

# **NASA Taxonomy 2.0 Project Overview**

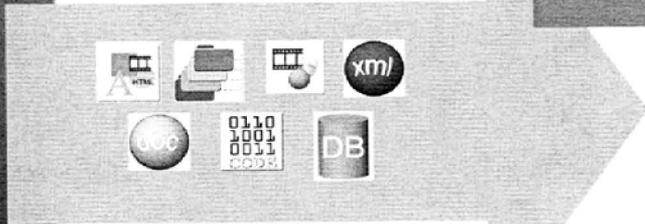
*Jayne Dutra, Jet Propulsion Laboratory,  
California Institute of Technology  
Joseph Busch, Taxonomy Strategies  
ASIS Conference  
Providence, Rhode Island  
November 16, 2004*



# Publishing Cycle in the Real Time Organization

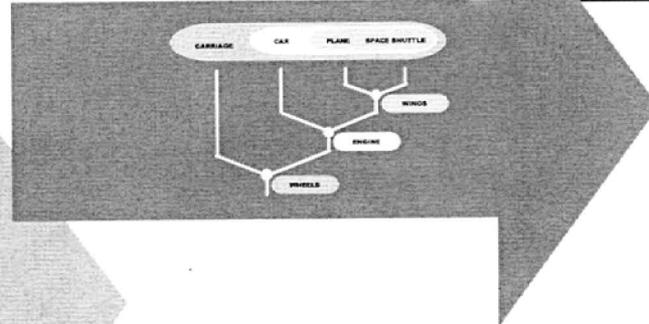


## Create



- Content
- Assets

## Classify



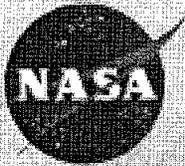
- Logical & Intuitive Filters
- Taxonomy

## Discover



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

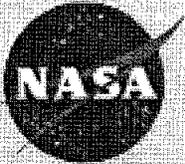
Finding the right information at the right time to solve the problem at hand



# Project Benefits: Enable Knowledge Discovery

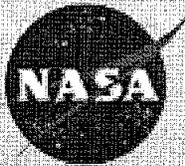


- Make it easy for various audiences to find relevant information from NASA programs quickly
  - Provide easy access for NASA Web resources
  - Information integration for unified queries and management reporting
  - Improve search results targeted to user interests
  - Enable the ability to move content through the enterprise to where it is needed most
  - Facilitate Records Management and Retention Requirements
- Comply with E-Government Act of 2002
- Be ready to participate in federal XML projects



## ***Best Practices increase interoperability and extensibility***

- **Faceted Classification Schema**
  - Facets give flexibility and power
  - Modular in nature for easier maintenance
  - Can tag what is appropriate to the use case
- **Polyhierarchy**
  - Concepts can appear more than once
  - Enables knowledge discovery from multiple viewpoints
  - User-centric organization



# NASA Taxonomy Best Practices

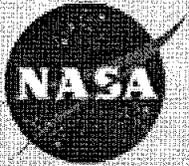


- **Hierarchical Granularity**

- Different levels of depth depending on attribute set and content
- The NASA taxonomy is broad in nature by design
- Integration points allow for mapping of local vocabulary terms back to larger semantic framework

- **Use of Standards**

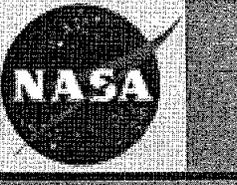
- Incorporates existing federal and industry terminology standards like NASA AFS, NASA CMS, FEA BRM, NAICS, and IEEE LOM
- Complies with metadata standards like Z39.19, ISO 2709, and Dublin Core
- Political buy in from COIs



- **Mapping Aliases**

- Librarian approach allows for a meaningful expression of relationships
- Synonyms, acronyms, related terms, broader and narrower terms
- Easily maintained in derived RDF files
  
- *Example: MER A= Spirit, MER B= Opportunity*
- *Example: Section 366 = Section 372*
- *Example: JIMO = Jupiter Icy Moons Orbiter = Prometheus 1*

***A means of holding our institutional memory***



## Project Progress: Initial Taxonomy Development

### Phase 1

- Audit Content practices
- Identify & survey stakeholders

### Phase 2

- Build community of practice
- Agree on comprehensive branches & taxonomy detail

