



Building the NNEW Weather Ontology

**Kelly Moran
Kajal Claypool**

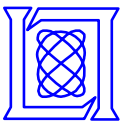
MIT Lincoln Laboratory

5 May 2010



Outline

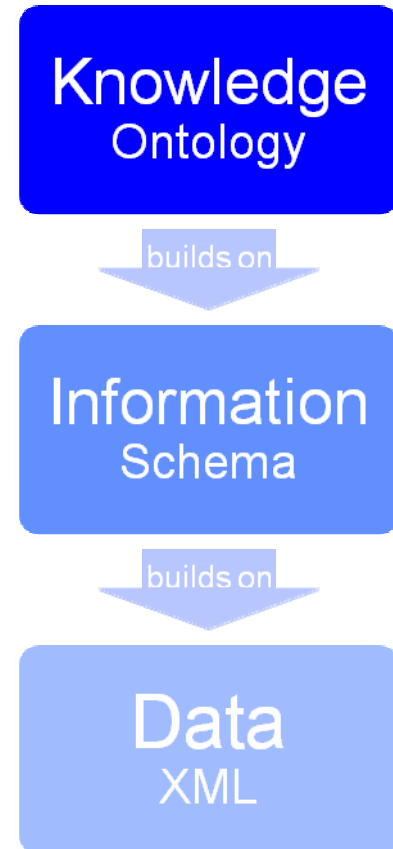
- **Introduction**
- **Ontology Development Methods & Tools**
- **NNEW Weather Ontology Design**
- **Application: Semantic Search**
- **Summary**



What is an Ontology?

“An ontology defines a **common vocabulary** for researchers who need to **share information in a domain**. It includes machine-interpretable definitions of basic concepts in the domain and relations among them.”
(Noy 2000)

- **Definition:** Set of elements and the relationships between them
- **Purpose:** Formally model a domain
 - Used for information integration and knowledge sharing
- **Formal specification:** Popular language is Web Ontology Language (OWL), a semantic markup language
- **Representation:** Modeled as a graph





Role of Ontologies within WXXM

- **Ontologies are used by Observations & Measurements model**

```
<om:observation>  
<om:observedProperty  
  xlink:href="http://sweet.jpl.nasa.gov/ontology.owl#Temperature"/>
```



External
JPL SWEET
Ontology

- Provide linkage from data instance to higher-level knowledge tree
- Enable semantically-enhanced data queries (SPARQL)

- **WXXM leverages and extends the Observations & Measurements model**

```
<wx:observation>  
<om:observedProperty  
  xlink:href="http://wmo.org/ont/wx/1.1/wx.owl#Precipitation"/>
```



External
Weather
Ontology



- Though some work in this area exists, no single ontology meets the needs of the WXXM community today



Relevant Ontologies / Vocabularies

- **Wordnet**
 - Large domain-independent lexical database
- **Suggested Upper Merged Ontology (SUMO)**
 - Upper and mid-level ontology containing a WordNet mapping
- **JPL's Semantic Web for Earth and Environmental Technology (SWEET 2.0)**
 - Mid-level ontology to enables scalable classification of Earth system science concepts
- **NetCDF Climate and Forecast Conventions (CF) Standard Names**
 - Non-hierarchical list of terms that represent quantitative measurements of different weather phenomena
- **Joint METOC Broker Language (JMBL) Parameter List**
 - Similar to CF but with the addition of statistical parameters

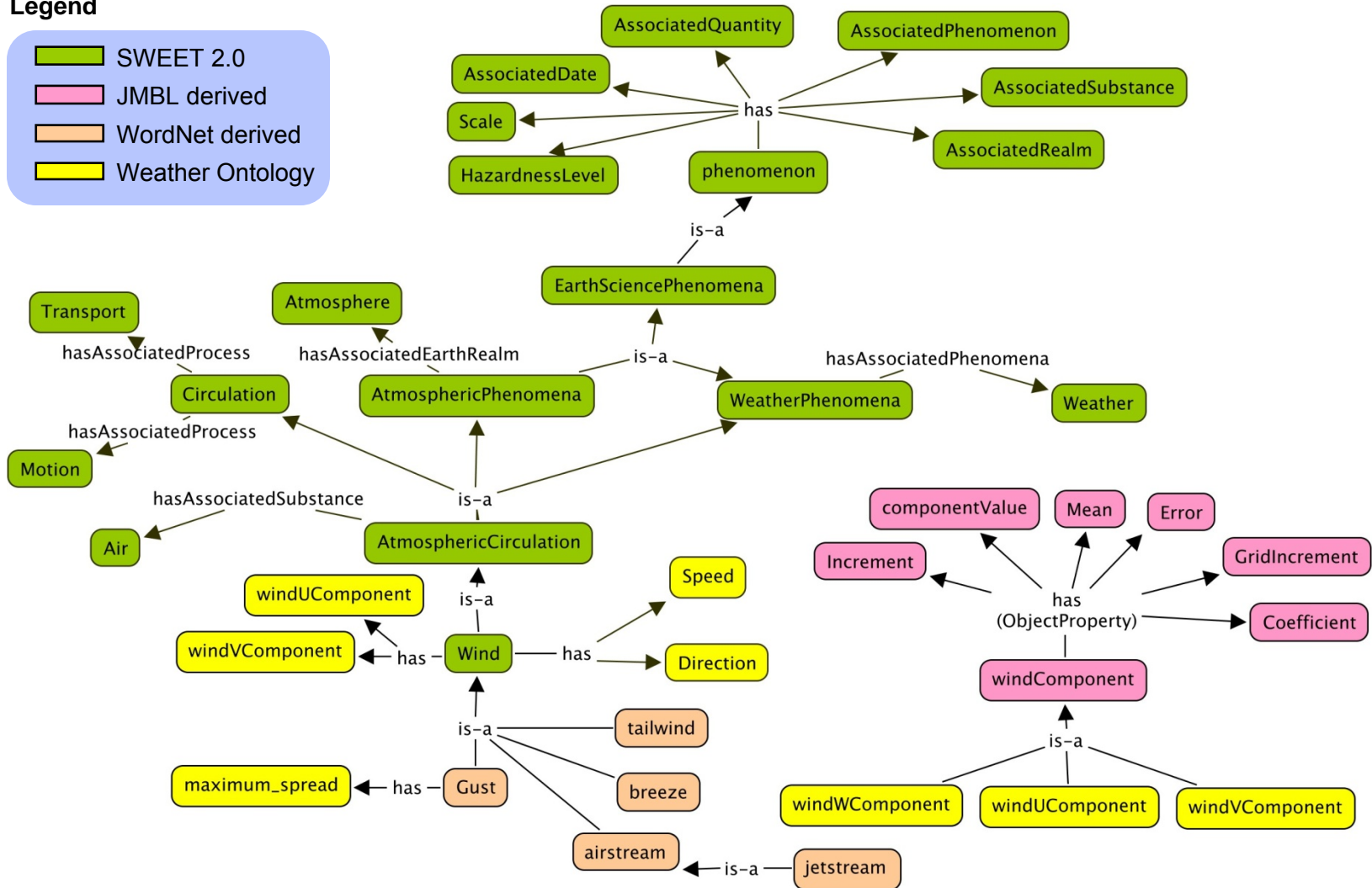
Constructing the Weather Ontology for NNEW is the art of combining these efforts in a coherent structure, and extending where necessary.



