Envisioning a Future of Computational Geoscience in a Data Rich Semantic World

Praveen Kumar[a], Mostafa M. Elag[a], Peishi Jiang[a], Luigi Marini[a], Scott D. Peckham[b]
Leslie Hsu[c], and Kim Miller[c]

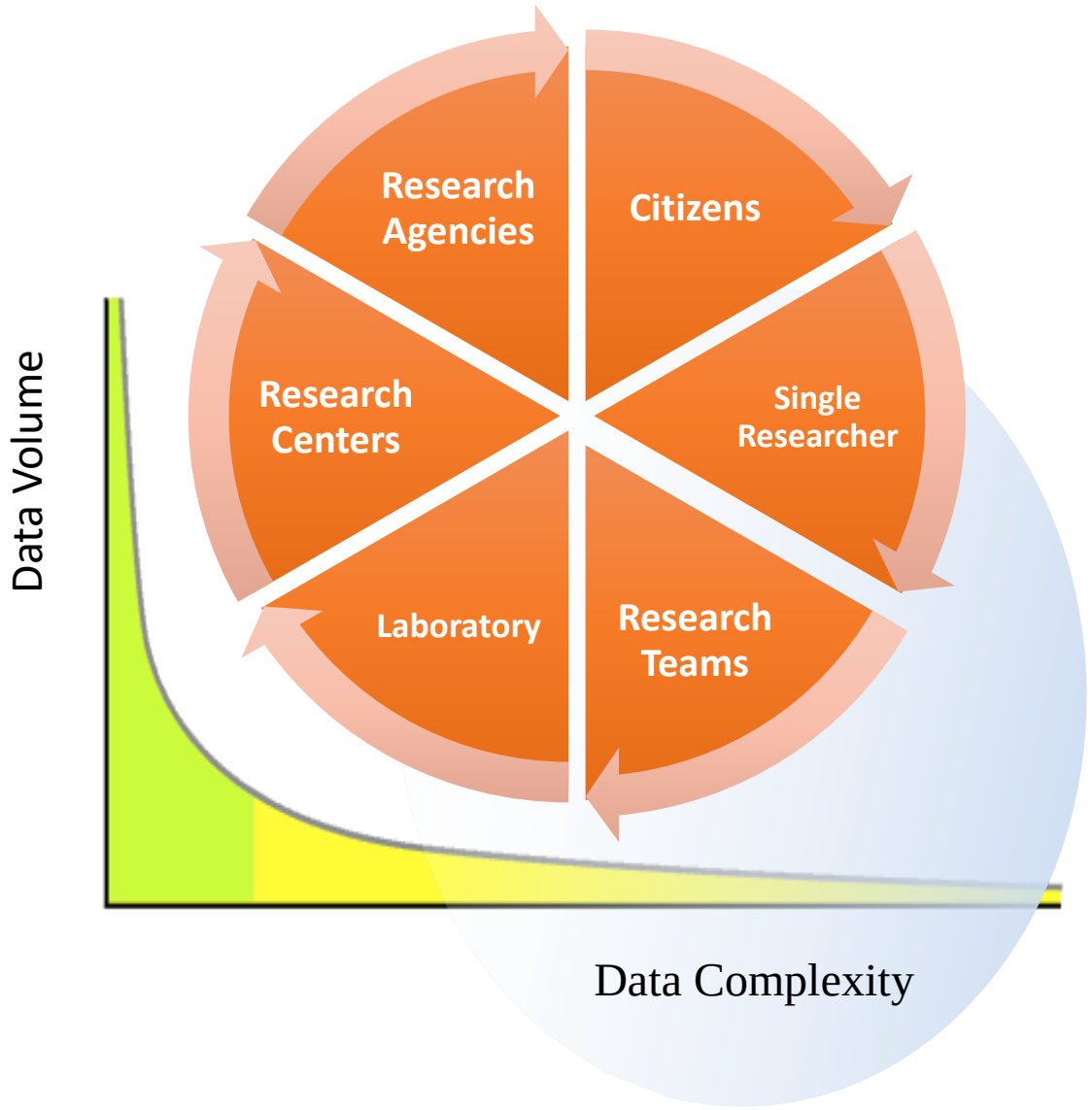
[a]University of Illinois at Urbana-Champaign

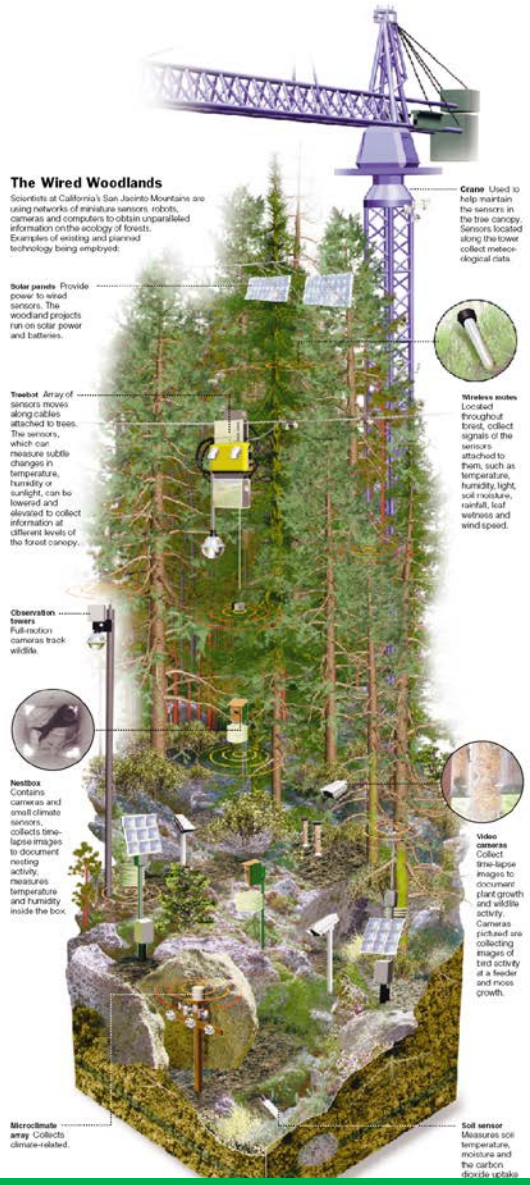
[b]University of Colorado Boulder

[c]Columbia University

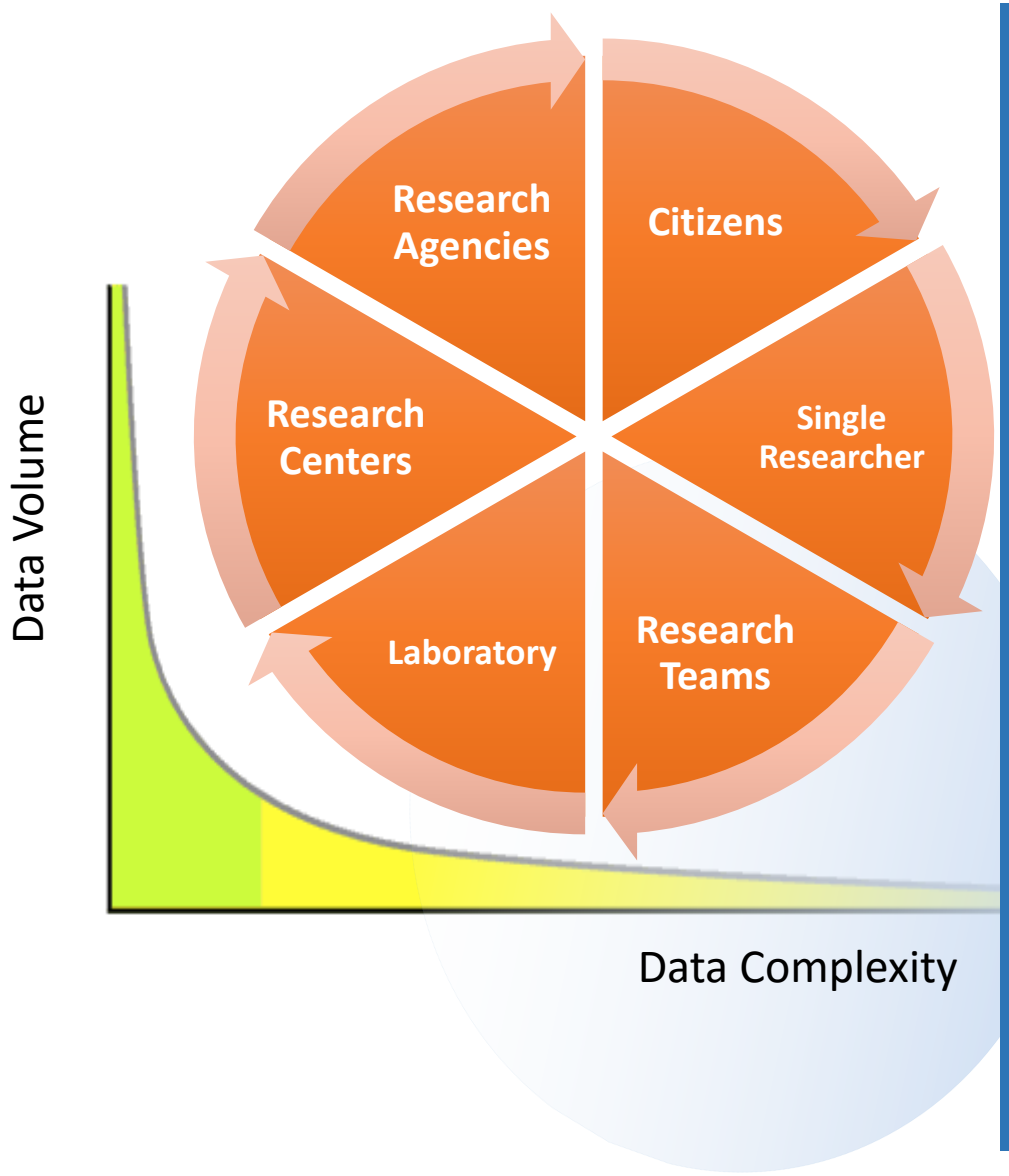
SEAD | Sustainable Environment Actionable Data

