# Overview of the FP7 Project EO2HEAVEN "Earth Observation and Environmental Modelling for the Mitigation of Health Risks"

Kym Watson[1], Jose Lorenzo[2], Ingo Simonis[3]

[1] Fraunhofer Institute of Optronics, System Technologies and Image Exploitation
IOSB, Fraunhoferstr. 1, 76131 Karlsruhe, Germany
kym.watson@iosb.fraunhofer.de
http://www.iosb.fraunhofer.de
[2] Atos
jose.lorenzo@atos.net
http://atos.net/en-us/home.html
[3] Open Geospatial Consortium
ingo.simonis@igsi.eu
http://www.opengeospatial.org

**Abstract.** The project EO2HEAVEN "Earth observation and environmental modelling for the mitigation of health risks" advanced knowledge on the impact of environmental factors on public health outcomes. The multidisciplinary and user-driven project approach focused on the effect of atmospheric pollution (in case studies in Durban / South Africa and Saxony / Germany) on cardiovascular and respiratory diseases and the waterborne disease cholera (in a case study in Uganda). EO2HEAVEN has developed methodologies, models, spatial data services (using OGC standards) and applications supporting the main activities involved in environmental health: discovery and acquisition of environmental data, integration of heterogeneous Earth observations (satellite, in-situ and field data), development of models of health effects, development of risk maps and predictions for early warning systems. EO2HEAVEN has specified and implemented a Spatial Information Infrastructure (SII). This is an open architecture based on international standards and geospatial web services supporting the large-scale initiative GEOSS of GEO.

**Keywords:** health risks, environmental models, spatial information infrastructure

## 1 The EO2HEAVEN Project

EO2HEAVEN was funded in the European Community's Seventh Framework Programme (FP7/2007-2013) in the Theme "Environment (including climate change)" under grant agreement 244100. The project was coordinated by Fraunhofer IOSB with 10 partners from the European Union and 3 from Africa. The project duration was 02/2010 - 05/2013. This large scale integrating project had a cost budget of about 8.7 Meuro and a EU-funding budget of about 6.3 Meuro.

The project EO2HEAVEN took a multidisciplinary and user-driven project approach focused on three case studies:

- The impact of air quality on respiratory and cardiovascular diseases in Saxony, Germany
- The relationship between industrial pollutant exposure and adverse respiratory outcomes in Durban, South Africa
- The links between environmental variables and cholera outbreaks in the Kasese district of Uganda.

Public health stakeholders and practitioners were actively involved in the project activities and worked with technology and service providers in the areas of Earth observation (EO) and environmental monitoring.

