Towards Seamless Planetary-Scale Services

Data Integration: Enabling the Acceleration of Science Through Connectivity, Collaboration, and Convergent Science

EGU, 2020-may-06

Peter Baumann

Jacobs University | rasdaman GmbH

Motivation

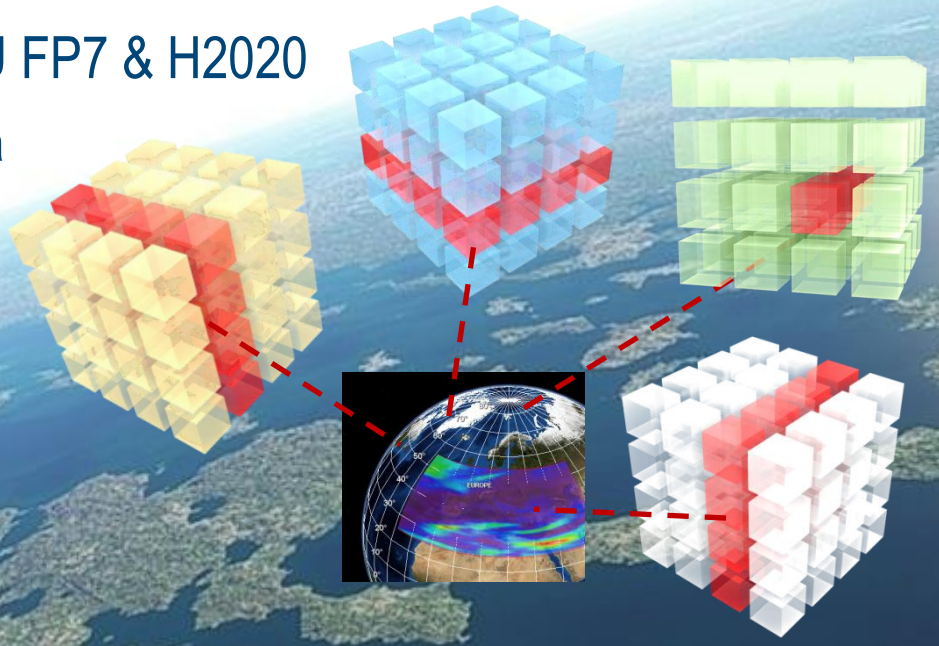
- FAIR (Findable, Accessible, Interoperable, Reusable) summarizes core requirements on data
- ...leaving obstacles:
  - FAIR is metadata-centric; how to transpose to data (such as pixels)?
  - FAIR should be easy, hiding technical details – how?
  - Analytics? Fusion of disparate, heterogeneous data?
  - Why should I want to find data? Just wanna use them!
    - *Data Broker is a task, not a solution*
- EarthServer = focus on data (with metadata), fusion ...location-transparent
  - Open data provider community, open standards, freedom in client choice

rasdaman
raster data manager

EarthServer

www.earthserver.xyz

- Agile, location-transparent analysis + fusion + visualization ready datacubes
- Open federation of large-scale data providers
  - DIASs, research institutes, agencies, universities, companies, ...
  - 20+ PB and growing: Sentinel SAR & hyperspectral, thematic, products, ...
  - open standards, community governance
- Intercontinental initiative, stated with EU FP7 & H2020
  - free of charge; no need to publish all data
  - Now accepting membership requests



