OGC Big Data Standards:

WCS, WCPS

FOSS4G-Europe, Bremen, 2014-07-14

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

Co-funded by
the European Union

Array DB Research @ Jacobs U

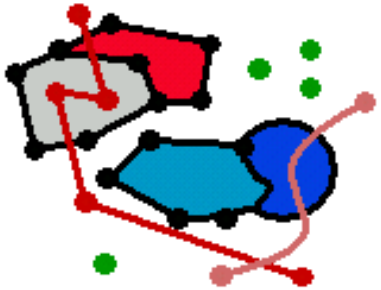
- Large-Scale Scientific Information Systems group
  - massive n-D array services
  - www.jacobs-university.de/lsis
- Main impact:
  - pioneer Array DBMS, rasdaman
  - Standardization: OGC WCS suite editor, ISO Array SQL

ISO: member, SC32 / WG3 SQL; SC32 Big Data Study Group;
OGC liaison, TC211
Open Geospatial Consortium: co-chair, BigData.DWG,
WCS.SWG, Coverages.DWG; co-founder, Temporal.DWG
Research Data Alliance: co-chair, Big Data Interest Group
and Geospatial Interest Group
Charter Member, OSGeo
member, ERCIM Expert Group Big Data
member, Belmont Forum, WP 3 Harmonization of global environmental data infrastructure
council member, CGI / IUGS
founding member and secretary, CODATA Germany
...

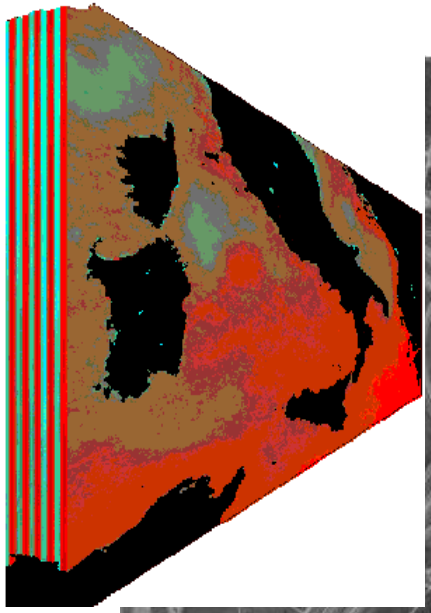
Overview

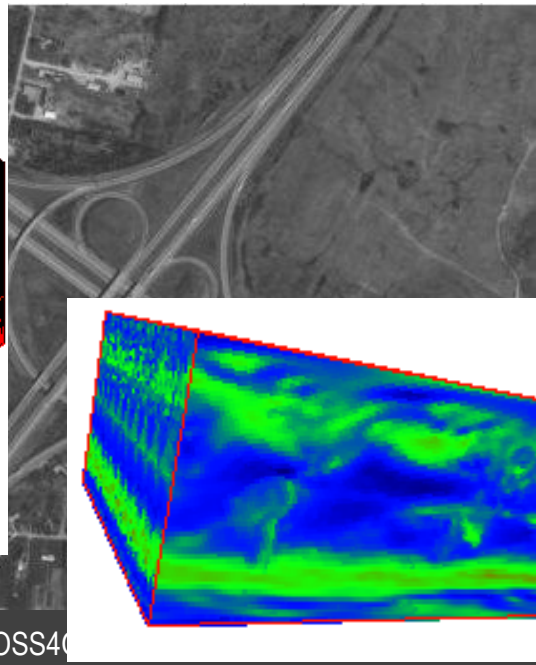
- OGC coverage data model
- OGC WCS service model
- OGC WCPS service for Agile Analytics

