SOS profile by facilitating its practical application.

In summary, this paper describes at hand of the EEA use case scenario how SWE technology can be applied to enable the exchange of near real-time data on a large scale and an approach is developed that answers the above defined research question. Challenges related to this approach are discussed and ways to address them are presented.

The remainder of this article is structured as follows. Section 2 introduces the relevant background of this work. Section 3 describes the underlying use cases and their associated requirements. The approach how to address the EEA's requirements by a lightweight SOS profile that forms the basis for the EEA's Sensor Web architecture is presented in Section 4. In Section 5 results are discussed. An outlook on remaining future work is given in Section 6. Finally, Section 7 draws conclusions from the conducted research.

# 2 Background

## 2.1 The European Environment Agency

The European Environment Agency (EEA) is an agency of the European Union (EU) with currently 32 member states (European Environment Agency 2011). The EEA's main task is to provide policy makers, stakeholders as well as the general public with independent information about the environment.

A core element of the EEA's work is *Eionet*[2], the European environment information and observation network (European Environment Agency 2005). Eionet serves as a partnership network for providing timely and quality-assured environmental data to the member states as well as cooperating countries. Thus, a central task of the EEA is to collect data from its member states, perform quality-assurance, and provide access to these data sets. For achieving this goal, the EEA is in close cooperation with so called National Focal Points (NFPs) in the member states (i.e., national environment agencies or environment ministries). Eye on Earth[3] is a global public information service, where EEA is one of the leading partners. This platform is primarily used as a dissemination platform and for sharing data between various organisations on a global scale. This also means that data collected within the Eionet community is intended to be shared via Eye on Earth.

This task leads to significant streams of environmental (observation) data between EEA and its member states. On the one hand, member states need to report their collected data to the EEA. On the other hand, the EEA gives member states, policy makers, stakeholders, and the general public access to collected and quality-assured data sets. This large number of different data flows relying on often different individual protocols is a key driver for the EEA to aim at establishing a common, standardised, and interoperable approach. By relying on a common standards based protocol, the EEA expects to significantly reduce the currently high overheads for harmonising incoming data formats as well as for adjusting its systems to new data access mechanisms of the member states.

## 2.2 Sensor Web Enablement

The Open Geospatial Consortium (OGC) is an international standardisation organization. The aim of the OGC is the integration of spatiotemporal data and services into the World Wide Web. To achieve this aim, the focus of the OGC is to establish interoperability and therefore develop standards for service interfaces, data models and formats (Percivall 2008). Within the framework of OGC standards, one group of standards concerns the integration of sensors and sensor data into spatial data infrastructures: the Sensor Web Enablement (SWE) framework, a set of standards for the interoperable integration of sensors and sensor data, which is especially designed to also support the transfer and the communication of environmental observations (Botts 2008).

The SWE architecture comprises information models, to represent data observed by sensors as well as metadata describing sensors, and service models which define the interfaces for sensor related

[2] http://www.eionet.europa.eu/
[3] http://www.eyeonearth.org/

functionalities (i.e., access, tasking, and alerting). The SWE framework is intentionally designed with a high degree of flexibility. It is capable of handling data measured by physical sensing devices, processed sensor data, as well as data derived from virtual sensors, e.g., simulation models. Furthermore, it is scale and domain independent. Also, real-time data access as well as access to archived historical observation data can be provided via SWE (Bröring et al. 2011).

For enabling interoperable data flows across Europe, three parts of the SWE framework are considered: the standardised models for sensor data (Observations and Measurement, O&M) and metadata (Sensor Model Language, SensorML), as well as the Sensor Observation Service (SOS), which defines an interface for sensor data access.

### 2.2.1 Sensor Observation Service

The Sensor Observation Service (SOS) provides a standardised interface for accessing sensor data in a common manner (Bröring et al. 2012). It provides operations for retrieving both observation data (*GetObservation*) and sensor metadata (*DescribeSensor*). Besides these data access operations, the SOS offers also a transactional interface which allows the insertion of sensors (*InsertSensor*) and their observed data (*InsertObservation*).

Within the EEA's SWE architecture, the SOS interface plays a central role. It is intended to be used by the EEA as well as the member states to provide access to their data sets.

### 2.2.2 Observations and Measurements

Observations and Measurements (O&M) defines a data model as well as encoding for observation data. While the conceptual model of O&M is defined by ISO (ISO 2011), the according XML encoding is openly available as an OGC standard (Cox 2011). O&M provides a model for associating observed values with all relevant properties necessary for their interpretation (e.g., location, time, observed parameter, and unit of measurement).

With regard to the SOS, O&M is used in two operations. First, O&M is the default output of the *GetObservation* operation. Second, O&M is used for encoding measured data that shall be inserted into a SOS through the *InsertObservation* operation.

Also for the official European INSPIRE initiative, O&M has significant relevance (INSPIRE Cross Thematic Working Group on Observations & Measurements 2011). Several of the INSPIRE Annex Themes have been specified so that their scope, in addition to the basic spatial information, includes measured, modelled, or simulated data. As O&M is ideally suited to cover this type of data, its usage is recommended by INSPIRE.

### 2.2.2 Sensor Model Language

While O&M is used for observed data, SensorML offers a model and an XML encoding for sensor metadata (Botts 2007). Information such as the owner of a sensor, the sensor type, identifiers of a sensor, inputs/outputs of a sensor, or geometric information about a sensor can be described in SensorML documents. SensorML is relevant for two SOS operations. First, it is the default output of the *DescribeSensor* operation. Second, it is used for encoding the descriptions of sensors that shall be registered at a SOS through the *InsertSensor* operation.

## 2.3 Related Sensor Web Projects

In the past years, a multitude of research projects and testbeds have been conducted to provide sensor data in a homogeneous way by utilizing Sensor Web technology. These projects analyzed the suitability of SWE standards and contributed to their development. In the following a non-exhaustive selection of projects is presented.

As part of the 6th Framework Programme, the projects OSIRIS[4] (Jirka et al. 2009a) and SANY[5] (Klopfer et al. 2009) were funded by the European Commission between 2006 and 2009. Both projects dealt with use cases from environmental monitoring and risk management which incorporated in-situ sensors. SWE infrastructures were setup to gain interoperable access to those sensors.