Complexity of Data: Synthesis and Reuse

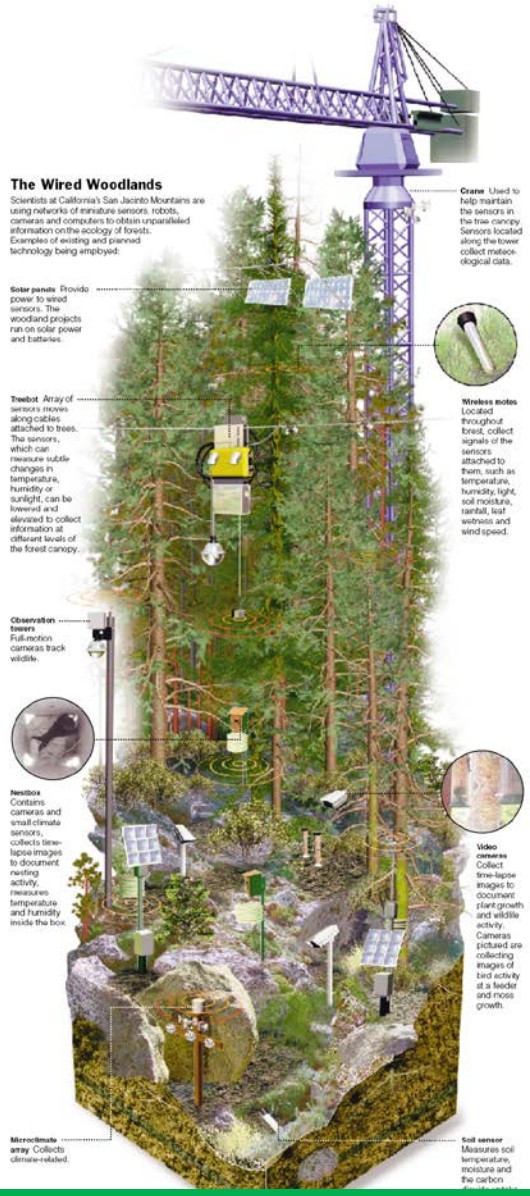




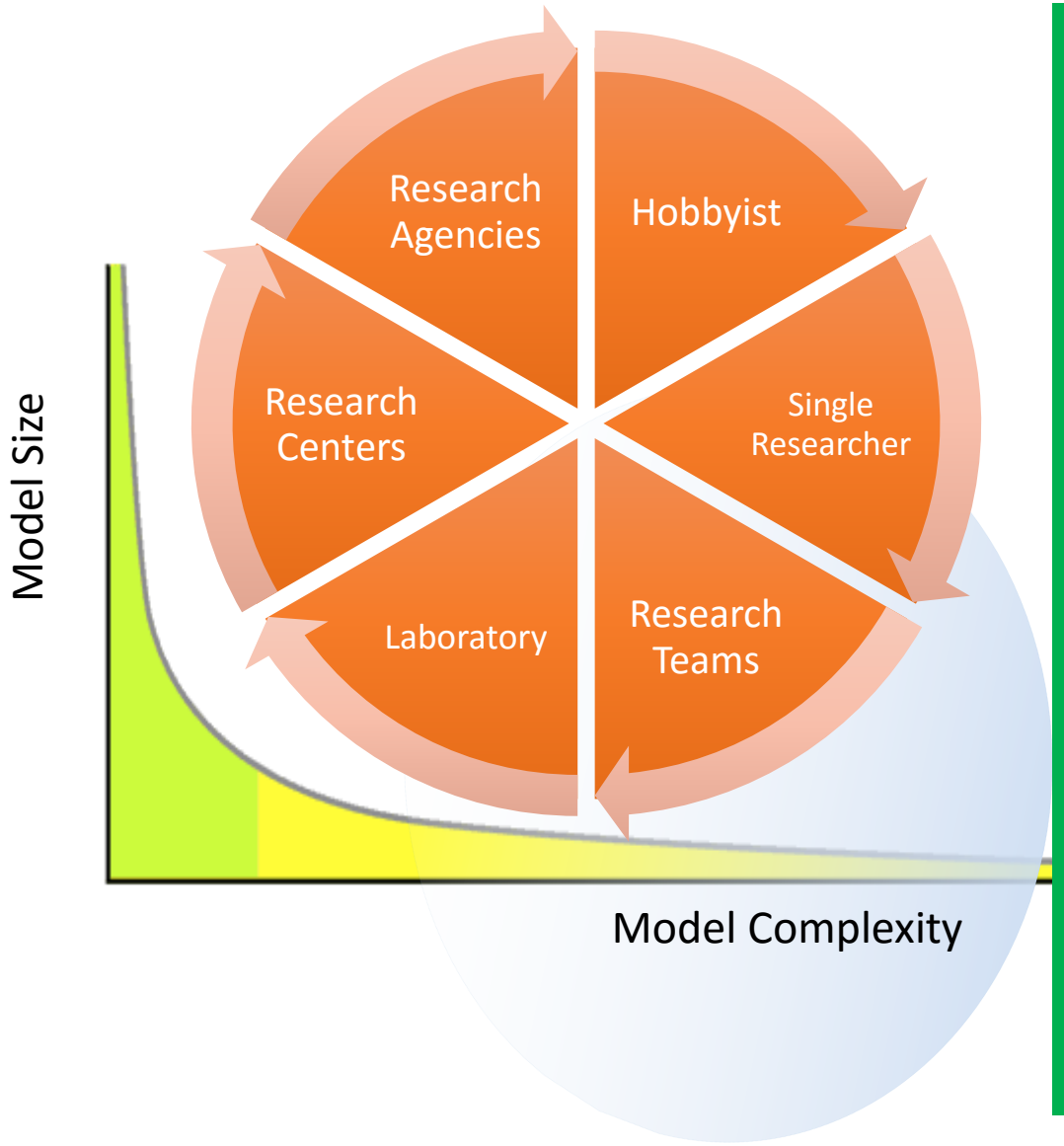
Complexity of Data: Synthesis and Reuse

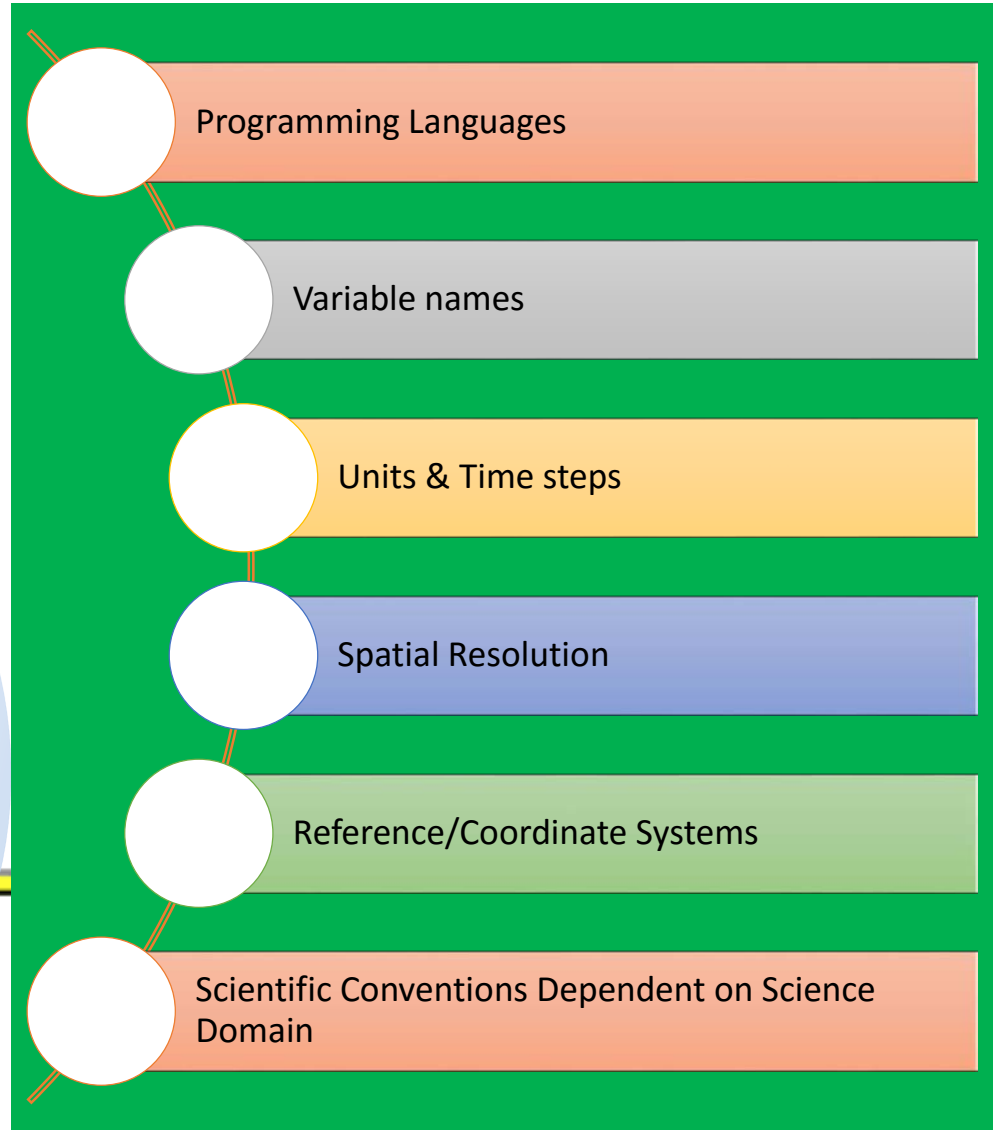


- **Collection**
  - Methods, instruments,…
- **Domains**
  - Physical, biological, chemical, remote sensing, …
- **Metadata**
  - Need I say more?
- **Storage** - Isolated, unorganized, distributed, networked, heterogeneous, …
- **Usage**
  - Analyses, modeling, …
- **Life cycle** - Use and hide, use and lose, use and share, use-share-reuse, use-share-archive-reuse, use-share-archive-enhance-reuse,…



Complexity of Models: Interoperability and Coupling





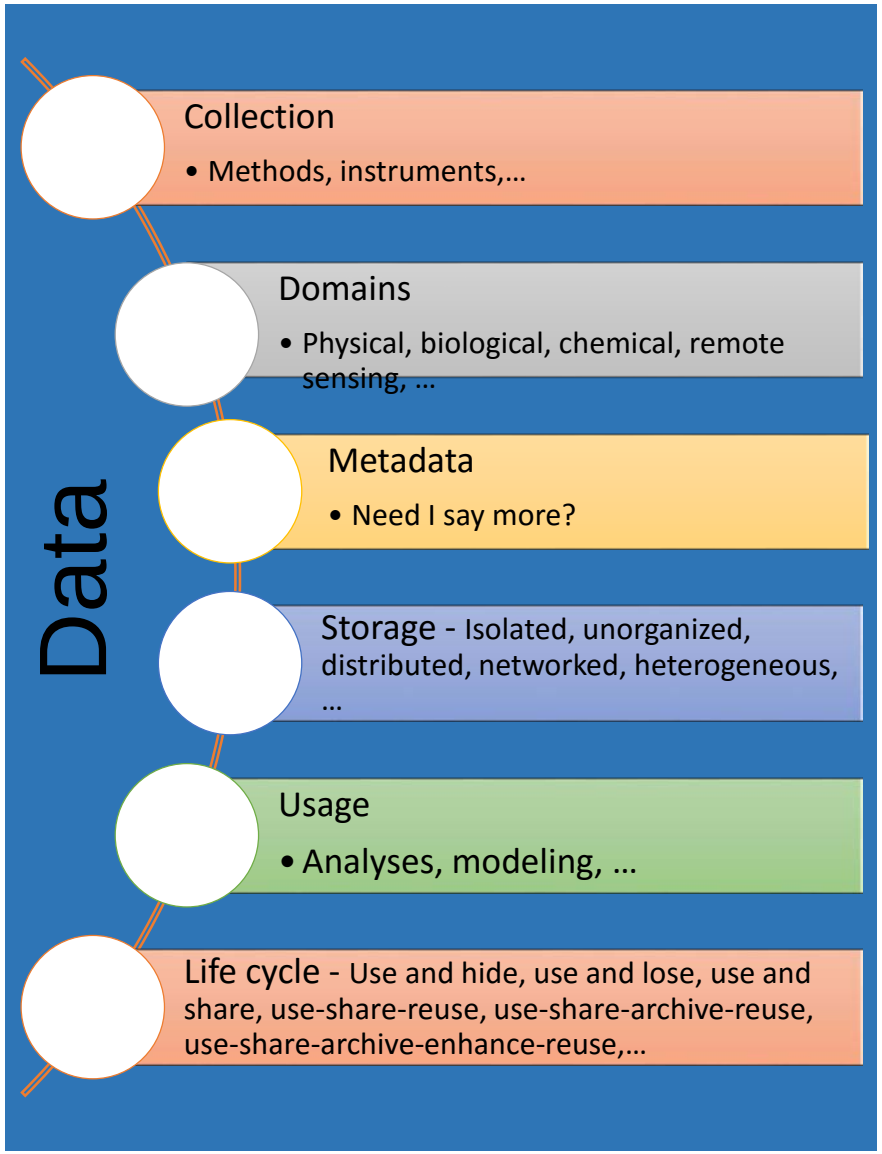
Modeling frameworks

Standalone Models

Model-as-a-Service (MaaS)

