A Framework for Ontology-Supported Intelligent Geospatial Feature Discovery Services

Liping Di, Peng Yue, Peisheng Zhao, Wenli Yang, and Weiguo Han

Center for Spatial Information Science and Systems (CSISS)

George Mason University

10519 Braddock Road STE 2900, Fairfax, VA 22032

Introduction

- Rapid increasing in the high resolution remote sensing images
- Analysis of geospatial imagery for feature discovery
  - a promising approach for characterizing Weapons of Mass Destruction (WMD) (including nuclear) proliferation
- The overwhelming volume of routine image acquisition has greatly outpaced the increase in the capacity of manual interpretation by intelligent analysts
- An automated system, which automatically identify suspicious WMD proliferation sites in images for intelligent analysts for further investigations, can
  - significantly reduce the workload of human intelligent analysts and increase the possibility of prompt detection of WMD proliferation.

The Problems

- Current methods in the geospatial image mining and feature extraction
  - detect isolated or elementary features, such as buildings and roads
  - Classification is often performed on per-pixel basis
  - region-based characterization has received increasing attention in the recent years
  - mainly exploit image features, such as, color and texture, and, to some extent, size and shape.
- Ignore important spatial (topological) relationships
- Geospatial images contain complex (compound) features and patterns.
  - e.g., WMD: buildings for hosting fuel concentration machines, cooling ponds, and transportation railways
- The identity of a compound geospatial feature can be derived from the spatial relationship among the elementary elements.

Complex Features

- A manufacturing facility consists of a group of elementary ground features (e.g., buildings for hosting fuel concentration machines, cooling ponds, transportation railways, fence, etc).
  - containment buildings are surrounded by fence and near to the cooling tower and switch yard
  - buildings are near to both the field and court

(a) Nuclear Power Plant : **C**-Containment Building, **S**-Switch Yard, **CT**-Cooling Tower, **F**-Fence

(b) School: **B**-Building, **F**-Field, **T**-Tennis Court

The Research Objective

- An ontology-supported approach for discovering complex features in a service oriented environment
  - The ontology for facilities helps find compound features that contain the specified spatial relationships among elementary features.
  - The use of service computing significantly enhance the ability to use online/near-line data over the Web and allowing the widespread automation of data analysis and computation.

Technology Foundation

- The geospatial interoperability standards developed by ISO, FGDC, and OGC;
  - enabling the development of chainable services
- Interoperable geospatial web services technology.
  - providing the framework for deploying distributed services.
- Ontology, Semantic Web and Artificial Intelligence (AI) technology
  - enabling the automatic construction of solutions to users' questions.

