Consider a Spherical Cow

aka “Semantic and Knowledge-Based Mechanisms for Accessing and Using Data”

Jim Myers, Chief Scientist
Computational Science and Mathematics Directorate
Computational and Information Sciences Division
Pacific Northwest National Laboratory

Why?

- There is a class of bovine-related problems for which shape is not important
- Yet shape is clearly needed in a general cow model

- Should we “reach consensus” here?

What is semantic/knowledge technology good for?

- A) Making information computer readable?
- B) Enable collaboration (by standardizing data models)?
- C) Decoupling?
  - Model from software
  - Model from model
- D) Standardization of model representation allowing model manipulation

- Only C & D, driven by requirements to do A) and B)

- All that is left coupled is the minimum necessary to do enabled A) and B)

DM Workshop Conclusions

- Key Unmet Requirements:
  - Data Description, Provenance, Annotation
  - Data Integration
  - Integrated Analysis and Visualization Environments

    … that scale in terms of data size, data heterogeneity, community size, and rapidly advancing science

NC Meeting Conclusions?

- Unmet challenges in
  - Tool Description, Provenance, Annotation
  - Tool Integration
  - Integrated Collaboration and Problem Solving Environments

  … that scale in terms of tool performance, tool heterogeneity, community size, and rapidly advancing technologies

A Broad View of Semantic Technologies

- Controlled vocabularies, schema, ontologies, inference, ontological mapping, …
- Metadata services – exposure of data through standardized mechanisms
- Publish/Subscribe Messaging
- Aspect-Oriented Design/Programming
- Translation/Mapping as a first class operation
- Agents…

metadata

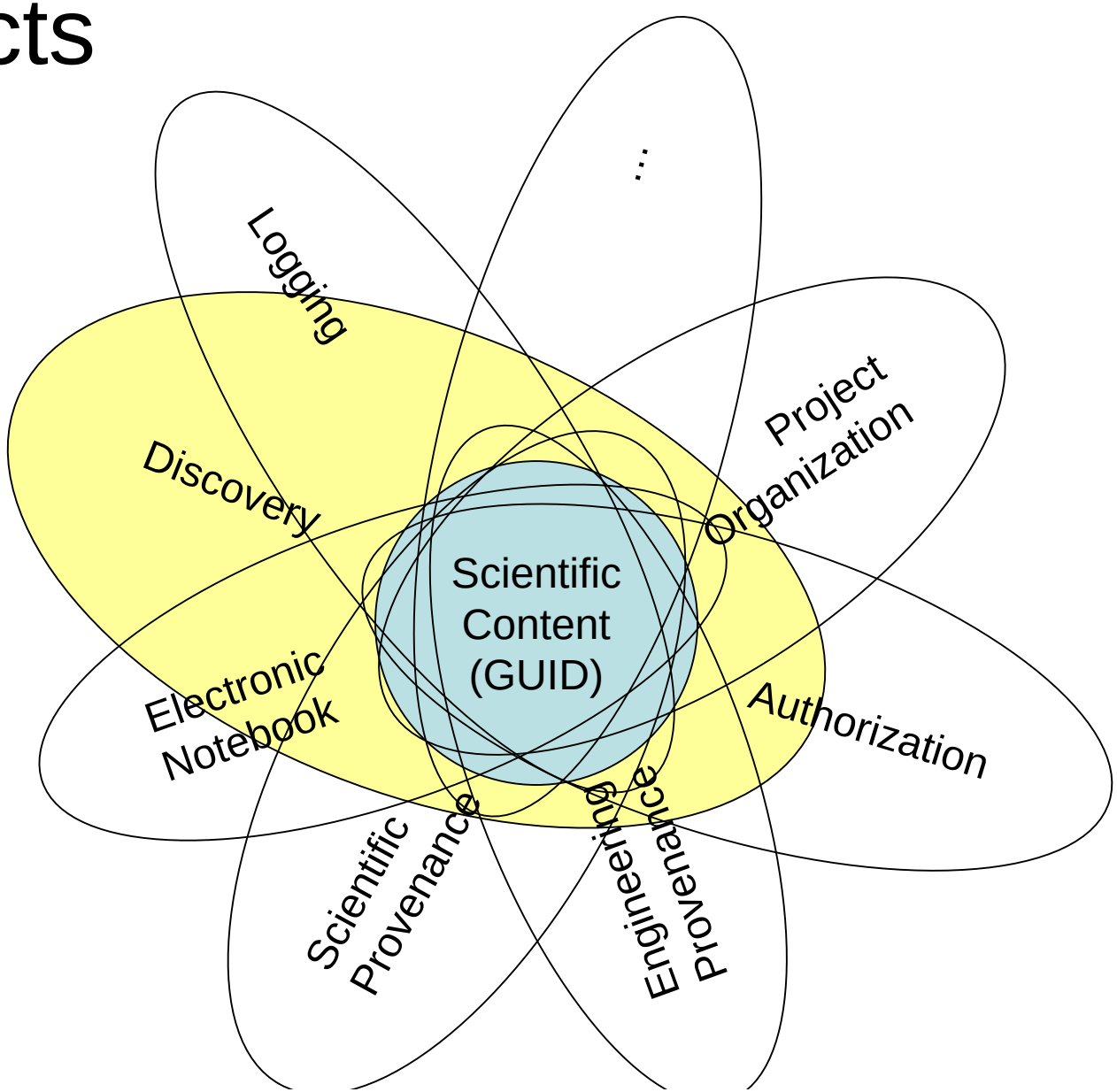
- www.dictionary.com:
  - n : data about data
  - Not to be confused with Metadata

- Jim:
  - somebody else's data model exposed for use outside their application

Aspects:

- allows programmers to modularize *crosscutting concerns*
- encapsulate behaviors that affect multiple types of data/objects into reusable modules (~orthogonal models)
- typical behaviors involving data do not fit naturally into a single program module or data model

Metadata – a way to deal with aspects



Examples

- JAAS
- Data Format Description Language
  - Identify structures within binary/ascii data
- Scientific Annotation Middleware, SRB, Data Turbine, Narada, CML, ESML, …

Semantics Across the NC Program

- C) Decoupling - model from software, model from model
- D) Standardization of model representation allowing model manipulation

- Data/Content
- Collaborative Suites (session Management, sharing mechanisms)
- Distributed Execution/Workflow
- Security
- Portals
- Problem Solving Environments/ IDAVEs

The Science Bazaar

"When I use a word," Humpty Dumpty said, in rather a scornful tone, "it means just what I choose it to mean--nether more nor less."

"The question is," said Alice, "whether you can make words mean so many different things.“

"The question is," said Humpty Dumpty, "which is to be the master--that's all."

- Semantic Technologies decouple concerns allowing everyone to be the master of their domain