Federation:
The Next Key Innovation After ARD

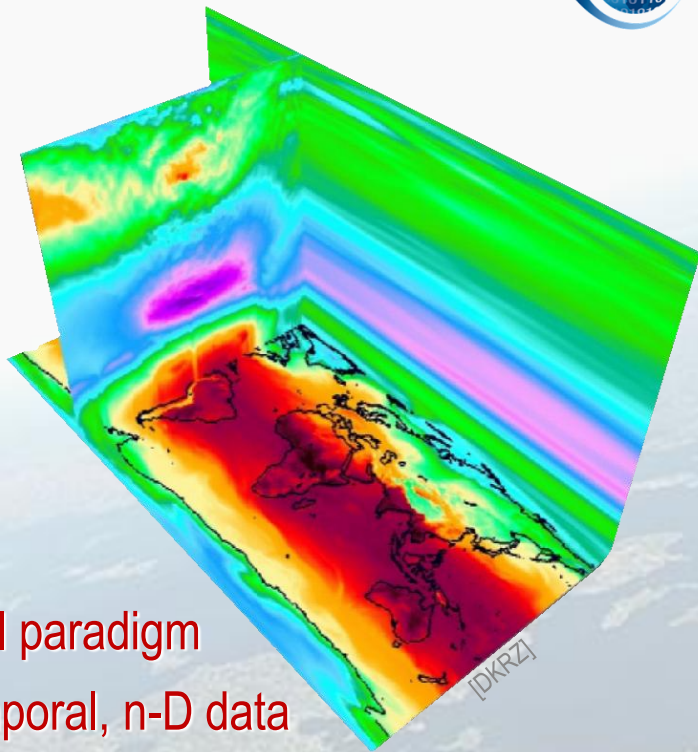EarthServer Side Event, ESA Phi-Week, 2021-10-14
Peter Baumann
Jacobs University | rasdaman GmbH

Datacubes?

- Sensors & simulations

  → gridded („raster") data

- Datacubes are the natural paradigm

  to interact with spatio-temporal, n-D data

- Pioneered by EarthServer since 2011

[DKRZ]

EarthServer Datacube Federation

- spatio-temporal analytics & fusion, dozens of PB

  - location-transparent

  - open standards, coding-free

  - Open for code & data

- Open, free, transparent, democratic

  - Open & private, free & commercial

  - quickly growing community

  - DIASs; NASA, EURAC to join

- Powered by rasdaman

PARSEC Big Data Toolbox

Helmh... Jacobs Univer... MundiDIAS Alfred Wegener Institut CODE-DE CreoDIAS ...hia Taiwan

EarthServer Datacube Federation
Understand. Space. Time.
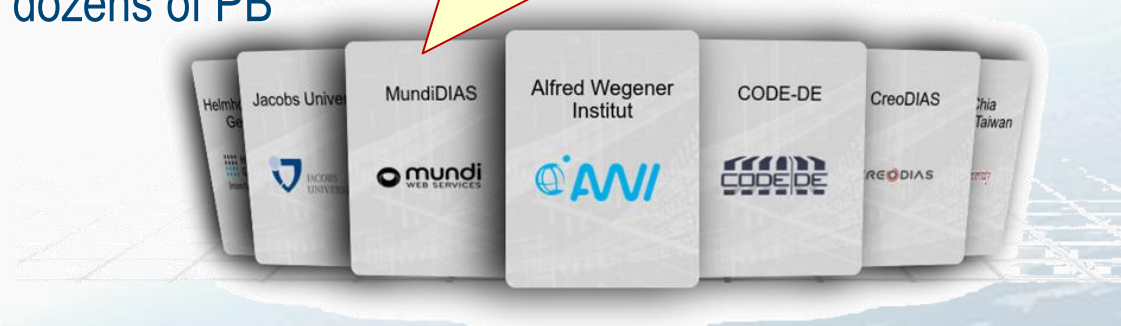
EarthServer Key Innovation Components

- **Strictly standards**
  - powerful datacube analytics with EU/INSPIRE WCPS

- **Federation**
  - single information space, 100% location transparent

- **Array DBMS power backend**
  - Flexibility, scalability, security
  - performance – 300x, 400x reported

