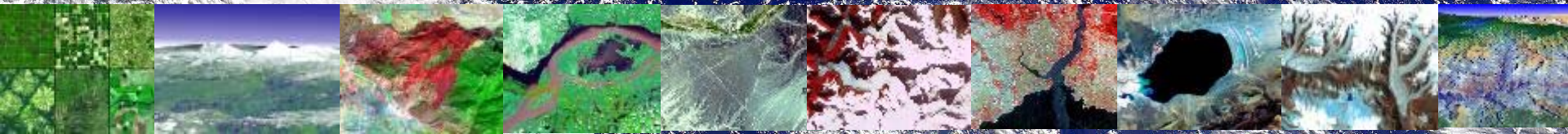


GEO Gridの概要と そのIT技術の現状・将来について

小島 功

情報技術研究部門
産業技術総合研究所
kojima@ni.aist.go.jp

一部資料作成協力&引用:
山本直孝、Steven Lynden,
岩男弘毅、山本浩万、児玉信介、松岡昌志
(順不同、産総研情報技術研究部門)



概要

- GEO Gridとは
 - 産総研地質部門[旧 地質調査所]と情報部門との分野連携プロジェクト
 - 地質や地理、衛星データ処理など分野の研究者にサービスを提供しつつ、その周辺でIT関連の研究を行う。
- 特徴
- アプリケーション・VO
- GEOにおけるIT技術の現状
- IT関連の研究開発課題



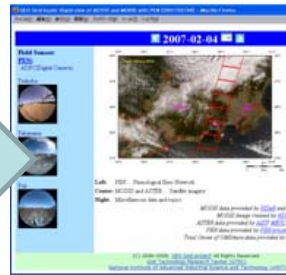
<http://www.geogrid.org>



GEO Grid とは？

- 衛星画像データベース (ASTERなど) を基礎とした、e-Science基盤のこと。

UIはWeb2.0的
アプローチ



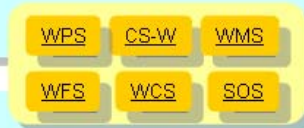
Application Model prediction, evaluation,
Alarm official announcement, etc.

Global Warming
Water Cycle
Ecosystem
Mitigation
Environment
Resources
Geology
Agriculture



AIST: OGC Associate member (since 2007)

OGC規格の
サポート



Open Geospatial Consortium (OGC)
compliant Web Service

User Interface
Workflow Tool

Open Grid Forum (OGF)
compliant infrastructure

Grid Security Infrastructure
Web Portal
Virtual Organization

AIST: OGF Gold sponsor (a founding member)

グリッド技術を
基盤

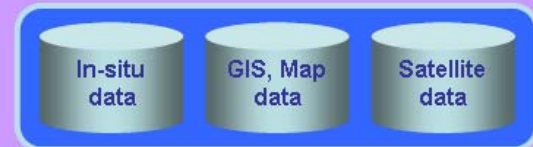


Computing Grid

Storage

ASTER
MODIS
PALSAR
JERS-1

Data Grid & Database Federation



Storage Grid



特徴

- **VO** (仮想組織) の概念に基づくセキュリティの担保
 - フリーデータだけでなく、地物関連データの特徴
- グリッド技術によるスケーラビリティの提供
 - 分散・並列・高性能化
- 標準の尊重と標準コミュニティへの貢献
 - OGC (Open Geospatial Consortium) & GEOSS

For Details, Please have a look our IEEE Systems Journal Paper “Design Principles and IT Overviews of the GEO Grid” Vol2. No.3, 2008.09 GEOSS特集号



GEO Gridで扱われるデータ

Linked by:

- Geolocation
- Observation/creation time
- Specific parameter

衛星画像データ

- ASTER
- MODIS
- PALSAR

→ **GeoTiff/WCS**

Raster DATA

センサーネット

- PEN

→ **SensorML/SOS**

Sensor DATA

GISデータ

- Geological maps
- Roads, railways, etc

→ **GML/WFS**

Vector DATA

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



GEO Grid におけるVO

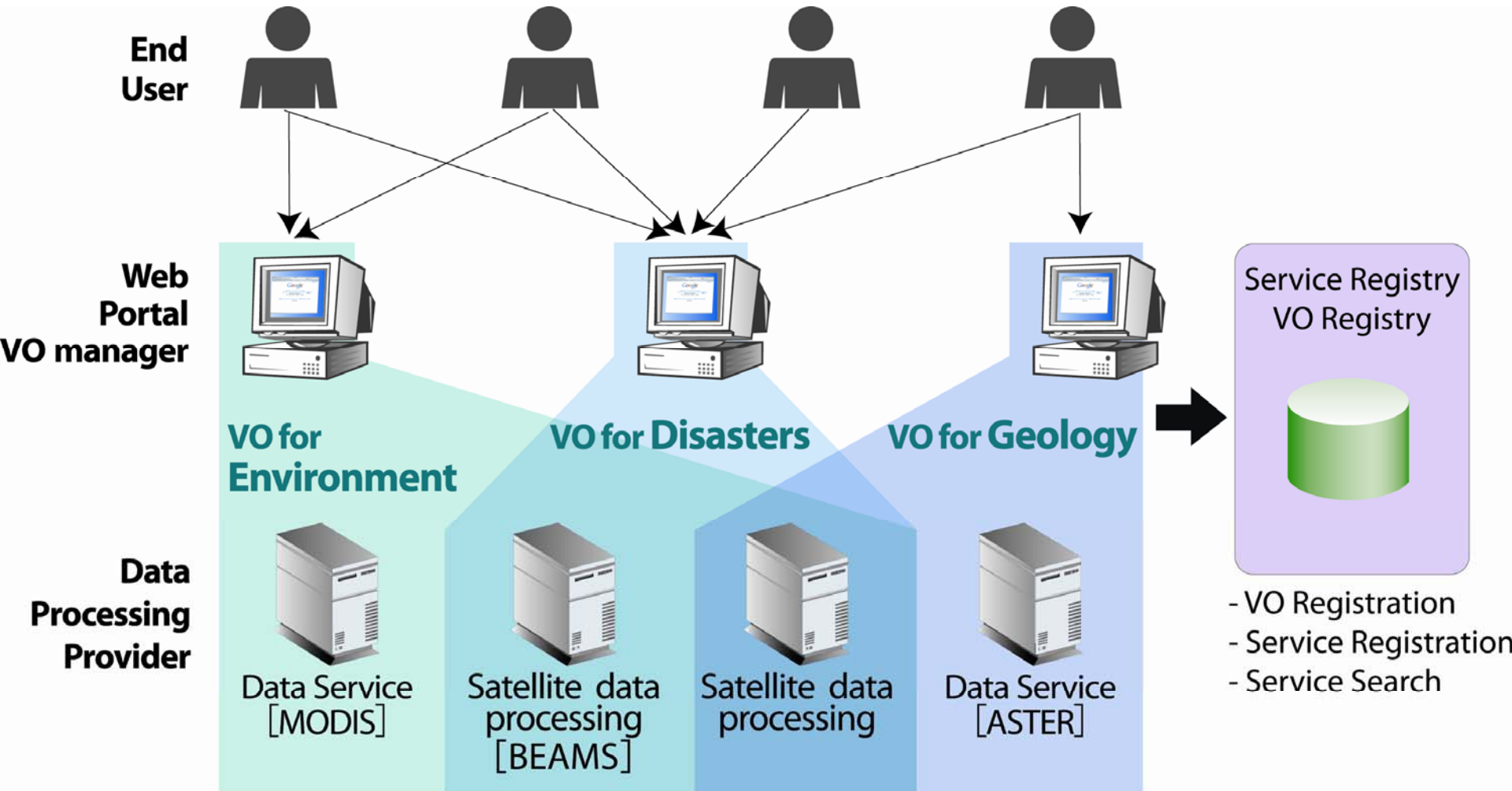
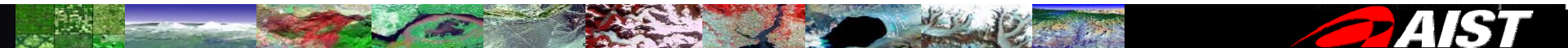


