Environmental data on demand – enabling dynamic live data flows with Sensor Web Enablement.

|  |  |
| --- | --- |
| **Title** |  |
| **Date Issued** | 07/01/2010 |
| **Publisher** | European Environment Agency (EEA) |
| **Description** | Strategy paper describing the EEA’s response to the increasing need for real-time environmental data flows through implementation of the OGC protocol Sensor Web Enablement |
| **Contributor** | European Environment Agency (EEA); Atkins Danmark a/s; Umweltbundesamt, Austria |
| **Format** | DOC |
| **Language** | en |

These are Dublin Core metadata elements.

Contact

Jan Bliki, EEA ([jan.bliki@eea.europa.eu](mailto:jan.bliki@eea.europa.eu))

Peter Kjeld, EEA ([peter.kjeld@eea.europa.eu](mailto:peter.kjeld@eea.europa.eu))

**Revision history**

|  |  |  |  |
| --- | --- | --- | --- |
| Version | Date | Author | Change |
| 0.9 | 09.12.09 | Jon Maidens (Atkins DK A/S) | Reworking based on feedback from PK and JB |
| 1.0 | 15.12.09 | Jon Maidens (Atkins), Katharina Schleidt, Christian Ansorge (UBA)  Jan Bliki, Peter Kjeld (EEA) | Reworking of doc and further review by EEA and UBA |
|  |  |  |  |

Table of Contents

[1. Summary 5](#_Toc248640334)

[2. The challenge of live data flows 5](#_Toc248640335)

[2.1 Serving timely environmental data 5](#_Toc248640336)

[2.2 The barrier to flow 6](#_Toc248640337)

[2.3 Environmental data without borders 6](#_Toc248640338)

[3. Sensor Web Enablement 7](#_Toc248640339)

[3.1 What is Sensor Web Enablement? 7](#_Toc248640340)

[3.2 What is Sensor Web Enablement good for? 8](#_Toc248640341)

[3.3 What does SWE mean to the EEA? 9](#_Toc248640342)

[4. Implementation plan 11](#_Toc248640343)

[4.1 EEA roadmap for 2010 11](#_Toc248640344)

[5. Further reading 12](#_Toc248640345)

# Summary

The demand for near real-time data flows is increasing. The EEA needs to stimulate the flow of environmental data, and increase possibilities for use, in providing real-time indicators of the state of the environment and in-situ monitoring and satellite observation.

This paper proposes that through the EEA-driven adoption of the Sensor Web Enablement (SWE) protocol these goals can be addressed. SWE is a set of open standards for interoperable communication provided by the Open Geospatial Consortium (OGC). The standards are directed at the transfer and the communication of environmental observations and data among distributed systems.

Through implementation of one open standard, which is thematically independent, a high level of interoperability can be achieved between data providers. This translates in to a benefit of increased and improved access to environmental data for the public, industry and government, which falls in line with the EEA’s remit under SEIS and INSPIRE. Also EEA can cement is position as the data and information node for Europe. Furthermore, using an established open standard is targeted at costs of setting up data flows by recommending an architecture which can be implemented across different themes and levels.

The EEA can lead the way in the implementation of the protocol through enabling a SWE data flow for an existing data flow as a proof of concept. The building blocks are already there and with a carefully planned balance of targeted implementation, demonstration and stakeholder community building over the next 5 and 10 years, the EEA can promote a viable data sharing architecture. Member State and institutional support implementation will be encouraged through an FP7 announcement.

# The challenge of live data flows

## Serving timely environmental data

Environmental data comes from many sources. It can come from a mechanical or electrical device measuring some parameter in-situ, it can be a human observer out in the field on a survey reporting something in a standardised way, or it can be analysis equipment in a laboratory analysing samples. This data may be aggregated or harmonised, but the common factor to all these various data sources is that they are linked to place and time.

From the data we extract our information through analysis or combination with other data and on the derived information we communicate messages or base decisions. These decisions may be long term strategic or they might be short term response necessitating fast action. Decision- makers demand faster availability of underlying environmental datasets for use in deriving economical and statistical information. Making the right decisions requires access to accurate and timely information, which can only be achieved through a system delivering accurate and timely data.