Modules and Libraries

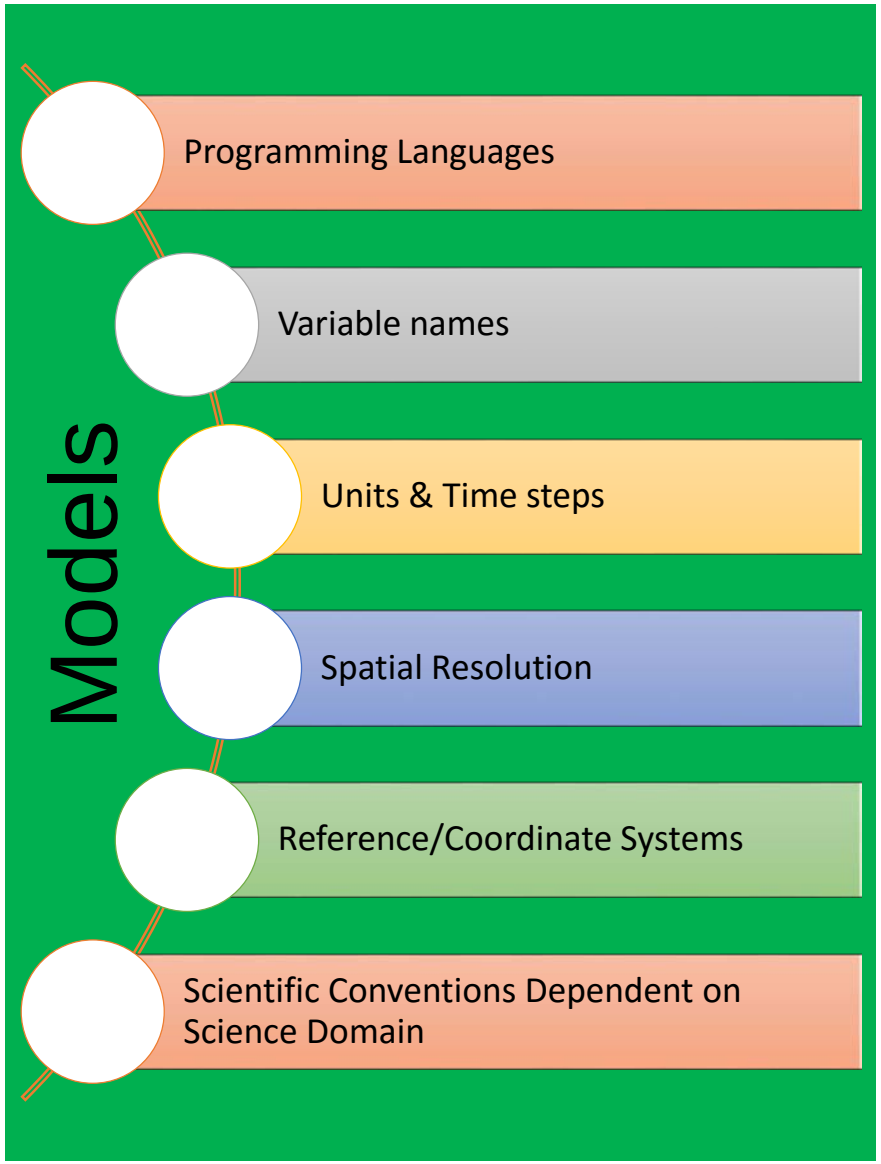
Data-Model Interoperability



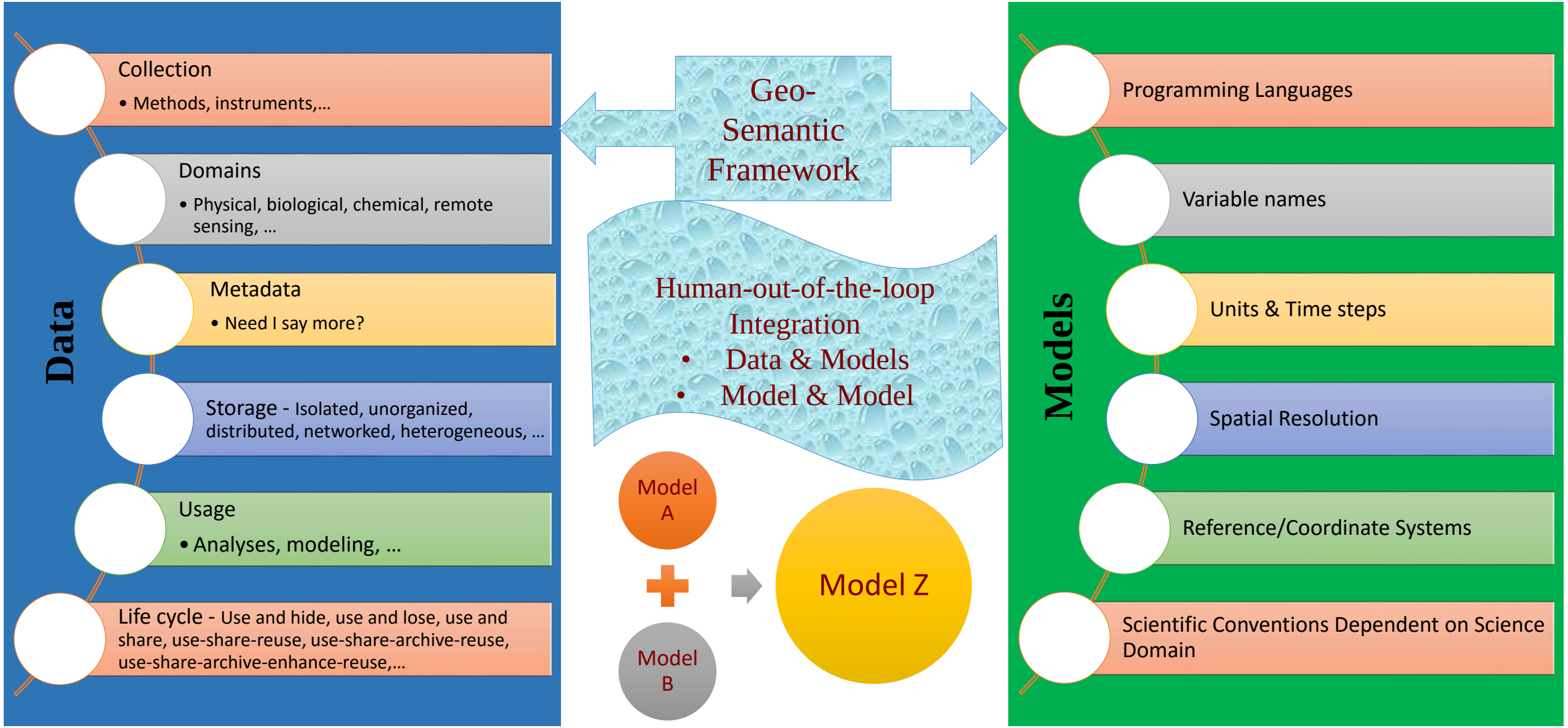
Interoperability Levels

- L0: None
- L1: Technical
- L2: Syntactic
- **L3: Semantic**
- L4: Pragmatic
- L5: Dynamics
- L6: Conceptual
- L7: Automatic

adapted from Wang, et al., 2009



Data-Model Semantic Integration



Data-Model Semantic Integration

SEAD | Sustainable Environment Actionable Data

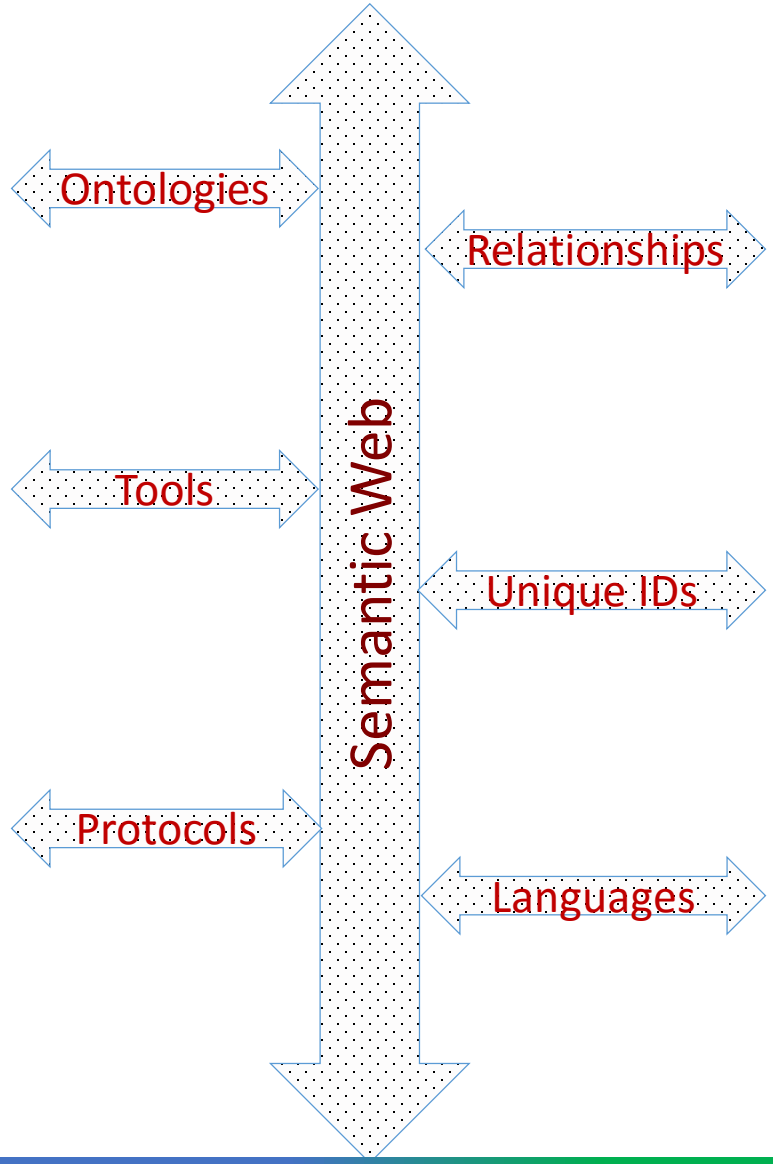
Clowder

Current

CUAHSI UNIVERSITIES ALLIED FOR WATER RESEARCH

HYDROSHARE

OTHERS



CSDMS
COMMUNITY SURFACE DYNAMICS MODELING SYSTEM

Model-as-a-Service (MaaS) Semantically Enabled Models [Integration of Unstructured Models]