Figure 2. VO design



VO提供の理由と特徴

- 地質データや衛星画像、解析プログラムなど
 - 商用データなど、必ずしもフリーで提供できないものが多い。
 - 利用における権利関係が複雑。
 - データ提供にまだまだ抵抗感のあるデータ提供者が多い
 - 高価な計算機やサービスをフリーで使わせたくない。
- VO(Virtual Organization)の支援。
 - 仮想コミュニティの提供
 - 応用サービス・データをすべてVO単位で管理。

VO & グリッド技術の利用

- データベース統合
 - **OGSA-DAI**を用いた複数データベースの同時検索、結果の統合
- データ転送
 - **GridFTP**を用いたファイル・データの第三者転送
- アプリケーション実行
 - **GRAM**によるアプリケーションの実行・管理
- セキュリティ
 - PKI (Public Key Infrastructure)とX.509証明書及びその拡張である代理証明書(**プロキシ証明書**)を用いたグリッドセキュリティ基盤 (Grid Security Infrastructure, GSI)
 - シングルサインオンや権限移譲 (delegation) が実現され、頑健かつ多機能なセキュリティ基盤

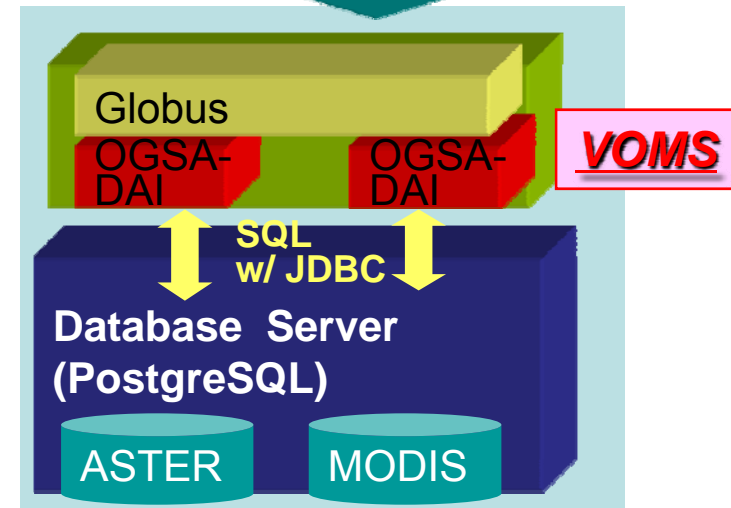
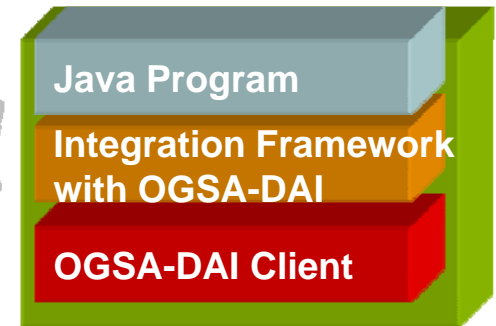
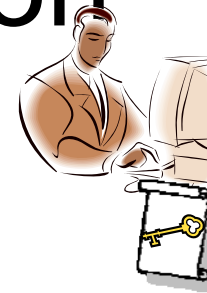
Grid Solution

- **Grid Security Infrastructure (GSI)**

- Security infrastructure based on PKI and X.509 certificates.
 - Users use various database with single authentication / SSO.

- **Virtual Organization Membership Service (VOMS)**

- Users and Resources are controlled under VO
 - Administrators of service provider do not need to configure access control for each user
- VO単位での認証ができる。
- 個人単位のアカウント管理をしなくていいので圧倒的に楽。



VOMS

Execute simulation



transfer all data on-demand



GEO Grid Toolkit(開発中)

提供者等のロールに応じてツールを使用

Service Provider

GEO Grid Service Development Kit
(GEO Grid SDK)

VO manager

GEO Grid Virtual Organization Tools
(GEO Grid VOT)

Portal manager

GEO Grid Portal Development Kit
(GEO Grid PDK)



gridsphere portlets with VOMS

naotaka's GEO Grid Portal - Mozilla Firefox

http://gfm45.apgrid.org:12088/gridsphere/gridsphere?cid=OGCProxy&gs_action=getContents&URL=https://geodata2.geogrid.org/mapserv/basemap/bmng&

2008/02/21

設定 管理 Satellite DATA IMS Credential Service Grid Science DCP Landslide

satellite data ims On-demand DEM mosaic job results

SIMSportlet

You must specify * entries.

Satellite Se
 ASTER
 Observation

From:
 To:

Search Are
 Degree

W: -180

Data Quality

ASTER Higher Level Products

Return to Search

Filename	ASTL1A_0806140139460806189056.dat
Scene Upper Left (Lon, Lat)	(139.7811, 36.5529)
Scene Upper Right (Lon, Lat)	(139.4622, 36.4515)
Scene Lower Left (Lon, Lat)	(138.6193, 35.9995)
Scene Lower Right (Lon, Lat)	(139.2957, 35.8988)

DEM/ORTHO product

完了

The screenshot displays the GEO Grid Portal interface. On the left, a world map is shown with a red border. The main content area features a search form for satellite data, with 'ASTER' selected under 'Satellite Selection'. Below the search form, search results for 'ASTER Higher Level Products' are displayed, including a table with file names and coordinates. A small thumbnail image of a satellite scene is also visible. The browser's address bar shows the URL: http://gfm45.apgrid.org:12088/gridsphere/gridsphere?cid=OGCProxy&gs_action=getContents&URL=https://geodata2.geogrid.org/mapserv/basemap/bmng&. The page includes navigation tabs for 'Satellite DATA IMS', 'Credential Service', 'Grid', 'Science DCP', and 'Landslide'. The date '2008/02/21' is displayed at the top right.

