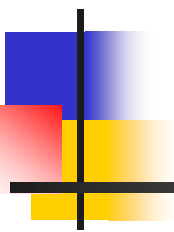


Knowledge as Cyberinfrastructure: Toward Geographic Knowledge Systems (GKS)



Rob Raskin
NASA/Jet Propulsion Laboratory
Pasadena, CA

Feb 7, 2008



Data to Knowledge



Data

Information

Knowledge

Basic Elements

Bytes

Numbers

Models

Facts

Services

Save

Visualize

Infer

Understand

Predict

Storage

File

Database

GIS

Ontology

Mind

Volume/Density

High

Low

Density

Low

High

Syntax

Semantics



What is Knowledge?

- Facts, relations, meanings, contexts, common sense
 - Organized information
- Shared understanding of meaning
- Suitable for reasoning/inference
- Dynamic, expandable
- Information with context



GIS with Knowledge

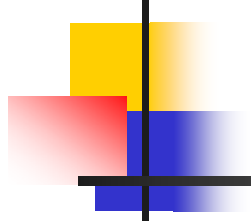
- Data layers registered not just *spatially*, but also **semantically** to features and objects
 - Embedded in metadata, analogous to world files
- Open access to such a system-wide list of potential features/objects



Intelligent Search for Layers

- Consults knowledge base to find alternative meanings
 - Clustered by: synonyms, parent, children
- Enables discovery of resources without exact keyword match
- Semantic understanding is crucial

Semantic Understanding is Difficult!



Variable t: temperature

Data quality= 5 ??????

Variable t: time

Sea surface temperature: measured 3 m above surface

Sea surface temperature: measured at surface

Let's eat, Grandma.

Let's eat Grandma.

Time flies like an arrow.

Fruit flies like a pie.

“Mission accomplished. Major combat operations in Iraq have ended” -G. Bush

**Low-Profile
Rice Has
Bush's Ear**

LA Times headline



Knowledge Base of Facts as Triples

Subject-Verb-Object representation

- Flood is a WeatherPhenomena
- GeoTIFF is a FileFormat
- Soil Type is a PhysicalProperty
- Pacific Ocean is a Ocean

- Ocean has substance Water
- Sensor measures Temperature



Semantic Web Vision

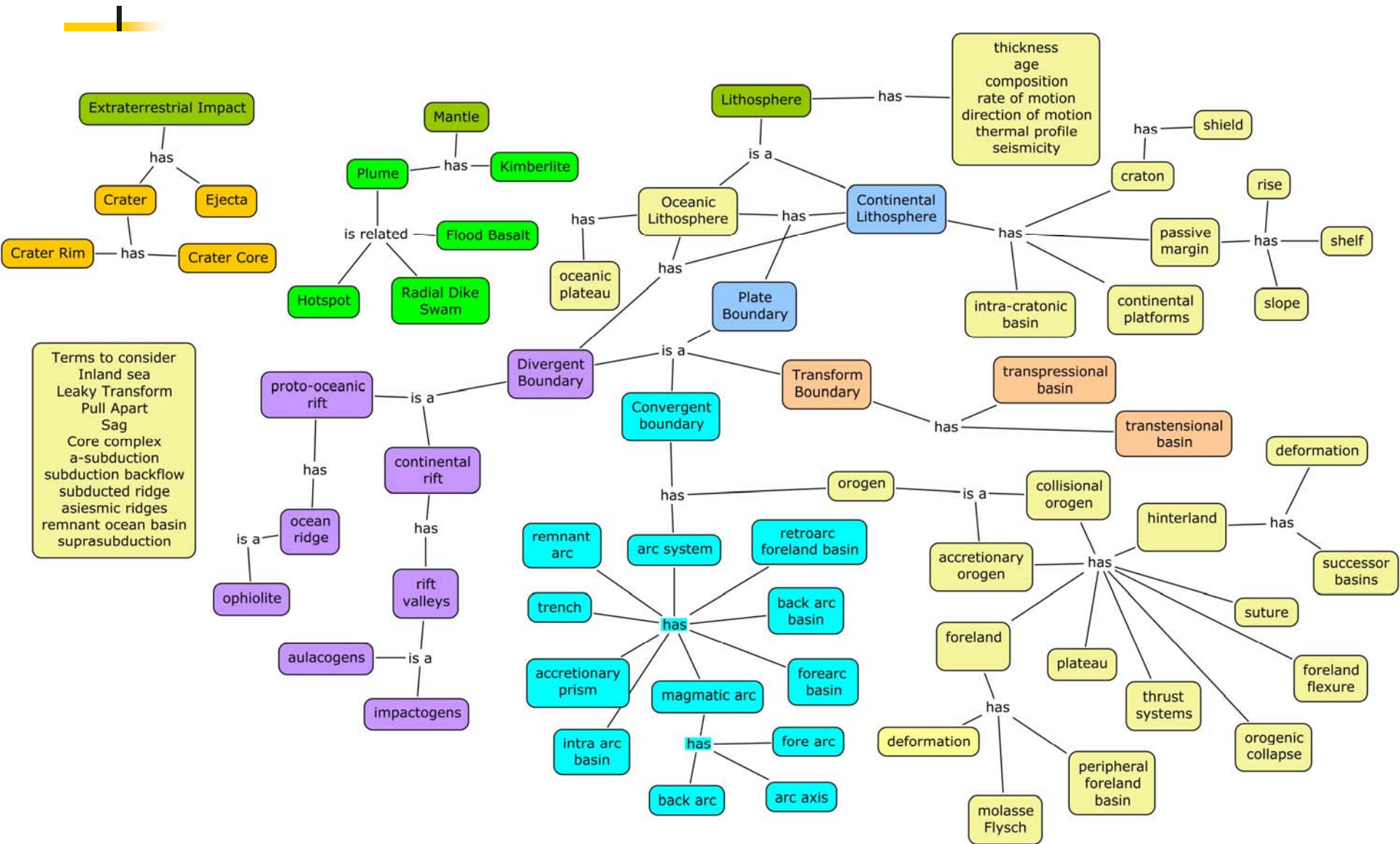
- Web page creators place XML tags around technical terms on web pages
- XML tags point to knowledge base where term is “defined”
- Search tools use this information to provide value-added services
 - Common search engines (Google) use these capabilities only minimally, at present