Current

OTHERS

Data-Model Semantic Integration

SEAD | Sustainable Environment Actionable Data

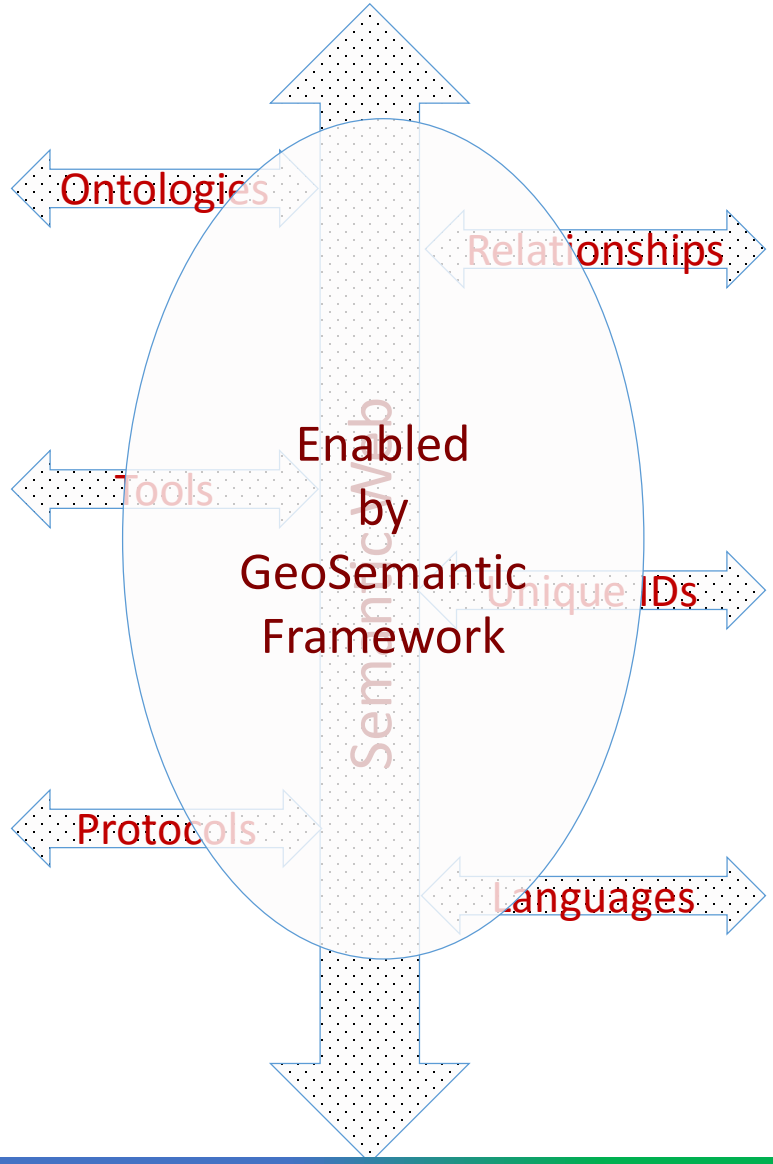
Clowder

Current

CUAHSI UNIVERSITIES ALLIED FOR WATER RESEARCH

HYDROSHARE

OTHERS



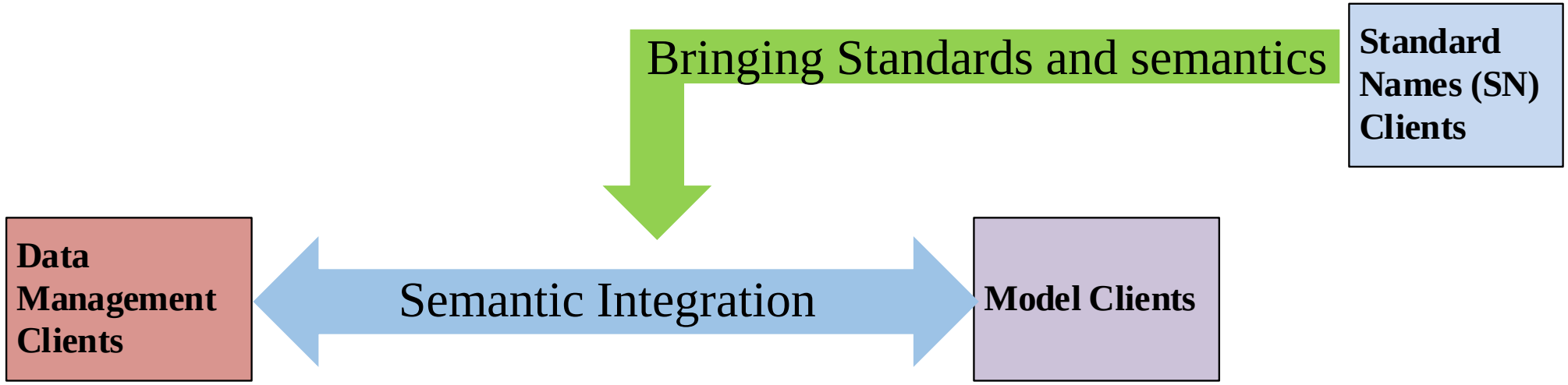
CSDMS
COMMUNITY SURFACE DYNAMICS MODELING SYSTEM

Model-as-a-Service (MaaS) Semantically Enabled Models [Integration of Unstructured Models]

Current

OTHERS

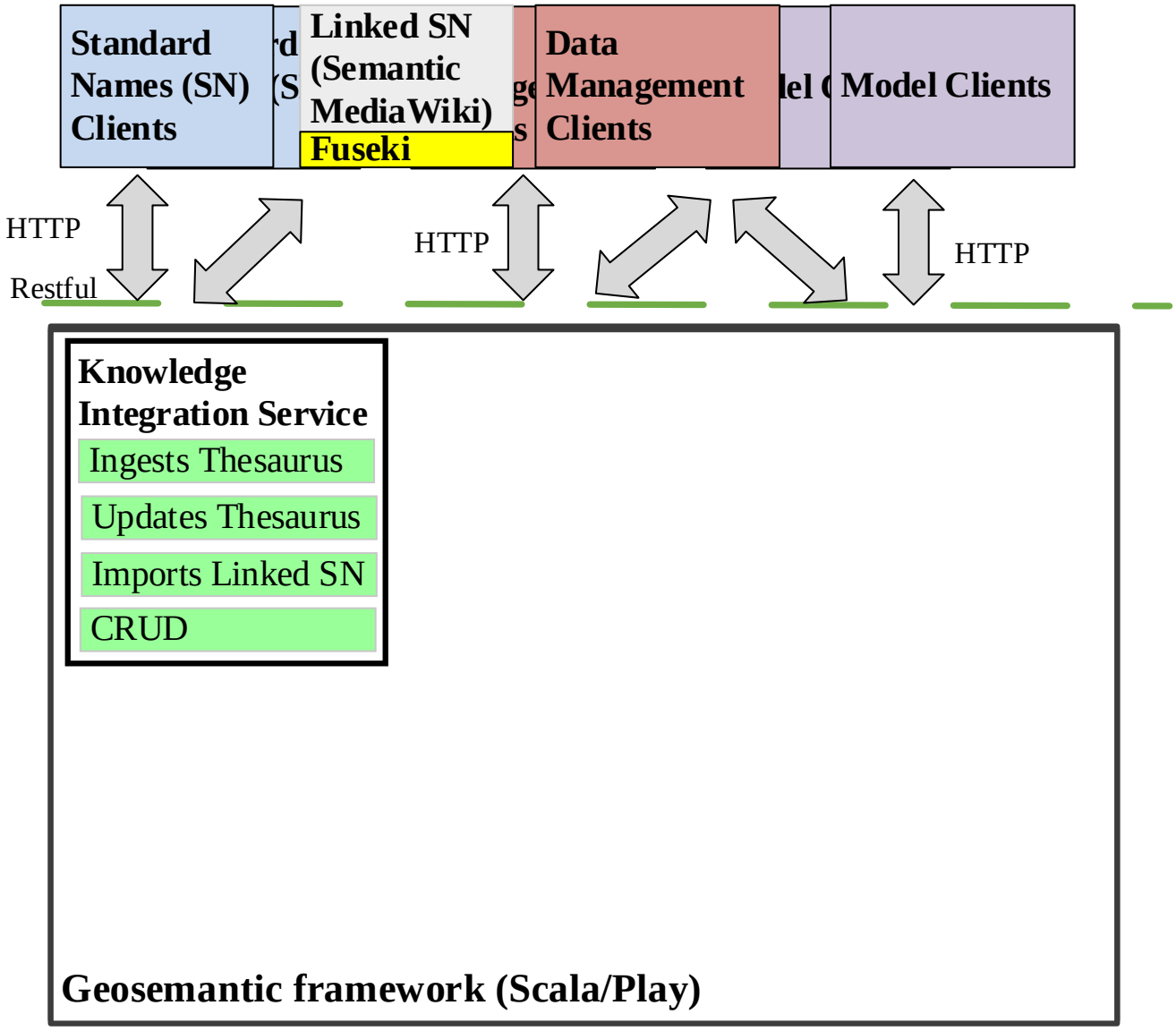
GeoSemantic Framework Architecture



GeoSemantic framework bridges the gap between Semantic Web standards and resources life cycle by lowering the technology barrier for incorporating semantic in their life cycle.