The Key Research Issues

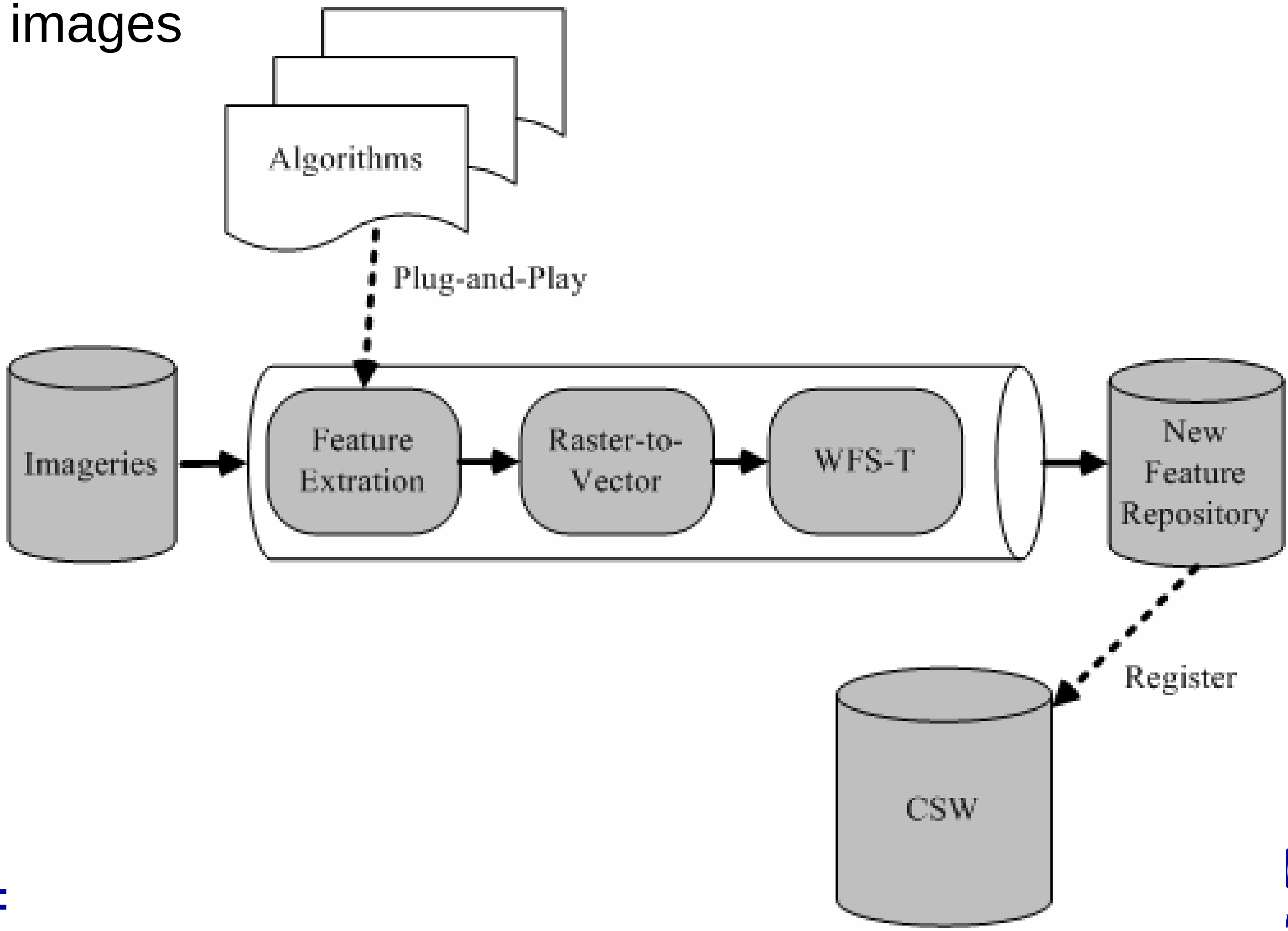
- Discovery of elementary geospatial features
- Geospatial ontologies for compound features
- Automatic decomposition of feature detection tasks into workflows
- Intelligent service framework for complex feature discovery

Discovery of elementary geospatial features

- Discovery of elementary features can be conducted by
  - 1) either performing new feature extraction from high resolution remote sensing images
  - 2) or accessing existing feature repositories.
- The Catalogue Service for the Web (CSW) from the Open Geospatial Consortium (OGC) can be used by front-end users to find the features of interest.

Feature Extraction and Registration

Using Web services to make existing algorithms capable of plug-in-and-play for on-demand feature extraction from images



Geospatial ontologies

- Encoding the knowledge of the subject matter experts
- Allowing automated inference and retrieval of knowledge items to facilitate matching between the subject area concepts and classes of the features extracted from the imageries

Geospatial ontologies for compound features

- Spatial and spatiotemporal ontologies for compound geospatial features associated with interested facilities.
  - For example, spatial predicate such as surrounded by, near, and cross can be created in the ontologies.
- Ontologies for describing semantic concepts and assigning semantic labels to compound geospatial features and their constituents for application domains such as manufacturing domain scenarios.
  - The elementary features, such as building, fences, bridges, railway, railway stations, and airports, can be defined here.

A Layered Approach can be used

- Upper-level ontologies such as DOLCE or BFO
- Concepts specific to geographic domain and spatial relations
  - Using the SWEET ontology and Ordnance Survey Ontologies
- Application Ontologies
  - Ontologies pertaining to manufacturing facilities can be developed by extending existing ontologies.

Task Decomposition for Complex Feature Discovery

- The task of detecting a manufacturing facility in an intelligent system can be decomposed into a series of queries of geospatial features based on the spatial relationship among the features.
- Examples of the queries are:
  - Find groups of buildings surrounded by fences
  - Find a building near metro station in a specific city
  - Find roads cross parks

Query Generalization

- The queries are constructed based on spatial relationships between two groups of geographic features (binary query).