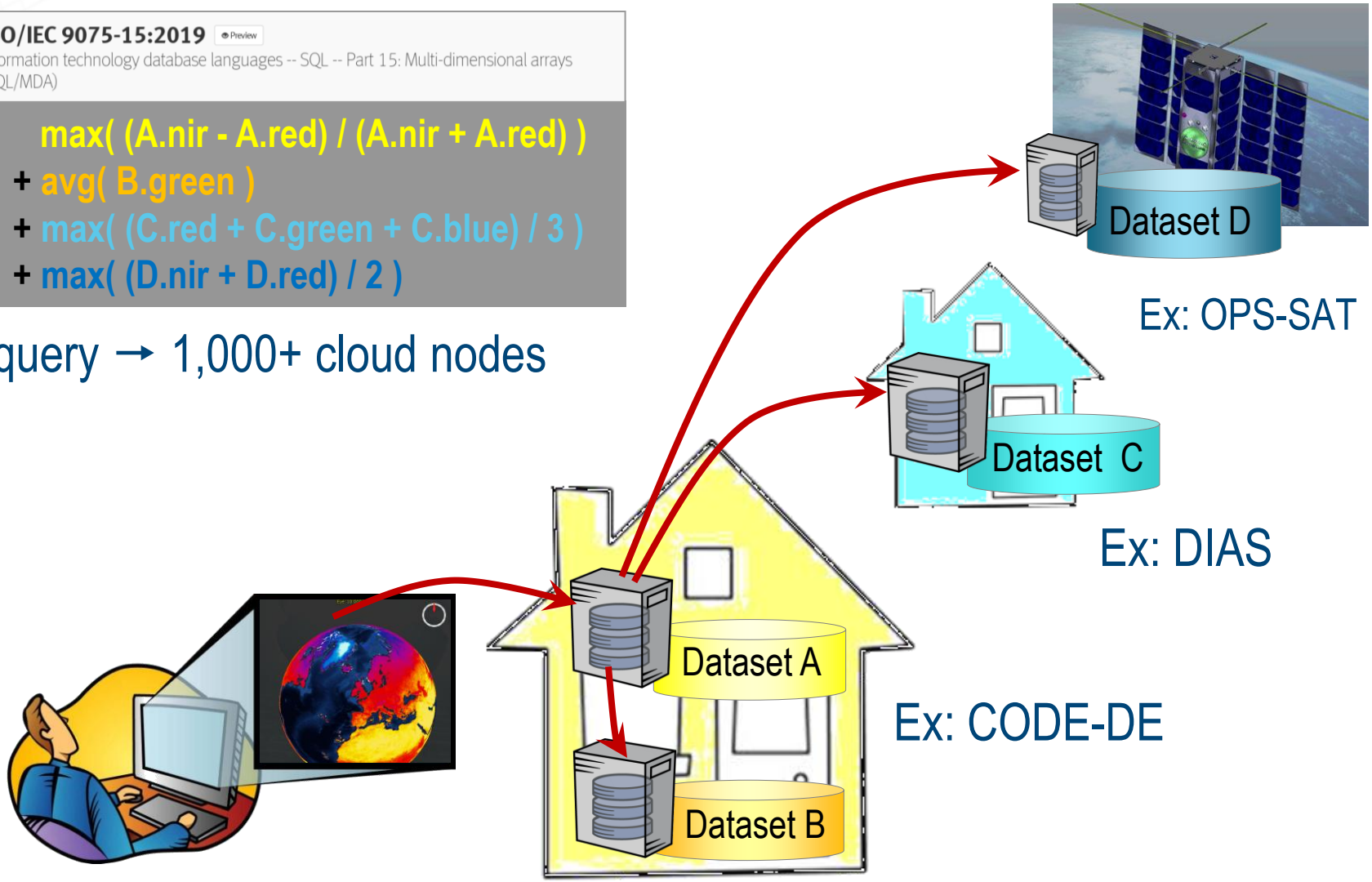
Reviewers & EC:
"proven evidence", will "*significantly transform [how to] access and use data*" …and *"with no doubt has been shaping the Big Earth Data landscape"*

