

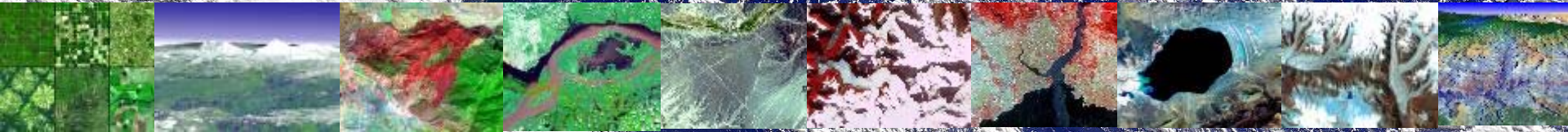
Data Management in GEO Grid



Isao Kojima

**National Institute of Advanced Science
and Technology(AIST),JAPAN**

kojima@ni.aist.go.jp



Geographical Survey
Institute

University of Tsukuba

Mt. Tsukuba

GEO Grid: An e-Science Project @AIST(National Research Institute of Advanced Industrial Science and Technology), METI Japan

JAXA

Japan Aerospace
eXploration
Agency

Geological Survey of Japan of AIST

Information Technology Research
Institute of AIST

*Doing collaborative works with
ASIA-Pacific
research institutes*

AIST

What is GEO Grid?

e-Science infrastructure on heterogeneous data archives.



<http://www.geogrid.org>

- *Aims to contribute GEOSS 10years plan*
- **Core archive contents: our satellite sensor data**
 - ASTER satellite images \approx 150TB(1500,000 scenes, y2000->)
 - Now extending to manage (Petabyte-Scale) PALSAR,PRISM etc. (y2006->)
- **Can Federate with other data resources**
 - Satellite sensors: FORMOSAT-II,MODIS,Landsat, etc
 - Heterogeneous GIS data: Geological maps, Sensors.

- **Supports Standard Technology**

- OGC(GIS) and OGF(Grid)

- **Supports VO (Virtual Organization)-VO manager**

based security (VOMS)

- Organize research communities
- Protect important data

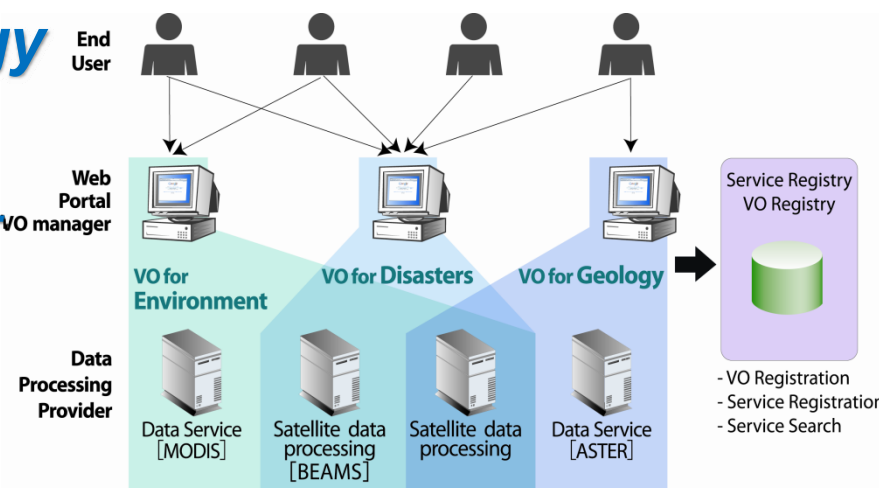
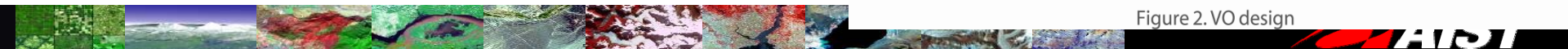
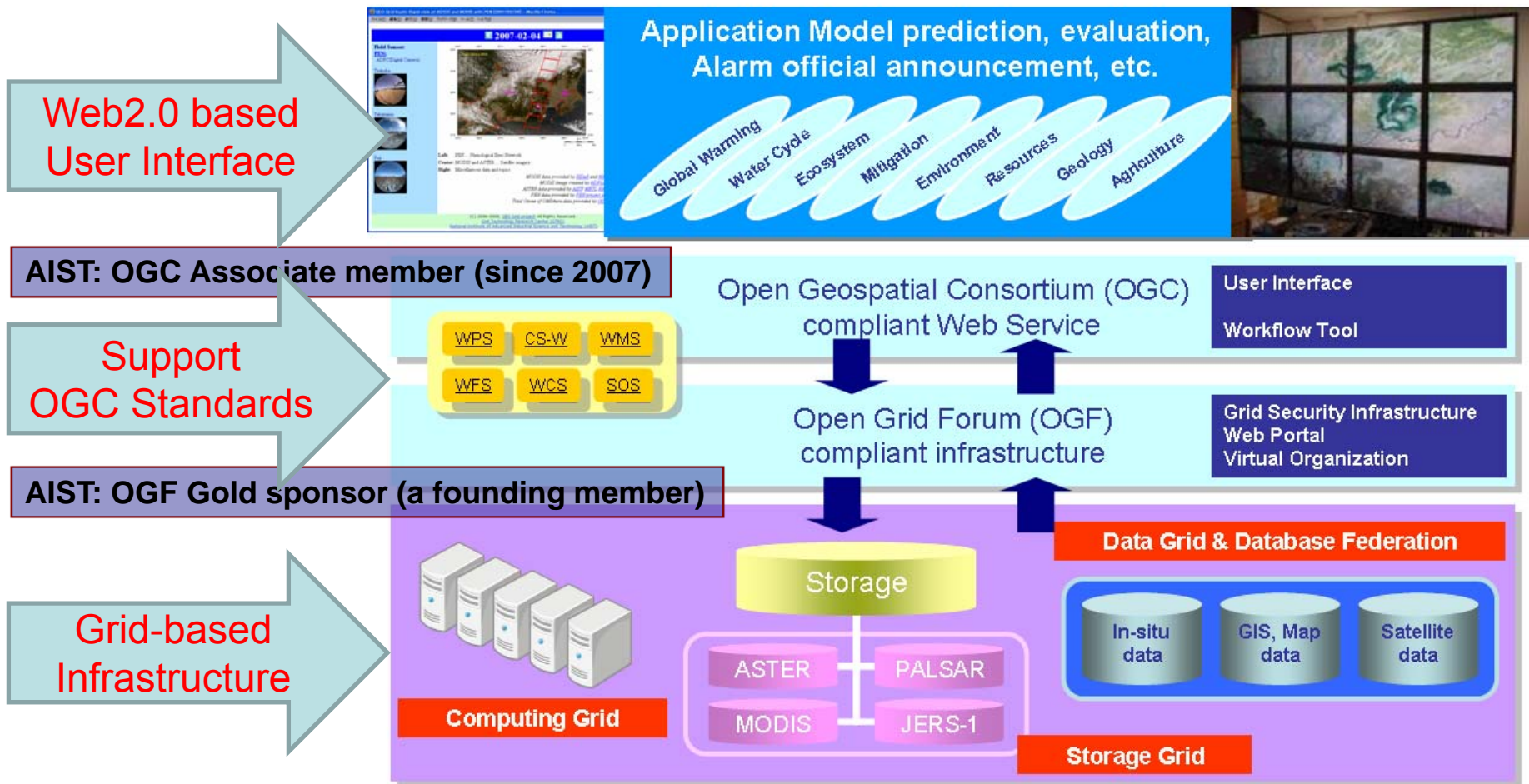


Figure 2. VO design





GEO Grid Architecture & Features





Various Data in GEO Grid

Linked by:

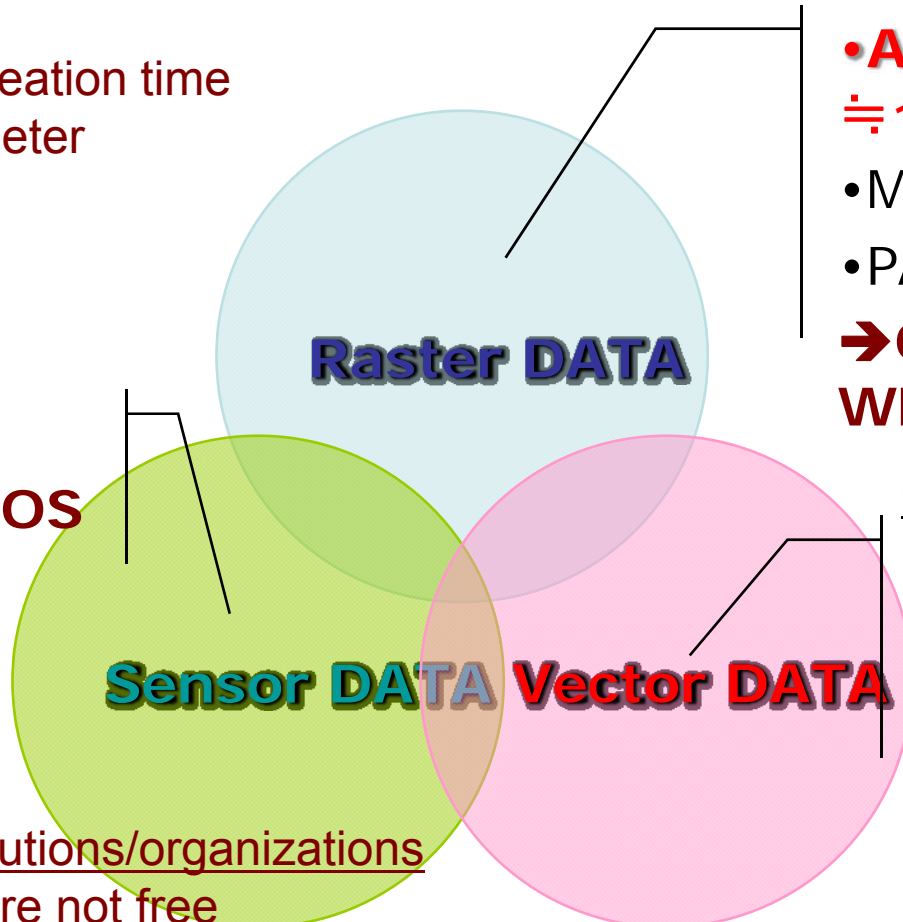
- Geolocation
- Observation/creation time
- Specific parameter

Satellite Images

- **ASTER**
≡ 150TB, 1,500,000 scenes
- MODIS
- PALSAR
- ➔ **GeoTiff/WCS or WMS**

Sensors

- PEN
- ➔ **SensorML/SOS**



GIS Data

- Geological maps
- Roads, railways, etc
- ➔ **GML/WFS**

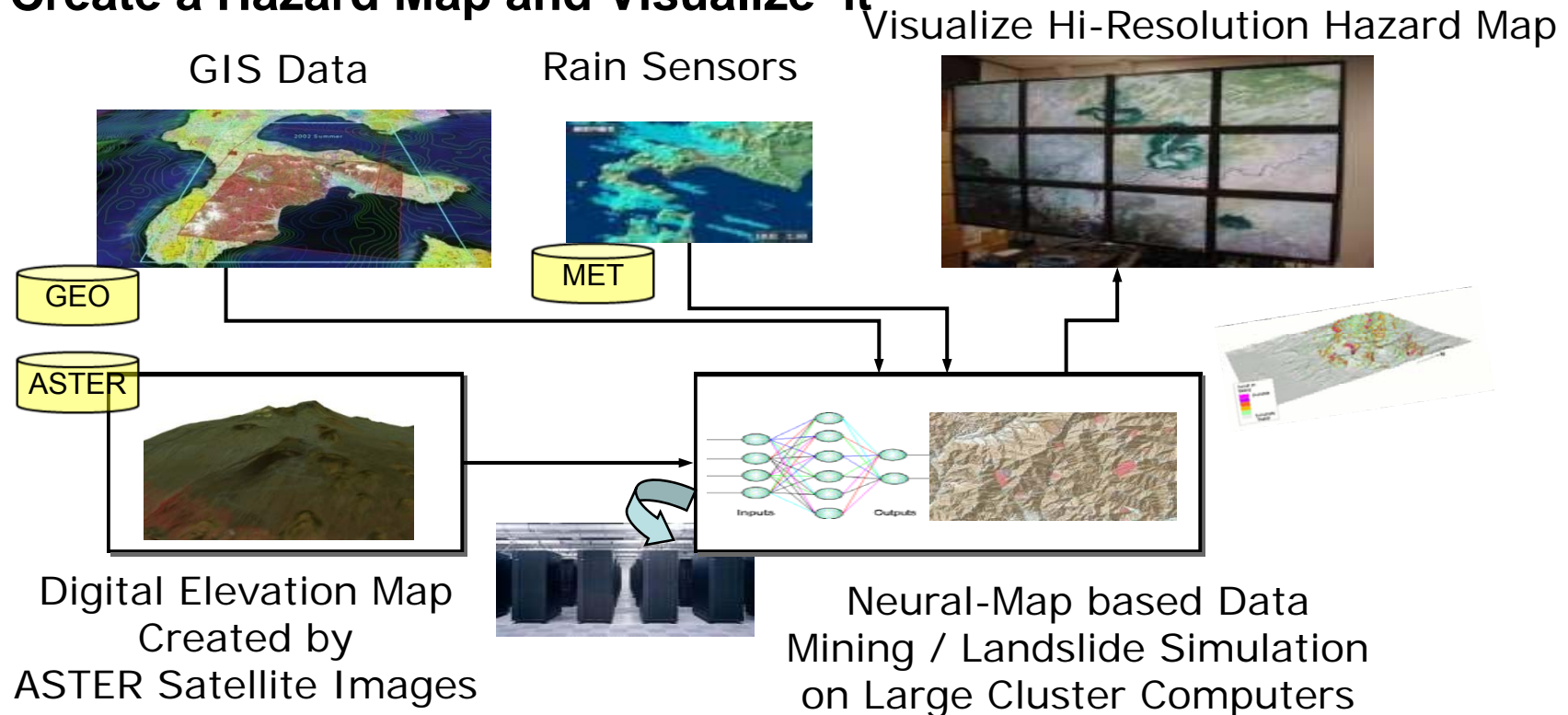
- Owned by institutions/organizations
- Many of them are not free



An Example Application in GEO Grid

Creating a Hazard Map for Landslide

1. Integrate various databases over multiple & distributed organizations
2. Perform large-scale simulation using the data
3. Create a Hazard Map and Visualize it



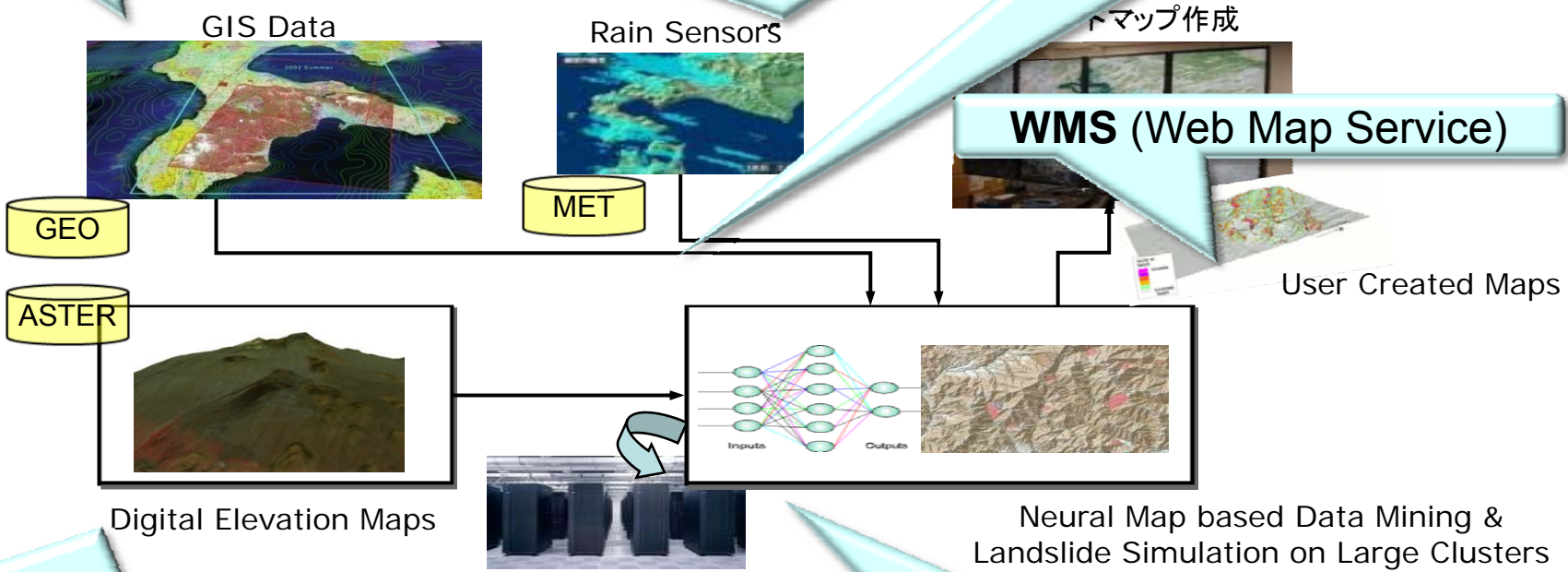
GEO Grid : *VOMS-enabled OGC Services*

- All OGC Services are being VOMS-enabled(GridSite-based solution)
- AIST Original Implementation of CSW