- A generalized form of the query:

Find [<quantifier1>] <feature1> <operator> [<quantifier2>] <feature2> [in <spatial area>] [at <time>]

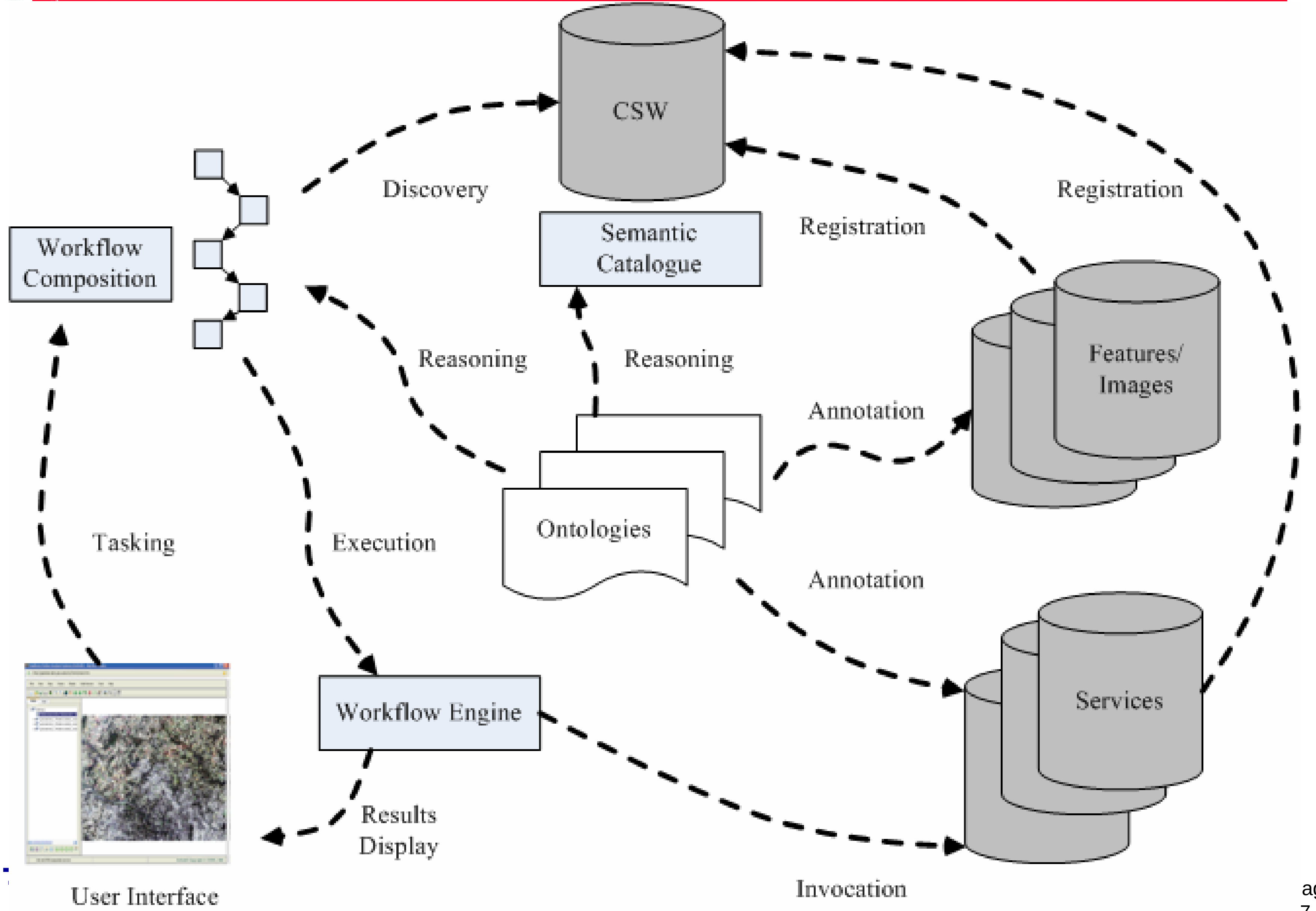
Complex queries

- A more complex query may involve multiple geographic features.
  - For example: *Find a group of buildings surrounded by fence and near a railway station*.
- A generalized form
  - *Find (Find groups of building surrounded by fence) near a railway station.*

