STANDARDIZED DATA ACCESS SERVICES FOR GOME-2/METOP ATMOSPHERIC TRACE GAS PRODUCTS

T. Heinen [(1)], B. Buckl [(1)], T. Erbertseder [(1)], S. Kiemle [(1)], D. Loyola [(2)]

(1) German Aerospace Center (DLR), German Remote Sensing Data Center (DFD)
Oberpfaffenhofen, D-82234 Wessling, Germany
(2) German Aerospace Center (DLR), Remote Sensing Technology Institute (IMF)
Oberpfaffenhofen, D-82234 Wessling, Germany

Abstract

Established for land use applications, standardized data access services such as the Web Map Service (WMS) and the Web Coverage Service (WCS) are widely used to provide interoperable access to any kind of earth observation data and any other kind of geo-referenced information. These existing service standards come along with multiple data sources and software tools, making the interpretation and combination of different data layers for various purposes at global and regional scale very easy. However, the use of these standards in the atmospheric data domain is still very restricted. This paper presents DLR's initiative of developing standardized data access services for atmospheric trace gas data products continuously processed from GOME-2/MetOp data.

1. INTRODUCTION

Within the distributed EUMETSAT Polar System (EPS) ground segment, DLR acts as an integral part of the Ozone and Atmospheric Chemistry Monitoring Satellite Application Facility (O3M-SAF) for the processing, archiving and delivery of GOME-2 atmospheric products. Based on the continuous reception of GOME-2 level 1b data different processing chains for generation and dissemination of near-real-time and off-line data products are triggered. The resulting level 2 datasets provide the total and tropospheric column density of different trace gases such as O3, NO2, BrO, SO2, OClO and HCHO, as well as cloud properties (Pruin & Kiemle, 2004). Aggregated (level 3) and assimilated (level 4) data products are generated in the context of the World Data Center for Remote Sensing of the Atmosphere (WDC/RSAT)[1] (Erbertseder et al, 2006). The GOME-2 Data Processor Version 4.2 (GDP 4.2) algorithm is used to generate these value added products and is described in Valks et al. 2008.

Recognizing the importance of interoperability and harmonized access to geospatial information, the German Remote Sensing Data Center (DFD) introduces standardized service-oriented data access to GOME-2 atmospheric products within DLR's multi-mission earth observation ground segment facility. These new user services extend the already established data services provided by DLR with EOWEB®[2] and the WDC/RSAT in order to enhance interoperability between different international initiatives such as 'Heterogeneous Mission Accessibility' (HMA-ESA) and 'Global Monitoring for Environment and Security' (GMES/KOPERNIKUS-EU). Furthermore the aim is to foster collaboration and support Spatial Data Infrastructures (SDI) on different levels.

Especially for GOME-2-based air quality products the requirements for individual and customized applications are increasing. Users ask for defining and displaying their own area of interest from regional to urban scales and analysing the data by combining it with information on population density, land use, administrative boundaries and traffic density to name a few. These requirements are crucial when e.g. excess levels of air pollutants are quantified and statistically analysed.

This paper shows how these requirements can be addressed using service-oriented data access and presents the extension of DLR's existing multi-mission web gateway EOWEB® with OGC Web Map Service (WMS) and Web Coverage Service (WCS) compliant interfaces. In this way the GOME-2 total column trace gas products are easily accessible not only for weather forecasting and climate applications, but they can be seamlessly integrated into existing processing systems, GIS environments widely used in applications like aviation control and air quality monitoring as well as popular virtual globe software such as NASA World Wind and Google Earth.

2. USER SERVICES AT DLR RELATED TO GOME2 TRACE GAS PRODUCTS

GOME-2 level 2 trace gas products are processed at the DLR in the O3M-SAF infrastructure based on input data of the EUMETSAT EPS ground segment. Immediately after receiving L1b input data, higher level data products are generated within a value-adding processing chain integrated into WDC-RSAT. In addition to the processing chain, the DLR operates different user services offering access to multi-level GOME-2 data products. The services mainly differ in the way to discover, evaluate and access the data and are tailored for its target users. Processors and user services are based on DLR's multi-mission Data Information and Management System (DIMS, Kiemle 2002). Figure 1 gives an overview of the existing services.