VO User Interface

- User interface is constructed as JSR168 porlet on GridSphere

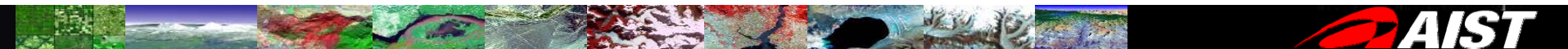
日本、台湾、USなどのサーバを
Webサービス経由で統合、しかもセキュリティを
担保して接続

FORMOSAT-2

MODIS

ASTER

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



OGC規格のサポートと貢献

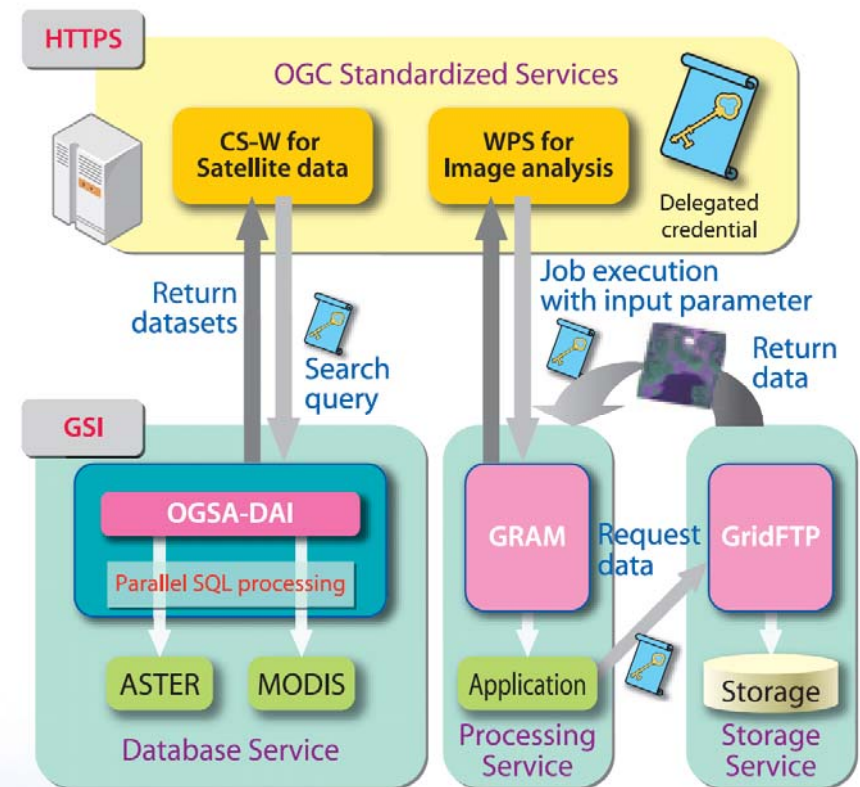
地理・GIS分野の特有の事情：

- WMS(画像の検索と配信)WFS(地物の検索と配信)CSW,(メタデータの登録・収集・検索)WPS(データ処理)など独自の標準が一通りそろっている。
- カタログでメタデータを検索、サービスのURLを得てデータを検索・取得し、その結果を別サービスで加工、、、
- スケーラビリティや性能に問題

1. バックエンドでサポートする場合の問題を解決

2. その知見をベースにOGC-TCやOWSへ参加

- セキュリティ
- ワークフロー
- WPS



自前のソフトウェア実装

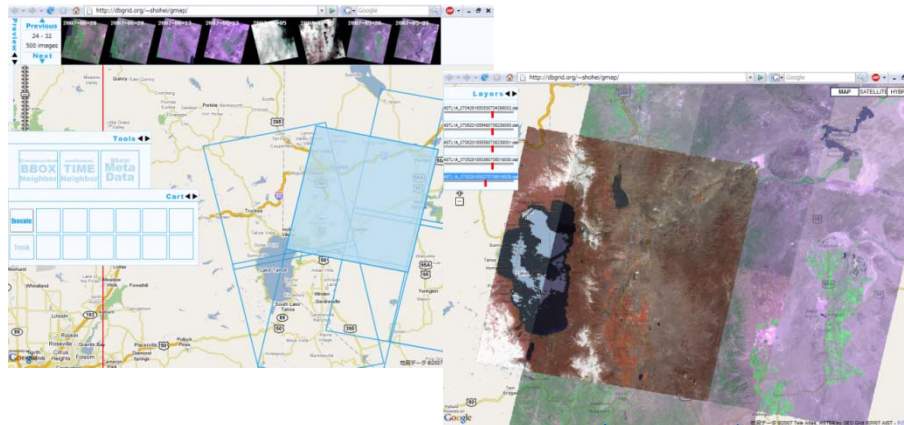
- WFS、WMSやCSW、WPSなどで独自の実装や改良を試みている。
 - パブリックドメインのソフトウェア
 - **高速化・大規模化に難点**
 - 画像サーバ
 - オンデマンドの処理(リアルタイムの画像切り取り・配信など)が弱い
 - カタログサービス(後述)
 - フリーソフトでは200万件ぐらいからのポリゴンのマッチング検索が結果の数が増えるに従って、うまく動かなくなる。
 - グリッド環境との統合・接続のため
 - セキュリティ
 - 独自の知見からの機能改良
 - 異種データの扱いや並列データ転送など



rinzo.ma: User Interface for OGC Web Services

Browser-Based W*S Interface & Mashup Environment

- Highly customizable Plug-In architecture
- Utilizes javascript/Ajax/Openlayers
- Supports WMS ,CSW, etc.



Tiled Wall Browsers (フルスクリーンのブラウザたちが連動して1枚の画面！)
(This isn't Tiled Wall Displays)

特許出願中



VOMS-enabled OGC Services

- OGCサービス=HTTPベースで簡潔かつ高機能
 - セキュリティに対する枠組みがない！
- VOMS-enabled OGC servicesの必要性
 - 証明書のdelegation
 - クライアント(Webブラウザ)にユーザ証明書を期待できない、、
 - OGC Proxyの開発
 - Single Sign On Tool の開発など

アプリケーション & VO のリスト (例)

Global roads development using WPS (ECO VO)

Field sensor data integration using SOS (FON VO)

Shake amplitude estimation using WPS (GHz VO)