NNEW Weather Ontology Fragment

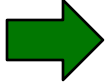
Legend

- SWEET 2.0
- JMBL derived
- WordNet derived
- Weather Ontology





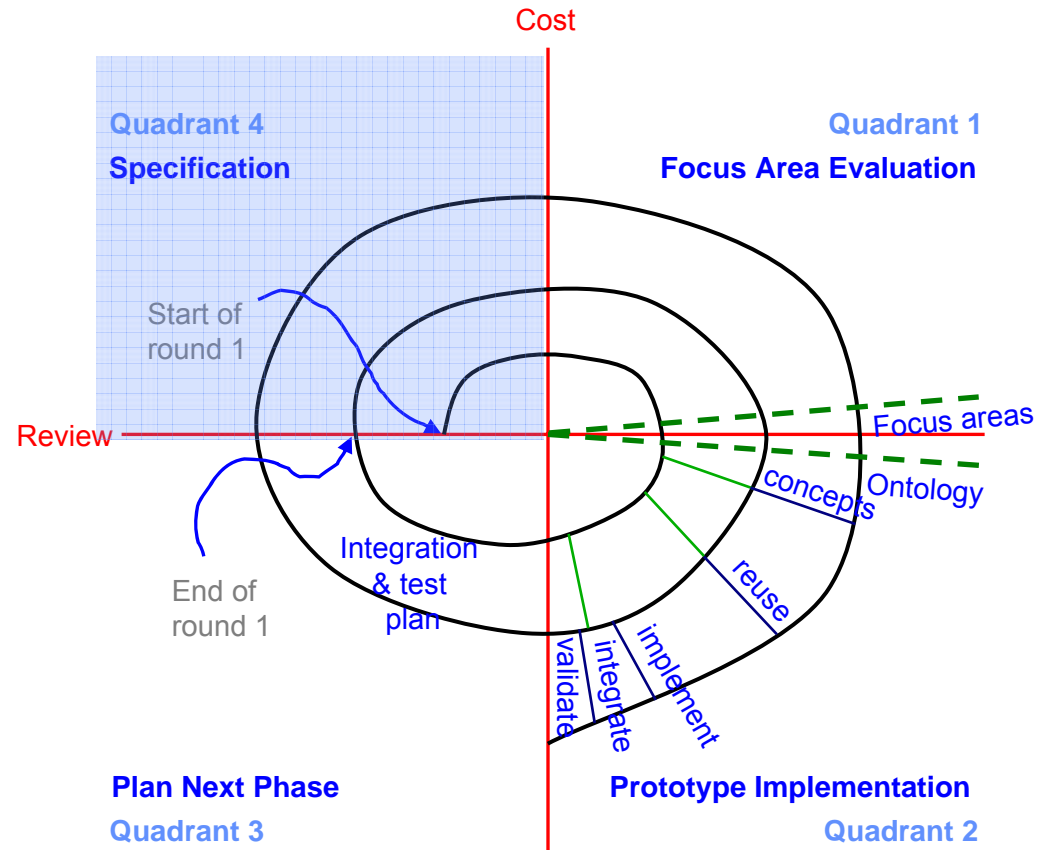
Outline

- Introduction
-  • **Ontology Development Methods & Tools**
- **NEW Weather Ontology Design**
- **Application: Semantic Search**
- **Summary**



Ontology Development Methods

- **Ontology-level method:**
 - Spiral development methodology
 - **Specification:** Define the domain and scope of the ontology
 - **Focus Area Evaluation:** Segment the overall domain and scope of ontology into smaller focus areas. Prioritize the focus area.

