Building bridges: experiences and lessons learned from the implementation of INSPIRE and e-reporting of air quality data in Europe

4 AUTHORS:

Alexander Kotsev
European Commission

11 PUBLICATIONS   14 CITATIONS

SEE PROFILE

Olav Peeters
Belgian Interregional Environment Agency

1 PUBLICATION   0 CITATIONS

SEE PROFILE

P. C. Smits
European Commission

89 PUBLICATIONS   805 CITATIONS

SEE PROFILE

Michel Grothe
Geonovum

16 PUBLICATIONS   21 CITATIONS

SEE PROFILE

RESEARCH ARTICLE

Building bridges: experiences and lessons learned from the implementation of INSPIRE and e-reporting of air quality data in Europe

Alexander Kotsev · Olav Peeters · Paul Smits · Michel Grothe

Abstract The collection, exchange and use of air quality data require diverse monitoring, processing and dissemination systems to work seamlessly together. These systems should supply data in a format usable for other applications such as planning, assessing population exposure and environmental impact assessment. As air quality does not change abruptly near national borders, international cooperation within this domain is highly desirable. This manuscript reports lessons learned during the implementation phase of one the world's largest data harmonisation effort of environmental information infrastructures, INSPIRE, a backbone of a European wide spatial data reporting system which involves an unprecedented number of actors and volumes of data. It is important within the context of the Digital Earth concept and the establishment of a global spatial data infrastructure (SDI) through the Global Earth Observation System of Systems (GEOSS), as the quality of ambient air is among the most pressing contemporary environmental problems. We summarise our findings from the perspective of national public authorities, obliged by law to transmit standardised spatio-temporal data to streamline reporting and facilitate the use of information, while keeping public expenditure under control. To develop implementation strategies for these type of e-reporting data flows we established a cross-border case study, looking at the process of harmonisation and exchange of data in Belgium and the Netherlands based on interoperable standards. Our results cover the legal, semantic, technological and organisational aspects of reporting. They are relevant to a cross-thematic audience having to undergo similar processes of reporting, such as climate change, environmental noise, marine, biodiversity, water management, etc.

Keywords Air quality e-reporting · Infrastructure for Spatial Information in the European Community (INSPIRE) · Shared Environmental Information System (SEIS) · Service oriented architecture (SOA) · Spatial data interoperability · Semantic interoperability

A. Kotsev (✉) · P. Smits
Digital Earth and Reference Data Unit, Institute for Environment and Sustainability, Joint Research Centre, European Commission, Via Enrico Fermi, 2749 Ispra, Italy
e-mail: alexander.kotsev@jrc.ec.europa.eu

P. Smits
e-mail: paul.smits@jrc.ec.europa.eu

O. Peeters
Belgian Interregional Environment Agency (IRCEL—CELINE), Brussels, Belgium
e-mail: peeters@irceline.be

M. Grothe
Geonovum, Amersfoort, Netherlands
e-mail: m.grothe@geonovum.nl

Introduction and background

Clear understanding and actions targeted at the improvement of the ambient air quality are among the key environmental challenges in an increasingly urbanized world. Therefore, local, regional, and global air quality issues, and regional and global environmental impacts, including climate change, should be viewed in an integrated manner (Akimoto 2003). Informed policy making about air quality related problems is dependent on huge volumes of spatio-temporal data, which in many cases need to be available across borders. Furthermore, the local release of pollutants into the atmosphere can have regional or global consequences. The source-receptor tables, the so-called "blame matrixes", annually published by European Monitoring and Evaluation Programme (EMEP) (see e.g. Schulz et al. 2013—Appendix C) quantify the enormous magnitude of this cross-border exchange of emissions most

systematically with a solid scientific base. The availability of (cross-border) air quality data is hence of fundamental importance to policy makers taking adequate measures, the experts in charge of forecasts predicting e.g. high pollution episodes, etc. Differences among monitoring systems and data availability are however also an important limitation in environmental management (D'Amore et al. 2012). In order to leverage the use such data for decision making, it should be seen within the context of the Global Earth Observation Systems (GEOSS) as a "system of systems", helping to refocus and making the existing systems and Earth observation coordination mechanisms more effective, thus making technology as interrelated as the planet it observes (Lautenbacher 2006), and unlocking all available data that is gathered to a centrally available resource like the internet, which permit access to use the data for the respective user's purpose (Butterfield et al. 2008).

This manuscript reports on the experiences and lessons learned during the implementation of the world's single largest data harmonisation effort for an environmental information infrastructure—the infrastructure for spatial information in Europe (INSPIRE), in support of the reporting requirements under the European air quality directive (2008/50/EC) and its implementing provisions regarding the reciprocal exchange of information and reporting on ambient air quality (2011/850/EU). "Introduction and background" provides background information and defines the legal and organisational context in which European countries have to exchange authoritative information on air quality. This section also describes the new mechanism for sharing air quality data using these new developments of information and communication technologies (ICT), further referred to as e-reporting. The subsequent section provides details on a cross-border case study between the Netherlands and Belgium in the harmonisation and exchange of air quality information to pilot test the overall process and available technology with real world data. The third section presents our results covering legal, semantic, technological and organisational aspects together, thus being fully aligned with the interoperability levels defined within the European Interoperability Framework for European public services (European Commission 2010).

Background

Achieving good spatial and temporal coverage of air quality data is difficult, as in-situ measurements stations are not dense enough. These measurement stations provide data for discrete point locations with a limited representativeness which necessitates the further application of modelling techniques for a better understanding of the spatial distribution of pollutants. The collection of volunteer information by citizens through smartphones and all sorts of different gadgets is at the same time fundamentally changing the traditional ways of data collection. Data for air quality coming from such devices however is not really comparable and hardly reliable. This leaves the existing governmental reporting systems and remote sensing platforms the only well-established source of reliable information on the ambient air quality. Hence the inclusion of air quality data coming from the existing well established authoritative sources is very important for addressing the challenges defined within the global environmental agenda and the related GEOSS. Furthermore, as air quality does not change abruptly on country borders and there is a well-documented cross-border exchange of emissions, it is essential that data for the policy making process are available in an easy to use and accessible form across national borders.

From a European perspective the decision on the European Parliament and of the Council on a EU Environment Action Programme to 2020 "Living well, within the limits of our planet"[1] - also known as the 7th Environmental Action Plan (EAP)—has as one of the priority objectives to improve the knowledge and evidence base for EU environment policy. As part of this objective, the use of common approaches and standards for acquisition and collation of consistent spatial information under the INSPIRE Directive[2] are promoted. Also promoted are the efforts to streamline reporting obligations under different relevant pieces of legislation; all of this contributes to minimising duplication of reporting efforts and increases transparency. In fact, the effective collection, sharing and reuse of information are key to implementing the 7th EAP. INSPIRE, SEIS—the European Shared Environmental Information System, and Copernicus—the European Earth Observation Programme—all contribute to this. This section of the manuscript provides an overview of all these initiatives, established on a European scale to leverage the discovery and usage of environmental information.

