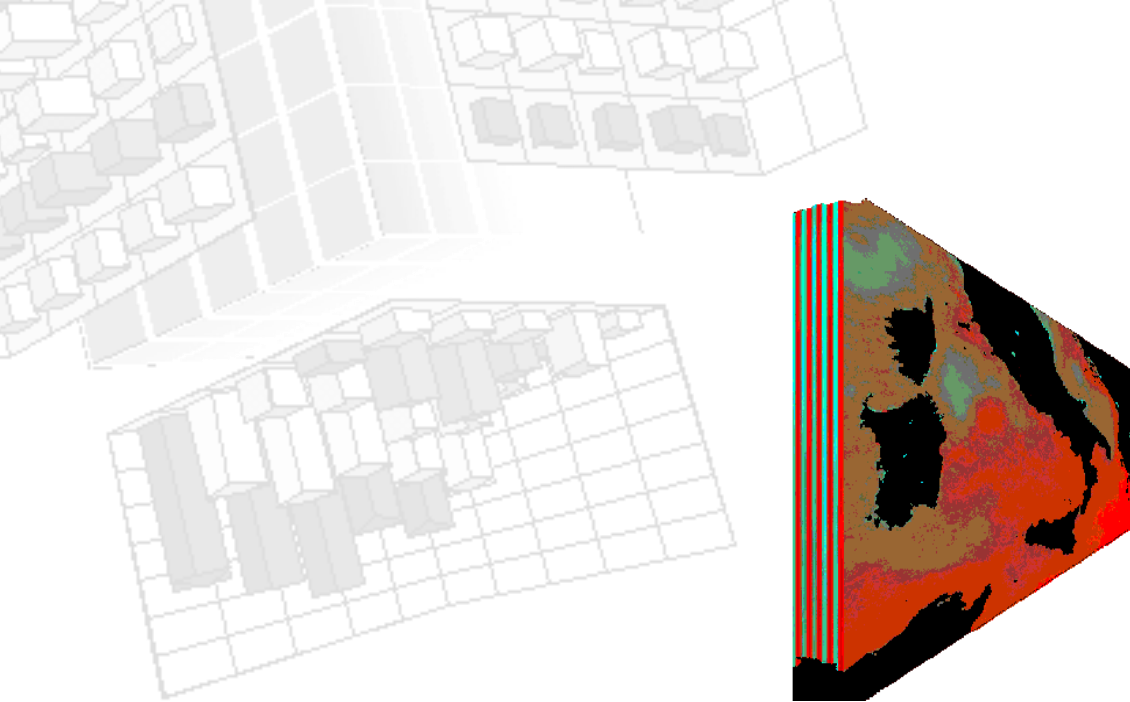




Standards für Big Geo Data

GiN Forum “Big Geo Data & INSPIRE”
14.7.2014

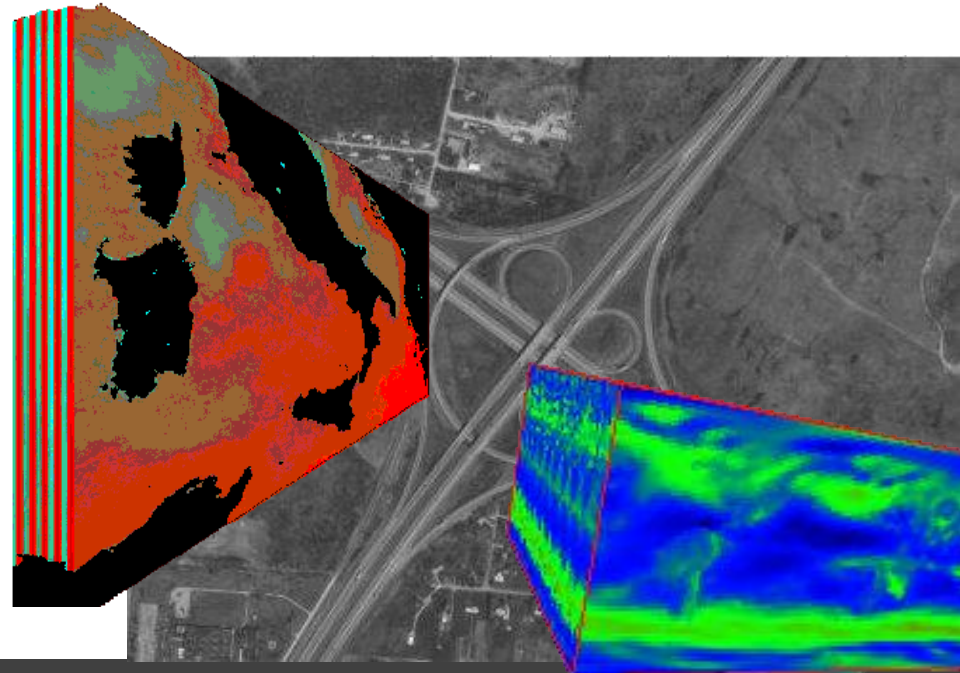
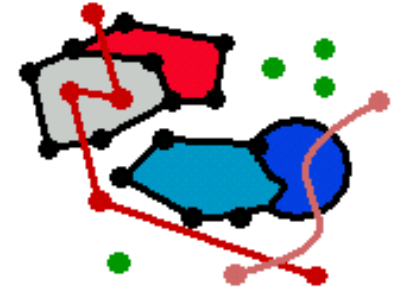
Peter Baumann
Jacobs University | rasdaman GmbH
p.baumann@jacobs-university.de

Overview

- Big Data – why should we care?
- OGC
- INSPIRE
- ISO
- Wrap-up

Geo Data, as per OGC & ISO

- geographic object = **feature**
- Special kind of feature: **coverage**
 - Raster data & more...
- Typically, coverages are the **Big Geo Data**

