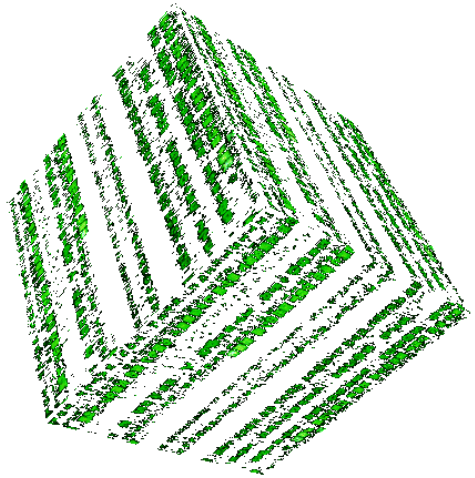


Big Spatio-Temporal Datacubes on Steroids …and Standards

FOSDEM 2017, Brussels

Dimitar Misev, Peter Baumann

Jacobs University | rasdaman GmbH

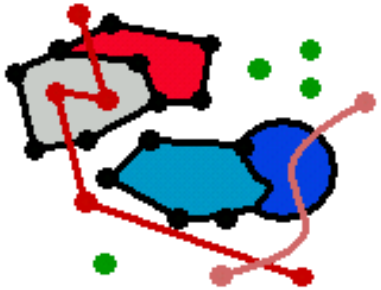
baumann@rasdaman.com

[gamingfeeds.com]

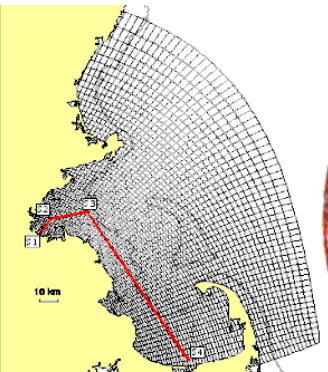
[co-funded through EU EarthServer-2]

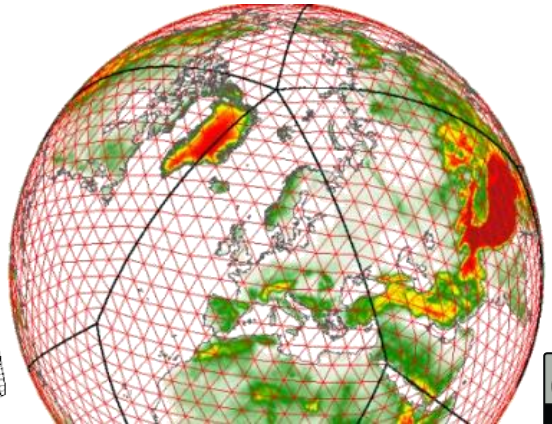
Modeling Coverages

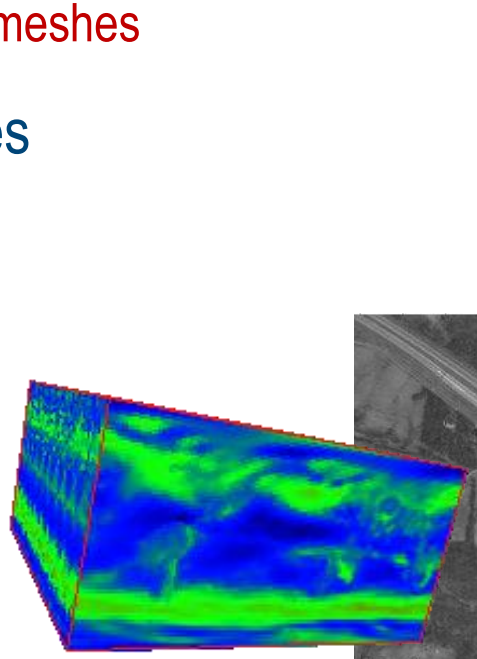
Features & Coverages

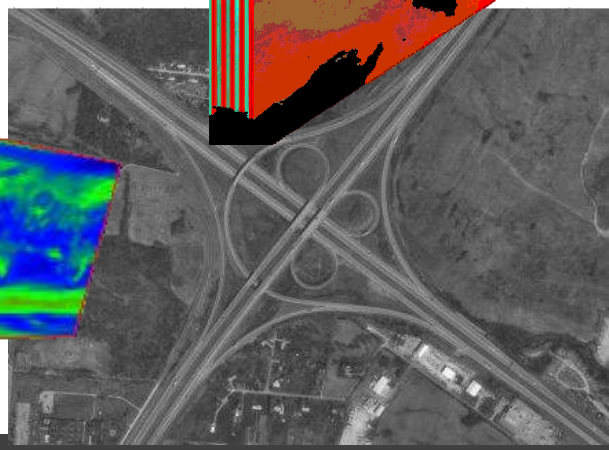


- The basis of all: geographic feature
- Special kind of feature: coverage
  - aka space-time varying phenomenon
  - regular & irregular grids, point clouds, meshes
- Usually, Big Geo Data are coverages