Geo Data, as per OGC & ISO



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





rasdaman raster data management

JACOBS UNIVERSITY

Big Data Variety: Coverages

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

«FeatureType» *Abstract Coverage*

Grid Coverage

Rectified GridCoverage

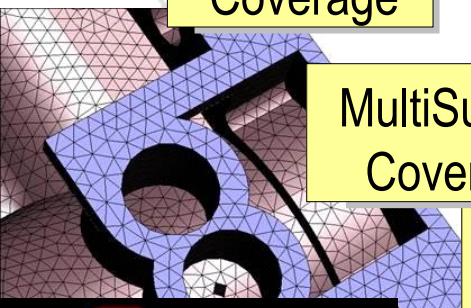
Referenceable GridCoverage

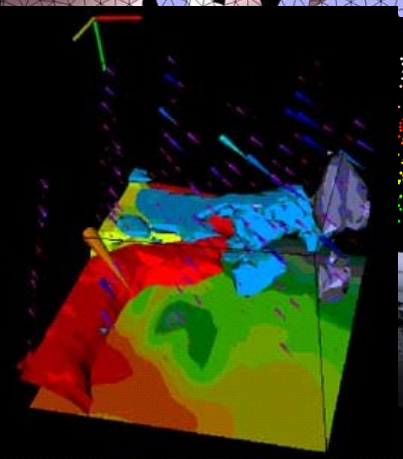
MultiSolid Coverage

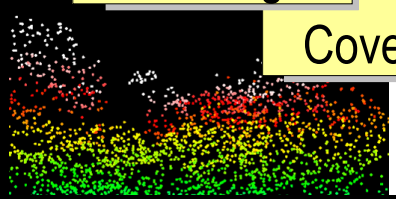
MultiSurface Coverage

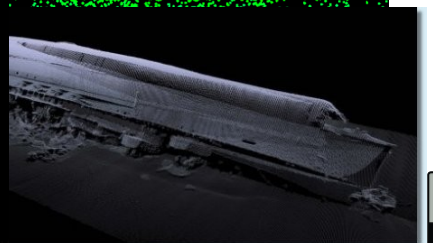
MultiCurve Coverage

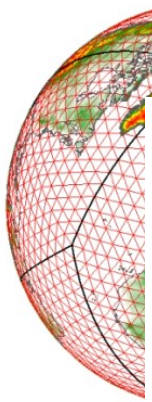
ltiPoint Coverage

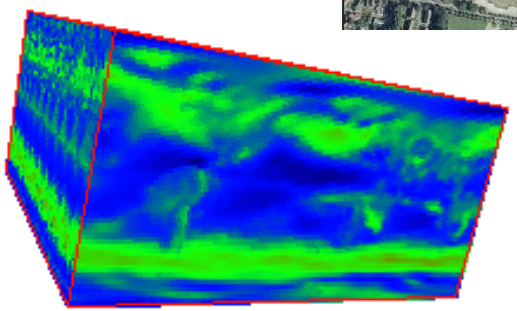


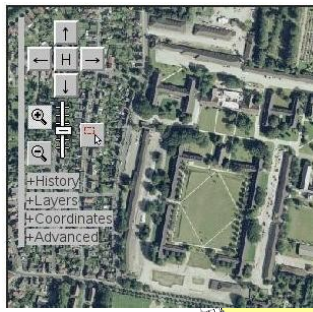




