



# Developing Integrated Taxonomies for a Tiered Information Architecture

---

Jayne Dutra, Enterprise Information  
Architecture

Online Information 2006 Conference

November 30, 2006



# Who Am I?

---



- Core Capabilities:
  - Search, web design, portal technology, web content management, knowledge management, information architecture, taxonomy development, metadata design, business analysis for IT applications
- Past Credits:
  - Inside JPL Portal (home page including enterprise search), Manager for NASA Taxonomy and Core Metadata Specification, Co-Chair of NASA Web Managers, JPL KM Process Owner, JPL Engineering Taxonomy and Metadata Core Spec Task Manager, team member: NASA Taxonomy for Problem Reporting Systems, faceted navigation pilot for flight projects
- Current:
  - Most of my work is for the CIO's Office on EA tasks, with an emphasis on information architecture problems

# Problem Statement: JPL Today



Parts  
Catalogues



Engineering  
Repositories



Electronic  
Libraries

*What did I call it?  
Where did I put it?  
How do I find it?*



Problem  
Reporting



E-Mail Archives



Financial  
Data



# The Solution: An Integrated Enterprise Information Architecture

---

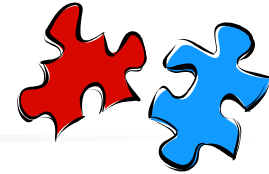
## A **Unified View** of our information space

- Integrated with business processes and best practices
- Based on institutional policies and common architecture
- Independent of any specific repository or technology
- Able to deal with harmonization of information according to a larger point of view
- Relevant and useful to our customers

*Expressed with a shareable metadata  
and its values (taxonomy)*



# Information Building Blocks



An integrated information architecture made up of several components:

- **Common Metadata Specification**
  - Core Metadata Specification for JPL Information Objects
- **Common language or controlled vocabularies**
  - By discipline, product, life cycle, process, etc.
  - NASA Taxonomy, JPL Taxonomy, Partner Taxonomies
- **Business Rules for data reconciliation**
  - You say “tomato”.....

# Nested Taxonomies

## Term criteria at the Enterprise level

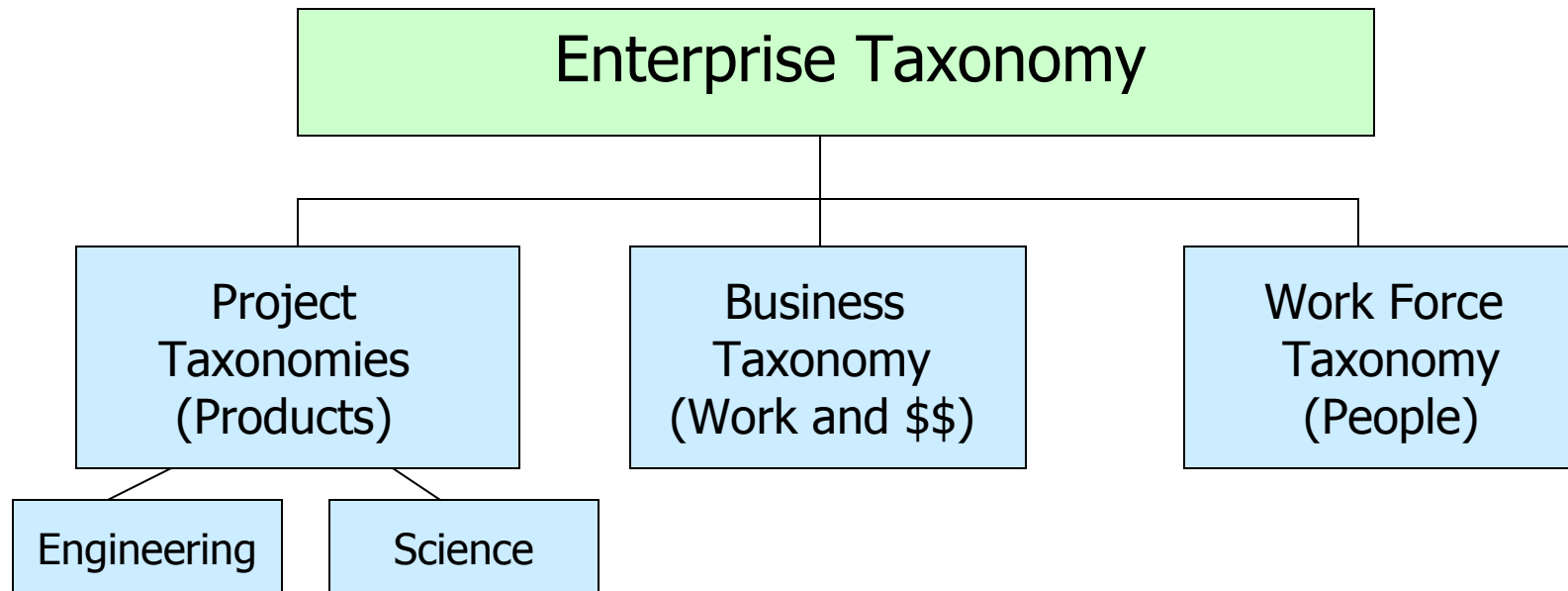
- Need for **broad** buckets of information that many different systems can map into
- Patterns tied to business process emerge over time
- Integrated with the NASA Taxonomy for interoperability