Decomposition Process: an example

1) Users submit a request to detect an instance of facility whose concept has been described in an ontology;
2) The request is converted to a complex spatial relationship query based on the constituent features and their spatial relationship defined in the ontology for manufacturing facilities;
3) The complex query is transformed into a chain of the binary queries;
4) The chain of the binary queries is used to construct a workflow;
5) The workflow is transformed into a service chain, consisting of services and data in the CSW.
6) The workflow engine executes the service chain to generate the answer, which is displayed in the graphical user interface overlaid with the source images to show locations of matched geographic features.

System Framework

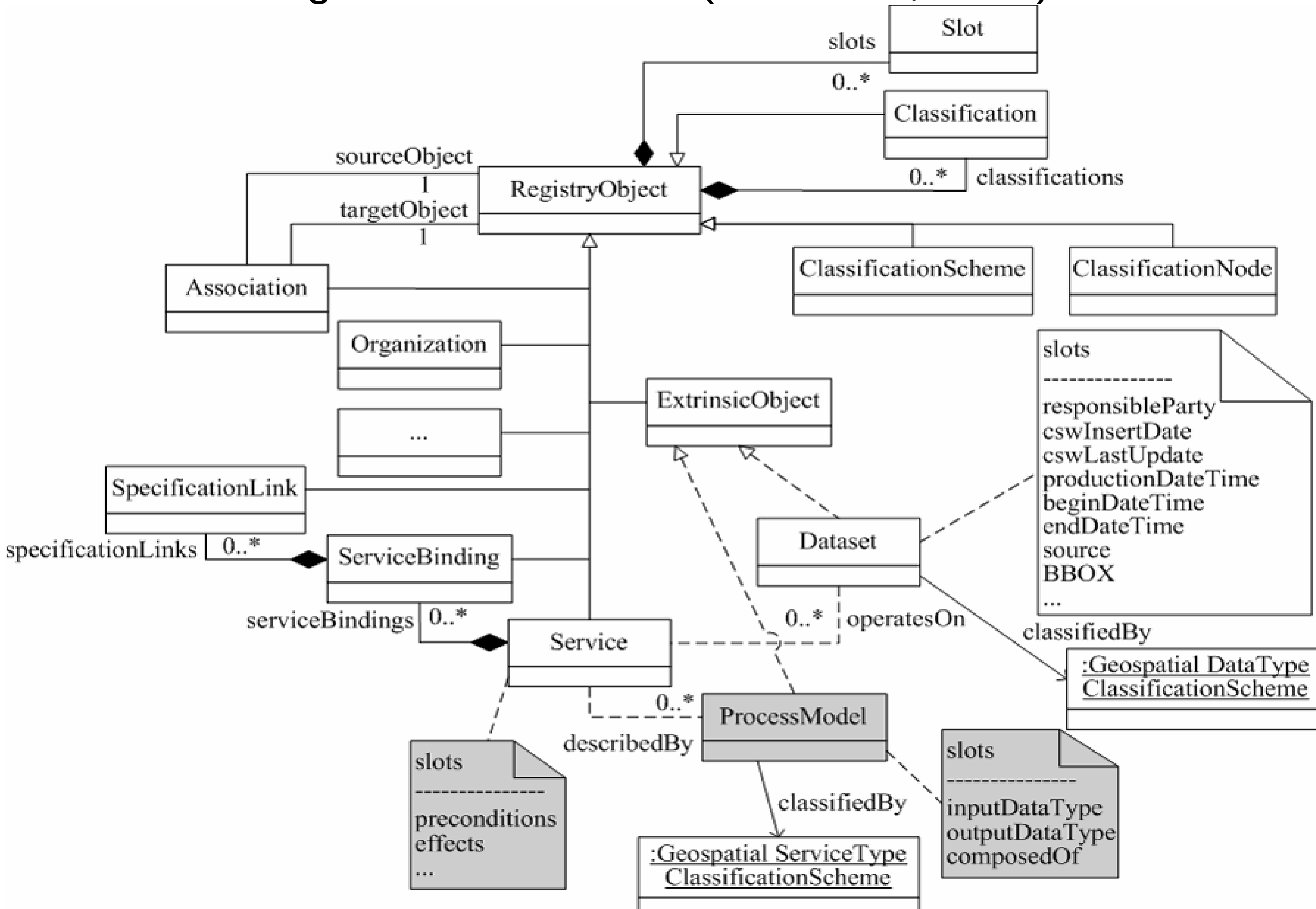


Ontoloiges

- The system can be used in different applications as long as ontologies for such applications are available.
- The ontologies are described using Web Ontology Language (OWL). The inference engine is used to find all relevant entailments.