The availability of timely and accurate data is of critical importance for European policy making. The European Union has come a long way since the use of paper-based reporting which was used for instance in the early days of the Habitats and Birds EU Directives. In fact, data flows for a large number of reporting obligations are in place in Europe. The processes on a European level should hence, in the broader context of Digital Earth be seen as a valuable example for the rest of the world. Moreover, there are opportunities and synergies between the current policy priorities in Europe and the establishment of the Digital Earth (Annoni et al. 2011). However, on EU level there are still issues that hamper an easy, effective and forward-looking reuse of data and information which is needed to implement the objectives of the 7th EAP. These issues can be categorised in four levels of interoperability: legal, organisational, semantic and technical interoperability. As part of the Digital Agenda for Europe, the EU institutions and the Member States have been adopting a series of

---

[1] http://ec.europa.eu/environment/newprg/pdf/PE00064_en.pdf
[2] Infrastructure for Spatial Information in the European Community (INSPIRE), http://inspire.ec.europa.eu.

measures to improve these types of interoperability. EU Open Data strategy, Communication from the Commission towards interoperability for European public services (European Commission 2010), Interoperability framework and Interoperability Solutions for European Public Administrations (ISA) program are several examples providing solutions which address the problem on different levels.

Infrastructure for spatial information in the european community—INSPIRE

On October 21st 2013, the European Commission adopted what is probably the world's single largest data harmonisation effort for environmental information. It is the result of an effort of hundreds of experts from across Europe that have been working together for several years to agree upon common definitions in important policy areas such as energy, climate change, biodiversity, the marine environment and human health (cf Fig. 1). This legal act complements other INSPIRE legal acts and standards, that together form the basis of the Infrastructure for Spatial Information in the European Community (INSPIRE) envisioned by Directive 2007/2/EC. Now that most documents needed for the establishment of the infrastructure are agreed upon, the implementation of INSPIRE continues with Member State implementation and maintenance. Considering the deadlines, INSPIRE will be fully implemented by 2020.

Harmonisation of data within the INSPIRE infrastructure has to go beyond any particular community, but take the various cross-community information needs into account. In order to achieve that The Directive and Implementing Rules (Commission Regulation (EU) No 1253/2013) include legally binding requirements that describe, usually on an abstract level, what Member States must implement and additional detail is provided by non-binding technical guidance documentation.

The INSPIRE data themes with high relevance to air quality e-reporting are provided within Table 1. However, they should all be seen as parts of a single infrastructure with common encoding rules, and interdependencies between the individual themes, and the content of the table is provided in

order to illustrate the type of phenomena which are being represented within the spatial data infrastructure. Other relevant components, provided by INSPIRE which are shared between several or all data themes include the Generic Conceptual Model, use of Observations and Measurements encoding (ISO 19156:2011) for spatio-temporal data, etc.

It is also worth noting that INSPIRE uses international standards as building blocks of the European interoperability infrastructure, thus by implementation authorities add value to existing systems through harmonisation. A significant share of the standards which are being used in INSPIRE are created and maintained within the Open Geospatial Consortium (OGC) and the world's largest developer of standards, the International Standardization Organisation (ISO). The OGC is an international consortium of more than 400 companies, government bodies, research institutions and universities which are developing consensus-based standards predominantly in the geospatial sector.

In a broader context, to address the diverse issues in the domain of environmental information and their sharing, the European Commission together with EU Member States is working on the establishment of European Earth Observation programme - Copernicus (Regulation 911/2010 of the European Parliament and the Council. http://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:32010R0911) and SEIS. A SEIS communication document (European Commission 2008) lays down the foundations of how to better utilize and exchange environmental information on different scales, ranging from local to European. This framework document sets out the principles for better management of data to be adopted by Member States and European institutions. These cover topics related to the use of standards, open source software, citizen participation, and reusability of data. A SEIS staff working document of 2013 (http://ec.europa.eu/environment/archives/seis/pdf/seis_implementation_en.pdf) 18 final) goes into further detail, particularly identifying the shortcomings and priorities associated with the process of streamlining the exchange of environmental information.

Fig. 1 The INSPIRE legal acts and data standards harmonise spatial objects, key attributes and exchange formats of 34 data themes

Annex I
1. Coordinate reference systems
2. Geographical grid systems
3. Geographical names
4. Administrative units
5. Addresses
6. Cadastral parcels
7. Transport networks
8. Hydrography
9. Protected sites

Annex II
1. Elevation
2. Land cover
3. Ortho-imagery
4. Geology

Annex III
1. Statistical units
2. Buildings
3. Soil
4. Land use
5. Human health and safety
6. Utility and governmental services
7. Environmental monitoring facilities
8. Production and industrial facilities
9. Agricultural and aquaculture facilities
10. Population distribution – demography
11. Area management/restriction/regulation zones & reporting units
12. Natural risk zones
13. Atmospheric conditions
14. Meteorological geographical features
15. Oceanographic geographical features
16. Sea regions
17. Bio-geographical regions
18. Habitats and biotopes
19. Species distribution
20. Energy Resources
21. Mineral resources

Table 1 INSPIRE data themes which are related to air quality e-reporting

Data theme	Relevance
Human health and safety	Geographical distribution of dominance of pathologies and information indicating the effect of environmental conditions on human health and well-being.
Environmental monitoring facilities	Representation of the established stations for observing the quality of ambient air, organized in monitoring networks.
Population distribution - demography	Spatial distribution and quantitative information about the socio-demographic characteristics of population. Possible input for air pollution exposure modelling.
Area management/restriction/regulation zones & reporting units	Air quality zones which are defined in 2011/850/EU and are subject to reporting to the European Environmental Agency (EEA).
Administrative units	Representation of the administrative units and how they correspond to the air quality zones defined in 2011/850/EU.
Atmospheric conditions and meteorological geographical features	Spatio-temporal characteristics of the atmosphere which comes from the established meteorological monitoring networks.

E-reporting

Thematic legal acts part of the European Community environmental acquis together with the above mentioned initiatives create the framework for harmonisation and exchange of environmental information, covering at least the topics of air quality, waste, biodiversity, habitats, water quality and availability, climate change and noise. Air quality reporting authorities are among the first in Europe with the obligation to electronically report data they collect in a coordinated manner considering the requirements of the EU environmental acquis. Authorities have to meet their air quality obligations at the same time following the legally binding requirements of two directives—INSPIRE (2007/2/EC) and air quality (2008/50/EC), in accordance with implementing provisions with regards to the reciprocal exchange of information and reporting on ambient air quality (2011/850/EU). Institutions have to align their data to a common reporting model, encode data into the interoperable Geography Markup Language - GML (Open Geospatial Consortium 2007) and Extensible Markup Language—XML (World Wide Web Consortium 2008), establish data transformation routines and start providing air quality information to the EEA by the end of 2013. Satisfying what the two directives demand, while interfering as little as possible with existing workflows is an ambitious task and the added value is only visible in the mid, or even the long term.