# Tiered Taxonomy Work

- Enterprise Information Architecture is tiered
- Work already exists in the Project Engineering and Science Domains
- Work is starting up in the Business Domain
- Work is waiting in the Work Force Domain





## Work Already Completed: Metadata Core Spec for the Project Domain

*Sponsored by the JPL Office of the CIO*

### Goals

- Create Lab wide information standards for engineering products
- Design and document a Lab wide information content model
- Improve information access and retrieval, including cross repository search
- Integrated with NASA metadata and taxonomy specs

### Deliverable

- JPL Core Metadata Specification for Project Documentation, v1.0
- Developed by a cross organizational team
- Delivered March, 2005





# JPL Enterprise Project Taxonomy Effort Goals

---

- Tightly integrated with JPL Project Metadata Core Specification
  - Provides vocabulary values for the appropriate metadata fields
  - Designed to be relevant for a domain of users (communities of practice)
- Provides a classification scheme for identifying content
  - A means for tagging content so it can be used and reused in different contexts
- Derived from JPL standard gold sources
  - Vocabulary owners are Subject Matter Experts in their domains
  - Discoverable with SOAs or Semantic Web technologies

# JPL Flight Project Engineering Taxonomy



- [Audiences](#)
- [Business Purpose](#)
- [Competencies](#)
- [Content Types](#)
- [Document Levels](#)
- [Document States](#)
- [Instruments](#)
- [Phase](#)
- [Locations](#)
- [Missions and Projects](#)
- [Organizations](#)
- [Repositories](#)
- [Roles](#)
- [Sensitivity Levels](#)
- [Subject Categories](#)
- [System/Subsystem](#)
- [Work Breakdown Structures](#)

Tips on using the JPL Taxonomy  
**What is the JPL taxonomy?**

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

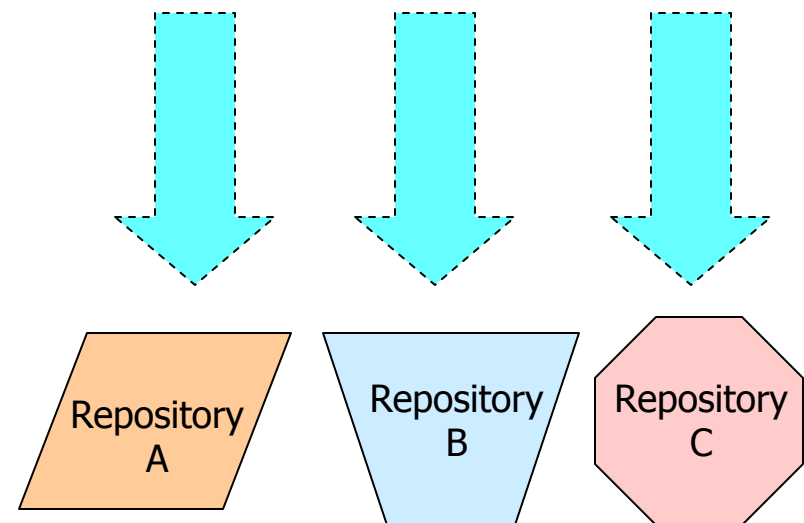
- Taxonomy values, equivalencies
- Metadata Core Specification
- RDF files for easy re-use
- FAQ files to assist developers