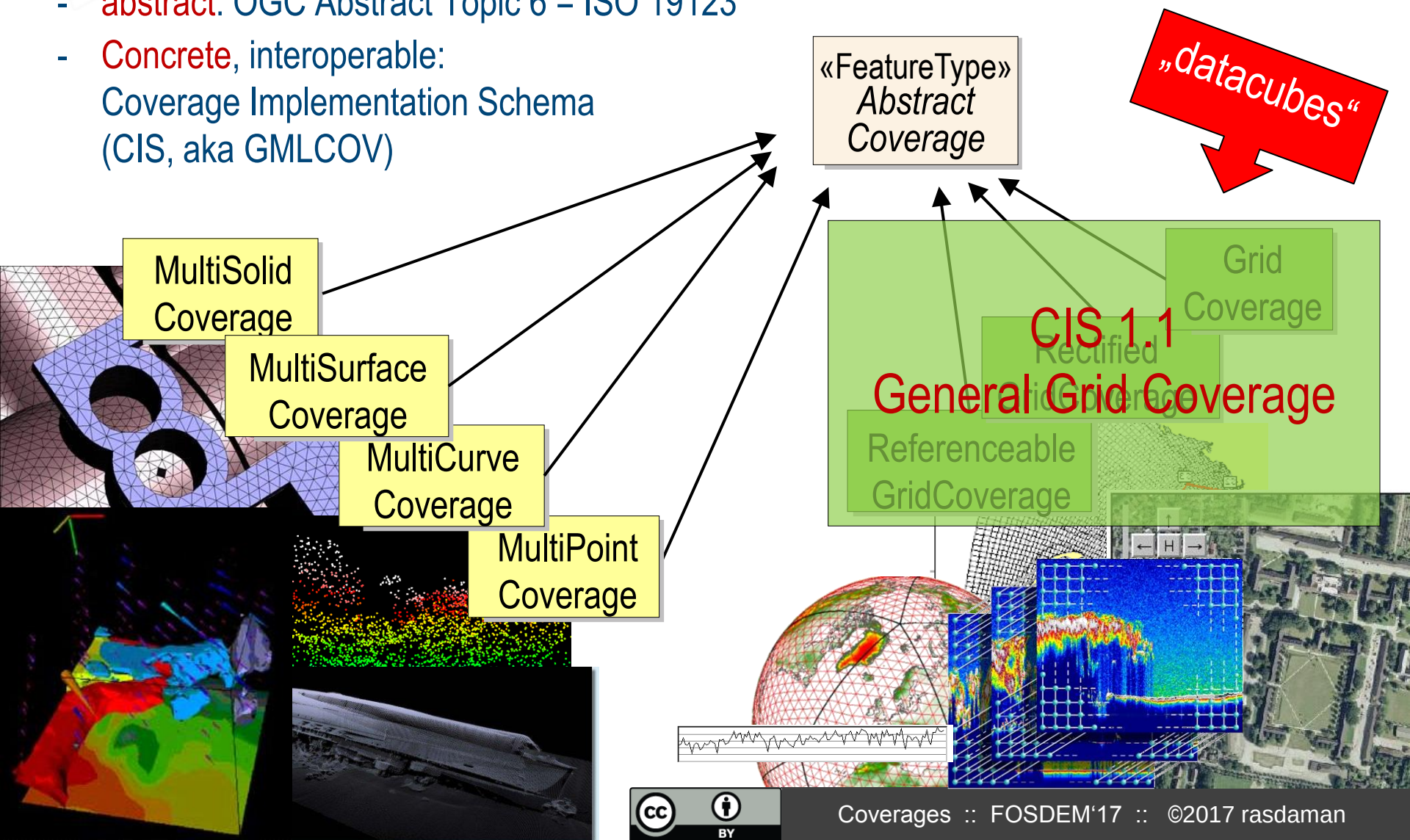




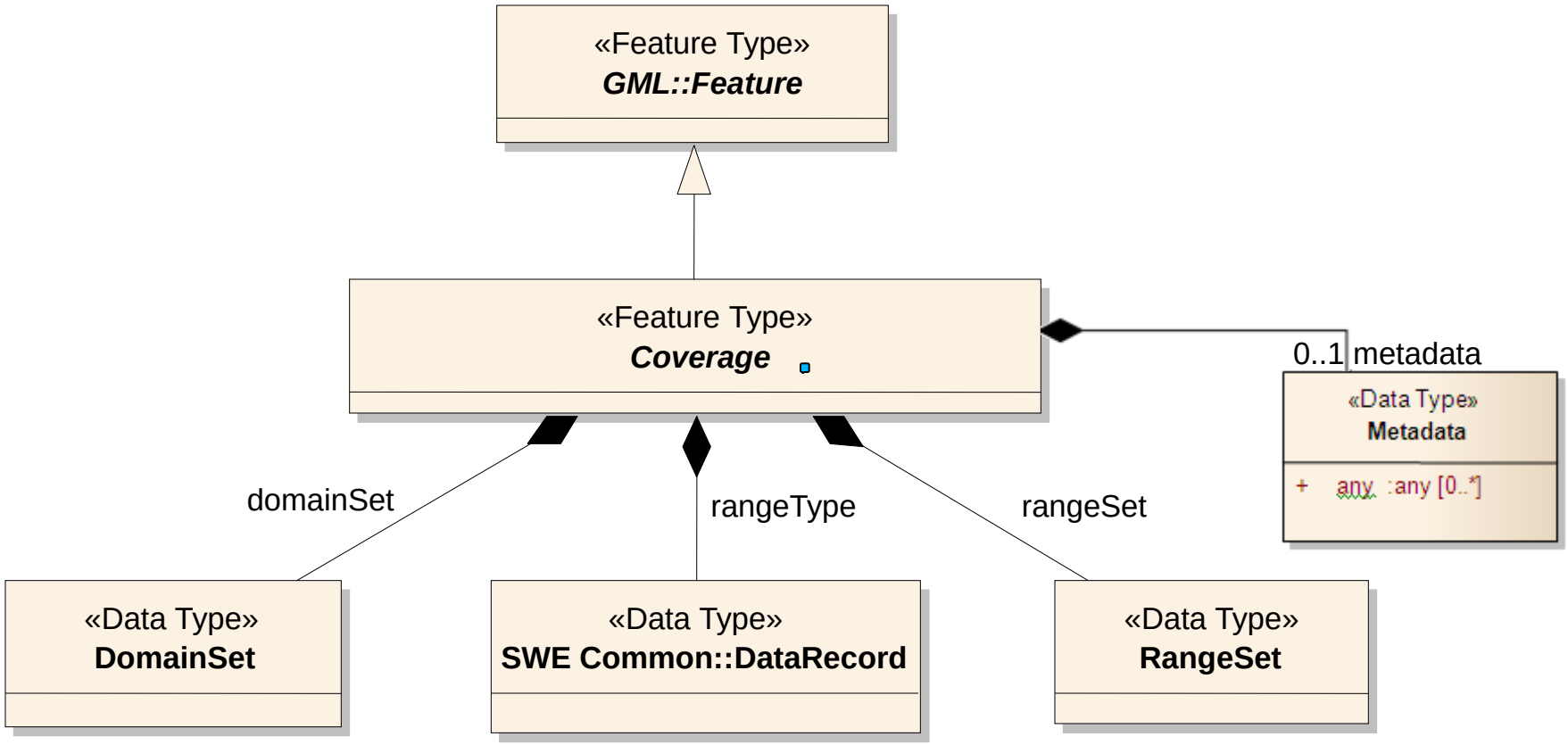


Coverages= Grids, Point Clouds, Meshes

- abstract: OGC Abstract Topic 6 = ISO 19123
- Concrete, interoperable:
  Coverage Implementation Schema
  (CIS, aka GMLCOV)



Coverage Definition



[OGC 09-146r2]

A Simple Coverage, in GML

```
<generalGridCoverage ...  gml:id="CIS_001">

    <domainSet>
        <generalGrid srsName="http://www.opengis.net/def/crs-compound?
                              1=http://www.opengis.net/def/crs/EPSG/0/4979
                            &amp;2=http://www.opengis.net/def/crs/OGC/0/AnsiDate"
            axisLabels="Lat Long h date">
            <regularAxis   axisLabel="Lat"  uomLabel="deg" lowerBound="40"  upperBound="60" resolution="10"/>
            <regularAxis   axisLabel="Long" uomLabel="deg" lowerBound="-10" upperBound="10"  resolution="10"/>
            <irregularAxis axisLabel="h"    uomLabel="m">
                <c>  0</c>
                <c>100</c>
            </irregularAxis>
            <irregularAxis axisLabel="date" uomLabel="d">
                <c>2015-12-01</c>
                <c>2015-12-02</c>
            </irregularAxis>
            <gridLimits srsName="http://www.opengis.net/def/crs/OGC/0/Index4D" axisLabels="i j k l">