Ontologies

- Method to store “facts”
- General definition: “all that is known”
- Computer science definition: Machine-readable definition of terms and how they relate to one another
 - As with a dictionary, terms are defined in terms of other terms
- Provide shared understanding of concepts
- Support **knowledge reuse**
- Support machine-to-machine communications with deeper semantics than controlled vocabulary


Plate Tectonics Ontology





Ontology Languages: RDF and OWL

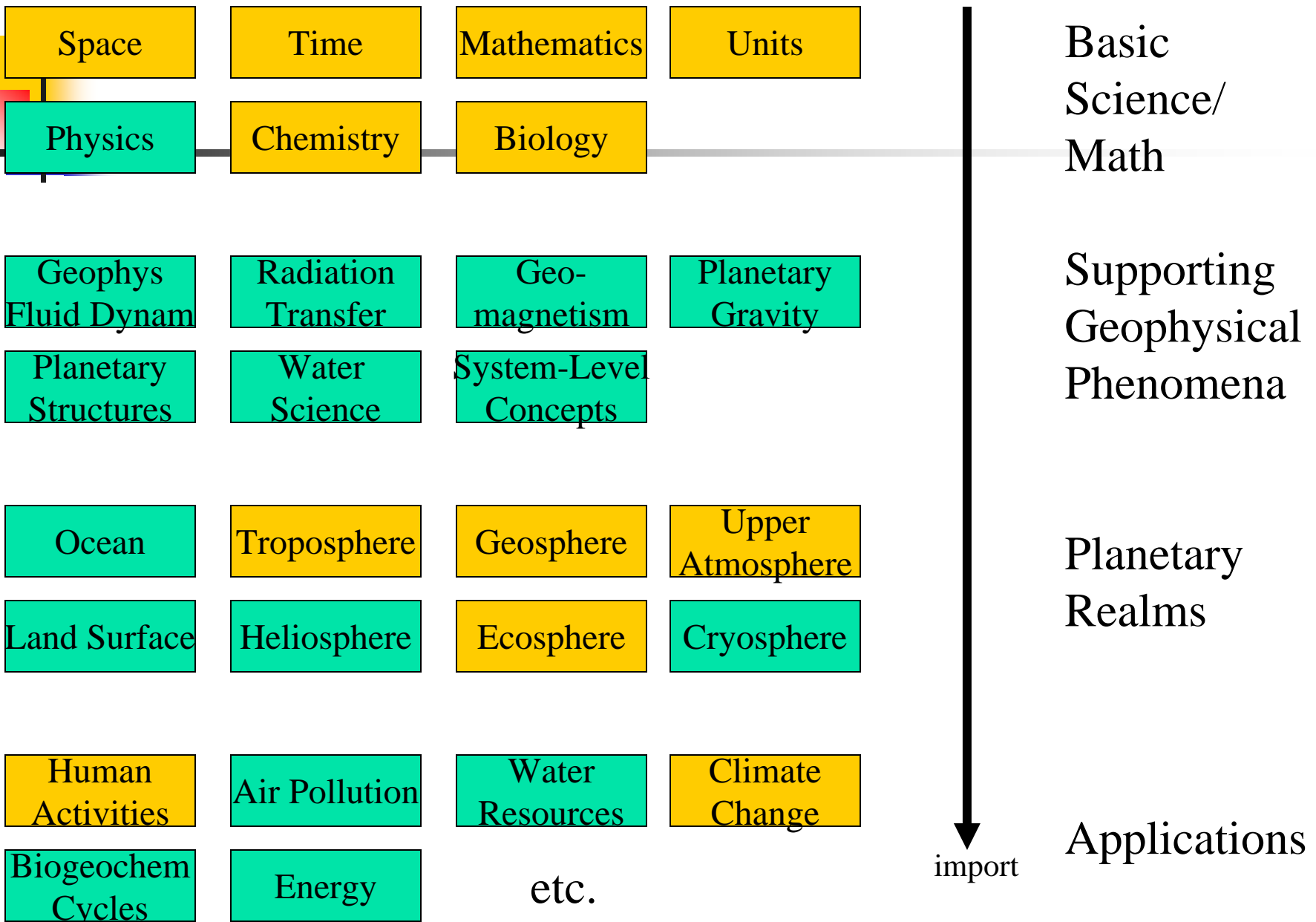
- Use of standard languages make it easy to extend (specialize) concepts developed by others
- World Wide Consortium (W3C) has adopted languages that specialize XML
 - Resource Description Formulation (RDF)
 - Ontology Web Language (OWL)