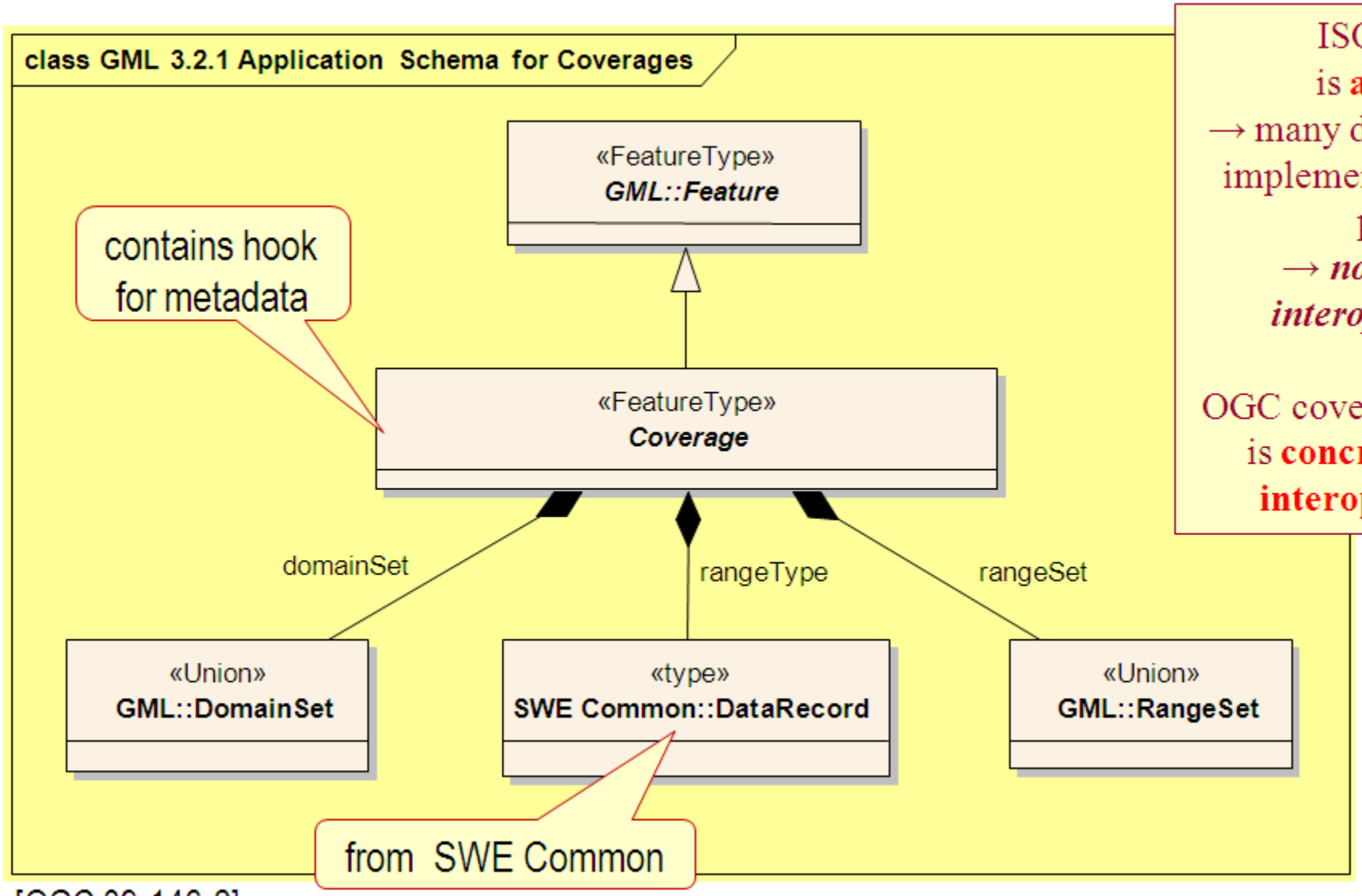








Coverage Data Structure



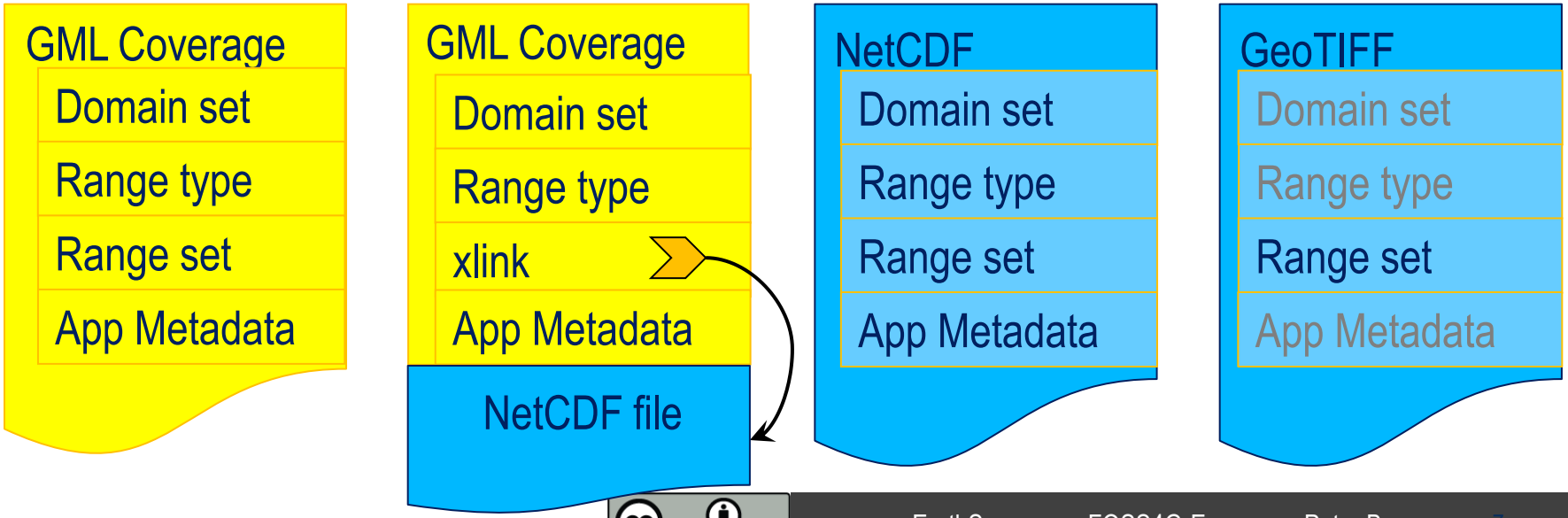
ISO 19123 is **abstract** → many different implementations possible → ***not per se interoperable***

OGC coverage std is **concrete** and **interoperable**

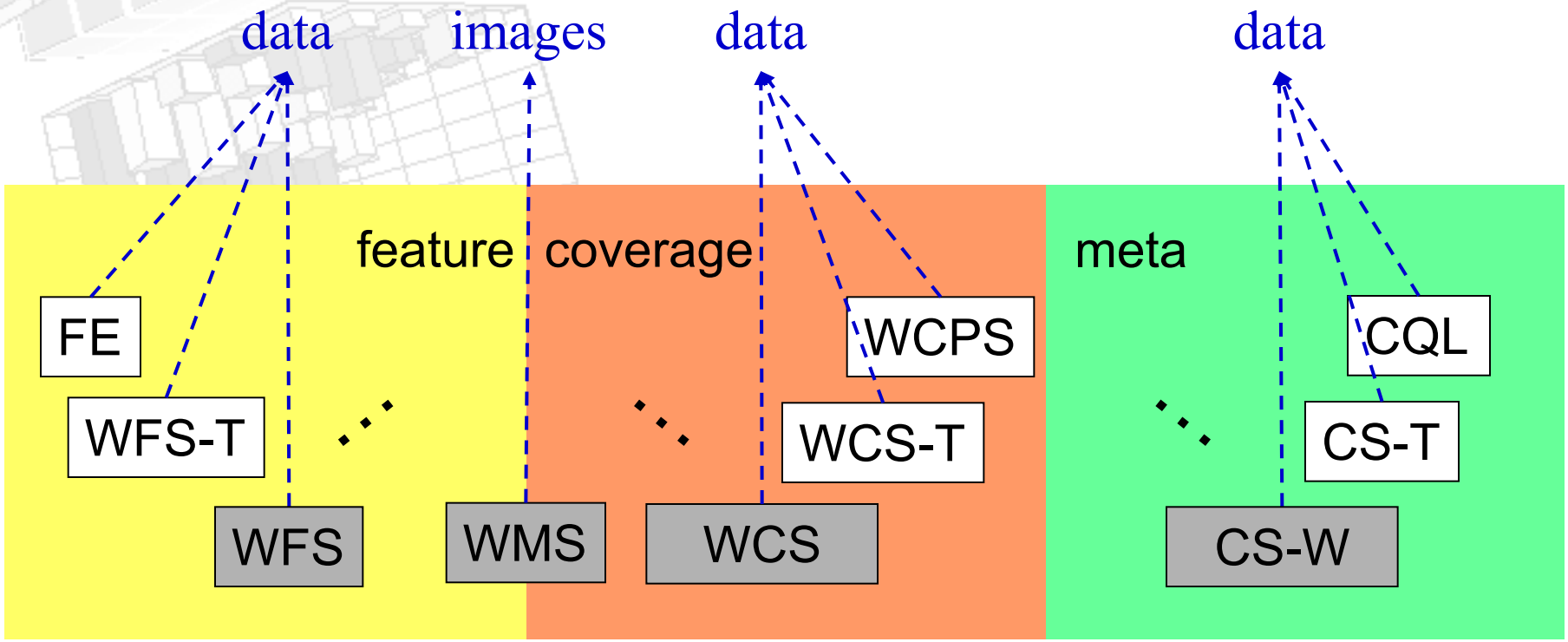
[OGC 09-146r2]