Tech: Language-Based Federated Processing

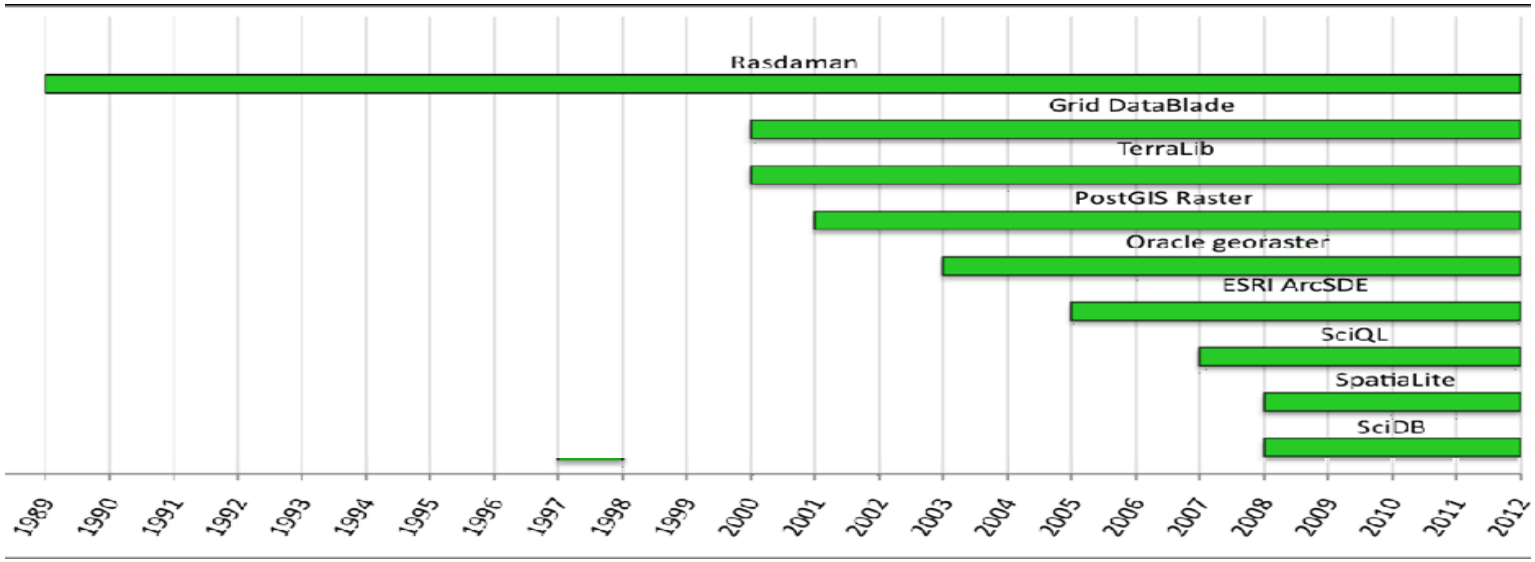
ISO/IEC 9075-15:2019 Preview
Information technology database languages -- SQL -- Part 15: Multi-dimensional arrays (SQL/MDA)

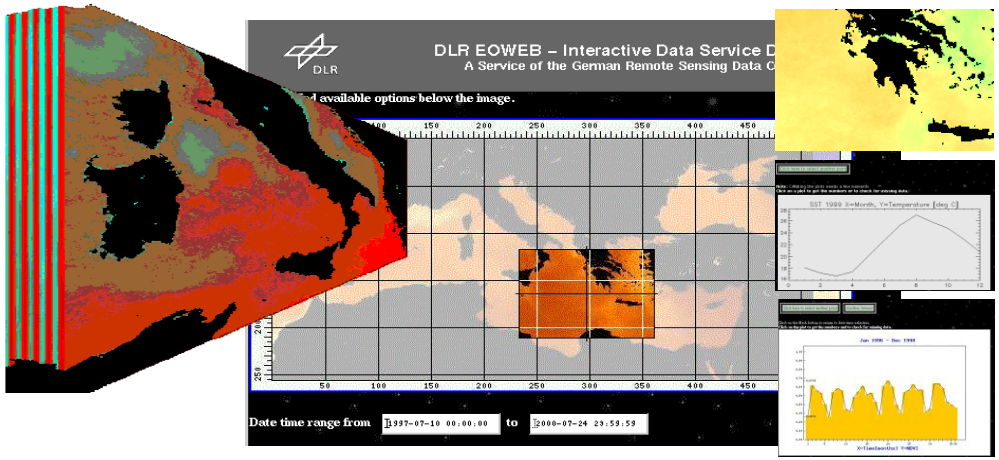
```
    max( (A.nir - A.red) / (A.nir + A.red) )
  + avg( B.green )
  + max( (C.red + C.green + C.blue) / 3 )
  + max( (D.nir + D.red) / 2 )
```