In order to implement what the legal acts require ca. Fifteen volunteer governmental agencies responsible for the national report of their respective EU member state, the Directorate General 'Environment' of the European Commission, the European Environment Agency (EEA) and the European Topic Centre for Air Pollution and Climate Change Mitigation cooperated within an air quality pilot project. This included the development of a common e-

reporting data model and synchronized activities towards data exchange on a European level. A detailed overview of the approach undertaken on data modelling level is provided by Schleidt (2013). A parallel activity by the technical coordinator of the INSPIRE directive implementation—the EC Joint Research Centre (JRC) has been established in parallel. Particular attention in this activity is paid on the proper integration of INSPIRE into the reporting and the means to exchange data across borders within a service-oriented architecture (SOA), defined as an environment where loosely-coupled network resources are operationalised as independent services that can be accessed without any knowledge of their underlying implementation platform (Zhao et al. 2007).

E-reporting is a first encounter with INSPIRE of a large community, which is not traditionally involved in standardisation and spatial data interoperability. At the same time the reporting authorities are obliged by law to do one of the first mass implementations of INSPIRE. Considering this, we summarise our findings and lessons learned to support further implementations and improve e-reporting. Our findings can be beneficial in at least four directions:

- Support air quality reporting authorities through the identification of problems and present potential solutions;
- Document experiences for other communities which will be undergoing a similar process of coordinated e-reporting (marine, environmental noise, water management, etc.);
- Guide European institutions, most of all EEA and Directorate General "Environment" of the European Commission on how to better manage data flows within a complex multinational data infrastructure;
- Provides input for the evolution of GEOSS through identification of the applicability of a SDI-centred approach for sharing and the reuse of environmental data on national and international levels.

Approach

We looked, for the purposes of this paper, at the implementation of INSPIRE and e-reporting of ambient air quality data from the perspective of the responsible authorities in two Member State (Belgium and the Netherlands) and considered together the existing legal, organisational, semantic and technological challenges. This use case is important, as the two legally binded authorities—Belgian Interregional Environmental Agency (IRCEL—CELINE) and National Institute for Public Health and the Environment of the Netherlands (RIVM) decided to address the problem together, with the support of the Joint Research Centre of the European Commission as technical coordinator of the INSPIRE implementation. This approach provided us with a use case for the identification of all necessary changes to existing reporting workflows, necessary for compliance with the reporting obligations and exchange of environmental information by means of e-reporting. It also offered valuable insights into what does and what does not work in sharing air quality data in a cross-border context. Both Belgium and the Netherland are among the countries with the most sophisticated spatial data infrastructures in Europe. Details of the current state of SDI and the alignment with the requirements of the INSPIRE Directive are provided by Vandenbroucke and Biliouris (2011a, b). IRCEL—CELINE is responsible for aggregating air quality data collected by three regional networks to a coherent national report. Peeters and Fierens (2013) give overview of the internal architecture and the activities which IRCEL—CELINE undertaken to align the institutional infrastructure to the new reporting obligations. RIVM is the equivalent institution in the Netherlands mandated to report on air quality. RIVM experts are aided in aligning their internal procedures to INSPIRE e-reporting by the Dutch National Spatial Data Infrastructure executive committee—Geonovum. Both institutions extensively use free and open source software, thus benefiting from the private-collective model of innovation (Hippel and Krogh 2003) that open source delivers to users and innovators.

Redesigned dataflow

The internal workflow in both countries before implementing INSPIRE compliant e-reporting is composed of networks of measuring stations which collect information on air quality in near-real time and are organised in regional networks and urban agglomerations. The next steps include validation and quality assurance on regional level and consequent aggregation on national level. The aggregated information is then used for the preparation of a national report, being used to transmit information to the EEA (Fig. 2.). It is important within the illustrated data flow to clearly distinguished between data acquisition within the regional/sub-national measurement networks and e-reporting. This paper entirely focuses on the latter, hence we do not address questions of representativeness and data collection.

The preparation of the national air quality report (Fig. 2.) does not meet the criteria of the new INSPIRE compliant e-reporting, as it misses at least the following components:

- Common rules for encoding of data, once acquired for all air quality related features;
- Interdependence and synergy with other data providers, forming a building block of a spatial data infrastructure;
- Interoperable network services organized within a SOA allowing easy discovery, view and download of information, also in near-real time conditions.

These issues needed to be resolved within the workflows of both IRCEL—CELINE and RIVM and the following activities were undertaken in order to comply with the new e-reporting and INSPIRE requirements:

- Mapping elements of the existing production databases to the newly established common reporting model, described by Schleidt (2013) and available at www.eionet.europa.eu/aqportal/datamodel;
- Identification and testing of technology which offers interoperable network services for easy discovery, viewing and downloading of near-real time information, incl. for e-reporting purposes;
- Application of a reporting layer which encodes data into compliant GML/XML and establishes network services;
- Provision of feedback to the European Environmental Agency and the European Commission.

Accomplishment of the above activities resulted in a new organisational setup (Fig. 3.), put in place in order to meet e-reporting requirements and facilitate the use of web services as major means for data dissemination on the provider level. Data within this setup are made available through the following OGC web services: Web Feature Service—WFS (OGC 2010) for phenomena which don't change too frequently (e.g. annually) and Sensor Observation Service—SOS (OGC 2012) for data with high spatio-temporal variability. Through the use of interoperable web services users are free to choose what solution to use for consuming the resources. The use of services furthermore enhances the reusability of data.