Coverage Encoding

- Pure GML: complete coverage represented by GML
- Special Format: other suitable file format (ex: MIME type “image/tiff”)
  - Usually specific to particular data sets (dimensions!); may lose some metadata
- Multipart-Mixed: multipart MIME, type “multipart/mixed”
  - GML domain set + range type + xlink reference to file in same package
    - *depending on format chosen this may repeat some domain set & range type info*



(Part of) The OGC Standards Quilt

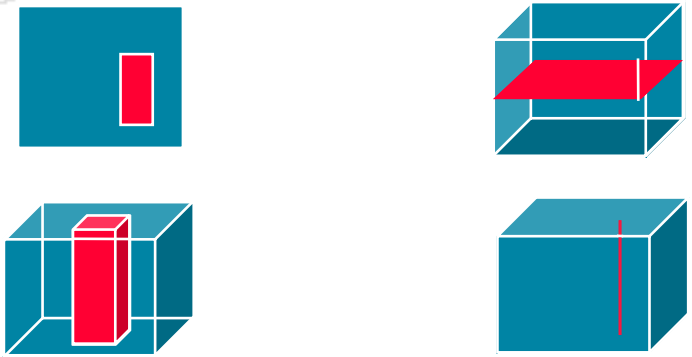


- WMS "portrays spatial data → pictures"
- WCS: "provides data + descriptions; data with original semantics, may be interpreted, extrapolated, etc.“

[OGC 09-110r3]

Web Coverage Service (WCS)

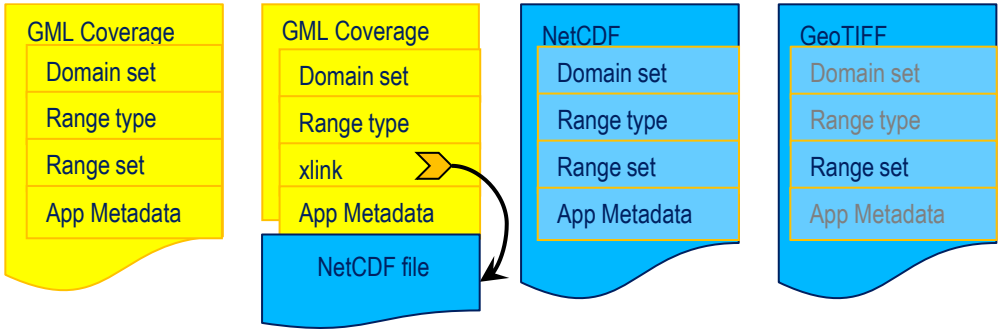
- Core: Simple & efficient access to n-D spatio-temporal coverages
  - plus format encoding
- subset = trim | slice



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

Encoding Coverages

- Remember coverage encoding:
  - pure GML
  - Special-format
  - Multipart



- *GetCoverage* returns Native Format by default, or:

  Ex: FORMAT=application/gml+xml

  FORMAT=image/tiff & MEDIATYPE=multipart/related

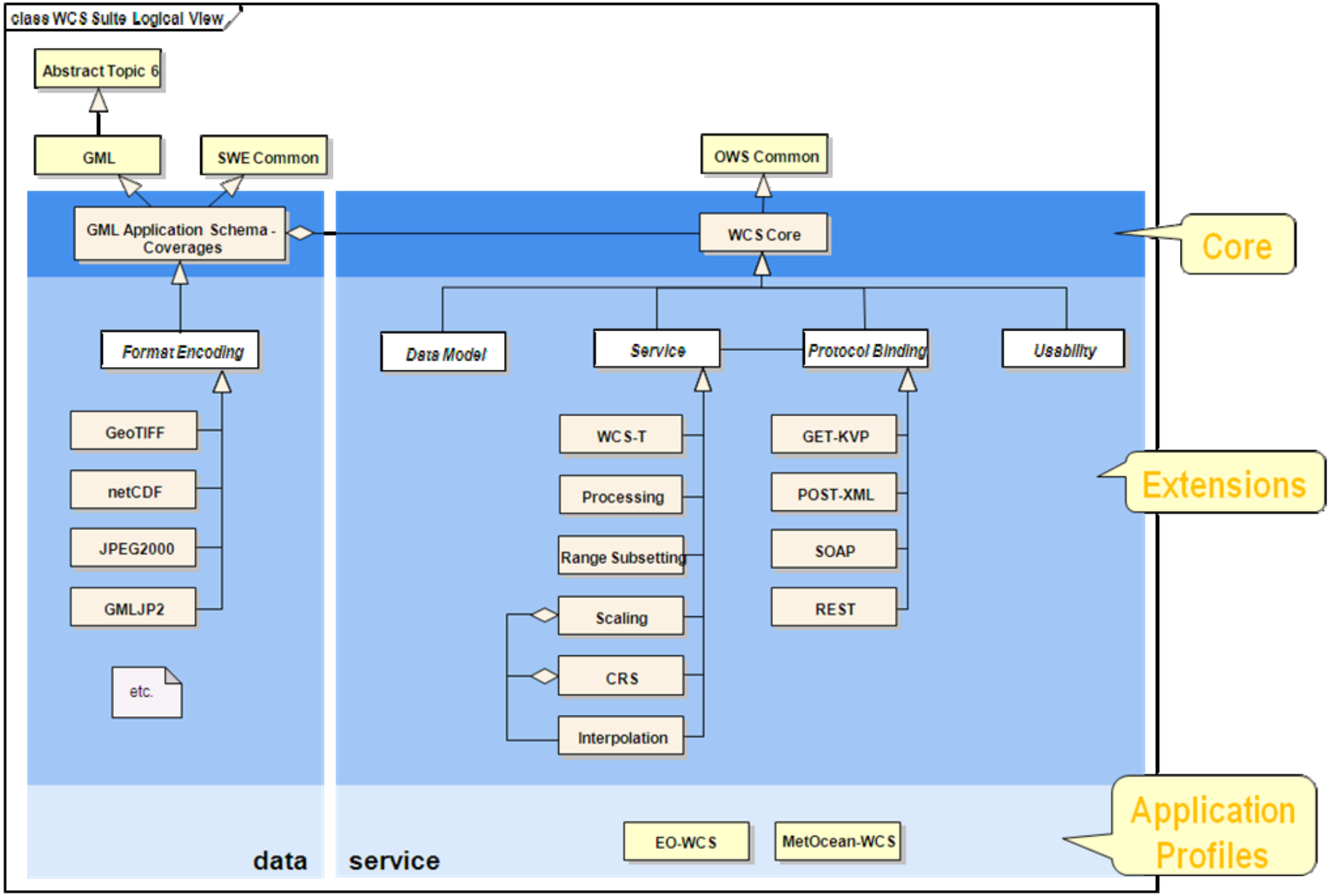
  FORMAT=image/tiff

- What formats are supported?

