OGC Geodatacube Standard Working Group Charter

To: OGC members & interested parties

A new OGC Standards Working Group is being formed. The OGC members listed below have proposed the OGC GeoDataCube SWG. The SWG proposal provided in this document meets the requirements of the OGC Technical Committee (TC) Policies and Procedures.

The SWG name, statement of purpose, scope, list of deliverables, audience, and language specified in the proposal will constitute the SWG's official charter. Technical discussions may occur no sooner than the SWG's first meeting.

This SWG will operate under the OGC IPR Policy. The eligibility requirements for becoming a participant in the SWG at the first meeting (see details below) are that:

- You must be an employee of an OGC member organization or an individual member of OGC;
- The OGC member must have signed the OGC Membership agreement;
- You must notify the SWG chair of your intent to participate to the first meeting. Members may do so by logging onto the OGC Portal and navigating to the Observer page and clicking on the link for the SWG they wish to join and;
- You must attend meetings of the SWG. The first meeting of this SWG is at the time and date fixed below. Attendance may be by teleconference.

Of course, participants also may join the SWG at any time. The OGC and the SWG welcomes all interested parties.

Non-OGC members who wish to participate may contact us about joining the OGC. In addition, the public may access some of the resources maintained for each SWG: the SWG public description, the SWG Charter, Change Requests, and public comments, which will be linked from the SWG's page.

Please feel free to forward this announcement to any other appropriate lists. The OGC is an open standards organization; we encourage your feedback.

Chapter 1. Purpose of the Standards Working Group

In a very agile approach, this Standards Working Group (SWG) has the goal to facilitate the handling of different GeoDataCubes (GDCs). For this purpose, an Application Programming Interface (API) will be defined that serves the core functionalities of GDCs. GDC users will be able to handle different GDCs according to the same principles, as interoperability between GDCs will be achieved.

The GDC SWG will define a metadata model including provenance and data lineage information to describe all details about a GDC. Furthermore, the SWG will identify formats to be used for data exchange. If existing formats do not meet the requirements, the SWG will extend its work to the development of a GDC exchange format.

The agile process will go through multiple iterations in order to produce standards of growing complexity, starting with the minimum viable solution that is then extended as required. The SWG will first define the basic characteristics of GDCs as well as core requirements for the use of GDCs and the exchange of GDC data. These requirements need to be met by the GDC API, metadata model, and exchange formats.

In the next step, the SWG will decide which existing standards or standards under development should be used or extended to achieve these core requirements. Afterwards, the GDC API, the GDC metadata model, and the data exchange format(s) will be defined and implemented. The goal is to develop a solution that supports a majority of typical GDC application scenarios. The simplicity of the solution will be prioritized over its ability to meet all possible use cases and be compatible with all available standards.

Chapter 2. Business value proposition

Over the past decade, GDCs have been developed independently, resulting in a lack of interoperability between them. By improving interoperability, the vendor community will be able to proceed with specific GDC variants, while at the same time the consumer community will be able to interact much more effectively with different instances. With the increase in available data products served as GDCs, it is becoming increasingly important to understand exactly what a GDC entails and how it was created. This trend is also evident in other OGC working groups. The increasing demand for standardized Analysis Ready Data (ARD) formats has led to the creation of the OGC ARD SWG. The ARD SWG also deals with the topics of data provenance and lineage. By simultaneously standardizing the interaction mechanisms with GDCs, numerous synergy effects can be exploited.

The creation of a GDC API, metadata model, and exchange format will enhance the interoperability between existing solutions, simplify the interaction with different data cubes, and facilitate the integration of data from multiple datacube sources.

The GDC SWG will enable interoperability and scalable data access across different GDCs. By following a user-centric approach, it will develop solutions that meet the needs of various types of users, including:

- Scientists (basic use of API) – user-friendly
- App/client developers (expert use of API) – easy to learn, interoperable, functionally rich, efficient
- GDC developers (API implementers) – easy to implement and integrate, efficient

Chapter 3. Scope of work

The GDC SWG will develop standards to interact with various GDCs. For this purpose, the following steps will be performed:

- Identification of real-life use cases with industrial relevance
- Definition of the GDC API (which may be a profile of (an) existing OGC API(s) or a new development)
- Definition of exchange format recommendations, profiles, or new developments
- Definition of the GDC metadata model (in particular information about how the GDC was built; similar to ARD concepts and vision around data provenance and lineage)
- The GDC API shall support accessing and processing at minimum
- Analysis of the usability of existing standards

The development shall start with a domain-independent core API that will simplify the integration of data provided by different communities. With the focus on simplicity rather than expressivity for the core API, the work of the SWG will allow for a stepwise integration of existing technologies across domains and communities. API extensions will further allow high levels of interoperability, address reproducibility of complex GDC workflows and toolchains, and thus enable new levels of interoperability and cost efficiency when dealing with data stored in and served as GDCs.