# Adding Richness Over Time

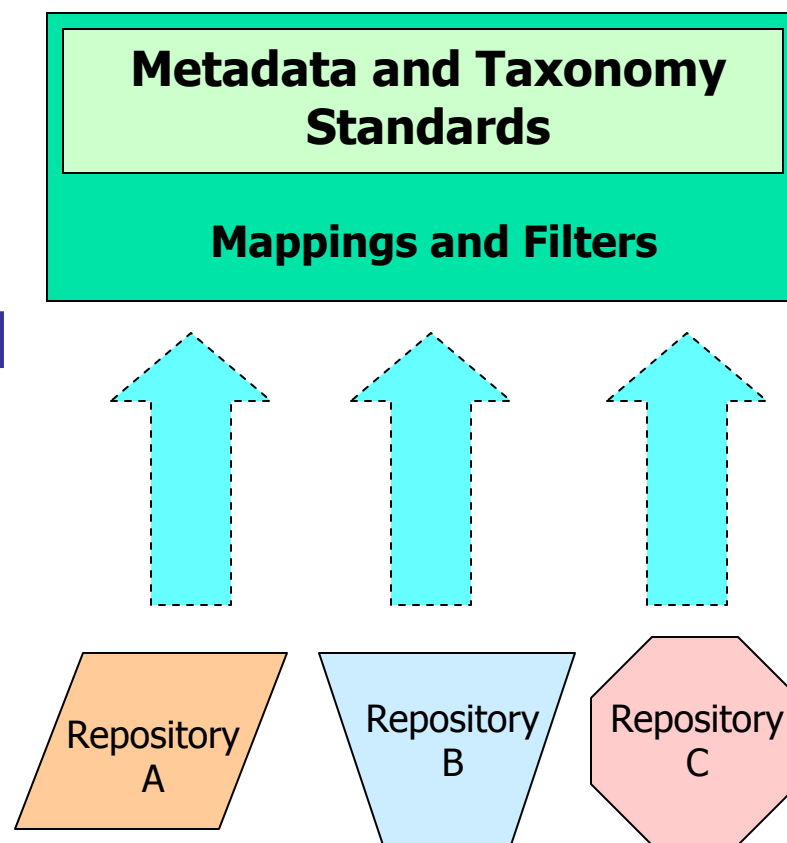
- Metadata Core Specification was a top down effort

**Metadata and Taxonomy Standards**



# Adding Richness Over Time

- Metadata Core Specification was a top down effort
- For the Phoenix Faceted Navigation Pilot, we used a bottom up approach





# Validation: Faceted Navigation Prototype for Engineers

---

## Semantic Search/faceted navigation for Flight Projects

- JPL Engineering Domain
- Pilot funded by JPL Chief Engineer
- For flight teams on Phoenix and CloudSat missions
- Completed in 6 weeks
- Goal: Provide cross repository search from a single interface based on
  - Relationships of information objects
  - Life cycles – mission and content
  - Task analysis for specific roles
  - Engineering processes

**14,981 items**

by Collection	by Date	by Controlled Item	by Content Type
Docushare Heritage Collection 1633	1980s 40	No 453	Data 132
PDMS:Sherpa 167	1990s 5085	Unspecified 1058	Design or Specification 90
Problem Failure Reporting System 10658	2000s 6882	Yes 13470	Meeting File 1558
Problem Reporting System 34			Plan 155
TeamCenter Community 2489			Procurement Record 544
			Quality Control Record 10697
			Report 511
			Requirement 231

by Organization	by Mission and Project	by Instrument	by System/Subsystem
2x - Business Operations 1	Aerospace Technology 2	Accelerometers 118	Flight System 7594
3x - Engineering and Science 837	Planetary Missions 14981	Antennae 217	Ground Data System 778
4x - Solar System Exploration 88	Space Sciences 15	Cameras 730	Ground Support Equipment 1392
5x - Office of Safety and Mission Success 98	Unspecified 2	Laser Altimeters 31	Integration and Test Equipment 22
6x - Mars Exploration 13947		Meteorology Instruments 57	Launch System 32
8x - Earth Science and Technology 49		Radar 45	Non-Flight System 59
		Robotics 50	Payload 3483
		Spectrometers 149	Science Data System
		Telescopes 352	

by Competency	by Subject	by Mission Phase	by Document State
Business 857	Astronautics 1116	Phase A 2	Archived 1
Engineering 10910	Chemistry and Materials 12	Phase B 37	Draft 10
Mission 2196	Engineering 10887	Phase C 71	In Preparation 47
Scientific 354	General 49	Phase D 105	Released 1075
Technical 2837	Geosciences 4	Phase E 7	Uncategorized States 13848
Unspecified 4206	Mathematical and Computer Sciences 660	Pre-Phase A 78	
	Physics 83	Unspecified 14656	
	Social Sciences 204		
	Space Sciences 6		
	Unspecified 3889		
	<a href="#">2 more</a>		

by Document Level	by WBS	by Sensitivity Level
Level 1 - Program 45	Flight System 1636	Discreet 1
Level 2 - Project 2	Launch System 214	Export Controlled 1121
Level 3 - System/Segment 33	Mission Design 182	None/Unlimited Distribution 52
Level 4 - Subsystem/Element 12117	Mission Operations System 580	Not Assessed 543
Unspecified 2793	Payload System 67	Proprietary 10860
	Project Management 138	