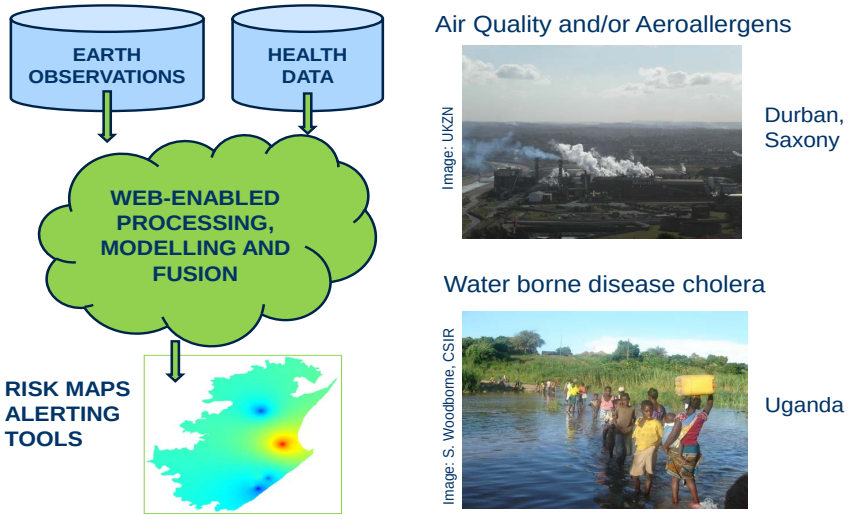


**Fig. 1.** Combining Earth observations and health data to produce risk maps and early warning systems

As illustrated in Fig. 1, EO2HEAVEN has developed methodologies, correlation models, spatial data services (using Standards of the Open Geospatial Consortium) and applications supporting the main activities involved in environmental health: discovery and acquisition of environmental data, integration of heterogeneous Earth observations (satellite, in-situ and field data) and development of models of health effects in correlation with the environment. The final result is the development of risk maps and predictions for early warning systems.

EO2HEAVEN has specified and implemented a Spatial Information Infrastructure (SII) to support the above methodological steps. This is an open architecture based on international standards and geospatial web services supporting the large-scale initiative GEOSS of GEO (Group on Earth Observations). These results are documented in publicly available reports and best practice documents, cf. www.eo2heaven.org.

EO2HEAVEN contributed best practices and proof of concept implementations to GEOSS pilot activities. EO2HEAVEN had strong interactions with the GEO Community of Practice "Health and Environment" and led the SBA health activities in the GEO Architecture Implementation Pilot phase 5 (AIP-5). Special emphasis was placed on achieving sustainability through capacity building in stakeholder and user training workshops in Uganda, South Africa and Germany. For details see the project web site www.eo2heaven.org. The EO2HEAVEN book gives an easy-to-read overview of all project work [1].

## 2 Project objectives and main results

The two high-level project objectives of EO2HEAVEN were 1) the development of a methodological approach for cross-domain analysis of environment and health questions; and 2) to contribute to and support the large-scale initiatives GEOSS and INSPIRE. A number of specific objectives were defined to achieve measurable progress towards these high-level objectives as described below. The public reports referred to are available at `http://www.eo2heaven.org/category/documents-categories/public-deliverables`.

1. To develop a shared understanding of the problems between the interdisciplinary research groups In order to facilitate the cross-domain work several tasks and associated reports were achieved:
   - The project took a multidisciplinary approach involving experts from the domains of health, epidemiology, microbiology, geo-informatics, ICT, modelling and statistics.
   - The three Case Studies shared a common structure and delivery planning split into three iterations. In collaboration with users, a detailed specification of the use cases for each of the EO2HEAVEN development cycles was provided. The specifications included functional and non-functional requirements, criteria for the validation and descriptions of the technical environment. In all, three full iterations of the Case Studies and use cases specifications were completed. For collecting and harmonising the user requirements from each use case specification there was a specific task producing a Harmonised use case specification and user requirements report after each cycle.
   - To organise cross-domain workshops that address research topics from an environmental, health and ethical perspective as well as the geospatial technology and data viewpoint. The first training and stakeholder workshops of the project were held in South Africa on 7-10 November

