rasdaman raster data management

JACOBS UNIVERSITY

Towards Comprehensive Big Data Support: the EarthServer Project

Findability Challenge Workshop
Taormina, 2012-may-10

Peter Baumann

Jacobs University | rasdaman GmbH

EU FP7-INFRA
EarthServer
grant #283610

Big Data Research @ Jacobs U

- Jacobs University: international, multi-cultural
  - 110 nations, english official language
- Large-Scale Scientific Information Systems research group
  - large-scale n-D raster services & beyond: theory, practice, application, standardization
  - Main outcomes:
    - rasdaman: n-D Array DBMS
    - Standards
  - www.jacobs-university.de/lsis
- Hiring

EU FP7-INFRA
EarthServer
grant #283610

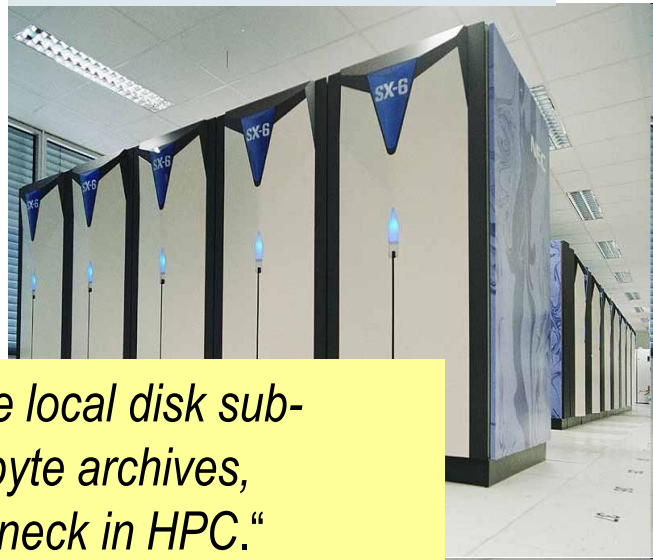
Roadmap

- Motivation
- Arrays in Databases
- EarthServer
- Conclusion

Climate Modelling

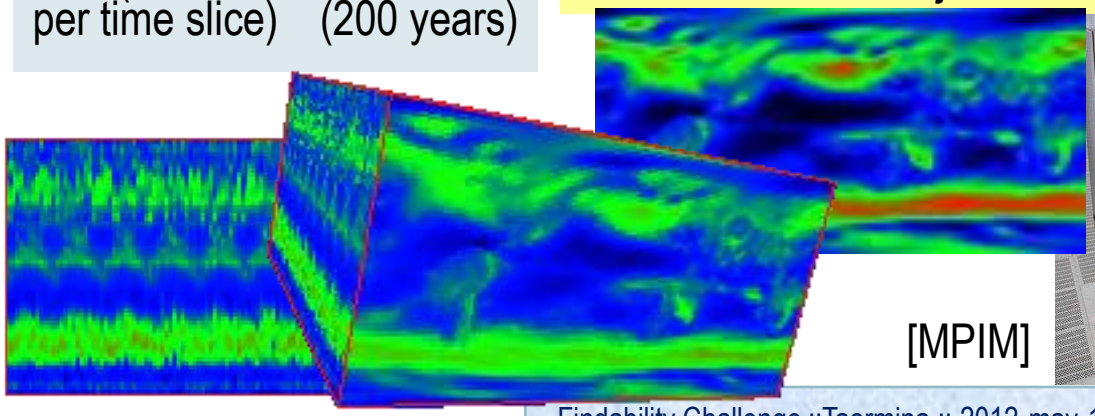
- Example: ECHAM T42
  - 50+ physical parameters („variables“): temperature, wind speed x/y, humidity, pressure, CO2, ...
  - 2.5 TB per variable

DKRZ: 24-node NEC SX-6



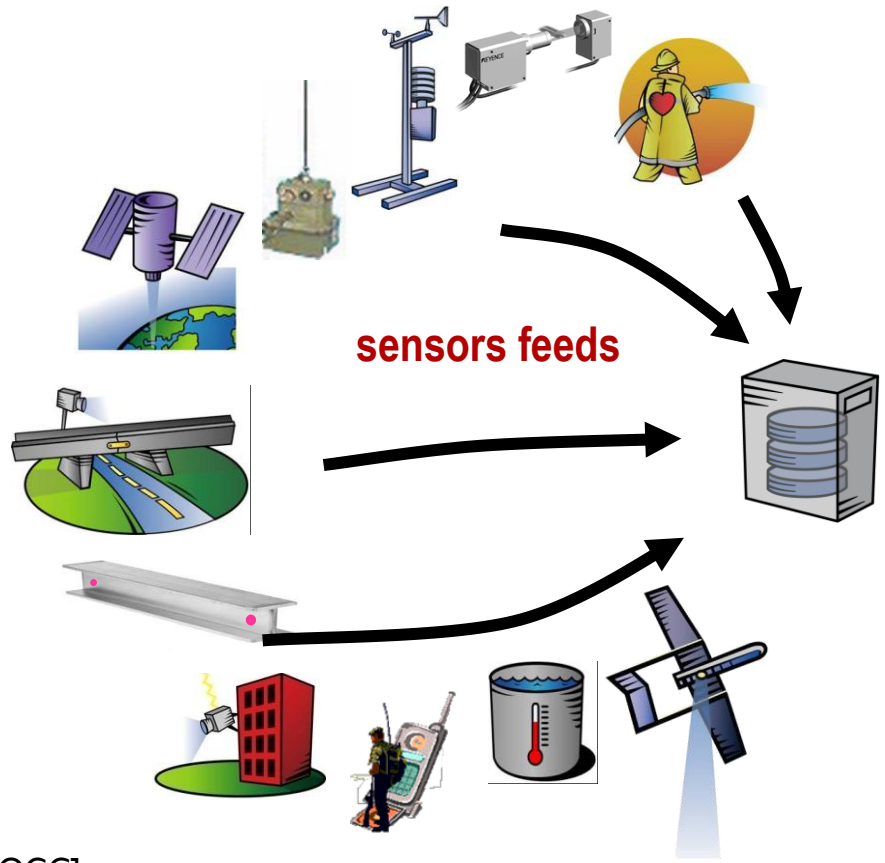
dimension	extent
Longitude	128
Latitude	64
Elevation	17
time (24 min per time slice)	2,190,000 (200 years)

„Even with multi-terabyte local disk sub-systems and multi-petabyte archives, I/O can become a bottleneck in HPC.“
-- Jeanette Jenness,
LLNL, ASCI-Project, 1998