1 query → 1,000+ cloud nodes



Datacubes: Experience Background

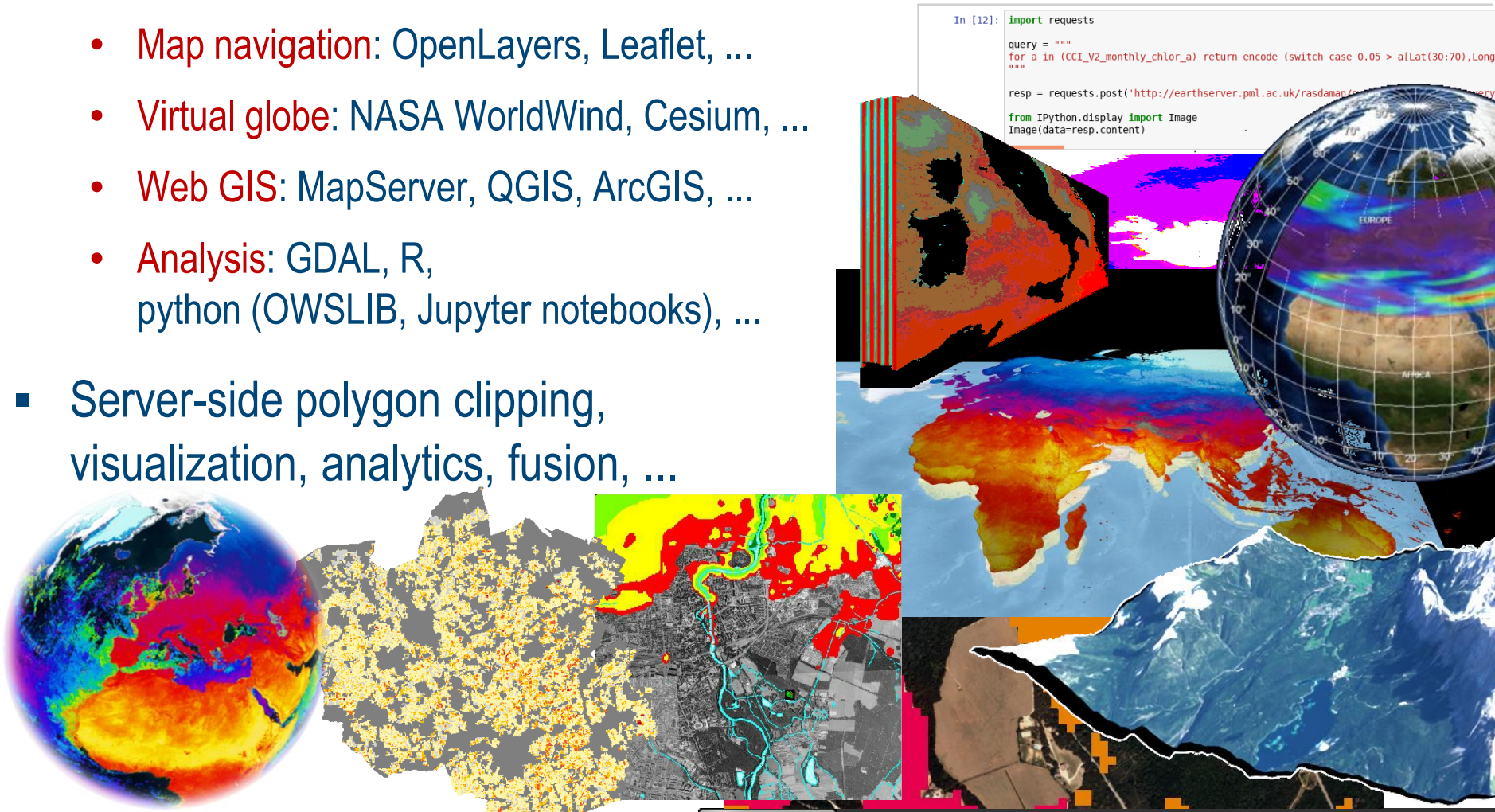




[Diedrich et al 2001]

Back to the User

- OGC W*S → users remain in comfort zone of well-known tools
  - Map navigation: OpenLayers, Leaflet, ...
  - Virtual globe: NASA WorldWind, Cesium, ...
  - Web GIS: MapServer, QGIS, ArcGIS, ...
  - Analysis: GDAL, R,
    python (OWSLIB, Jupyter notebooks), ...
- Server-side polygon clipping,
  visualization, analytics, fusion, ...



[rasdaman-based portals]