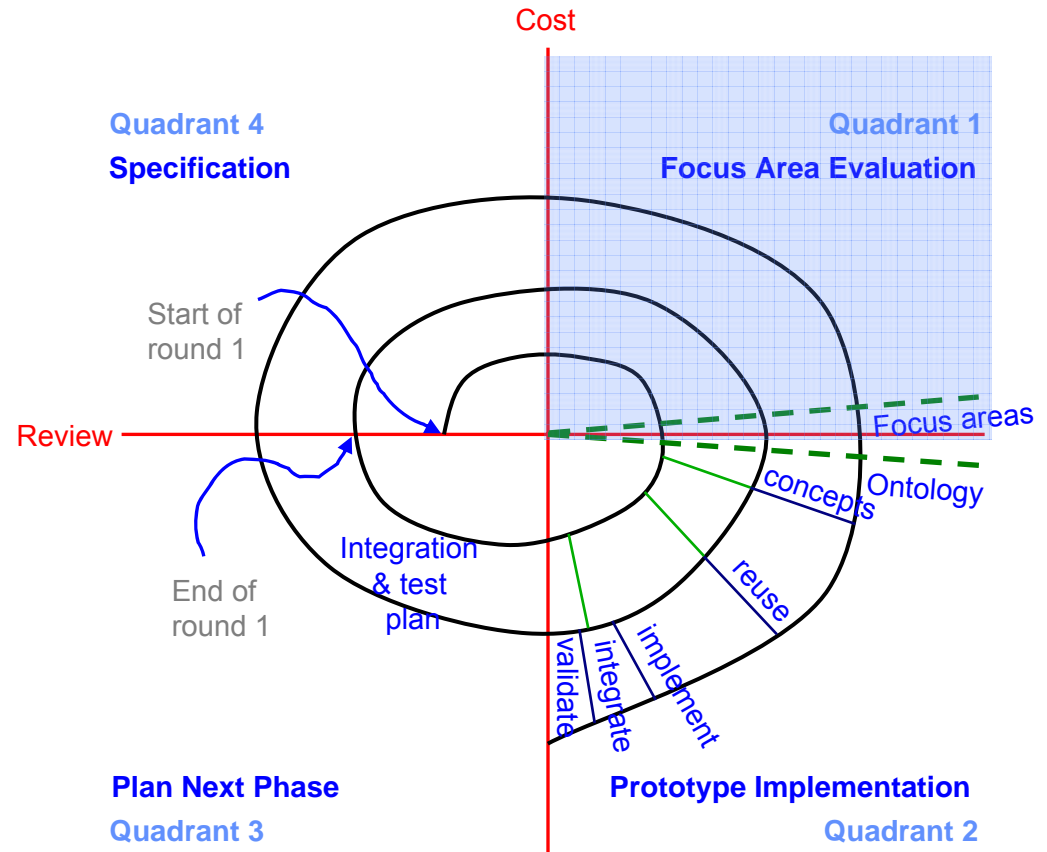




Ontology Development Methods

- **Ontology-level method:**

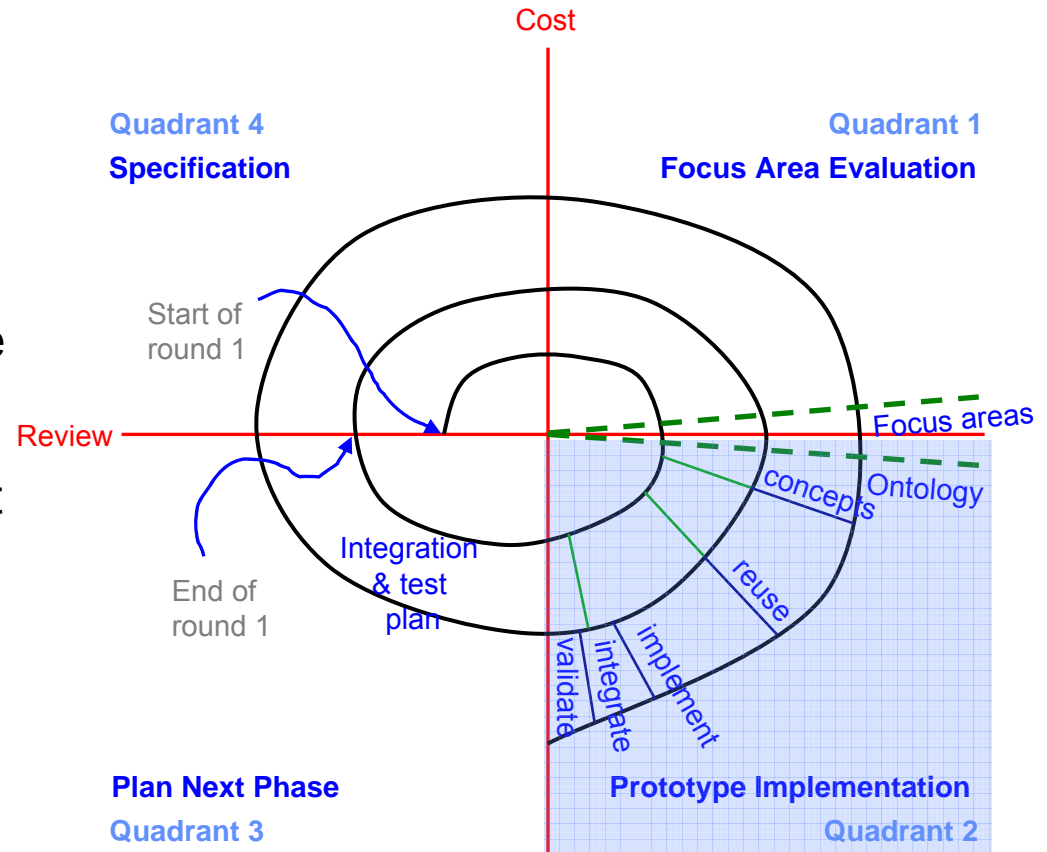
- Spiral development methodology
- **Specification:** Define the domain and scope of the ontology
- **Focus Area Evaluation:** Segment the overall domain and scope of ontology into smaller focus areas. Prioritize the focus area.