For the European Environment Agency (EEA), data traditionally finds its way up the chain to the European level through reporting obligations and priority dataflows to be presented in various information products (indicators, reports) and applications (online map viewers).

However, there are increasing demands for access to more immediate or near real-time data. Successful prototypes such as OzoneWeb, EyeOnEarth (bathing and airquality - http://eyeonearth.eu/) are already demonstrating the value and interest for these data flows. Further live information flows will be required in the near future through legislation such as the CAFE (Clean Air For Europe) Directive and the European initiative Global Monitoring for Environment and Security (GMES). GMES needs to link a system of in situ monitoring and satellite observation for an effective information exchange framework.

The European Commission has clearly expressed that to enable better policy making, the timeliness and periodicity of environmental data has to be improved with the goal to integrate it with economic (and social) data (“GDP and beyond: Measuring progress in a changing world” COM/2009/433/final).

Additionally, EEA does not envisage its position as a consumer of environmental data, solely focused on pushing out derived products to the general community, but also as a provider - an environmental data and information node for Europe. The Agency can serve this environmental data up to global strategies, such as the Global Earth Observation System of Systems (GEOSS), and also position itself as a environmental data and service broker for other organizations.

## The barrier to flow

The challenge is that environmental data is available in a multitude of formats and media, which are not harmonised across countries and media. Therefore there is a fundamental barrier to enabling a flow of data. Too many resources are spent trying to align the data in a common format, so that the content of the data becomes secondary in the allocation of effort and resources in extracting the information on which to base the decisions.

Additionally, the provider and user of information are currently tightly coupled, with data following disparate flows as a means to an end (e.g. to serve an application or meet an obligation). The same data is provided in different forms and formats to meet different requirements. A user typically requires information for a certain purpose and the format of the information will be dictated by the needs, application or a bilateral agreement. Therefore each data flow has to be dealt with independently.

It is costly for the EEA to manage multiple data flows in such a tightly coupled way when each flow includes overhead to maintain. Resource availability cannot keep up with that demand, and resources need to be better targeted than to a multitude of IT applications all targeted at data flow.

Therefore the challenges are to break the tight coupling, make data more readily available and counter the cost of ‘business as usual’.

## Environmental data without borders

To address these challenges, the EEA needs to strive for a vision of information exchange as defined within the principles of the Shared Environmental Information System (SEIS) and within the boundaries of INSPIRE. This requires open standards for interoperability of environmental data which is thematically and, in the long term, vendor independent as well as a common way of communicating between receiver and provider. In turn, this will reduce the cost of both data access and provision.

In formulating this vision, EEA has the opportunity to drive for a common approach which will address its own needs and have a value-added effect to stimulate the availability of further flows of data, through ease of re-use and rise in demand. Environmental monitoring and reporting is costly, and a vision based on mitigating costs will give greater success and filter throughout the system.

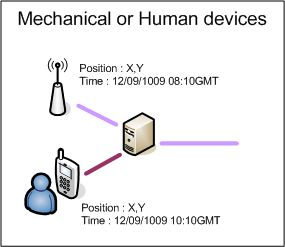
Building blocks are already being set in place and with effective coordination will lead to a shift in the definition of environmental live data flows in Europe. The success will not come through dictation but by demonstrating and collaborating with the data owners. Political issues will not be removed, but they can be tempered through a common language which this vision foresees, leading to a win-win situation.

The solution comes from mapping the data at any level of processing into an accepted model which means it in turn can then be used to serve a number of purposes. The solution is found in the OGC standard Sensor Web Enablement (SWE). SWE has been identified as the most promising technology to address the EEA's needs for live flows and as a good practice for Member States, as it lays down a standard for data interoperability and communication across all environmental themes.

# Sensor Web Enablement

## What is Sensor Web Enablement?

Sensor Web Enablement (SWE) is an Open Geospatial Consortium (OGC) initiative for interoperable communication especially designed for the transfer and the communication of environmental observations and data. The initiative is a framework covering several standards and specifications.