2011. They were directed towards training on the project's intermediate results related to data integration and software use. The results of the workshops are documented in detail in specific reports with the On-line Training materials and Training Workshops.
   - A second series of cross-domain stakeholder and training workshops addressing research topics from an environmental, health and ethical perspective as well as the geospatial technology and data viewpoint was organized and carried out with success in Uganda and South Africa. The participants encompassed IT and GIS experts, environmental scientists and experts in the field of health and the environment, as well as medical and health practitioners. An international group of cholera experts met under the auspices of EO2HEAVEN and now plan a longer term collaboration.
   - EO2HEAVEN participated in a number of GEO related meetings that were attended by experts from many of the SBA domains. This promoted the exchange of experience and approaches.
2. Establish a new quality for environmental health studies. To develop new data fusion methodologies to integrate environmental data from EO and various other sources into an information product that is best suited for environmental health studies in the EO2HEAVEN Case Studies
   - The work package about Environmental Monitoring for Health Applications provided recommendations on fit-for-purpose environmental information products and suitable and robust methods and models for risk map production or disease propagation simulations. The overall experience and gained knowledge in dealing with environmental data for health applications has been compiled into a set of public reports:
     - *D3.11 Catalogue of EO/EI Data Products*
     - *D3.13 Methodologies for Health and Environment data fusion and data mining*
     - *D3.14 Processing Chains Prototypes*
     - *D3.15 Environmental Monitoring for Health Applications* (provided at the end of the project as an overall compilation of results)
   - The task Synthesis and Recommendations put a strong emphasize on summarising the results in relation to the Case Studies. The objective was to collect information on the current status and experiences and provide an overview of activities and results which cross the boundaries of the Work Packages, but need to be treated as unified topics. This report is structured into three major sections: Definitions, project approach and case study outline; overview of health data and environmental data in the context of the case studies; and best practice recommendations for environment and health data correlation. This is available as public deliverable *D3.15 Environmental Monitoring for Health Applications.*
   - The final project results have been compiled in a published book available as PDF or as a print medium [1].
3. Ensure that stakeholders know how to use the EO2HEAVEN data products and tools: organise training workshops that address the EO2HEAVEN Case

Studies and provide practical tool exercises, and provide on-line materials for e-learning
   - For the stakeholder training workshops held in South Africa during 7-10 November 2011, a collection of presentations and exercises was developed. Cf. http://www.eo2heaven.org, subfolder "workshops".
   - A second iteration of stakeholder and training workshops was held in Uganda and South Africa in February 2013:
     - The workshops in the cholera prone Kasese district of Uganda targeted local health workers.
     - The workshops in Kampala and Pretoria were designed for health and environment experts and responsible people at system organisation level. They focussed on methods and issues around the establishment of information and warning systems.
     - The workshop in Durban was directed at the local stakeholders from the eThekwini municipality, providing training on the use of an environmental health system
     - These events allowed us to strengthen the links established with the organisations and explore how to make the results sustainable. The various training materials and presentations are available on-line.
4. Specify, develop and validate methods for extraction of Environmental parameters based on EO data and In-Situ sensors
   - Progress in these tasks is mainly reflected in the public report *D3.13 Methodologies for Health and Environment data fusion and data mining*. This document describes health and environment data fusion and data mining methodologies that are considered useful in the field of health and the environment. It provides guidance for the appropriate preparation, use and fusion of environmental in-situ and remote sensing data, and both forms of environmental data with health data. More precisely, the objective of this document is to identify and present relevant methods and tools to detect these correlations. The first section on the methods for health and environmental data fusion and mining describes time series extractions. Subsequent sections focus on spatial data fusion; air quality modelling; statistics such as descriptive statistics, non-spatial cluster analysis); validation, e.g. validation of remote sensing data, validation of health data, cross covariance analysis and the assessment of environment-health relationships; spatial statistics and proxy data. Information is also given on pre-processing methodologies, such as interpolation, aggregation, analysis of field and laboratory data.
   - The description of scenarios includes the specific workflows and processing steps of each scenario. Thus each scenario contains all necessary input data, the processing steps and finally the results.
5. Contribution to the objectives of the health task HE-01 Tools and Information for Health Decision-Making in the GEO workplan 2012-2015 (`http://www.earthobservations.org/geoss_imp.php`):
   - Develop tools and information systems for the environment and human health.

- Advance the integration of Earth observations and forecasts into health decision-making processes.
- Engage with health users and decision-makers to identify needs.
- Carry out capacity building and the promotion and sustainable use of Earth information by the health user-community.
- Establish linkages with other Societal Benefit Areas such as Ecosystems, Biodiversity, Climate and Disasters