Data transformation

From a technical perspective the redesign of the data flows requires the consequent use of several tools. The extract-transform-load (ETL) tool called HUMBOLDT Alignment Editor (HALE), described by Reitz and Templar (2012) has been the revelation of the inquiry in terms of mapping
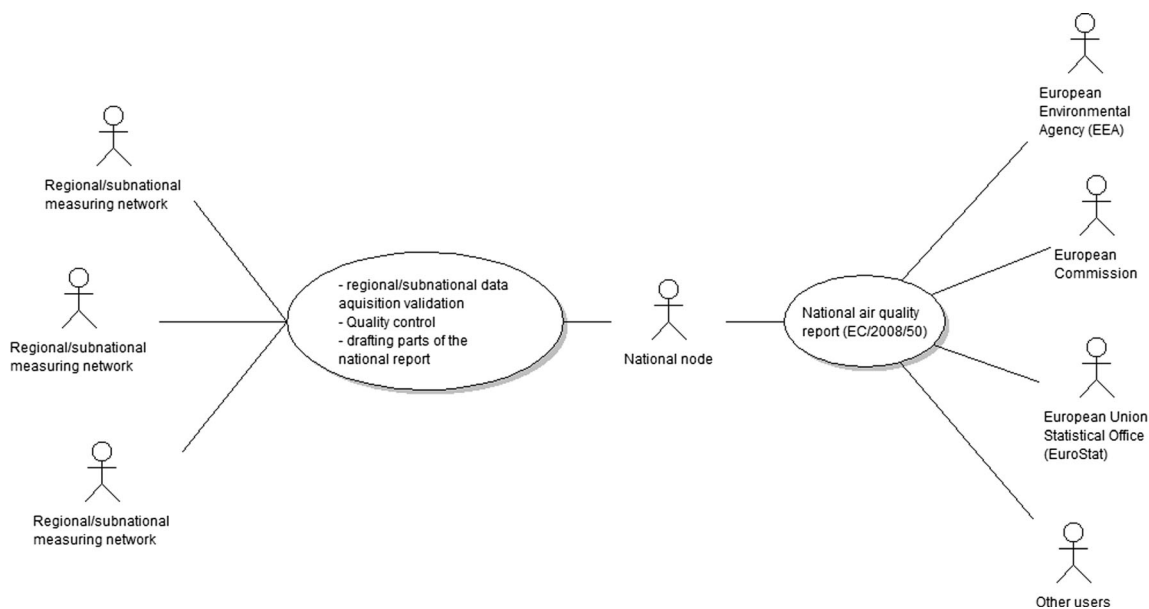
Fig. 2 Existing data flows at IRCEL-CELINE and RIVM before e-reporting

between the elements of the existing system and the redesigned INSPIRE compliant architecture (Fig. 4). HALE is platform independent, reliant on only the Java programming language, thus executables are distributable for the operating systems Linux, OSX and Windows. A configuration file can be exchanged conveniently between machines and platforms and data aggregation can be scheduled (e.g. via command line tools like crontab). Furthermore, HALE could also be extended to transform data from a SOS in a similar fashion, e.g. scheduled for hourly transformations of the latest available data. The configuration files for HALE which we used in order to transform source data in accordance with the e-

reporting model (Fig. 4) are made available at http://inspireaq.jrc.ec.europa.eu/hale/

Web services

To illustrate the value added from the establishment of web services and in order to verify the applicability of the established interoperable infrastructures in Belgium and the Netherlands we developed an experiment of data collection in near-realtime through the use of readily available technology (Fig. 5).
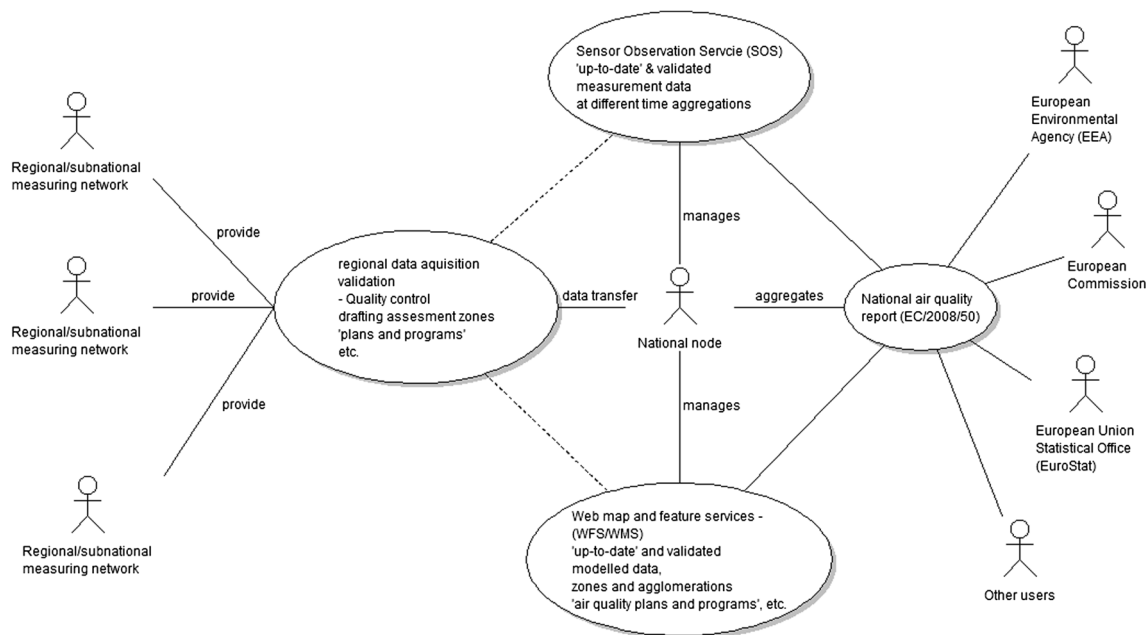

Fig. 3 Data flows after aligning to INSPIRE compliant e-reporting requirements
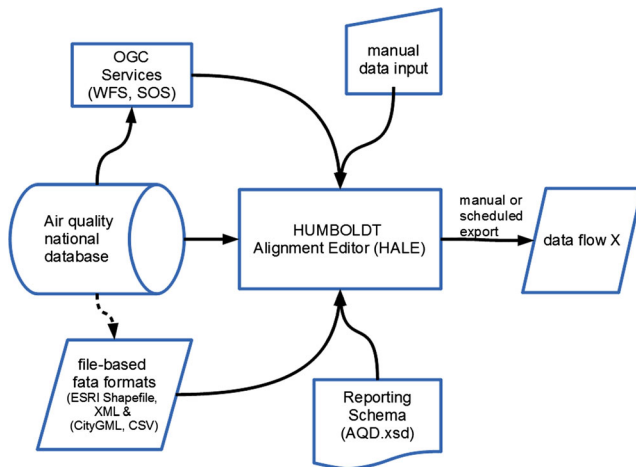
Fig. 4 Flowchart of a possible set-up of using the HUMBOLDT Alignment Editor for data integration and transformation necessary to compile the different data flows for e-reporting under the EU Air Quality Directive

We implemented a harvesting algorithm, capable of revisiting web service endpoints on an hourly basis, and collect observation data from the two exposed SOS instances in near-real time . The command line curl was identified as a solution to interact with SOS endpoints. In the case of our experiment curl uses xml, containing the request and storing the response within the client filesystem encoded as GML (Fig. 6).