Ontology development methods

- **Prototype implementation:**
 - **Conceptualize:** Enumerate important concepts
 - **Reuse:** Identify reuse opportunities at upper/mid/low ontologies for straight reuse or as starting point
 - **Implement:** Define the classes, class hierarchy, and properties for the concept
 - **Validate:** Validate the ontology focus area





Design Principles

- **Design principles:**
 - **Expressive representation**
Model concepts with hierarchies and relationships, not with flat term concatenation
 - **Internal concept reuse**
Reusing concepts *within* an ontology ensures consistency and reduces ambiguity
 - **Consistent scoping**
Converge on a common granularity for each sub-domain



Ontology Development Tools

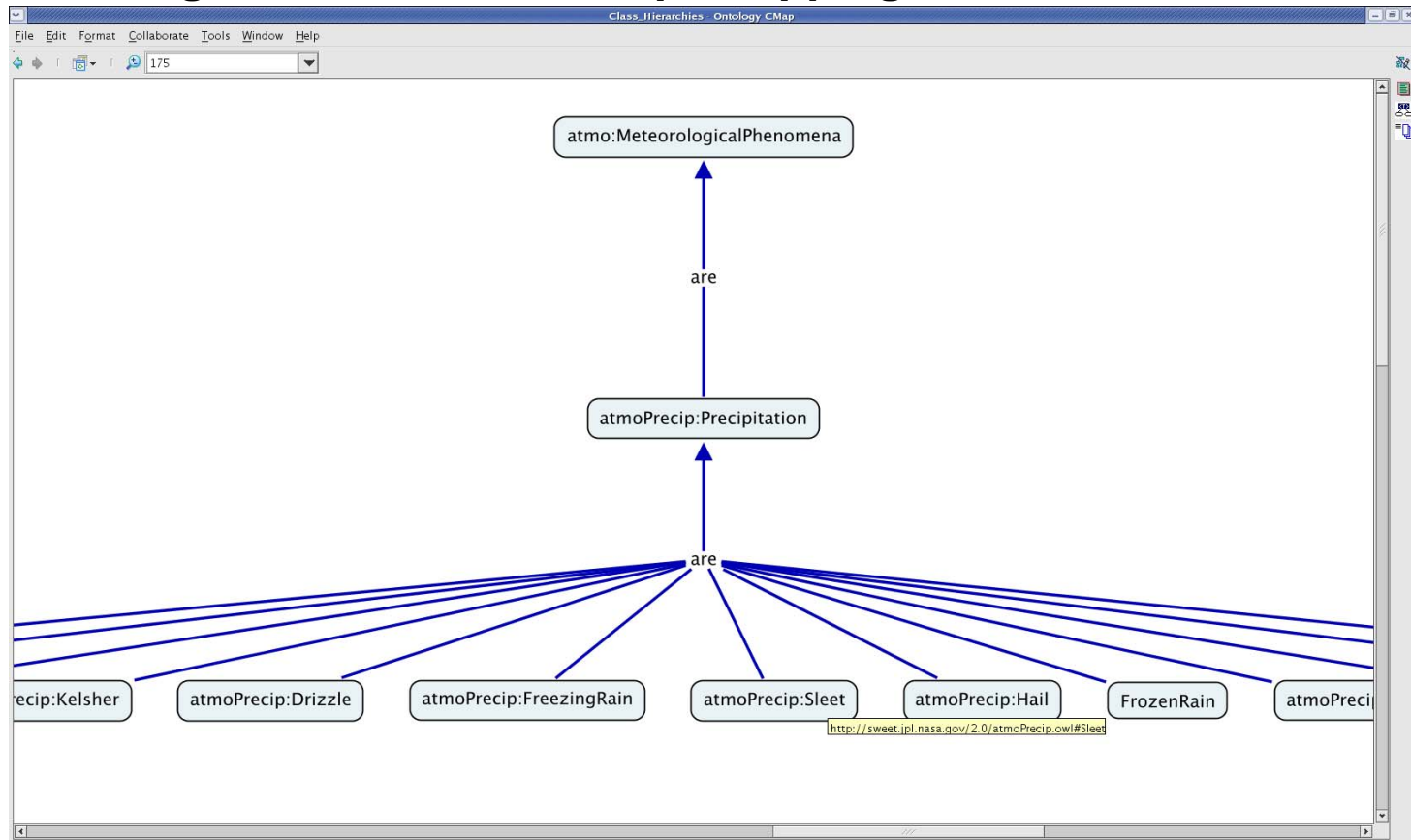
- Viewing & editing ontologies
 - Protégé: Free, open-source ontology edit from Stanford

The screenshot displays the Protégé ontology editor interface. The main window shows a class hierarchy on the left, with 'Lightning' selected. The right pane shows the 'Class Annotations' for 'Lightning', including a description and source. The bottom pane shows the 'Object property hierarchy' for 'Lightning', listing various properties like 'hasUnit', 'hasMeasurement', etc. Three blue callout boxes with white text and a small 'S' icon are overlaid on the image: 'Concept' points to the class hierarchy, 'Annotation' points to the class annotations pane, and 'Relationship' points to the object property hierarchy pane.