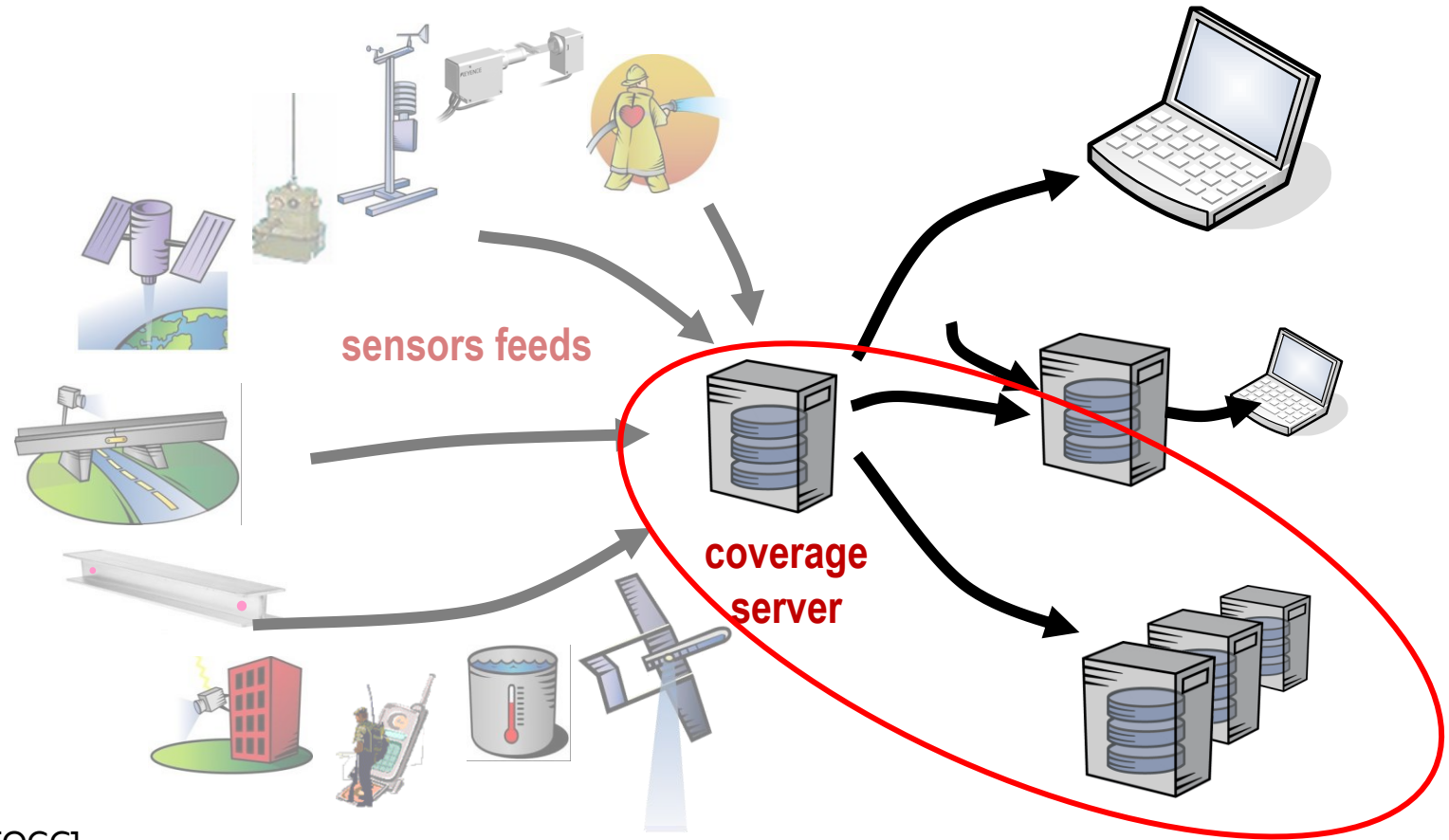
[MPIM]

The Geo Data Tsunami



[OGC]

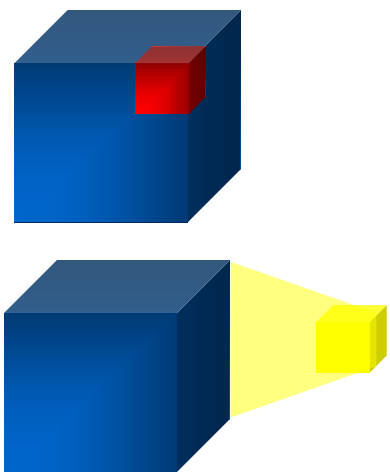
The Geo Data Tsunami

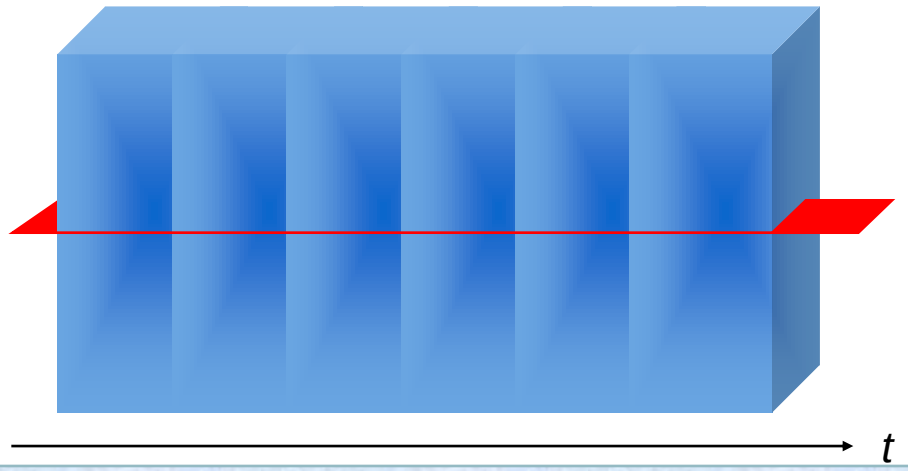


[OGC]

