ESSI Lab
http://essi-lab.eu

SEMANTIC BROKER FOR ENHANCING RESOURCES DISCOVERY

S. Nativi[1], M. Santoro[1], C. Fugazza[2], M. Craglia[2]

1. Institute of Atmospheric Pollution Research - CNR, Rome, Italy.
2. European Commission Joint Research Centre, Ispra, Italy.

Outline

- Rationale
- Geospatial Discovery Services
- Proposed Semantics-enabled Discovery Framework
- Experimentation
- Conclusion and Future Work

Rationale

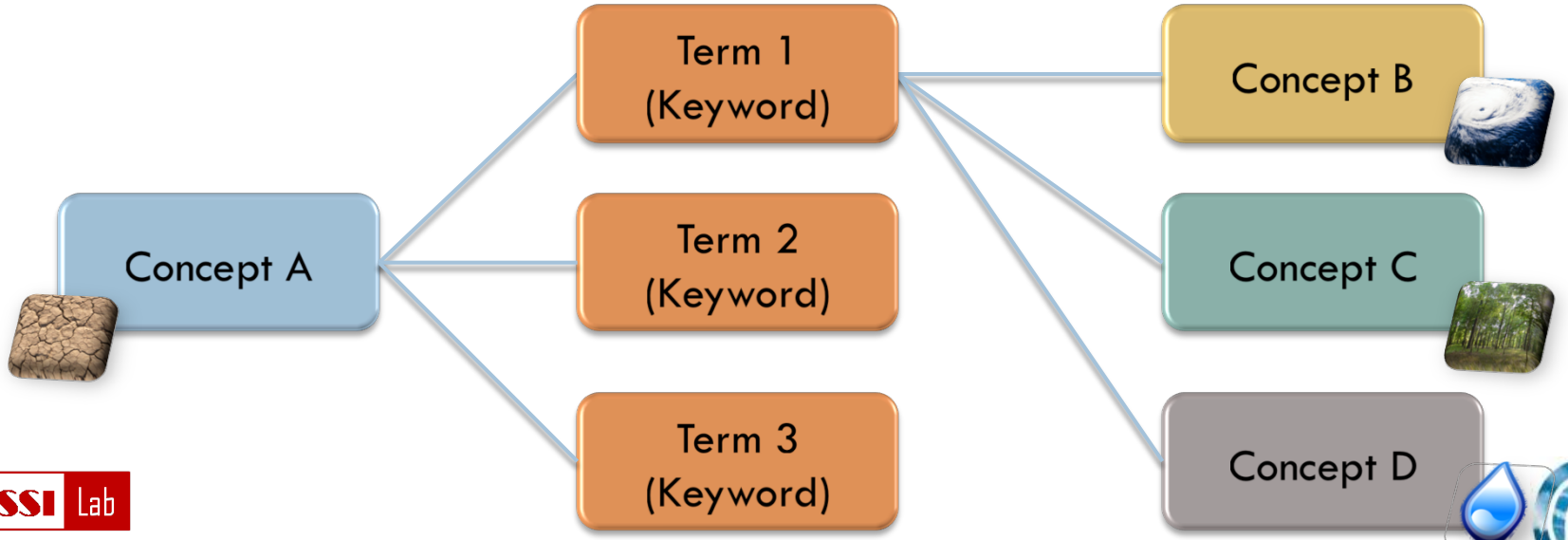
- Multidisciplinary Spatial Data Infrastructures (SDIs) manage large amount of heterogeneous geospatial data (from satellites, in-situ stations, marine cruises, etc.)
- Typical environment: heterogeneous resources (from different scientific domains) and data sources in a distributed environment
- Discovery solutions are still mainly based on syntactic matching rather than on conceptual matching
- There exist semantic gaps among geoscience domains
- Conceptual matching permits to overcome translations issues as well

Geospatial Discovery Services

- Discovery plays a central role in all SDI scenarios
- General frameworks for implementing discovery services are the Catalog, Inventory, Listing, etc. service (e.g. OGC Catalog Service Specification, THREDDS Data Server, etc.)
  - These provide specifications for geospatial resource discovery
- Possible constraints to build queries include:
  - Where
  - When
  - What
  - Who
  - How

