NASA Taxonomy 2.0
Project Overview

Jayne Dutra, Jet Propulsion Laboratory,
California Institute of Technology, NASA
KM Gov Semantics and Taxonomies SIG
March 8, 2004
Purpose of the NASA Taxonomy

Create

- Content
- Assets

Classify

- Logical & Intuitive Filters
- Taxonomy

Discover

FIRSTGov

- Site Maps
- Search Engine
- NASA Portals
- Content Integration Networks

Finding the right information at the right time to solve the problem at hand

California Institute of Technology
NASA/ Taxonomy Strategies
Project Benefits:
Enable Knowledge Discovery

- Make it easy for various audiences to find relevant information from NASA programs quickly
  - Provide easy access for NASA resources found on the Web
  - Share knowledge by enabling users to easily find links to databases and tools
  - Provide search results targeted to user interests
  - Enable the ability to move content through the enterprise to where it is needed most

- Comply with E-Government Act of 2002
- Be a leading participant in federal XML projects
Project Benefits:
NASA Taxonomy Best Practices

- Design process that:
  - Incorporates existing federal and industry terminology standards like NASA AFS, NASA CMS, FEA BRM, NAICS, and IEEE LOM
  - Provides a product for the NASA XML namespace registry
  - Complies with metadata standards like Z39.19, ISO 2709, and Dublin Core

- Practices increase interoperability and extensibility
What is the NASA Taxonomy?

- The classification scheme is meant to encompass all of NASA web content (NASA web space) including internal as well as external material. It is a means for tagging content so it can be used and reused in different contexts.

How to Use the NASA Taxonomy

- This is a generic taxonomy from which specializations can be derived for specific purposes.
  - Not all facets need to be used in each instance
  - A facet is repeatable
  - The taxonomy is modular and dynamic
Project Progress: NASA Taxonomy 2.0 Revision Completed

Follow-on Work
➢ Integrate with applications

Phase 4
➢ Dublin Core mapping
➢ XML schema development

Phase 3
➢ Test & validate Taxonomy

Phase 2
➢ Build community of practice
➢ Agree on comprehensive branches & taxonomy detail

Project 1

Project 2

California Institute of Technology
NASA/ Taxonomy Strategies
Objectives of Phases 1-2

- Understand current strategies and practices for creating, collecting, and organizing information across NASA
- Observe how information is used and organized, the audiences for this information, and the information needs of these audiences
- Elicit goals, hopes, and concerns for an information architecture solution
- Start building a community of interest
- Mainly concentrated on content for outreach (NASA portal)
Key Findings

- Most (70%) NASA content already has some tagging or is categorized.
- Most (70%) owners add tag content with metadata.
- Almost half (45%) use a standard metadata schema, but many different standards are used.
- Most (60%) use a controlled vocabulary, but lots of different controlled vocabularies are used.

Different NASA constituencies care deeply about what schemas are specified, and what vocabularies are used because ...
## Audience Uses Vary Widely

- Better understand the program in total, and obtain scheduling information, project status and best practices.
- Access procurement rules and examples, and procurement action synopses.

- Engineering specifications.
- Scholarly research, competitive intelligence, and general aerospace research.

- Develop educational products, support current products, learn, etc.
- Topic research and fact finding, topic background research, and downloading curriculum support materials.
- In the classroom as stand alone items, hands-on learning opportunities, class projects, to expand on a student's learning potential.
- Find NASA contact information on services, information about student opportunities, information about career opportunities, and latest educational news.
Test and Validation Phase

- Qualitative validation
  - Confirm stakeholders and communities
  - Focus on Projects, Engineering & Science

- Quantitative validation
  - Select and build test collection
  - Stratify automated categorizer – Ames support

- Extend taxonomy value space as needed

- Review results with stakeholders and report to CIO Council
Engaged Cross-Section of NASA Community

78 interviews across the Agency

... by location

California Institute of Technology
NASA/ Taxonomy Strategies
Focused on Projects, Engineering & Science

... by primary audience served.