Within the 7th Framework Programme of the EU, the projects GENESIS, EO2Heaven, and ESS have built up SWE-based architectures to integrate sensors with applications. In GENESIS[6], SWE services have been used to share air as well as water quality data to investigate the correlation between public health and environmental factors. EO2Heaven[7] addresses human health and utilizes the SOS to monitor infections as well as human exposure to environmental pollution. The ESS[8] project makes use of SWE services to enhance crisis management by providing real time information to decision makers and ground forces, such as fire fighters.

As a joint project between Germany and Indonesia, GITEWS[9] (Raape et al. 2009) developed a SWE based Tsunami early warning system. Thereby, the SOS is used to provide access to the various sensors such as terrestrial seismic, marine, or virtual sensors, i.e., simulations. In South Africa, the national power supplier funded the project AFIS (Terhorst et al. 2008) which integrates through a SWE infrastructure in-situ weather sensors with remote sensing data for the detection of bush fires. Under the umbrella of the OGC, multiple testbeds have been conducted to evaluate and enhance the SOS and its associated encodings. First, the Ocean Science Interoperability Experiments (Oceans IE) 1 & 2 (Bermudez et al. 2006) have been conducted; followed by the Groundwater Interoperability Experiment[10] as well as the Surface Water Interoperability Experiment[11].

This work is similar to and builds up on the above listed projects and testbeds. However, due to the need for a productive system at the EEA, the practical requirements for this work as a foundation to answer the research question are clearer defined. A key requirement is the profile for the SOS to facilitate interoperability. The definition of this profile goes beyond the related work and is the contribution of this research.

# 3 Use Cases

This section introduces the use cases of the EEA for the transfer of environmental observation data across Europe. The following two sub-sections address two directions of data flow. First the flow of environmental data from member states to the EEA is described (Section 3.1). Second, the use case of providing access to the collected, processed, and quality-assured observation data is described (Section 3.2). Then, Section 3.3 derives the key requirements of the EEA from those use cases.

[4] http://www.osiris-fp6.eu/
[5] http://sany-ip.eu/
[6] http://www.genesis-fp7.eu/
[7] http://www.eo2heaven.org/
[8] http://www.ess-project.eu/
[9] http://www.gitews.de/
[10] http://www.opengeospatial.org/projects/initiatives/gwie
[11] http://www.opengeospatial.org/projects/initiatives/swie

## 3.1 Reporting by Member States

An example where the data collected from the EEA member states is used is Eye on Earth[12]. Eye on Earth is a global public information service with focus on data sharing between various organisations but also a place for the general public to find relevant and reliable environmental information and data. Air quality (Ozone, PM10, and NO2) is among the many thematic topics found in Eye on Earth. This data is transmitted on an hourly basis by the EEA from its member states using a proprietary data format as described in (Targa et al. 2009). At the EEA the data is received by a Microsoft Biztalk server setup which validates, processes, and finally inserts the received data into a SQL Server SDE database. From here the data can be accessed for viewing in Eye on Earth using ArcGIS Server map services. Another topic, bathing water quality, relies on a similar mechanism.

With regard to Eye on Earth, the EEA would significantly benefit from a standardised approach for the data exchange with its member states. Relying on standardised SWE formats and interfaces the EEA would be able to discontinue the use of proprietary software that is bound to the currently used data formats. Instead, the EEA and its member states would be able to rely on any software supporting the SWE standards so that the maintenance of a proprietary software infrastructure is no longer necessary. Member states could get into a position where they can flexibly choose software components as long as the SWE standards are supported. As there are multiple data flows between the EEA and its member states, the ultimate goal is that a multitude of currently used (proprietary) data formats can be narrowed down to one single standard format that is supported by a broad range of software systems which can handle all occurring data flows of environmental monitoring data.

For enabling the flow of data between member states of the EEA, two use cases have to be considered: active reporting of environmental data by the member states to the EEA (*pushing the data*) and harvesting of observation data by the EEA from its member states (*pulling the data*).

In the first case, the member states rely on a server provided by the EEA that offers a standardised interface for *pushing* new near real-time measurement data into the systems of the EEA. For example, a data provider in a member state would be responsible for setting up a national measurement network for air quality. This data provider collects the data measured by the sensor network and performs first processing steps (e.g. filtering out erroneous data). Subsequently, the data provider connects to the server located at the EEA and submits the freshly collected data using the transactional SOS operations InsertSensor and InsertObservation (see Figure 2). The advantages of this approach are that new data can be pushed by the member states directly to the EEA as soon as it is available and that it frees the member states from providing data access servers.

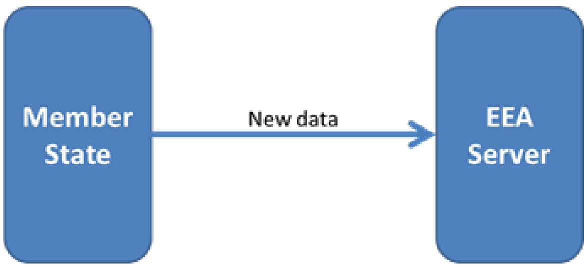


Figure 2 - Pushing Environmental Data by the EEA Member States

The second case is a *pull*-based approach. Thereby, the EEA regularly queries servers of the member states for new data sets. The member states would be responsible to set up SOS servers enabling the

12 http://www.eyeonearth.org/

access to their environmental observation data through a standardised interface. By querying for new measurement data through the SOS operation GetObservation in regular intervals, the EEA is able to retrieve the new data sets (see Figure 3). . The data transfer is initiated and controlled by the EEA. Thus, delays may occur as new data is only transferred to the EEA systems when the regularly scheduled query for new data is executed by the EEA. Other than in the previous approach, new data is not automatically pushed to the EEA.

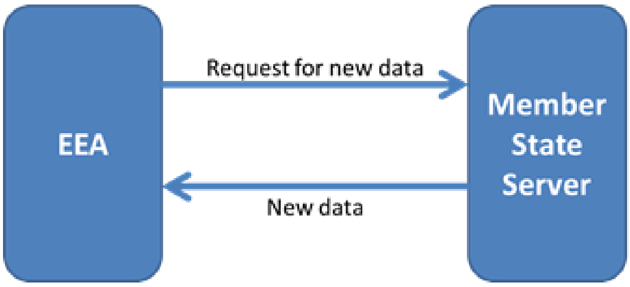


Figure 3 - Collection of Environmental Data by the EEA

## 3.2 Provide Access to Environmental Data

The second group of use cases concerns the provision of data by the EEA to external consumers, for example member states, policy makers, research projects or the general public. The following examples illustrate typical scenarios in which data consumers benefit from the provision of environmental data by the EEA through standardised interfaces.