Ontology Development Tools

- Visualizing ontologies
 - CMap Tools COE: Florida Institute for Human & Machine Cognition tool for concept mapping in OWL





Ontology Development Tools

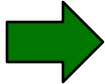


Ontology Development Tools

The screenshot displays an ontology development interface with a central mapping area. On the left, a tree view shows the source ontology structure, including 'Lightning', 'GreenElf', 'FirstStroke', 'Probability', 'PositiveStreamer', 'Stroke', and 'HighAltitudeLightning'. On the right, a tree view shows the target ontology structure, including 'weatherConditionSevere', 'heightOptimumCoupling', 'deviceIdentifier', 'windShearCode', 'soilMoistureSaturationMinimum', 'precipitationAccumulatedSnowNonConvective', and 'lightning'. A red arrow points from the 'Lightning' node in the source to the 'lightning' node in the target. A blue arrow points from the 'HighAltitudeLightning' node in the source to the 'lightning' node in the target. The interface includes a menu bar (Ontologies, Algorithms, Help), a toolbar (Set Source, Set Target, Remove), and a 'Mappings' section (crn.rdf) with buttons for Open, Save, and Save As. A 'Tools' section contains various actions like 'Show composite', 'Delete Line', and 'Update'.

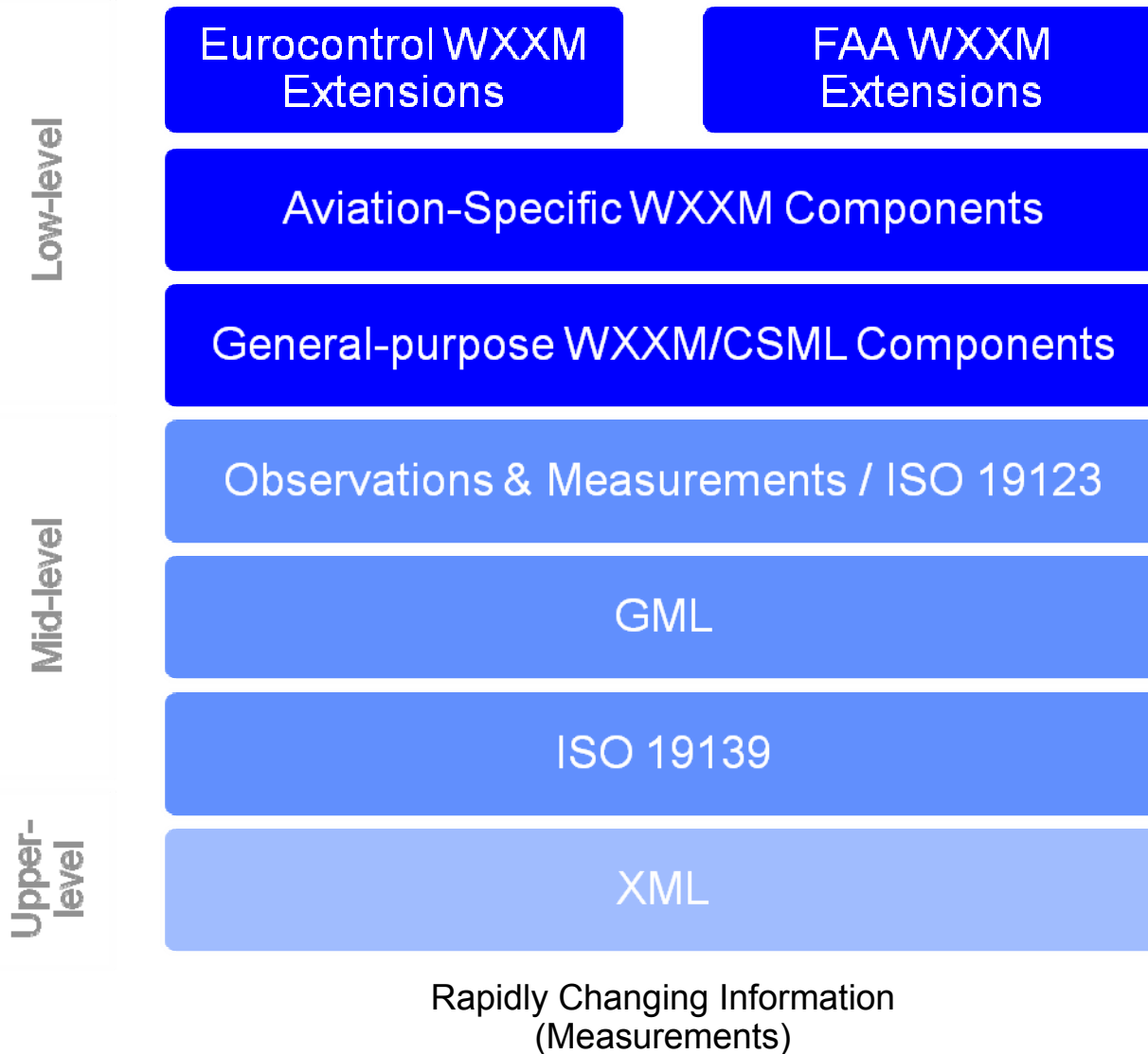


Outline

- Introduction
- Ontology Development Methods & Tools
-  • **NEW Weather Ontology Design**
- Application: Semantic Search
- Summary

