Federating CODE-DE and DIAS Datacubes: Successes and Lessons Learnt

Peter Baumann
Jacobs University

Datacubes today are accepted as a quantum leap towards user-centric services on spatio-temporal mass data; this goal is also pursued by Analysis-Ready Data, and both approaches largely overlap in that they lead to handy paradigms for space/time integration, homogenization of data, and independence from low-level properties such as the internal tiling or format encoding of Earth data.

In this contribution we report on two datacube activities where large-scale queryable datacubes have been established based on the open datacube standards of OGC. A DIAS file archive has been augmented with a datacube frontend as part of the European Open Science Cloud (EOSC) initiative. Independently, the batch-oriented Hadoop interface of the German Sentinel hub, CODE-DE, has been enhanced with an interactive datacube API. Then, the idea came up to connect both services. Today federated queries are possible where users just fire their query against any of the nodes, and these team up to transparently find the answer for the client. This location transparency frees clients from substantial work in particular when it comes to cross-data-center fusion. However, this added service quality harbours new challenges, one of them being access control across autonomous data centers and down to the level of single pixels.

In our talk we present the services, their federation, challenges faced and lessons learnt.