An important aspect is the exchange of environmental data not only on a regional, national, or continental level, but also globally. An example of this is the exchange of air quality data between the US Environment Protection Agency (US EPA) [13] and the EEA. The aim of this initiative is to provide more near real-time data for air quality modelling but also the possibility for dissemination of the US EPA data into Eye on Earth. Currently the data exchange is based on existing non-standard services, but it shows the importance and necessity of supporting open standards in order to ease and make data exchange robust and stable. By relying on an open and standards based approach, it would become possible to easily integrate also other parties into a similar data exchange. Based on this example it can be shown that the EEA would benefit from a standards based approach as it allows to provide environmental data to multiple partners through a single system.

## 3.3 Requirements of EEA

Based on the previously described use cases, the following characteristics of the intended SWE architecture can be derived from the EEA use case, but can also be seen as the general requirements for spatial data infrastructures today:

- *Interoperability*: An approach is needed which ensures that different systems (at the EEA, member states, and other data consumers) are able to communicate with each other without the need for system adjustments.
- *Vendor independency*: The EEA cannot prescribe a specific software tool to its member states. Thus, the use of an open standard is essential so that it can be implemented within any system used by partners of the EEA.

[13] http://www.epa.gov/

- *Domain independency*: The approach for exchanging different types of data (e.g. air quality data and water quality data) shall be possible with the same protocol.

The following section introduces an approach to address these requirements by relying on the standardised OGC Sensor Web Enablement framework.

# 4 Approach

This section describes the approach that was chosen in order to fulfil the requirements of the EEA. Section 4.1 outlines the overall architecture. Section 4.2 introduces the main contribution of this work, the lightweight profile for the SOS which aims at increasing interoperability, in this case, between the EEA and its member states. This profile is a core element for realising the overall architecture by specifying the concrete data formats and interfaces. Section 4.3 presents an example how the profiles described in Section 4.2 can be used for implementing the developed architectural approach. More specifically, an extension for ArcGIS Server is introduced which is a core element for achieving the integration of SWE into the existing EEA infrastructure.

## 4.1 Architecture

The designed architecture for the EEA infrastructure is shown in Figure 4. It makes use of the OGC SWE standards for enabling the relevant data flows. These data flows rely on the O&M XML encoding and the SOS interface. Both, push based as well as pull based communication shall be possible with the architecture, although the first iteration of the presented work has focused on the second option whereas the first variant will be subject to future work.

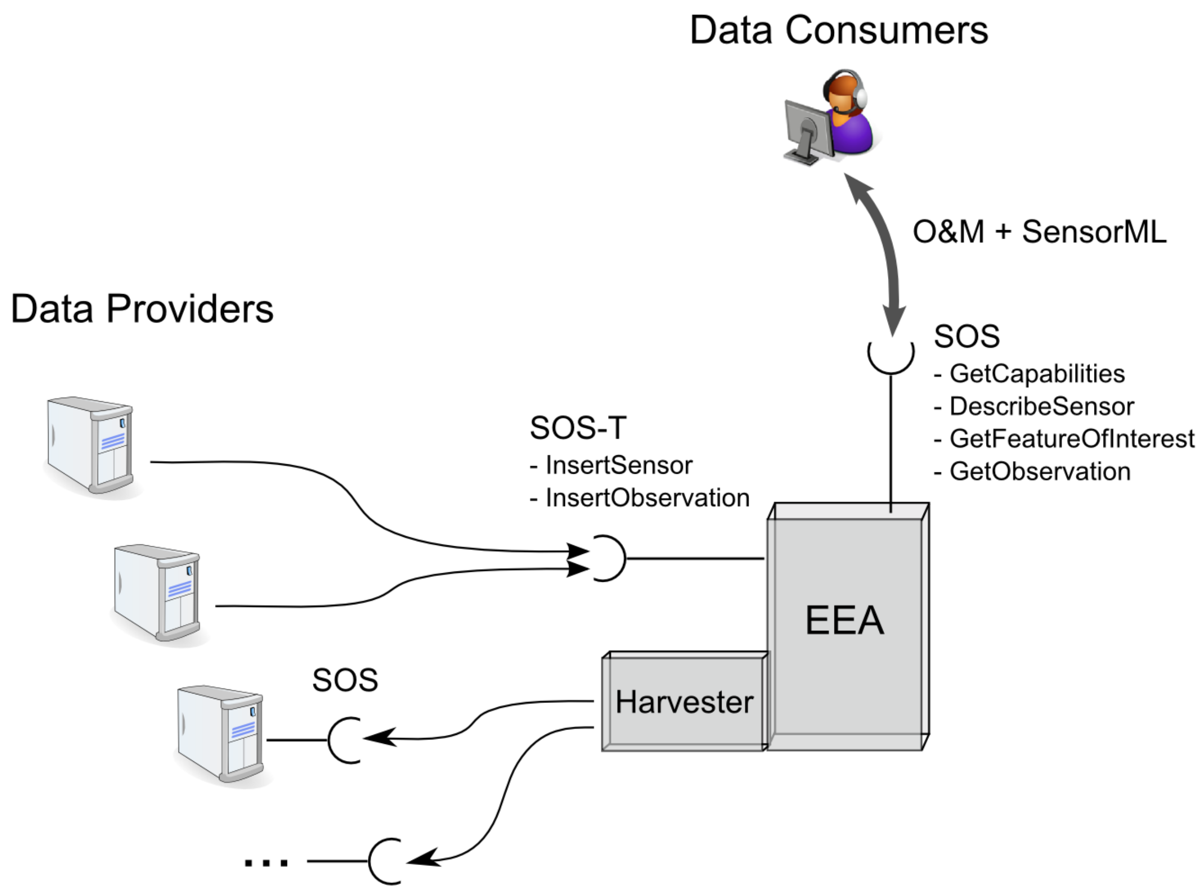


Figure 4 - Overview of the Architecture

In the first case, data providers actively transmit (i.e., *push*) through implemented adapter components for their data sources new data to an SOS server at the EEA. Therefore, the SOS at the EEA has to support the transactional extension of the SOS interface, so that it is able to receive sensor registrations and observation insertions. This way, the internal infrastructure of the EEA is hidden by the SOS and all data can be delivered through the same interface.