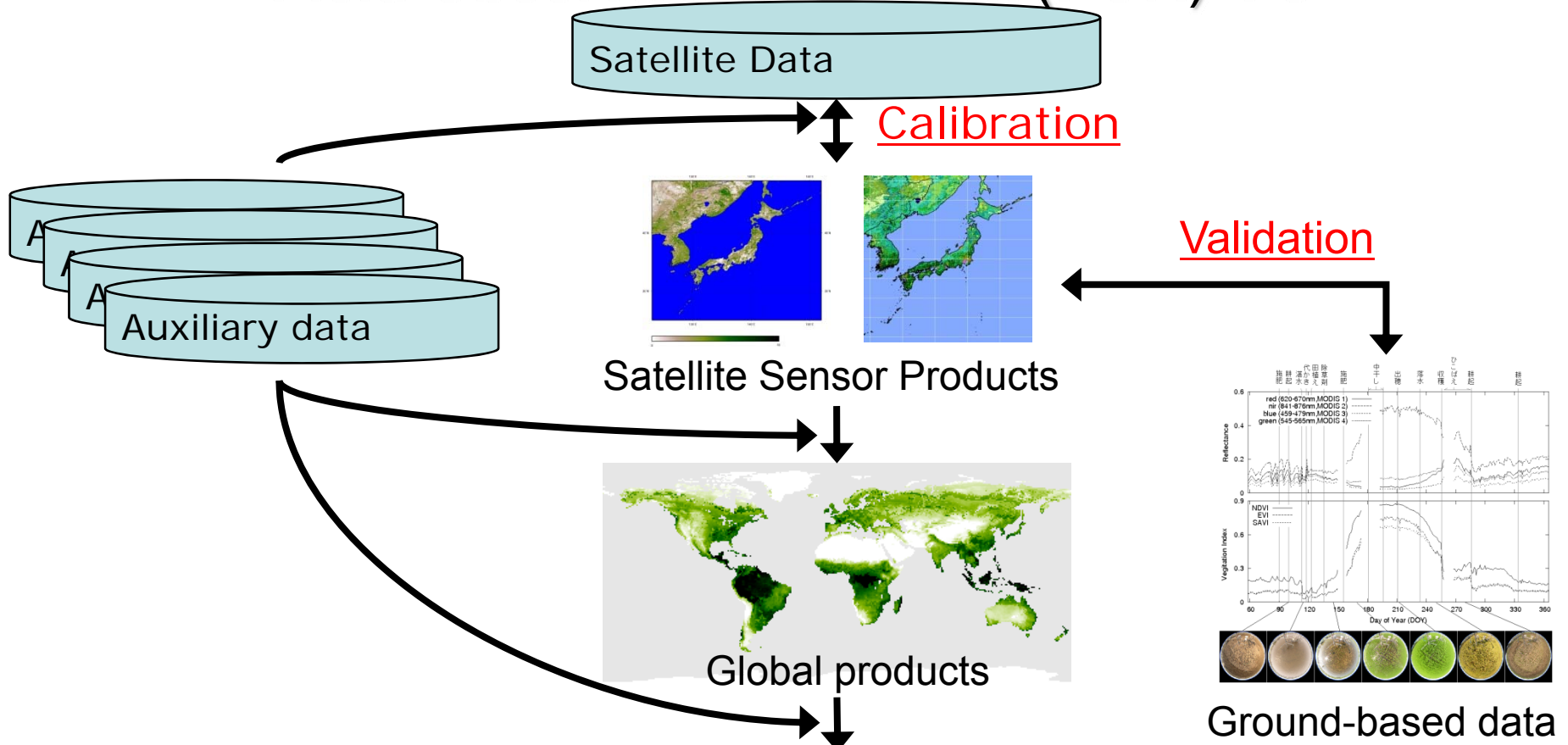
Landslide application using WFS and WCS (GHz VO)

Registry for geological data (GSJ VO)

(Accumulating data and metadata for application...)



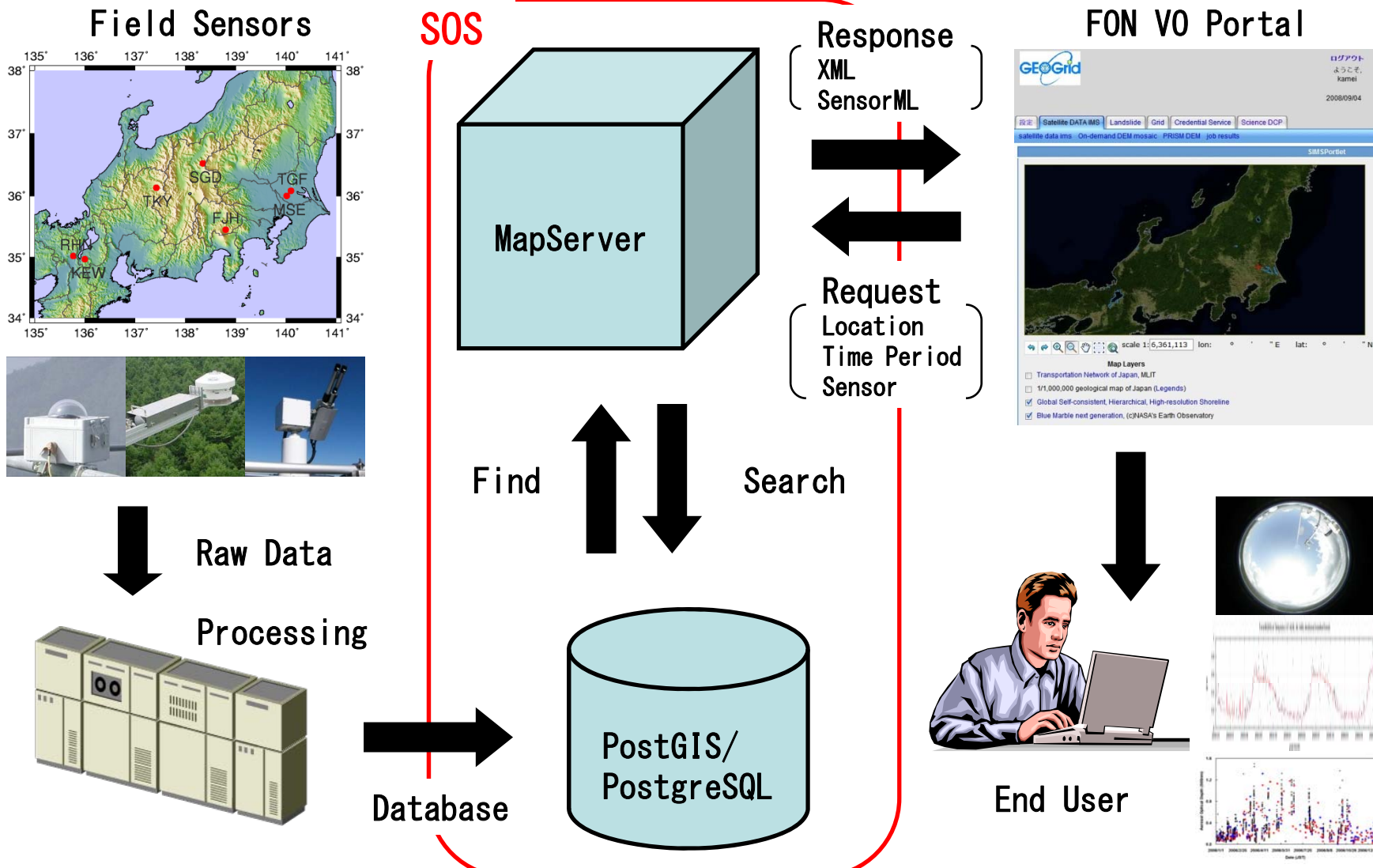
Field sensor data integration at Field Observation Network (FON) VO