The Need for Semantics

- Discovery solutions usually provide syntactic matching on the *what* constraint (keyword)
  - This limits the effectiveness of geospatial data discovery due to the limitations of searching words in free text instead of searching through matching concepts



The Need for Semantics

- Discovery solutions usually provide syntactic matching on the *what* constraint (keyword)
  - This limits the effectiveness of geospatial data discovery due to the limitations of searching words in free text instead of searching through matching concepts
- Multidisciplinary interoperability solutions must be able to address semantic gap/mismatch among disciplines
- To achieve this we propose a *Discovery Augmentation* approach to enhance discovery functionalities

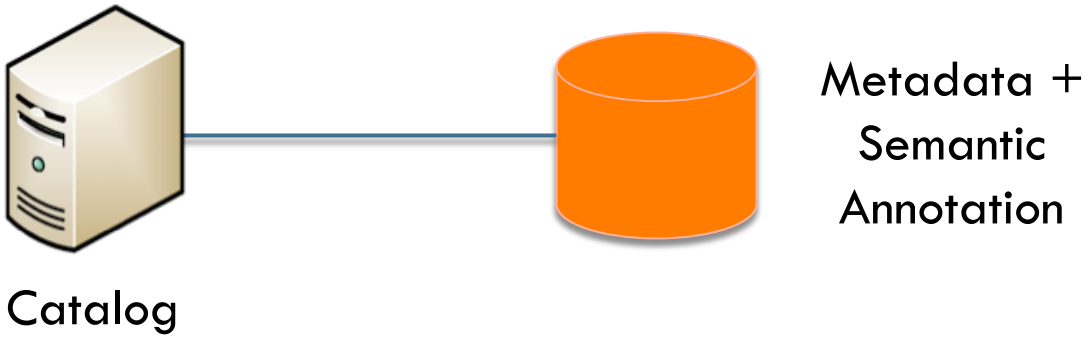
Discovery Augmentation

- To achieve the enablement of semantic-aware discovery we propose a Discovery Augmentation approach
  - Enrich searchable information associated with geospatial resources

- Several approaches can be followed to cope with the required additional description
  - Provider-based
  - User-based
  - Third-party

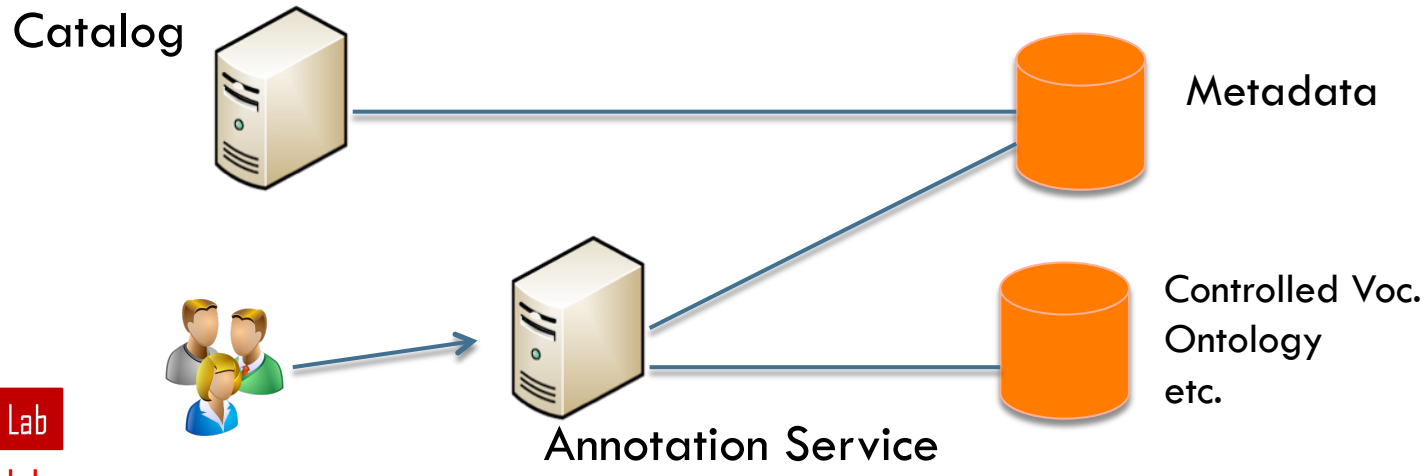
Provider-based Approach

- Data providers enrich resource metadata by adding related semantic information based on controlled vocabularies or even on full ontologies
  - Very accurate because it allows to directly perform queries on the semantic content
  - Low extensibility
  - Need to change present catalog specifications