EO2HEAVEN contributed to the activities on components C1 (Air-borne Diseases, Air Quality and Aeroallergens) and C2 (Water-borne Diseases, Water Quality and Risk) in this HE-01 task. The EO2HEAVEN results are recorded in the corresponding GEO component sheets. EO2HEAVEN contributed to the definition of HE-01 and also led the health thread in AIP-5 (part of task IN-05 GEOSS Design and Interoperability) in the GEO Workplan 2012-2015. The EO2HEAVEN capacity building activities are recorded in the results of the GEO task ID-01 "Developing Institutional and Individual Capacity". The realisation of a cholera early warning system with components and knowledge from EO2HEAVEN is now being actively pursued together with the pivotal organisations WHO (coordinator of GEO task HE-01) and NOAA (point of contact for the component HE-01-C2 on water-borne diseases). This activity initiated in EO2HEAVEN will continue beyond the project. EO2HEAVEN has shown involvement in the health task HE-01 through its continued participation in the GEO Health and Environment Community of Practice by representing the Health and Environment CoP at a meeting of the GEO Integrated Global Water Cycle Observations CoP. This will strengthen the linkage between the SBAs Health and Water. The participation in other GEO events has contributed to an increased awareness of the thematic links to other SBAs.

6. To provide an open and generic Spatial Information Infrastructure (SII) architecture: base the architecture on open specifications; continue and expand the architectures specified by the FP6 projects ORCHESTRA and SANY; assure that the SII architecture complies with international standards; active and early contribution to standardisation working groups and feed specification and implementation experience back to standards bodies; include innovative technologies into the SII architecture and infrastructure
   - The EO2HEAVEN Spatial Information Infrastructure (SII) is designed as a multipart document structured in heavily interlinked reports *Specification of the Spatial Information Infrastructure*, *Advanced Sensor Web Enablement Concepts* and *Advanced Distributed Geo-Processing and Spatial Decision Support*, such that there is a coherent and redundancy-free set of architectural specifications.
   - The SII continues the series of architecture specifications of the previous FP6 European project ORCHESTRA and it builds upon agreed specifications of a geospatial service-oriented architecture (SOA) provided by ORCHESTRA and extended by a Sensor Service Architecture of the SANY project. The EO2HEAVEN extensions include topics on remote sensing, health data (access control, integration with environmental

data) and mobile applications. EO2HEAVEN has submitted part of the SII to OGC as a proposal for an OGC Best Practices Document for Sensor Web Enablement - "Provision of Observations through an OGC Sensor Observation Service (SOS)" (OGC 13-015, status: positive voting result). This paper provides recommendations to simplify the provision of observation data based on the SOS 2.0 standard, especially for scientists and non-IT experts.
  - The fourth issue of the SII was released as a set of public documents in April 2013: *D4.13 Specification of the SII Implementation Architecture (issue 4)*, *D4.14 Specification of the Advanced SWE Concepts (issue 4)* and *D4.15 Specification of the Advanced Geo-processing Services (issue 4)*.
7. To provide an operational infrastructure: iterative and cyclic development with explicit validation by end-users in each cycle
  - The SW- Component Development & Integration work package structured its work into generic component development and specific component development for the three case studies. Attention was paid to technology for the different user classes (scientist vs. end user such as a policy or decision maker) and with different end devices (desktop with good internet connection vs. mobile platform such as smart phone or tablet, possibly with intermittent connectivity). Several generic components are available as open source software, thus setting up an environment to foster their sustainable development.
  - Aligned with project work, EO2HEAVEN responded to the GEOSS AIP-5 Call, following the collaboration already initiated within the previous Architecture Implementation Pilots AIP-3 and AIP-4, with contributions to the following threads: (1) to lead the SBA Health Air Quality and Waterborne thread and (2) to provide a number of components and data sets to be used in the same SBA Health thread. The main aim was to facilitate the work of scientists by making data sets directly available via standardized interfaces. The activities have been complemented by a contribution of tutorials and best practice guides.
  - A mobile application for health data collection in Uganda was specified and implemented [2]. Currently, the recording of new cholera cases is an error prone and inefficient process. The data gets copied from one handwritten paper form to another, sometimes digitized and aggregated several times, and reported in weekly or monthly intervals only. The EO2HEAVEN client to record cholera cases on a mobile tablet computer shall facilitate the data acquisition process during the registration and further reporting of cholera patients.