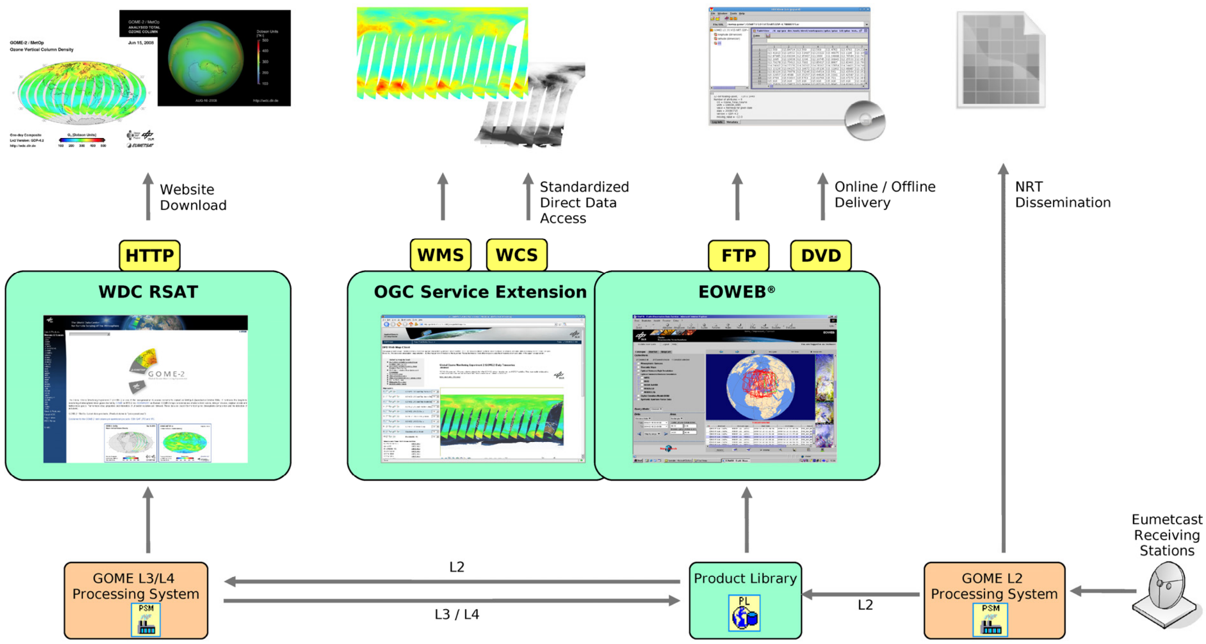

Figure 1: **Overview of GOME-2 related user services at DLR**

On the lower side of the processing chain is the dissemination of the near-real-time level 2 data products. Available 2 hours after sensing, these products are broadcasted via EUMETCast and uploaded to FTP servers of subscribed users such as the SAF network partners. Although such subscription service may provide means to filter some data products, it does not provide any user-oriented functionality such as filtering/selecting, subsetting/mosaicking or visualisation of the data products. Instead it is assumed that the subscribed user maintains specialized software for these purposes.

The next higher level of user service is driven by the catalog and ordering use case. An example of this user service is the central UMARF catalog provided by EUMETSAT or the DLR EOWEB® gateway. Both services provide a web interface implemented with Java applet technology that allows for interactive discovery of data products through a catalog interface. Evaluation of specific datasets is supported through metadata browsing and viewing of quicklooks. Access of data products can be either online through FTP or offline through ordering of products stored on CD/DVD. As with the subscription service, processing, visualisation and analysis of the data products in the user's context usually require specialized software.

On the other side the WDC-RSAT is an example of a web-based user service that generates products from value-added L3/L4 data products. Although L3/L4 data products can be downloaded from the website, the services focus on producing end-products such as daily composite maps, chemical analyses by data assimilation and animations/movies of those datasets. Products can be interactively discovered, evaluated and accessed through the website. End products can then easily be consumed by users with standard software and tools but are not suitable for further processing or overlaying with the user's own datasets. Neither does it allow for displaying a specific subset or a mosaic of single data products.

While the method to access the data products differs, the user usually gets the data in the native HDF5 or BUFR format (Loyola et al., 2008). As indicated before, this makes it difficult for non-experts to use the data in other systems and standard off-the-shelf software. In order to facilitate the evaluation of and access to GOME-2 related data products, EOWEB® introduces standardized interfaces to access selected datasets. This approach is outlined in the next chapter.