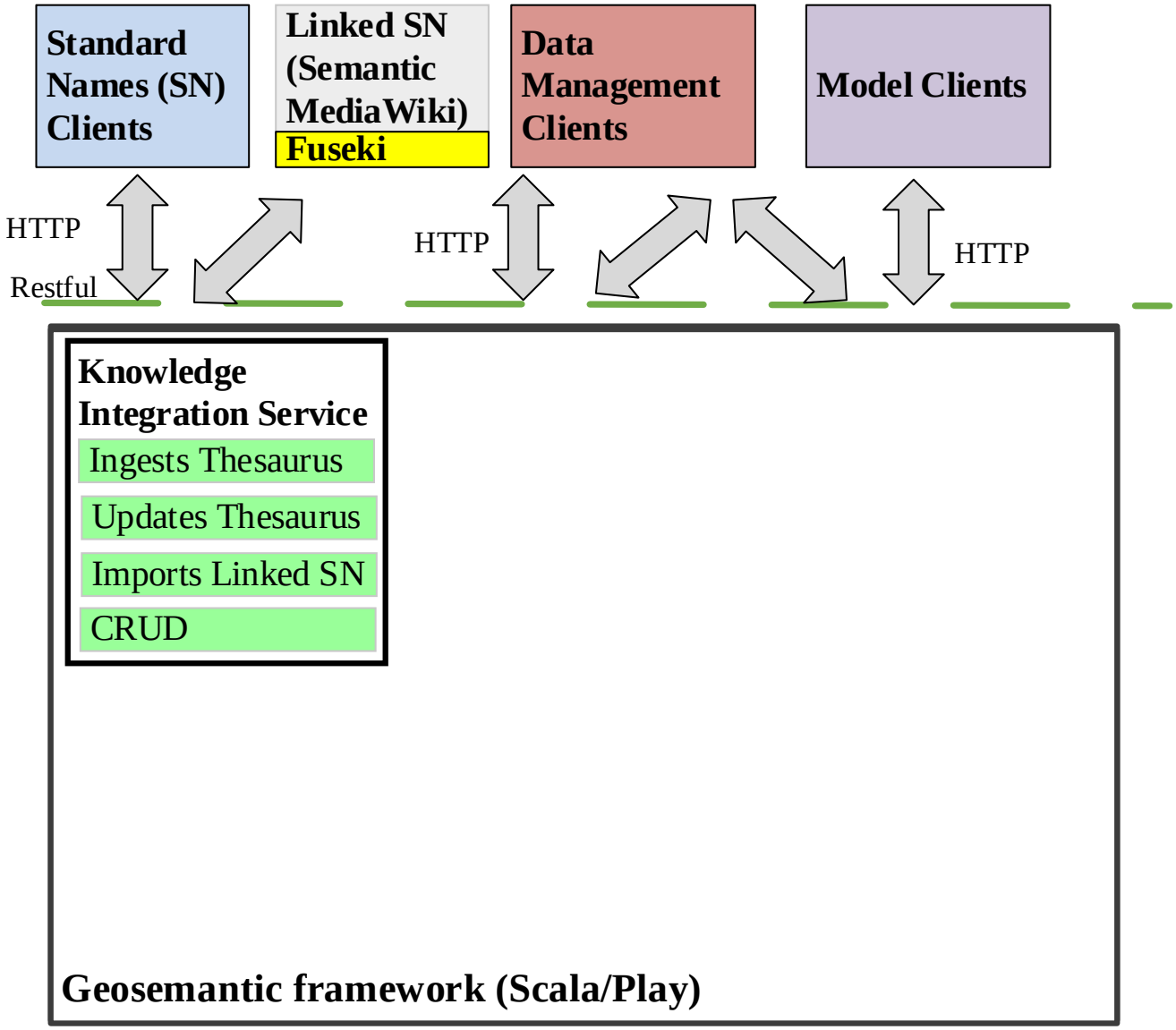
GeoSemantic Framework Architecture

- It uses Micro-service architecture, where each service focuses on specific function and has a unique endpoint.
- Knowledge Integration Services: It is the information entry point for the framework. It ingests standards and can annotate them back.
- Semantic MediaWiki is used to create semantic relationships among SN.



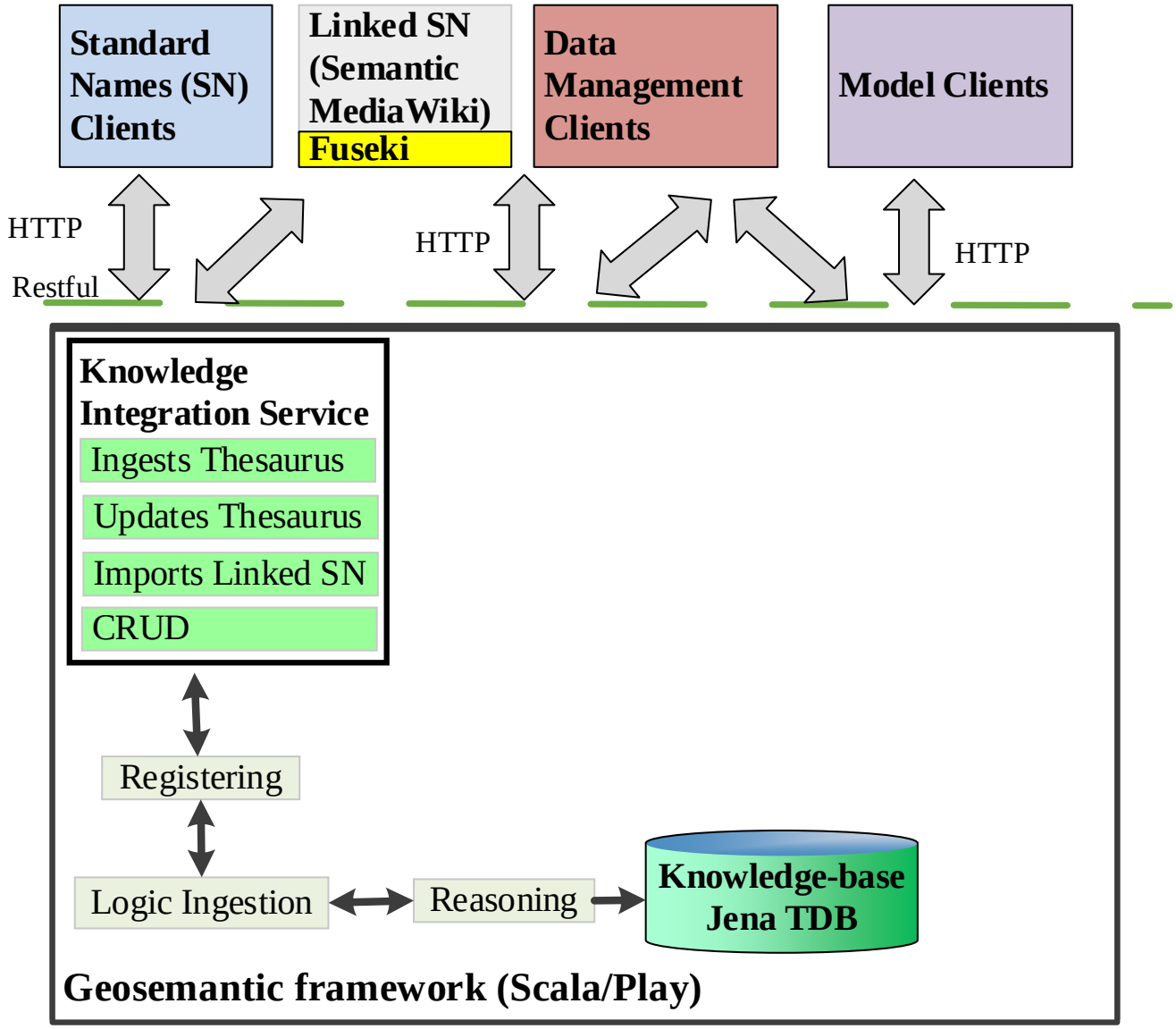
GeoSemantic Framework Architecture

- It uses Micro-service architecture, where each service focus on specific function and has a unique endpoint.
- Knowledge Integration Services: It is the information entry point for the framework. It ingests standards and can annotate them back.
- Semantic MediaWiki is used to create semantic relationships among SN.



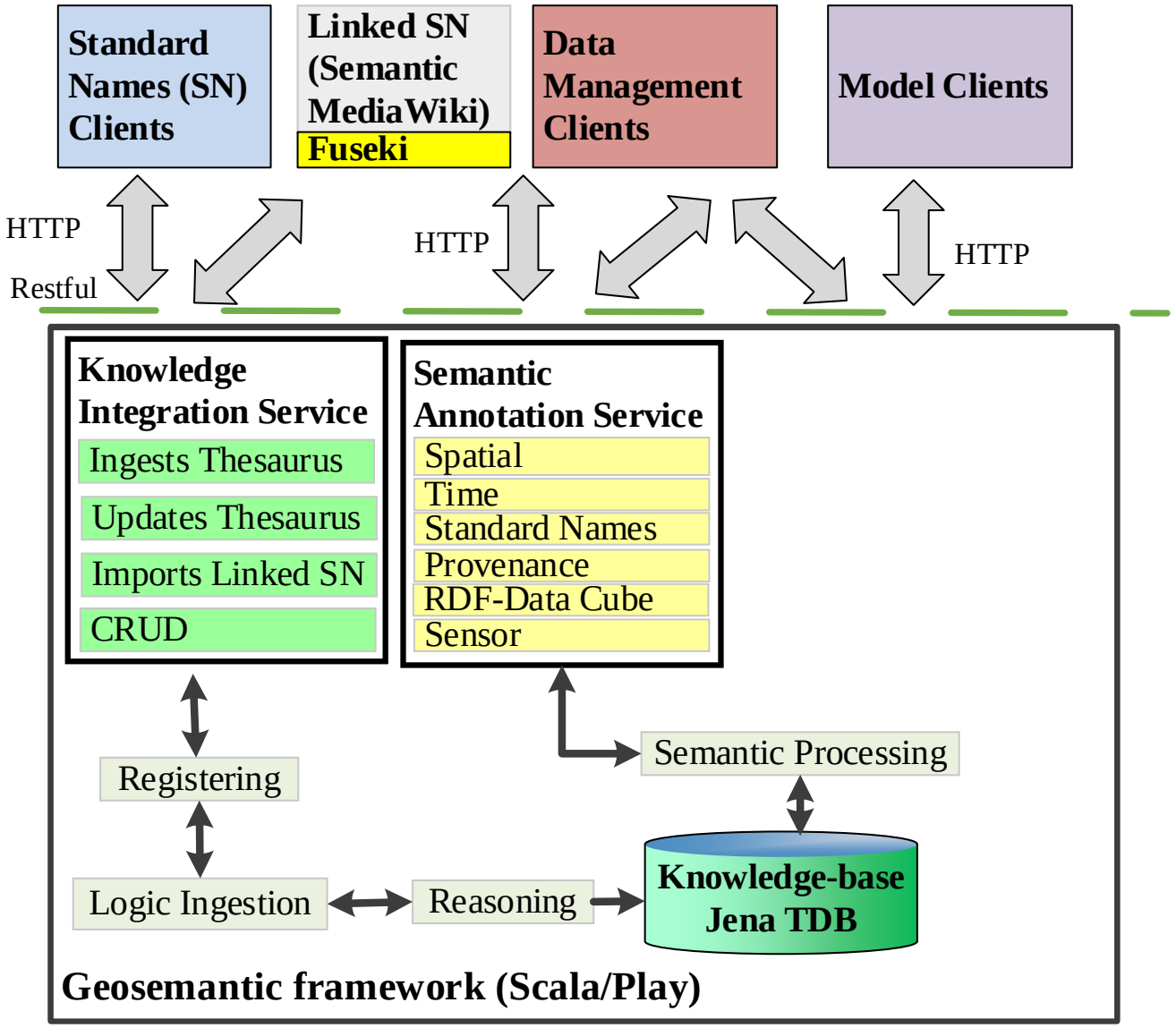
GeoSemantic Framework Architecture

- It uses Micro-service architecture, where each service focus on specific function and has a unique endpoint.
- Knowledge Integration Services: It is the information entry point for the framework. It ingests standards and can annotate them back.
- Semantic MediaWiki is used to create semantic relationships among SN.
- knowledge base: stores the standard graphs after processing.