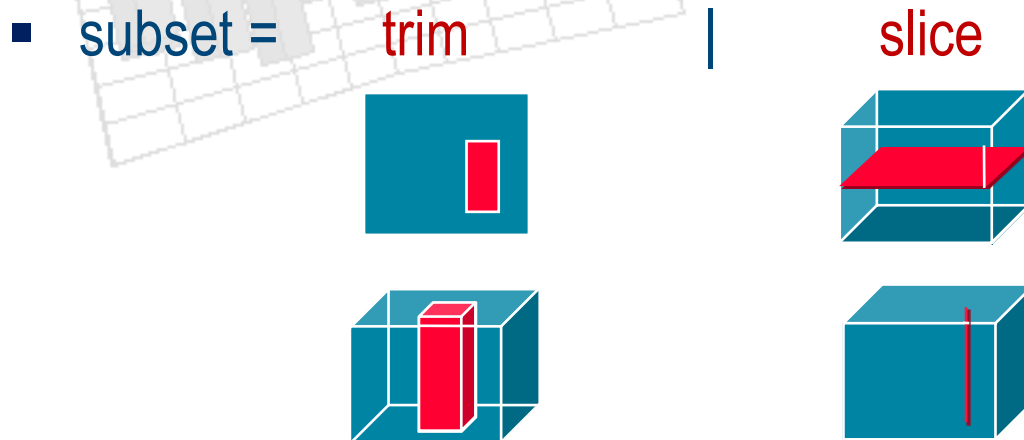


- regular & irregular grids, point clouds, meshes
 - All spatio-temporal



Web Coverage Service (WCS)

- **Core**: Simple & efficient access to n-D spatio-temporal coverages
 - plus format encoding

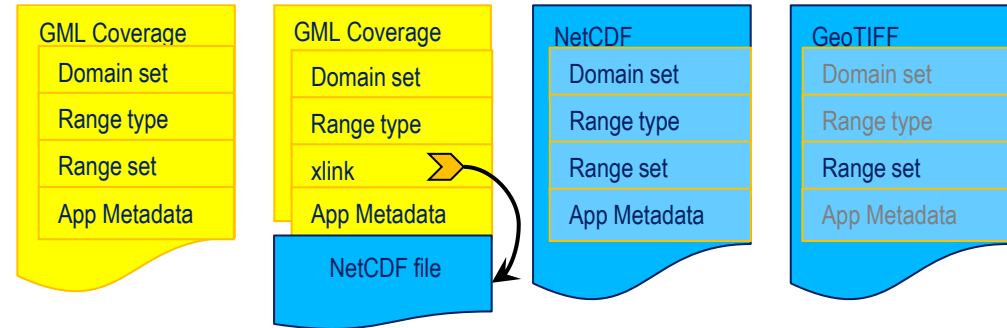


- **Extensions** add functionality facets
- **Application Profiles** bundle purpose-specific functionality

Encoding Coverages

- coverage encoding variants:

- pure GML
- Special-format
- Multipart



6

- GetCoverage* returns Native Format by default, or:

Ex:

FORMAT=application/gml+xml

FORMAT=image/tiff & MEDIATYPE=multipart/related

FORMAT=image/tiff

- What formats are supported?

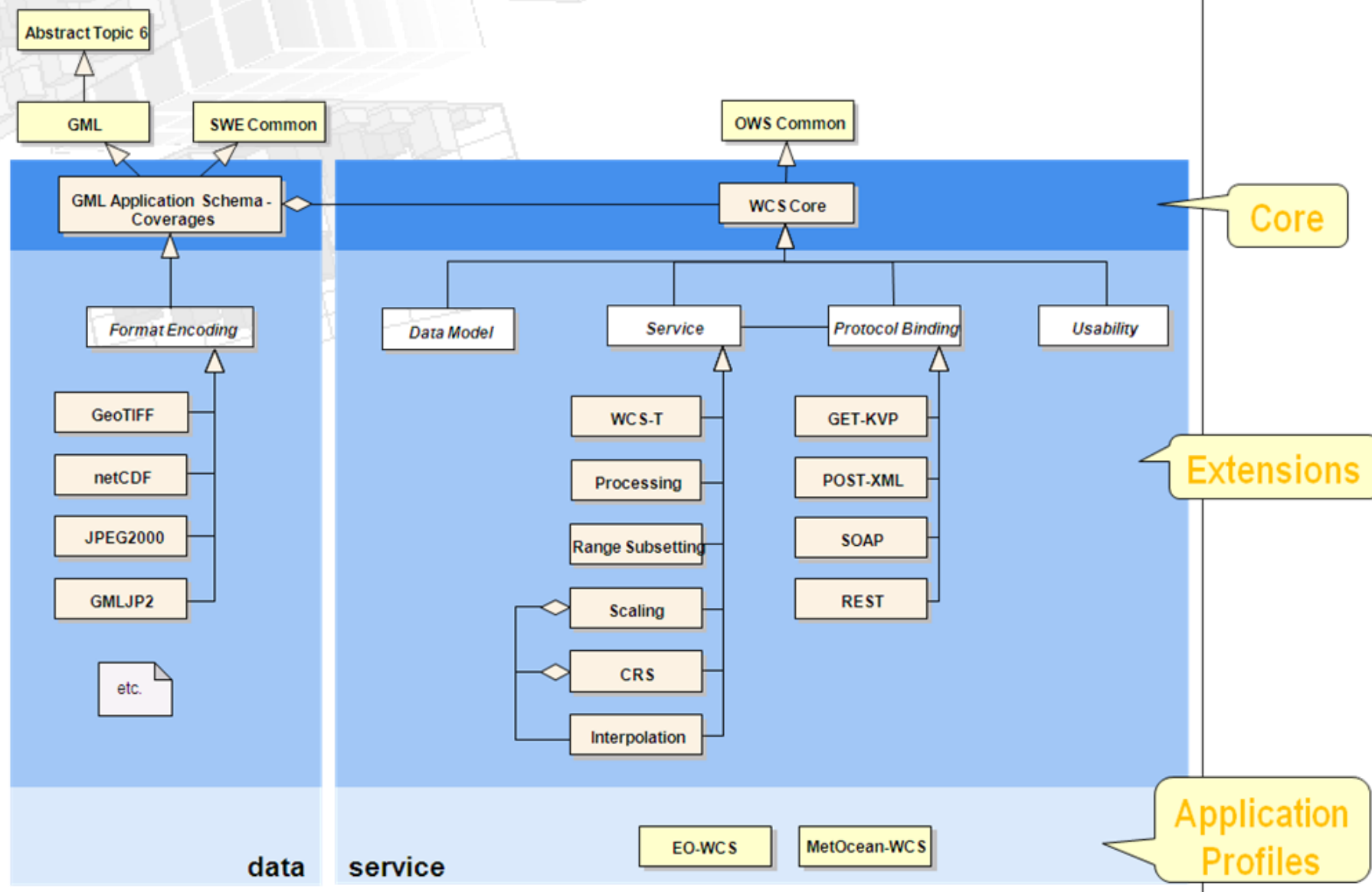
GetCapabilities: service metadata report formats available