                <indexAxis axisLabel="i" lowerBound="0" upperBound="2"/>
                <indexAxis axisLabel="j" lowerBound="0" upperBound="2"/>
                <indexAxis axisLabel="k" lowerBound="0" upperBound="1"/>
                <indexAxis axisLabel="l" lowerBound="0" upperBound="1"/>
            </gridLimits>
        </generalGrid>
    </domainSet>

    <rangeSet>
        <dataBlock>
              <v>01</v> <v>02</v> <v>03</v> <v>04</v> <v>05</v> <v>06</v> <v>07</v> <v>08</v> <v>09</v>
              <v>01</v> <v>02</v> <v>03</v> <v>04</v> <v>05</v> <v>06</v> <v>07</v> <v>08</v> <v>09</v>
              <v>01</v> <v>02</v> <v>03</v> <v>04</v> <v>05</v> <v>06</v> <v>07</v> <v>08</v> <v>09</v>
              <v>01</v> <v>02</v> <v>03</v> <v>04</v> <v>05</v> <v>06</v> <v>07</v> <v>08</v> <v>09</v>
        </dataBlock>
    </rangeSet>

    <rangeType>
        <swe:DataRecord>
            <swe:field name="panchromatic">
                <swe:Quantity definition="http://opengis.net/def/property/OGC/0/Radiance">
                    <swe:uom code="W.m-2.sr-1.nm-1"/>
                </swe:Quantity>
            </swe:field>
        </swe:DataRecord>
    </rangeType>

</generalGridCoverage>
```