  GetCapabilities: service metadata report formats available

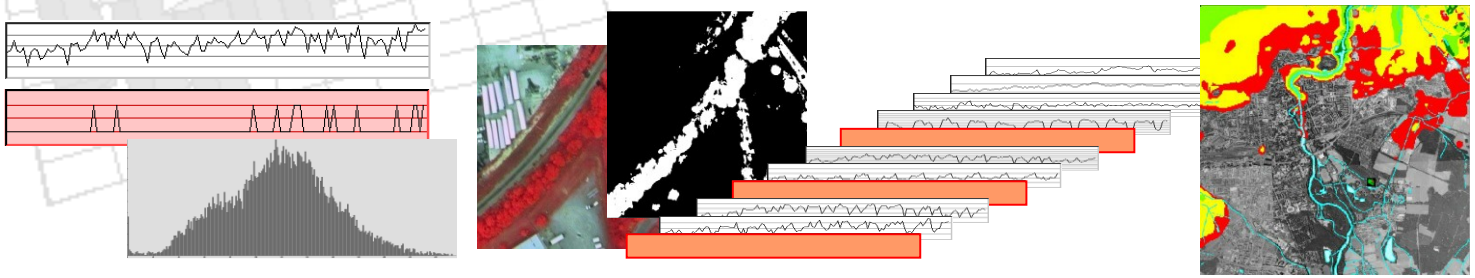
  - Ex: <wcs:Format>image/tiff</Format>

WCS: The Big Picture



OGC WCPS

- 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"
    )
```

→ $(\text{tiff}_A, \text{tiff}_C)$

OGC WCPS 2.0: outlook

- WCPS++ integrated with XQuery
  - more flexible & efficient than OGC *GetCapabilities*
- "Identifiers of all coverages offered"

```
/CoverageOfferings/OfferedCoverage/coverage/@id
```

- "All formats supported by this server"

```
/CoverageOfferings/ServiceIdentification/ServiceMetadata/
  formatSupported/text()
```

- "spatial extent of coverage X"

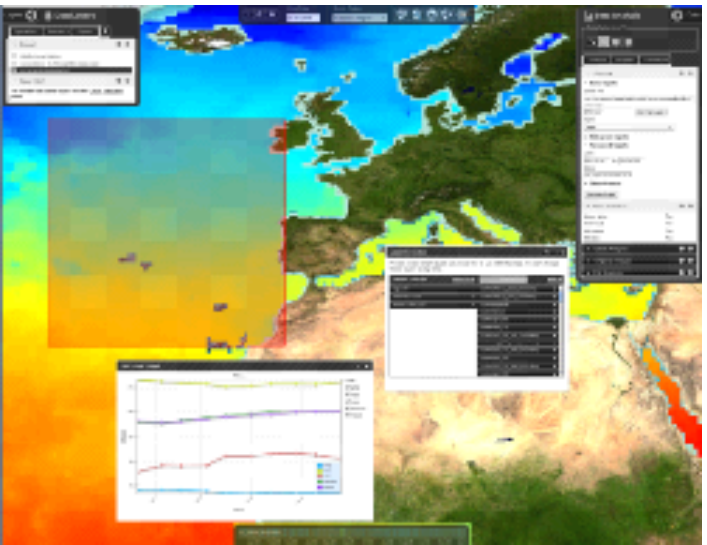
```
/CoverageOfferings/OfferedCoverage/coverage[@id="X"]/domainSet
```

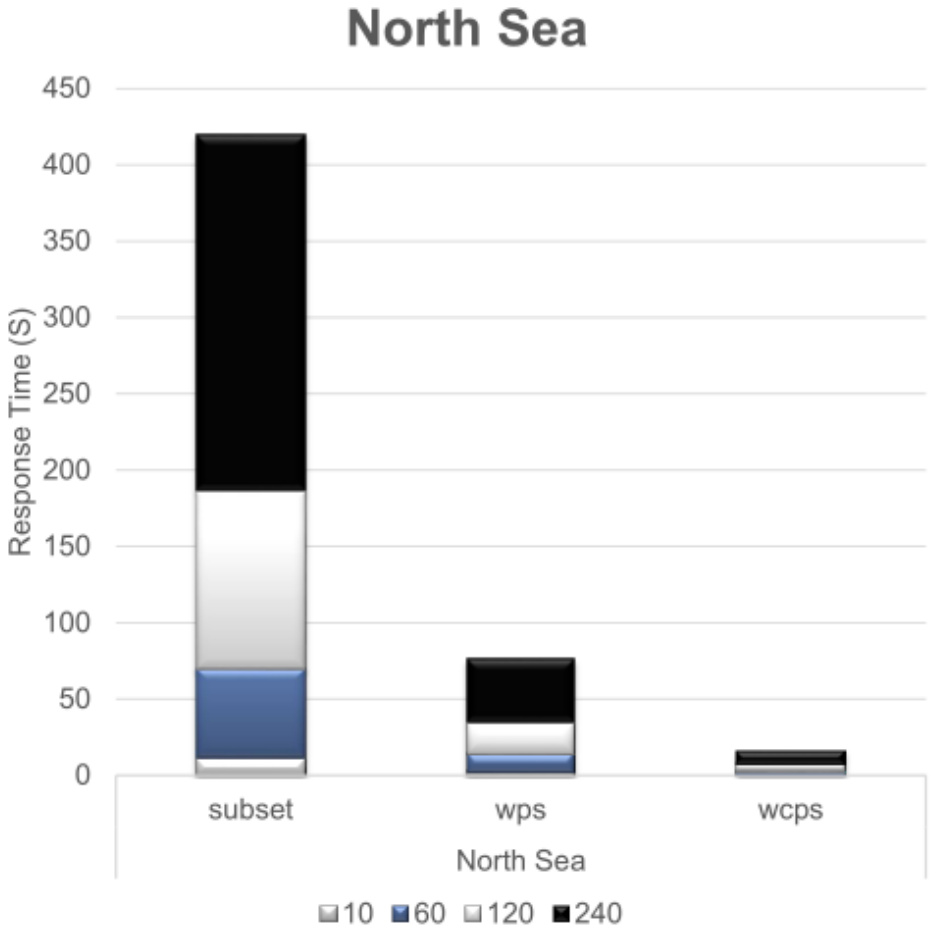
- "...and its pixel values"