In the second case, a harvesting component at the EEA regularly *pulls* sensor data from SOS servers which are operated by agencies in the member states. The harvester subsequently feeds the received data into the infrastructure of the EEA.

The data handling within the EEA infrastructure needs to be embedded into the existing software stack, which, in case of the EEA, relies on Python scripts and on Microsoft BizTalk (Vasters 2001). For storing the collected environmental data, a MS SQL Server is used. Special focus has been put on the development of an according data model supporting the storage and management of near real-time data in an optimal manner (i.e. efficient storage and fast query for typical request parameters).

The SQL Server is also the data source used by the SOS extension of the ArcGIS Server (Section 4.3). This extension makes it possible to offer the data in the underlying database to internal and external data consumers through SOS servers. Thus, the EEA is not only able to collect the data from the member states in a standardised way but makes it also accessible in an interoperable manner.

The elements of this architecture are being implemented in an iterative manner. The lightweight SOS profile described in section 4.2 is the basis for instantiating this architecture, as the profile offers a detailed approach how the different SWE interfaces, data models and data formats shall be applied.

## 4.2 Lightweight SOS Profile

The OGC SWE framework plays a central role in the system architecture. For achieving interoperability between different parties, profiles of the relevant SWE standards are necessary. The reason for this lies in the high flexibility of the SWE framework. This flexibility makes it possible to apply SWE standards to a very large number of different scenarios, domains and sensor types. At the same time, this flexibility requires that several parts of the OGC SWE standards are optional so that these elements are not supported by every implementation. In addition, it is possible that there is more than one way to encode certain sensor observations and metadata. Thus, profiles are important to restrict the SWE standards to a minimum set of elements which have to be implemented by every profile compliant software package. In this work, a lightweight SOS profile for stationary in-situ sensors was developed. As the second generation of the SWE standards has been published or is in the final stages of the specification process (Bröring et al. 2011), the 2.0 version of the different standards are used. Detailed information about this profile can be found in a dedicated OGC specification (Jirka et al. 2011).

From the SOS, four operations are included in the profile. The *GetCapabilities* operation allows retrieving a description of the content and allowed operation parameters of a SOS server. To load measurement data from an SOS server the *GetObservation* operation can be used whereas metadata about the sensors can be accessed by sending *DescribeSensor* requests. Finally, the *GetFeatureOfInterest* operation is intended for serving the geometric descriptions of features associated with observations. A more detailed description which parameters are supported for the different operations of the lightweight SOS profile can be found in Table 1 (the GetCapabilities operation is not shown in this table as it does not include any restrictions). Because the realisation of active data pushing from member states to the EEA has not been included in the first phase of this work, the transactional SOS operations have been included in the architectural concept but not yet in the lightweight SOS profile. To reduce the complexity, this work has focused on the core operations of the SOS specification.

Developing a profile for the O&M standard is a core element of this work as it defines the format of environmental data flowing from the member states to the EEA and from the EEA to data consumers. The most important restriction of the O&M profile is the focus on a limited set of observation types. This avoids large overheads for software developers to support exotic observation types while still covering all of the common requirements occurring in practice. The observation type with the highest relevance within the profile is the *Measurement* which can be used for encoding scalar values. However, also other data types are supported. For integer values the *CountObservation* is available, for Boolean values the *TruthObservation*, for categorical values the *CategoryObservation,* and for textual values the *TextObservation* can be used. A detailed overview of the elements that need to be included in an O&M document according to the lghtweight SOS profile is described in Table 2. Figure 5 shows an example how such an O&M document is encoded.

```
<sos:GetObservationResponse xmlns="http://www.opengis.net/sos/2.0">
  <observationData>
    <om:OM_Measurement gml:id="o1">
      <gml:identifier codeSpace="not_used">obsTestWeatherStation</gml:identifier>
      <om:phenomenonTime>
        <gml:TimeInstant gml:id="p1">
          <gml:timePosition>2009-01-11T16:22:25.00Z</gml:timePosition>
        </gml:TimeInstant>
      </om:phenomenonTime>
      <om:resultTime xlink:href="#p1"/>
      <om:procedure xlink:href="urn:ogc:object:feature:Sensor:52North:weatherStation123"/>
      <om:observedProperty xlink:href=
                               "http://sweet.jpl.nasa.gov/2.3/propTemperature.owl#Temperature"/>
      <om:featureOfInterest xlink:href=
                            "http://myServer.org/features/SamplingPointAt52NorthHeadquarters"/>
      <om:result xsi:type="gml:MeasureType" uom="Cel">28</om:result>
    </om:OM_Measurement>
  </observationData>
</sos:GetObservationResponse>
```

Figure 5 - O&M Excerpt

With regard to SensorML a minimum set of mandatory metadata was defined to define how sensor metadata provided by the member states to the EEA and by the EEA to its consumers has to be encoded. On the one hand, the SensorML profile comprises general descriptive elements (e.g. textual descriptions, keywords, classifiers for sensor types). On the other hand, more technical information such as identifiers of the sensor, relationships to geographic features and the outputs of a sensor are covered by the profile. The detailed description of these elements is provided in Table 3.

The SOS operations, their parameters as well as the used encoding profiles are essential for the defined architecture. The next section (4.3) introduces an example how the described lightweight SOS profile is implemented in practice for the ArcGIS Server environment at EEA.

## 4.3 ArcGIS Server SOS Extension

The existing EEA infrastructure utilizes ArcGIS Server (Bader 2005) to provide environmental data. Also, many EEA member state organizations make use of this technology. To enable an easy integration of the SOS into those legacy systems, the starting point for implementing the developed SWE architecture based on the lightweight SOS profile was the design of an extension for ArcGIS Server which provides the functionality of the SOS interface and implements the lightweight SOS profile. In the future, further implementations of the lightweight SOS profile will be available (e.g. within Open Source SOS implementations such as the 52°North SOS).

On top of a database, hosting the environmental data, an ArcGIS Server and its typically used Map Service are set up with the default deployment routines. Then, the new SOS Extension is added to the Map Service to provide the additional interface functionality. Without this extension, the data is accessible via the GeoServices REST API (Esri 2010), an interface which is currently going through the OGC standardization process. By adding the SOS Extension, the data becomes available via the interface of the OGC SOS standard (Section 2.2.1).