- **Zero-coding**
  - Democratization of data insight

- **Clients, clients, clients**
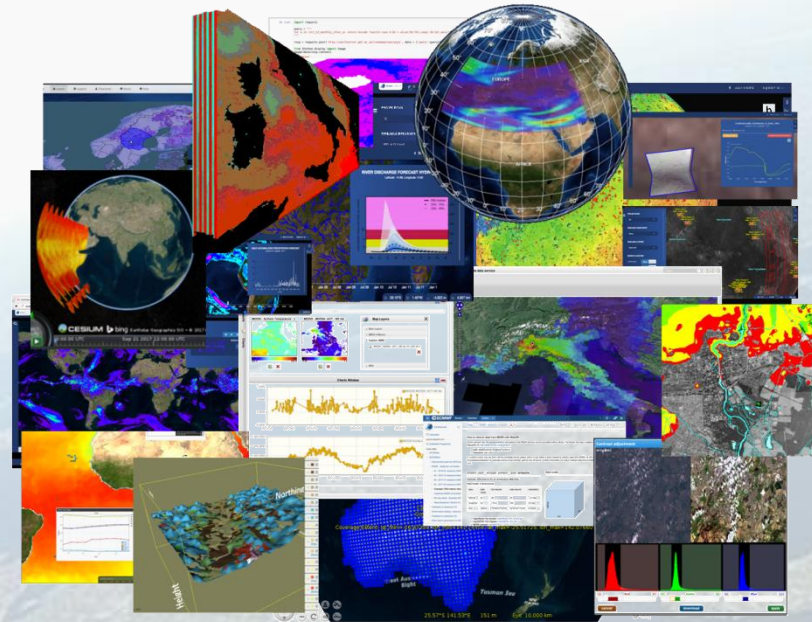
rasdaman

„raster data manager" - pioneered actionable datacubes: 160+ publications

- High-level datacube query language (now ISO SQL/MDA) + coverage QL (WCPS)

- massively scalable Datacube Management & Analytics engine, full-stack implementation

  - Dozens of PB; 1000x parallelization, planetary-scale federation

  - Intelligent ETL: automated datacube maintenance & optimization

- Reference implementation, multi-award winning
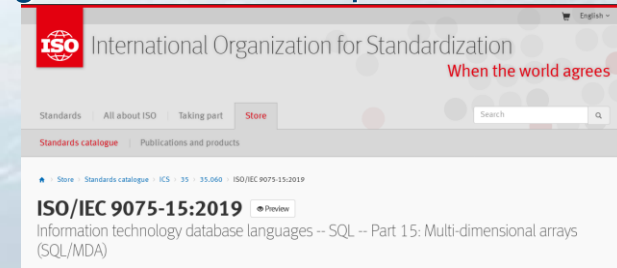
Optimized Distributed Processing

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

[ACM SIGMOD, VLDB]

BigPicture
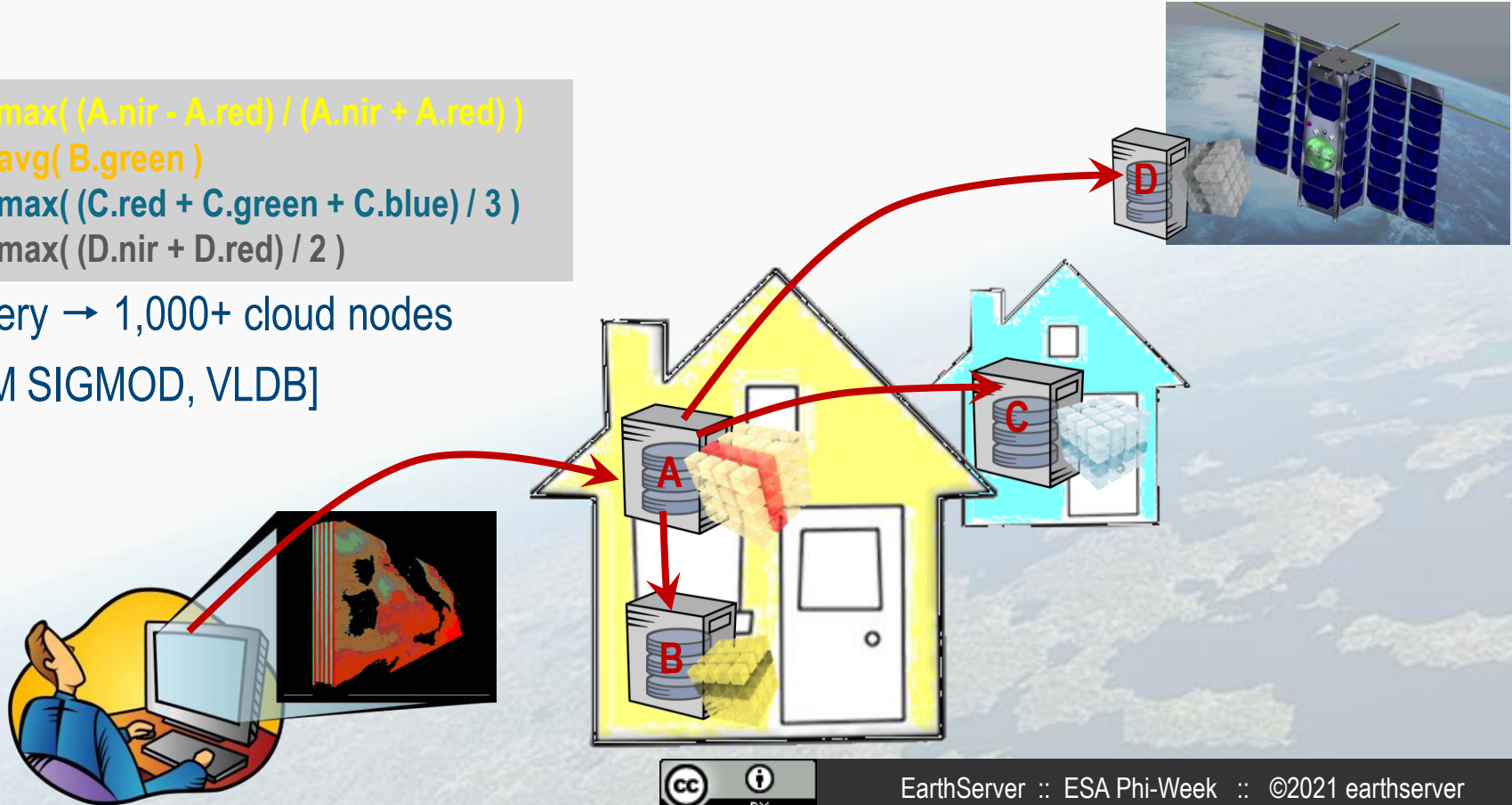
- **Rule-based classification** of field health, per crop type