• Ex:

<wcs:Format>image/tiff</Format>

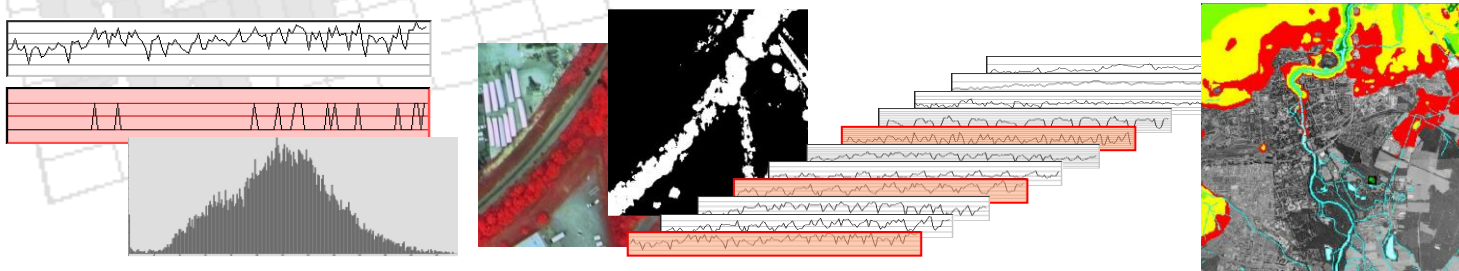
WCS: The Big Picture

class WCS Suite Logical View



OGC Web Coverage Processing Service

- OGC Web Coverage Processing Service (WCPS) - adopted 2008
= high-level grid coverage filtering & processing language



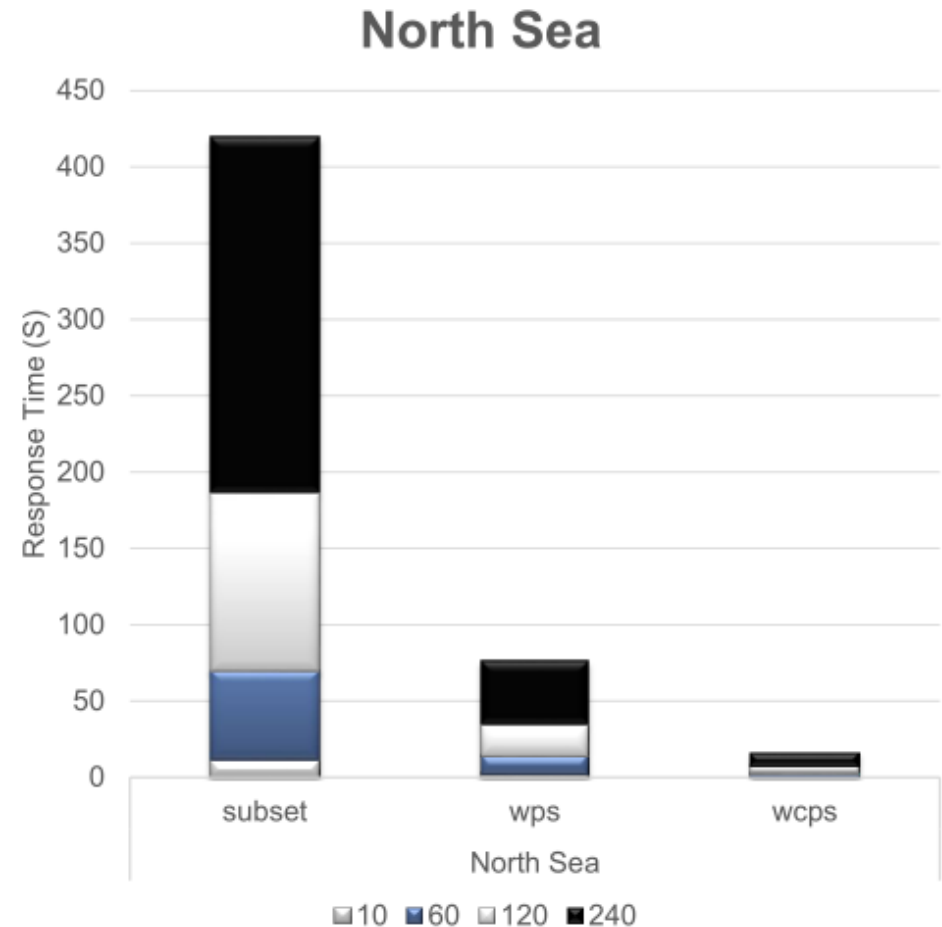
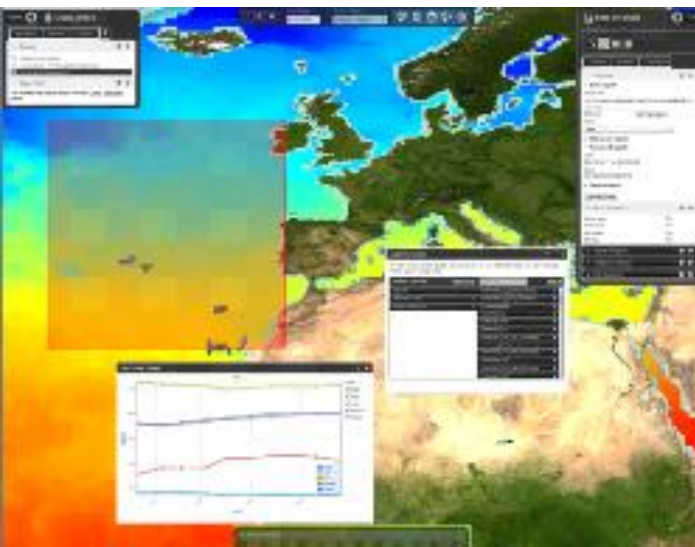
- "From MODIS scenes M1, M2, M3: difference between red & nir, as TIFF"
 - ...but only those where nir exceeds 127 somewhere

```
for $c in ( M1, M2, M3 )
where
    some( $c.nir > 127 )
return
    encode(
        $c.red - $c.nir,
        "image/tiff"
    )
```