For this ArcGIS Server SOS Extension two key development steps have been undertaken. At first, a data model for near real-time observation data used by the database was designed. This data model is aligned with the O&M 2.0 standard (Section 2.2.2) as well as the here proposed lightweight profile (Section 4.2). The elements of this data model reflect the key types of O&M, such as *observation*, *feature of interest*, and *procedure*. By grounding the data model in O&M, it is generic enough to support various data sets from differing domains.

In a second step, the interface of the ArcGIS Server SOS Extension was developed.  Before a fully OGC SOS compliant interface has been implemented, a simpler resource-oriented interface was designed, which follows the principles of the GeoServices REST API (Esri 2010) that is natively supported by ArcGIS Server. An overview of the resource-oriented SOS interface is presented in Figure 6. Observations, features, and procedures are the key resources that can be directly accessed, since each entity has its own URL, or query operations can be used to filter resources temporally, spatially or thematically.

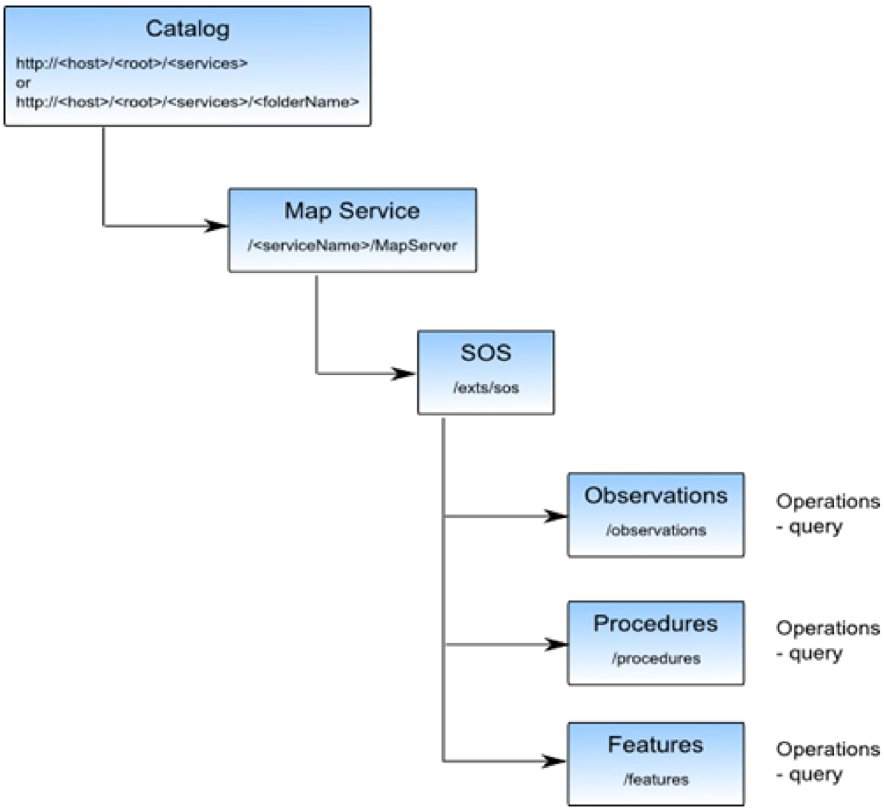


Figure 6 - Overview of resources accessible via the SOS Extension interface

On top of this interface for resources of the environmental data model, the SOS operations proposed by the lightweight profile are implemented. The *GetObservation* operation is returning observation entities, the *DescribeSensor* operation returns metadata about procedures, and the *GetFeatureOfInterest* operation is used to give access to the features resource.

The resulting ArcGIS Server SOS extension is not limited to the EEA case. Instead it is made publically available so that it can be used by other data providers for enabling their ArcGIS infrastructures to support the SWE standards. Thus, it can be considered a first practical implementation of the described SWE architecture and profiles for enabling the cross-European exchange of environmental data. However, further implementations in different software environments will follow in the future.

# 5 Application and Discussion

For verifying the lightweight SOS profile (Section 4.2), it was integrated into the open source implementation of the 52°North Sensor Observation Service[14]. The operations of the lightweight SOS profile, *GetCapabilities*, *GetObservation*, *DescribeSensor* and *GetFeatureOfInterest*, were implemented supporting the parameters listed in the profile specification. By linking the SOS to the 52°North Thin SOS Client[15] it was possible to verify that the operations and parameters included in the lightweight profile are sufficient to support typical functionality such as sensor data access and visualisation.

In addition, the ArcGIS Server SOS Extension (Section 4.3) has been deployed and tested for a data set provided by the EEA[16] which has been generated by a network of around 1500 air quality stations all across Europe, as shown in Figure 7. In this example setup, 30 days of data comprising over 1 million observations were loaded into an SQL Server database[17]. Via the SOS Extension, observations as well as metadata about the sensor network stations can be accessed.

[14] http://52north.org/sos

[15] http://52north.org/communities/sensorweb/clients/Thin_SWE_Client

[16] http://www.eea.europa.eu/maps/ozone/resources/about-the-data

[17] The endpoint URL of this example deployment of the ArcGIS Server SOS Extension is: http://ags.dev.52north.org:6080/arcgis/rest/services/AirQualityEurope/MapServer/exts/SosSOE