The example xml syntax (Fig. 7) shows a HTTP POST method, which would request from the SOS server data for the latest available concentration of particulate matter (PM10) for all measurement stations in Belgium.

Through a consequent step (Fig. 8) the output GML was loaded within a spatially aware PostgreSQL/PostGIS database through the ogr2ogr tool, part of the OGR library (www.gdal.org/ogr).

Finally we deployed the 52° North implementation of SOS (www.52north.org/sos) for serving air quality data from the PostgreSQL/PostGIS database, and Geoserver (www.geoserver.org), used for deployment of other OGC services such as WMS and WFS for serving data for the point locations of the measurement stations and air quality zones subject to reporting.

Findings

We subdivide our findings from the implementation process into legal, organisational, semantic and technological, following thereby the categories used by the Interoperability Solutions for European Public Administrations programme (http://ec.europa.eu/isa).

Legal

We reviewed the legislation on air quality reporting in Europe together with what the INSPIRE implementation demands in order to identify how well the two correspond to each other. Results show that the advantage which INSPIRE and the spatial data infrastructure approach provide is considered within the air quality legislation through article 7 of 2011/850/EU. This is however the only reference to INSPIRE from the legal acts on air quality, and the rest of the content on the exchange of information is following an own logic. This is most evident in the expected mechanisms for exchange of data. We consider that the approaches for establishment of a successful spatial data infrastructure should be considered more on the stage of drafting environmental legislation which has to do with the exchange and coordinated use of data. Craglia et al. (2003) and Tóth et al. (2012) both provide a good starting point for principles and methods which should be taken into account for the establishment of spatial data infrastructures. The following INSPIRE founding principles should be considered throughout the environmental legislation in order to make better use of the advantages which data and service interoperability provides for domain specific policies[3]:

- Data should be collected only once and kept where it can be maintained most effectively.
- It should be possible to combine seamless spatial information from different sources across Europe and share it with many users and applications.
- It should be possible for information collected at one level/scale to be shared with all levels/scales; detailed for thorough investigations, general for strategic purposes.
- Geographic information needed for good governance at all levels should be readily and transparently available.
- Easy to find what geographic information is available, how it can be used to meet a particular need, and under which conditions it can be acquired and used.

Organisational

Firstly, the necessity to satisfy together the requirements of the two Directives stimulate collaboration in cross-border and cross-thematic context. This is evident by the Belgium-Netherlands cross-border case study, where involved institutions partnered by the European institutions established a partnership benefiting all involved parties.

Secondly, from an organisational perspective e-reporting has impact on the institutions and people and their knowledge.

[3] Source: INSPIRE web site: http://inspire.ec.europa.eu/index.cfm/pageid/48
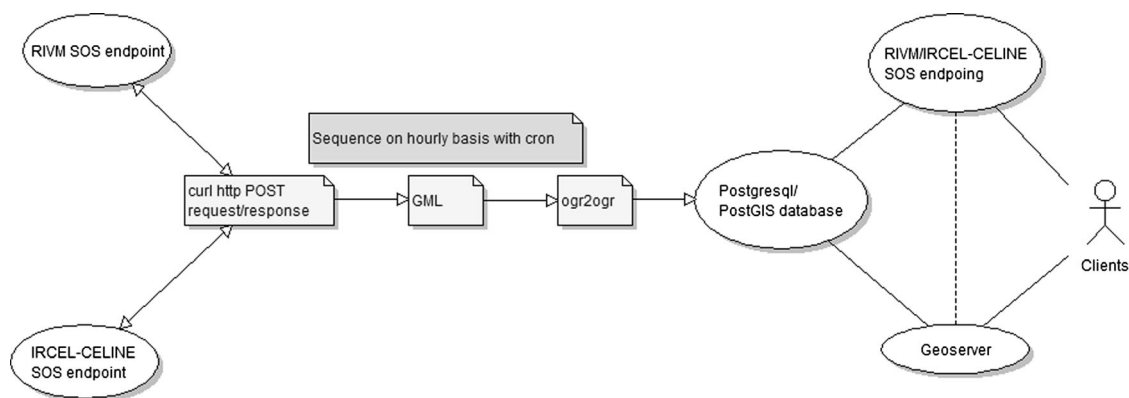
Fig. 5 Harvesting of SOS services through free and open source software, configured to work in a sequence through a cron job

The new e-reporting and interoperability principles, architecture and technology all need to be adopted by organisations and their employees, which is the most (resource) intensive and time-consuming aspect of adopting new technologies. Based on the experiences in this cross-border study, the following organisational aspects are of importance:

- Continuous involvement of the management board as resources in the described reporting organisations are becoming scarce, especially for data management and ICT;
- ICT experts are not always familiar with service oriented architectures and more specifically OGC and INSPIRE network services demand. Sometimes even completely new software choices need to be made and specific installations are needed, for instances for serving observations through SOS;
- Abstract models such as "Observations and Measurement" (ISO 19156:2011) in the e-reporting information needs practice and experience to understand and to work with.

Thirdly, currently e-reporting interferes with existing data flows, which contradicts one of the guiding principles of the establishment of Spatial Data Infrastructures, and most particularly the logic of INSPIRE. That is why we think that in the long term, reporting should consider superimposing an interoperability layer on top of existing architectures, rather than the require reporting authorities to have to redesign the existing workflows.

Semantic

We consider the ability to transfer data with an unambiguous meaning of critical importance for the success of e-reporting. Semantic interoperability in the presented case is directly dependent on the definitions within the involved legislative

acts and the common air quality reporting model that reflects them. We discovered discrepancies which need to be addressed in order to ensure the transmission of meaning together with data. The most striking difference we identified is related to the completely different perspective and meaning of the term 'metadata' within the air quality and INSPIRE communities, which is also reflected in the correspondent legislation, and can therefore not be easily interconverted.

Ensuring semantic interoperability inevitably requires mappings on several levels (between requirements of legal acts, and consequently between reporting systems and databases on European and national level). It is important to be able to conduct them in an unambiguous and easy to use manner to keep the burden at minimum. The open source HALE which we describe in Chapter 2 is proven to be fully capable of doing logically and semantically consistent mappings and consequent geodata transformation. It is in addition an interesting example where an ETL-tool is developed via a European Project (under the Sixth Framework Programme for Research and Technological Development) to facilitate the implementation of INSPIRE. As the common open source code base became more stable and complete it may well become a very handy "Swiss Army Knife" which can be used to bridge the gap between legacy (SDI) infrastructures within the agencies and the new requirements for e-reporting. It remains to be seen if the development community will get the necessary support in the long run to maintain the code and add new features as e-reporting evolves.