The GDC SWG's approach will be to create a new standard from existing OGC Building Blocks and related geospatial initiatives/developments with minimal extension. This approach is intended for the creation of new standards within the OGC, but has so far only been applied in rudimentary form. The GDC SWG will thus break new ground.

3.1. Statement of relationship of planned work to the current OGC Standards baseline

GDCs are - broadly defined - multidimensional arrays with one or more spatial or temporal dimension(s). Multidimensional arrays have been addressed in the context of OGC Coverages with corresponding services, APIs, and information models. The SWG may analyze in particular the following existing standards:

- OGC API Environmental Data Retrieval (EDR)
- OGC Coverage Implementation Schema (CIS)
- OGC API Processes Part 2 and 3
- OGC API Coverages
- ZARR, including proposed extensions (e.g., tiles)
- GeoTIFF
- HDF5
- OGC Coverage Implementation Schema, in particular: CIS 1.1 General Grid Coverage which models n-D geo datacubes

The above list is not exhaustive and new standards may be considered if they are relevant to the goals of the SWG.

3.2. What is out of scope?

- This group will not define specific implementations of processes or products. Instead, it will generate a general mechanism to interact with GDCs.

- The SWG will not define the storage format of GDCs.

- The SWG will not define the coding language for GDCs.

- The SWG will not define implementation details or put constraints on storage solutions.

The scope will be further refined in future discussions with the ARD SWG.

3.3. Specific existing work used as starting point

The GDC SWG will follow an agile methodology with the goal to create a first core standard within the first year. Subsequent iterations may add additional functionality. The GDC SWG will start with a use case collection and analysis phase that further informs the selection of additional starting points or other work to be considered. The targeted use cases shall reflect real world scenarios, though should allow for a rapid implementation of the GDC standards without adding unnecessary complexity.

In the next step, existing state-of-the-art standards are examined with regard to their usability for GDCs. In parallel, the GDC SWG will review two Engineering Reports resulting from OGC Innovation Program initiatives as well as specifications developed by the openEO initiative:

- The OGC Testbed-17: Geo Data Cube API Engineering Report (21-027) is, like all Innovation Program Engineering Reports, available from the OGC public engineering reports webpage. The OGC Testbed-17 Engineering Report (ER) documents the results and recommendations of the Geo Data Cube API task in 2021. The Testbed-17 Call for Participation provides additional insights into requirements and use cases. The ER defines a draft specification for an interoperable GDC API leveraging OGC API building blocks, details implementation of the draft API, and explores various aspects including data retrieval and discovery, cloud computing and Machine Learning. Implementations of the draft GDC API have been demonstrated with use cases including the integration of terrestrial and marine elevation data and forestry information for Canadian wetlands.

- OGC Testbed-16: Data Access and Processing Engineering Report (20-016) summarizes the results of the 2020 Testbed-16 Data Access and Processing task. The task had the primary goal to develop methods and apparatus to simplify access to, processing of, and exchange of environmental and Earth Observation (EO) data from an end-user perspective.

- The openEO initiative develops an open API to connect R, Python, JavaScript and other clients to big Earth observation cloud back-ends in a simple and unified way

In addition to these documents and specifications, the SWG will review current developments, e.g., the Microsoft Planetary Computer API or the Google Earth Engine API.

It is emphasized that the GDC SWG will start with a use case analysis that further informs the selection of additional starting points or other work to be considered.

3.4. Is this a persistent SWG

[x] YES

[ ] NO

3.5. When can the SWG be inactivated

The GeoDataCube SWG can be inactivated when the Standards managed by the SWG require no further update or maintenance.

Chapter 4. Description of deliverables

The deliverables of this SWG activity include a new GDC API, the GDC metadata model, and recommendations for GDC exchange formats.