A Simple Coverage, in JSON

```
{   "type": "CoverageByDomainAndRangeType",
    "domainSet":{
        "type": "DomainSetType",
        "generalGrid":{
            "type":          "GeneralGridCoverageType",
            "srsName":    "http://www.opengis.net/def/crs/OGC/0/Index2D",
            "axisLabels": ["i", "j"],
            "axis": [{      "type": "IndexAxisType“, "axisLabel": "i", "lowerBound": 0, "upperBound": 2
            },{               "type": "IndexAxisType“, "axisLabel": "j", "lowerBound": 0, "upperBound": 2  }]
        }
    },
    "rangeSet": {      "type": "RangeSetType",
                       "dataBlock": { "type": "VDataBlockType",  "values": [1,2,3,4,5,6,7,8,9] }     },
    "rangeType": {   "type": "DataRecordType",
                       "field":[{     "type": "QuantityType",
                                      "definition": "ogcType:unsignedInt",
                                      "uom": { "type": "UnitReference", "code": "10^0“ }     }]
    }
}
```

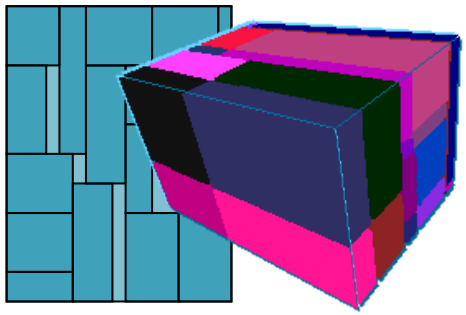
A Simple Coverage, in RDF

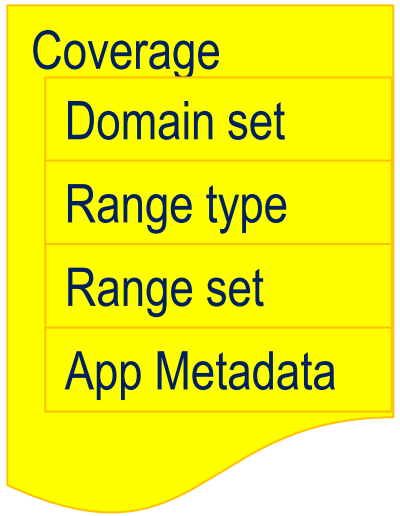
```
<http://www.opengis.net/cis/1.1/examples/CIS_05_2D>
 <http://www.w3.org/1999/02/22-rdf-syntax-ns#type>
 <http://www.opengis.net/cis/1.1/CoverageByDomainAndRangeType> .

<http://www.opengis.net/cis/1.1/examples/CIS_05_2D>
 <http://www.opengis.net/cis/1.1/domainSet>
 <http://www.opengis.net/cis/1.1/examples/CIS_DS_05_2D> .
<http://www.opengis.net/cis/1.1/examples/CIS_DS_05_2D>
 <http://www.opengis.net/cis/1.1/generalGrid>
 <http://www.opengis.net/cis/1.1/examples/CIS_DS_GG_05_2D> .
<http://www.opengis.net/cis/1.1/examples/CIS_DS_05_2D>
 <http://www.w3.org/1999/02/22-rdf-syntax-ns#type>
 <http://www.opengis.net/cis/1.1/DomainSetType> .
<http://www.opengis.net/cis/1.1/examples/CIS_DS_GG_05_2D>
 <http://www.opengis.net/cis/1.1/axis>
 <http://www.opengis.net/cis/1.1/examples/CIS_DS_GG_I_05_2D> .
<http://www.opengis.net/cis/1.1/examples/CIS_DS_GG_05_2D>
 <http://www.opengis.net/cis/1.1/axis>
 <http://www.opengis.net/cis/1.1/examples/CIS_DS_GG_J_05_2D> .
<http://www.opengis.net/cis/1.1/examples/CIS_DS_GG_05_2D>
 <http://www.opengis.net/cis/1.1/axisLabels>
 <http://www.opengis.net/cis/1.1/axisLabels0> .
<http://www.opengis.net/cis/1.1/axisLabels0> <http://www.w3.org/1999/02/22-rdf-syntax-ns#first> "i" .
<http://www.opengis.net/cis/1.1/axisLabels0> <http://www.w3.org/1999/02/22-rdf-syntax-ns#rest> <http://www.opengis.net/cis/1.1/axisLa
<http://www.opengis.net/cis/1.1/axisLabels1> <http://www.w3.org/1999/02/22-rdf-syntax-ns#first> "j" .
<http://www.opengis.net/cis/1.1/axisLabels1> <http://www.w3.org/1999/02/22-rdf-syntax-ns#rest> <http://www.w3.org/1999/02/22-rdf-sy
```