  - In-field anomalies, frost, drought

  - complex criteria on timeseries: Sentinel-2a,
    CORINE land cover, soil data, climate water balance

- Lead: Spatial Business Integration, Germany

- Using rasdaman in realtime; Germany ~1h
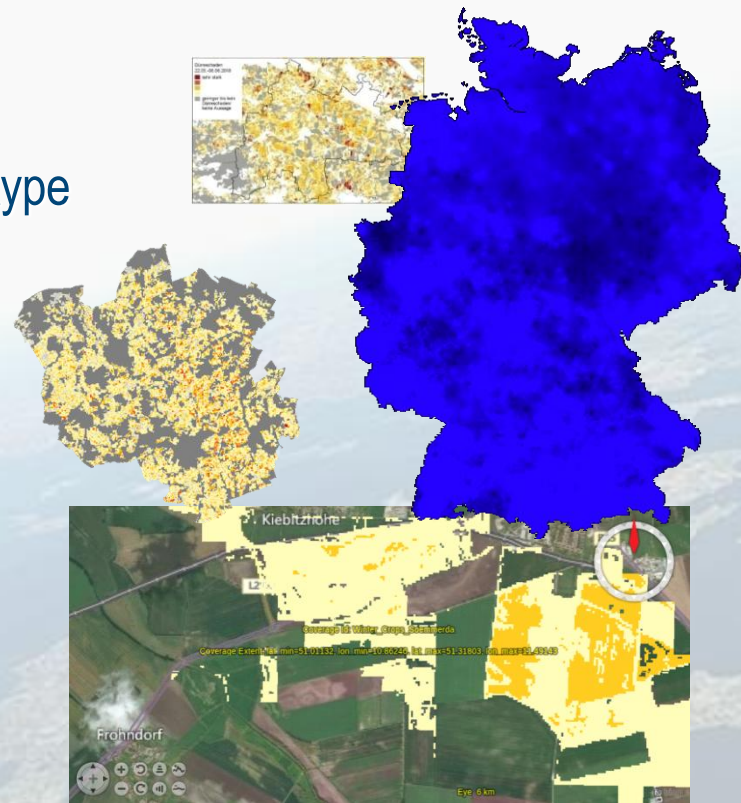
Radio Networks Planning

- Telefonica:

  series of 3D x/y/t datacubes

  of their German mobile networks

  - original ground resolution

  - space/time coverage analysis

- Analytics & marketing (!)

[o2online.de/netz]

Über 44 Mio. Mobilfunkanschlüsse. Das O₂ Netz verbindet die meisten Menschen in Deutschland.

Warum wollen wir über das Netz sprechen? Das O₂ Netz ist so gut wie noch nie.

Beim größten 4G-Ausbau der Unternehmensgeschichte hat O₂ alleine im Jahr 2020 11.000 zusätzliche 4G-Sender installiert.

Dank dieser Investitionen versorgt das O₂ Netz bundesweit zusätzliche 7 Millionen Menschen mit schnellem 4G und erreicht insgesamt 99 Prozent der Haushalte in Deutschland.

Überzeuge dich selbst und sieh hier die Entwicklung der LTE-Versorgung über alle Frequenzen, die Telefónica für LTE einsetzt. Die Daten beziehen sich auf den Zeitraum vom Januar 2020 bis April 2021. Je dunkler die Flächen desto länger liegt LTE-Verfügbarkeit vor.