(tiff_A,
tiff_C)

Use Case: Plymouth Marine Laboratory

- “Avg chlorophyll concentration for given area & time period, from x/y/t cube”
 - 10, 60, 120, 240 days
- Conclusions:
 - „we must minimise data transfer as well as [client-side] processing”
 - “standards such as WCPS provide the greatest benefit”



Semantic Interoperability: WPS vs WCPS

- WCPS: semantics in query → machine understandable

```
for $c in ( M1, M2, M3 )
return encode abs( $c.red - $c.nir ), "hdf" )
```

- WPS: semantics in human-readable text

```
<ProcessDescriptions ...>
  <ProcessDescription processVersion="2" storeSupported="true" statusSupported="false">
    <ows:Identifier>Buffer</ows:Identifier>
    <ows:Title>Create a buffer around a polygon.</ows:Title>
    <ows:Abstract>Create a buffer around a single polygon. Accepts the polygon as GML and
provides GML output for the buffered feature. </ows:Abstract>
    <ows:Metadata xlink:title="spatial" />
    <ows:Metadata xlink:title="geometry" />
    <ows:Metadata xlink:title="buffer" />
    <ows:Metadata xlink:title="GML" />
    <DataInputs>
```

WCS

data access

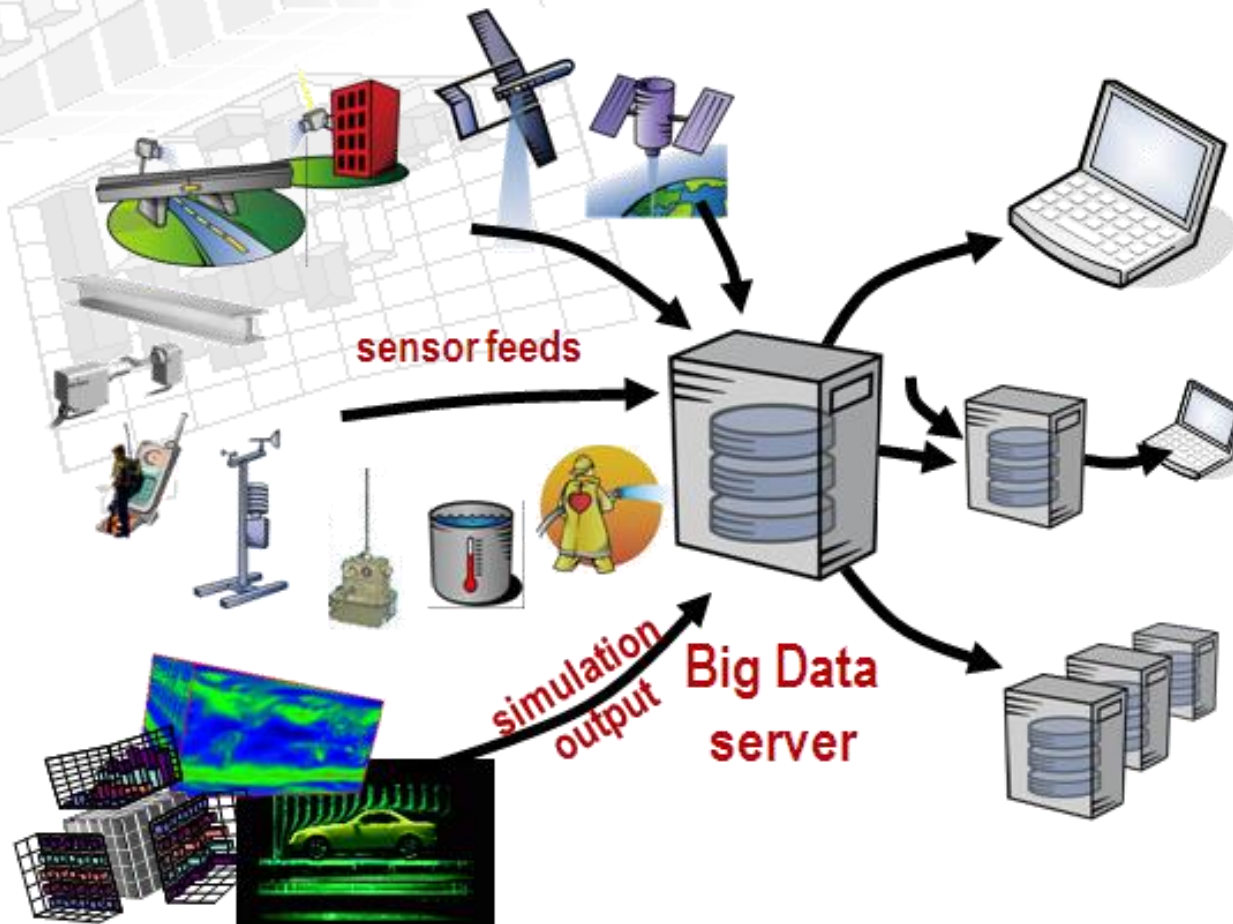
WCPS

ad-hoc retrieval

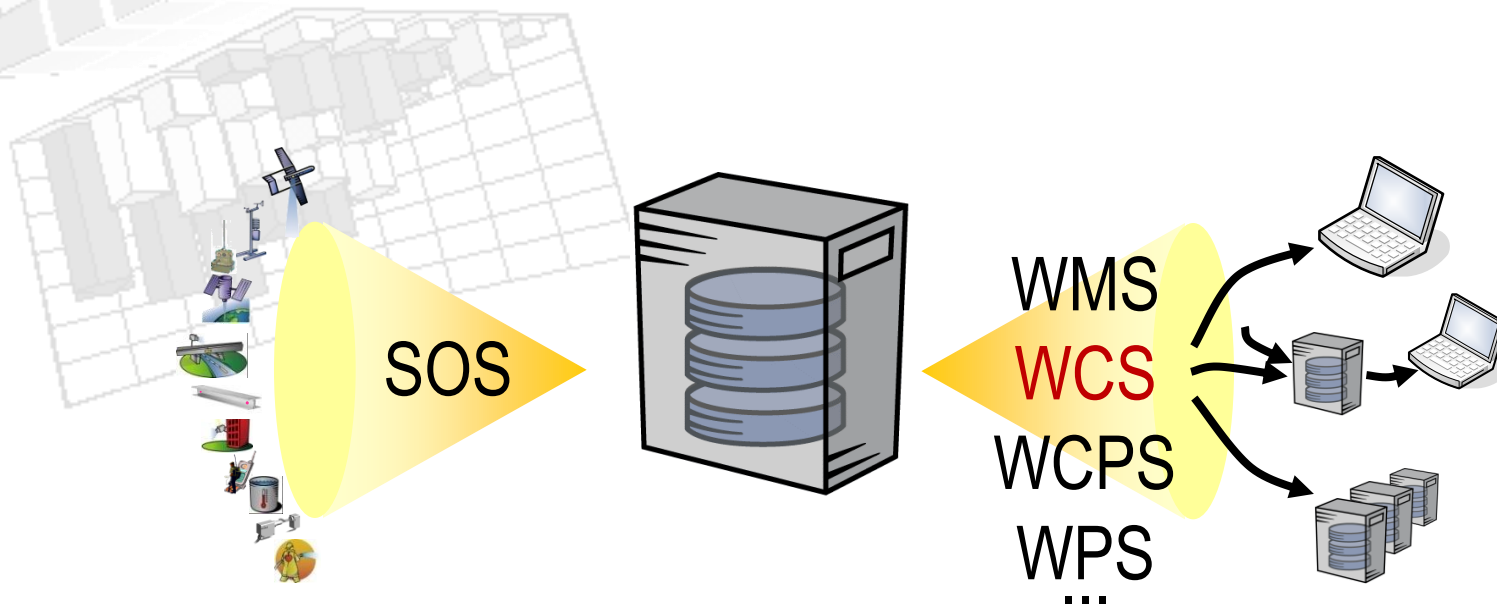
WPS

predefined process

End-to-End Architecture for Coverages



End-to-End Architecture for Coverages



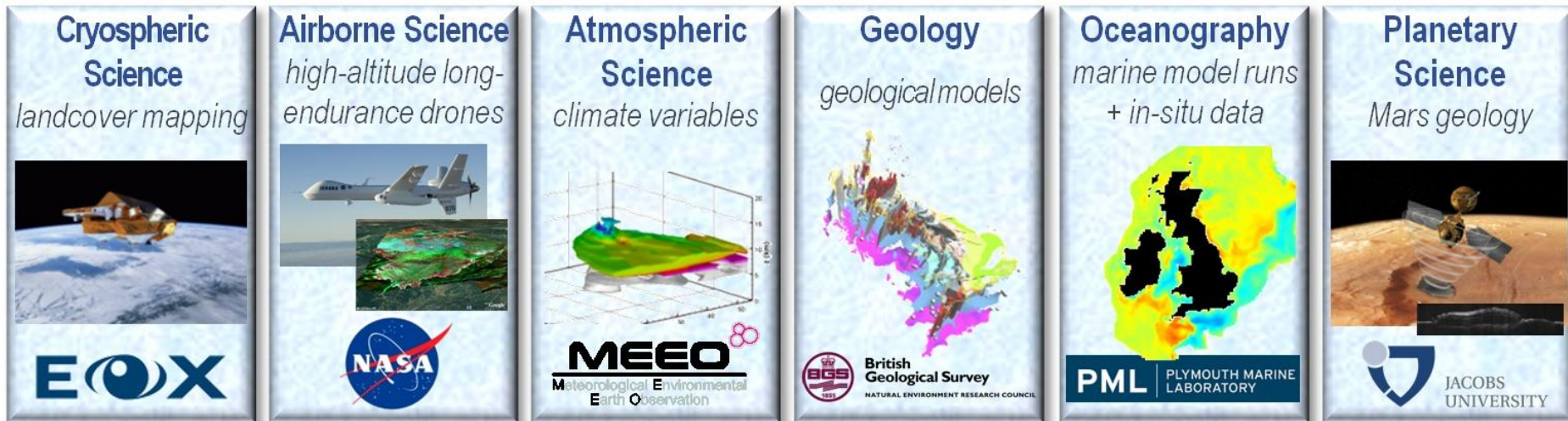
SWE, SOS: upstream
sensor data capturing