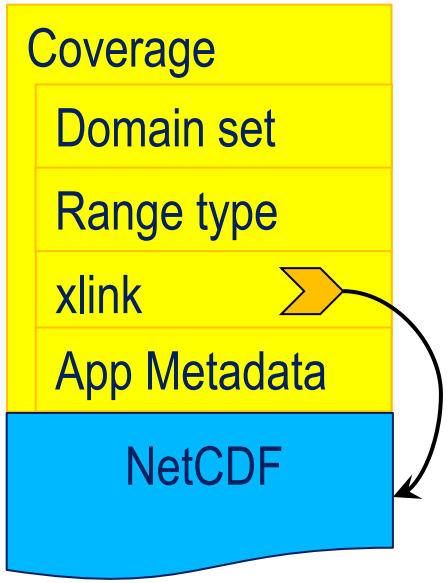
Encoding CIS 1.1 Coverages

- Single file encoding:
  - Informationally complete: GML, JSON, RDF, …
  - Further formats: GeoTIFF, NetCDF, JPEG2000, GRIB, …

- Multipart: container(“header” + file1 + file2 + …)
  - Multipart/MIME, zip, GMLJP2, SAFE, GeoPackage, ...
  - Built-in collections / tiling



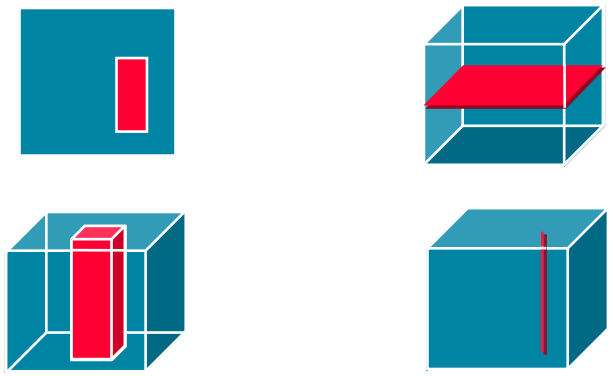




Serving Coverages

Web Coverage Service (WCS)

- WCS Core: access to spatio-temporal coverages & subsets
  - subset = trim | slice



- WCS Extensions: optional facets
- WCS Application Profiles: bespoke fct packaging
- *rasdaman supporting all WCS Extensions*

Large, growing implementation basis: rasdaman, GDAL, QGIS, OpenLayers, OPeNDAP, MapServer, GeoServer, GMU, NASA WorldWind, EOxServer; Pyxis, ERDAS, ArcGIS, ...

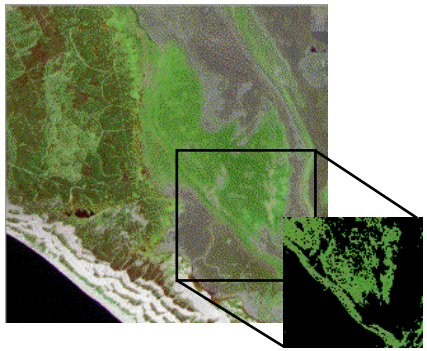
rasdaman

www.rasdaman.org

= „raster data manager“: SQL+ n-D arrays
  - Scalable parallel “tile streaming” peer architecture

- Mature, in operational use, on OSGeo Live
  - Ex: www.planetserver.eu

- ESA 2017: “world leading environment”, “standard working horse for OGC standardisation on these innovative data access interfaces“



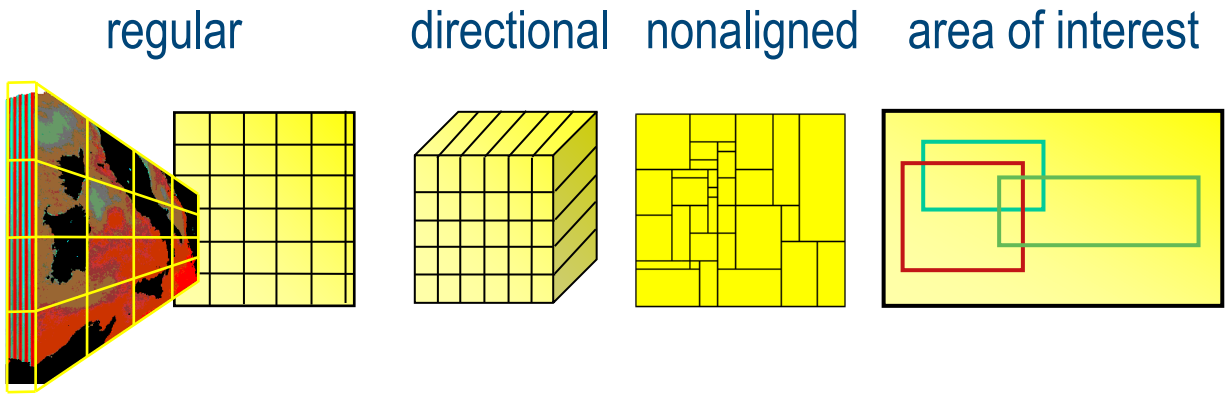
CIO 100 MOST PROMISING BIG DATA Review SOLUTION PROVIDERS 2016