Jan.'20   Juli '20   Dez.'20   April '21

AI + Datacubes

Big **Data** Analytics

Datacubes:
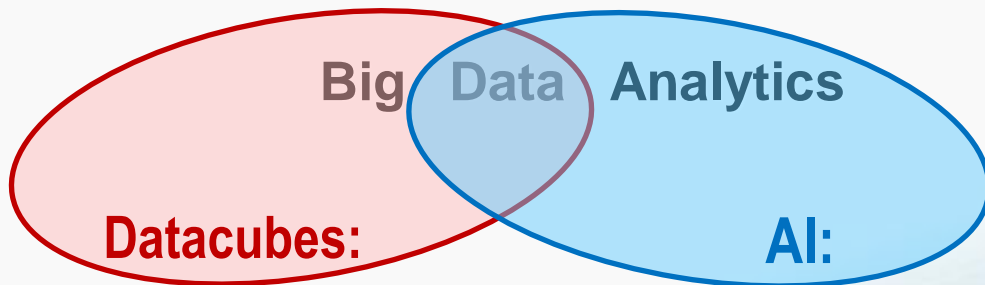
– Flexibility of search

– Federation

– ARD

– Scalability

AI:

– Flexibility of topics

– Enhanced Insight

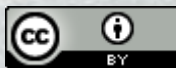– Large library of training data

- Tech: disruptive QL-based tensor algebra approach

- User benefit: easier, faster, flexible → basis for Digital Twins: any-time analysis, what-if, ...

Conclusion

- Datacubes accepted ARD cornerstone

  - *Single* datacube = homogeneous = ARD

  - *Federations* heterogeneous ➜ EarthServer as enabler

- EarthServer: Critical mass, continuously growing

  - Multi-dozen PB of data, function rich, extensible

  - location-transparent, open standards, coding-free

- Open, free, transparent, democratic

  - quickly growing community

  - *Have data?* *Join!* – *Wanna benefit from data?* *Join!*
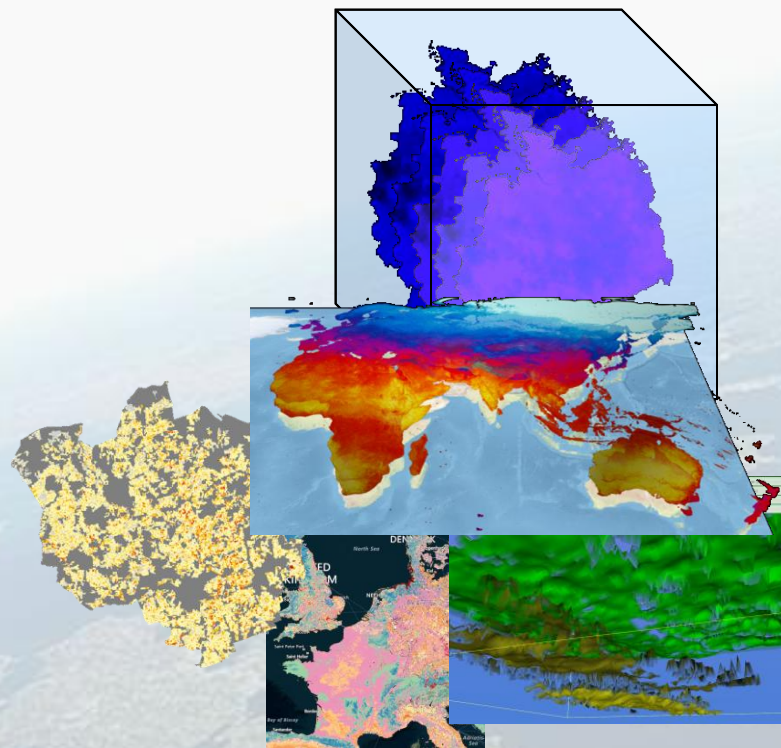
Side Event Program

17:30 – 17:45   **Federation: the Next Key Innovation After ARD**
Peter Baumann

17:45 – 18:00   **A Federated National Datacube Repository**
Chen-Yu Hao, GIS Taiwan

18:00 – 18:15   **Datacubes: a Cloud Operator's Perspective**
Jurry del la Mar, T-Systems

18:15 – 18:30   **The Geospatial Data Infrastructure of AWI's Observation to Analytics & Archive Framework**
Tilman Dinter, Alfred-Wegener-Institut

18:30 – 18:45   **Datacubes + AI Knowledge Packs = New Service Types**
Taras Matselyukh, OPT/NET

18:45 – 19:00   **LANDSUPPORT: Towards a Free Integrated Land Decision Support System**
Fabio Terribile, Universita di Napoli