### Phase 3

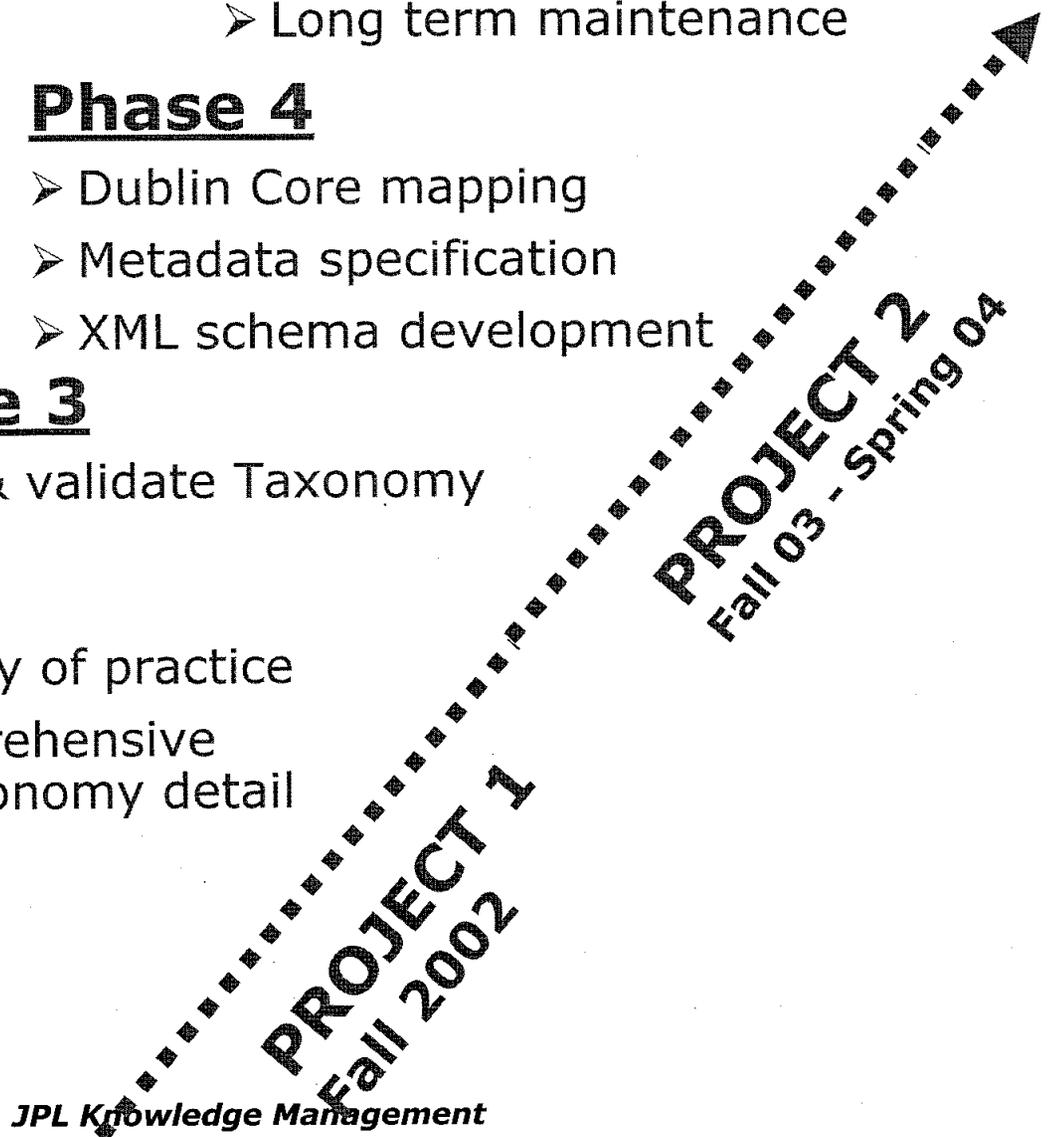
- Test & validate Taxonomy

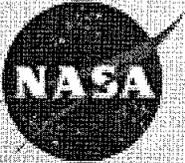
### Phase 4

- Dublin Core mapping
- Metadata specification
- XML schema development

### Follow-on Work

- Integrate with applications
- Long term maintenance



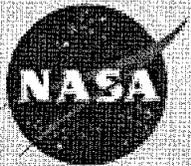


# Extend Taxonomy Value Space as Needed



## NASA Taxonomy Facets (Top Level)

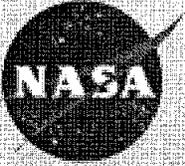
- Access Security Requirements (*new*)
  - Audiences
  - Business Purpose (*formerly Functions*)
  - Competencies (*formerly Disciplines*)
  - Content Types (*formerly Information*)
  - Industries
  - Instruments (*new*)
  - Locations
  - Missions and Projects
  - Organizations
  - Subject Categories (*new*)
- 
- Dates (*formerly Chronology*)
  - Collections
- 
- <http://nasataxonomy.jpl.nasa.gov>