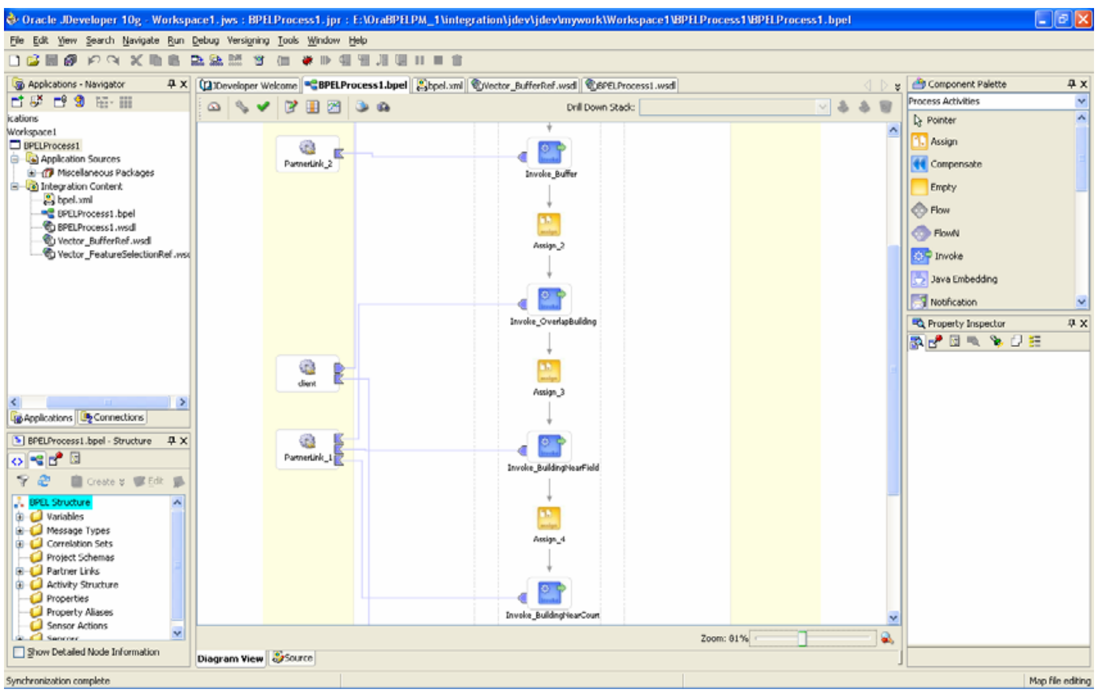
Catalogue

- Extending a semantic CSW (Yue et al., 2011)



