EVOLUTION OF EARTH OBSERVATION ONLINE DATA ACCESS SERVICES

Torsten Heinen[1], Bernhard Buckl[1], Simone Giannecchini[2], Stephan Kiemle[1], Stephan Meissl[3]

(1) German Aerospace Center DLR, German Remote Sensing Data Center, Oberpfaffenhofen, Germany
(2) GeoSolutions S.A.S., Camaiore, Italy
(3) EOX IT Services GmbH, Vienna, Austria

ABSTRACT

With the Heterogeneous Missions Accessibility (HMA) initiative, the OGC standard "Web Coverage Service (WCS) Earth Observation Application Profile" has been developed to harmonize online access to very large primary Earth Observation data holdings. Although its use in web mapping servers has proven valuable capabilities, this standard is not yet widely adopted. Its acceptance for data download by end users is hampered by the lack of interpretation guidelines and its complexity requiring considerable server and client implementation efforts.

In this context, the project "Evolution of EO Online Data Access Services" funded by the European Space Agency (ESA) and presented in this paper analyses relevant scenarios and technologies for data publication and access, identifies potential for improvements of standards and their implementations, prototypes and evaluates selected improvements and proposes standard extensions for future releases. We hope hereby to considerably improve the acceptance of online EO data access services and standards and to promote their evolution and diffusion.

Index Terms— Web Coverage Service, Earth Observation, Online Data Access

1. INTRODUCTION

The challenges for Earth Observation (EO) data providers have increased considerably over the last decade. On the one hand, particularly Payload Data Ground Segments (PDGS) face ever increasing data rates of new sensor generations while maintaining and operating large archives of historical, heterogeneous sensor families. On the other hand, demand and requirements for instant access to large volumes of full resolution EO datasets have increased dramatically with rising processing capabilities of EO users such as Copernicus Downstream Services and value adding service providers in general [1].

In this context, the WCS Earth Observation Application Profile (EO-WCS) has been developed within HMA as a complementary interface to traditional file based download interfaces such as HTTP and FTP. Among other features, EO-WCS allows for efficient access to EO datasets by enabling download of spatiotemporal subsets and band constraints of full datasets and thus save bandwidth and storage requirements on the client.

The ESA technology project "Evolution of EO **O**nline **D**ata **A**ccess **S**ervices" (EVO-ODAS) conducted by DLR, EOX, and GeoSolutions specifically addresses current challenges and limitations of this interface in a broader context by analysing, prototyping, and evaluating different scenarios with stakeholders involved in current real world projects. The EVO-ODAS project has been started in October 2015 with duration of 18 months.

2. OBJECTIVES & METHODOLOGY

The aim of EVO-ODAS is to foster the evolution and usage of existing online data access standards by

- analysing relevant scenarios of data dissemination and related technological state-of-the-art
- identifying potential for improvements in these scenarios, through consideration (and possibly evolution) of the standards and their implementations
- demonstrational prototyping and evaluation of selected improvements and
- proposing standards extensions for future releases.

To meet these objectives the project is split in two phases as depicted in Figure 1. This allows for early and regular feedback from Standards Organization (OGC Technical Committee, Working Groups and other relevant programs) as well as Stakeholders contributing real world challenges in their respective field of work (see chapter 3).

The project workflow is following these main activities.

- Definition of online data access scenarios in dialog with identified project stakeholders using specific user stories and generic use cases
- Derivation of system requirements for prototype design and evaluation
- Identification of evolution potentials, issues, and limitations in existing standards and tools
- Setup of infrastructure for prototyping including state-of-the-art technologies for demonstration and validation