WFS(Web Feature Service)

SOS(Sensor Observation Service)

CSW (Catalog Service Web)
AIST-CSW implementation

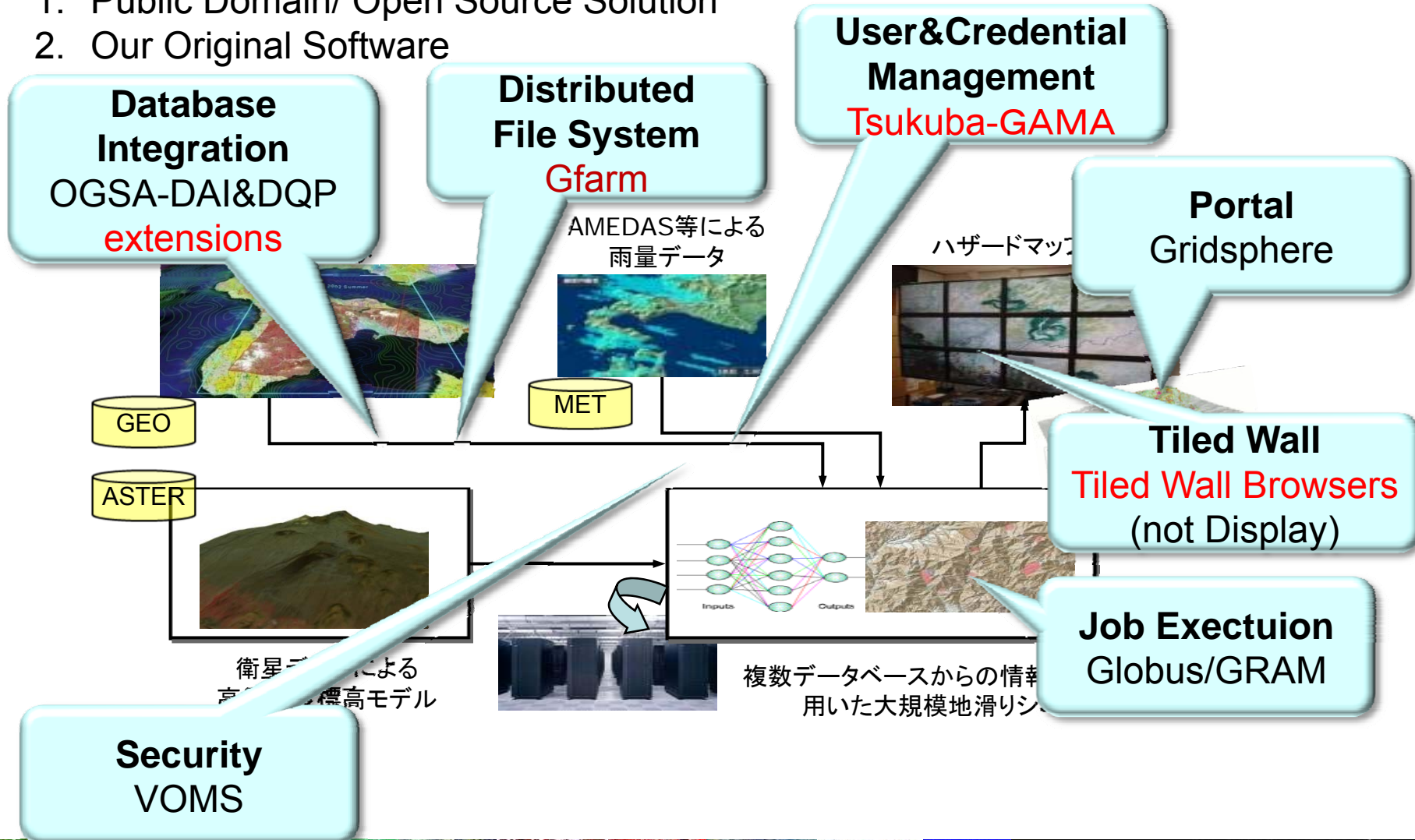


WCS(Web Coverage Service)

WPS (Web Processing Service)

GEO Grid : *Grid Layer*

1. Public Domain/ Open Source Solution
2. Our Original Software



Other Application Examples

Global roads development using WPS (ECO VO)

Field sensor data integration using SOS (FON VO)

Shake amplitude estimation using WPS (GHz VO)

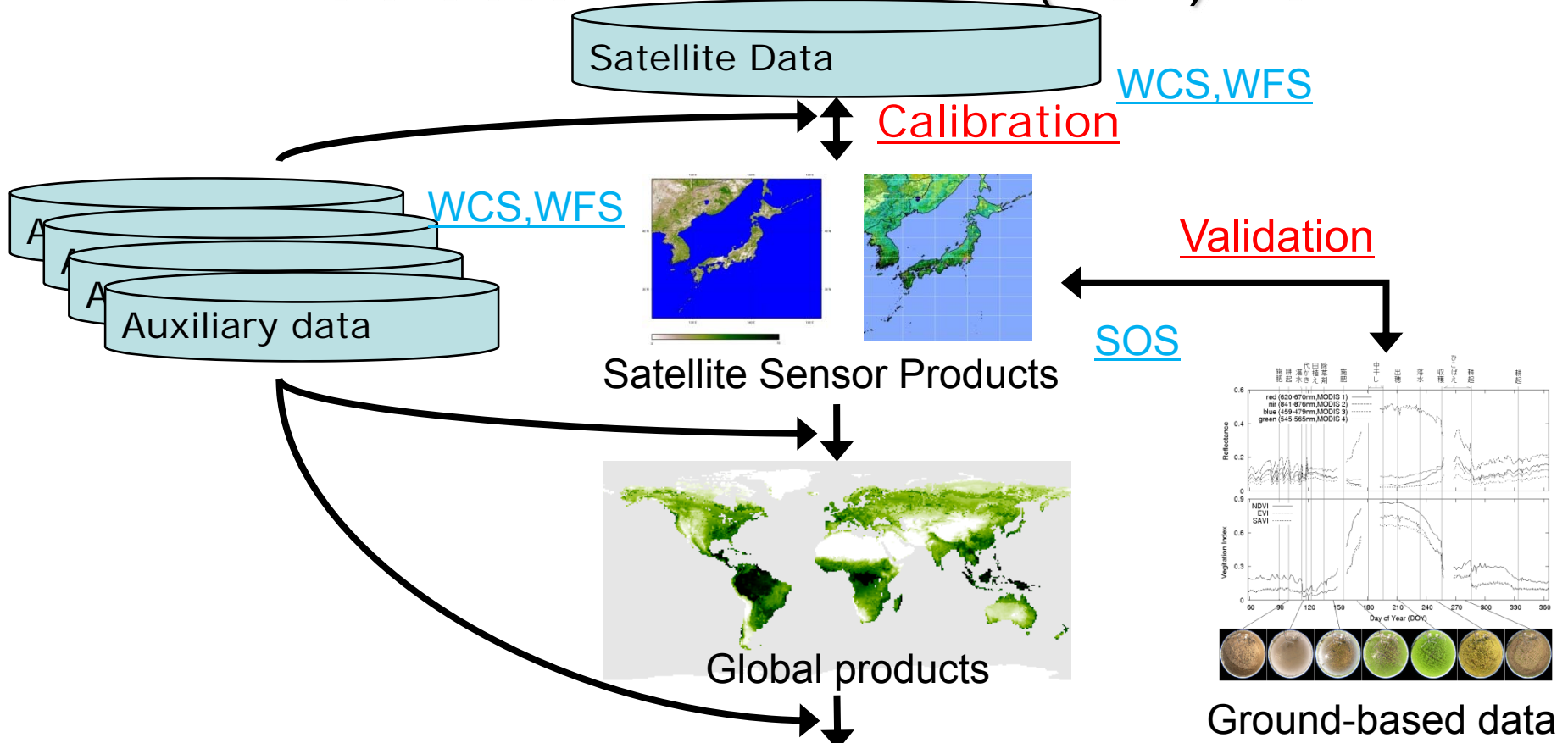
Landslide application using WFS and WCS (GHz VO)

Registry for geological data (GSJ VO)