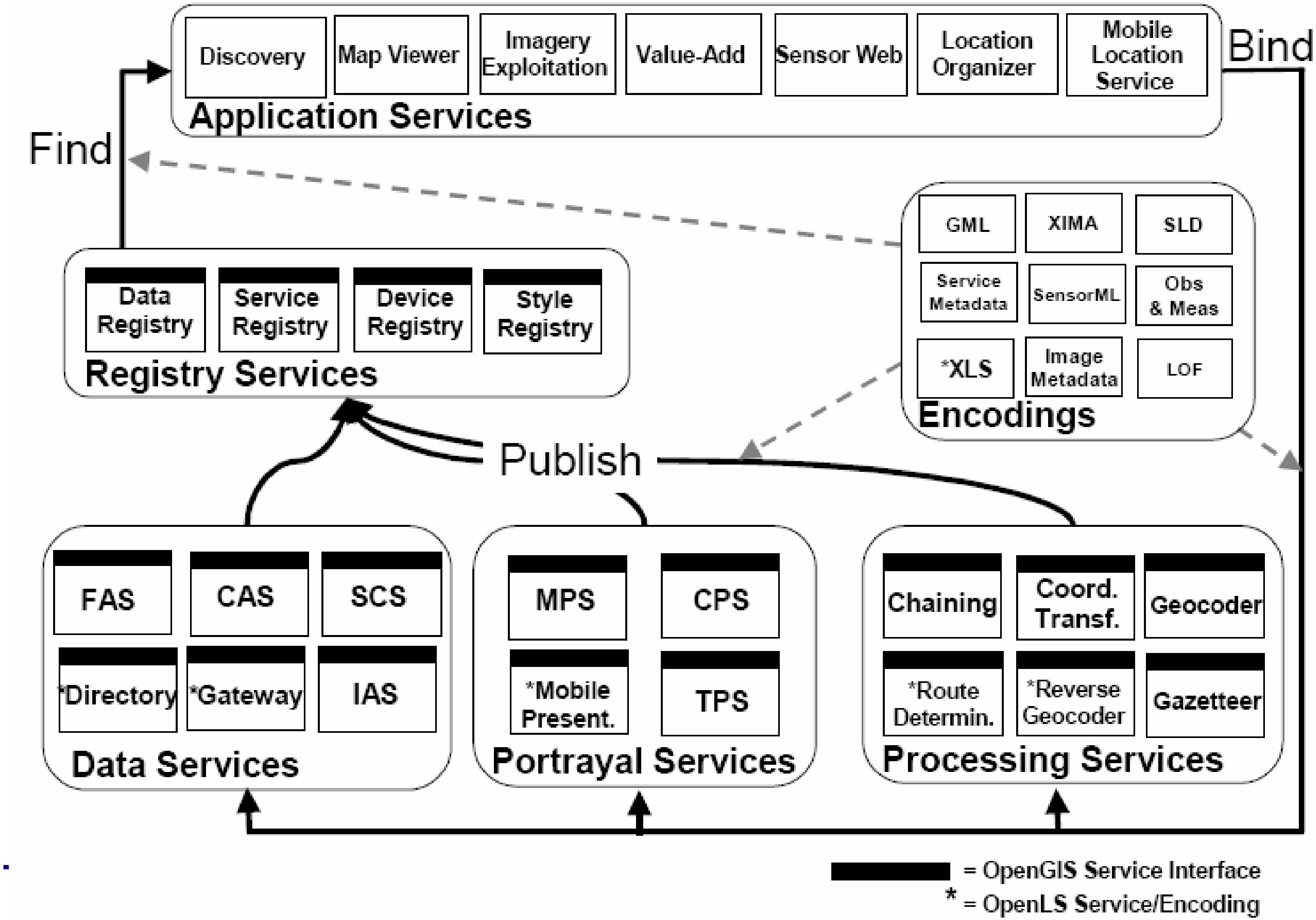
Workflow

- The facility detection task can be decomposed into a series of binary queries (spatial operations) on elementary features using the ontological reasoning and planning in the workflow composition service.
- The series of operations are encoded as a service chain, which will be executed by a workflow engine.