3. SERVICE-ORIENTED DATA ACCESS TO GOME-2 PRODUCTS

This chapter gives an overview of service-oriented data access that is about to be incorporated into DLR's EOWEB® user service by comparing it to the traditional catalog ordering use case. The comparison focuses on the actual access and not the discovery of data products. Also the interfaces based on the WMS and WCS specifications are explained briefly. Finally, the chapter shows how GOME-2 datasets will be made accessible through these services.

Comparing service-oriented data access and traditional catalog ordering

Unlike catalog ordering, the service-oriented approach offers another layer of abstraction between the data provider and consumer (see Figure 2). This additional layer not only abstracts from specific files and data stores but also allows for value-adding service functions offered by the data provider.

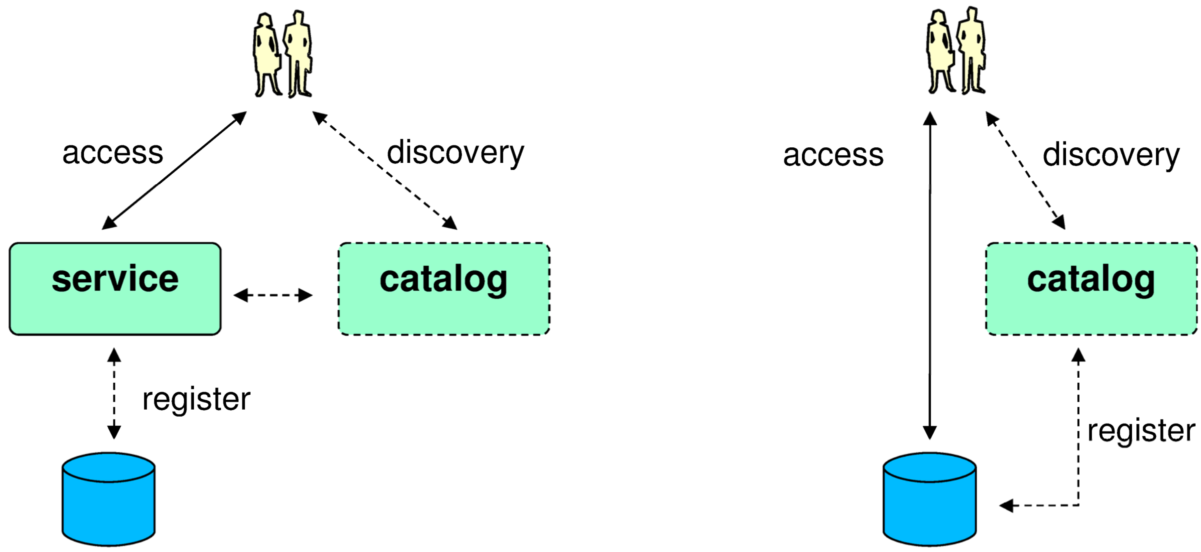

Figure 2: **Comparison between service-oriented data access (left) and traditional catalog ordering access (right)**

In case of the catalog ordering, the user discovers and evaluates data of interest in the catalog. The data is then ordered and possibly processed by the system. Online access to the data is usually conducted asynchronously through FTP. This powerful approach allows complex, time-consuming processing and tailoring of data before it is delivered to the user. However the interaction between users and the systems differs and is therefore not interoperable, forcing the user to interactively access different portals for each data provider if these are not integrated in a common infrastructure. Furthermore exploitation of data, such as visualisation on top of user datasets is often impeded by using file formats not appropriate for common off-the-shelf software. This might not be a problem for scientists and experts who have access to special resources and facilities, but poses a major task for non-experts such as private and public as well as governmental users.

In contrast, service-oriented data access is synchronous and can offer basic filtering, processing and visualisation functions that facilitate rapid access to data products. These functions and the service in general are executed through a well-defined request-response pattern. The user passes parameters with the request and can expect to receive the desired response accordingly. Accessing data through a service layer instead of direct access to a set of files also enhances flexibility in data management for the data provider. Datasets can e.g. be renamed, moved and processed without forcing the user to change his access behaviour or software. However, a service-oriented approach alone still suffers from lack of interoperability. The user would still need special software for a specific service. Therefore standardization of data access policies is an important part to homogenize data access across different data provider.