# Phoenix Semantic Search Pilot

- 15,000 items
- 5 repositories
- Tagged with JPL Taxonomy facets
- Unified search from one interface
- Faceted navigation for intuitive info discovery
- Sponsored by the JPL Chief Engineer



## Some Further Use Cases: Faceted Navigation Using Relationships

---

- Project Manager:
  - “I’d like to check all documents needed to complete my Certification for Launch to see what state they’re in, no matter where they are.”
- Cognizant Engineer:
  - “I’d like to see all problem failure reports on a sub-system I designed and flew 5 years ago so I can incorporate the lessons learned into my current mission.”
- Project Information Management Engineer:
  - “I’d like to see the status of all Phase B documents that I need to prep for an upcoming gate review so I know we’re ready.”



# Work In the Financial Arena

---

## JPL Business Domain

- Term clarification for vocabulary used in financial reporting, work force planning, estimating and proposals

### **Sample Vocabulary Problem Term: Project**

Oracle project number – financial costing system

IBF/NSM project number – NASA reference

PRL project number – JPL flight project name

*How can we efficiently search for information about a "project"?*





# IA for the Business Domain

## Describing Work for Forecasting and Strategic Planning

Strategic Question	Metadata Elements
<i>What work are we doing?</i> <i>What type of business is it?</i>	<ul style="list-style-type: none"><li>■ Work Type</li><li>■ Business Type</li><li>■ Work ID/ Project Name</li></ul>
<i>How did we acquire it?</i>	<ul style="list-style-type: none"><li>■ Competition Type</li></ul>
<i>Who is funding the work?</i> <i>Who are we delivering it to?</i>	<ul style="list-style-type: none"><li>■ Role (inc customer, sponsor, partner, line roles and project roles)</li></ul>
<i>Who is doing the work?</i> <i>What capabilities does it take?</i>	<ul style="list-style-type: none"><li>■ Organization</li><li>■ Competency, Capabilities</li><li>■ Work description</li></ul>
<i>What phase is it in?</i> <i>Where is it taking place?</i>	<ul style="list-style-type: none"><li>■ Work State</li><li>■ Work Maturation Level</li><li>■ Implementation Mode</li></ul>



## Work Already Completed: Metadata Working Group for Business Domain

---

### Goals

- Create Lab wide information standards for business/work products
- Design a Lab wide business information content model
- Improve information access and retrieval, including cross repository search and reporting
- Identify “Gold Source” owners of data

### Deliverable

- JPL Core Metadata Specification for Business, alpha version
- Developed by a cross organizational team
- Delivered May, 2006
- Next steps: validation and assigned owners

# JPL Business Taxonomy - Alpha



## JPL Business Taxonomy - Top Level Facets

- [Business Purpose](#)
- [Business Type](#)
- [Commitment State](#)
- [Competencies](#)
- [Content Types](#)
- [Funding Mechanism](#)
- [Implementation Mode](#)
- [Missions/Projects](#)
- [Organizations](#)
- [Repositories/Applications](#)
- [Roles](#)
- [Work Acquisition Type](#)
- [Work Breakdown Structures](#)
- [Work Maturation Level](#)
- [Work Type](#)

