OGC Geospatial Coverages Data Cube Community Practice

Marie-Francoise Voidrot\(^{(1)}\) and George Percivall\(^{(2)}\)

(1) Open Geospatial Consortium Europe, Toulouse, France; Email: mvoidrot@ogc.org

(2) Open Geospatial Consortium, Washington, USA; Email: gpercivall@ogc.org

Abstract. Data Cubes for geospatial information provide the means to integrate observations and other types of geospatial data for use in multiple applications through simplified access and efficient analytics. Acknowledging that diverse implementations already exist, this Community Best Practice, developed by the OGC members and supported by at least 6 independent implementations listed in the document, defines requirements for a core Geospatial Coverage Data Cube infrastructure and guidelines for enhancements and extensions to the basic core. With this paper and other activities, OGC is contributing to the advancement of Data Cubes more generally and as a foundation for federations of Data Cubes for geospatial information.

1. Objectives for Geospatial Data Cubes
The OGC Geospatial Coverages Data Cube Community Practice document codifies a set of community practices regarding geospatial data cubes in order to promote development, proliferation and federation of data cubes for the purpose of rendering geospatial data more readily available for analysis and decision support.

Key themes in this document are:
- Multidimensional spatial-temporal data cube structure based on the Coverages concepts with interoperable services for access and analytics.
- Geospatial Analytics: A Geospatial Data Cube is a digital infrastructure to manage geospatial data and provide access and analytics on the data.
- Federation of Data Cubes: Coordination of accessibility and interoperability at multiple, distributed data cubes provided by different organizations. Users should be able to access any data cube in the federation without knowing the location.

Data Cube related technologies are demonstrating to be an efficient technical mean to achieve the needed bridge from large amounts of observational (e.g. remote sensing) and other geospatial data to the wider geospatial domain. Data Cubes technology can support analytics through:
- Providing harmonized, aggregated access to large amounts of geospatial data by offering a logical view of single, multi-dimensional objects;
- Drilling through long time series of different satellite series more efficiently;
- Efficiently merging different information sources (space, airborne, in-situ, GIS)

Several existing solutions implement the concept with different approaches. The document identifies 13 requirements common to 6 independent implementations as a baseline for interoperability.

2. Content of the “OGC Geospatial Coverages Data Cube Community Practice” document
2.1. Data Cubes commonalities
After defining a Data Cube more generally, the document details why and how the OGC Coverage Implementation Schema (CIS), can be used as a reference Information model for Geospatial Coverages Data Cubes.

To “Provide harmonized, aggregated access to large amounts of geospatial data by offering a logical view of single, multi-dimensional objects;”, Geospatial Coverage Data Cubes require data preparation and structuring to transform raw data as made available by the production workflow into organized information enabling efficient search, analytics and visualization.
As “Drilling through long time series of different satellite series more efficiently” and “Efficiently merging different information sources” are other goals of Geospatial Data Cubes, data optimization has to be performed when ingesting the data into a Geospatial Data Cube Service in advance to any user query and access based on administrator specifications. Metadata are then critical to describe the data that are offered by a Data Cube to let the user understand the data, make it usable a queryable: data provider, geospatial and temporal resolution, parameter and unit, quality …. The significant value of Geospatial Data Cubes Services is the ability to perform analytics that meet user needs. Key to the value is to be user centric and that the user does not need to have detailed knowledge of the source data structure and observation artefacts. Thus, the data cube analytics will follow specific requirements in terms of data access methods and queries, unifying data management to trim and slice using the same syntax along all axes. Despite the data preparation made at the ingestion time before any user query, the Data Cubes may need to offer data transformation capacities on the fly at the retrieval time to serve a broader set of usages.

2.2. Interoperability: APIs and Encodings

Interoperability and scalable fusion of spatial information across different data cubes is crucial and highly dependent on the use of robust international standards governing the access and transfer protocols for communication between client and server as well as among different servers. [Strobl 2017]. For interoperability and federation, data cubes need to provide access and analytics using well known interfaces. To serve the largest communities possible, Application Programming Interfaces (APIs) defined by open consensus standards are needed. The OGC Web Service standards have been implemented by the examples listed in Annex C of the document. A Data Cube service interface for data access shall be compliant with OGC Web Coverage Service (OGC WCS) and support encodings that are defined in open consensus standard and that are used by the community targeted.

2.3. Interoperability for Analysis and Visualization

A Data Cube service interface shall support OGC WCS-P Processing extension and/or OGC Web Processing Service (WPS) to run an analysis tool and return the results to the user without having to download the entire dataset.

3. Further developments

Geospatial Data Cubes defined in this paper provides a basis for a network of interoperable Geospatial Data Cubes usable as a Federation of Data Cubes. Experiments of federations are ongoing as they are already needed. OGC pushes several threads of work on DGGS, Met Ocean Profiles to WCS, OGC APIs and STAC specification to better prepare future interoperability challenges. Analytics on arrays will be a topic of discussion in OGC in conjunction with other communities that will lead to increased analysis capabilities for Geospatial Coverage Data Cubes.

References

[1] [OGC 2019] The OGC Geospatial Coverages Data Cube Community Practice, OGC Document 18-095r4 , Currently under public review on https://www.opengeospatial.org/standards/requests/191
[2] [Strobl 2017] Strobl, Peter, et.al., The Six Faces of the Data Cube, Proc. of the 2017 conference on Big Data from Space (BiDS’17) Toulouse, France 28–30 November 2017 doi: 10.2760/383579.

Acknowledgments

Authors acknowledge the editors of the [OGC 2019] document for their efforts to reach consensus on a challenging topic: George Percivall (OGC), David Gavin (Digital Earth Australia), Peter Baumann (Jacobs University and rasdaman), Peter Strobl (EC-JRC), Syed R. Rizvi and Andrew Cherry (ama-inc.com), Simone Mantovani (MEEO), Grega Milcinski (Sinergise), Stephan Meissl (EOX) and Marie-Francoise Voidrot (OGC). Thank you also to all those who have help mature the goals and concepts via their fruitful discussions and reviews.