The Challenge for Data Centers

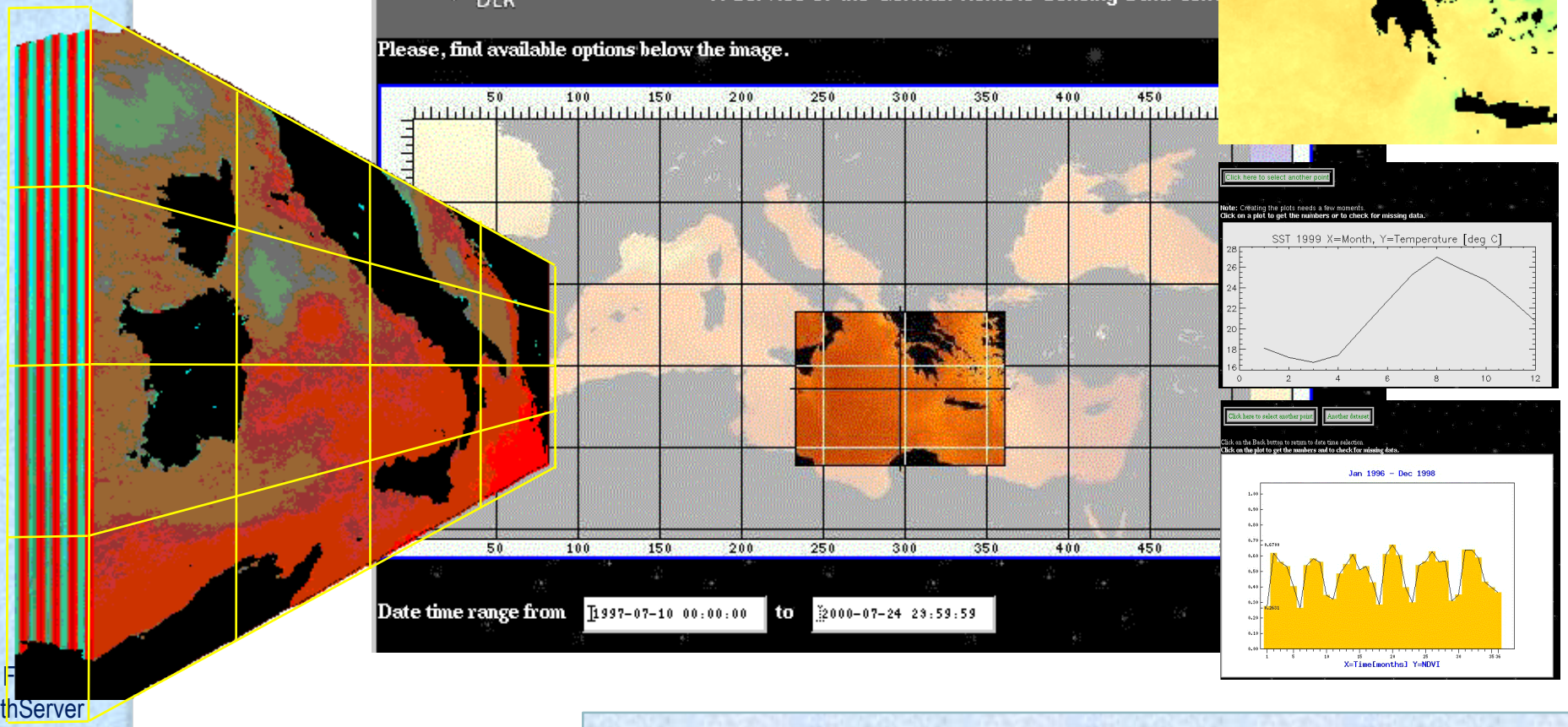
- Serving *data* is not enough
- trend: service quality as differentiating criterion
- transition from data stewardship to service stewardship
- Specifically with high-volume data, what can this mean?





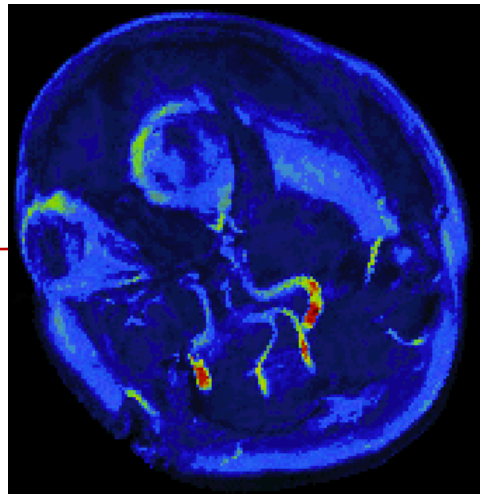
Motivational Example: EO Image Time Series Archive

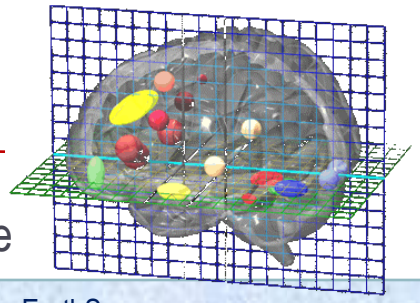
[DFD-DLR, Diederich et al]



Human Brain Imaging

- goal: understand structural-functional relations in human brain
- Experiments capture activity patterns (PET, fMRI)
  - Temperature, electrical, oxygen consumption, ...
  - → lots of computations → „activation maps“
- Example: “*a parasagittal view of all scans containing critical Hippocampus activations, TIFF-coded*.“