OGC Web Services

The mission of the Open Geospatial Consortium (OGC)[3] is to develop in consensus with international industry, government and academic members publicly available interface specifications (called OpenGIS® Specifications) that support interoperable solutions for exchanging geospatial information. The following paragraphs give a brief overview of two specifications, namely the Web Map Service (WMS, Beaujardiere 2006) and Web Coverage Service (WCS, Whiteside & Evans 2006). It should be noted that the following description does not explain the service interface and concepts in detail, please refer to the specification for more information.

While the WMS interface enables the access to portrayed geospatial information (maps or „pictures of geospatial data“), the WCS interface allows the retrieval of geospatial datasets with original semantics representing space-varying phenomena (Whiteside & Evans 2006). The latter is mostly, but not solely, multi-dimensional raster data such as satellite data, model calculations or digital elevation models. One of the most popular output formats of the WCS for such datasets is GeoTIFF, but others can be supported by the service provider. Each pixel of the service response contains a value representing a geophysical phenomenon (e.g. a value of 320.5 for total column ozone concentration given in Dobson units). Other values may exist for the same pixel, e.g. the estimated error for the ozone concentration or another trace gas concentration, resulting in additional data layers in the WCS. In contrast, the WMS delivers image formats such as JPEG or PNG, which stores one RGB value (e.g. [0,0,255]=blue) per pixel. The colors have been applied by a WMS styling rule translating physical values into color values.

The WMS and WCS specification defines the specific service operations and parameters. Those operations are invoked through HTTP GET and POST requests. One service operation common to all OpenGIS® specifications is „GetCapabilities“. The „GetCapabilities“ operation delivers a schema-based XML document that describes the supported service operations and available datasets

The most important WMS operation is „GetMap“. This operation is refined with spatial, temporal and other constraints. For example the BBOX parameter allows a user to specify a spatial subset; the SRS parameter allows the definition of the resulting map projection and the TIME parameter controls the point (or period) in time. Other parameters control the styling rule and properties such as size, format and transparency of the output image. A simple WMS request (version 1.1.1) might look like this:

http://server.com/mywms?SERVICE=**WMS**&REQUEST=**GetMap**
&LAYERS=**ozone&**STYLE=**o3_colortable**
&SRS=**EPSG:4326**&BBOX=**-25,34,35,72**
&TIME=**2008-09-25**
&FORMAT=image/png&WIDTH=600&HEIGHT=380

The WCS operation „GetCoverage“ is similar but allows retrieval of the actual data not a portrayed image. It supports the same spatial and temporal parameters to constrain the resulting dataset. In addition it allows a user to specify parameters that are defined for the specific dataset. For example a parameter OZONE=100/300 parameter would only output values between 100 and 300 Dobson units; a BAND=1,2 parameter select only the first two bands of the dataset; the parameters RESX and RESY specify the ground resolution (in units of the coordinate reference system) of the output dataset. The following WCS example request (Version 1.0.0) illustrates this:

http://server.com/mywcs?SERVICE=**WCS**&REQUEST=**GetCoverage**
&COVERAGE=**ozone**
&CRS=**EPSG:4326**&BBOX=**-25,34,35,72**
&TIME=**2008-09-25**&OZONE=**100/300**&BAND=**2**
&FORMAT=GeoTIFF&RESX=**30**&RESY=**30**

GOME-2 WMS and WCS access

Value-added GOME-2/MetOp data products processed for the WDC-RSAT processor have been chosen to be the first candidates for a pilot project on the emerging service-oriented data access. Although the spatial resolution is relatively low compared to land-use data sets (1440x720 pixels, i.e. approx. 0.25° per pixel), the data products offer different species at a global scale and multiple temporal resolutions.