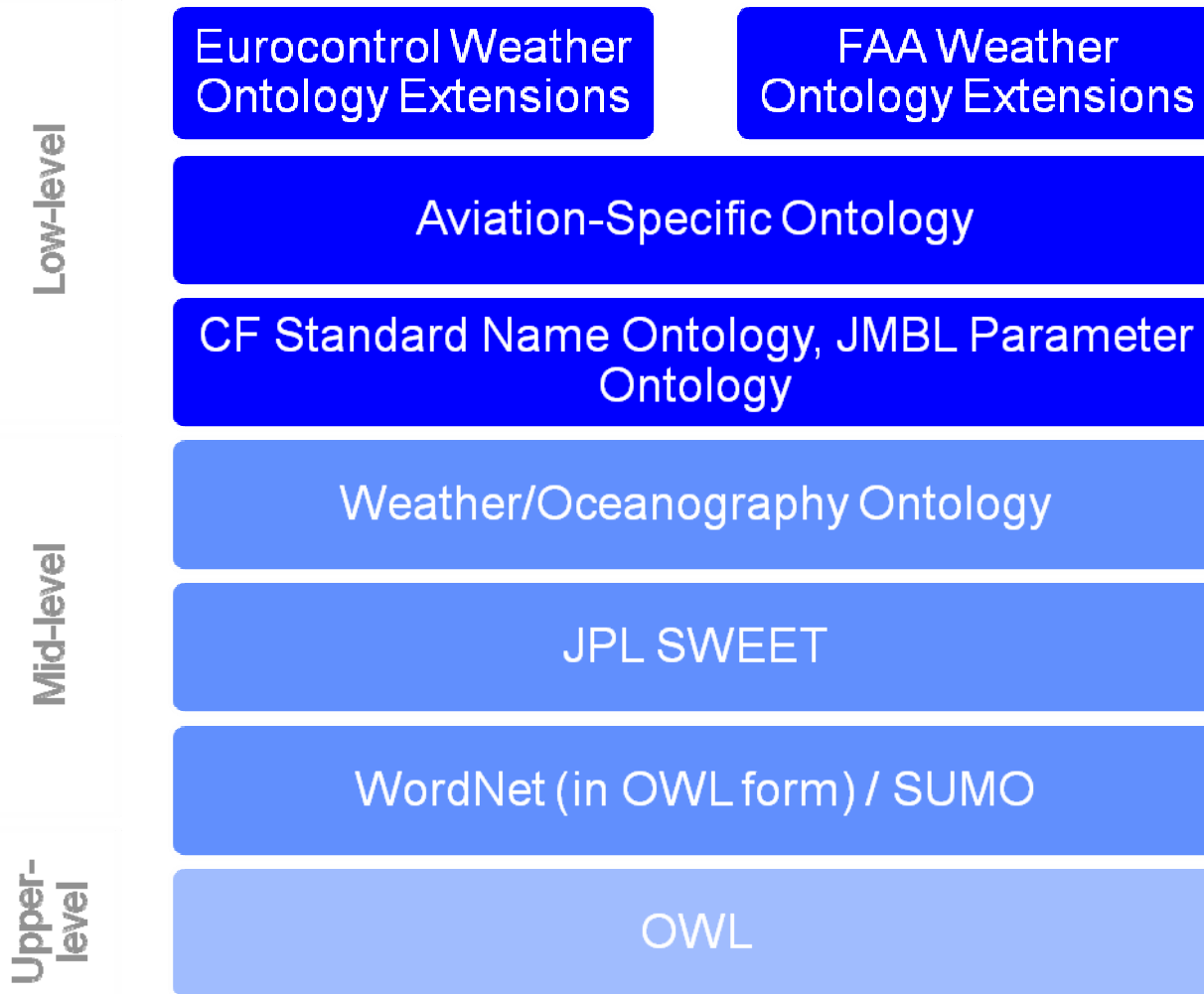


Operational Domain (XML)





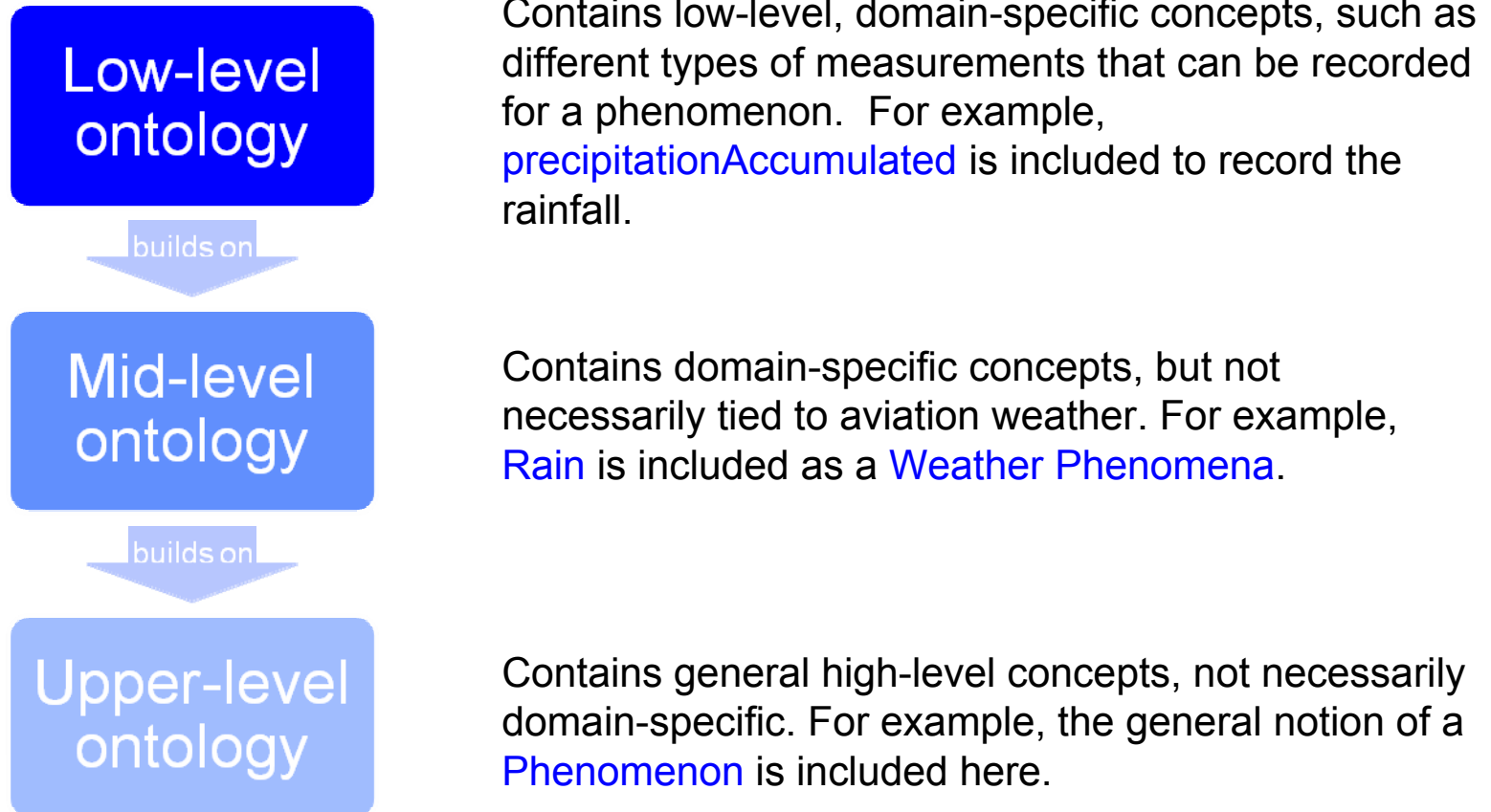
Knowledge Domain (OWL)