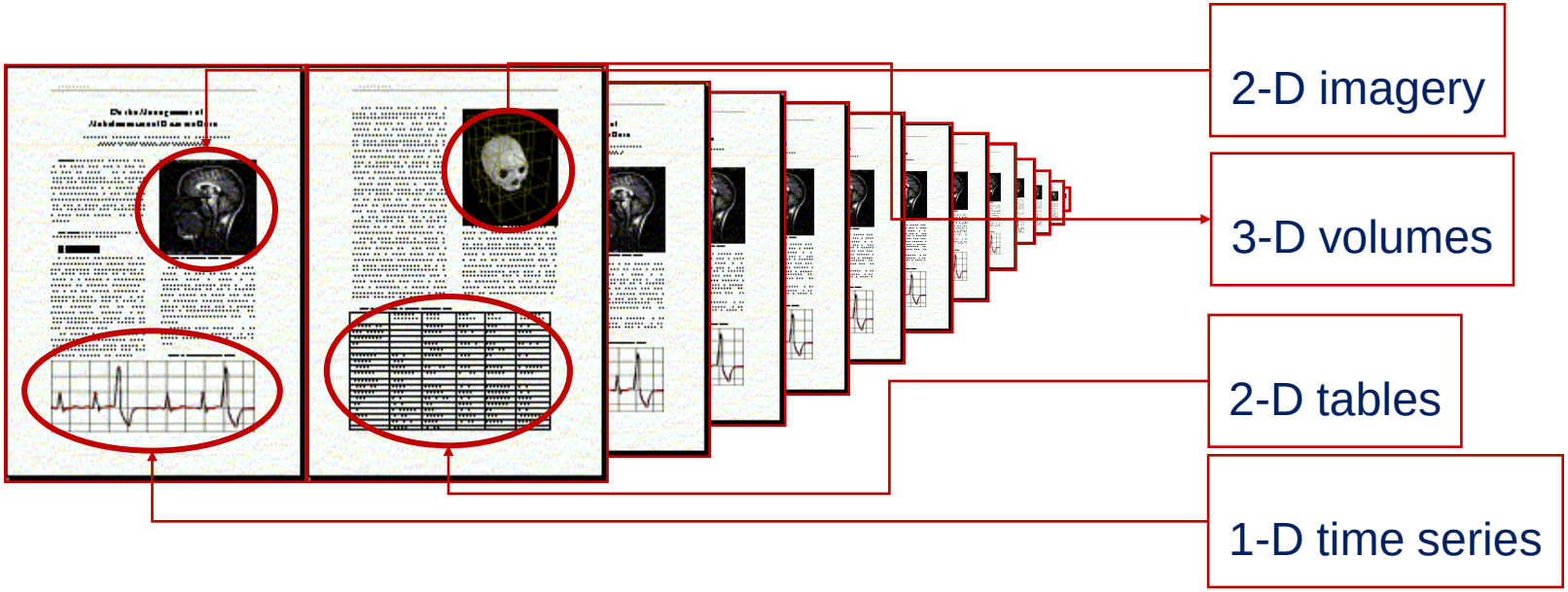




```
select tiff( ht[ $1, *:*, *:* ] )
from   HeadTomograms as ht,
       Hippocampus as mask
where  count_cells( ht > $2 and mask )
       / count_cells( mask )
       > $3
```

`$1` = slicing position, `$2` = intensity threshold value, `$3` = confidence

It's Not Always About Big Return: Reverse Lookup Scenario



„all clinical trials of drug X
where patient temperature > 40° C
within the first 48 hours."

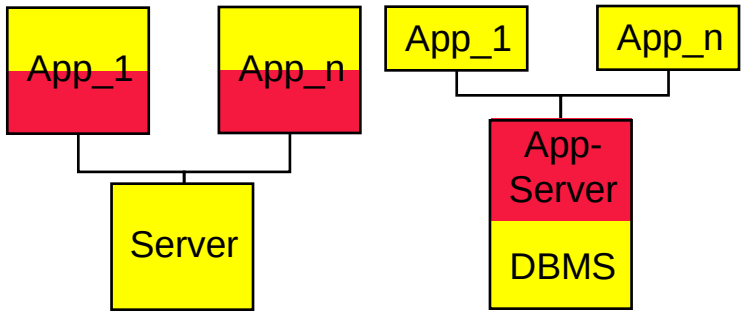
Who Needs Array Databases?

Sensor, image, statistics data:

- **Life Science:** Pharma/chem, healthcare / bio research, bio statistics, genetics
- **Geo:** Geodesy, geology, hydrology, oceanography, meteorology, earth system research, ...
- **Space:** optical astronomy, radio astronomy, cosmological simulation
- **Engineering & research:** Simulation & experimental data in automotive/shipbuilding/ aerospace industry, turbines, process industry, astronomy, experimental physics, high energy physics, ...
- **Management/Controlling:** Decision Support, OLAP, Data Warehousing, census, statistics in industry and public administration, ...

Why Array *Databases*?

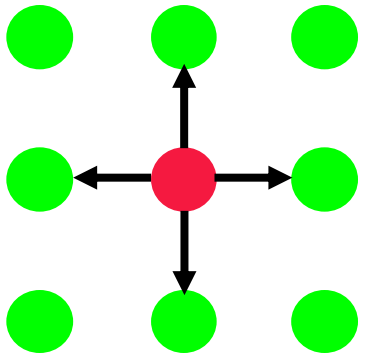
- "classical“ database benefits for raster data:
  - data integration
  - flexibility
  - scalability
  - *...plus all further assets, like off-the-shelf tool support*



- Unfortunately database people have been soooo conservative
  - "images are matrices [...] which are stored as byte strings, ie, BLOBs"
  - „this is NOT SQL!“

Array Analytics

- Array Analytics := Efficient analysis on multi-dimensional arrays of a size several orders of magnitude above main memory of evaluation engine
  - Array := n-D sequence (cf programming languages) := raster
  - Cf 1st Workshop on Array Databases, Uppsala 2011 www.rasdaman.com/ArrayDatabases_Workshop
- Earth, Space, Life, Social sciences; business (OLAP!)
- Research issues:
  - Concepts: modeling, access interfaces (QLs!), ...
  - Architecture: storage, processing, optimization, ...
  - Scalability, usability, applications, standards, ...



Related (DB) Work

- Precursors
  - Image partitioning, API access library [Tamura 1980]
  - Fixed set of imaging operators [Chang, Fu 1980; Stucky, Menzi 1989; Neumann et al 1992]
  - PICDMS [Chock, Cardenas 1984]
- Algebra & models
  - rasdaman model & algebra [Baumann 1991]
  - „Call to order“ [Maier 1993]
  - AQL [Libkin,Machlin 1996]
  - AML [Marathe, Salem 2002]
- Components & systems
  - rasdaman [Baumann+ 1992+]
  - tertiary storage for arrays [Sarawagi, Stonebraker 1994]
  - ESRI ArcSDE, Oracle GeoRaster [~2004]
  - TerraLib [Camara et al, 2003]
  - MonetDB: RAM, SciQL [~2004]
  - PostGIS Raster [~2007]
  - SciDB [~2008]

The rasdaman Array DBMS

www.rasdaman.org

- C/S Array DBMS for massive n-D raster data
  - new attribute type: array<celltype,extent>
  - In operational use on dozen-TB objects

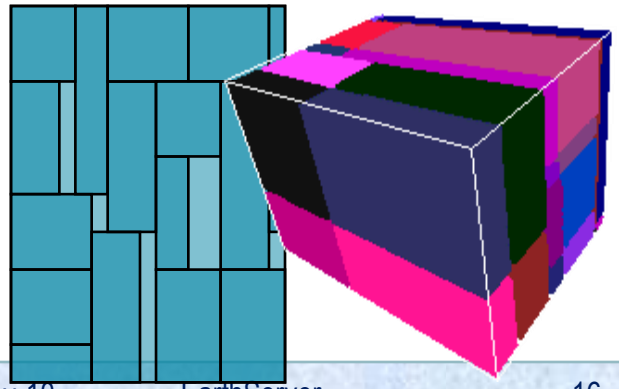
metadata	att 1	att 2	att n
	key1	...	oid 1
	key2	...	oid 2
	key3	...	oid 3