Technological

As e-reporting completely depends on technology it is critical to make the right choices. Our perspective in this context is that EU requirements for reporting should consider first the availability and maturity of technology with stimulation of innovation in mind, and avoid imposing a burden on Member

```
curl -X POST -d @GetObservationRequest.xml http://sos.irceline.be/sos > GetObservation.Response.xml
```

Fig. 6 Curl syntax for interaction with SOS server endpoints

Fig. 7 Syntax of a
GetObservation request to an
OGC SOS server

```
<GetObservation xmlns="http://www.opengis.net/sos/1.0"
 xmlns:ows="http://www.opengis.net/ows/1.1"
 xmlns:gml="http://www.opengis.net/gml"
 xmlns:ogc="http://www.opengis.net/ogc"
 xmlns:om="http://www.opengis.net/om/1.0"
 xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
 xsi:schemaLocation="http://www.opengis.net/sos/1.0
 http://schemas.opengis.net/sos/1.0.0/sosGetObservation.xsd"
 service="SOS" version="1.0.0" srsName="urn:ogc:def:crs:EPSG::4326">
 <!-- offering ID -->
 <offering>81102 - PM10</offering>
 <!-- select last time element -->
 <eventTime>
         <ogc:TM_Equals>
         <ogc:PropertyName>om:samplingTime</ogc:PropertyName>
         <gml:TimeInstant>
         <gml:timePosition>latest</gml:timePosition>
         </gml:TimeInstant>
         </ogc:TM_Equals>
 </eventTime>
 <observedProperty>81102 - PM10</observedProperty>
 <responseFormat>text/xml;subtype=&quot;om/1.0.0&quot;</responseFormat>
</GetObservation>
```

State authorities which bring little or no added value either to them, or to the European agenda. In this section of the paper we summarise our findings related to ICT.

What we discovered with the experiment described in Chapter 2 is that standards should be used in a clear and unambiguous manner in order to make the most of adopting them. This is important, as even a slight deviation from the use of a standard, for example for the encoding of data, or a web service will immediately lead to a decrease in interoperability. Such a decision is also very likely to limit the number of readily available software products, thus might impose additional burden on reporting authorities which would have to invest in developing their own solutions. We observed that the extended reporting model cannot be served through readily available implementations of OGC standards (WFS and SOS), as they require encoding which is different.

Secondly, we consider that the 'push' model for data transmission, currently foreseen as the preferred way of reporting air quality data, limits reuse of data, particularly when not explicitly following a widely used standard. Considering that both 'push' and 'pull' models have advantages and disadvantages we see the latter, accomplished through the establishment of a SOA and exposing necessary data where it is being produced, as the more suitable way of exchange of information when numerous data providers with different infrastructures are involved. The flexibility of the 'pull' model through readily available technology is illustrated by the example

provided in chapter 2.3. The 'push' model is however the preferred mode of data transmission for many Member States, also reflected in the relevant air quality legislation through Article 3(1) of 2011/850/EU, which states that "*The Commission, assisted by the European Environment Agency, shall establish a data repository and make it accessible through the ambient air quality portal*" and "*Each Member State shall nominate a person or persons responsible for the delivery on its behalf to the data repository of each reported and exchanged information*". This imposes the responsibility for aligning the deliveries and acts as an obstacle for the establishment of SOA. It limits the use of innovative solution such as SOS as major means for transmission of air quality data in near real time conditions, described by Jirka et al. (2012) and Kjeld et al. (2012). It is important to note that in a SOA one can technically develop a 'push' routine and ingest information in a transactional service or remote repository following a schedule from a 'pullable' service.

Thirdly, if governments want to keep their independence of commercial interests in data transformations, a lively community around an ETL-tool which is part of the public domain is crucial. We argue on the one hand that one should not lose sight of the available tools for the task during the development of a schema for reporting, but on the other hand, sometimes innovation is needed to better cater for all possible use cases a schema is needed for. If the necessary innovation is not seen as commercially interesting, proprietary ETL-tools will not

```
ogr2ogr -f "Postgresql" PG:"host=hostname dbname=sosdb user=username password=passwd" GetObservation.xml  -a srs EPSG:31370 -append
```

Fig. 8 Syntax for using ogr2ogr for loading data into PostgreSQL/PostGIS database

facilitate a solution. With the example of using HALE (Chapter 2.2), itself a product of a European research programme, we demonstrate how this innovation within an open source context can be fast and efficient. We consider these type of community controlled tools as indispensable for the implementation of INSPIRE and e-reporting. Furthermore, the use of a fully featured ETL-tool for e-reporting which is part of the public domain is that also commercially uninteresting functionalities necessary for small peculiarities of a reporting schema can be included into the code base with a very fast development cycle. It is not so much the 'free' of 'free of cost', but rather the 'free' of 'free speech' (Stallman 1999) which is important in this context. This freedom gives sufficient room for innovation to overcome the gap between existing infrastructures and the technical requirements of interoperable data for e-reporting. During our trails several functionalities related to the extensive use of external vocabularies were extended and released within a matter of days as new releases of the HALE-application.

Discussion and conclusions

The points below provide an overview of the challenges which should be addressed in order to streamline e-reporting and add value from it for both the institutions concerned and their users on national and international level.

General considerations for effective air quality e-reporting

As result of the inclusion of air quality data into a wider SDI providers involved benefit in several directions related to improved networking opportunities and visibility for work (Heumesser et al. 2011). However, as air quality e-reporting in the case of Europe inevitably imposes an extra responsibility over a large number of institutions, the requirements for reporting and the level of burden should be very well studied prior to the definition of the obligation. It should be preceded by a thorough study phase which considers capacities of authorities together with the state of their infrastructures and existing business processes. In an ideal situation e-reporting should not interfere with the existing workflows, but only provide a 'reporting' layer on top of existing infrastructures. The architecture of the reporting infrastructures should also clearly distinguish between the processes of data acquisition, validation and verification, and the subsequent reporting.

A clear focus should in addition be put on such practices which stimulate innovation and provide value added not only for the central level, but also for reporting institutions and their users. Moreover, readily available technology and the possibility for reusability of components should always be consider in order to keep investments of the reporting authorities to a

minimum. In our study we identified that open source software as the preferred development paradigm. It is easier to add any missing functionalities in cooperation with the development community maintaining these applications. The open source software, used during our trials is mature enough and has proven to be fully capable of meeting complex reporting obligations. It has successfully been implemented for mapping, transforming and serving data as described in Chapter 2. It can act as a real alternative to proprietary solutions providing all necessary means for building solutions based on internationally recognized standards. Moreover, it's use provides a good example which can be followed by other EU Member State public authorities, wanting to innovate and at the same time cut the costs of ICT. As e.g. Bouras et al. (2014) argue, open source can provide new business opportunities, create jobs, minimize research and development costs, and contribute to the development of ICT skills and e-inclusion in Europe in reaching its full potential.