The GDC SWG applies an agile development methodology. The first iteration that will only include the core functionality will be made available for consideration at the September 2023 OGC Member Meeting, with a goal of approving them in late 2023/early 2024.

Subsequent iterations are planned thereafter.

4.1. Initial deliverables

GeoDataCube API standard (OGC API-Geodatacubes) and GDC metadata model.

4.2. Additional SWG tasks

To be completed as the SWG continues its operation.

Chapter 5. IPR Policy for this SWG

[x] RAND-Royalty Free

[ ] RAND for fee

Chapter 6. Anticipated audience / participants

- Main industrial worldwide players

- Institutions and international organizations

- User communities for multi-dimensional data

Chapter 7. Domain Working Group endorsement

This draft charter will be presented to Earth Observation Exploitation Platform Domain Working Group with a request for endorsement.

Chapter 8. Other informative information about the work of this SWG

8.1. Collaboration

The SWG intends to use public GitHub repositories as the primary means for its work.

8.2. Similar or applicable standards work (OGC and elsewhere)

The following standards and projects may be relevant to the SWG's planned work, although none currently provide the functionality anticipated by this committee's deliverables:

- OGC DAPA: combining access with limited processing

- openEO: scalable datacube process API

- STAC

- Cloud Optimized GeoTIFF (COG)

- OGC API - Processes Part 3: Workflows and Chaining

- OGC Analysis Ready Data SWG (proposal)

- OGC Discrete Global Grid Systems (DGGS) work

- OGC Abstract Specification Topic 21 - DGGS, as it contains a very valuable introduction to 'common spatio-temporal classes to support temporal and spatio-temporal geometry, topology, zones, zonal identifiers, zonal query, and RSs using temporal or zonal identifiers'

- OGC Observations, Measurements & Samples (OMS) SWG results

- Storage format specifics that influence the design of the API

The SWG intends to seek and if possible maintain liaison with each of the organizations maintaining the above works.

8.3. Details of first meeting

The first physical meeting of the SWG will be held during the February 2023 OGC Member Meeting. Call-in information will be provided to the SWG's e-mail list and on the OGC portal calendar in advance of the meeting.

8.4. Projected on-going meeting schedule

The work of the SWG will be primarily via email and conference calls, preferably every two weeks, with face-to-face meetings ideally taking place at each OGC member meeting.

8.5. Supporters of this Charter

The following people support this proposal and are committed to the Charter and projected meeting schedule. These members are known as SWG Founding or Charter members. The charter members agree to the SoW and IPR terms as defined in this charter. The charter members have voting rights beginning the day the SWG is officially formed. Charter Members are shown on the public SWG page. Extend the table as necessary.

Name	Organization
Claudio Iacopino	European Space Agency (ESA)
Stephan Meißl	EOX IT Services GmbH
Miruna Stoicescu	European Organisation for the Exploitation of Meteorological Satellites (EUMETSAT)
Jonas Eberle	German Aerospace Center (DLR)
Ryan Ahola	Natural Resources Canada (NRCan)
Pedro Gonçalves	Terradue Srl
Jérôme Jacovella-St-Louis	Ecere Corporation
Panagiotis (Peter) A. Vretanos	MariaDB (CubeWerx Inc.)
Peter Baumann	Constructor University and rasdaman GmbH
Alexander Jacob	Eurac Research
Paul Haesler	Geoscience Australia and Open Data Cube (ODC)
Matthias Mohr	WWU Münster
Alexey Shiklomanov	NASA Goddard Space Flight Center (GSFC)
Brian Terry	Analytical Mechanics Associates (CEOS-SEO)
Roberto Alacevich	European Space Agency (ESA)
Samantha Lavender	Pixalytics Ltd
Jeroen Dries	VITO
Peter Strobl	European Commission, DG JRC

8.6. Conveners

- Claudio Iacopino, European Space Agency (ESA)
- Miruna Stoicescu, European Organisation for the Exploitation of Meteorological Satellites (EUMETSAT)
- Ingo Simonis, OGC

Chapter 9. References

- Draft charter of the Analysis Ready Data (ARD) SWG

- Testbed-17 Call for Participation section "DataCube"

- OGC Testbed-17: Geo Data Cube API Engineering Report (21-027)

- OGC Testbed-16: Data Access and Processing Engineering Report (20-016)

- Data Access and Processing API (DAPA) for Geospatial Data

- openEO API v1.1.0