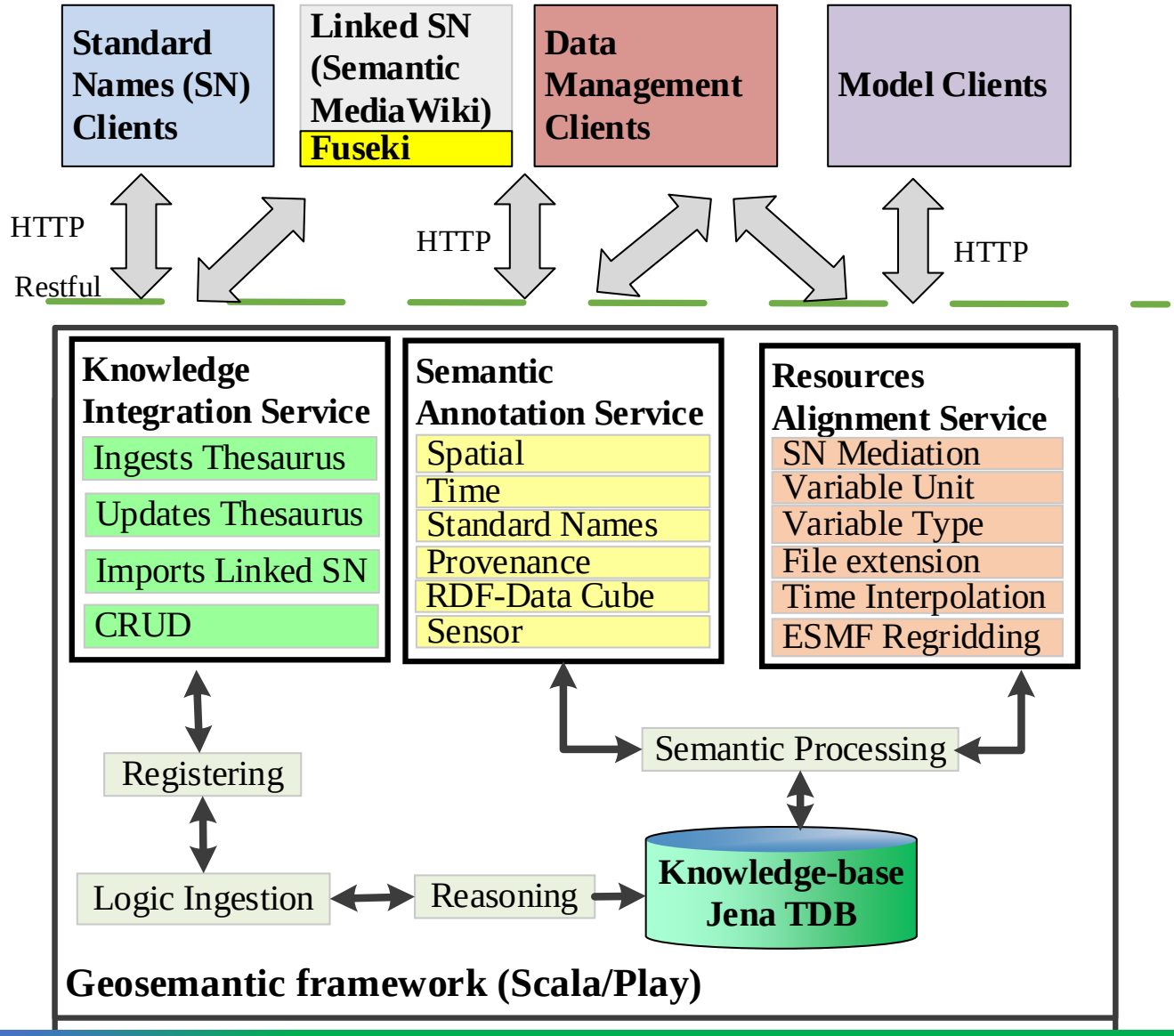
GeoSemantic Framework Architecture

- It uses Micro-service architecture, where each service focus on specific function and has a unique endpoint.
- Knowledge Integration Services: It is the information entry point for the framework. It ingests standards and can annotate them back.
- Semantic MediaWiki is used to create semantic relationships among SN.
- knowledge base: stores the standard graphs after processing.
- Semantic Annotation Services: enrich the data and models using scientific annotations.



GeoSemantic Framework Architecture

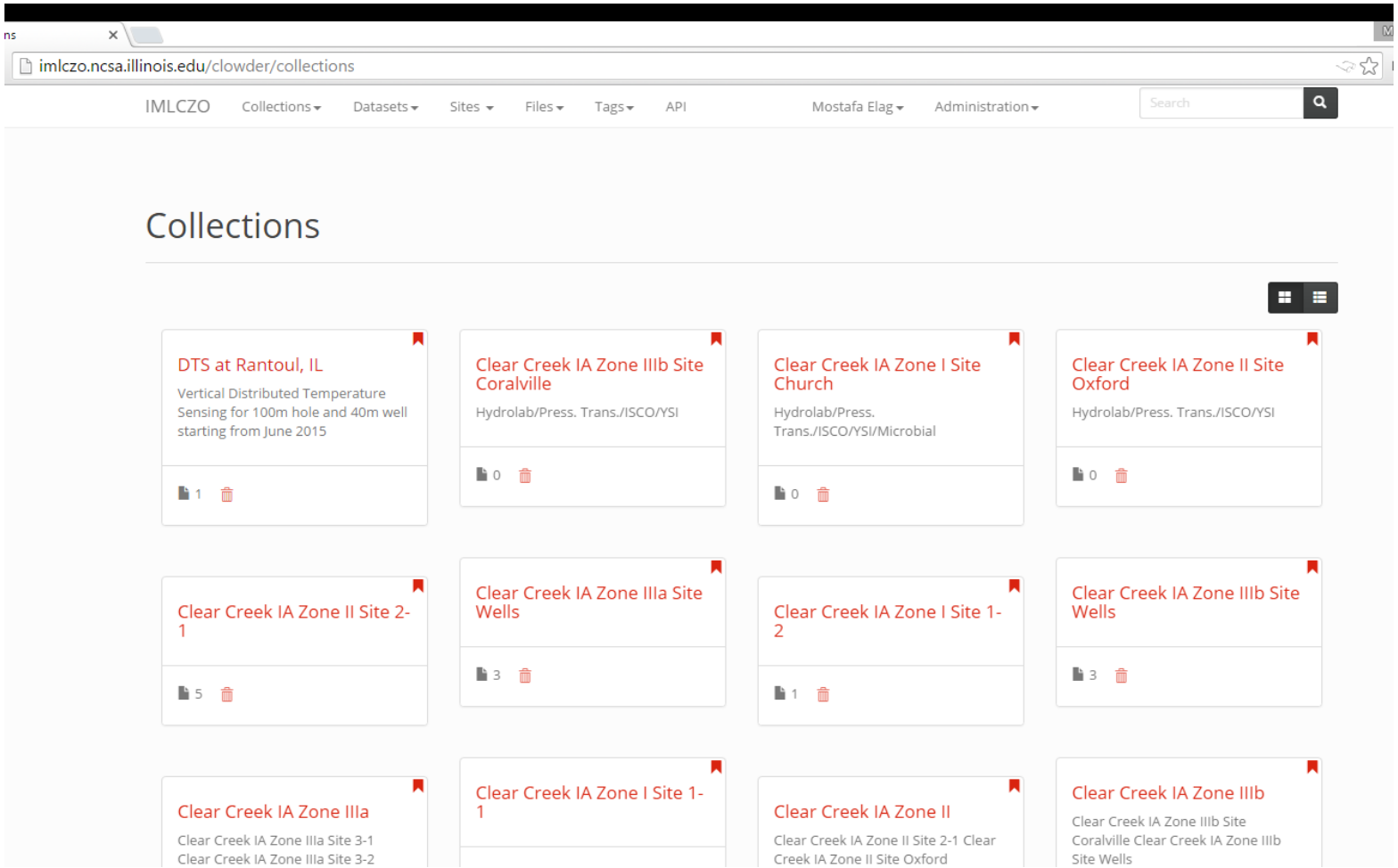
- It uses Micro-service architecture, where each service focus on specific function and has a unique endpoint.
- Knowledge Integration Services: It is the information entry point for the framework. It ingests standards and can annotate them back.
- Semantic MediaWiki is used to create semantic relationships among SN.
- knowledge base: stores the standard graphs after processing.
- Semantic Annotation Services: enrich the data and models using scientific annotations.
- Resource Alignment Services: ensures semantic consistency between coupled models.



Clients: IML-CZO data and Serviced BMI models

Semantic enrichment of IML-CZO data

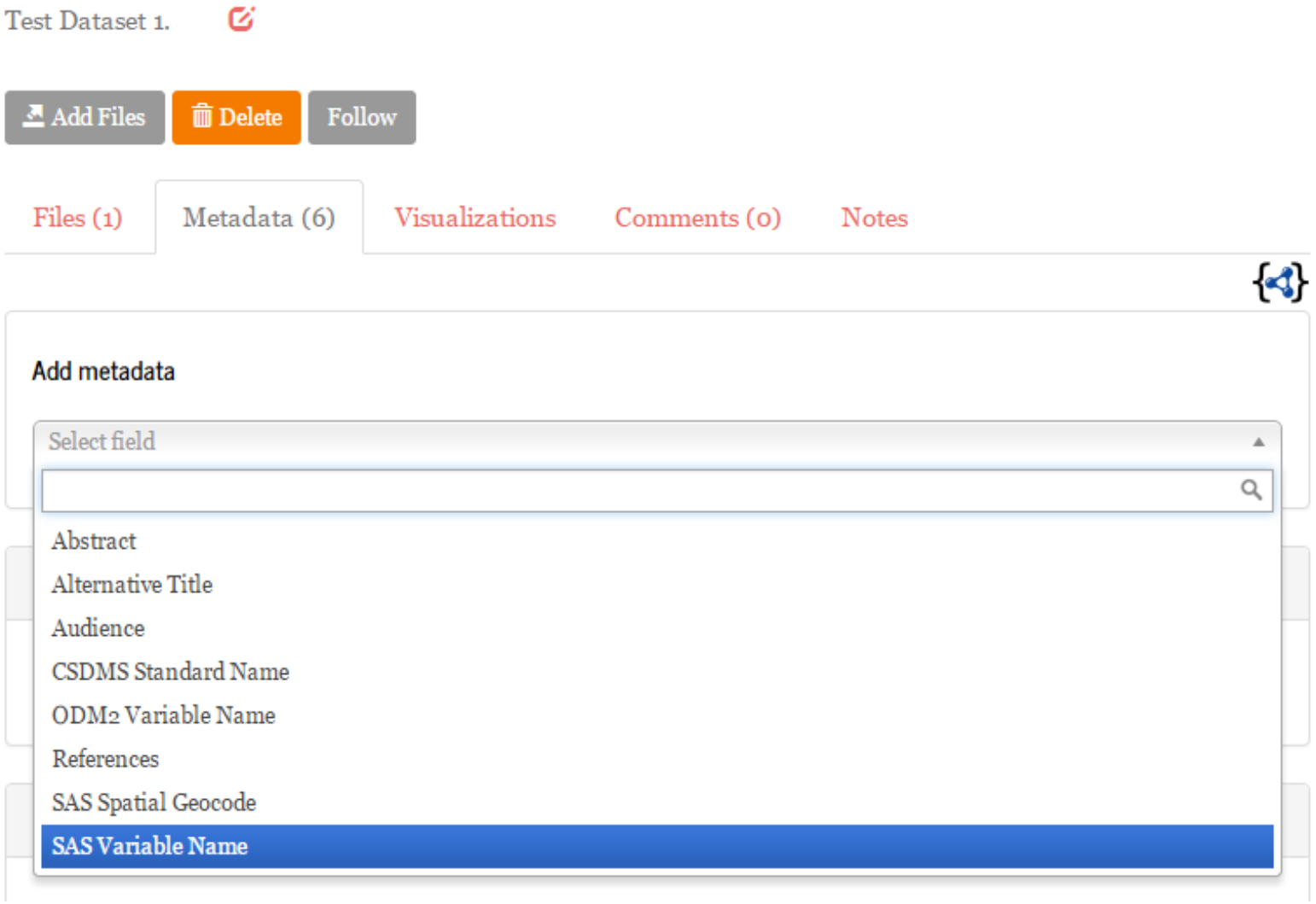
- SEAD is a client for the framework.