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



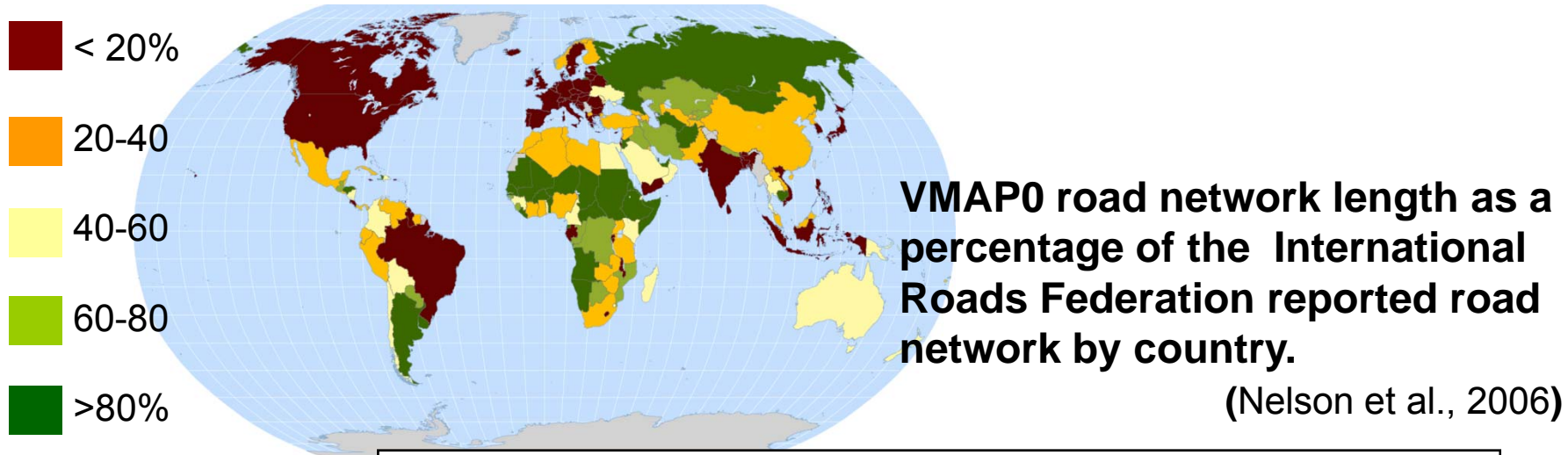
AIST's SOS Architecture



Global roads development

Why road map ?

- Best available data covers only one-quarter to one-third of the existing road networks
- Accuracy varies considerably by region.