52%—Projects, Engineering & Science

California Institute of Technology
NASA/ Taxonomy Strategies
Extend Taxonomy
Value Space as Needed

NASA Taxonomy Facets (Top Level)

- Access Requirements (new)
- Audiences
- Business Purpose (formerly Functions)
- Competencies (formerly Disciplines)
- Content Types (formerly Information)
- Industries
- Instruments (new)
- Locations
- Missions and Projects
- Organizations
- Subject Categories (new)
- Dates (formerly Chronology)
- Collections

## Taxonomy Depth and Breadth

<table>
<thead>
<tr>
<th>Facets</th>
<th># Terms</th>
<th># Levels Deep</th>
</tr>
</thead>
<tbody>
<tr>
<td>Access Requirements</td>
<td>5</td>
<td>2</td>
</tr>
<tr>
<td>Audiences</td>
<td>62</td>
<td>6</td>
</tr>
<tr>
<td>Business Purpose</td>
<td>96</td>
<td>4</td>
</tr>
<tr>
<td>Competencies</td>
<td>169</td>
<td>4</td>
</tr>
<tr>
<td>Content Types</td>
<td>96</td>
<td>4</td>
</tr>
<tr>
<td>Industries</td>
<td>22</td>
<td>3</td>
</tr>
<tr>
<td>Instruments</td>
<td>56</td>
<td>3</td>
</tr>
<tr>
<td>Locations</td>
<td>106</td>
<td>8</td>
</tr>
<tr>
<td>Missions/Projects</td>
<td>648</td>
<td>6</td>
</tr>
<tr>
<td>Organizations</td>
<td>323</td>
<td>6</td>
</tr>
<tr>
<td>Subject Categories</td>
<td>78</td>
<td>3</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>1661</strong></td>
<td></td>
</tr>
</tbody>
</table>

California Institute of Technology  
NASA/ Taxonomy Strategies
• **Content Types**
  - Announcements
  - Press Kits
  - Press Releases
  - Articles, Notes, and Papers
  - Calendars and Schedules
  - Agendas
  - Case Studies
  - Catalogs and Databases
  - Correspondence
  - e-Mails
  - Memos
  - Databases
  - Bibliographic Databases
  - Image Databases
  - Designs and Specifications
  - Configuration Controls
  - Notebooks
  - Quality Control
  - Requirements
  - Drawings
  - Educational Materials
  - Activity Guides
  - Educational Toys
  - Educator's Guides

---

**Configuration Controls**

**Broader Terms:**
- Designs and Specifications

**Scope Note:**
Records of changes to documentation or hardware, including engineering change requests and waivers.

**Term Number:**
52
NASA Taxonomy Web Site Features

- Query link into FirstGov
- Quick access to Google
- A-Z index of facet branch
- Easy navigation to top level facet and term index

California Institute of Technology
NASA/ Taxonomy Strategies
<table>
<thead>
<tr>
<th>Collection</th>
<th>Source URL</th>
<th>No of Docs</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lessons Learned Database</td>
<td><a href="http://llis.nasa.gov">http://llis.nasa.gov</a></td>
<td>1,370</td>
</tr>
<tr>
<td>SIRTF (Space Infrared Telescope Facility)</td>
<td><a href="http://sirtifweb.jpl.nasa.gov">http://sirtifweb.jpl.nasa.gov</a></td>
<td>4,054</td>
</tr>
</tbody>
</table>
...that provides common access framework across test collections
Search on “rover” Returns 595 Matches ...Organized by Taxonomy

by Organization
- NASA Affiliated Institutions 6
- NASA Centers 304
- NASA Contractors 18
- NASA Enterprises 3
- NASA Headquarters 7
- Other NASA Partners 14

by Subject
- Aeronautics 13
- Astronautics 109
- Engineering 106
- Geosciences 26
- Life Sciences 31
- Mathematical and Computer Sciences 65
- Space Sciences 208

by Missions and Projects
- Data 1
- Earth Sciences 3
- Human Exploration and Development... 130
- Planetary Missions 125
- Space Sciences 7

by Date
- 1989 37
- 1991 40
- 1992 44
- 1993 44
- 1999 61
- 2000 60
- 2001 40
- 29 more

by Competencies
- Business 4
- Engineering 4
- Mission 3
- Scientific 3

by Information Type
- Reviews and Lessons Learned 5
- Status Reports 1
- Technical Reports 3

by Collection
- LessonsLearned 4
- NTRS 591

California Institute of Technology
NASA/ Taxonomy Strategies
Taxonomy Enables Discovery of Unknown but Related Content

5 items matching
Text contains rover

Information Type:
Reviews and Lessons Learned

by Organization
- NASA Centers 5
- NASA Enterprises 3

by Subject
- Aeronautics 4
- Astronautics 4
- Chemistry and Materials 1
- Engineering 1
- Mathematical and Computer Sciences 1

by Date
- 1997 1
- 2002 1
- 2003 3

by Competencies
- Business 4
- Engineering 4
- Mission 3
- Scientific 3

by Collection
- Lessons Learned 4
- NTRS 1

California Institute of Technology
NASA/ Taxonomy Strategies
Search & Browse Demo Site