The resulting products consist of different species, temporal aggregations, processing levels as well as the spatial and temporal dimension. See Figure 3 for two examples. Currently the following products are available:

- Species: Ozone (O3), Sulfur Dioxid (SO2), Nitrogen Dioxid (NO2), Tropospheric Nitrogen Nioxid (NO2Tropo), Cloud Fraction (CFR), Cloud Top Albedo (CTA), Cloud Top Height (CTH), Cloud Top Pressure (CTP).
- Aggregation (temporal): Latest, Daily, Monthly, Forecast
- Processing levels: L3, L4

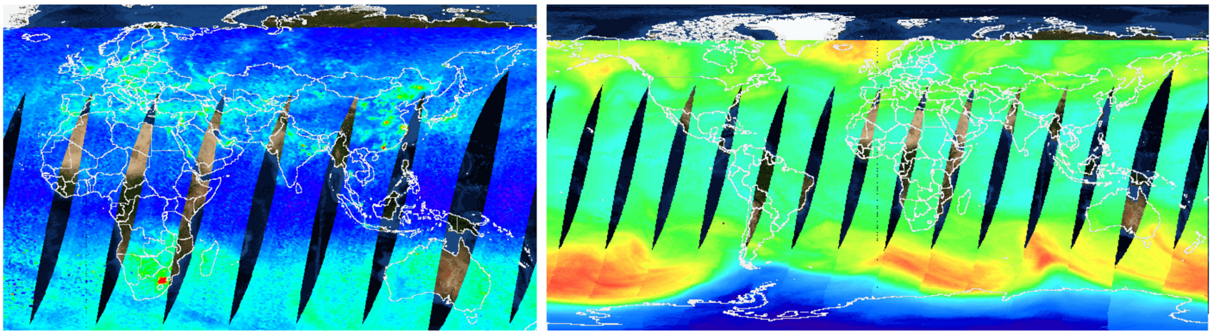

Figure 3: **Examples of GOME-2 WMS layers combined with background and country boundaries; left: daily aggregated nitrogen dioxid ("gome2.no2.l3.daily") and right: daily aggregated total ozone ("gome2.o3.l3.daily")**

While access to the spatial and temporal dimensions are controlled by the respective BBOX and TIME parameters of the WMS and WCS, the other dimensions have been made selectable through different layers. L3 and L4 GOME-2 datasets are identified by a layer name that is composed from the following convention: gome2.{species}.{level}.{aggregation}. E.g. "gome2.o3.l4.monthly" denotes the layer representing monthly assimilated ozone values. This pragmatic approach results from the fact that the L3/L4 processor produces one file for each point in time, processing level, temporal aggregation and specie. Furthermore new species are introduced from time to time, so this approach does not affect the existing data layers. The disadvantage is obviously a quite high number of layers.

A preferable approach from a data management point of view would have been to represent the different species in one file with multiple bands. This would have reduced the number of files and published data layers. The WMS STYLE (e.g. STYLE=ozone) or WCS range subset parameter (e.g. BAND=so2) could have then been used to select the specific species within one single layer (say "gome2.l3.daily").

4. CONSUMING GOME-2 DATA PRODUCTS THROUGH SERVICES

After introducing the service-oriented access to GOME-2 data products, this chapter points out how the service-oriented approach could affect the user and provider of datasets. Exemplary scenarios also show how applications and services can be used to access WMS and WCS servers and consume GOME-2 datasets.

It is obvious that new data provider interfaces require new client software to enable data access. The difference to proprietary systems is that in case of the WMS and WCS interfaces those software tools not only already exist but support is steadily growing. The spectrum ranges from geographic information system (GIS), web-based online mapping tools over 3-dimensional virtual globes to low-level software libraries available as commercial off-the-shelf (COTS) and free and open source software (FOSS).

Depending on the need and scenario, users can choose to utilize and integrate existing software into their systems. The following paragraphs describe simple use cases for specific software categories. All share the ability to request geospatial datasets from WMS and sometimes also WCS servers (Figure 4).