SWE does not necessarily imply an observation from a single mechanical or electronic device; the important criteria are that observations have a relation to time and place. It could also be a set of observations related to rivers, regions, counties or Europe as a whole. We are able to define our data sources at the desired geographical level and make that data available in the same format, regardless of the geographical level.

SWE defines a data model which provides the basic structure of how data should be communicated. In addition, services are available which allow us to work with the environmental data from a particular data source, e.g database or sensor. The protocol suite supports access to data, aggregation of data, alerts when measured values are changing or exceeding a threshold. The suite also supports other procedures regarding controlling sensors and data collection. Real-time and archived observations from these sensors and models are accessible in standard encodings, through standard interfaces, and are capable of being geo-referenced and processed on-demand without a priori knowledge of the underlying sensor system.

For the purposes of live data flows, we are concerned with only two of the above mentioned aspects of SWE and those are the fundamental data model and the Sensor Observation Service (SOS).

The OGC is behind the concept of Sensor Web Enablement and a key criteria for the EEA is the support of open standards. Importantly, the SWE effort involving OGC members in developing the global framework of standards and best practices is in continual progress, and therefore EEA can influence the formulation of the standards to include its needs within the broader application of SWE.

## What is Sensor Web Enablement good for?

SWE provides a common interface (protocol) for interoperable communication of data and overcomes our main challenge to sharing environmental data. The data structure for communication is independent of the content within the boundaries of any kind of observation. This does not impact the thematic data structures established on-site. This means once we have our system in place to deal with this common structure, we can work with different geographical scales, or different themes because we are assured that the two dimensions of time are location are available and the data is available in a common format.

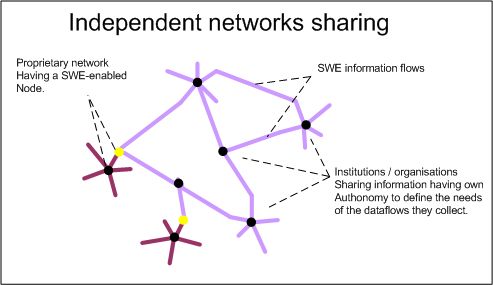


Figure 1: SWE facilities the transfer and the communication of environmental observations and data among distributed systems.

An important extension of SWE is data can be cascaded through a series of levels (developed under the FP6 S@NY IP project). A local network makes the monitoring data available as a validated SWE stream which is subscribed to at a regional or national level. This next body in the chain will have its own requirements for tracking this stream and other neighbouring streams. This second body also releases a validated stream derived from its set of providers but this stream is aggregated in some way, for example temporally, so that it is a manageable flow of validated data. This stream can then be subscribed to by the next level institution, such as the EEA, which wants to work with this information for its own needs and again would include some level of aggregation and harmonization before releasing its own data.

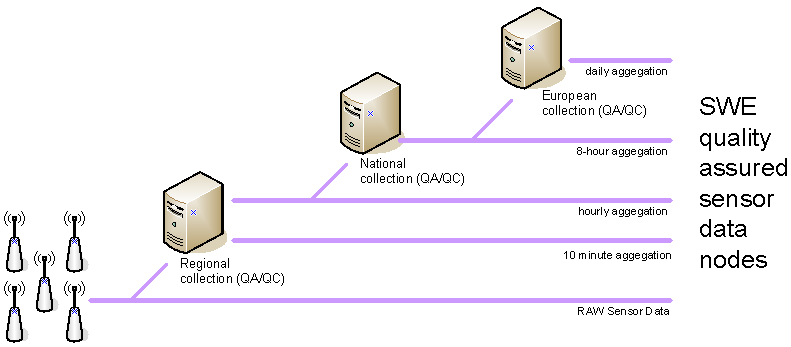
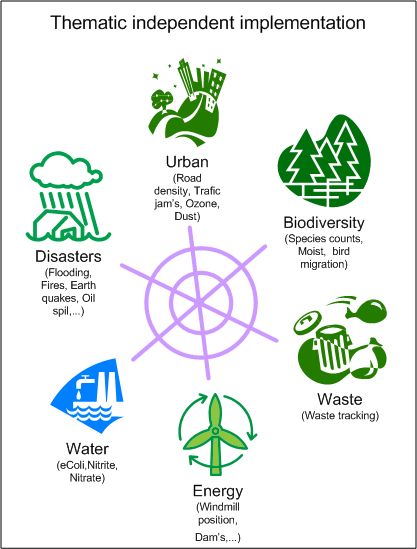


Figure : cascading of SWE data through cascading and/or filtering

Under a service chain scenario of environmental data built upon a series of distributed, cascading SWE data flows, changes to data at source will pass up through the system much faster. This process is achieved because all service levels are following the same standard - all that is changing along the way are certain describing parameters and the metadata describing the data. A near real-time data flow is enabled.