GEOSPATIAL WORLD FORUM

The Earth Monitoring Competition
Copernicus masters
WINNER
T-SYSTEMS BIG DATA CHALLENGE
2014
T · · Systems ·

Adaptive Partitioning & Parallelization

- Any partitioning („tiling“), canned into strategies [ICDE 1999]

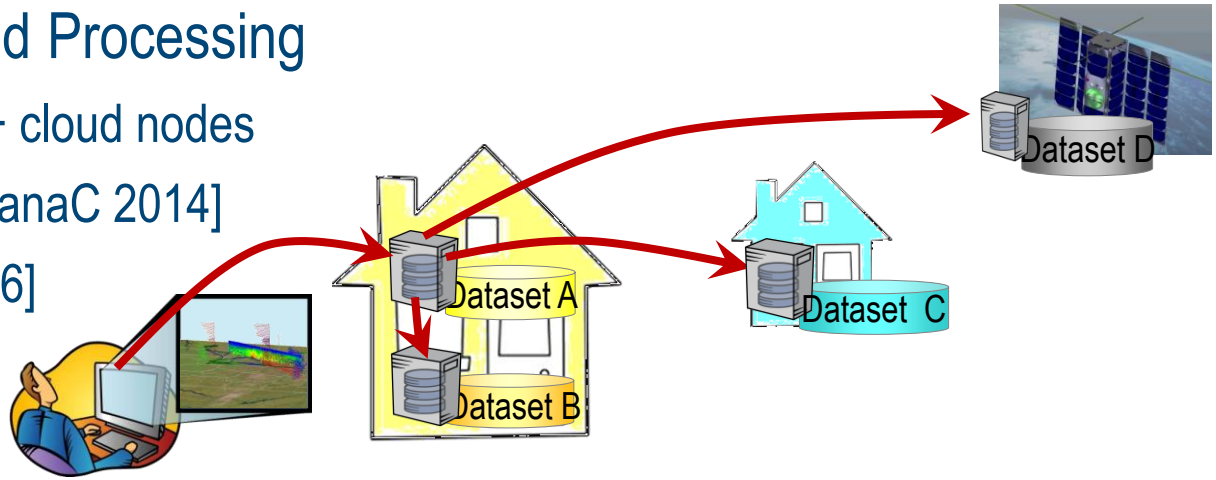


- Parallel, Distributed Processing

  1 query → 1,000+ cloud nodes

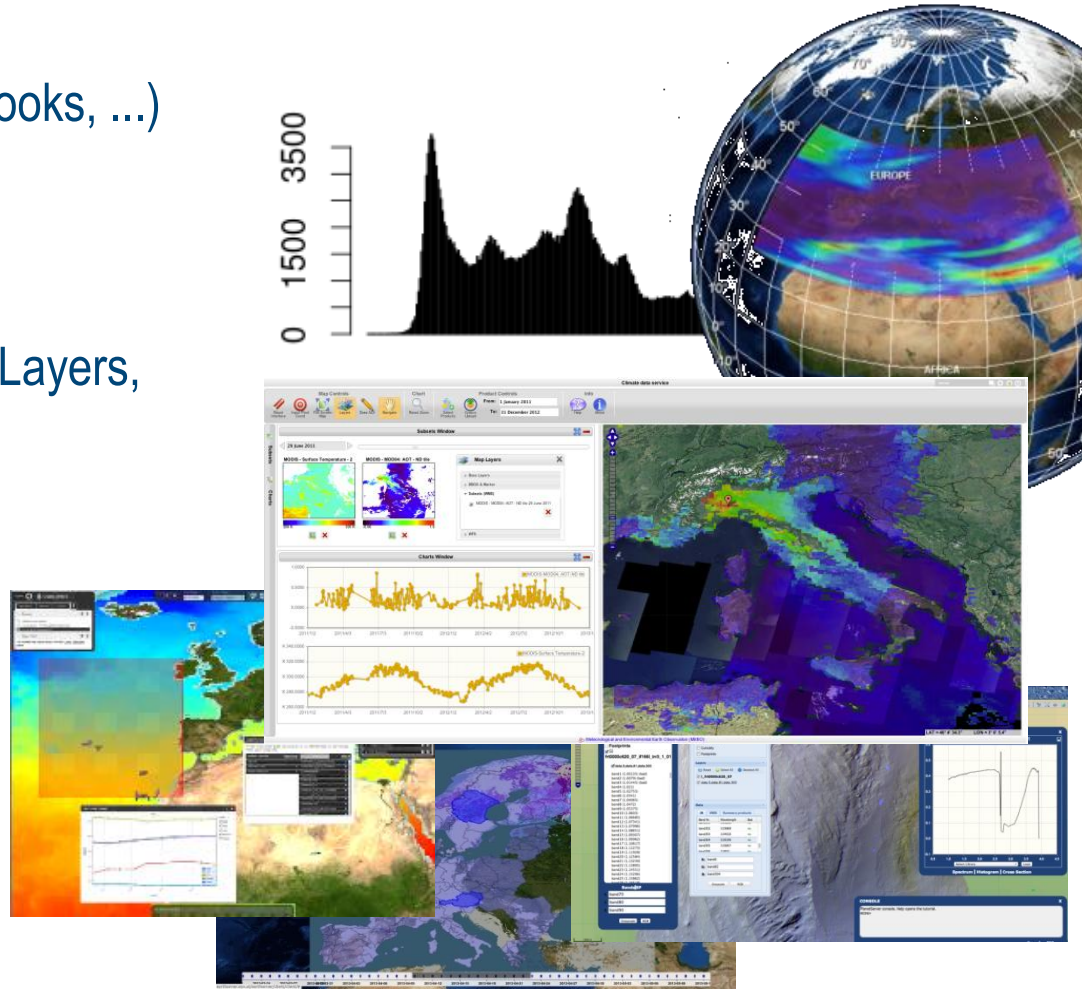
  [ACM SIGMOD DanaC 2014]