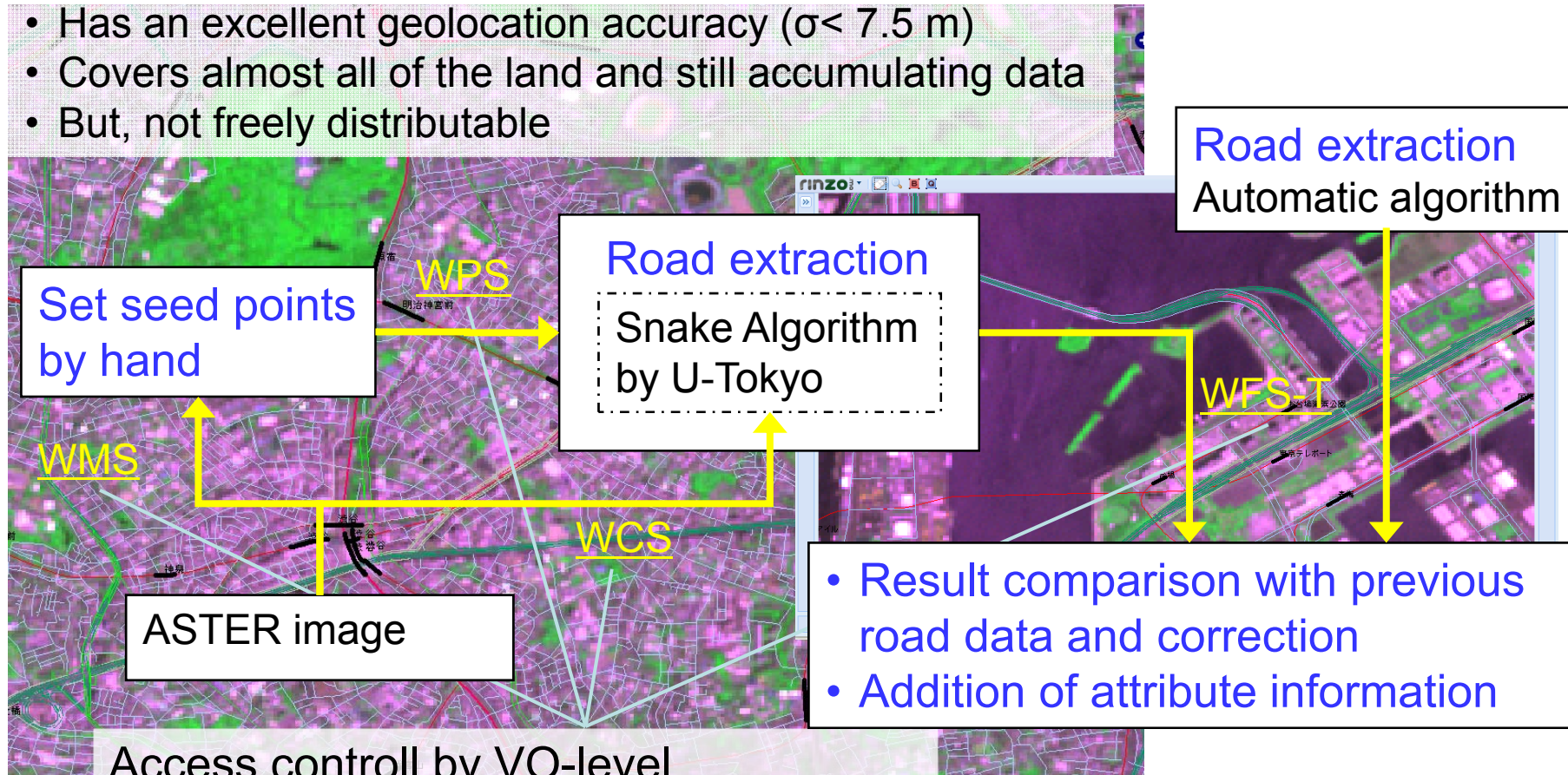


The Demand

- Pre- and post-disaster planning
- Economic development Analysis

ASTER image for 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



Access control by VO-level authentication and authorization

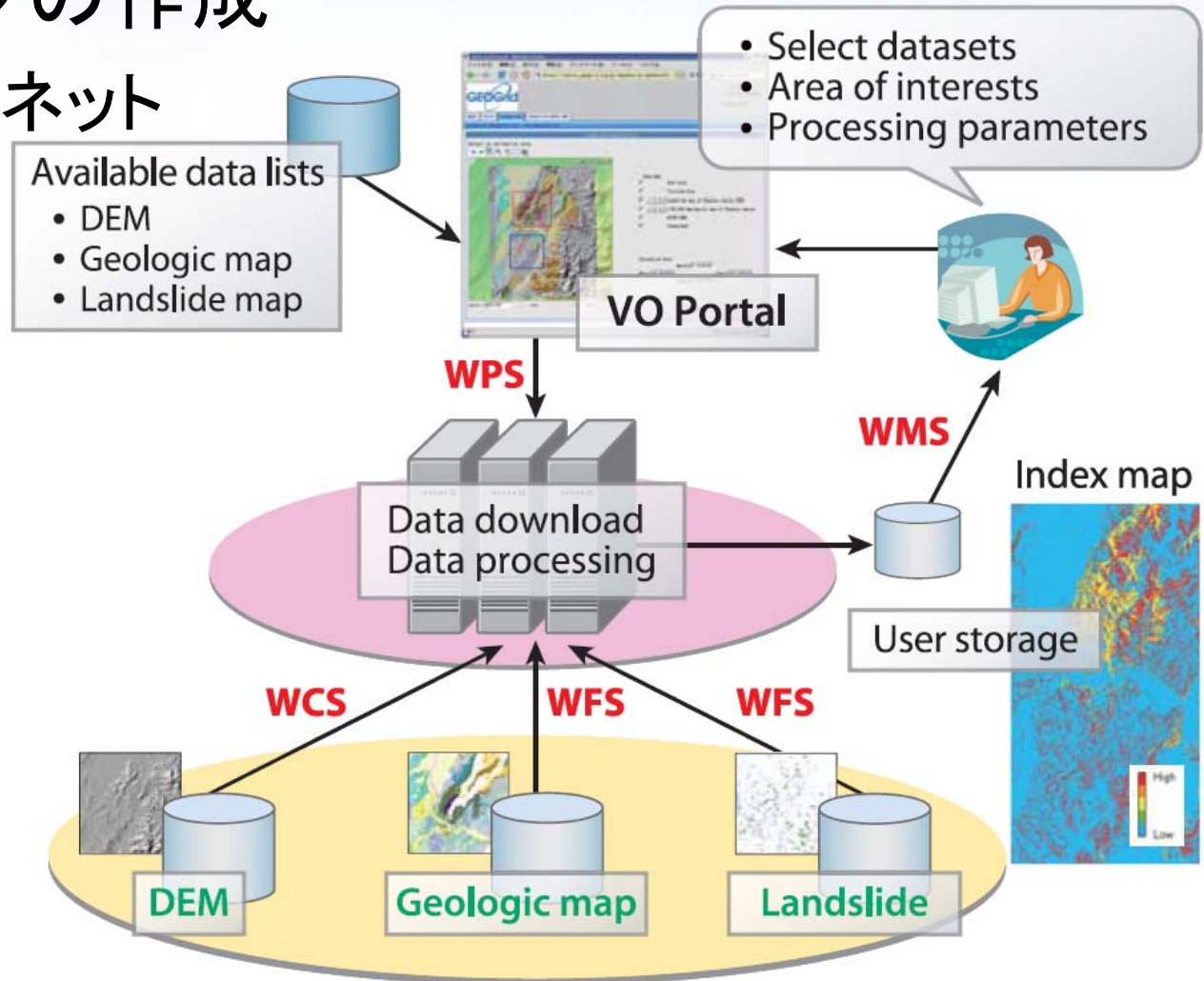
Land Use/Land Cover Change

- 衛星画像からのDEM/LCマップ作成
- Digital Confluence Project による校正



防災

- 地滑りマップの作成
 - ニューラルネット



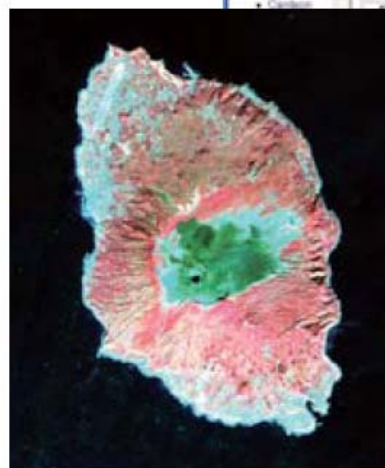


防災

- 自然災害 (火山)



ASTER archive

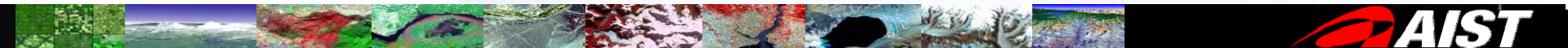
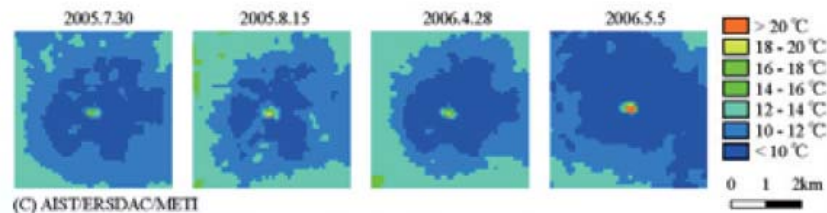


ASTER image database for volcanoes (Official Version 1.2)

Volcano name : Merapi
 Latitude : -7540 degree
 Longitude : 110440 degree
 Elevation : 2911 m
 Link to the Smithsonian Global Volcanism Program

Time series of ASTER image data Click any of thumbnail images to view the full resolution image.