ATable	OID	array
	oid 1	
	oid 2	
	oid 3	
	oid 4	
	oid 5	

- rasql = declarative array query language

```
select img.green[x0:x1,y0:y1] > 130
from   LandsatArchive as img
```

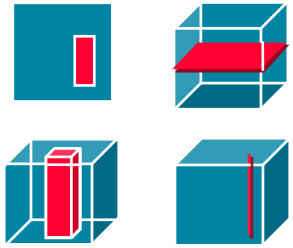
  - algebraic foundation
- Tile-based architecture
  - n-D array → set of n-D tiles → DB blobs
  - evaluation based on “tile streaming”
  - extensive storage & query optimization



Some Array Query Operators

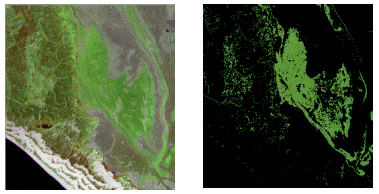
- trimming & slicing

```
select c.img[ *:*, 100:200, *:*, 42 ]
from   ClimateSimulations as c c
```



- result processing

```
select img * (img.green > 130)
from   LandsatArchive as img
```

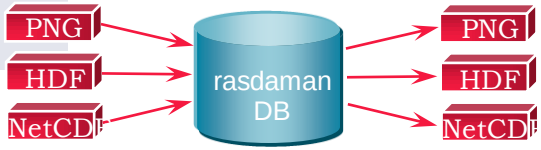


- search & aggregation

```
select mri.img
from   MRI as mri, masks as am
where  some_cells( mri.img > 250 and m.img )
```

- data format conversion

```
select png( c[ *:*, *:*, 100, 42 ] )
from   ClimateSimulations as c
```



Examples

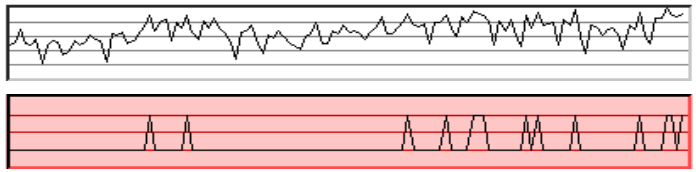
- Histogram
  - select marray bucket in [0:255]
    values count_cells(a = bucket)
    from a, b

- Matrix multiplication
  - select marray x in sdom(a)[0], y in sdom (b)[1]
    values condense +
    over z in sdom(a)[1]
    using a[x,z] * b[z, y])
    from a, b

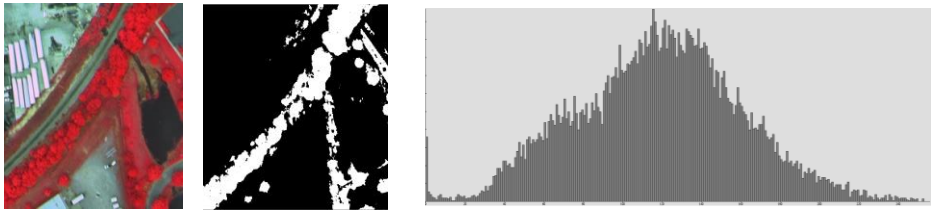
N-D Sensor, Image, and Statistics Queries