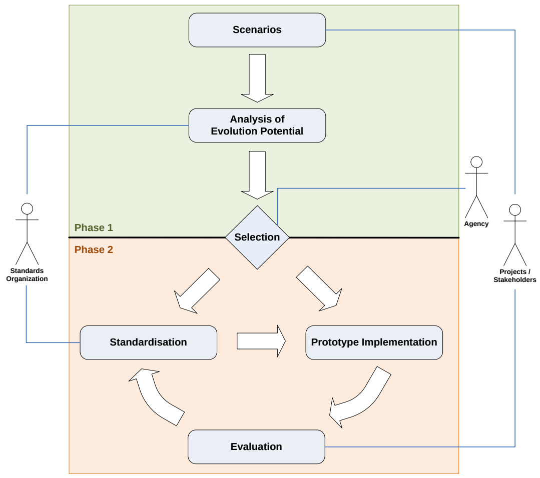

Figure 1: Project workflow and external actor interaction

- Selection of scenarios to be focused on in phase 2 of the project based on the following criteria
    - relevance for one or more use cases,
    - expected benefit in terms of functional improvement, usability, scalability, and performance
    - feasibility for prototype implementation and validation.
- Selection of evolution potentials and issues to be addressed in the prototype
- In parallel solutions to the selected issues are discussed and submitted to OGC
- Design and implementation of prototype including client and server components of different origins
- Prototype validation based on selected scenarios jointly with Projects and Stakeholders

3. STAKEHOLDERS

The project team has identified stakeholder projects serving as source of scenarios and requirements as well as taking profit of the outcome of the EVO-ODAS project and thus allowing a sound evaluation. All the stakeholder projects handle large amounts and long timeseries of EO data. They often have specific requirements concerning the efficient ingestion of data, the discovery of data and access services, the visualization, server-side processing and download of tailored coverages, and the support for federated authentication and content-aware access control and accounting of the end users' activity. A short introduction to the stakeholder's background is given below.

TIMELINE: The DLR project TIMELINE [2] aims at the generation of a 30-year time series of global change relevant variables like surface temperature for Europe. These time series will be used for change and trend detection and a better description and parameterization of land surface processes and their interaction with atmospheric processes.

UKIS: The Environmental and Crisis Information System project (the German abbreviation is UKIS) aims at setting up a framework of modularised and generalized components to build the next-generation information systems. They are integrating access to low-level EO archives and processing chains, extracting and fusing EO-derived information, and visualising air quality forecast simulation results.

EOC-WIS Visualisation System: The work of the Science Communication and Visualization Unit in DLR's Earth Observation Center (EOC-WIS) is primarily intended for non-experts, the media, and scientists active in fields other than remote sensing. Complex science topics are made accessible in a manner appropriate to these target groups and new, intuitively comprehensible forms of presentation are developed.

OPUS: The Scope of OPUS is to implement and demonstrate reusable ODA services integrating large amounts of near real time and offline earth observation data for further usage in collaborative infrastructure activities as planned for Copernicus services [3].

D-SDA: In the context of EVO-ODAS the German Satellite Data Archive (D-SDA) represents a Long Term Archive for EO products. The D-SDA is the entity preserving EO data of many ongoing and historic missions for the long term [4].

4. USE CASES & SCENARIOS

Five use cases describing generic system functions and interactions have been generalized from the stakeholder's user stories and have been aggregated into five end-to-end scenarios.

4.1. Use Cases

EO Data Ingestion: EO data emerge in data centres either in continuous (near real-time) processing streams or in bulk re-processing campaigns. The registration of new datasets and metadata in online data access servers currently follows system-specific methods. A harmonized backend system interface with comprehensive data management functions as well as best practices for its use and integration into EO payload data ground segments are highly desirable.

EO Data Discovery: An efficient data discovery for end users requires metadata to be linked bidirectional to related metadata, dataset series, and services to allow quick evaluation of the datasets different representations by users.

EO Data Access: This use case describes how clients and tools can visualize, download, and perform certain analysis on data, from single datasets up to very large timeseries. Inefficient access to large coverage descriptions, missing asynchronous download requests and basic transformations, and the missing linking of coverage descriptions with other services are examples of already known deficiencies.