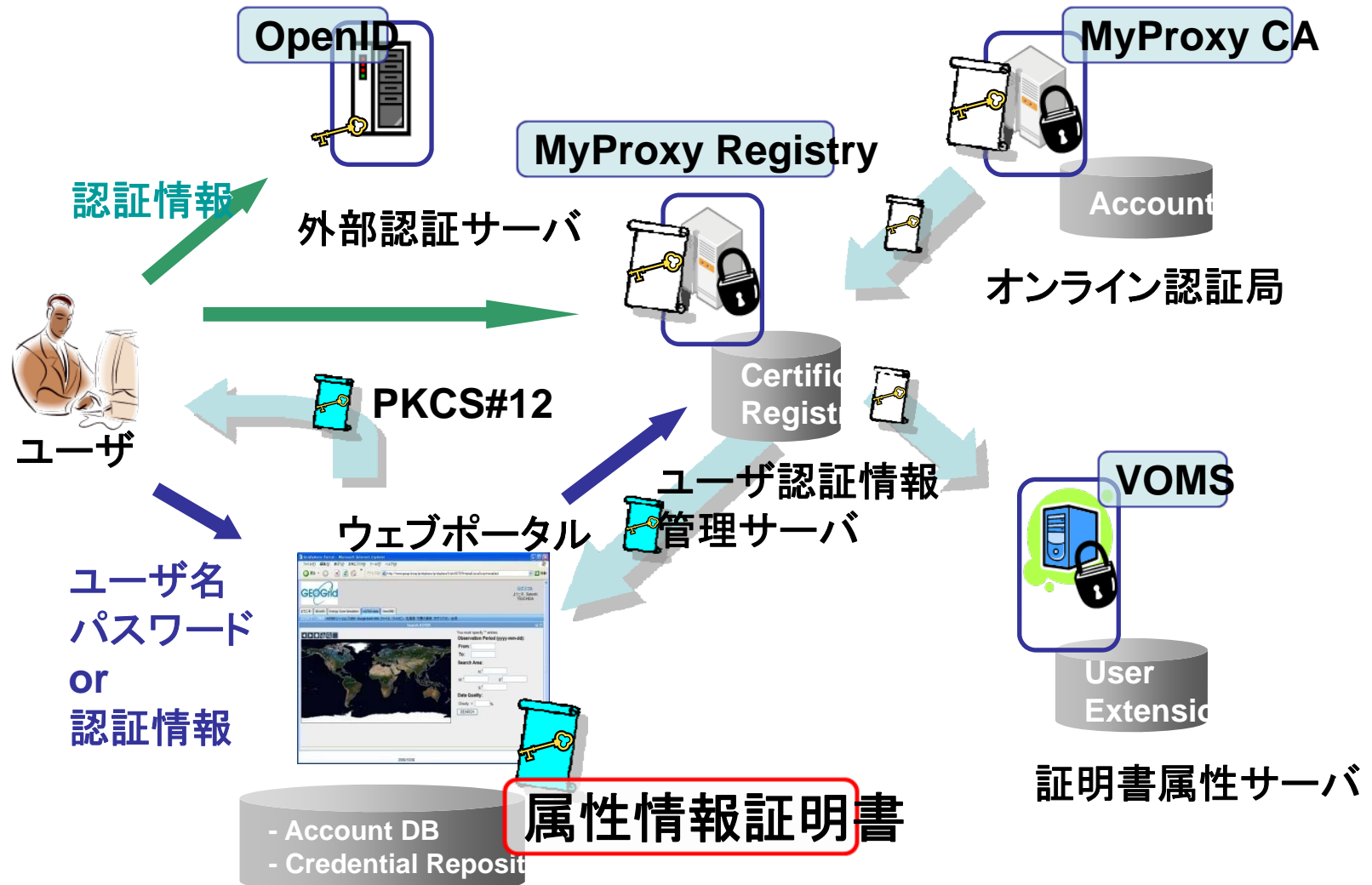
File name and Path (url)	VNR image	SWIR image	TIR image	Nadir view	Backward view
4_120					
6_119					
7_120					
1_119					



研究開発(1) セキュリティ

- Tsukuba-GAMA
 - 頑健なグリッドの認証と、コミュニティですでに使われている認証を連携させたい。
 1. 敷居が低いが安全性が低い認証
 2. 頑健だが管理の手間がかかる認証
 3. 既にインターネットなどで普及している認証
 4. など
- これらをグリッドの認証と効果的に結び付けたい。
 - 証明書の動的な生成と統合的な管理で、さまざまな認証方式を連携させる。

Tsukuba-GAMA



Login Flow for OpenID user

OpenID Login

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。

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

現在

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

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。

一度だけ認証

認証状態を保持

認証拒否

**Accept Once
for GEO Grid Portal**

ifyRequestで、



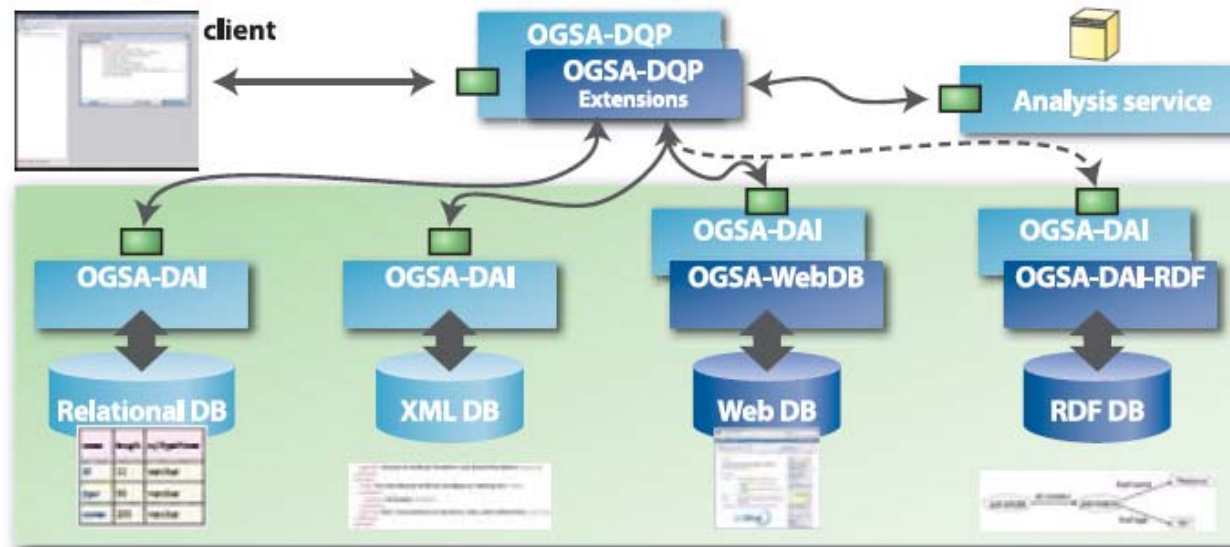
研究開発(2) データ統合

- 1) データベース統合の課題
 - 異種データベースの統合
 - 関係データベースだけでなく、XMLやRDFの統合
- 2) OGCカタログサービスの課題
 - 大規模なデータへの対応
 - データ量が100万個のオーダー
 - 検索結果が数10万個の場合
 - ユーザの意図に反して、膨大な結果が出たり、逆に一つもなかったり、
 - 異種スキーマの統合的利用
 - ebRIMプロファイル
 - ISOプロファイル
 - JMPプロファイル

Heterogeneous Database Integration Framework

Based on OGSA-DAI and OGF WS-DAI

Extend it to provide heterogeneity 異種データベースの統合ツール



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

Example

広域のDBをWebサービス基盤上で宣言的に結合

- 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

column name	column type
author	string
title	string

PUBLICATION

関係データベース

column name	
recid	
title	
authors	
snapshot	
url	string

CITSEER

Webデータベース





OpenXML

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

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

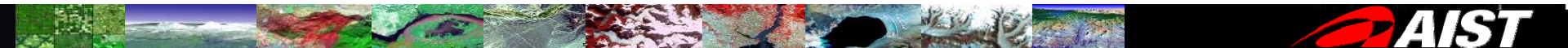
where publication.title=citeseer.title
and xauthor.name=publication.author;

name	title	url	field
...
...