# Implementing Strategic Value Using Semantic Frameworks



Facets	Strategic Value
Access Requirements	Sensitivity and access control
Audiences	Who is the content intended for
Business Purpose	Why the content was created
Competencies	What field or discipline is relevant
Content Types	The genre of the content
Industries	External partners & businesses
Instruments	Flight payloads that yield science
Locations	Sites where work occurs – on and off Earth
Missions/Projects	NASA's lines of business
Organizations	NASA organizations
Subject Categories	The topic of the content



# NASA Specific Metadata Fields



*NASA has data sets unique to its business*

<b>Definition</b>	<b>NASA Taxonomy Mapping</b>	<b>Values</b>
<b>NASA missions and projects</b>	<b>nasa:missionsProjects</b>	<b>Taxonomy: Missions and Projects</b>
<b>Business purpose</b>	<b>nasa:businessPurpose</b>	<b>Taxonomy: Business Purpose</b>
<b>Technical competencies</b>	<b>nasa:competencies</b>	<b>Taxonomy: Competencies</b>
<b>Standard industry categories</b>	<b>nasa:industries</b>	<b>Taxonomy: Industries</b> <i>From NAICS standard</i>



# NASA Taxonomy Website and Resource



**Link to Metadata Specification**

**Link to XML DTDs and Schema**



NATIONAL AERONAUTICS AND SPACE ADMINISTRATION

+ Visit NASA.gov  
+ Contact the NASA Curator

- TAXONOMY TOP LEVEL FACETS

+ FAQs

+ NASA METADATA

+ NASA TAXONOMY XML

+ NASA XML PROJECT

## NASA Taxonomy - Top Level Facets

Tips on using the NASA Taxonomy

- [Access Security Requirements](#)
- [Audiences](#)
- [Business Purpose](#)
- [Competencies](#)
- [Content Types](#)
- [Industries](#)
- [Instruments](#)
- [Locations](#)
- [Missions and Projects](#)
- [Organizations](#)
- [Subject Categories](#)

**Background and training materials**

### What is the NASA taxonomy?

The NASA taxonomy is a controlled vocabulary that is designed to populate the [NASA metadata core specification](#).

It is also a means of tagging NASA content so that it can be used and reused in many different contexts.

**Links to Controlled Vocabularies**



+ Freedom of Information Act  
+ The President's Management Agenda  
+ FY2002 Agency Performance Report



Contact the NASA Curator  
NASA Official: Jayne Dutra  
Last Updated: May 25, 2004

<http://nasataxonomy.jpl.nasa.gov>

**JPL Knowledge Management**



# Built Demonstration of Taxonomy Value in Search and Navigation



NATIONAL AERONAUTICS  
AND SPACE ADMINISTRATION

Search



219958 items

### by Organization

- [NASA Affiliated Institutions](#) 1378
- [NASA Centers](#) 76545
- [NASA Contractors](#) 10108
- [NASA Enterprises](#) 815
- [NASA Headquarters](#) 4042
- [Other NASA Partners](#) 999

### by Competencies

- [Business](#) 386
- [Engineering](#) 393
- [Mission](#) 555
- [Scientific](#) 410
- [Technical](#) 218

### by Subject

- [Aeronautics](#) 26532
- [Astronautics](#) 31758
- [Chemistry and Materials](#) 17086
- [Engineering](#) 39631
- [Geosciences](#) 30770
- [Mathematical and Computer Sciences](#) 13286
- [Space Sciences](#) 22685
- [4 more](#)

### by Information Type

- [Catalogs and Databases](#) 32
- [Designs and Specifications](#) 62
- [Plans and Agendas](#) 158
- [Results and Analyses](#) 260
- [Reviews and Lessons Learned](#) 1819
- [Status Reports](#) 119
- [Technical Reports](#) 229
- [6 more](#)

### by Missions and Projects

- [Aerospace Technology](#) 60
- [Biological and Physical Research](#) 68
- [Data](#) 140
- [Earth Sciences](#) 1497
- [Human Exploration and Development...](#) 10680
- [Planetary Missions](#) 4819
- [Space Sciences](#) 9467

### by Collection

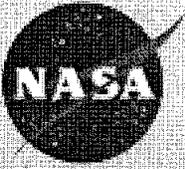
- [Lessons Learned](#) 1370
- [NTRS](#) 213900
- [SIRTF](#) 4054
- [Webb](#) 634

### by Date

- [1972](#) 8392
- [1973](#) 8512
- [1974](#) 7828
- [1975](#) 7704
- [1992](#) 8131
- [1993](#) 8519
- [1994](#) 7712
- [74 more](#)