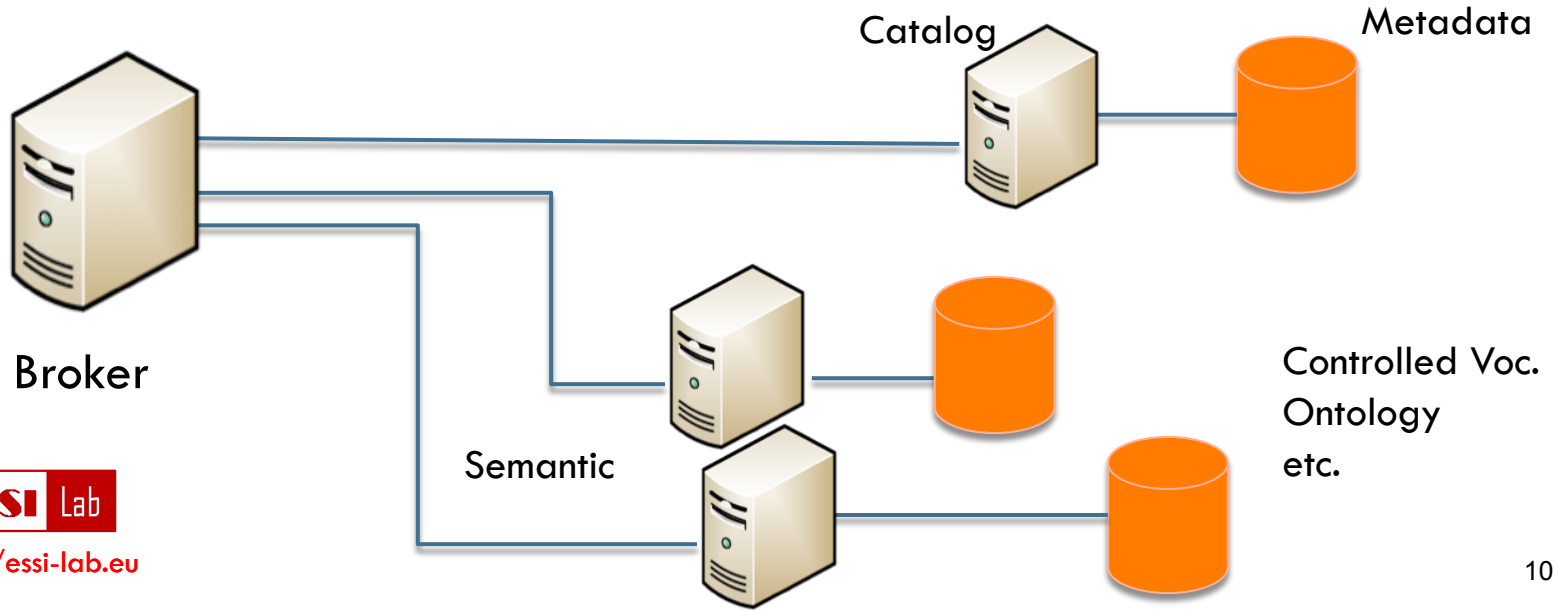
User-based Approach

- Delegate to users the enrichment of resource description by adding the so-called resource annotation capability
  - Scalability: this approach distributes the task of enriching large repositories of metadata to a wide range of users, scaling and making potentially use of a much higher amount of knowledge.
  - Concern: Quality Control



Third-party Approach (Brokering)