„Raster SQL“: navigation, extraction, aggregation, analysis

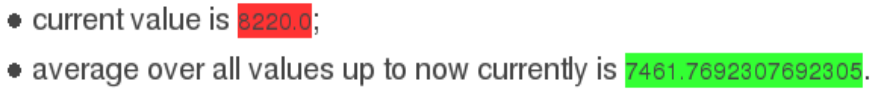
- Time series



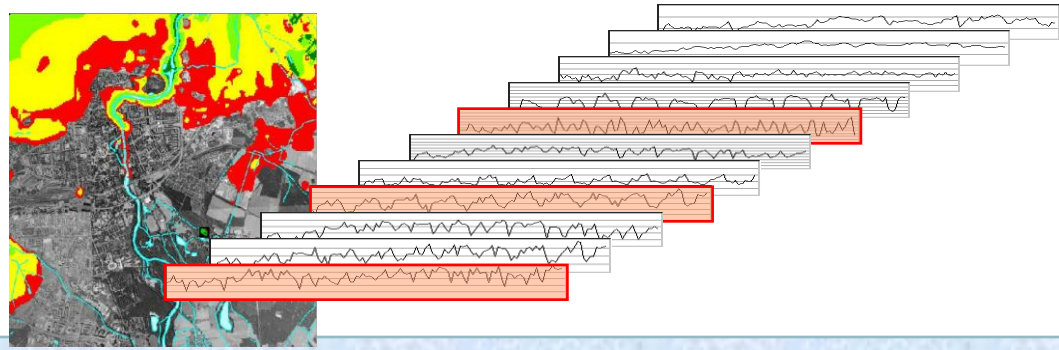
- Image processing



- Summary data



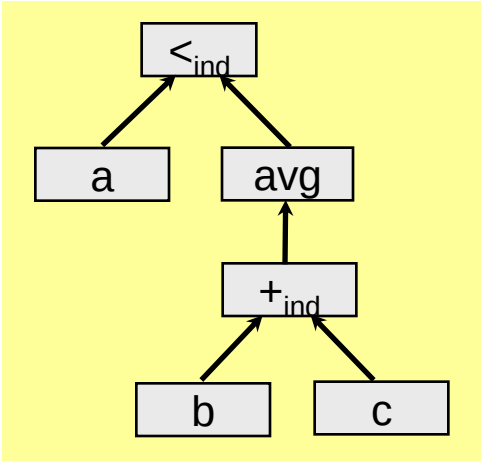
- Sensor fusion & pattern mining



Query Processing: Overview

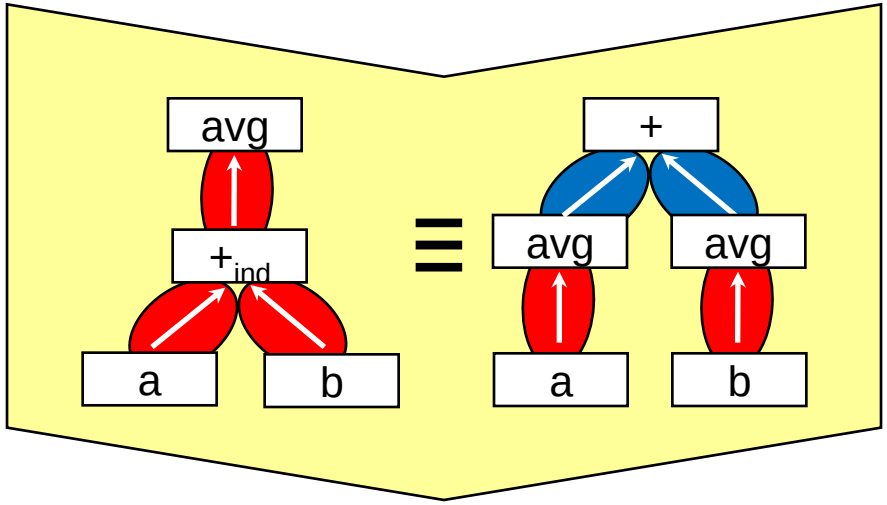
- Parsing
- Normalisation
- Optimization
  - Common subexpression elimination
- [Generate query plan]
- Tile-based evaluation

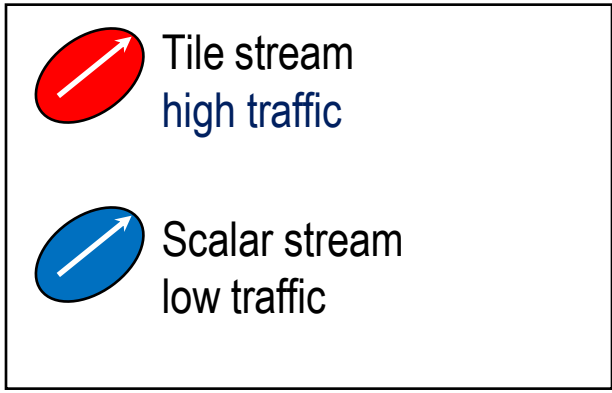
```
select a < avg_cells( b + c )
from   a, b, c
```



Optimization/1: Query Rewriting

```
select avg_cells( a + b )
from   a, b
```





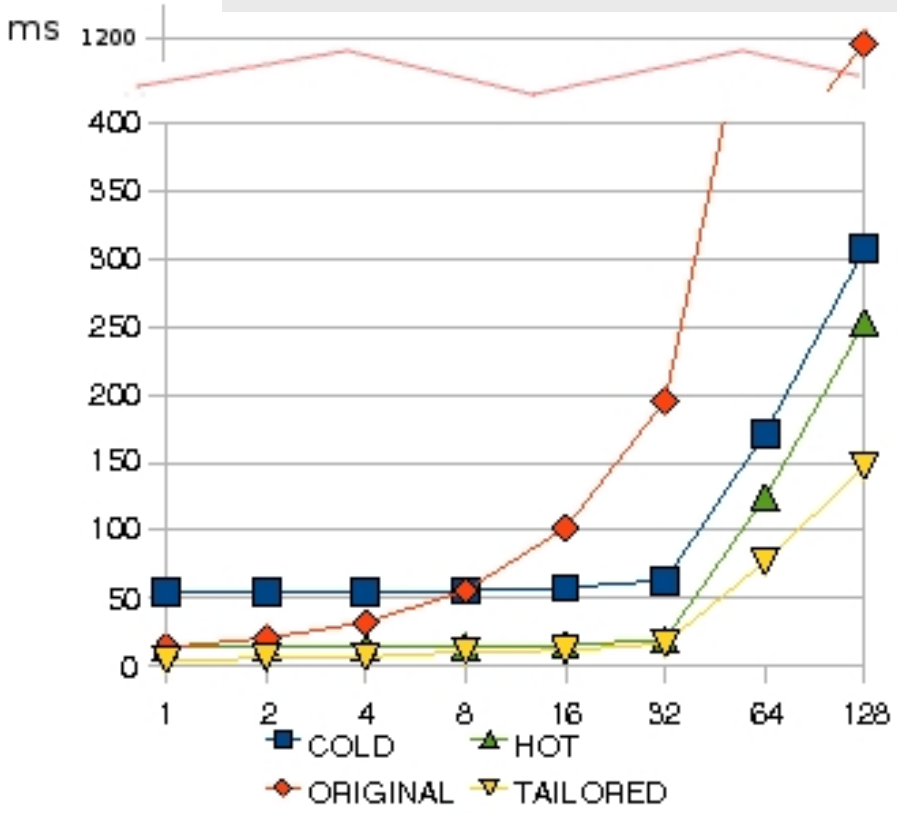

```
select avg_cells( a )
     + avg_cells( b )
from   a, b
```

- *understood:*
  heuristic optimization – 150 rules in rasdaman [Ritsch 2002]
- *partially understood:*
  cost-based optimization

Optimization/2: Just-In-Time Compilation

- Observation: interpreted mode slows down
- Approach:
  - cluster suitable operations
  - compile & dynamically bind
- Benefit:
  - Speed up complex, repeated c
- Variation:
  - compile code for GPU

```
select x*x*...*x
from float_matrix as x
```



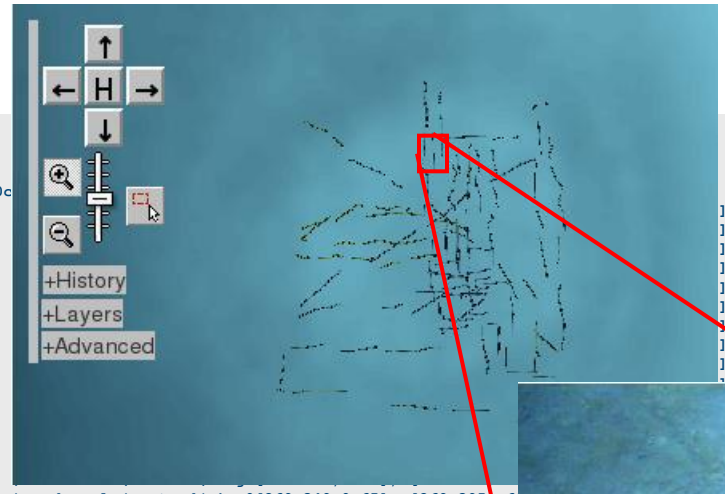
Times [ms] for 512^2 * n ops

[Jucovschi, Stancu-Mara 2008]

Optimization Does Pay Off!