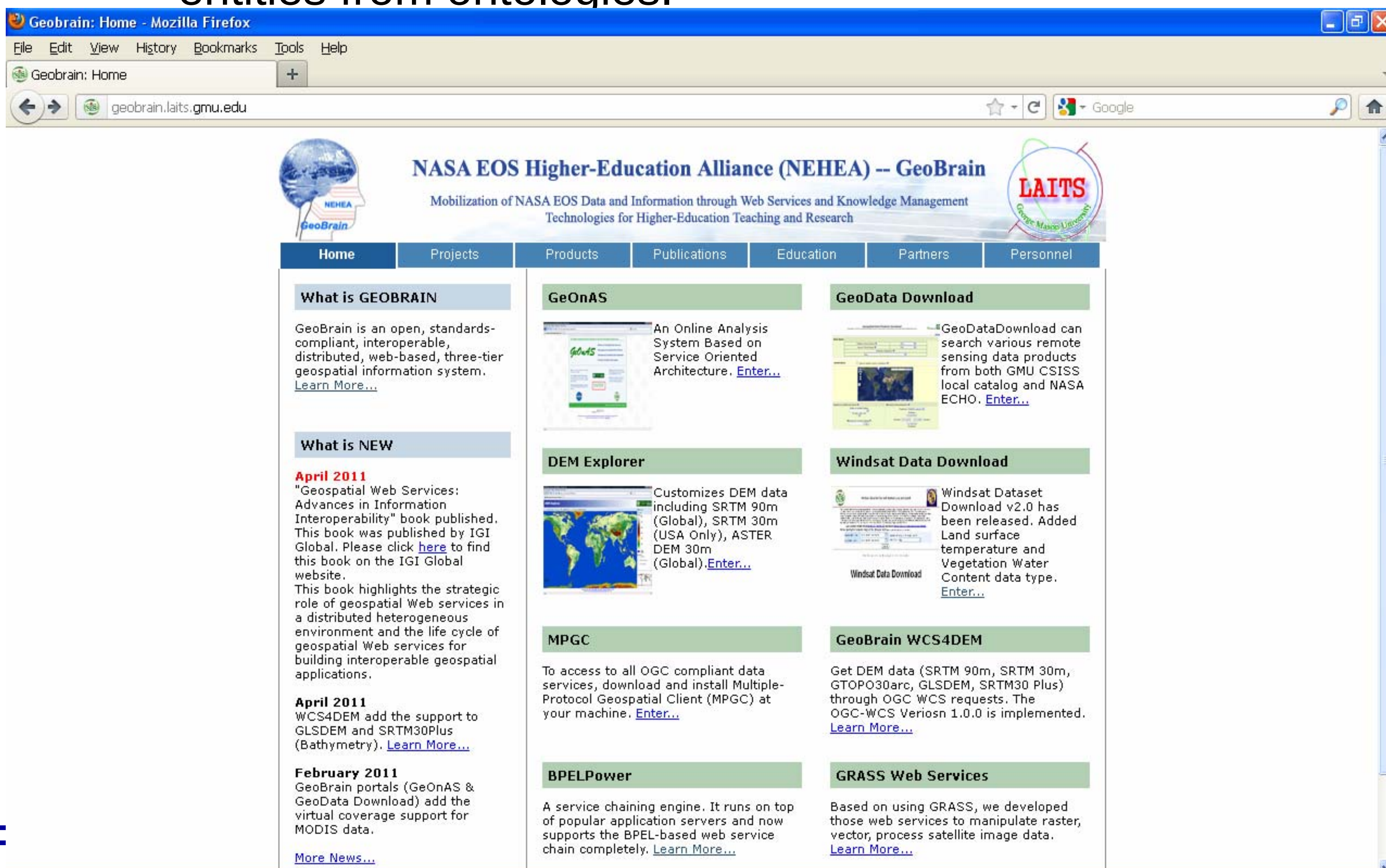
Geospatial Interoperable Standards in the System

- Open Geospatial Consortium (OGC) OWS Framework(OGC,2003)