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

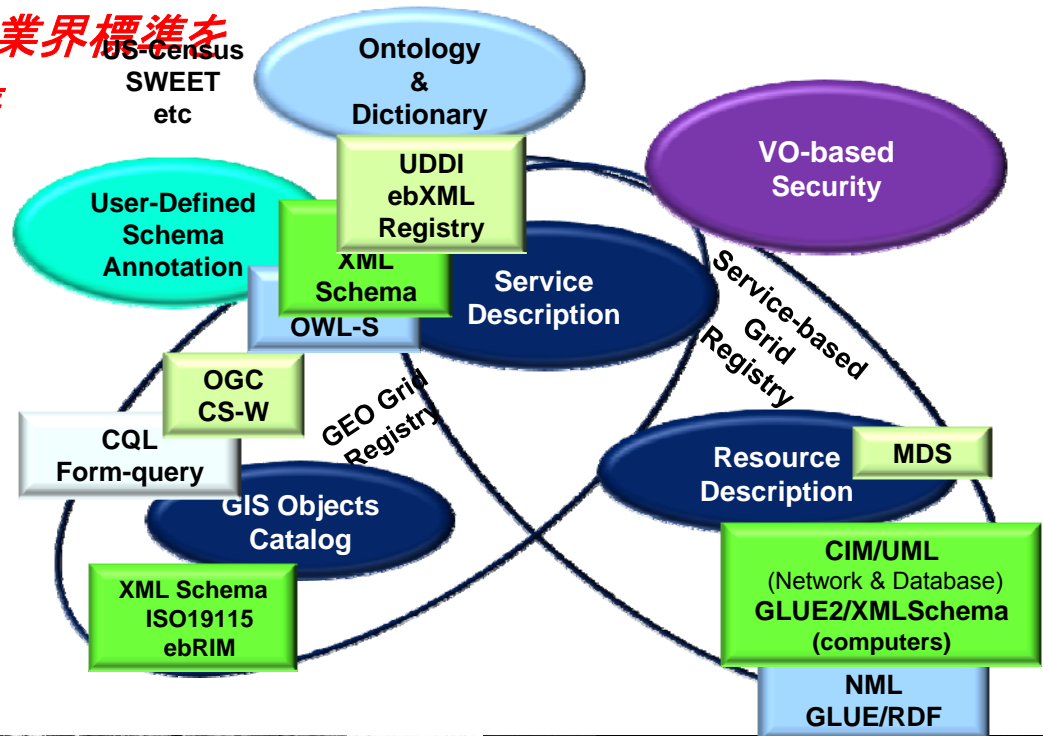
```
<author>
  <name>Steven Lynden</name>
  <field>Grid Computing</field>
</author>
```

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



Heterogeneous Service/Object Registry for GEO Grid メタデータカタログ & サービスレジストリ Catalog Service Web

- **Complies OGC(Open Geospatial Consortium) catalog service CS-W(Catalog Service Web) specification. OGCの規定するメタデータフォーマットとアクセス法**
 - Can handle Heterogeneous XML schema (ISO, ebRim, GLUE etc).
 - **異種のデータスキーマ[プロファイル]を扱わねばならない。**
 - **Webサービスレジストリなど、業界標準を越えた部分のメタデータとの連携**
- **Developed with Search Engine software**
 - DBMSを使わない**
 - 全文検索(スキーマレス)が基礎**
 - 検索は距離とランキングが基礎**
 - URLを教えるとクローラが収集**
 - Web2.0 Approach
 - User-defined Schema,
 - Ontologies and Dictionaries





現在のシステム: Dublin Core, ISO19115 & JMP2.0

ebRIM EO (Earth Observation) Profile

GEO GridとしてはISOからebRIMに重点を置いているが、両者を統合して検索できねばならない。(ebRIMからRDFへのAlignは議論がある)

属性間の対応関係を統一あるいは記述して異種データ統合を可能に(RDF)

<JMP><003>=<createddate>

緯度・経度・ポリゴンなども

検索側のスキーマと

データのスキーマを分離
(CSWの要求)

検索エンジン的な結果返戻の手法

Exactマッチだけでなく、
距離とランキングを考慮

マルチ・プロトコル:

OGC CSWのHarvestingと、
OAIのHarvestingをサポート

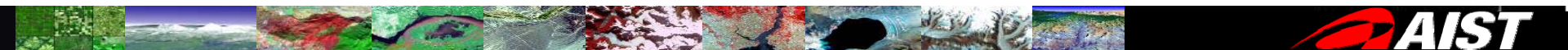
The screenshot shows a Mozilla Firefox browser window displaying a map of Japan. On the left, there is a search interface with the following fields and options:

- MENU
- カタログ検索
- 検索(最小化)
- 検索対象: JMP (dropdown menu)
- 検索開始日: (text input)
- 検索終了日: (text input)
- キーワード: 名古屋 (text input)
- 変量: (text input)
- 検索 (button)
- 全選択 追加 << >> (buttons)
- 100 of 1 (text)
- 20万分の1数値地図 幅「名古屋」 (checkbox and text)
- WMSレイヤー追加 (button)
- ベースマップ (button)
- 選択レイヤー (button)

At the bottom right, a WMS GetCapabilities window is open, displaying XML metadata:

```

WMS GetCapabilities
</Style>
<ScaleHint min="4.98902848429637e-09" max="49890.2848429637"/>
</Layer>
<Layer queryable="0" opaque="0" cascaded="0">
  <Name>boring</Name>
  <Title>Boring</Title>
  <Abstract>Boring points (point data)</Abstract>
  <SRSTitle>EPSG:4326</SRSTitle>
  <Style>
    <Name>default</Name>
    <Title>default</Title>
    <LegendURL width="20" height="15">
      <Format>image/png</Format>
      <OnlineResource xmlns:xlink="http://www.w3.org/1999/xlink" xlink:type="simple" xlink:href="http://geodata2.geogrid.org/mapserv/geod
    /g25k_tsukubaj?version=1.1.1&service=WMS&request=GetLegendGraphic&
  
```



研究開発(3)不均質科研

- 基盤A) 不均質なペタバイト級時空間センサデータの統合利用基盤
 - 産総研・筑波大・静岡大の共同研究
 - 時間解像度・空間解像度の異なる情報の利用
 - 条件に対して答えが返ってこない。あるいは膨大。
 - 解像度の異なる情報の付き合わせ:
 - 異種の情報の付き合わせ



問題のフレーム：画像系の情報とテキスト系の情報の連携・融合による知見の創出

画像データ・センサデータなど

Wikipediaから
イベントやキーワードの
緯・経度の抽出して
画像を発見(静大)

画像処理(変化や差分の検出など)
をトリガとしてテキスト系の情報の
発見・統合
そのためのイベントベースの
情報基盤?
(筑波)

メタデータの統合基盤
(AIST)
オントロジ・辞書などの利用
アノテーション

テキスト系の情報・辞書



メタデータの基盤について

- GEO応用におけるサービス・データのレジストリ
 - どこにどんなサービスやデータ、資源があるか。
 - OGCによる規格・ソフト
 - Webサービス規格・ソフト
 - Gridにおける規格・ソフト
- 利用者によるアノテーション追加
 - メタデータの作成追加、スキーマの定義・拡張
- 高機能なデータ検索・異種データ統合
 - オントロジの利用
 - 地球科学関連、国勢調査、地形、行政関連
 - 整備された辞書
 - いわゆる、SemanticWeb技術の利用
 - マッチメーカー・推論など

今後の課題

- GEO Grid ITとして

- ワークフローの実現

- Pegasus
- Kepler
- BPEL
- MapReduce

- ツールキットの整備と配布

- E-Science基盤として

- 大規模計算との**連携**・仮想化・クラウドへの展開
- デジタル・レポジトリやアーカイブとしての実現。
- 機関レポジトリとの統合
- 広域データベース**連携**・統合・メタデータ基盤