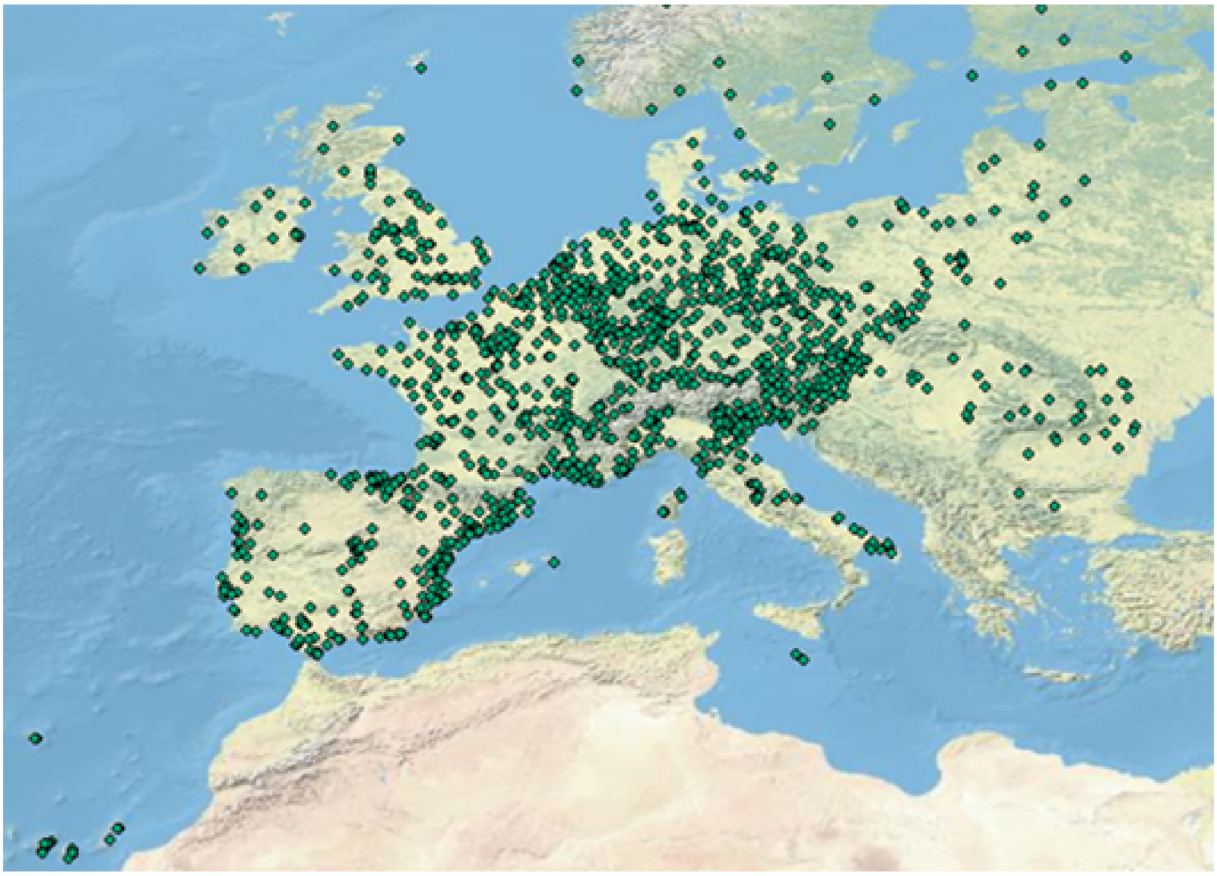

Figure 7 - Network of air quality sensor stations across Europe

The first test results have shown the applicability of the SWE framework and the lightweight SOS profile to near real-time data streams as they occur in practice at the EEA. However, several aspects need further work. Investigations are necessary to determine if additional observation types for more compact encodings of the measured values should be included in the lightweight SOS profile. Currently, the volume of the transmitted O&M documents is rather large due to the verbose XML encoding. Single O&M observations are used for each measured value. The structure of those single O&M observations is plain XML and thus easier to parse by client applications. O&M also allows the encoding of data as comma separated values (CSV) by utilizing the SWE Common Data Model standard (Robin 2010). Figure 8 shows the difference in size of these two alternative O&M encoding styles. It is clear that the CSV way of encoding the data is much more efficient. However, with today's high bandwidth networks, also the transmission of around 700 kilo bytes for 1000 data values in the XML style encoding does not necessarily result in latency issues.

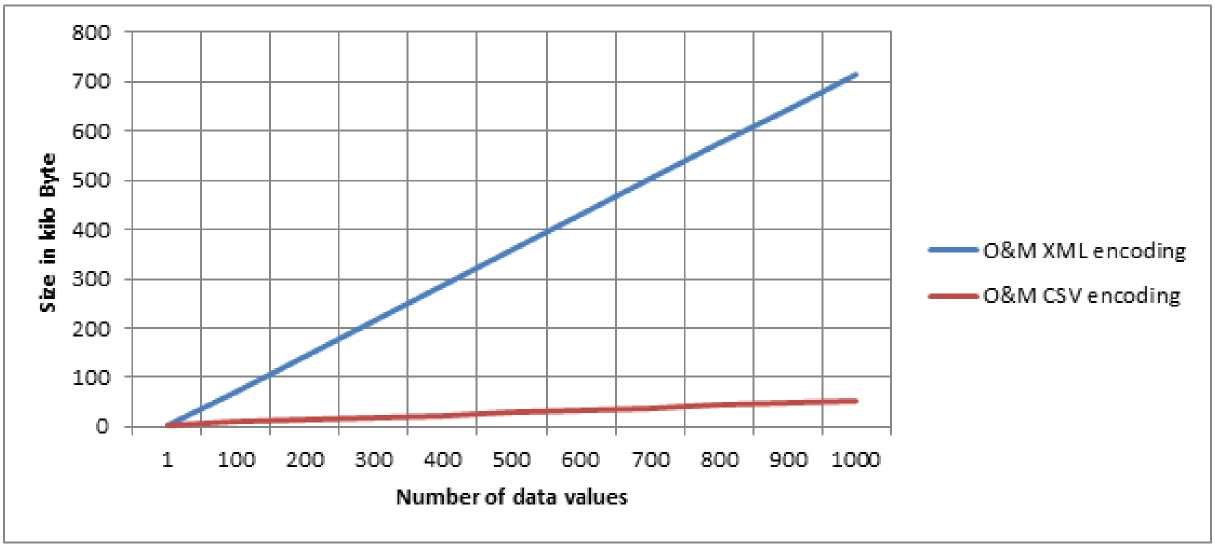


Figure 8 - The size of different O&M encoding styles

Although the lightweight SOS profile offers a very simple approach for implementing SWE, it still has some inherent complexity. Thus, the more and more popular REST concepts (Richardson & Ruby 2007) have to be considered. As REST based interfaces offer a very simple way to communicate with web services, a REST approach for the SOS 2.0 standard may need to be considered in the future to facilitate the linking of applications to the EEA's SOS servers. The here developed SOS extension for the GeoServices REST API (Section 4.3) is a candidate for such a REST based interface design.

Finally, it is important to gain acceptance of the developed infrastructure by further stakeholders from practice. Although the EEA is at the core of the system, the practical application within its member states is critical. This may also lead to further requirements and experiences that will help to improve the current version of the architecture as well as its implementations.