Applications are constructed with OGC services and associated with specific VOs



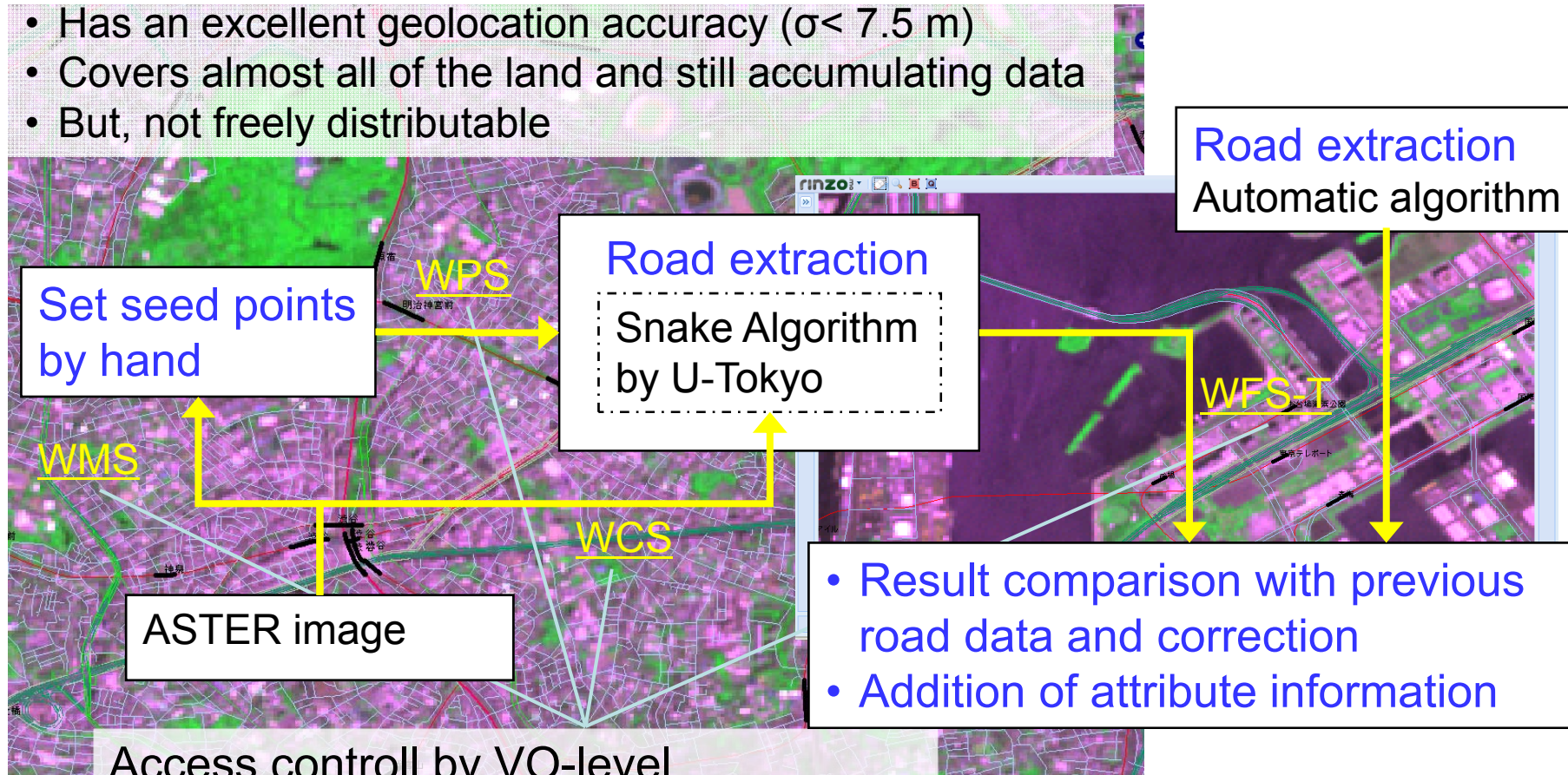
Field sensor data integration at Field Observation Network (FON) VO



Global carbon cycle
Climate change predictions
Evaluation of global warming
etc...

ASTER image processing to detect global roads development

- Has high spatial resolution (15m/pixel)
- Has an excellent geolocation accuracy ($\sigma < 7.5$ m)
- Covers almost all of the land and still accumulating data
- But, not freely distributable



Land Use/Land Cover Change

- Create Land Cover maps from satellites
- Calibrate with Digital Confluence Project



An aerial photograph of a city, likely Tsukuba, Japan, showing a dense urban area with a grid-like street pattern, green spaces, and a large mountain range in the background under a clear blue sky.

Data Management

OGSA-DQP/WebDB/XML extensions

AIST-CSW implementation

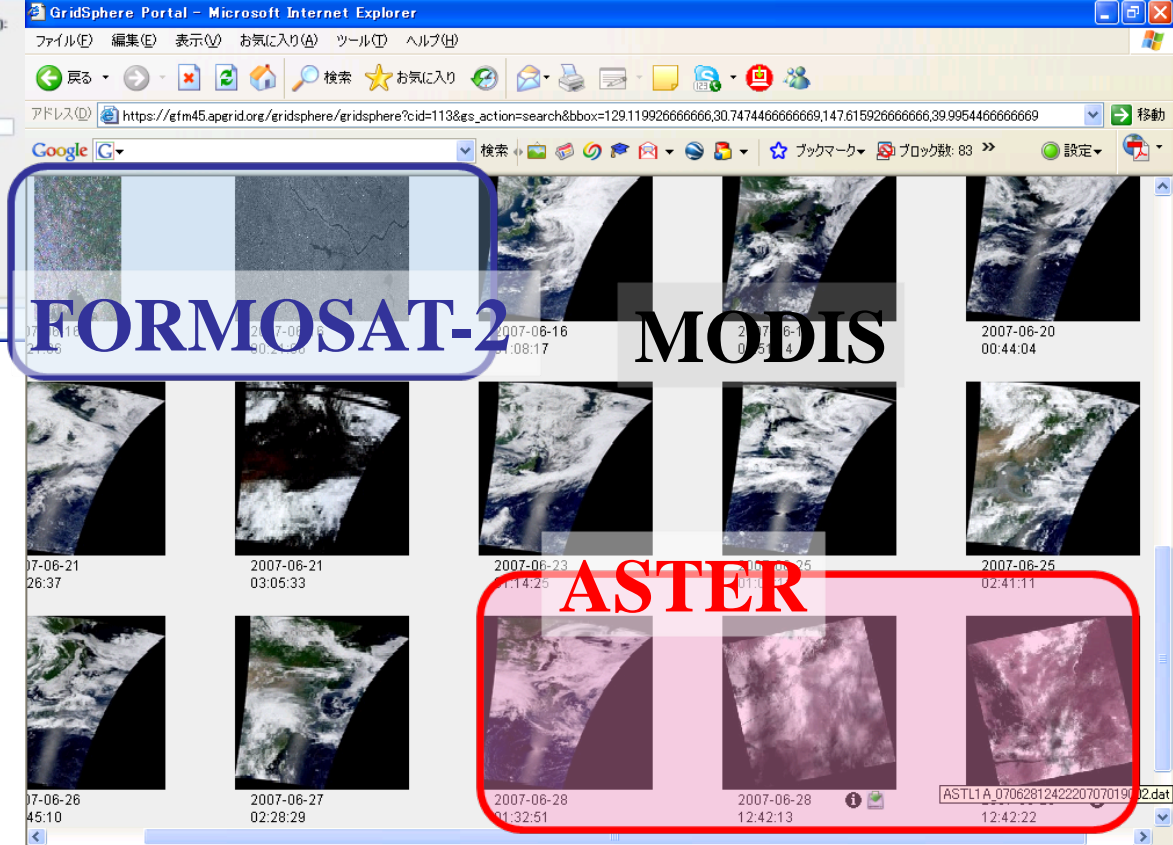
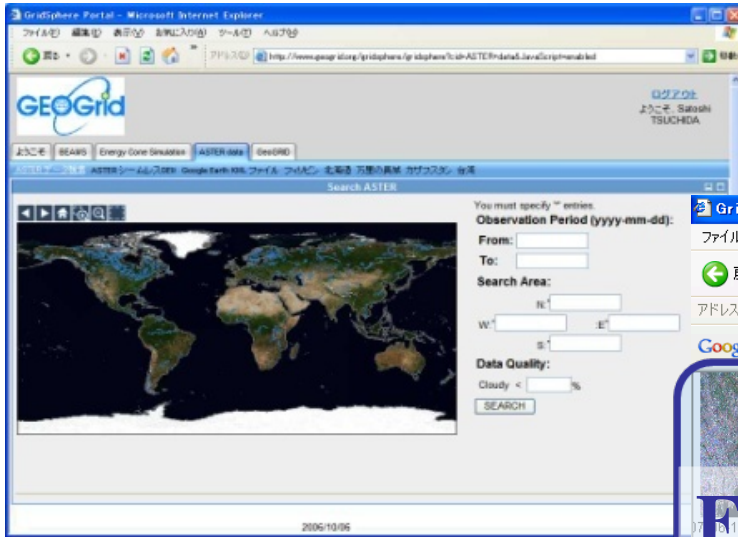
Gfarm