Relatively Static Information
(Types, Associations, Classifications)



NNEW Weather Ontology Design





NNEW Weather Ontology Statistics

Humidity

Lightning

Measurement

Precipitation

Pressure

Storm

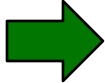
Visibility

Wind

- 8 current focus areas and growing
- 514 unique terms (concepts and properties)
- Weather ontology terms correspond to ≈ 72 (12%) of CF terms and ≈ 330 (24%) of JMBL terms



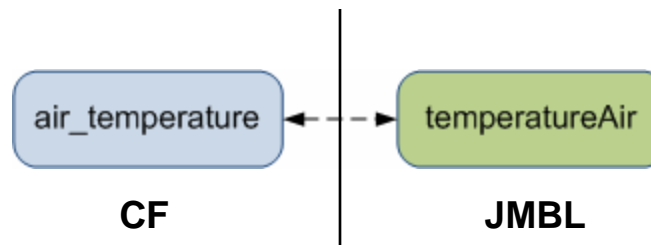
Outline

- Introduction
- Ontology Development Methods & Tools
- NNEW Weather Ontology Design
-  • Application: Semantic Search
- Summary



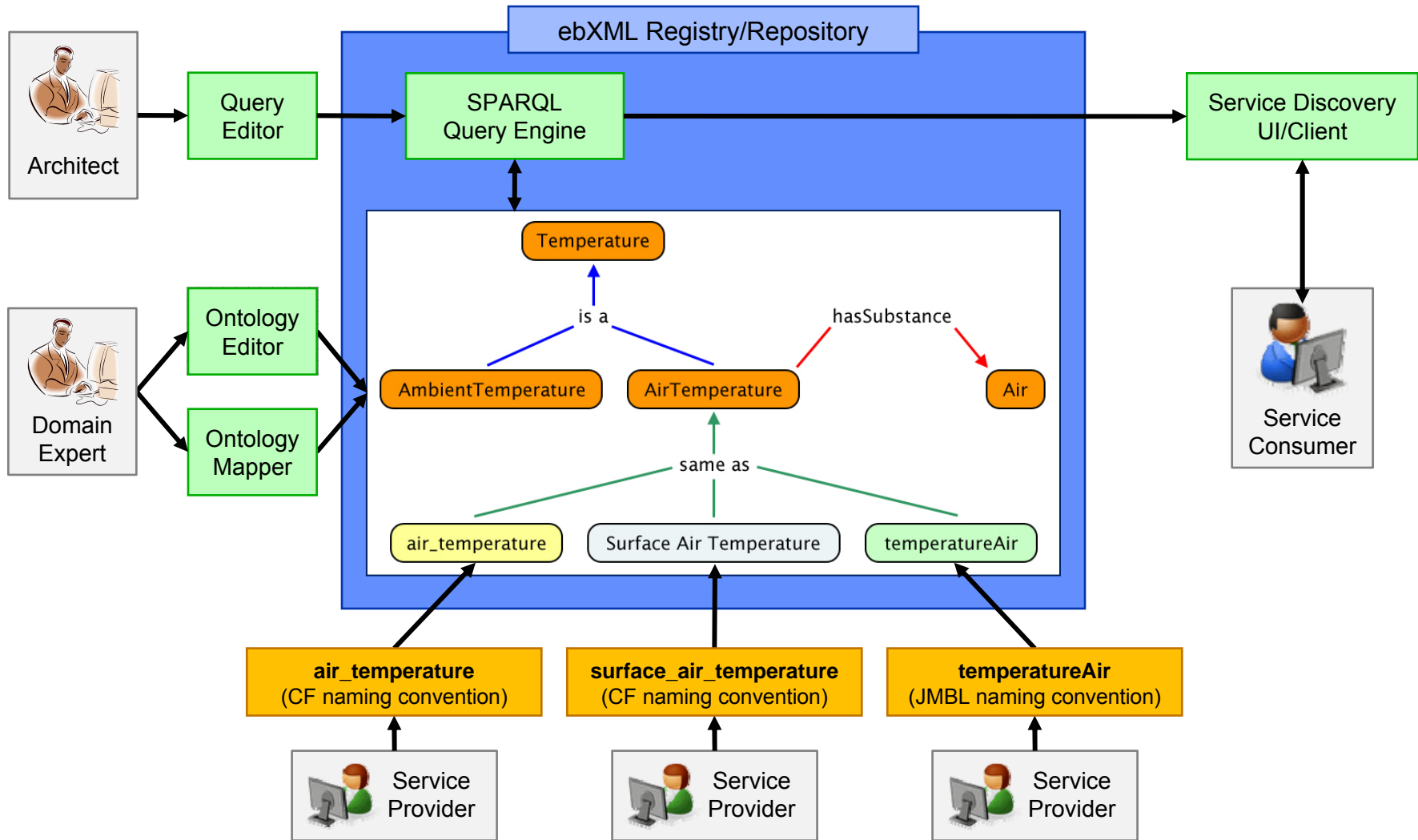
Application: Semantic Search

- Ontologies also enable data integration and semantic searching
- **4-D Wx Data Cube:** Virtual repository that stores distributed weather data
 - Users request data from repository via search
- Producers use different terminology and naming conventions
- Repository must be able to “translate” between terms to return all **semantically similar** data
- **Intelligent discovery**
 - Consults knowledge base/ontology to find alternative meanings
 - Enables discovery of resources without exact keyword match





Application: Semantic Search





Application: Semantic Search

The screenshot shows a web browser window displaying a search application. The search criteria on the left include:

- Search for:** Data Set (OK)
- Field:** temperatureAir
- Threshold:** [empty]
- Search Criteria:** Fuzzy, Title, Abstract, Keywords, Classification, Status, Category
- Where (place)?** Relationship: Please select, Bounding Box

The search results on the right show:

Result 165423

1 result(s) found

| Select | Title | Description | Type | Status |
|--------------------------|---------------------------|--|---------|-----------|
| <input type="checkbox"/> | DOD Model Air Temperature | Air Temperature produced by DOD-run weather model { Types: JMBL.temperatureAir } | Dataset | Submitted |

Annotations with arrows point to:

- Term used for searching:** points to the search field containing "Data Set (OK)".
- Semantic search turned off:** points to the search criteria section.
- Results returned:** points to the search results table.



Application: Semantic Search

Search for: Data Set (OK) [Search]

Field: temperatureAir
Threshold: 0.5

Search Criteria

| Select | Title | Description | Type | Status |
|--------------------------|---------------------------|--|---------|-----------|