Clients: IML-CZO data and Serviced BMI models

Semantic enrichment of IML-CZO data

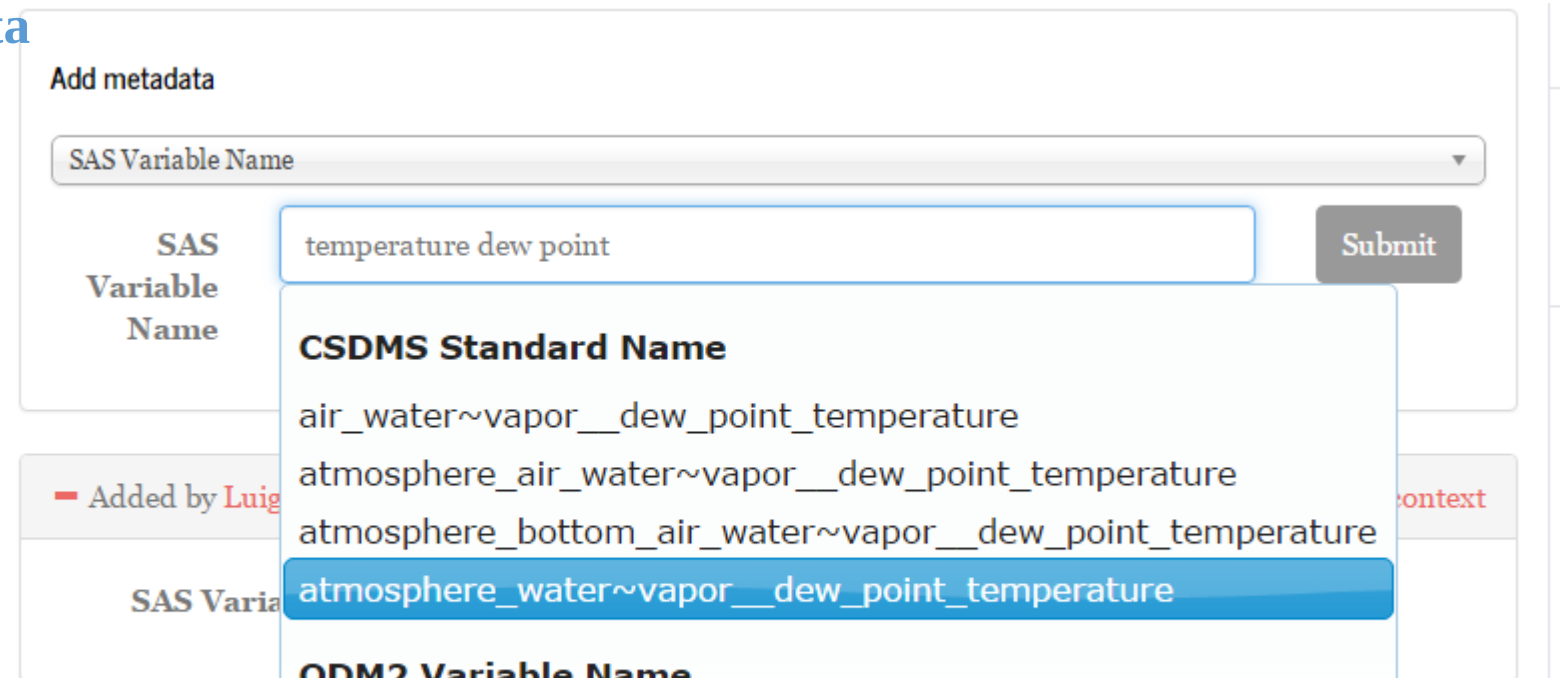
- SEAD is a client for the framework.
- IML-CZO data space has over than 200 data collections and consumes various Standard Names

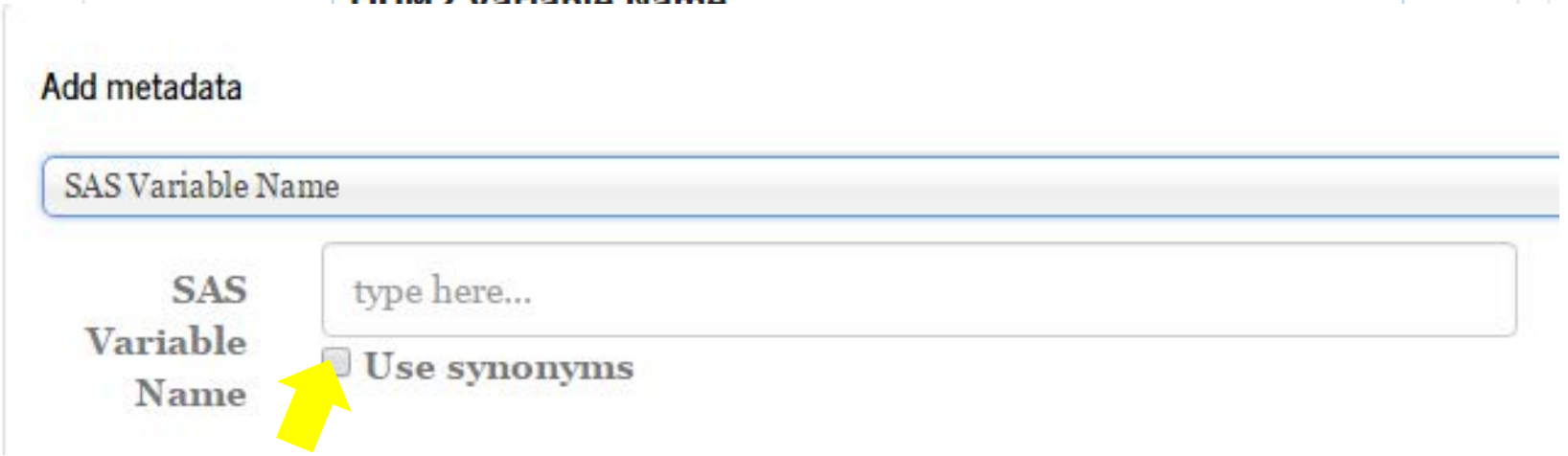


Clients: IML-CZO data and Serviced BMI models

Semantic enrichment of IML-CZO data

- SEAD is a client for the framework.
- IML-CZO space has over than 200 data collections and consumes various Standard Names
- Search across stored SN graphs using our search widget.
- Choose from synonyms based on Linked CVs.

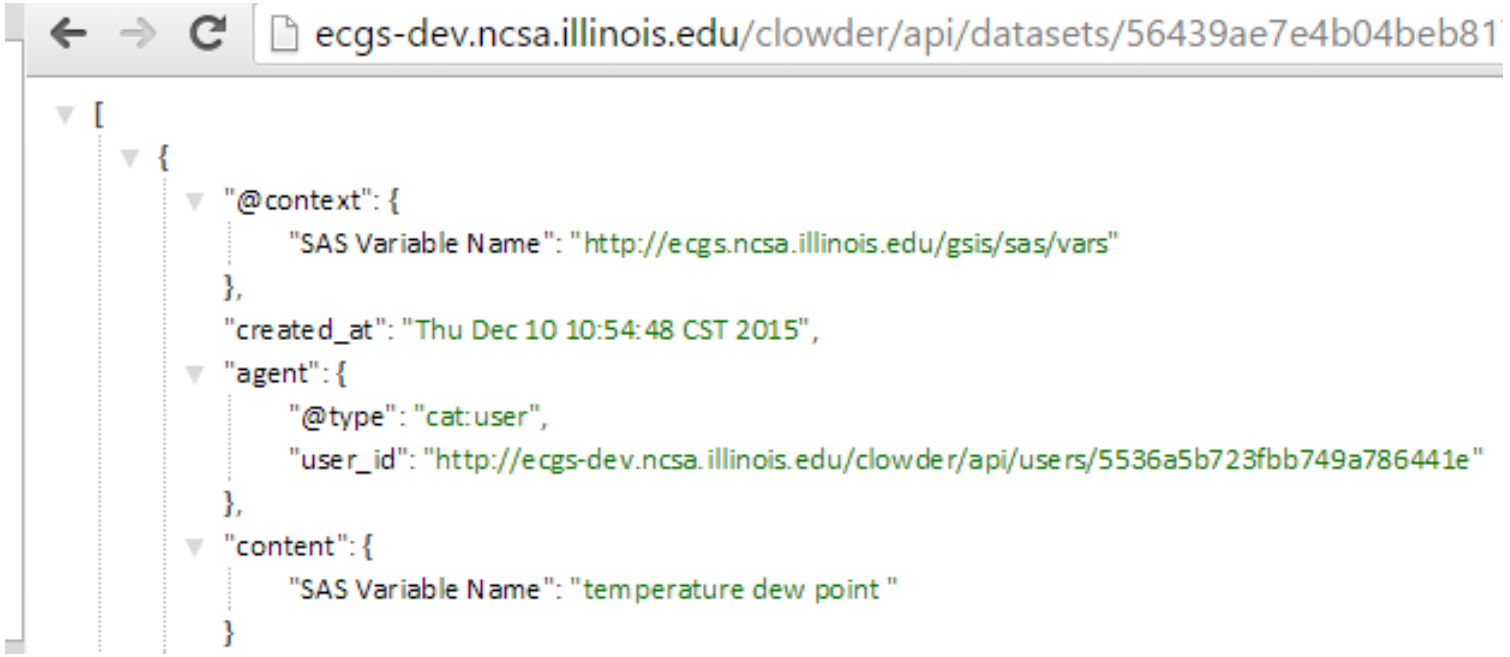




Clients: IML-CZO data and Serviced BMI models

Semantic enrichment of IML-CZO data

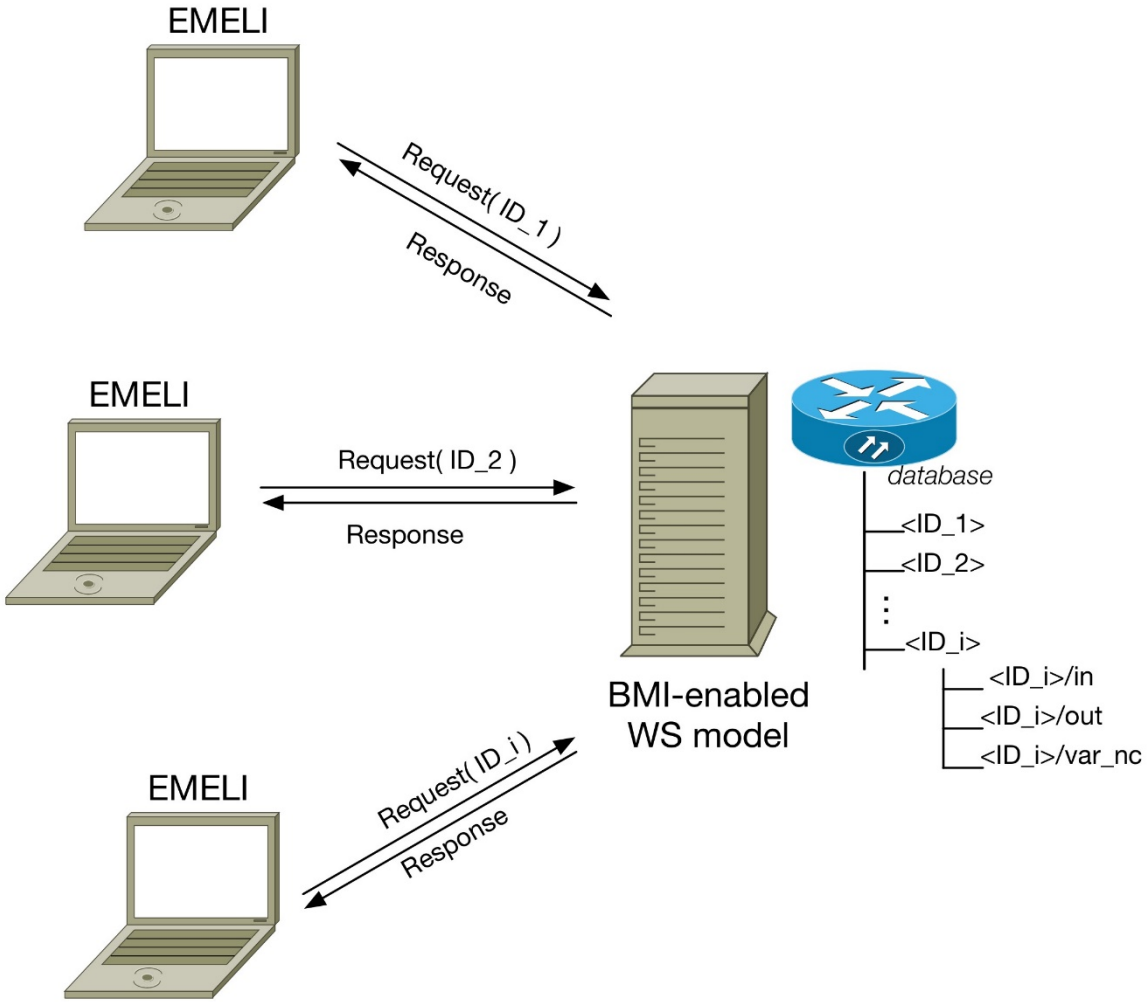
- SEAD is a client for the framework.
- IML-CZO space has over than 200 data collections and consumes various Standard Names
- Search across stored SN graphs using our search widget.
- Choose from synonyms based on Linked CVs.
- All metadata is serialized as JSON-LD