Tsukuba-GAMA



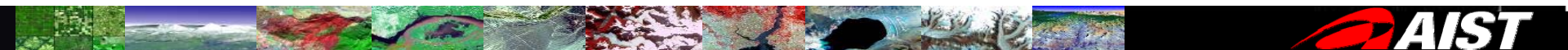
SIMS: VO based User Interface for Federating Databases

- User interface is constructed as JSR168 portlet on GridSphere
- Provides WMS Interface(in SIMS-II)

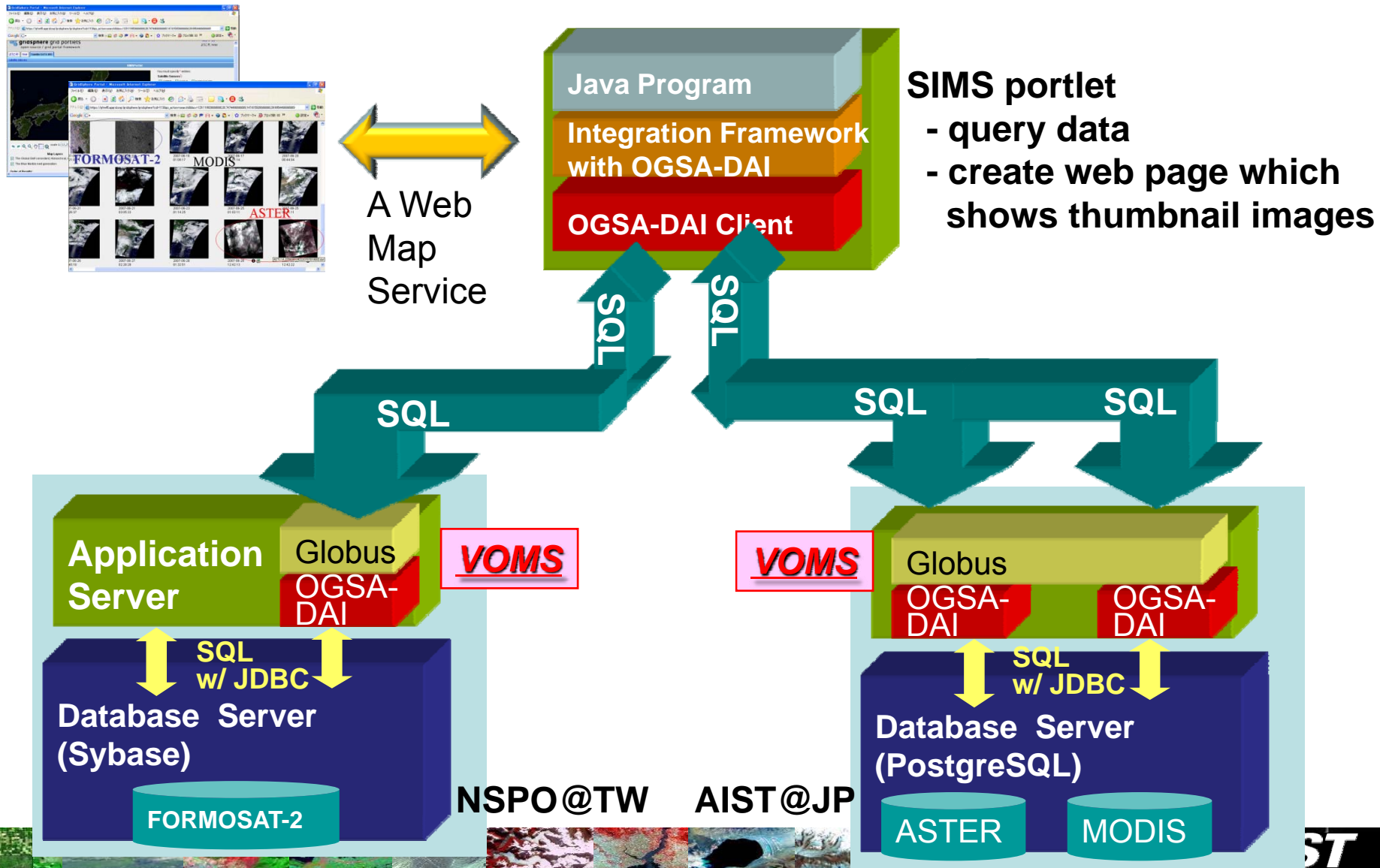


Within the Service

- Access database using OGSA-DAI Java API
- Submit image analysis via GRAM
- Retrieve input data from GridFTP server



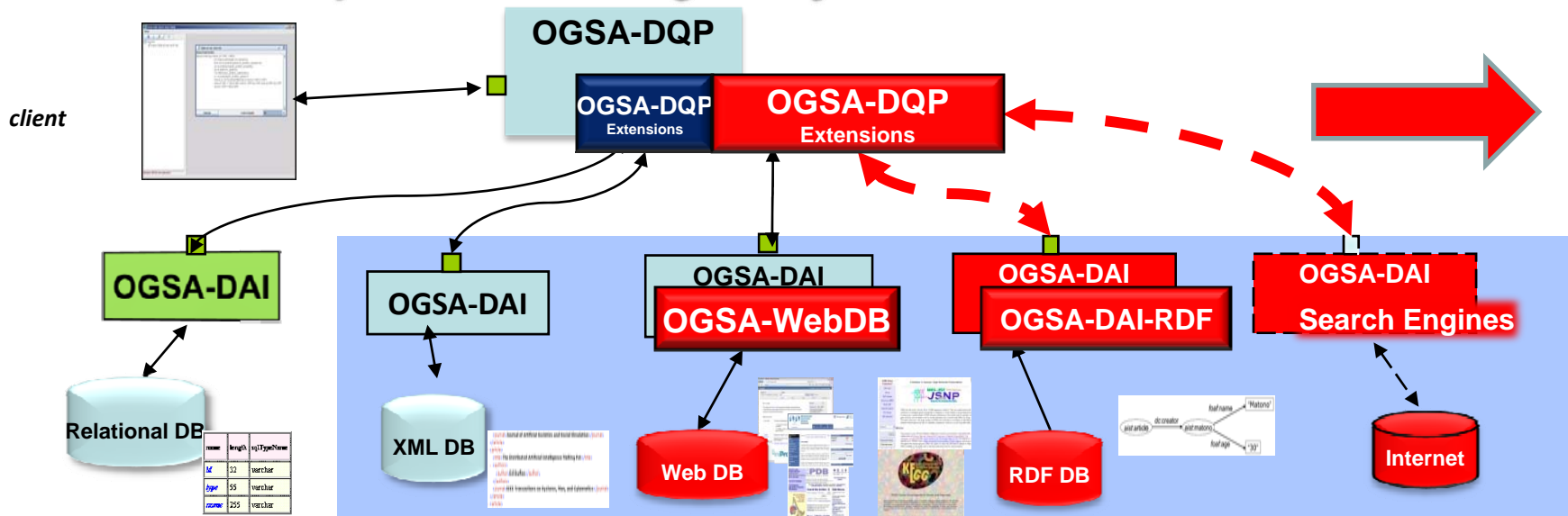
Database Federation



Current Work(1): Heterogeneous Database Integration Framework

Based on OGSA-DAI and OGF WS-DAI

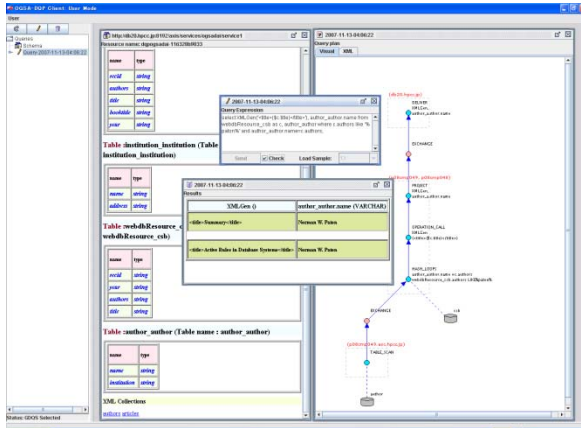
Extend it to provide heterogeneity



- **Our Feature:** Supports **XML** Databases, **Web** Databases and **RDF** Databases