| <input type="checkbox"/> | DOD Model Air Temperature | Air Temperature produced by DOD-run weather model { Types: JMBL.temperatureAir } | Dataset | Submitted |
| <input type="checkbox"/> | METARS | Aviation Routine Weather Reports { Types: CF.air_temperature, CF.eastward_wind, CF.northward_wind, CF.dew_point_temperature, CF.surface_air_pressure } | Dataset | Submitted |
| <input type="checkbox"/> | PIREPS | Pilot Reports { Types: CF.air_temperature, CF.eastward_wind, CF.northward_wind, CF.turbulence } | Dataset | |
| <input type="checkbox"/> | PIREPS (ARPS) | Pilot Reports { Types: CF.air_temperature, CF.eastward_wind, CF.turbulence } | Dataset | |
| <input type="checkbox"/> | model air temperature | Air temperature from Rapid Refresh model. {Types: CF.air_temperature} | Dataset | Submitted |
| <input type="checkbox"/> | RUC Model Air Temperature | Air Temperature produced by the RUC-20 model { Types: CF.air_temperature } | Dataset | Submitted |
| <input type="checkbox"/> | | from Real-Time Mesoscale Analysis. {Types: CF.air_temperature} | Dataset | Submitted |
| <input type="checkbox"/> | TAFS | Terminal Area Forecast Reports { Types: CF.air_temperature, CF.eastward_wind, CF.northward_wind, CF.dew_point_temperature, CF.surface_air_pressure } | Dataset | |

Annotations:

- Term used for searching
- Semantic search turned on
- Results returned



Summary

- Ontologies can be used in conjunction with other data modeling methods to enhance their semantic interoperability
 - WXXM can benefit from the inclusion of ontologies
 - Provide semantics for otherwise context-free data
 - Provide consistent use of terminology
 - Enable reuse of domain knowledge (representation of notions of time, measures, etc)
 - Allow for convergence on distinct and mutually agreed-upon definitions
 - Allow for cross-implementation interoperability
 - Promotes quality information sharing
- Utility of ontology demonstrated in conjunction within ebXML registry/repository OWL profile demonstration
 - Enable information integration for related applications



Resources

- Protégé: <http://protege.stanford.edu>
- NNEW Ontology wiki page:
<https://wiki.ucar.edu/display/NNEWD/Data+Models+and+Formats>
- Kelly Moran: kmoran@ll.mit.edu
- Kajal Claypool: claypool@ll.mit.edu



References

- **N. F. Noy and D. L. McGuinness, "Ontology development 101: A guide to creating your first ontology," Online, 2001. [Online]. Available:**
<http://www.ksl.stanford.edu/people/dlm/papers/ontology101/ontology101-noy-mcguinness.html>