  [VLDB BOSS 2016]



Science & GIS Tool Interfacing

- General-purpose scientist tools:
  - Java, C++
  - python (OWSLIB, Jupyter notebooks, ...)
  - R (under work)
- Geo tools:
  - MapServer, GDAL, QGIS, OpenLayers, Leaflet, NASA WorldWind, ...
- OGC, INSPIRE WCS
  Reference Implementation
  - interface all tools supporting OGC's „Big Geo Data" standards suite

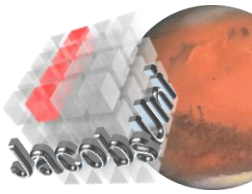


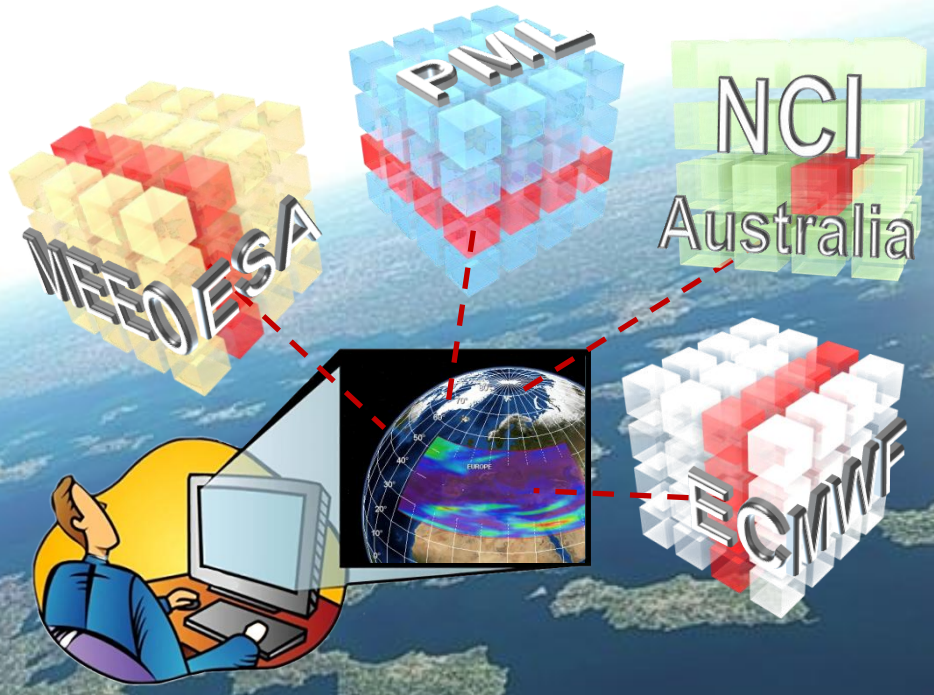
/etc

rasdaman raster data management | JACOBS UNIVERSITY

EarthServer: Datacubes At Your Fingertips

- Agile Analytics on x/y/t + x/y/z/t Earth & Planetary datacubes
  - EU rasdaman + NASA WorldWind
  - 100s of TB sites now, next: 1+ Petabyte per cube
- Intercontinental initiative, 3+3 years:
  EU + US + AUS
- Global data federation
  - Access, extract, aggregate, combine any-size datacubes
  - Common basis: OGC WCS





Co-funded by the European Union

www.earthserver.eu

OSGeo Experiences & Thoughts

- OSGeo important association, but opportunities for improvement
  - observations based on:
    - *engaged in open-source software since 2008*
    - *OSGeo Charter Member*
    - *Executed OSGeo incubation procedure with rasdaman*