Example(DQP extension@AIST)

- Integration of WebDB, relational, XML data



```

<authors>
  <author>
    <name>Steven Lynden</name>
    <field>Grid Computing</field>
  </author>
  <author>
    <name>...</name>
    <field>...</field>
  </author>
  ...
</authors>
    
```

AUTHOR

XML Databases

column name	column type
author	string
title	string

PUBLICATION

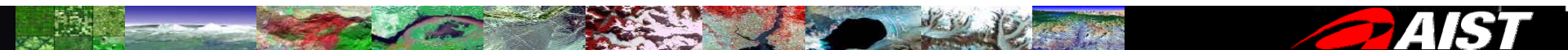
Relational Databases

A screenshot of the CiteSeer.IST website, a scientific literature digital library. The page shows a search bar, navigation links, and a list of documents. A table on the left side of the page lists columns and their types:

column name	column type
recid	
title	
authors	
snapshot	
url	string

CITeseer

Web Databases





OpenXML

```
select xauthor.name, publication.title,
       citeseer.url, xauthor.field
from publication, citeseer,
```

```
OpenXML(
  author,
  '//author',
  '//name/text() name, //field/text() field'
) as xauthor
```

```
where publication.title=citeseer.title
and xauthor.name=publication.author;
```

name	title	url	field
...
...

```
<authors>
  <author>
    <name>Steven Lynden</name>
    <field>Grid Computing</field>
  </author>
  <author>
    <name>...</name>
    <field>...</field>
  </author>
  ...
</authors>
```



```
<author>
  <name>Steven Lynden</name>
  <field>Grid Computing</field>
</author>

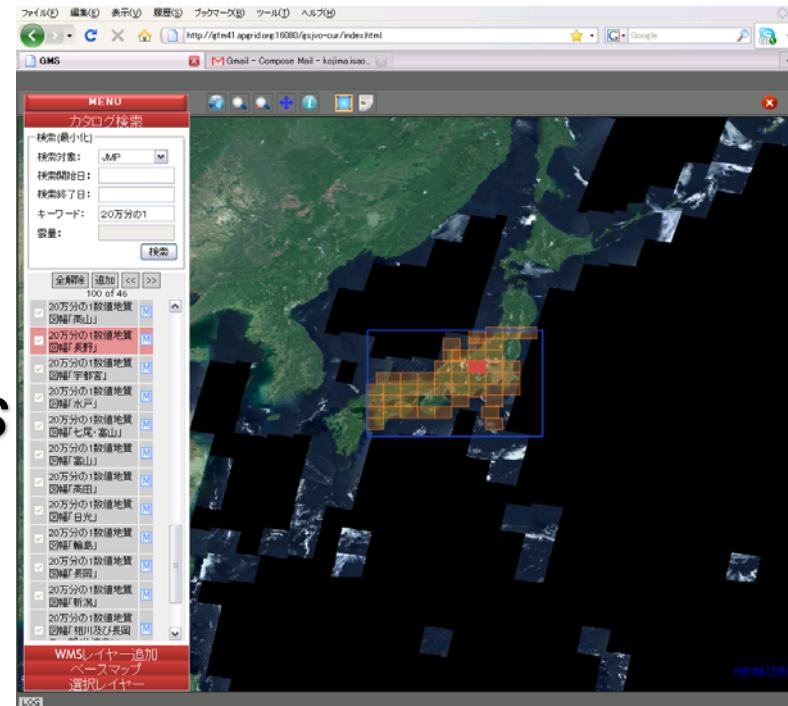
<author>
  <name>...</name>
  <field>...</field>
</author>
```



Current Work(2)

AIST-CSW: OGC Catalog Service Web Implementation

- *Supports CSW2.0.2*
 - FILTER & CQL(Subset)
 - Transaction/Harvest
 - Distributed Search
(Implementing)
- *Various Schema Profiles*
 - (Dublin Core)
 - ISO19115
 - ebRim EO Profile
 - JMP(Japan Metadata Profile)
- *Extensions*
 - Ontological Search



Our CSW Client

- VOMS-enabled
- OpenLayers

児玉、宮脇@産総研情報技術研究部門

AIST-CSW: Search engine approach (same as GENESI-DR)

- **Constructed on top of Search Engine Software**
 - WISE <http://www.bsearchtech.com>
- **Why not RDBMS? :Simple, Flexible and Fast**
 1. Query is relatively simple
 - CSW Query does not require joins as in SQL
 2. Search is Information Retrieval Style
 - Ranking/Scoring is useful
 3. Keyword match based uniform Access to various schema profiles
 - Dublin Core, ISO, ebRim EO, JMP(japan profile),,
 - Flexibility for schema is essential= > Text Search
 4. Search Performance is good
 - Metadata is append only(almost no updates)
 - Especially when the result set is huge



```
<?xml version="1.0" encoding="UTF-8" ?>
<getRecordsResponse xmlns="http://www.opengis.net/cat/csw/2.0.2" xmlns:ns2="http://www.opengis.net/ogc"
xmlns:ns3="http://www.opengis.net/gml" xmlns:ns4="http://www.w3.org/1999/xlink" xmlns:ns5="http://www.opengis.net/ows"
xmlns:ns6="http://purl.org/dc/elements/1.1/" xmlns:ns7="http://purl.org/dc/terms/" xmlns:ns8="http://www.isotc211.org/2005/gco"
xmlns:ns9="http://www.isotc211.org/2005/gmd" xmlns:ns10="http://www.opengis.net/gml/3.2"
xmlns:ns11="http://www.isotc211.org/2005/gts" xmlns:jmp20="http://zgate.gsi.go.jp/ch/jmp/">
<SearchResults numberOfRecordsReturned="100" numberOfRecordsMatched="98"> <jmp20:MD_Metadata>
  <jmp20:identificationInfo>
    <jmp20:MD_DataIdentification> <jmp20:citation>
      <jmp20:title>20万分の1数値地質図幅「鹿児島」</jmp20:title> <jmp20:date>
        <jmp20:date>2004-12-01</jmp20:date> <jmp20:dateType>002</jmp20:dateType>
        </jmp20:date> </jmp20:citation>
      <jmp20:abstract>20万分の1地質図幅は、既存の地質資料に基づいて、国土地理院発行の20万分の1地勢図のそれぞれの区画毎に編集
        した地質図です。この数値地質図は、原資料である地質図を数値化したものです。</jmp20:abstract>
      <jmp20:pointOfContact>
        <jmp20:individualName>CD-ROM編集 総括:鹿野和彦・長谷川 功、表示システム作成: 巖谷敏光・川畑 晶、データ編集:宮崎純
        一・巖谷敏光、ベクトルデータ作成:宮崎純一・中島和敏、ラスター画像データ作成:川畑 晶、メタデータ作成:渡辺和明、装丁:中島和敏
        </jmp20:individualName>
        <jmp20:organisationName>産業技術総合研究所地質調査総合センター</jmp20:organisationName> <jmp20:contactInfo>
        <jmp20:onlineResource> <jmp20:linkage>http://www.gsj.jp/HomePageJP.html</jmp20:linkage> </jmp20:onlineResource>
        </jmp20:contactInfo> <jmp20:role>009</jmp20:role> </jmp20:pointOfContact>
      <jmp20:descriptiveKeywords> <jmp20:MD_Keywords>
        <jmp20:keyword>20万分の1、ベクトル、数値地質図、鹿児島、中国西部、九州及び南西諸島、日本</jmp20:keyword>
        </jmp20:MD_Keywords> </jmp20:descriptiveKeywords>
      <jmp20:language> <jmp20:isoCode>jpn</jmp20:isoCode> </jmp20:language>
      <jmp20:characterSet>023</jmp20:characterSet>
      <jmp20:topicCategory>008</jmp20:topicCategory> <jmp20:extent> <jmp20:geographicElement>
        <jmp20:EX_CoordinateBoundingBox> <jmp20:extentReferenceSystem> <jmp20:authority> <jmp20:title>測量法</jmp20:title>
        <jmp20:date> <jmp20:date>1960-07-01</jmp20:date> <jmp20:dateType>003</jmp20:dateType> </jmp20:date> </jmp20:authority>
        <jmp20:code>TD / (B, L)</jmp20:code> </jmp20:extentReferenceSystem>
          <jmp20:westBoundCoordinate>130</jmp20:westBoundCoordinate>
          <jmp20:eastBoundCoordinate>131</jmp20:eastBoundCoordinate>
          <jmp20:southBoundCoordinate>31.333333</jmp20:southBoundCoordinate>
          <jmp20:northBoundCoordinate>32</jmp20:northBoundCoordinate>
        </jmp20:EX_CoordinateBoundingBox>
```