More work continuing with Integrated Business team for validation and expansion



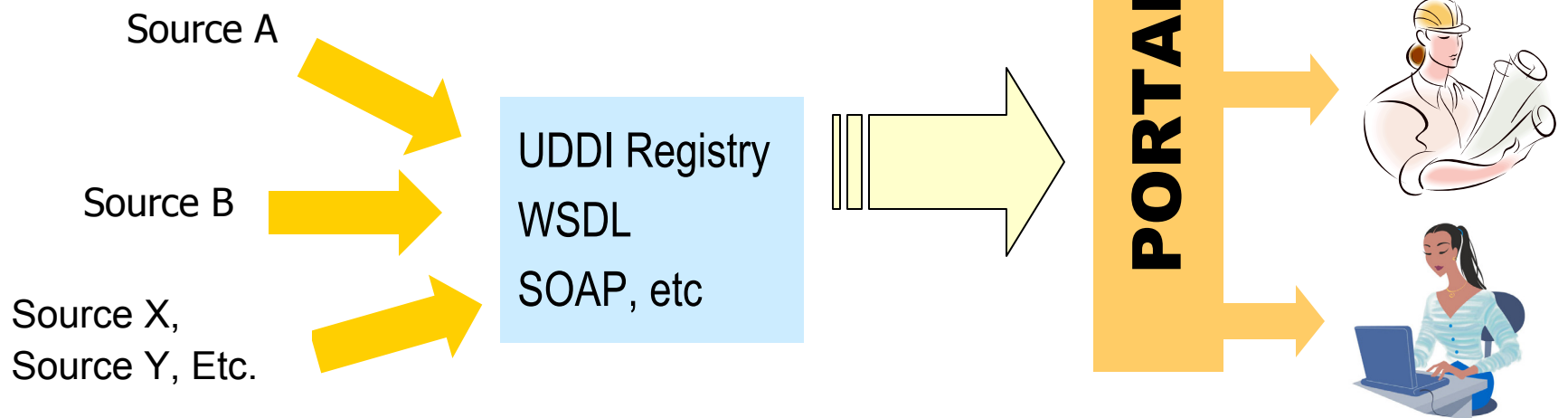
# Work Force Metadata

---

- Collect and document attributes about people
- Some use cases
  - Access management – who are you and what can you see?
  - Targeted content delivery – what content helps you get your work done?
  - Work force planning – what skill sets do you have that we can apply to work?
  - IT Services – How can we provision you with the proper IT services?

# Using People Metadata for Targeted Content Delivery in Portals and Dashboards

- Make content available to delivery mechanisms using Service Oriented Architectures
- Data streams presented as services and available for consumption by workers in portals, dashboards and other devices





# But What Goes Where?

---

## Attributes That Describe People

- An Engineer
- Specialty is Electrical Engineering
- Works on propulsion systems
- Worked on projects X, Y, Z
- Currently working on A
- As a Cog E
- On propulsion subsystem
- Project is in Phase C
- Has published papers on propulsion systems

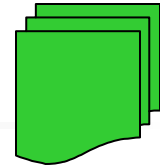
## Corresponding Taxonomy Facet

- Competency
- Capability
- Topic or Subject Matter
- Past Assignments
- Current Assignment
- Role
- System/Subsystem
- Project Phase
- Topic or Subject Matter

*We can codify and track certain attributes for re-use*

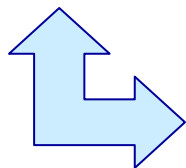


# Matching Attributes for People to Attributes for Content



## Attributes About People

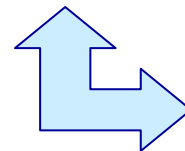
- Competency/Discipline
- Subject Matter
- Past Task Assignment - Role
- Current Task Assignment - Role
- Subsystem/WBS
- Task Phase
- Associations to objects as Author or Reader/Subscriber



**People Metadata**

## Attributes About Info Objects

- Objects related to a Competency
- Interest in Subject Matter Areas
- Objects associated with Role
- Information on a Subsystem/WBS
- Objects associated with a project phase
- Information on project products
- Information on technologies