OSGeo Experiences & Thoughts

- OSGeo important association, but opportunities for improvement
- Organizational Maturity: Process definition & implementation, QM
  - No inventory of decisions taken
  - Incubation management: pyWPS 8 years; rasdaman 6.5 years
  - Apply incubation criteria to itself (first)
  - In elections, typically insiders recommend each other

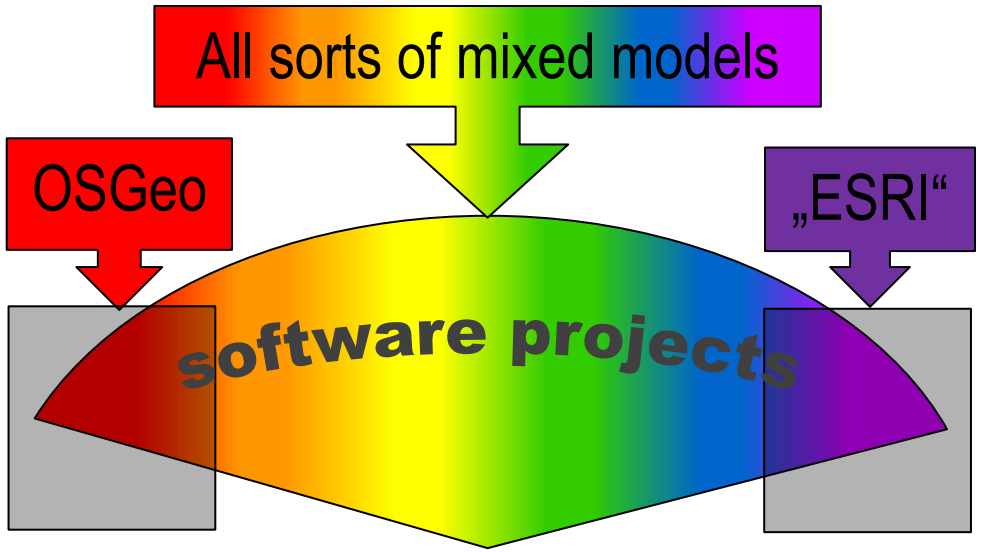
OSGeo Experiences & Thoughts

- OSGeo important association, but opportunities for improvement
- Organizational Maturity: Process definition & implementation, QM
- Focus on Core Mission
  - Should brand „good software“, not conquer project
  - „design by committee“ over „expert leadership“

OSGeo Experiences & Thoughts

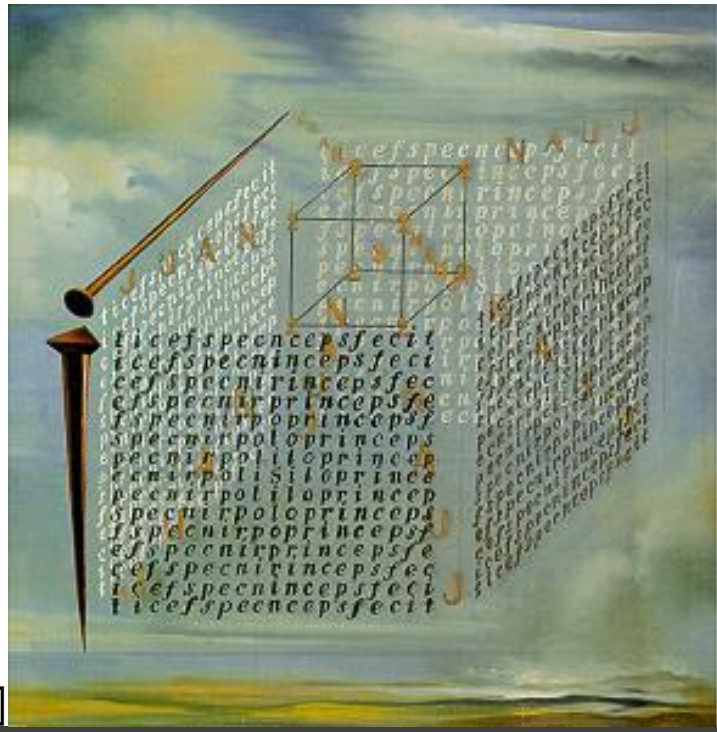
- OSGeo important association, but opportunities for improvement
- Organizational Maturity: Process definition & implementation, QM
- Focus on Core Mission
- Dogmatic „Software Communism“
  - „all software free“ - why?
  - Large companies don‘t care, small companies vulnerable

  → Need inclusive approach



Conclusion

- OGC Coverages for pixel-level interoperability: CIS & WCS
  - robust, scalable, mature
  - consensus: OGC + ISO + INSPIRE std bodies, major tools & vendors
- rasdaman community: scalable array engine for spatio-temporal regular & irregular grids
  - OGC, INSPIRE reference implementation
  - blueprint for datacube standards
- See us:
  - www.earthserver.eu
  - www.rasdaman.org
  - www.jacobs-university.de/lsis



[Dali]