(以下略)

Metadata Example ISO/JMP

```
<rim:ExtrinsicObject id="urn:uuid:bce71bb1-d71b-40a9-ae91-201cbfdc61e7" objectType="urn:x-ogc:specification:csw-ebri-  
cim:ObjectType:DataMetadata">
```

```
<rim:Slot name="modified" slotType="dateTime">
```

```
<rim:ValueList> <rim:Value>2006-06-15T15:00:00Z</rim:Value>
```

```
</rim:ValueList>
```

```
</rim:Slot>
```

```
<rim:Slot name="envelope" slotType="geometry">
```

```
<wrs:AnyValue xmlns:gml="http://www.opengis.net/gml" <gml:Polygon srsName="EPSG:4326">
```

```
<gml:outerBoundaryIs>
```

```
<gml:LinearRing>
```

```
<gml:coordinates>139.4282,35.4882 140.2614,35.4882 140.2614,36.1379 139.4282,36.1379  
139.4282,35.4882</gml:coordinates>
```

```
</gml:LinearRing>
```

```
</gml:outerBoundaryIs>
```

```
</gml:Polygon>
```

```
</wrs:AnyValue>
```

```
</wrs:ValueList>
```

```
</rim:Slot>
```

```
<rim:Slot name="title" slotType="string">
```

```
<rim:ValueList> <rim:Value>URI: {
```

```
"uri": "http://maps.geogrid.org/mapserv/ms_aster.pl?",
```

```
"option": {
```

```
"LAYERS": "ASTL1A_0606161247510612129002.dat",
```

```
"SERVICE": "WMS",
```

```
"VERSION": "1.1.1"
```

```
}
```

```
<rim:Value>THUMBNAIL-URI: {
```

```
"uri": "http://www.geogrid.org/cgi-bin/thumb.pl?",
```

```
"option": {
```

```
"res": "small",
```

```
"type": "jpeg",
```

```
"filename": "ASTL1A_0606161247510612129002.dat"
```

```
}
```

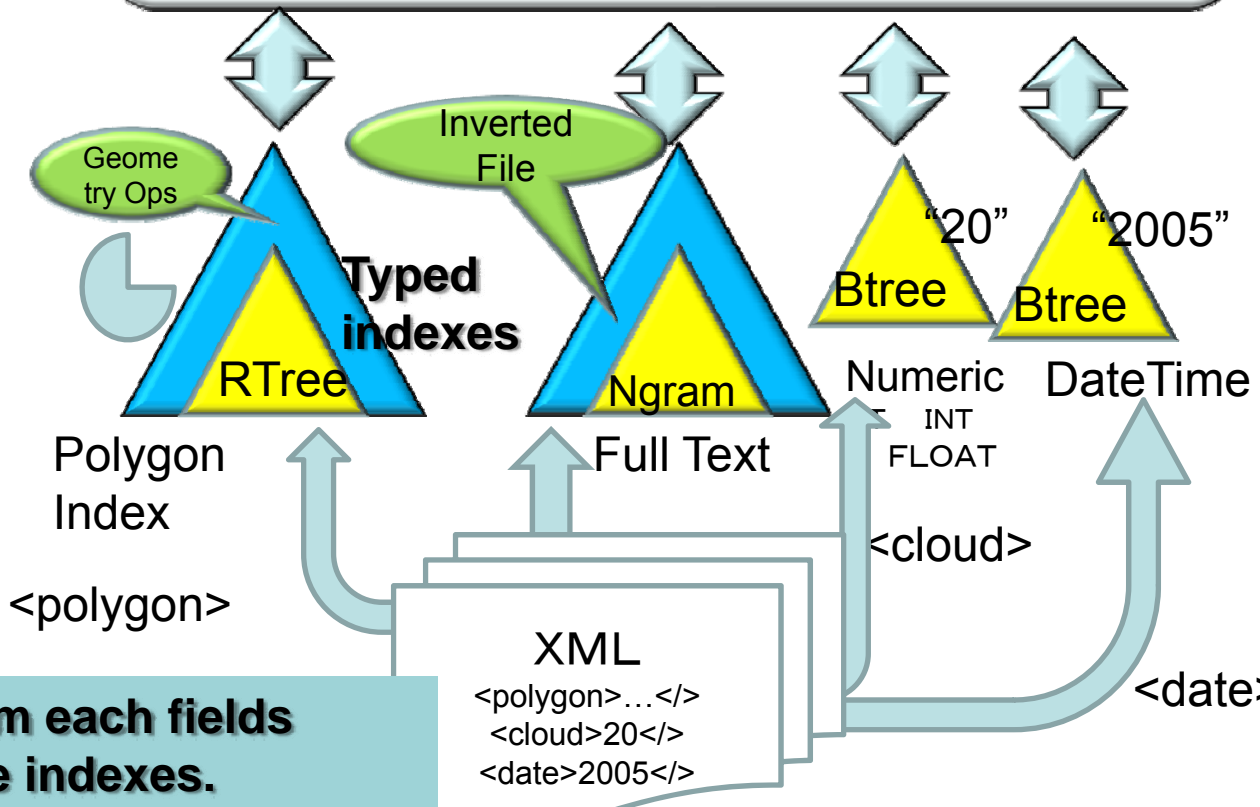
```
</rim:Value>
```

Metadata Example ebRim

OGC Filter:
AND&OR of filed conditions
(No Joins as in SQL)

Form
Query

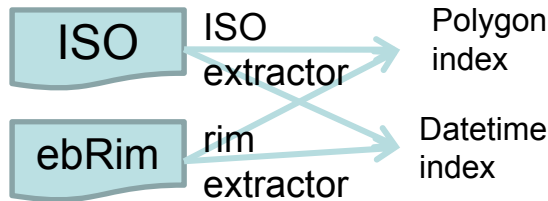
- Parse the Query and decompose
- Send Subqueries to each indexes
- AND/OR operations of intermediate results



Architecture

- **Extensible set of indexes**

- FullText
- Polygon (Rtree and/or geometric library)
- Numeric
- etc



Information is extracted from each fields and stored into the separate indexes.

rinzo.ma: User Interface for OGC Web Services

<http://www.rinzo.ma>

*Browser-Based W*S Interface & Mashup Environment*

- Highly customizable Plug-In architecture
- Utilizes javascript/Ajax/Openlayers
- Supports WMS ,CSW, etc.
- Development started at AIST, now at Shizuoka Univ.



Works with our Tiled Wall Browsers(TWB) (NOT Tiled Wall Displays!)

- *TWB: Federation of Full-Screen Web Browsers*

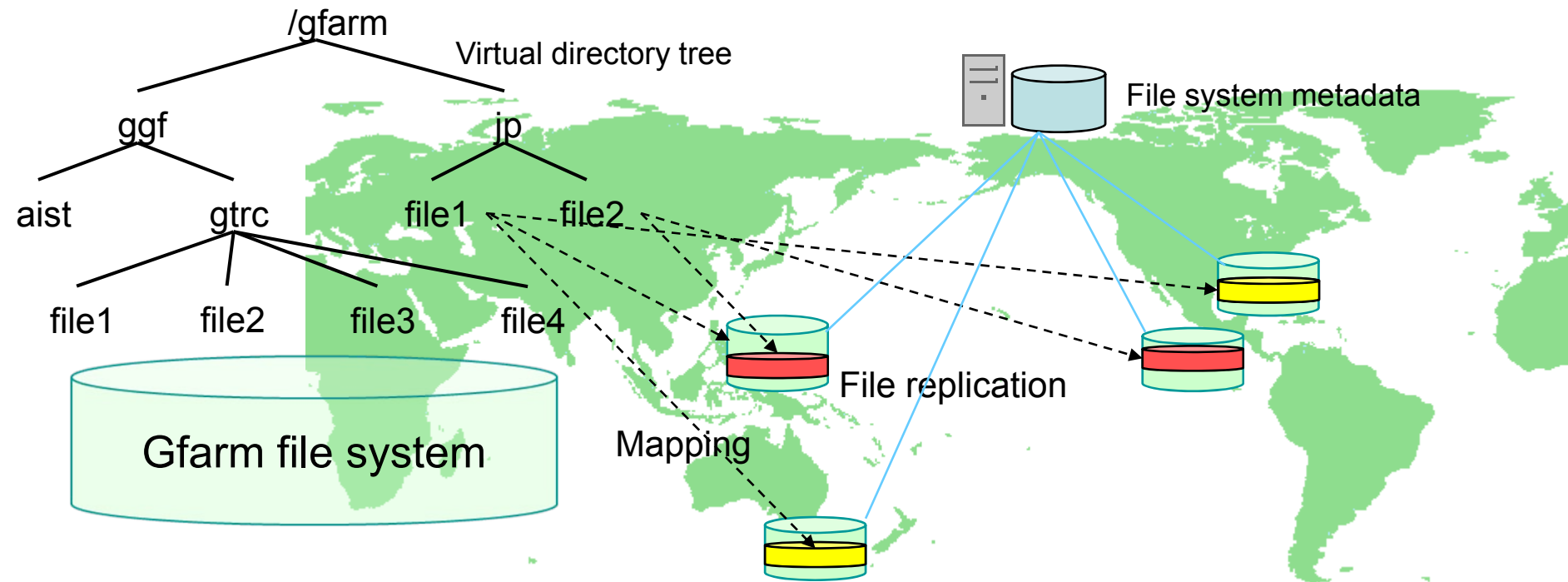


Gfarm Distributed/Cluster File System

<http://datafarm.apgrid.org/> (Initially developed at AIST, now at Univ. of Tsukuba)