Ex: real-time WMS zoom/pan/styling

- 1 background, 1 bathymetry, 3*RGB
- www.earthlook.org



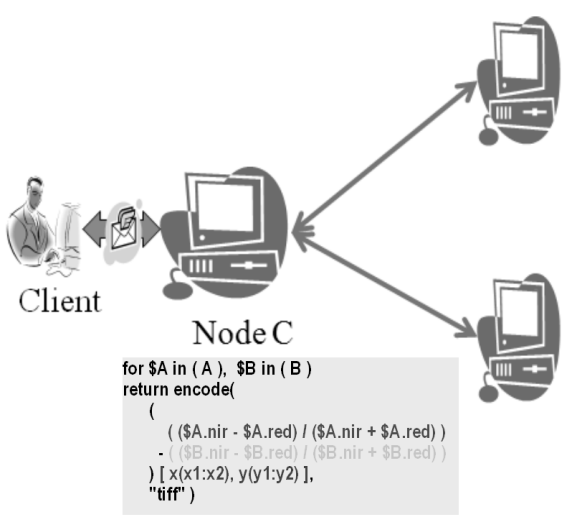
```
SELECT png(
(marray x in [0:399,0:399] values {255c,255c,255c})
overlay
((scale( extend( img0[269:349,0:65], [269:395,-60:65] ), [0:399,0:399] ) < -1300)*{0c
+(-1300.000000< scale( extend( img0[269:349,0:65], [269:395,-60:65] ), [0:399,0:399]
+(-1289.999999< scale( extend( img0[269:349,0:65], [269:395,-60:65] ), [0:399,0:399]
+(-1281.999999< scale( extend( img0[269:349,0:65], [269:395,-60:65] ), [0:399,0:399]
+(-1274.999999< scale( extend( img0[269:349,0:65], [269:395,-60:65] ), [0:399,0:399]
+(-1272.499999< scale( extend( img0[269:349,0:65], [269:395,-60:65] ), [0:399,0:399]
+(-1270.999999< scale( extend( img0[269:349,0:65], [269:395,-60:65] ), [0:399,0:399]
+(-1270.499999< scale( extend( img0[269:349,0:65], [269:395,-60:65] ), [0:399,0:399]
+(-1269.999999< scale( extend( img0[269:349,0:65], [269:395,-60:65] ), [0:399,0:399]
+(-1269.499999< scale( extend( img0[269:349,0:65], [269:395,-60:65] ), [0:399,0:399]
+(-1268.999999< scale( extend( img0[269:349,0:65], [269:395,-60:65] ), [0:399,0:399]
+(-1268.499999< scale( extend( img0[269:349,0:65], [269:395,-60:65] ), [0:399,0:399]
+(-1267.999999< scale( extend( img0[269:349,0:65], [269:395,-60:65] ), [0:399,0:399]
+(-1267.499999< scale( extend( img0[269:349,0:65], [269:395,-60:65] ), [0:399,0:399]
+(-1266.999999< scale( extend( img0[269:349,0:65], [269:395,-60:65] ), [0:399,0:399]
+(-1266.499999< scale( extend( img0[269:349,0:65], [269:395,-60:65] ), [0:399,0:399]
+(-1265.999999< scale( extend( img0[269:349,0:65], [269:395,-60:65] ), [0:399,0:399] ) and scale( extend( img0[269:349,0:65], [269:395,-6
+(-1265.499999< scale( extend( img0[269:349,0:65], [269:395,-60:65] ), [0:399,0:399] ) and scale( extend( img0[269:349,0:65], [269:395,-6
+(-1264.999999< scale( extend( img0[269:349,0:65], [269:395,-60:65] ), [0:399,0:399] ) and scale( extend( img0[269:349,0:65], [269:395,-6
+(-1264.499999< scale( extend( img0[269:349,0:65], [269:395,-60:65] ), [0:399,0:399] ) and scale( extend( img0[269:349,0:65], [269:395,-6
+(-1263.999999< scale( extend( img0[269:349,0:65], [269:395,-60:65] ), [0:399,0:399] ) and scale( extend( img0[269:349,0:65], [269:395,-6
+(-1263.499999< scale( extend( img0[269:349,0:65], [269:395,-60:65] ), [0:399,0:399] ) and scale( extend( img0[269:349,0:65], [269:395,-6
+(-1262.999999< scale( extend( img0[269:349,0:65], [269:395,-60:65] ), [0:399,0:399] ) and scale( extend( img0[269:349,0:65], [269:395,-6
+(-1261.999999< scale( extend( img0[269:349,0:65], [269:395,-60:65] ), [0:399,0:399] ) and scale( extend( img0[269:349,0:65], [269:395,-6
+(-1260.999999< scale( extend( img0[269:349,0:65], [269:395,-60:65] ), [0:399,0:399] ) and scale( extend( img0[269:349,0:65], [269:395,-6
+(-1259.999999< scale( extend( img0[269:349,0:65], [269:395,-60:65] ), [0:399,0:399] ) and scale( extend( img0[269:349,0:65], [269:395,-6
+(-1256.999999< scale( extend( img0[269:349,0:65], [269:395,-60:65] ), [0:399,0:399] ) and scale( extend( img0[269:349,0:65], [269:395,-6
+(-1249.999999< scale( extend( img0[269:349,0:65], [269:395,-60:65] ), [0:399,0:399] ) and scale( extend( img0[269:349,0:65], [269:395,-6
+(-1239.999999< scale( extend( img0[269:349,0:65], [269:395,-60:65] ), [0:399,0:399] ) and scale( extend( img0[269:349,0:65], [269:395,-6
+(-1229.999999< scale( extend( img0[269:349,0:65], [269:395,-60:65] ), [0:399,0:399] ) and scale( extend( img0[269:349,0:65], [269:395,-6
+ (-126.5 < scale( extend( img0[269:349,0:65], [269:395,-60:65] ), [0:399,0:399] ))*{255c,255c,255c})
overlay (scale( extend( img2[124:468,0:578], [124:717,-14:578] ), [0:399,0:399] ))
overlay (scale( extend( img3[11375:11578,0:120], [11375:11968,-473:120] ), [0:399,0:399] )) )
FROM Hakon_Bathy AS img0, Hakoon_Dive1_8 AS img1, Hakoon_Dive2_8 AS img2, Hakoon_Dive2b_8 AS img3
```