Multiple collections  
viewed through a  
unified interface

...that provides common access framework across test collections



# Search & Browse Demo Site

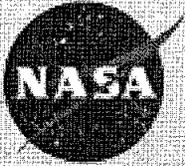


- <http://tb1.siderean.com:7880/test/test2query3.jsp>

Logon: nasa

Password: facets

- 
- Hosted by Siderean [www.siderean.com](http://www.siderean.com) with Seamark software



# Project Outcomes: NASA Taxonomy Benefits



## **... at the NASA Level**

- Metadata specification for all NASA content publishers
- RDF schema published 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)

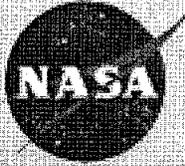


# Project Outcomes: NASA Taxonomy Benefits



## ... at the Federal Level

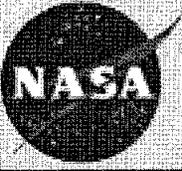
- NASA taxonomy development in accordance with e-Gov Act of 2002
- Integration with FEA at the BRM & DRM level
- Increased interoperability with other federal agencies through common data models and standards
- Better interoperability with industry partners for increased speed of mission development
- 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**
  - Public and Engineering portals
  - Search integration
    - Verity K2
    - Faceted search and navigation - Seamark
  - Content integration networks for real time delivery
- **NASA Enterprise Architecture Group standards development**
  - Agency Registries, RSS Syndication, Semantic Web components



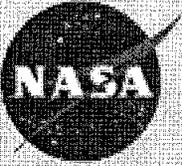
# Wrap Up and Discussion



**Thanks for your time!**

***Jayne Dutra and Joseph Busch***

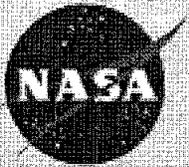
[Jayne.E.Dutra@jpl.nasa.gov](mailto:Jayne.E.Dutra@jpl.nasa.gov)  
[j.busch@worldnet.att.net](mailto:j.busch@worldnet.att.net)



# White Papers and Supporting Documentation



- NASA Taxonomy –  
<http://nasataxonomy.jpl.nasa.gov> 0204-0844
- White Paper: Taxonomy Development With NASA, Dutra and Busch, 2003
  - [https://pub-lib.jpl.nasa.gov/pub-lib/dscgi/ds.py/Get/File-22/NASA Taxonomy Overview-0203.ppt](https://pub-lib.jpl.nasa.gov/pub-lib/dscgi/ds.py/Get/File-22/NASA%20Taxonomy%20Overview-0203.ppt)
- White Paper: Implementing the NASA Taxonomy Through Service Oriented Architectures, Dutra and Xiao, 2/2004
  - [https://pub-lib.jpl.nasa.gov/pub-lib/dscgi/ds.py/Get/File-118/Content Integration Networks WP 02 04 04.doc](https://pub-lib.jpl.nasa.gov/pub-lib/dscgi/ds.py/Get/File-118/Content%20Integration%20Networks%20WP%2002%2004%2004.doc)



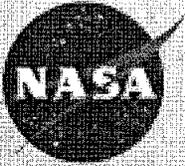
# Back Up Slides



# NASA Metadata Specification and Dublin Core Mapping - I



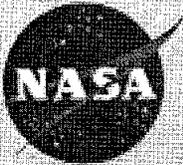
Dublin Core Elements	Definition	NASA Taxonomy Mapping	Values
Title	A name given to the resource	dc:title	String
Creator Creator Affiliation	Content maker Organization employing creator	dc:creator dc:creator.affiliation	String
Subject	Content topic	dc.subject	Taxonomy: Subjects
Publisher	Entity responsible for making the resource available	dc:publisher	Taxonomy: Organizations
Format	File format of the resource	dc:format	String – Internet Media Types
Type	Content Genre	dc:type	Taxonomy: Content Types



# NASA Metadata Specification and Dublin Core Mapping - II



Dublin Core Elements	Definition	NASA Taxonomy Mapping	Values
Description	An account of the content of the resource	dc:description	String
Identifier	Unique searchable identifier	dc.identifier	String
Date	Date the resource was published	dc:date	Format: yyyy-mm-dd
Access Rights	Access clearance info	dcterms:accessRights	Taxonomy: Access Security Reqmts
Language	Language of the resource	dc:language	String – RFC 3066
Rights	Copyright info	dc:rights	String
Coverage	Spatial locations	dc:coverage	Taxonomy: Locations
Audience	Entity for whom the resource is intended	dcterms:audience	Taxonomy: Audience



## Recommendations

- Create NASA Taxonomy Team
  - Taxonomy Editor as facilitator and manager
  - Steering Committee
- Institute a process for management of source controlled vocabularies
- Team meets on a regular basis to decide:
  - When vocabularies should be modified
  - How vocabularies should be modified
  - Tracking of earlier versions