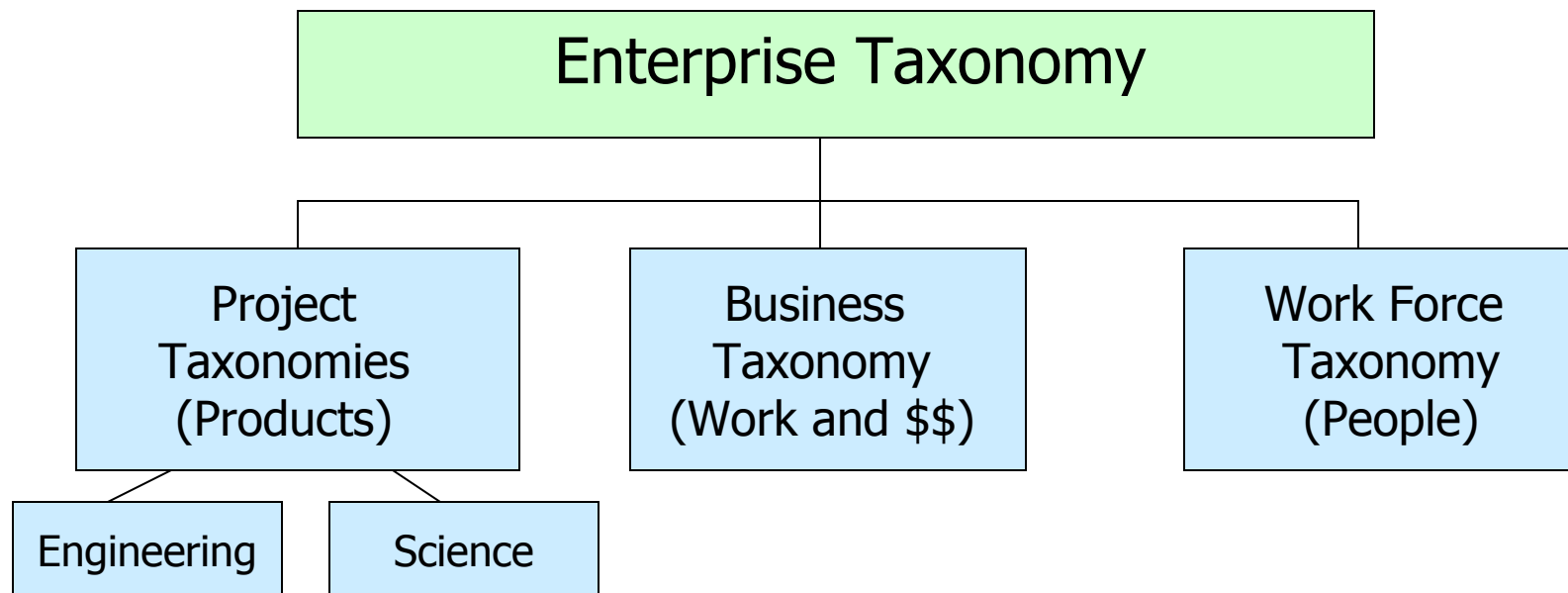


**Project Metadata**



# Tiered Taxonomy Work

Enterprise Information Architecture is tiered







# How An Enterprise Taxonomy is Born: The Cream Rises to the Top

---

## Enterprise Taxonomy

- Organization
- Mission/Projects

### Project Taxonomy

- Content Type
- Mission Phase
- Organization
- Document State
- Missions/Projects
- System/Subsystem

### Business Taxonomy

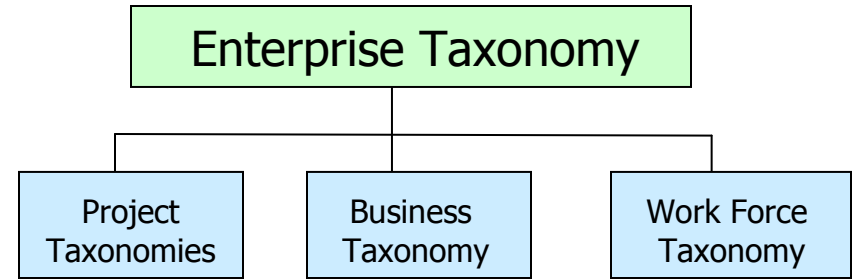
- Work Type
- Funding Mechanism
- Business Type
- Organization
- Acquisition Type
- Missions/Projects

### Work Force Taxonomy

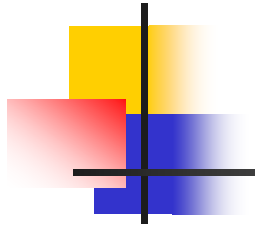
- Organization
- Competency/Skills
- Capability
- Missions/Projects
- Role



# Summing Up



- Iterative process in a tiered model
  - Direct connection between IA and use cases, business questions
  - Metadata specification defined
  - Taxonomy definition: Values expressed with controlled terms where appropriate
    - Domain specific taxonomies add business value
    - Transition points enable interoperability
    - Look for patterns of usage tied to business process
  - Integrate across domains and upward to a meta level
  - Implement and validate
  - Long term stewardship and configuration management



# Questions, Discussion

---

Thanks!

*Jayne.E.Dutra@jpl.nasa.gov*



# Information and IT Services Delivery by Mission (and Content) Life Cycle Based on Role

---

## Leverage Mission Life Cycle

- What roles are active at each stage
- What IT services are needed at each phase of a mission
- What capabilities are most in use at each phase of a mission