EarthServer: *Big Earth Data Analytics*

- **Mission:** to enable standards-based ad-hoc analytics on the Web for Earth science data
  - scalable to Petabyte/Exabyte volumes
  - directly manipulate, analyze & remix any-size geospatial data
- **Core idea:** integrated query language for all spatio-temporal coverage data
  - standards based ↳
  - Server + clients
- **Starting point:** the rasdaman Array DBMS
- EU FP7-INFRA: operational services by end of project ↳

EarthServer Technical Approach

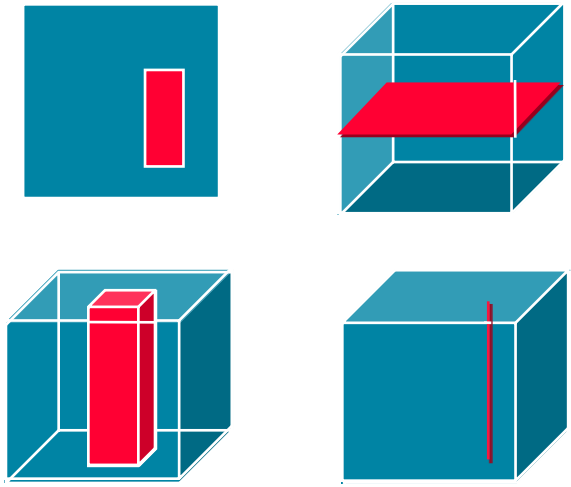
- Starting point: rasdaman
- Core DB Challenges:
  - query distribution
  - cloud virtualization
  - In-situ databases
- Clients, clients, clients!
  - from mobile devices
    to Web GIS
    to high-end immersive VR
  - Challenge: how to map query paradigm to user interaction?
- standards
  - WCS & WCPS + XQuery



Standards Perspective: Open Geospatial Consortium

- "coverages" based on ISO 19123, GML [OGC 09-146]
- Web Coverage Service (WCS) [OGC 09-110] Core:

subset = trim | slice



- further functionality in extension standards
- Web Coverage Processing Service (WCPS) [OGC 08-068r2]
  - Declarative geo raster processing language

WCPS By Example

- "From MODIS scenes M1, M2, and M3, the absolute of the difference between red and nir, in HDF-EOS"

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

(hdf_A, hdf_B, hdf_C)

WCPS By Example

- "From MODIS scenes M1, M2, and M3, the absolute of the difference between red and nir, in HDF-EOS"
  - …but only those where nir exceeds 127 somewhere

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

→ (hdf_A, hdf_C)

WCPS By Example

- "From MODIS scenes M1, M2, and M3, the absolute of the difference between red and nir, in HDF-EOS"
  - …but only those where nir exceeds 127 somewhere
  - …inside region R

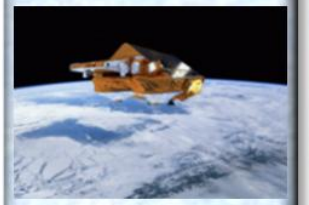
```
for $c in ( M1, M2, M3 ),
    $r in ( R )
where
    some( $c.nir > 127 and $r )
return
    encode(
       abs( $c.red - $c.nir ),
       "hdf"
    )
```


(hdf_A)

EarthServer Lighthouse Applications

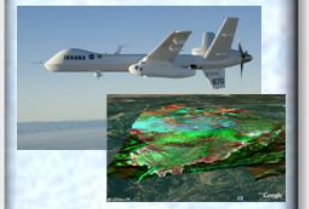
- Each 100+ TB ultimately
- front-end to existing archives - no new archives

Cryospheric Science
landcover mapping



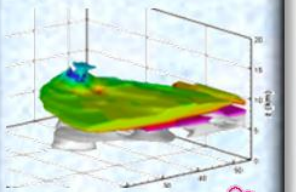
EOX

Airborne Science
high-altitude long-endurance drones



NASA

Atmospheric Science
climate variables

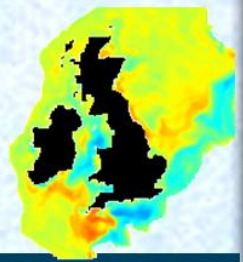


MEEO
Meteorological Environmental Earth Observation

Geology
geological models

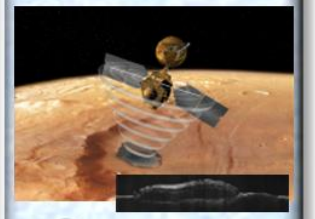
British Geological Survey
NATURAL ENVIRONMENT RESEARCH COUNCIL

Oceanography
marine model runs + in-situ data



PML PLYMOUTH MARINE LABORATORY

Planetary Science
Mars geology



JACOBS UNIVERSITY

Partners

Activity leaders:

JACOBS UNIVERSITY

EOX

Consiglio Nazionale delle Ricerche

Fraunhofer

rasdaman raster data manager

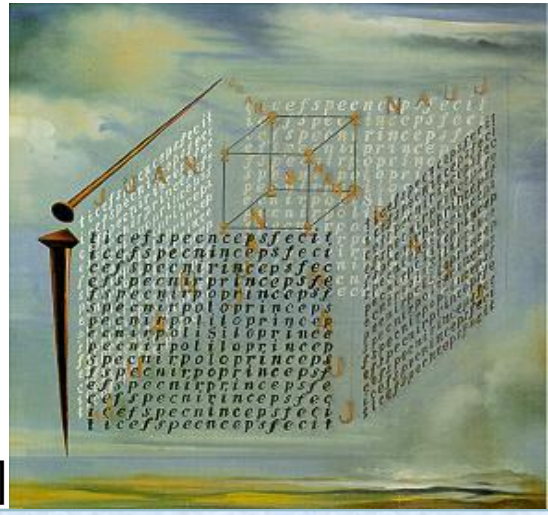
PML PLYMOUTH MARINE LABORATORY

MEEO Meteorological Environmental Earth Observation

BGS 1835 British Geological Survey NATURAL ENVIRONMENT RESEARCH COUNCIL

Conclusion

- Core Issues in Earth science information services:
  - Ad-hoc and „long-tail“ analysis, best near-realtime
  - Unified handling of data & metadata
    - Large array support with „processing & filtering“
  - Fusion: different data types, different locations; database or in-situ
  - Scalability: query complexity, data complexity, data volume, users
- EarthServer vision: „mix & match“multi-source, any-size geo data
  - integrated data/metadata QL; scalability & query distribution; standards
  - Operational services for all Earth science domains
  - www.earthserver.eu



[Dali]