WCS: downstream
download
& processing
& visualization

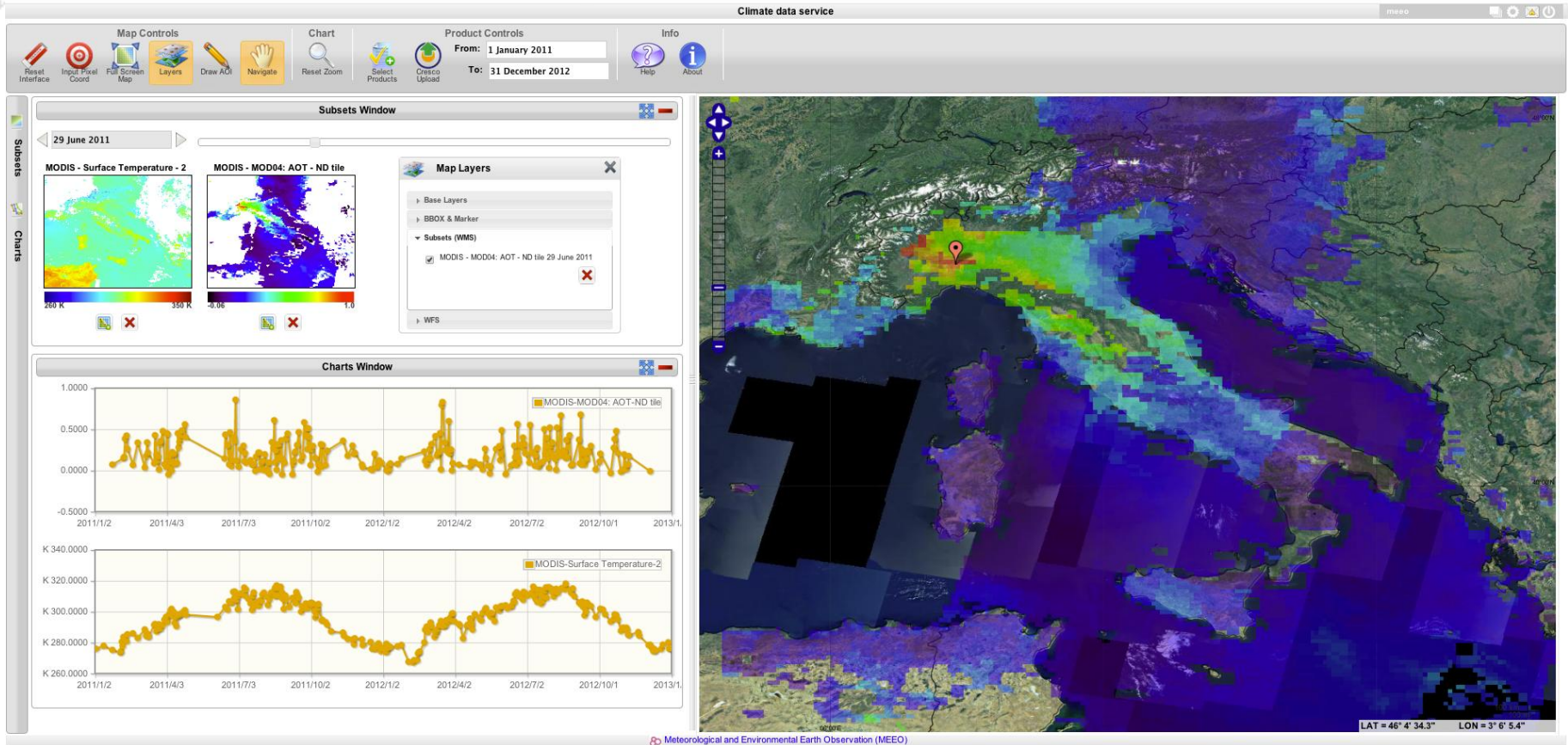


EarthServer: *Big Earth Data Analytics*

- Scalable On-Demand Processing for the Earth Sciences
 - EU FP7-INFRA, 3 years, 5.85 mEUR
- 260 TB databases for Earth & Planetary sciences
- Platform: rasdaman Array DBMS



