GEOSS Common Infrastructure (GCI) and the GEO Discovery and Access Broker (DAB)

Stefano Nativi

National Research Council of Italy (CNR-IIA)
The GCI revolutions

1st Revolution
[Service driven]

2nd Revolution
[Brokering pattern]

3rd Revolution
[Big Data]

4th Revolution
[... ]

Discovery
[functionality]

Discovery and Access
[functionality]

(re-)Use
[functionality]
Service-driven collaborative Approach

Virtual-Collaborative Management approach

Discovery [functionality]

[Governance]
Generally Recognized Barriers (especially for GEOSS)

- **PROBLEM**: service **users need to know** the nature and location of service providers,
  - making it **difficult to bind** and **dynamically change the bindings** between users and providers

- **SOLUTION**: The **broker pattern** separates users of services (clients) from providers of services (servers) by inserting an intermediary, called a broker.
  - When a client needs a service, it queries a broker via a service interface. The broker then forwards the client’s service request to a server, which processes the request
Collaborative-Acknowledged Management approach

Discovery and Access

Collaborative-Acknowledged Management approach

[functionality]
GEOSS Information System

More than 35 brokered systems

Data Providers Brokered (capacities, systems, networks, etc.)

GEO Home Page
GEOSS Portal
Systems Registry
Semantic engines

Private sector (testing, interoperability)
More than 35 brokered Data Providers – capacities, systems, Communities

Publish

About 14 Million (2 Million GEOSS Data Core) Discoverable and potentially Accessible first level resources (mix of data collections, datasets and individual images)

Contain [source: data providers]

More than 82 Million (more than 50 Million GEOSS Data Core) Discoverable and potentially Accessible individual resources (e.g. satellite scenes, rain gauge records)

Resources
### Ongoing interoperability tests:

- IGN
- UK data.gov
- Geoscience Australia
- GBIF
- DigitalGlobe
- FP7 GeoCarbon DBs
- e-GEOS Cosmo Sky Med
- …
Flexible, adaptable, and extensible interoperability Environment

Collaborative-Acknowledged Management approach

[Governance]
Interoperability Standards

Interdisciplinary and Global Interoperability

Standardization

Common Technological and Semantic baseline

Federation

Mediation/Brokerage

Global infrastructures (e.g. SoS, NoN)

Domain/discipline

Cohese infrastructures

Global infrastructures (e.g. SoS, NoN)
Big Data challenges for GEOSS

- **Volume:** Millions of discoverable (small & medium) products; Long EO time/space series, ...

- **Variety:** Different product types (data, services, models, documents); Data models; Protocols; Semantics; Granularity levels; Organizations; Maturity level, ...

- **Veracity/Validity/Value:** Evaluation support, Essential Variables, Discovery Ranking, User Feedbacks, ...

- **Velocity:** serve Countries with limited Internet access, Fast Discovery & Access, Data Transformation and Analytics

- **Visualization:** preview, Tiling services, visual exploration, ...

**Scalability and Reliability**
Big Data Challenges in building the Global Earth Observation System of Systems

Stefano Nativi, Paolo Mazzetti, Mattia Santoro, Fabrizio Papeschi, Max Craglia, Osamu Ochiai
Hybrid Cloud services
(integrating public and private clouds)

(Cloud-based) Software Ecosystem
GCI (hybrid) IaaS and PaaS

- Computing
- Storage
- Monitoring
- Auto Scaling
- Load Balancing
- Routing
- Clustering
GEO DAB in the Cloud

More than 32 virtual machines on a public cloud

GCI provides scalable and reliable services
Scalability benefits: Discovery Ranking Metrics

- A weight-based algorithm rewarding a set of criteria belonging to four high-level principles

Record score depends on 4 main Aspects:
- Record quality
- Data accessibility
- Textual constraints Matching
- Domain constraints Matching

Essential Variables come first

Metrics is: Configurable and Flexible
DAB Monitoring and Statistics

- **Statistics**
  - Year, month, week, day
- Most used **discovery constraints**
- Most used **remote data systems**
- Number of **discovery requests**
- Number of **access requests**
GEO DAB Authentication

- Support existing and well-used users' credentials (e.g. Facebook, Google, Twitter, etc.)
- Support OpenID and OAuth protocols
- Support GEO DAB credentials

Authentication Steps:
1. User accesses the Client Application and requires Login
2. The GEO DAB generates an appropriate request (OpenID, OAuth) to the Credential Provider Login web page (e.g. Google, Twitter, Facebook, etc.)
3. User executes the Login on Credential Provider web page
4. The provider authenticates the user and sends user info back the GEO DAB
5. The GEO DAB decodes user info and sends it back to the Client Application
6. The Client Application displays successful Login
GEO Discovery and Access Broker APIs

- GEOAPI are **high level client-side Open APIs** (Application Program Interface)
- DAB users are typically **software agents**, such as web-based or desktop client applications

Designed and developed in **JavaScript** to simplify the development of applications and clients making use of the DAB
GCI and DAB Achievements

- (From a “Catalog of Catalogs” to a) **multi-disciplinary Brokering Platform**
- (From discoverability to) **accessibility and harmonization services**
- (From an single infrastructure to a) **cloud-based software ecosystem**
- (From a virtual governance to a) **collaborative-acknowledged governance**
Thank you!

Questions?