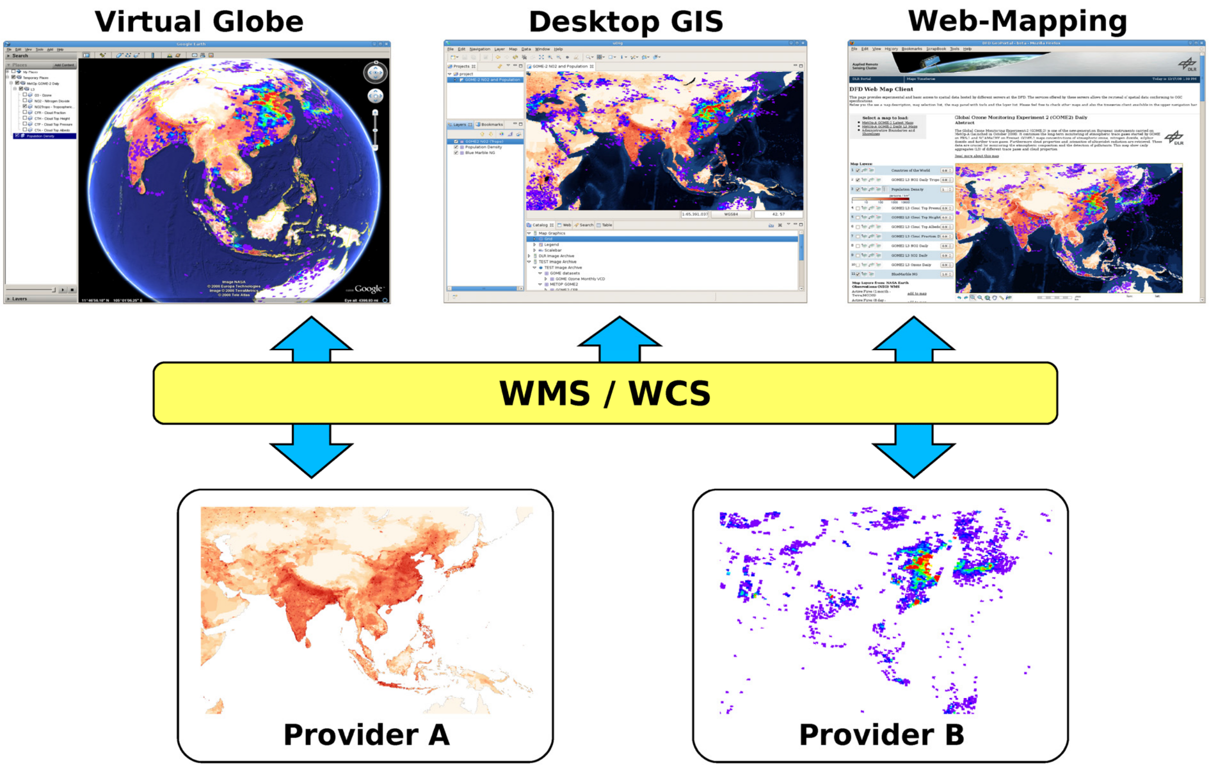

Figure 4: **Interoperable access to different data providers through WMS-compliant off-the-shelf software; from left to right: Google Earth, uDig and Community Mapbuilder; datasets taken from NASA Earth Observations WMS (population density) and DLR EOWEB® WMS (GOME-2 tropospheric NO2)**

Online mapping or web-mapping tools allow the generation and publishing of maps through internet-based technologies within a web browser, e.g. Community Mapbuilder[4], OpenLayers[5] and the lightweight WMS JavaScript Library (wms-map)[6] are FOSS examples of such tools. These tools often allow user interaction known from GIS such as zooming, panning and layering of datasets from WMS servers. Access to WCS servers however is often not supported, due to the restricted environment of a web browser. Integration of interactive maps is useful for portals and web pages showing specific thematic contexts, e.g. combining different datasets to visualize the temporal tropospheric NO_2 evolution in urban areas under negative weather condition (smog). Thus web-mapping is very well suited for rapid and easy data evaluation and visualization of geospatial information, especially for non-expert users.

GIS applications provide a broader range of analysis and processing functions then web-mapping tools. While ESRI's ArcGIS[7] and ERDAS Imagine[8] are popular COTS examples, FOSS GIS such as uDig[9], QGIS[10], OpenJUMP[11] and gvSIG[12] are powerful alternatives. WMS access is virtually always supported; however the retrieval of data through WCS for further processing makes it very valuable for solving specific problems, e.g. the observation of chemical plume surrounding the globe after volcanic eruptions. Furthermore the generation of buffers around high concentrations of SO2 in combination with current flight paths could be used to disseminate warning to nearby airplanes or calculate alternative routes.

Virtual globe applications such as NASA World Wind[13] and Google Earth[14] applications support WMS server as a data source and thus allow overlaying and combining datasets from different data providers similar to web mapping solutions. The high degree of popularity, attractive user interface and a broad range of available datasets makes it good platform for publication and presentation of datasets to a large number of users.