Server Side EO Data Processing: Handling big EO data efficiently requires on-the-fly computations as well as the execution of long-running processes. Improvements of standards and implementations include the processing of datasets not available online, the support of asynchronous processing, and the capability to integrate scientific tools and infrastructures for remote processing.

Access Control and Accounting: EO data centres need to account and report on users' access activities for the verification of assured service levels and the identification of potentials performance issues. Access policies need to support the enforcement of region-dependent restrictions. Users should be able to use single-sign-on for overarching services.

4.2. Scenarios

Rolling Data Repository: This scenario demonstrates an online repository that allows discovery of all available product but keeps just a subset available for download (e.g. last 3 month) to cope with limited storage space. Already evicted products can be downloaded by reloading the products on demand.

Timeseries Animation Renderer: In this scenario large spatiotemporal datasets ingested in ODA services are used to feed distributed rendering farms that produce high quality 3D movie presentations.

On-Demand Coverage Analysis: This scenario represents processing of coverages already ingested in ODA services with user specified analytical methods such as spatiotemporal average computations by generating results as new coverages, plots, or videos.

Particle Forecast (and Alerting) Service: This scenario demonstrates server side processing of particle dispersion forecast, like volcanic or fire emissions (Figure 2) based on real-time EO datasets. The 4D (3D + time) dispersion model is published in ODA services and intersected with user specified n-D spaces such as administrative boundaries or flight paths to trigger an alerting service.

Figure 2: Illustration of fire hotspots and their particle dispersion, Source: DLR (CC-BY 3.0)

Interactive Timeseries Browser: This scenario covers the interactive visualization and downloading of large timeseries published in ODA services through web-based tools such as the EVO-ODAS demonstrator client (Figure 3).

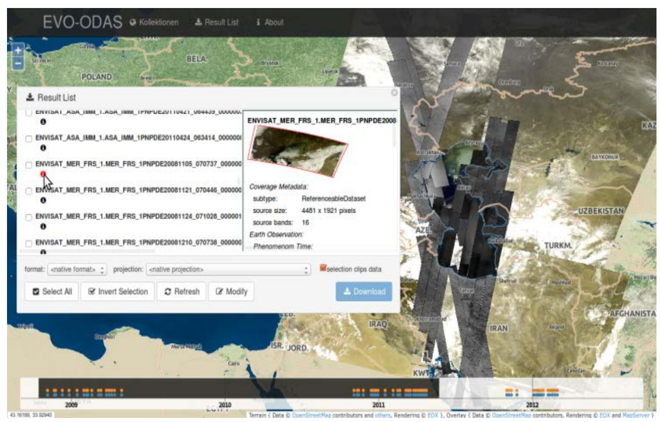

Figure 3: Web-based tool to inspect and browse large EO dataset timeseries. Source: EOX

5. STANDARDIZATION

The adoption of OGC standards for EO online data access services helps to get interoperable access to heterogeneous data holdings without requiring different proprietary tools. Thus, EVO-ODAS carefully analyses existing OGC specifications and documents being relevant for the scenarios identified above in order to find issues, limitations, and evolution potentials to be addressed.

An identified limitation is in ***cross service interaction*** since it is often necessary to link from one service to another, holding different views on the same data. For example, a user might first be presented with an RGB view on some data via WMS before downloading the actual data via WCS. Another limitation concerns the ***grouping of associated data***. The current coverage definition allows exactly one domainSet per coverage which prohibits, for example, the modelling of an EO satellite scene holding bands with different spatial resolutions, like in Sentinel 2 or Landsat, as a single coverage. Additionally, a typical EO satellite scene is associated with a number of auxiliary data like masks. ***General coverage groupings*** are an identified evolution potential. EO-WCS adds homogeneous and heterogeneous coverage groupings for 2D coverages to WCS. A general coverage grouping mechanism would overcome the 2D limitation and the binding to EO data.