Considering the architectures for e-reporting, traditionally many of the existing practices have been implemented with centralization of data and redistribution of a final product on behalf of those who contributed the data. Typical examples of this type of data dissemination in the case of Europe are Eurostat and the EEA. Now, with the emergence of the SOA as architecture it is technically completely feasible to adopt similar practices for data transmission on both country and international level. In order to do so international institutions should lead by example, even though that in many cases they are not legally binded to do so as are individual countries. This implies a common sense approach, considering the importance of leading the process by example. In this context European institutions should consider how to harmonize their data and provide access to them through compliant interoperable solutions. That would justify the requirements which are imposed on Member States in an unambiguous manner, not to mention the added value for data users form the provision of uniform data dissemination on different levels. A good example in this sense would be the EEA, supported by the Joint Research Centre of the European Commission to provide INSPIRE compliant network services which transmit environmental data, thus act as reference for Member State implementations. Such a process would be eased by the fact that the EEA data infrastructure is considering the ISO and OGC standards which are among the building blocks of INSPIRE.

As proven by the description of workflows in Chapter 2, making e-reporting possible involves numerous cross-cutting issues to be resolved, which require capacity of all actors involved to be raised. This covers both the political aspects and the understanding of the value added from connecting air quality data within the framework of a broader spatial data infrastructure, as well as all technological aspects which make that actually possible.

Evolution of existing initiatives

Significant amounts of manual reporting related activities would no longer be necessary after the initial investment of setting up an e-reporting system. Besides potential resource savings to comply with reporting obligations an important advantage of e-reporting is that this is a huge source of recently updated comparable information which INSPIRE itself needs to demonstrates it's added value. In the course of preparation of that paper however we discovered that IN-SPIRE and the requirements stemming out of the air quality specific legislation are not fully aligned, particularly with respect to whether data should be exposed through the establishment of distributed service oriented architecture, or centralized and disseminated through the EEA on behalf of the member countries. That is why supranational environmental legislation having to do with sharing and transmission of information should be better coordinated and lay down data exchange principles which are unambiguous. In the case of Europe, because of the broad perspective involved in synchronizing legislation on the exchange of environmental data we see the SEIS as the natural environment where such problems should be dealt with. In any case this should be discussed considering the broader question of how to better align the individual legal acts of the European Community environmental acquis.

There are several important lessons to be learned which can lead to improvement of INSPIRE implementation in the mid to long term which are discussed below and are related to (1) extending the focus of the data infrastructure beyond the geospatial domain, and (2) provision of working examples with data and services for their viewing, downloading and discovery. In relation to (1), INSPIRE is originally created as a framework directive targeted at 'unlocking' geospatial information. Being able to precisely localize objects in geographic space is important, but in many cases the temporal dimension is at least equally important. This is particularly true for phenomena such as air quality with significant spatial, but also temporal variability. That is why the use of technology which extends beyond the spatial domain and provides a spatio-temporal data model for data encoding and manipulation would be particularly useful for the INSPIRE evolution. Good candidates to be considered for provision of information in near-real time are the standards organized within the Sensor Web Enablement (SWE) framework. SWE in the Open Geospatial Consortium context refers to web accessible sensor networks and archived sensor data that can be discovered, accessed and, where applicable, controlled using open standard protocols and interfaces (Botts et al. 2008). The SOS is the component of SWE which is created for provision of observations and measurements in space and time. It is well suited for flexible delivery of data with spatio-temporal variability such as marine observations, meteorological data

and air quality. The SOS standard is already considered as a candidate by the Technical Guidance document for the establishment of INSPIRE download services as a possible extension (http://inspire.jrc.ec.europa.eu/documents/ Network_Services/Technical_Guidance_Download_ Services_3.0.pdf). The mapping of the requirements of the INSPIRE Directive and the specifications of the SOS, as they are defined by the Open Geospatial Consortium is already been proposed by Bröring et al. (2013). As for (2), real working examples providing harmonized data through interoperable services would be beneficial for those who have to implement INSPIRE. Considering the huge volume of legal and technical documentation which are being produced, the evolution of INSPIRE needs to be complemented by real examples, to ease implementations, which would often be conducted by groups for which the topics of spatial data infrastructures and interoperability are completely unknown or are a side activity. That would be fully aligned with one of the principles of SEIS which states that '*information should be supported through common, free and open software standards*', and lead to synergies with other EU-wide initiatives, such as EU initiatives for building open ICT systems by making better use of standards (European Commission 2013).

In conclusion, identifying a suitable way of sharing the lessons learned by institutions implementing INSPIRE can significantly benefit other communities or equivalent institutions in other regions of the world. As was stated in Chapter 3 there are many similarities between infrastructures in different countries which make it feasible to share developments and lessons learned. A wiki approach might be helpful in that sense, but the question of incentives for those who contribute remains to further be elaborated.

Finally, we consider the experiences from the implementation of INSPIRE and e-reporting for air quality information provides a valuable example for other EU-wide reporting initiatives on the governance model around data on a European level. In the broader context of Digital Earth the manuscript contributes to a better understanding of how governmental institutions should address the exchange of environmental information in order to be able to 'plug' it into the framework of GEOSS. The lessons learned should be used as a baseline for planning and implementation of cross-border and/or crossthematic e-reporting and at the same time complying with the requirements of supranational legislation. Furthermore, we consider the findings made in Chapter 3 of high relevance to both the air quality community and other domains having to report data to centralized bodies, and at the same time satisfying a diverse set of other requirements related to data dissemination.

References

Akimoto H (2003) Global air quality and pollution. Science 302(5651): 1716–1719. doi:10.1126/science.1092666

Annoni A, Craglia M, Ehlers M, Georgiadou Y, Giacomelli A, Konecny M, Ostlaender N, Remetey-Fülöpp G, Rhind D, Smits P, Schade S (2011) A European perspective on Digital Earth. Int J Digit Earth 4(4):271–284. doi:10.1080/17538947.2011.582888

Botts M, Percivall G, Reed C, Davidson J (2008) OGC® Sensor Web Enablement: overview and high level architecture. GeoSens Netw Lect Notes Comput Sci 4540:175–190. doi:10.1007/978-3-540-79996-2_10

Bouras C, Filopoulos A, Kokkinos V, Michalopoulos S, Papadopoulos D, Tseliou G (2014) Policy recommendations for public administrators on free and open source software usage. Telematics Inform 31(2): 237–252. doi:10.1016/j.tele.2013.06.003

Bröring, Jirka AS, Kotsev A, Spinsanti L (2013) Making the sensor observation service INSPIRE compliant, INSPIRE Conference 2013, WWW document http://inspire.jrc.ec.europa.eu/events/conferences/inspire_2013/schedule/submissions/236.pdf