Nevertheless custom applications are needed to implement specific use cases or server side processing system. Different low level libraries such a GDAL[15] and GeoTools[16] facilitate building custom applications on top of WMS/WCS services. Those applications can benefit from server side processing functions, making manual aggregation, subsetting or selection of specific values obsolete. Furthermore applications that are compliant to those interfaces can be reused by different data providers and are not tightly coupled to a specific service implementation and data format.

5. CONCLUSION AND OUTLOOK

This paper has shown how GOME-2 datasets will be made accessible by extending EOWEB® with OGC-compliant interfaces. The DLR is committed to establish service-oriented data access to its existing multi-mission infrastructure in order to promote interoperability between initiatives, communities and projects. This includes support of and contribution to national and international Spatial Data Infrastructures (SDI). Providing WMS and WCS interfaces for earth observation datasets furthermore enables users to choose from a broad range of third-party software.

The services are currently being tested and are expected to be publicly available beginning of 2009 with aggregated and assimilated GOME-2 data products. Development and extension of these services will continue in conjunction with provision of additional datasets such as GOME-2 level 2, GOME-1/ERS and 10 years of value added NOAA AVHRR data for Europe. These datasets will be published in EOWEB and a web mapping portal. Medium term goals also include user management and access control in order to meet the requirements of a new German "satellite data security regulation" (Satellitendatensicherheitsverordnung, SatDSiV) and potential commercial datasets.

REFERENCES

Beaujardiere, J., (2006) OpenGIS® Web Map Server Implementation Specification, Version 1.3, OGC 06-042. http://portal.opengeospatial.org/files/?artifact_id=14416

Erbertseder, T., Baier, F., Meyer-Arnek, J., Elbern, H., Schwinger, J., Loyola, D., (2006) Assimilation of METOP/GOME-2 and IASI chemical observations. Proceedings of the METOP RAO Workshop, ESA/ESRIN, Frascati, 15-17 May 2006

Kiemle, S., (2002) From Digital Archive to Digital Library – a Middleware for Earth-Observation Data Management. 6th European Conference on Research and Advanced Technology for Digital Libraries, ECDL 2002 Proceedings, Lecture Notes in Computer Science, Springer Verlag, Berlin Heidelberg.

Loyola, D., Zimmer, W., Kiemle, S., Valks, P. , Ruppert, T., (2008) Product User Manual for GOME Total Column Products of Ozone, Minor Trace Gases, and Cloud Properties, DLR/GOME/PUM/01, Iss./Rev. 1/E. http://wdc.dlr.de/sensors/gome2/DLR_GOME_PUM_1E.pdf

Pruin, B., Kiemle, S., (2004) Processing and Product Management for GOME-2 Atmospheric Trace Gas Monitoring. Proceedings of the 2004 EUMETSAT Meteorological Satellite Conference, EUMETSAT Publ. EUM P 41, 2004.

Valks P., Loyola, D., Hao, N., Rix, M., (2008) Algorithm Theoretical Basis Document for GOME-2 Total Column Products of Ozone, Minor Trace Gases, and Cloud Properties (GDP 4.2 for O3M-SAF OTO and NTO), DLR/GOME-2/ATBD/01, Iss./Rev. 1/C. http://wdc.dlr.de/sensors/gome2/DLR_GOME-2_ATBD_1C.pdf

Whiteside, A., Evans, J.D., (2008) Web Coverage Service (WCS) Implementation Standard, Version 1.1.2, OGC 07-067r5. http://portal.opengeospatial.org/files/?artifact_id=27297

ENDNOTES

1. http://wdc.dlr.de
2. http://www.eoweb.de
3. http://www.opengeospatial.org
4. http://communitymapbuilder.osgeo.org/
5. http://openlayers.org/
6. http://wms-map.sourceforge.net/
7. http://www.esri.com/products/index.html
8. http://www.erdas.com
9. http://udig.refractions.net/
10. http://www.qgis.org/
11. http://openjump.org/
12. http://www.gvsig.gva.es/
13. http://worldwind.arc.nasa.gov/
14. http://earth.google.com/
15. http://www.gdal.org/
16. http://geotools.codehaus.org/