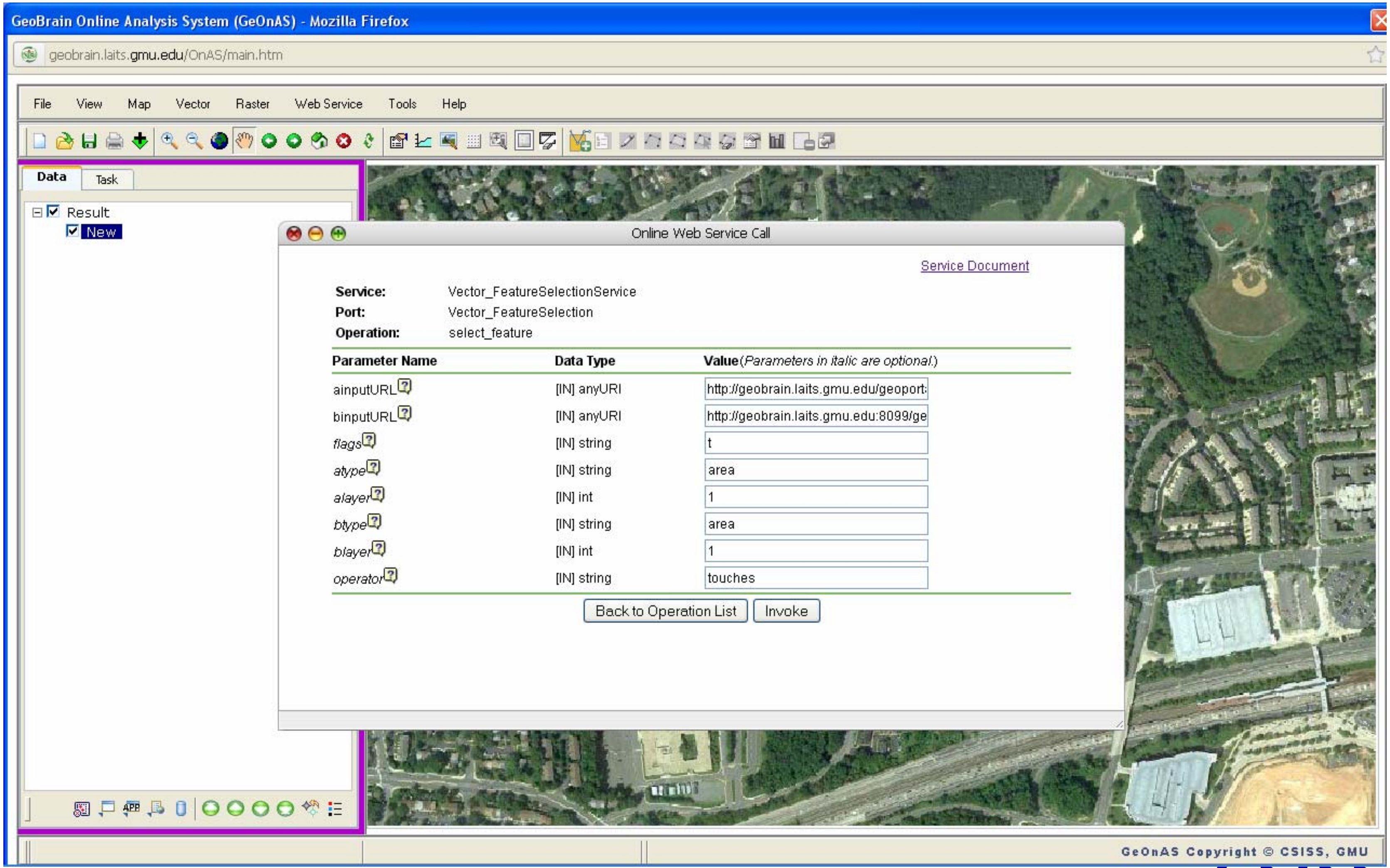
Data and Services

- GeoBrain (http://geobrain.laits.gmu.edu/)
  - The semantics for data and services are annotated using entities from ontologies.



User Interface

- GeoBrain Online Analysis System (GeOnAS)



Discussion

- Currently extraction of semantic information and semantic labeling of the features - a one-step approach
  - use training data in the form of image segments with known objects and then use various statistics to match training data with the imagery.
  - such one-step approaches are likely to fail when there are subtle differences between the complex features on an image.
    - presence of an industrial chimney is a salient feature distinguishing nuclear power plant from a coal-firing plant. However, chimney is a relatively small feature in the planar view that is unlikely to produce distinguishable effect in matching statistics.
- The use of formally encoded semantic information in the form of ontologies can solve these problems by explicitly identifying salient features of the compound objects and their constituents.

Discussion

- The approach described in this paper is a two-step approach, with
  - step 1 to identify the location and type of elementary ground features (such as building, road, etc) from high-resolution images, which has mature technology already, and
  - step 2 to extract high-level semantic information (such as nuclear fuel concentration facilities, nuclear test sites, missile test sites, etc) through discovering compound ground features based on spatial (topological) relationships among the elementary features.

Conclusions and Future Work

- The development of the system follows the SOA paradigm and using Web services and Semantic Web technologies.
- It allows discovery and composition of features and services in a distributed environment.
- The automatic discovery of compound features using workflow composition and execution can reduce the workload of human intelligent analysts and provide valuable information in decision making.
- The next step will be the proof-of-concept implementation, and evaluation the ontology approach.

Q&A

Thank you!