Logon: NASA
Password: facets

Hosted by Seamark www.seamark.com with Siderean software
Taxonomy Phase 4

Dublin Core Mapping and XML Schema Development

- Complete Dublin Core mapping
- Create any necessary NASA specific tags
- Develop XML schema from metadata
- Review results with stakeholders and report to CIO Council
- Register schemas in NASA XML Registry
- Educate and train publishing communities
<table>
<thead>
<tr>
<th>Dublin Core Elements</th>
<th>Definition</th>
<th>NASA Taxonomy Mapping</th>
</tr>
</thead>
<tbody>
<tr>
<td>Creator</td>
<td>Content maker.</td>
<td>dc:creator</td>
</tr>
<tr>
<td></td>
<td></td>
<td>dc:creator.employee</td>
</tr>
<tr>
<td></td>
<td></td>
<td>dc:creator.organization</td>
</tr>
<tr>
<td>Subject</td>
<td>Content topic.</td>
<td>dc:subject.organization</td>
</tr>
<tr>
<td></td>
<td></td>
<td>dc:subject.missionsProjects</td>
</tr>
<tr>
<td></td>
<td></td>
<td>dc:subject.disciplines</td>
</tr>
<tr>
<td>Publisher</td>
<td>Publisher of this</td>
<td>dc:creator</td>
</tr>
<tr>
<td></td>
<td>manifestation.</td>
<td>dc:creator.employee</td>
</tr>
<tr>
<td></td>
<td></td>
<td>dc:creator.organization</td>
</tr>
<tr>
<td>Contributor</td>
<td>Content contributor.</td>
<td>dc:creator</td>
</tr>
<tr>
<td></td>
<td></td>
<td>dc:creator.employee</td>
</tr>
<tr>
<td></td>
<td></td>
<td>dc:creator.organization</td>
</tr>
<tr>
<td></td>
<td>Genre.</td>
<td>dc:creator</td>
</tr>
<tr>
<td></td>
<td></td>
<td>dc:creator.employee</td>
</tr>
<tr>
<td></td>
<td></td>
<td>dc:creator.organization</td>
</tr>
<tr>
<td>Coverage</td>
<td>Space, period, date,</td>
<td>dc:coverage.locations</td>
</tr>
<tr>
<td></td>
<td>jurisdiction, etc.</td>
<td>dc:coverage.chronology</td>
</tr>
<tr>
<td>Audience</td>
<td>Content audience.</td>
<td>dc:coverage.locations</td>
</tr>
<tr>
<td></td>
<td></td>
<td>dc:coverage.chronology</td>
</tr>
<tr>
<td>Non DC</td>
<td>NASA missions and projects.</td>
<td>nasa:missionsProjects</td>
</tr>
<tr>
<td></td>
<td>Business functions.</td>
<td>nasa:functions</td>
</tr>
<tr>
<td>Non DC</td>
<td>Technical specialties.</td>
<td>nasa:disciplines</td>
</tr>
<tr>
<td></td>
<td>Standard industry categories.</td>
<td>naics:industries</td>
</tr>
</tbody>
</table>
Project Outcomes: NASA Taxonomy Benefits

... at the NASA Level

- Metadata specification for all NASA content publishers
- XML schema registered in accordance with agreed standards (to enable appropriate use and reuse)
- Enhancement of Agency Web publishing processes
- Integration with NASA public portal content management system for:
  - Reduced publishing cycles
  - Better quality of Web materials – coordinated themes
- Integration with NASA Search Engine, Web Site Registration System
- Application in many technical areas, including engineering and science disciplines (STEP and science data dictionaries)
Project Outcomes: NASA Taxonomy Benefits

... at the Federal Level

- NASA taxonomy development in accordance with e-Gov Act of 2002
- Integration with FEA at the BRM & DRM level
- Increased interoperability with other federal agencies through common data models and standards
- Better interoperability with industry partners for increased speed of mission development
- Enhanced results in First Gov search engine
- Readiness to actively participate in e-Gov initiatives
Taxonomy Follow-on Work