This does not mean the right flow will be instantly available to meet the final needs. Service chains require political and organizational cooperation. However it becomes much easier for a particular provider to meet the needs of multiple demands on its information flow as those requests can be fulfilled through a common format and means of communication.

Therefore environmental indicators (analytical results) could be produced or map viewers published at the moment new data has been received from a distributed system of data providers. Additionally very different streams can be fused together because they are bound by a common protocol. Delivering seamless merging of cross-thematic data with viable results, e.g. traffic data and ozone readings can be achieved. All flows communicate with each other using the same method – SWE.

The meaning of the data is only derived in the making of user products. The flow itself remains free of that and so maintains its flexibility and thus the possibilities which can be opened up for the use of that flow. The possibilities further up the chain are lower because of harmonization or aggregation, but the volume and thus cost of information processing is lower. It may also be that value has been added, but within such a system a user always has the option to subscribe lower in the chain to get the information that they want.

Figure : Data fusion

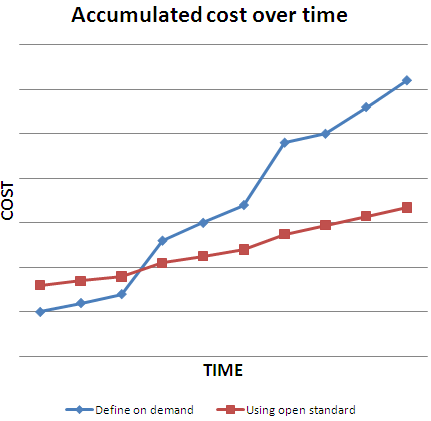
## What does SWE mean to the EEA?

Within a distributed network of SWE data providers, the EEA can gain increased and improved access to environmental data for the public, but without having to have knowledge of the sensors themselves. The existing chain of ownership of data is maintained. The QA/QC is provided by the owners of each of the nodes; the data owners will implement the appropriate filters and fail-safes so that the quality of the data meets an agreed standard. As stated previously, SWE does not dictate how the underlying data is stored, merely how it is shared.

Through the use of the other protocols available under SWE, further data flows can be achieved. A sensor from the ground or a sensor from space can communicate in the same way, allowing one to trigger the other using the same infrastructure. Therefore in a GMES scenario, we could use a Sensor Alert Service to tell us if a particular parameter breached a set threshold. On an occurrence of this event, this alert could trigger another sensor, such as a satellite, and through its Sensor Planning Service tell it to move and take an image of the area around the alerting sensor.

The wider benefit is increased and improved access to environmental data for the public, industry and government which falls in line with the EEA’s remit under SEIS and INSPIRE. SWE also lays the foundation for countering the costs of the increasing demand for the delivery of environmental information.

The value of following the SWE protocol is that the technological investment is not limited thematically so implementing new flows will become easier and implemented in a much shorter way. The process is independent of what is being measured. Therefore the same technological implementation could be transferred to meet the needs of another stream at a much cheaper cost than building a new data provision system. Also, the software implementation would be very similar at all levels of a data stream.



Reporting dictates a cost in resources, technology and time both upon the side of the provider and the user. This spans not just the set-up but also continual operational costs. The adoption of SWE is no different in that it requires initial overhead and subsequent overheads to maintain the system, and, as with any new technology, adoption has initial costs in terms of building expertise. (In these early stages market demand has not stimulated a range of competitive solutions from different software vendors.) Therefore custom development would be required on the supplier’s side.