Clients: IML-CZO data and Serviced BMI models

Serviced BMI models

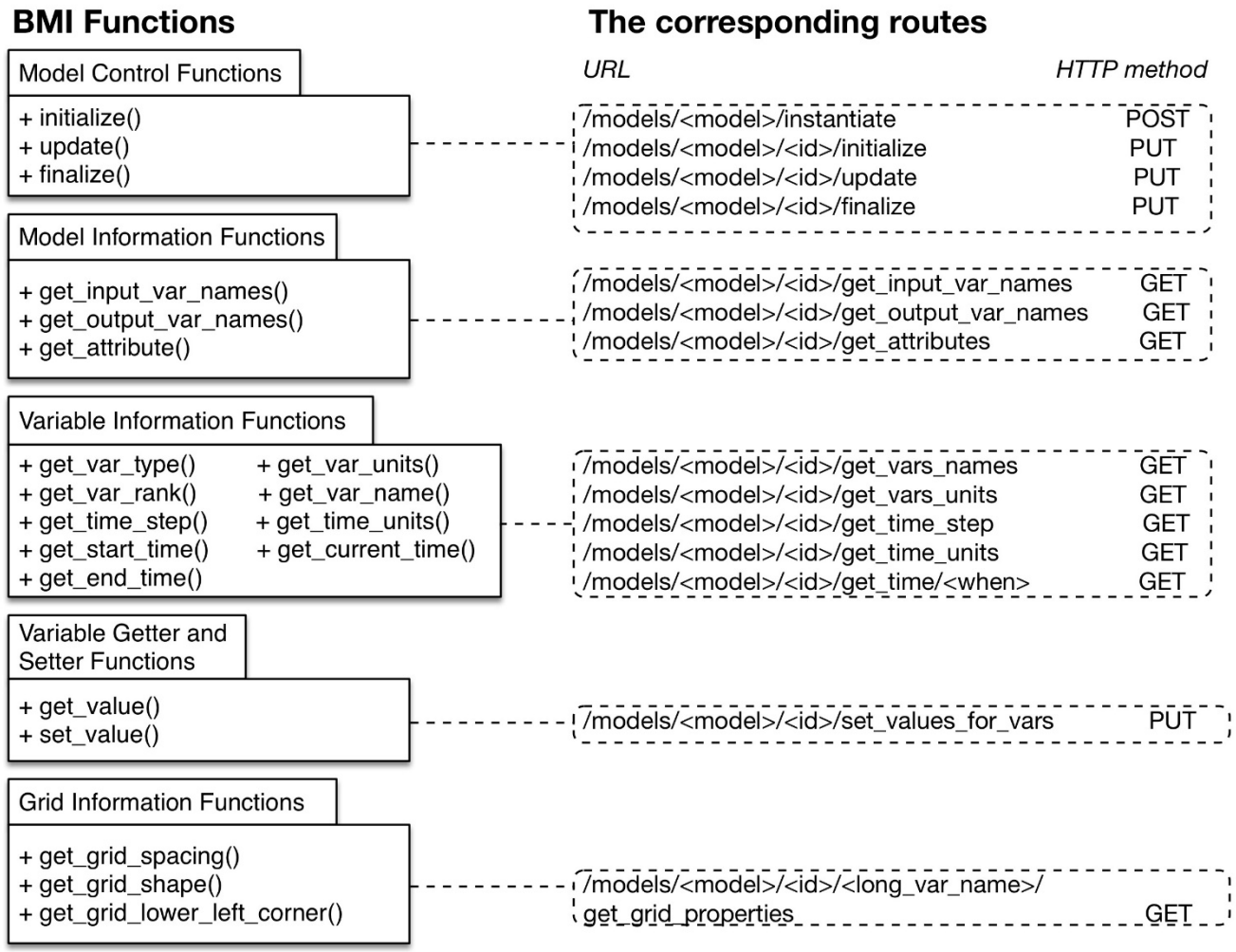
- BMI-enabled models are exposed as web services
- EMELI 1.0 (Experimental Modeling Environment for Linking and Interoperability) is promoted to EMELI 2.0 to serve Web serviced models.
- A unique UUID is assigned to each model execution task



Clients: IML-CZO data and Serviced BMI models

Serviced BMI models

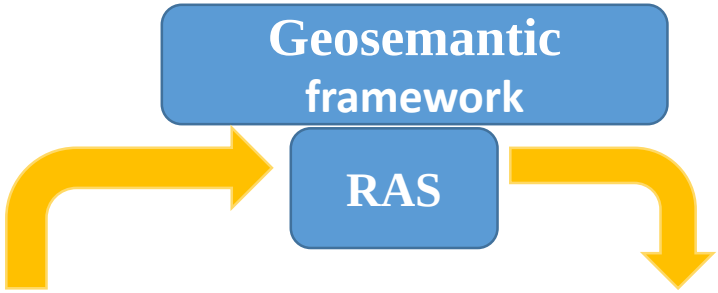
- BMI-enabled models are exposed as web services
- EMELI 1.0 (Experimental Modeling Environment for Linking and Interoperability) is promoted to EMELI 2.0 to serve Web serviced models.
-
- A unique UUID is assigned to each model execution task.
- Each BMI function has a corresponding route.



Integration between IML-CZO data and Serviced BMI models

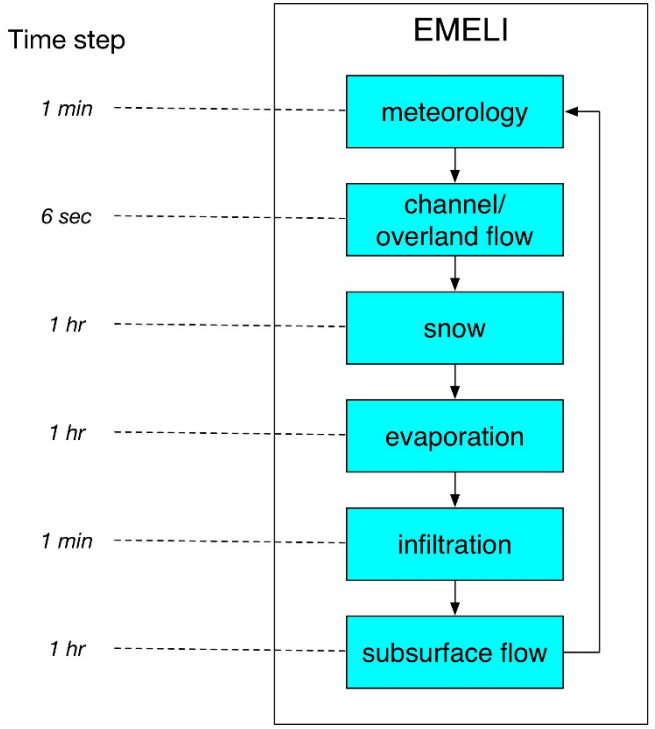
Goose creek watershed, IML-CZO


```
    },
    "created_at": "Tue Nov 24 10:24:22 CST 2015",
  ▼ "agent": {
        "@type": "cat:user",
        "user_id": "http://ecgs-dev.ncsa.illinois.edu/clowder/api/users/555a0c933f22ab256761eab0"
    },
  ▼ "content": {
        "SAS Variable Name": "Abs"
    }
},
▼ {
  ▼ "@context": {
        "SAS Spatial Geocode": "http://ecgs.ncsa.illinois.edu/gsis/sas/geocode"
    },
    "created_at": "Wed Nov 18 09:58:24 CST 2015",
  ▼ "agent": {
        "@type": "cat:user",
        "user_id": "http://ecgs-dev.ncsa.illinois.edu/clowder/api/users/553f9fc123fbb749a786443a"
    },
  ▼ "content": {
        "Location Name": "Denver, CO, USA",
        "Latitude": "39.74",
        "Longitude": "-104.99"
    }
},
▼ {
  ▼ "@context": {
        "SAS Variable Name": "http://ecgs.ncsa.illinois.edu/gsis/sas/vars"
    },
    "created_at": "Wed Nov 18 09:57:37 CST 2015",
  ▼ "agent": {
        "@type": "cat:user",
        "user_id": "http://ecgs-dev.ncsa.illinois.edu/clowder/api/users/553f9fc123fbb749a786443a"
    },
  ▼ "content": {
        "SAS Variable Name": "odm2:windGustSpeed"
    }
},
▼ {
  ▼ "@context": {
        "Alternative Title": "http://purl.org/dc/terms/alternative"
```

Resources Alignment between IML-CZO data and BMI models

Serviced Topoflow components



Summary and Future work

- GeoSemantic framework lowers the barrier for incorporating semantics in resource life cycle by:
  1. Ingesting standards and ontologies.
  2. Creating Standard Names semantic wiki for linking Controlled Vocabularies.
  3. Providing SAS for semantic enrichment of resources.
  4. Developing RAS to ensure semantic consistency between coupled resources.
- Future work will concentrate on:
  1. Adapting more standards and endpoints to satisfy different geoscience communities.
  2. Converting EMELI to be a web service.
  3. Creating MicroData templates to incorporate semantic annotation directly in HTML.
  4. Developing micro-services to annotate simulation models.

Project URLs

- We invite input and feedback from the Geoscience community at:

1. EarthCube: http://workspace.earthcube.org/geo-semantic
2. Confluence: https://opensource.ncsa.illinois.edu/confluence/display/ECGS/GeoSemantic

- We encourage developers to contribute to the framework source code at

  https://opensource.ncsa.illinois.edu/stash/projects/ECGS

- Geosemantic Wiki of Standard Names is available at
  http://ecgs.ncsa.illinois.edu/mediawiki/index.php/Main

Acknowledgments

Support from NSF grants ACI-0940824", ACI-1261582", EAR-1331906", and ICER-1440315" are gratefully acknowledged

THANK YOU