- Taxonomy stewardship
  - Maintenance, education and training
  - Facilitate standard adoption process
- Apply in public and internal portals
  - DM, PM, CM, and DAM tagging
  - Search integration
    - Verity K2
    - Faceted search and navigation
  - content integration networks for real time delivery
- Implement in website registration system
- Participate in NASA Enterprise Architecture Group standards development
  - Web Services infrastructure: Agency UDDI, RSS specifications, directory capability
Next Steps

- CIO presentation
- Define review and approval process
- Facilitate review and approval process
- Deliver metadata specification
- Formalize taxonomy as XML schema
- Delegate taxonomy ownership
  - Within Code V
- Plan follow-on work – phased implementation
White Papers and Supporting Documentation

- NASA Taxonomy Ver. 2.0, Presentation, Dutra 3/8/04 (*this presentation*)


*California Institute of Technology
NASA/ Taxonomy Strategies*
Wrap Up and Discussion

Thanks for your time!

Jayne.E.Dutra@jpl.nasa.gov

California Institute of Technology
NASA/ Taxonomy Strategies
Taxonomy Tagging

Examples
# 2001 Mars Odyssey Data Archives

http://wufs.wustl.edu/missions/odyssey/#Odyssey%20Data%20Sets

<table>
<thead>
<tr>
<th>Attribute</th>
<th>Values</th>
</tr>
</thead>
<tbody>
<tr>
<td>Content Types</td>
<td>Data Files; Web Sites</td>
</tr>
<tr>
<td>Audiences</td>
<td>Researchers; Scientists</td>
</tr>
<tr>
<td>Organizations</td>
<td>Jet Propulsion Laboratory</td>
</tr>
<tr>
<td>Missions and Projects</td>
<td>Mars Odyssey</td>
</tr>
<tr>
<td>Industries</td>
<td>N/A</td>
</tr>
<tr>
<td>Locations</td>
<td>Mars</td>
</tr>
<tr>
<td>Business Purpose</td>
<td>Scientific and Technical Information</td>
</tr>
<tr>
<td>Competencies</td>
<td>Planetary and Lunar Science</td>
</tr>
<tr>
<td>Dates</td>
<td>2002-present</td>
</tr>
</tbody>
</table>

California Institute of Technology
NASA/ Taxonomy Strategies
## Clementine – DSPSE

http://www.cmf.nrl.navy.mil/clementine/

<table>
<thead>
<tr>
<th>Attribute</th>
<th>Values</th>
</tr>
</thead>
<tbody>
<tr>
<td>Content Types</td>
<td>Web Sites; Data Files; Images</td>
</tr>
<tr>
<td>Audiences</td>
<td>Researchers; Scientists; Educators; Students</td>
</tr>
<tr>
<td>Organizations</td>
<td>Naval Research Laboratory</td>
</tr>
<tr>
<td>Missions and Projects</td>
<td>Clementine</td>
</tr>
<tr>
<td>Industries</td>
<td>N/A</td>
</tr>
<tr>
<td>Locations</td>
<td>The Moon</td>
</tr>
<tr>
<td>Business Purpose</td>
<td>Scientific and Technical Information</td>
</tr>
<tr>
<td>Competencies</td>
<td>Planetary and Lunar Science</td>
</tr>
<tr>
<td>Dates</td>
<td>1994</td>
</tr>
</tbody>
</table>

California Institute of Technology
NASA/ Taxonomy Strategies
# Jupiter's Ring System

http://ringmaster.arc.nasa.gov/jupiter/jupiter.html#index

<table>
<thead>
<tr>
<th>Attribute</th>
<th>Values</th>
</tr>
</thead>
<tbody>
<tr>
<td>Content Types</td>
<td>Web Sites; Animations; Images; Reference Sources</td>
</tr>
<tr>
<td>Audiences</td>
<td>Educators; Students</td>
</tr>
<tr>
<td>Organizations</td>
<td>Ames Research Center</td>
</tr>
<tr>
<td>Missions and Projects</td>
<td>Voyager; Galileo; Cassini; Hubble Space Telescope</td>
</tr>
<tr>
<td>Industries</td>
<td>N/A</td>
</tr>
<tr>
<td>Locations</td>
<td>Jupiter</td>
</tr>
<tr>
<td>Business Purpose</td>
<td>Scientific and Technical Information</td>
</tr>
<tr>
<td>Competencies</td>
<td>Planetary and Lunar Science</td>
</tr>
<tr>
<td>Dates</td>
<td>1979-1999</td>
</tr>
</tbody>
</table>

**California Institute of Technology**  
**NASA/ Taxonomy Strategies**