## Example:

Proposal Phase → System Engineering Capability

Activities → Trade Studies, Model Building, Requirements Formulation

IT Needs → Document storage, simulation and visualization, traceability

IT Apps → Docushare, MBED, Cradle, CORE

## Other Proposal Phase roles and activity types:

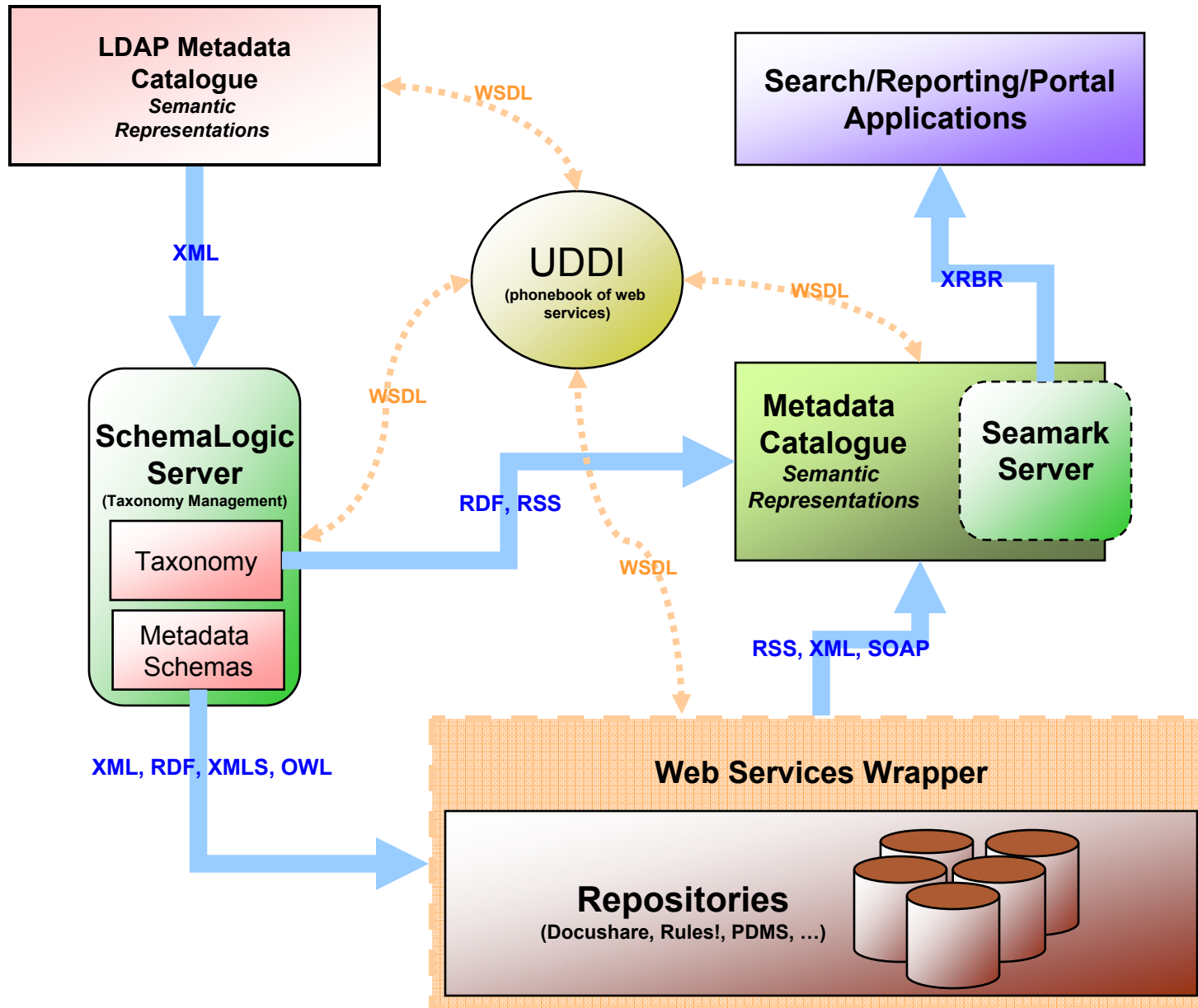
Project Managers – do planning

Principal Investigators – formulate science definition goals

