LOWERING
THE BARRIER TO ENTRY

THE OGC'S DATA ACCESS AND PROCESSING API WILL ENABLE SCIENTISTS AND GEOSPATIAL ANALYSTS TO ANALYSE DATA IN A WAY MOSTLY INDEPENDENT OF ITS LOCATION, SUPPORTING THE OPEN SCIENCE MOVEMENT. **INGO SIMONIS** EXPLAINS HOW

For the past few years, OGC has been modernising our standards to better align with web best-practices and end-user expectations, resulting in our growing OGC API family of standards. Part of this effort has also been to design our standards to better take advantage of cloud infrastructure, including being able to deploy and share spatial analysis workflows across different cloud providers.

This has beneficial ramifications for collaboration, transparency, and accessibility of scientific workflows – which are all cornerstones of the "Open Science" movement. I previously touched on these Earth observation processing packages in our column in the July/August 2020 issue of GeoConnexion International, which discussed our "applications-to-the-data" architecture and the Application Deployment and Execution Service (ADES) APIs.

Another part of the effort to simplify Earth observation data processing and analysis workflows is the development of the OGC Data Access and Processing API (DAPA). Developed as a draft specification during OGC Testbed-16 in 2020 and now being tested in real-world scenarios during this year's initiatives, DAPA is a so-called "convenience API" that enables scientists and other geospatial analysts to run several operations on Earth observation or other data using a single API call, in turn providing the data in a form directly ready for further analysis.

This differentiates DAPA from existing APIs, such as OGC API Features or OGC API Coverages. Where the latter two are data centric APIs – they focus on data access and subsetting – DAPA is a user-centric API that includes data access with processing. As such, it takes lots of processing burden away from the user.

DAPA does this in a way that is mostly independent of the data's location – meaning that the same call can access data stored in a local file, an in-memory structure (such as an xarray) or remotely in the cloud. Ultimately, this means that an end-user can initiate the process on one archive for one set of data, then just change the URL to have the same process(es) run on an entirely different dataset.

For example, with a single DAPA API call, you can say, "Please give me all the data you have, in a datacube, for this specific area and time window, with these fields and as a result of this map algebra." A datacube is then created on the fly and delivered to you in a form ready for you to work on in your software of choice. And, say you run that on a Landsat archive, you could then use the same API call to reproduce it on, say, a PeruSat-1 archive.

Open science

The reproducibility of the DAPA calls, just like the packagability of the ADES 'apps', makes them ideal for use in support of the Open Science movement. This movement aims to make scientific research and its dissemination more accessible, by professionals and amateurs alike, and to generate transparent and accessible knowledge that is shared and developed through collaborative networks.

The movement is receiving big support from NASA and ESA, which have also been instrumental in their sponsorship of the development of DAPA and ADES. Indeed, open standards in general, due to their enabling of the interoperability required for collaboration across organisations and disciplines, play a critical role in the open science movement, too.

The Open Science movement's desire to make scientific data and processes accessible by more than just scientists ties in well with OGC's recent efforts to design standards with a strong, end-user-centric, perspective, rather than the data-provider centric view that has come to dominate earlier standardisation work. This means simplifying and improving not just their form and function, but also their documentation.

User-friendly standards

Rather than reading standards documents – which are, by their nature, meticulously defined to reduce room for interpretation and therefore tedious to read – many developers favour an approach that starts with simple documentation and examples. From there, additional features are explored stepwise, with the actual standard document often being the last resource being consulted. As location tech grows outside of the traditional geospatial spheres of expertise, this user-centric view becomes critical if the benefits of widespread standards adoption are to be realised.

With this in mind, work undertaken in Testbed-17 aims to lower the barrier of entry to implementing and accessing DAPA and
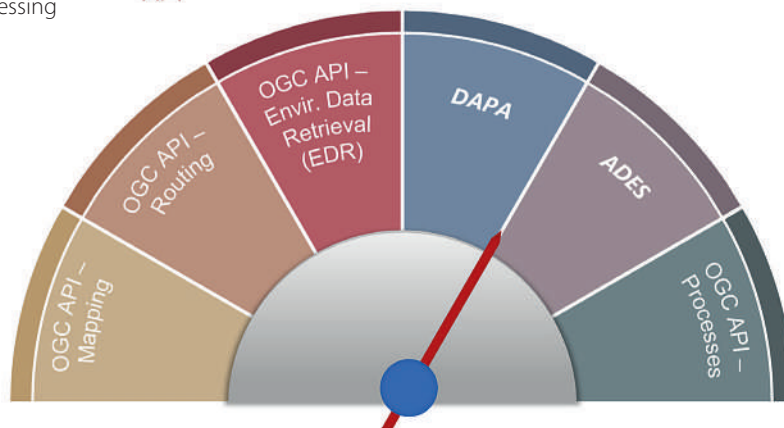
OGC is lowering the barrier of entry for developers new to standards and/or location data

OGC's data access and processing API (DAPA) enables selections of cloud-based data to be processed, packaged and delivered as a datacube

HOW CAN WE IMPROVE?

What do you think OGC can do to help make its standards easier to understand and implement? Let OGC and its incoming DevRel know at ogc.org/contact. More information on upcoming and published OGC APIs is available at ogcapi.org

other OGC APIs by creating sets of example code for both server- and client-side software, cloud deployment and installation scripts, and best practices. The testbed will deliver all knowledge necessary to develop, deploy and execute standards-based Web APIs to web developers, and follow a 'how to' philosophy with lots of hands-on experiments, examples and instructions. Better yet, by providing scripts that illustrate the deployment and operation of API instances on local machines as well as across different cloud environments, it will make the challenge of mapping software components to cloud infrastructure a smooth experience.

In addition to documentation and code examples meant to make users' lives easier, OGC will soon bring in a new staff member for developer relations, who will provide an interface between OGC and the developer community: What do developers need from OGC? Where do they struggle? What materials can we provide to help?

All these activities have come from OGC's ambitions to make our standards easier to understand and implement, and to feel more tangible than our earlier work. OGC stands behind its position that standards – through making data Findable, Accessible, Interoperable and Reusable – unlock tremendous value and power cross-collaborations that offer huge benefit to society. By making the standards themselves align with the FAIR principles, we are lowering the barriers to their adoption and spreading their value further.

Ingo Simonis is chief technology innovation officer of the OGC (www.ogc.org)

DAPA and ADES fit within a spectrum of different processing APIs available from OGC

Middle ground: APIs that encapsulate the set of functionalities required by the community by specifying an easy-to-learn and easy-to-use API

OGC API – Env. Data Retrieval (EDR)

OGC API – Routing

OGC API – Mapping

DAPA

ADES

OGC API – Processes

Specific APIs with limited functionality and well-defined behavior

Any process to be deployed on the Web and invoked in a standardized way with standardized job control for long-running processes and standardized interfaces for retrieving processing results