## 3 Aspects of the EO2HEAVEN SII Architecture

EO2HEAVEN considered requirements from both the environmental and health domains. In comparison to the environmental domain, there are several essential differences in the health domain:

– information models (e.g. including epidemiological and diagnostic data)
– humans as sensors (issuing a diagnosis)
– reporting procedures: health data is aggregated and passed through a chain of authorities
– quality and reliability of data
– depending on national law, health data is subject to strict security and access constraints

EO2HEAVEN also tackled the partly conflicting requirements of the main players in the health and environment field as shown in Table 1.

**Table 1.** Conflicting requirements of the players

| Group | Favourite Tools | Vision |
|---|---|---|
| IT experts | XML, Java, OGC standards, Databases | Generic results |
| Scientists | Excel, R, MATLAB, CSV, maps | Methods and models transferrable to similar problems |
| Practitioners | Web clients, black box Decision Support System | Optimized for own organization, but extendable when new requirements emerge |

The OGC services used by the IT specialists comprised SOS (Sensor Observation Service to store and access observations / measurements), SPS (Sensor Planning Service) and WPS (Web Processing Service) to parameterize and schedule models, WMS (Web Map Service) and WFS (Web Feature Service). The Observation & Measurement Model and Sensor Model Language were the two fundamental information models. The software components and architectures implemented in the three case studies are described in the public document *D5.17 Final specific components package (documentation).*

The two central questions were: a) How can these "elephant-like" standards be harnessed for data access, processing and publication? and b) And that with good performance in distributed systems, with poor internet connection?

The following gives a compact summary of the main topics covered in the SII architecture. It is a step towards distributed geo-processing and spatial decision support incorporating models from the environmental and health domains and enabling processing workflows for scientists. The use of sensor web technology for analysing correlations between health and environmental data is looked at in [3]. For further background on using OGC standards for modelling and scientific workflows see [4]. The challenges of environmental and health studies in the case of cholera research are explained in [5]. For challenges in analysing health risks due to air pollution, see [6]. A technique for the spatial interpolation of air quality data based on a regional classification is presented in [7].

– *Part I: Overview*

- *Part II: SII Enterprise Viewpoint* A summary of requirements, in particular from the case studies relating to issues of data and information, data quality, security and user management, privacy of health data, events, processing and fusion, decision support, integration of in-situ and EO data, and environmental and health data
- *Part III: Sensor Service Architecture* Based on FP6 SANY SensorSA, and adapted to latest standards, including models of interaction (request / reply, event based), management of resources, meta information approach, security with a focus on access control, considering specific EO2HEAVEN requirements, e.g. for health risk maps
- *Part IV: SII Engineering and Technology Viewpoint* Specification of the architectural patterns and policies for the setup and operation of the SII, e.g.: resource discovery for SOS observations and sensors, handling of large data sets, processing and fusion support, processing of quality information (in-situ and EO), attachment of quality information, handling and visualization of uncertainty information.
- *Part V: Advanced Sensor Web Enablement Concepts* The purpose of Part V was to facilitate the practical use of the OGC Sensor Web Enablement (SWE) framework, to find a balance between a) full generality with complete metadata and b) ease of use. With regard to the handling of remote sensing data, this part includes a guide on metadata, on mapping of data to regular grids for efficiency and on usage of SensorML for sensor description. A study on using drones for gathering EO Data is also in Part V. Part V also analyses the health data format SDMX-HD in conjunction with the SWE framework. A universal mapping of SDMX-HD to O&M is not possible.
  A user interface to define rules for event handling within the SWE Framework was defined.
  Semantic aspects going beyond the immediate needs of EO2HEAVEN were also considered with work relating to a Semantic Catalogue, the Semantic Sensor Network Ontology (SSNO) and linked data.
- *Part VI: Advanced Distributed Geo-Processing Services* A focus was placed on making the geo-processing functionality of Software environments such as GRASS, R and MATLAB available for use in SII (with a wrapper concept) and on strengthening SWE integration (parsing SOS data for the WPS). A concept of geo-processing repositories was introduced with a moving code paradigm (improving performance) [8], [9], download functionality for off-line usage and transparent and collaborative workflow development (researcher stays in control, no black box). Experiments to improve performance (Cloud Computing, using Graphical Processing Units GPU) are also reported on.