To contribute the findings of this work, the lightweight SOS profile was presented and published as a *Discussion Paper* (Jirka et al. 2011) during the OGC Technical Committee meeting in Brussels (Belgium) in November 2011. First feedback on the profile has been encouraging. Suggestions for further enhancements in the future concerned primarily the inclusion of additional observation types so that data arrays for sensor data can be encoded using comma separated values and to reduce the amount of necessary XML elements.

# 6 Outlook

In future, this work will focus on the transition of the developed approach into an operational mode. After the deployment in the EEA's infrastructure, it is important to establish SWE based reporting and data exchange with the EEA's member states. This step goes beyond the deployment at the EEA. On the one hand, it is important to make sure that member states interested in implementing SWE will have access to according SWE components. The support of the lightweight SOS profile within commonly used software (e.g. ArcGIS Server, ArcGIS Desktop, 52°North SOS) is a very important step towards this goal. On the other hand, to increase the acceptance and implementation of the SWE framework in EEA's member states, further support is necessary. This includes especially the provision of best practise documentation, tutorials as well as training workshops.

Also the practical implementation of the data flows between the member states and the EEA will require further work. Besides realising also the push based data transmission, which has been included in the general architecture but not yet included in the SOS profile, advanced aspects such as synchronisation of data sets are subject of future work.

Regarding the lightweight SOS profile for stationary in-situ sensors (Section 4.2), the aim is to advance it to an officially adopted OGC standard. In order to achieve this goal, further open discussion with the OGC community is necessary to fine-tune the existing Discussion Paper (Jirka et al. 2011) and to reach consensus on necessary enhancements (e.g., regarding the support of additional observation types).

The presented implementation of the ArcGIS Server SOS Extension (Section 4.3) enables the EEA and the member state organizations to easily extend their existing infrastructure to support the SOS interface. This extension supports not only an OGC SOS conform way of retrieving observations, and metadata about features and sensors, it also provides a resource-oriented perspective on the O&M model. This resource-oriented interface follows the principles of the GeoServices REST API (Esri 2010) which is currently making its way through the OGC standardisation process. This resource oriented, and REST-based interface of the SOS Extension is aligned with the OGC SOS standard and can be considered as a new implementation of the conceptual model of the SOS 2.0. In future, this extension of the GeoServices REST API for providing observation data will also be brought into the standardization process at the OGC.

Another challenge is the discovery of the available environmental data sets. The current scope of the work is focused on the exchange of environmental data through standardised data models and interfaces. Enabling discovery in such a Sensor Web architecture will be element of future work (Jirka et al. 2009b).

# 7 Conclusions

This paper addresses the research question of how to build a system that allows the sharing of environmental data on a large scale across multi-organizational environments by developing a standards-based approach. While this work is driven by the European Environment Agency (EEA), it is made sure that the requirements from practice are well covered. The key achievement of the presented work is the lightweight profile for OGC's Sensor Observation Service (SOS), which facilitates the implementation of SWE and enhances interoperability. We have implemented this lightweight profile in an SOS Extension for ArcGIS Server, which is needed to support many existing spatial data infrastructures, as seen in case of the EEA.

Although the presented work is strongly influenced by the EEA and its practical requirements, the transfer of the results to other use case scenarios is straightforward. The lightweight SOS profile has been developed in a use case independent manner so that stakeholders dealing with stationary in-situ sensors are able to re-use this key outcome of this research.

With regards to future developments, the practical deployment at the EEA as well as in first member states will be important steps. Thus, the presented results will be a core element for enabling the exchange of environmental data across Europe.

## Acknowledgements

This work is supported by the European Environment Agency (EEA).

## References

Bader E 2005 ArcGIS Server Administrator and Developer Guide. Redlands, CA, USA, Esri Inc.

Bermudez L, Bogden P, Bridger E, Creager G, Forrest D, Graybeal J 2006 Toward an Ocean Observing System of Systems. In Proceedings of OCEANS 2006, Boston, MA, USA, IEEE: New York, NY, USA

Botts M, Robin A 2007 OGC Implementation Specification: Sensor Model Language (SensorML) 1.0.0 (07-000). Wayland, MA, USA, Open Geospatial Consortium Inc.

Botts M, Percivall G, Reed C, Davidson J 2008 OGC Sensor Web Enablement: Overview and High Level Architecture. In Proceedings of GeoSensor Networks, Boston, MA, USA, Lecture Notes in Computer Science, Vol. 4540. Berlin, Germany, Springer

Bröring A, Echterhoff J, Jirka S, Simonis I, Everding T, Stasch C, Liang S and Lemmens R 2011 New Generation Sensor Web Enablement. Sensors 11: 2652-2699

Bröring A, Stasch C, Echterhoff J 2012 OGC Implementation Standard: OGC Sensor Observation Service, Version 2.0, Candidate Standard (12-006). Wayland, MA, USA, Open Geospatial Consortium Inc.

Cox S 2011 OGC Implementation Specification: Observations and Measurements (O&M) - XML Implementation 2.0 (10-025r1). Wayland, MA, USA, Open Geospatial Consortium Inc.

Esri 2010 GeoServices REST Specification, Version 1.0. Esri White Paper. Redlands, CA, USA, Esri Inc.

European Commission 2007 Directive 2007/2/EC of the European Parliament and of the Council of 14 March 2007 establishing an Infrastructure for Spatial Information in the European Community (INSPIRE). Brussels, Belgium, European Commission.

European Environment Agency 2005 Eionet connects. Copenhagen, Denmark, European Environment Agency

European Environment Agency 2011 EEA - Who we are. WWW document, http://www.eea.europa.eu/about-us

INSPIRE Cross Thematic Working Group on Observations & Measurements 2011 D2.9 Guidelines for the use of Observations & Measurements and Sensor Web Enablement-related standards in INSPIRE Annex II and III data specification development. WWW document, http://inspire.jrc.ec.europa.eu/documents/Data_Specifications/D2.9_O&M_Guidelines_V1.0.pdf

ISO 2011 ISO 19156:2011 - Geographic information -- Observations and measurements - International Standard. Geneva, Switzerland, International Organization for Standardization.

Jirka S, Bröring A, Stasch C 2009a Applying OGC Sensor Web Enablement to Risk Monitoring and Disaster Management. In Proceedings of the GSDI 11 World Conference, Workshop on Sensorweb Enablement: Strengthening the SDI. Rotterdam, The Netherlands

Jirka S, Bröring A, Stasch C 2009b Discovery Mechanisms for the Sensor Web. Sensors 9: 2661-2681

Jirka S, Stasch C, Bröring A 2011 OGC Lightweight SOS Profile for Stationary In-Situ Sensors Discussion Paper (OGC 11-169). Wayland, MA, USA, Open Geospatial Consortium Inc.