Enhancements and evolutions will be discussed and brought to OGC as necessary. In addition, a prototype ODAS will be set up to demonstrate the feasibility of improvements for the selected limitations and evolutions.

6. IMPLEMENTATION & EVALUATION

For the second phase of the project a subset of the scenarios and requirements will be selected to be addressed. The goal is to implement, deploy, and operate for a consistent period of time a prototype online data access system that demonstrates the results of the standardisation and implementation activities.

The implementation will be mainly performed leveraging on the Open Source tools GeoServer, MapStore, EOxServer, and EOxClient for which the consortium's partners bring together a broad range of complementary skills and real world expertise.

DLR currently operates an Online Data Access Service [5] which delivers atmosphere products containing a complete archive for a variety of EO missions covering a wider range of geophysical parameters (e.g. L2 MetOp A/B GOME-2 from 2007 to 2016). OGC WCS, WMS, WFS and WMTS services are available as well as test instances for WCS-EO and WMS-EO. We foresee that the EVO-ODAS prototype will be deployed on the infrastructure managed by DLR to work in addition to the current infrastructure, sharing some of the input datasets in order to reduce the impact on the storage infrastructure while still accounting for a realistic testing and benchmarking scenario.

The prototype implementation will be validated using large representative EO datasets provided by DLR as part of the selected stakeholders' scenarios. The validation will focus mainly on:

Functionalities: extensions and updates to the standards' interfaces as well as its server side implementations will be verified through the official OGC CITE Team Engine where applicable.

Standards Compliance: the server implementations will be evaluated through a dual setup in the EVO-ODAS prototype. All tests, performance benchmarks, and user acceptance tests will be run against both installations to ensure the syntactical and semantical interoperability of the developed standard extensions. Furthermore off-the-shelf Open Source software such as QGIS and GDAL will be used to validate interoperability.

Performance / Benchmarking: Automated and comparable benchmark tests will be conducted on the prototype to detect, compare, and optimize performance gaps in both server implementations.

With support from ESA, the project partners intend to further accompany the evolution of EO online data access services in the standardization process as well as in the open source implementations.

7. ACKNOWLEDGEMENT

This research is conducted under a programme of and funded by the European Space Agency. The view expressed herein can in no way be taken to reflect the official opinion of the European Space Agency.

8. REFERENCES

[1] Heterogeneous Missions Accessibility (HMA) ESA TM-21, April 2012 ISBN 978-92-9221-883-6, http://esamultimedia.esa.int/multimedia/publications/TM-21/TM-21.pdf

[2] Frey, Corinne and Dietz, Andreas and Bachmann, Martin and Bernhard, Eva-Maria and Ruppert, Thomas and Molch, Katrin and Wolfmüller, Meinhard and Künzer, Claudia and Müller, Andreas and Dech, Stefan (2014) TIMELINE - Processing of AVHRR time series over Europe and North Africa. 34th EARSeL Symposium, 16.-20. Jun. 2014, Warschau. http://elib.dlr.de/92337/

[3] Reck, Christoph and Campuzano, Gina and Dengler, Klaus and Heinen, Torsten (2015) High-Performance Online Data Access for Collaborative Infrastructures. PV 2015 Conference, 3-5 November 2015, Darmstadt, Germany. http://elib.dlr.de/100193/

[4] S. Kiemle, K. Molch, S. Schropp, N. Weiland, E. Mikusch, Big Data Management in Earth Observation – The German Satellite Data Archive at DLR, Big Data From Space BiDS'14, Frascati, ESA/ESRIN, (2014) http://elib.dlr.de/92405/

[5] Heinen, Torsten and Kiemle, Stephan and Buckl, Bernhard and Mikusch, Eberhard and Loyola, Diego (2009) The Geospatial Service Infrastructure for DLR's National Remote Sensing Data Library. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, Volume (4), pp. 260-269. IEEE Geoscience and Remote Sensing Society. http://elib.dlr.de/60172/