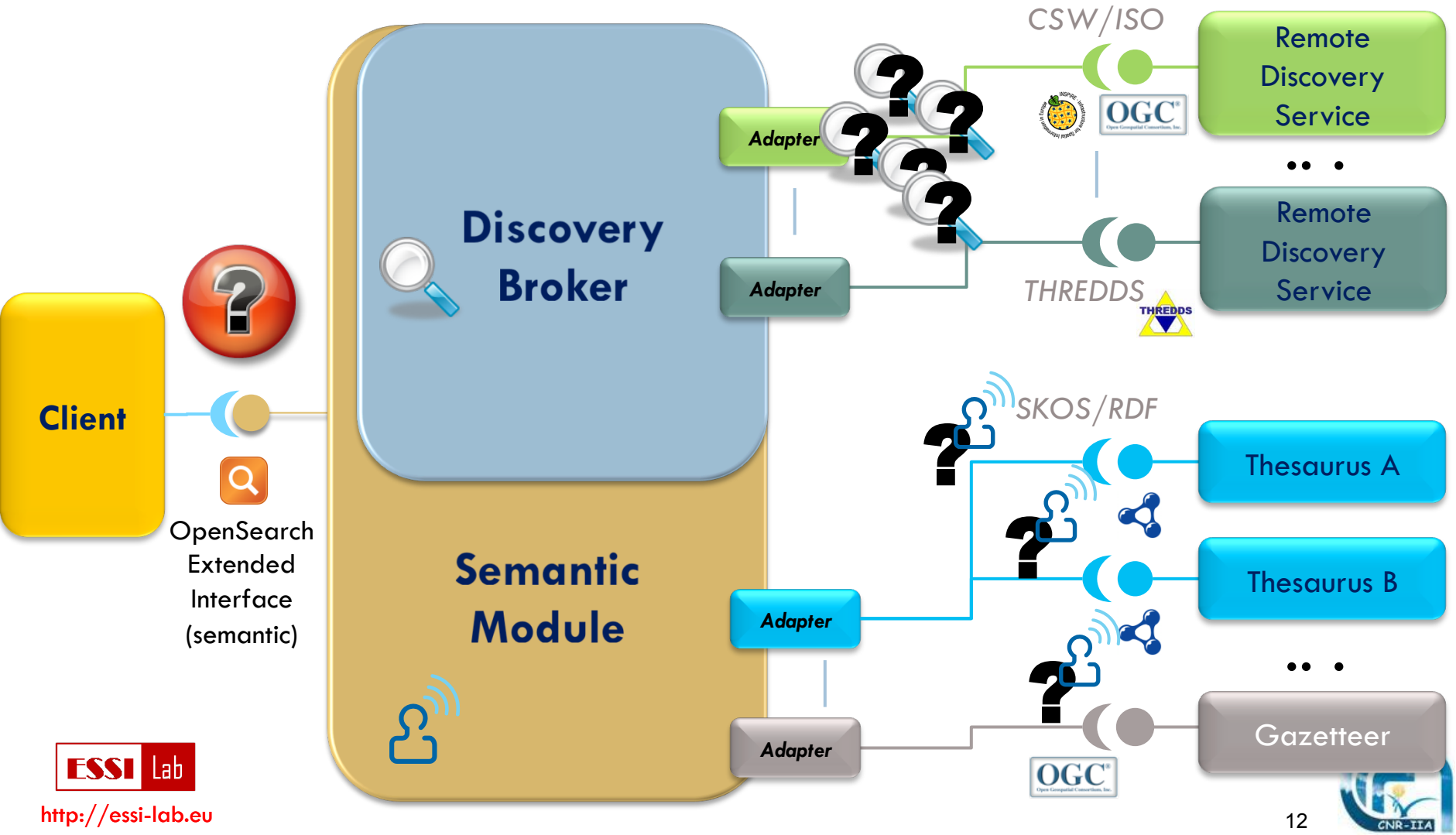
- Automatic semantic characterization from existing information and meta-information
  - Builds on existing systems (Discovery and Semantic services)
  - Extensibility: the business logic necessary to classify resources is concentrated in a separate component (provided by a third party) that can be adapted to satisfy new requirements without affecting the other existing systems
  - Flexibility: easy to add new heterogeneous semantic services
  - Concern: not as accurate as the provider-based one, mainly because the automatic mapping might be inaccurate in some cases



Architectural Principles for Semantic Broker Development

- **Separation of Concerns**
  - Each component is dedicated to one well-specified functionality: geospatial queries, semantic queries, aggregation of results, user interface.
- **Layered architecture**
  - Functionalities are layered according to their abstraction level
- **Flexibility**
  - Other services might be easily included in the proposed architecture (e.g. Gazetteer)
- **Extensibility**
  - New functionalities and/or association strategies can be implemented in the Broker without affecting other systems