```
/CoverageOfferings/OfferedCoverage/coverage[@id="X"]/rangeSet
```

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" )
```

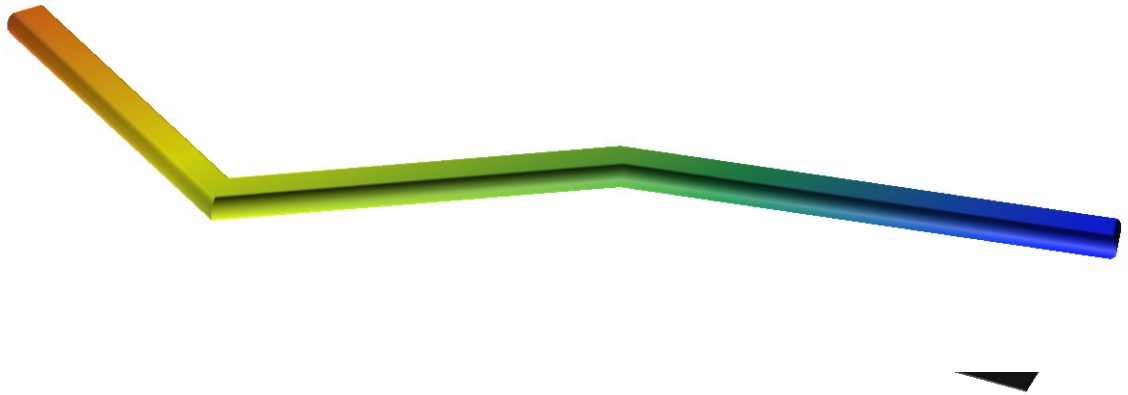
- WPS: semantics in human-readable text

```
<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	WCPS	WPS
data access	ad-hoc retrieval	predefined process

Outlook: MetOcean-WCS

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



Wrap-Up

- OGC standards foster interoperability, vendor independence
- WCS for simple, versatile data access
- WCPS for spatio-temporal Big Data Analytics
  - Flexibility + scalability + information integration
  - pictures → actionable data
- OGC WCS suite implementation proven
  - rasdaman (WCPS, WCS Core Reference Implementation)
    GeoServer, MapServer, (ESRI working), ...
  - OGC WCS & WCPS