Gfarm Provides:

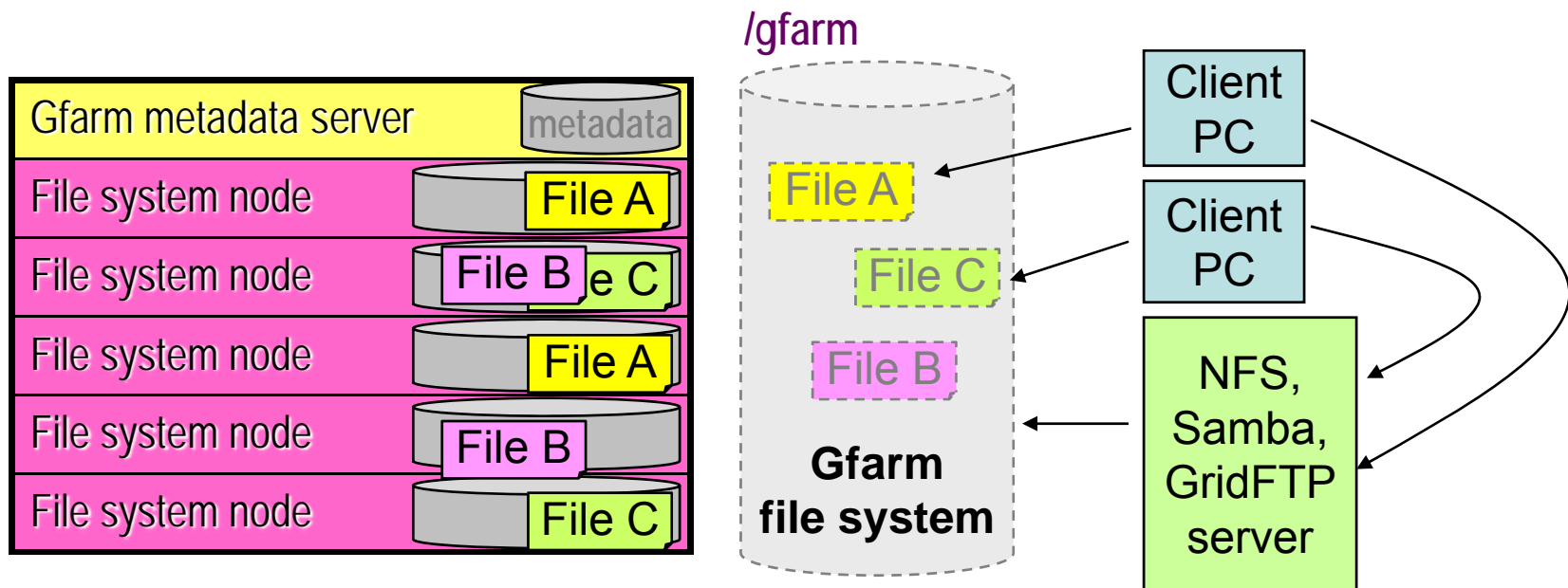
- Global naming space for distributed files.
- Replica Management functions.



Gfarm file system

Files are fragmented and distributed.

- Global Replica Management
- Interoperable with existing Unix applications(system calls).

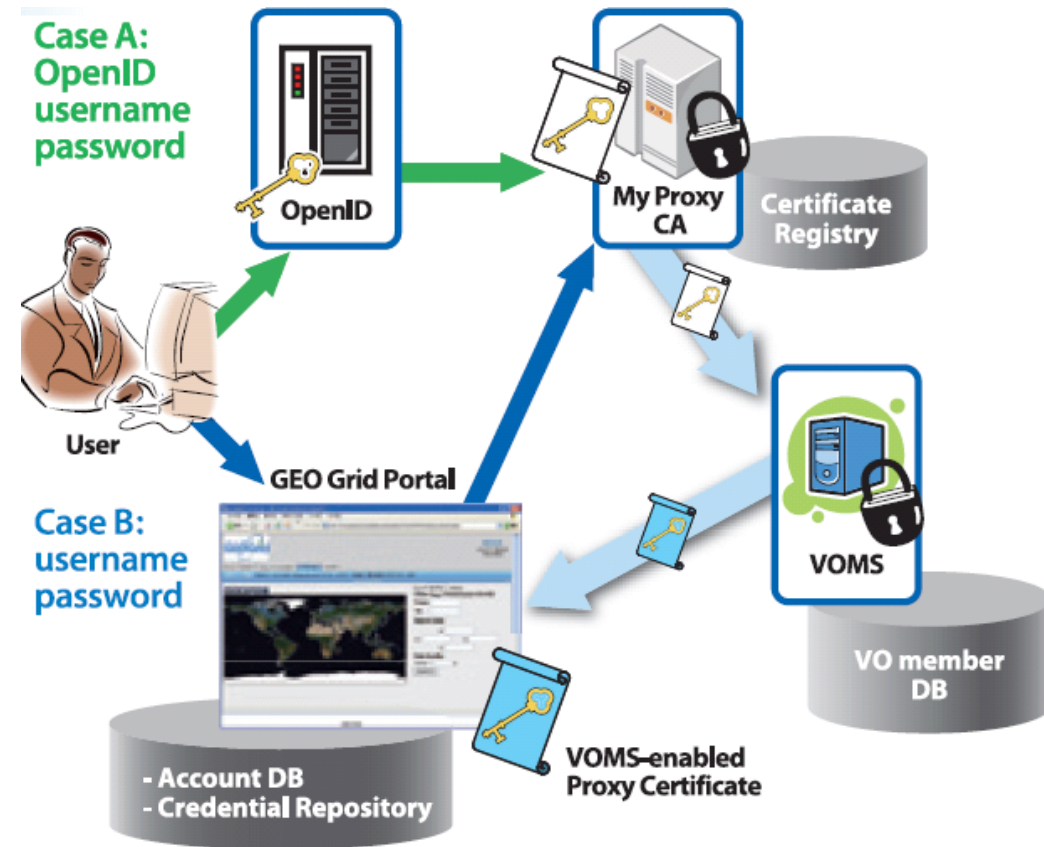


Security Framework for GEO Grid Tsukuba-GAMA

Provide an Integrated
Credential Management

Can cooperate with

- OpenID
- Username/password
etc





Login Flow for OpenID user

ログイン

ユーザー名

パスワード

ログイン |

OpenID Login

OpenID.ne.jp

1回の登録と1つのアカウントでOpenIDを認証するすべてのサイトに、すぐ登録することができます。

OpenID.ne.jp

1回の登録と1つのアカウントでOpenIDを認証するすべてのサイトに、すぐ登録することができます。

<http://naotaka.openid.ne.jp/>

[プロフィール](#) | [アカウント管理](#) | [ログアウト](#)

OpenID 確認

Accept Once for GEO Grid Portal

現在

https://gfm41.apgrid.org:13443/gridsphere/gridsphere?cid=openidlogin&gs_action=verifyRequest に

あなたのOpenid (<http://naotaka.openid.ne.jp/>)

あなたの個人情報と一致しているという証明を要求しています。

https://gfm41.apgrid.org:13443/gridsphere/gridsphere?cid=openidlogin&gs_action=verifyRequest に証明するための必要な個人情報を追加チェックし、認証してください。

ifyRequestで、



一度だけ認証

認証状態を保持

認証拒否



Summary

GEO Grid Implementation Features

- Supports OGC standards
 - *VOMS-enabled OGC services*
 - *AIST CSW implementation*
- Supports OGF standards
 - *OGSA-DAI/DQP/RDF extensions*
 - *Gfarm*
 - *Tsukuba-GAMA*

Extending the
Technology

We want to collaborate with other projects

Extending the
Community