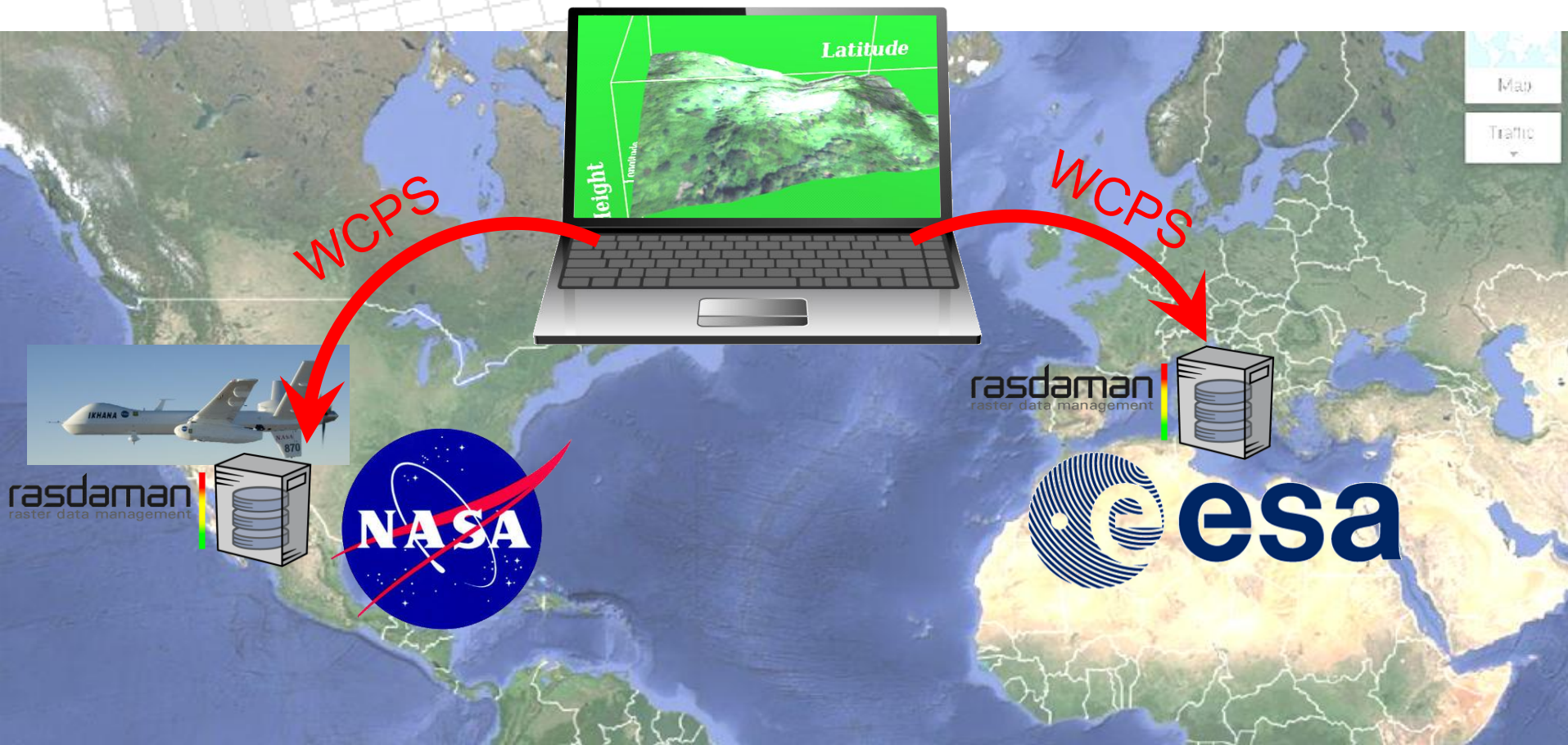
Use Case: Climate Data Service, MEEEO



[MEEEO 2013]

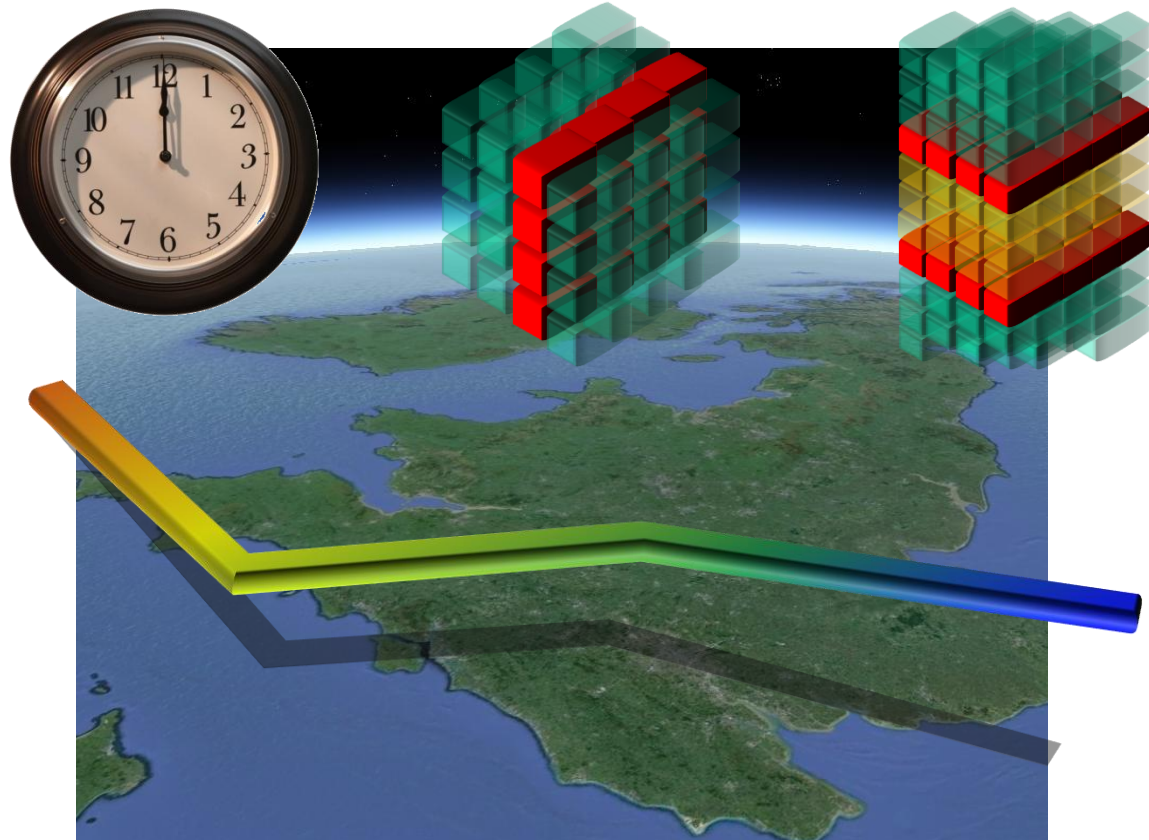
Secured Archive Integration

First-ever direct, **ad-hoc mix** from **protected** NASA & ESA services
in OGC WCS/WCPS Web client (EarthServer + CobWeb)



Outlook: MetOcean-WCS

- WCS Application Profile focussed on Meteorology, Climate, Aviation domains
- Data model:
 - 4-D x/y/z/t data cubes
- Service model:
 - Space-time extraction
 - curtains
 - corridors
- Status: drafting
 - JacobsU + UK MetOffice



Overview: Some Relevant OGC WGs

- WCS.SWG - WCS/WCPS maintenance
- Coverages.DWG - exchange with all interested domains
 - http://external.opengeospatial.org/twiki_public/CoveragesDWG/
- Temporal.DWG - establishing semantics for time & calendars
 - http://external.opengeospatial.org/twiki_public/TemporalDWG/
- BigData.DWG - well...you guess it 😊
 - http://external.opengeospatial.org/twiki_public/BigDataDwg/

INSPIRE

- Annex II + III frequently relate to coverages, Big Data
- Coverage Data Specification refers to OGC coverage std
- Download Service Specification
 - Inquiries since May 2013 remained unanswered
- Oral communication June 2014:
 - Download Services will not rely on only one standard
 - Project started to bring SOS up to speed for INSPIRE
 - „so how can INSPIRE work on WCS commence?“ – no information provided up to now