Butterfield M, Pearlman J, Vickroy S (2008) A system-of-systems engineering GEOSS: architectural approach. IEEE Syst J 2(3). doi:10.1109/JSYST.2008.925973

Commission Communication towards a Shared Environmental Information System - SEIS 2008. (COM(2008) 46 final). http://eur-lex.europa.eu/legalcontent/EN/TXT/?uri=CELEX:52008DC0046

Commission implementing decision of 12 December 2011, laying down rules for Directives 2004/107/EC and 2008/50/EC of the European Parliament and of the Council as regards the reciprocal exchange of information and reporting on ambient air quality (2011/850/EU). http://eurlex.europa.eu/legal-content/EN/TXT/?uri=CELEX:32011D0850&qid=1400686873702

Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Region, COM(2013) 455 final, 2013. Against lock-in: building open ICT systems by making better use of standards in public procurement. http://eur-lex.europa.eu/legalcontent/EN/TXT/?qid=1400686961073&uri=CELEX:52013SC0224

Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions towards Interoperability for European public services (COM(2010) 744 final). http://eur-lex.europa.eu/legalcontent/EN/TXT/?qid=1400687060664&uri=CELEX:52010DC0744

Craglia M et al. (Eds) (2003) Geographic Information in the Wider Europe. Final Technical Report of the Geographic Information Network in Europe project (GINIE) IST-2000-29493. http://www.ec-gis.org/docs/F22908/GINIE_BOOK.PDF

D'Amore F, Cinnirella S, Pirrone N (2012) ICT methodologies and spatial data infrastructure for air quality information management. IEEE J Sel Top Appl Earth Obs Remote Sens 5(6):1761–1771. doi:10.1109/JSTARS.2012.2191393, Article number 6291761

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). http://eurlex.europa.eu/legal-content/EN/TXT/?qid=1400687098123&uri=CELEX:32007L0002

EU Shared Environmental Information System - Implementation Outlook, Commission Staff working document. (SWD(2013) 18 final).

European Interoperability Framework (EIF) for European public services. 2010. Annex 2 to the Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of Regions 'Towards interoperability for European public services'. COM(2010) 744 final. http://ec.europa.eu/isa/documents/isa_annex_ii_eif_en.pdf

Heumesser C, Fritz S, Obersteiner M (2011) Identifying benefits and challenges arising from the voluntary provision of GEOSS. 34th International Symposium on Remote Sensing of Environment - The GEOSS Era: Towards Operational Environmental Monitoring. http://www.isprs.org/proceedings/2011/isrse-34/211104015Final00830.pdf

Hippel E, Krogh G (2003) Open source software and the 'private-collective' innovation model: issues for organization science. Organ Sci 14:209–223

Jirka S, Bröring A, Kjeld P, Maidens J, Wytzisk A (2012) A lightweight approach for the sensor observation service to share environmental data across Europe. Trans GIS 16(3):293–312. doi:10.1111/j.1467-9671.2012.01324.x

Kjeld P, Jirka S, Bröring A (2012) A sensor web for Europe. GEO Connexions 11(4):36–37

Lautenbacher C (2006) The global earth observation system of systems: science serving society. Space Policy 22(2006):8–11

Open Geospatial Consortium (2007) OpenGIS Geography Markup Language (GML) Encoding Standard. WWW document http://portal.opengeospatial.org/files/?artifact_id=20509

Open Geospatial Consortium (2010) OpenGIS Web Feature Service 2.0 Interface Standard. WWW document http://portal.opengeospatial.org/files/?artifact_id=39967

Open Geospatial Consortium (2012) OGC® Sensor Observation Service Interface Standard. WWW document https://portal.opengeospatial.org/files/?artifact_id=47599

Peeters O, Fierens F (2013) INSPIRE and e-reporting: building a SEIS for air quality data - a case from Belgium, INSPIRE Conference 2013, http://inspire.jrc.ec.europa.eu/events/conferences/inspire_2013/schedule/submissions/269.odt

Regulation (EU) No 911/2010 of the European Parliament and of the Council of 22 September 2010 on the European Earth monitoring programme (GMES) and its initial operations (2011 to 2013)

Reitz T, Templer S (2012) An environment for the conceptual harmonisation of geospatial schemas and data. In: Multidisciplinary research on geographical information in europe and beyond. Proceedings of the AGILE' 2012 International Conference on Geographic Information Science, Avignon, April, 24–27, 2012, pp 63–68. WWW document http://www.agile-online.org/Conference_Paper/CDs/agile_2012/proceedings/papers/Paper_Reitz_An_Environment_for_the_Conceptual_Harmonisation_of_Geospatial_Schemas_and_Data_2012.pdf

Schleidt K (2013) INSPIREd air quality reporting. European air quality e-reporting based on INSPIRE. IFIP Adv Inf Commun Technol 413: 439–450. doi:10.1007/978-3-642-41151-9_41

Schulz M, Gauss M, Benedictow A, Jonson JE, Tsyro S, Nyíri Á, Simpson D, Steensen BM, Klein H, Valdebenito Á, Wind P, Kirkevåg A, Griesfeller J, Bartnicki J, Olivié D, Grini A, Iversen T, Seland Ø, Semeena VS, Fagerli H (2013) Transboundary acidification, eutrophication and ground level ozone in Europe in 2011. EMEP Status Report 2013; August 7, 2013. WWW document http://emep.int/publ/reports/2013/EMEP_status_report_1_2013.pdf. See http://www.emep.int/publ/common_publications.html for updates of this status report

Stallman RM (1999) The GNU operating system and the free software movement. In: DiBona C, Ockman S, Stone M (eds) Open sources: voices from the open source revolution. O'Reilly, Sebastopol, pp 53–70

Technical Guidance for the implementation of INSPIRE Download Services, ver. 3.0, WWW document

Tóth, Portele KC, Illert A, Lutz M, Nunes de Lima M (2012) A conceptual model for developing interoperability specifications in spatial data infrastructures. JRC Ref Rep. doi:10.2788/20697

Vandenbroucke D, Biliouris D (2011) Spatial data infrastructures in The Netherlands: state of play, 2011, WWW document http://inspire.jrc.ec.europa.eu/reports/stateofplay2011/rcr11NLv123.pdf

Vandenbroucke D, Biliouris D (2011) Spatial data infrastructures in Belgium: state of play, 2011, WWW document http://inspire.jrc.ec.europa.eu/reports/stateofplay2011/rcr11BEv102.pdf

World Wide Web Consortium (2008) Extensible Markup Language (XML) 1.0 (Fifth Edition) WWW document http://www.w3.org/TR/REC-xml/

Zhao J, Tanniru M, Zhang L-J (2007) Services computing as the foundation of enterprise agility: overview of recent advances and introduction to the special issue. Inf Syst Front 9:1–8