Semantic Discovery Broker



Discovery Styles - 1

- Automatic query expansion
  - the query keywords (the "what" constraint) are **"expanded" with related concepts retrieved** from the set of federated semantic services. A default expansion regards the multilinguality relationship;
  - **Based on the discovered concepts,** a set of queries is created and submitted to the federated catalog services;
  - The Broker performs a **"smart" aggregation of the queries results** and provides them back to the client

Discovery Styles - 2

- User assisted query expansion
  - The user **browses** the federated semantic repositories and **selects the concepts of interest;**
  - The Broker creates the set of **geospatial queries based on the selected concepts** and submits them to the federated catalog services;
  - The Broker performs a **"smart" aggregation of the queries results** and provides them back to the client.

Prototype

- We developed a prototype in the FP7 EuroGEOSS Project
- The EuroGEOSS Semantic Discovery Broker uses the following set of Thesauri, aligned and hosted by the EC Joint Research Centre:
  - The General Multilingual Environmental Thesaurus (GEMET)
  - The INSPIRE Feature Concept Dictionary and Glossary
  - The ISO 19119 geographic service taxonomy
  - The GEOSS Societal Benefit Areas
  - The GEOSS Earth Observation parameters
  - The GEOSS AIP-3 Water Ontology
  - The thematic EuroGEOSS Drought Vocabulary
  - The Global Change Master Directory (GCMD) scientific keywords

Prototype


eurogeoss-broker.eu
www.eurogeoss-broker.eu
GEOSS Discovery and Access Broker
EuroGEOSS
Map
Concept Network
Clean Selection
More Specific
More General
Corresponding
Related
Concepts
Deuterium
Tritium
Chemical Parameters
Water Temperature
Snow Water Equivalent Air
Forms of Water
water stored in reservoir
Snow on Ground
Snow Water Equivalent Ground
Results Matching Keyword
water
water scarcity
Water
Results Matching URI
Water cycle research
water deficit
water stress
water runoff
Water
reservoir
Results Matching URI
Results Matching Keyword
Query constraints selection
My resources
- Keyword
water
Hide Semantic Options
Related Terms
Browse Concept Network
Selected Nodes
Selected Nodes
water runoff
water scarcity
Show Quality Options
Location
Enter a location name (case is ignored), e.g.: europe,italy,rome,etc...
Selected area
Overlaps
Contains
Disjoints
Click and Drag on the map holding the Shift key to select an area
- Time
From:
To:
- Results per page
10
Search results
Search results - All: 553
water scarcity
water runoff
Access/Use Constraints
Title
Helpers
Mean Monthly Water Balance Model Runoff 0.5-Degree 1986 (All Months)
Mean Monthly Water Balance Model Runoff 0.5-Degree 1987 (All Months)
Mean Monthly Water Balance Model Runoff 0.5-Degree 1988 (All Months)
Add

Experimentation

- We successfully tested the Semantic Discovery Broker in two GEOSS AIP-3 Use Scenarios
  - **e-Habitat Use Scenario:** multidisciplinary scenario where Biodiversity and Climate Change resources are discovered and accessed in order to run a model predicting habitat changes in a Protected Area
  - **EDO-based Comprehensive Drought Index Use Scenario:** multidisciplinary scenario where a decision maker needs to discover drought related resources in order to assess the effects of a drought hazard
- Presently, the Semantic Discovery Broker is part of the Brokering framework deployed and operated in the GEOSS Common Infrastructure

Technology

- The Semantic Discovery Broker and the GEOSS Brokering framework are based on technology solutions developed by the ESSI-Lab of the Institute of Atmospheric Pollution of the Italian National Research Council (IIA-CNR)
  - GI-cat (Broker Discovery Service)
  - GI-dac (Semantic Module)
  - GI-axe (Broker Access Service)

Conclusions and Future Work

- We have designed and implemented a loosely-coupled framework to integrate traditional geospatial discovery services with heterogeneous semantic services
- We tested the framework in two GEOSS AIP-3 multidisciplinary use scenarios
- The Semantic Discovery Broker is now part of the GCI
- Further work will mainly focus on the
  - Implementation of new adaptors for semantic services
  - Enhancement of the automatic query expansion
  - Strategies to improve the ResultSet presentation
  - Integration of automatic alignment tools

Thank you!