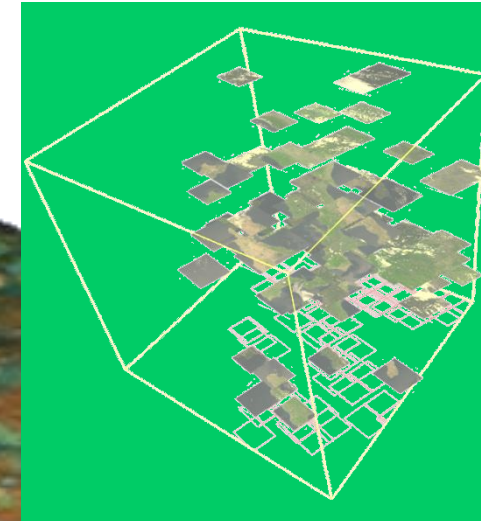
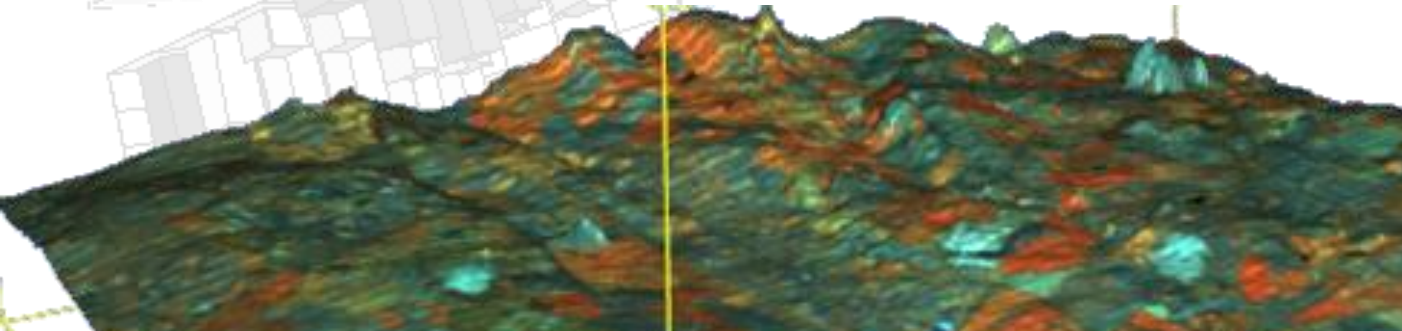


ISO Array SQL

- ISO 9075 Part 15: SQL/MDA
 - „Multi-Dimensional Arrays“
- Adding n-D arrays to Standard SQL
 - Target audience: science, engineering, ...
 - Based on concepts & experience of rasdaman
- ISO/IEC JTC1 SC32 WG3 SQL decision, June 2014
- Available in 2015 / 2016 ?



ISO Array SQL: Database Visualization



```
select
  encode(
    struct {
      red:    (char) s.image.b7[x0:x1,x0:x1],
      green:  (char) s.image.b5[x0:x1,x0:x1],
      blue:   (char) s.image.b0[x0:x1,x0:x1],
      alpha:  (char) scale( d.elev, 20 )
    },
    "image/png"
  )
from SatImage as s, DEM as d
```

ISO TC211 Select Activity

- ISO 19163 “Geographic information — Content components and encoding rules for imagery and gridded data”
- Optical + microwave + radar imagery
- ISO TC211 draft, comment period starts Aug 2014
- Controversial discussion: not aligned with OGC & INSPIRE
 - Coverages, metadata, time coordinates



Further Activities

■ Research Data Alliance (RDA)

- Big Data Analytics Interest Group
- Geospatial Interest Group
- 4th plenary: Amsterdam, 22 - 24 September 2014
- www.rd-alliance.org



■ ESA Big Data From Space conference 2014



→ CONFERENCE ON BIG DATA FROM SPACE

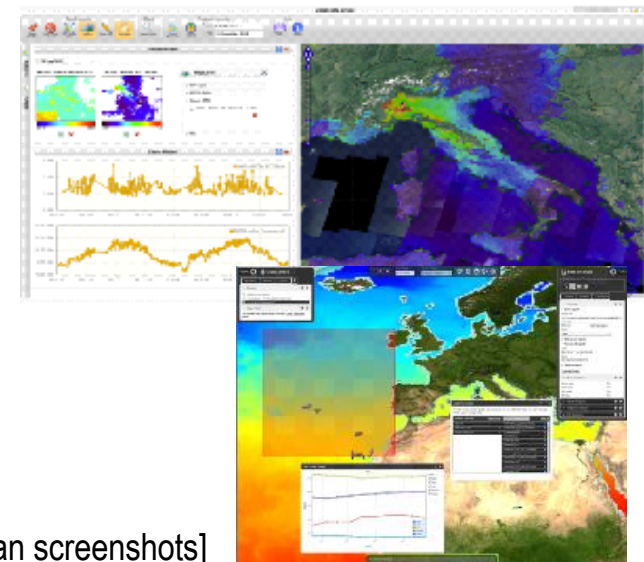
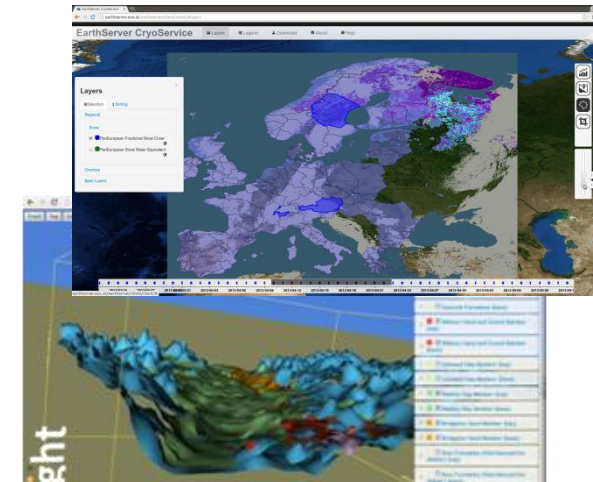
Research, Technology and Innovation (RT&I)

12-14 November 2014
ESA-ESRIN Frascati, Italy

congrexprojects.com/2014-events/BigDatafromSpace

Conclusion

- Standards leverage interoperability, vendor independence – also for Big Data
- Big Data Standardization
 - OGC.
 - Coverages: simple concept, harmonized across OGC
 - Coverage services: scalable + flexible – WCS currently most function-rich
 - ISO
 - In geo domain not yet completely Big Data aware
 - But strong horizontal Big Data activities
 - INSPIRE
 - Hard to say, currently somewhat intransparent
- Several international activities, gradual convergence visible



[rasdaman screenshots]