Semantic Web for Earth and Environmental Terminology (SWEET)

- SWEET is a concept space written in OWL
- Enables scalable *classification* of Earth system science concepts
- Funded by NASA
 - Initial focus to assist search for data resources

SWEET Ontologies: Modular Design





SWEET Spatial Ontology

- Concepts of 0-D, 1-D, 2-D, and 3-D objects
- Default coordinate system: lat/lon/up
- Polygons used to store spatial extents
 - Spatial attributes added (population, area, etc.)
- Spatial relations in 3-D
 - Above, inside, fuzzy concepts (“near”)



SWEET Data Ontology

- Dataset characteristics
 - Format, data model, dimensions, ...
- Special values
 - Missing, land, sea, ice, ...
- Parameters
 - Scale factors, offsets, ...
- Data services
 - Subsetting, reprojection, ...
- Quality measures

Fragment of SWEET

3DLayer

subClassOf

PlanetaryLayer

partOf

primarySubstance
="air"

Atmosphere

partOf

AtmosphereLayer

subClassOf

subClassOf

upperBoundary
=50 km

sameAs=

"Lower
Atmosphere"

Troposphere

Stratosphere

lowerBoundary
=15 km

isUpperBoundaryOf

isLowerBoundaryOf

Tropopause



Best Practices

- Keep ontologies small, modular
 - Be careful that “Owl:Import” imports everything
 - Use higher level ontologies where possible
- Identify hierarchy of concept spaces
 - Model schemas
 - Try to keep dependencies unidirectional



Ontology Development Tools: CMAP

- Free, downloadable tool for knowledge representation and ontology development
- Visual language with input/export to OWL
 - Supports subset of OWL language
- <http://cmap.ihmc.us/coe>



Resources

- SWEET
 - <http://sweet.jpl.nasa.gov>
- Ontology development/sharing site
 - <http://PlanetOnt.org>



Other Geospatial-CI Activities

- New AAG Cyberinfrastructure Specialty Group (CISG)
 - Approved Nov 2007
- AAG Annual Meeting
 - 5 CI Sessions plus Panel
- GIScience 2008 Workshops
 - Service-oriented Architectures for Geospatial Science
 - Cyberinfrastructure for GIS



AAG CI Specialty Group

- Purpose:
 - Encourage exchange of ideas and experiences in CI
 - Promote geography research in CI
 - Serve as a communication channel between CI funding agencies and geographers
 - Inform the AAG community of opportunities to fully exploit computational resources
 - Encourage broader access to computing resources, including to underrepresented groups



CI Specialty Group (cont.)

- 7 Officers needed!
 - Chair, Vice-Chair, Secretary/Treasurer, 3 additional Directors, Student Director
 - Election at April 2008 AAG Meeting
 - Business Meeting on Friday night
- Web site
 - <http://cisg.gmu.edu>