Kjeld P, Bliki J, Jirka S, and Wytzisk A 2011 Sensor Web Technology for Sharing Environmental Data across Europe. In Proceedings of the INSPIRE Conference 2011. Edinburgh, United Kingdom

Klopfer M, Simonis I 2009. SANY - An Open Service Architecture for Sensor Networks. WWW document, http://www.sany-ip.eu/publications/3317

Percivall G 2008. OGC Reference Model (OGC 08-062r4). Wayland, MA, USA, Open Geospatial Consortium Inc.

Raape U, Tessmann S, Wytzisk A, Steinmetz T, Wnuk M, Hunold M, Strobl C, Stasch C, Walkowski A C, Meyer O, Jirka S 2009 Decision Support for Tsunami Early Warning in Indonesia: The Role of Standards. Cartography and Geoinformatics for Early Warning and Emergency Management, Prague, Czech Republic, In: Lecture Notes in Geoinformation and Cartography, Vol. 2, pp. 233-247. Berlin, Germany, Springer

Richardson L, Ruby S 2007 RESTful Web Services. Sebastopol, CA, USA, O’Reilly Media, Inc.

Robin A 2010 SWE Common Data Model, Encoding Standard, Version 2.0 (OGC 08-094r1). Wayland, MA, USA, Open Geospatial Consortium Inc.

Targa J, Gabrielsen P 2009 Setting up an additional near real-time data exchange with a national or regional provider. Copenhagen, Denmark, European Environment Agency. WWW document, http://eea.eionet.europa.eu/Public/irc/eionet-circle/airclimate/library?l=/public/real-time_operational/real-time_maintenance/p7141a_additional/_EN_2.5_&a=d

Terhorst A, Moodley D, Simonis I, Frost P, McFerren G, Roos S, Bergh F 2008 Using the Sensor Web to Detect and Monitor the Spread of Vegetation Fires in Southern Africa. In Proceedings of GeoSensor Networks, Boston, MA, USA, Lecture Notes in Computer Science, Vol. 4540. Berlin, Germany, Springer

Vasters C F 2001 BizTalk Server 2000: A Beginner's Guide. McGraw-Hill Professional

# Tables

Table 1: Overview of the SOS operations and their parameters

| **Operation** | **Supported Parameters** | **Explanation** |
|---|---|---|
| DescribeSensor | procedure | The identifier of the sensor for which a description is requested. This parameter is mandatory. |
| GetObservation | offering | The identifier of the offering that contains the data requested by the user. This parameter is optional. |
| | temporalFilter | The time for which sensor data are requested. In the lightweight SOS profile this is restricted to a time period or a time instant. This parameter is optional. |
| | procedure | One or more identifiers of the sensors from which observations are requested. This parameter is optional. |
| | observedProperty | One or more identifiers of the phenomena which are of interest to the user. This parameter is optional. |
| | featureOfInterest | One or more identifiers of the geographic features for which observations are requested. This parameter is optional. |
| | spatialFilter | The area for which observations are requested. In the lightweight SOS profile the spatialFilter is restricted to bounding boxes. This parameter is optional. |
| GetFeatureOfInterest | featureOfInterest | One or more identifiers of the geographic features which shall be returned. This parameter is optional. |
| | spatialFilter | The area for which geometric features are requested. In the lightweight SOS profile the spatialFilter is restricted to bounding boxes. This parameter is optional. |
| | observedProperty | One or more identifiers of the phenomena that shall be measured at the geographic features returned to the user. This parameter is optional. |

Table 2: Overview of the elements of the different observation types (all elements are mandatory)

<table>
<tr><th>Element</th><th colspan="2">Explanation</th></tr>
<tr><td>identifier</td><td colspan="2">An identifier of the observation.</td></tr>
<tr><td>phenomenonTime</td><td colspan="2">The time when the observation was performed. This can be either a time instant or a time period.</td></tr>
<tr><td>resultTime</td><td colspan="2">The point in time when the observation became available (e.g. after all processing steps were completed). If no post-processing is performed this the resultTime is usually identical to the phenomenonTime.</td></tr>
<tr><td>procedure</td><td colspan="2">The identifier of the sensor that has produced the observation.</td></tr>
<tr><td>observedProperty</td><td colspan="2">The phenomenon for which the observation contains data.</td></tr>
<tr><td>featureOfInterest</td><td colspan="2">The identifier of the geometric feature to which the observation is associated. Within the lightweight SOS profile only sampling points are allowed as features of interest.</td></tr>
<tr><td rowspan="7">result</td><td colspan="2">The observed value. The specific data types differs between the observation types included in the lightweight SOS profile:</td></tr>
<tr><td>Observation Type</td><td>Type of the results</td></tr>
<tr><td>Measurement</td><td>A scalar value including the unit of measurement.</td></tr>
<tr><td>CountObservation</td><td>An integer value describing a count.</td></tr>
<tr><td>TruthObservation</td><td>A Boolean value (either true or false).</td></tr>
<tr><td>CategoryObservation</td><td>A term from a pre-defined vocabulary.</td></tr>
<tr><td>TextObservation</td><td>Any kind of free text.</td></tr>
</table>

Table 3: Overview of the elements to be included in a SensorML document (all elements are mandatory)

| Element | Explanation |
|---|---|
| description | Textual description of the sensor. |
| identifier | A unique identifier for the sensor. |
| keywords | Human readable keywords that describe the sensor. These keywords are of special relevance for enabling sensor discovery. |
| identification | Further names of the sensor (this shall be a short name as well as a long name). |
| classification | Classifiers characterising the sensor. There shall be one classifier describing the sensor type. |
| contacts | Contact information about the operator of the sensors. |
| featuresOfInterest | A list of references to features of interest which represent the geographic objects for which the sensor is delivering observations. Within the lightweight SOS profile only sampling points are allowed as features of interest. |
| outputs | The outputs of the sensor. This comprises on the one hand a list of observed properties and on the other hand the data types (if applicable including the unit of measurement). For restrictions on the supported data types please refer to Table 2 which describes the allowed observation types in the lightweight SOS profile. |