However, over the long term, because the flow is independent of the thematic content, the set-up costs for adding subsequent streams are much lower. Due to the fact that the protocol is independent of what is being measured, then the software and knowledge to manage the streams in a SWE enabled network can be reused. The cost of setting up a sensor network to monitor specific parameters is still there, but these sensor networks are not required because a network demands it, but because of a local need. Adoption of the SWE protocol for the management and communication of those sensors opens up the possibilities of the reuse of that information, therefore adding value to the investment.

Figure : Hypothetical scenario showing the increase of costs and time between the current implementation of defining on demand, against a standards-based approach to communication.

Cost efficiency comes from a common platform.

# Implementation plan

The general implementation plan has two phases. Using an existing data flow, such as ozone air quality data, the data stored at the EEA can be first made available using the SWE protocol. Secondly, incoming data can be accepted using the SWE protocol. Initially selected collaborating countries would implement modules in their own systems to transmit the information using the SWE protocol.

The present interface to EEA for receiving and storing environmental data and information is through the ReportNet portal. An additional component is needed for ReportNet which can receive this SWE enabled data alongside traditional methods.

The aim is a five to ten years implementation to achieve critical mass in terms of use. It is not necessary that all national, local or regional bodies follow this protocol or nor likely that SWE will be used throughout the reporting system. However, SWE enabled data on a higher level has many merits, even if the areas below it follow a different means of data flow.

It is important that industry and open source communities are involved in the building of the SWE module, as SWE requires community extension. The goal is pluggable software components which can fit to the existing systems.

Being able to foster Member State buy-in to SWE can only be partly achieved through demonstration and marketing. Financial support is necessary and it is proposed that an FP7 announcement will be made to help Member States buy in to it. The key to success is working directly with these stakeholders in their implementation to gain an assessment of the process.

Strong cooperation with the Joint Research Center (JRC) will also be an important element, particularly in fostering take-up amongst Member States through research opportunities. JRC has broad involvement in the recommendations for standards for SEIS, Inspire, GMES, and GEOSS. Also JRC has an interest in the future impact of ICTs on environmental sustainability, which Sensor Web Enablement could pay a role in.

The implementation plan is based upon demonstrating that SWE is viable for an existing reporting stream and demonstrating that the technology can be easily transposed to other reporting streams. In turn, industry and open source communities will see a demand and will respond with software products to facilitate the setting up and utilizing of such a flow.

## EEA roadmap for 2010

In 2010, a proof of concept project will be initiated by the EEA. The aim is to adopt the new protocol without changing the present infrastructure by opening up an additional communication channel, which works in parallel with the existing protocol. This will be the start to demonstrate the value/prove its worth. It will only work if it can be initiated as an interface to an existing system.

The work will be led by EEA and build on the results of the S@NY project where a SWE implementation has been demonstrated at the local level. The piloting phase involves a series of activities which are designed to assess the software against a pre-determined set of criteria, i.e. identify the issues. The first task is to implement the S@NY software at the EEA and integrate with an existing data flow in order to practically assess the service, specifically setting up a feed, for example Ozone exceedance reporting.

Austria has already implemented national air quality reporting through SWE technology and will be involved in the work as a test case. This test case will address predefined questions which would include:

* Can an existing monitoring infrastructure be upgraded at a payable cost?
* Can the concept be used without additional resources for maintenance?
* Will the technology make the international data flow more cost effective?
* What are the realised benefits?
* What are the risks and how can they be addressed?
* Does SWE have the needed potential for fulfilling data and metadata requirements for reporting obligations?

Any sensor system needs to maintain full control on who can get that information on the part of the data provider. Security is an essential part of the system.

One of the basic ideas behind this concept is that involved parties are loosely coupled together and as a consequence they will only participate if they can see an advantage in participation. A major part in evaluating the SWE concept is to see if national authorities are willing to adopt this technology. One premise for adoption is that there are available IT systems that the authorities can use, without starting large IT-projects without guarantee of success. The evaluation of SWE includes a survey of existing implementations and technology. This will require cooperation with the industry and the open source community.

# Further reading

A WIKI page has been established for the purpose of the SWE project at the EEA:

<http://svn.eionet.europa.eu/projects/SensorWeb/wiki>