## 4 Conclusions and future work

The integration of remote sensing and in-situ sensing data is still a major challenge due to the different measurement procedures and spatio-temporal resolutions. Handling the quality and privacy aspects of health data collection requires

further study. Scientists must learn and be able to publish their results together with the code and the data in order to allow other scientists to validate and build on their results. It is essential to bring scientists and practitioners together across disciplines in order to solve the complex issues around decision support systems in the health and environment domain. The rapid growth of data and proliferation of models calls for advanced geo-spatial processing capabilities in a true "Sensor and Model Web". Health related topics will be taken up by the newly founded OGC Health Domain Working Group.

## References


1. Klopfer (Ed.), M.: EO2HEAVEN mitigating environmental health risks. EO2HEAVEN Consortium (2013) Available from Fraunhofer IOSB, Karlsruhe, Germany and www.eo2heaven.org.
2. van der Schaaf, H., Rood, E., Watson, K.: A mobile application for reporting disease incidents. In: Proceedinds of ISESS 2013, IFIP Advances in Information and Communication Technology. (2013)
3. Jirka, S., Brauner, J., Bröring, A., Kunz, S., Simonis, I., Watson, K.: Application of Sensor Web Technology for Analysing Correlations between Health and Environmental Data. Light up the Ideas of Environmental Informatics (2012) 103–110 http://publica.fraunhofer.de/eprints/urn:nbn:de:0011-n-2281663.pdf.
4. Watson, V., Watson, K.: Design of a software framework based on geospatial standards to facilitate environmental modelling workflows. In Seppelt, R., Voinov, A., Lange, S., Bankamp, D., eds.: International Environmental Modelling and Software Society, Leipzig, Germany, iEMSs (July 2012) 1216–1223 http://www.iemss.org/society/index.php/iemss-2012-proceedings. ISBN: 978-88-9035-742-8.
5. Simonis, I., van der Merwe, M., Vahed, A.: Remote sensing and health data fusion: Methodological challenges in cholera research . In: Proceedings of the 34th International Symposium on Remote Sensing of Environment, Sydney, Australia. (2011) http://www.isprs.org/proceedings/2011/isrse-34/211104015Final00208.pdf.
6. Richter, S., Wiemann, S., Karrasch, P., Kadner, D., Brauner, J., Siegert, J., Rossmann, J., Elsner, B., Arloth, J.: Analysing Health Risks from Air Pollution Effects in Saxony, Germany. Light up the Ideas of Environmental Informatics (2012) 69–75
7. Wiemann, S., Richter, S., Karrasch, P., Brauner, J., Pech, K., Bernard, L.: Classification-driven air pollution mapping as for environment and health analysis. In Seppelt, R., Voinov, A., Lange, S., Bankamp, D., eds.: International Environmental Modelling and Software Society, Leipzig, Germany, iEMSs (July 2012) 467–474 http://www.iemss.org/society/index.php/iemss-2012-proceedings. ISBN: 978-88-9035-742-8.
8. Brauner, J.: Ad-hoc Geoprocessing in Spatial Data Infrastructures - Formalizations for Geooperators. In: Proceedings 1st AGILE PhD School. (2012)
9. Müller, M., Bernard, L., Kadner, D.: Moving code - Sharing geo-processing logic on the Web. ISPRS Journal of Photogrammetry and Remote Sensing (2013) http://dx.doi.org/10.1016/j.isprsjprs.2013.02.011.