Project Resource Administrators – create budget and schedules

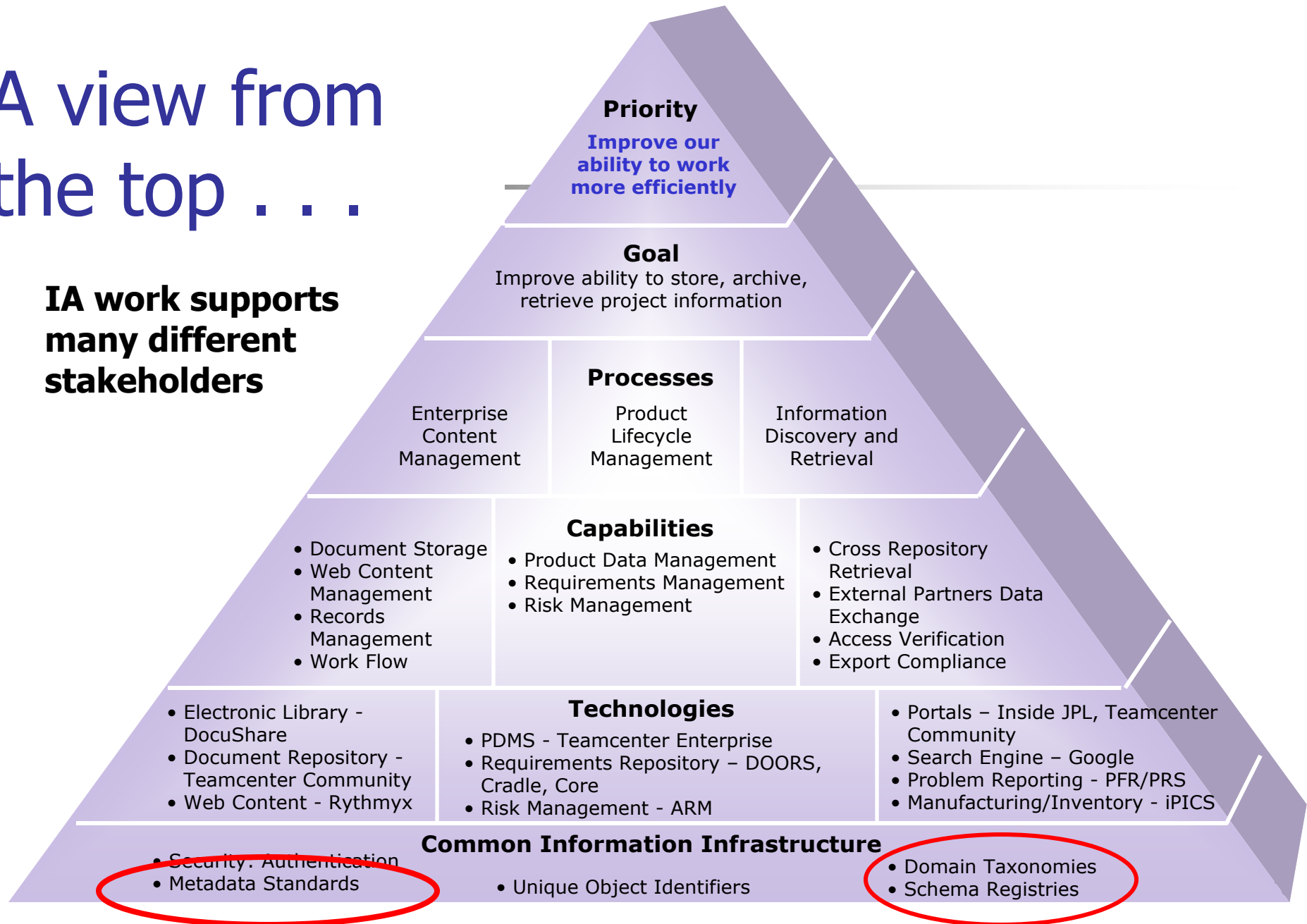
*This is a very rich area for us to mine!*

# Notional IA Architecture – High Level

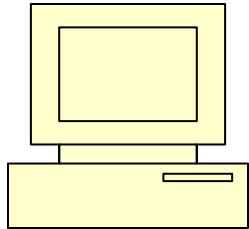


# A view from the top . . .

**IA work supports many different stakeholders**



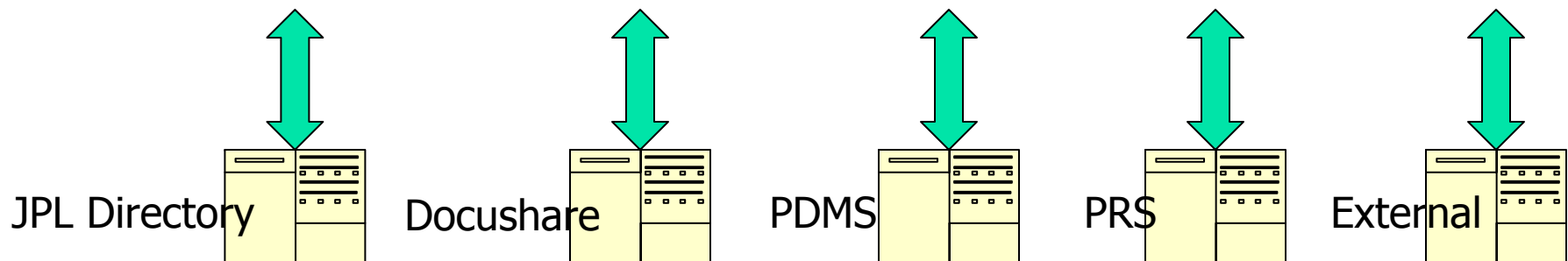
# Unified Search for JPL Projects: Goal



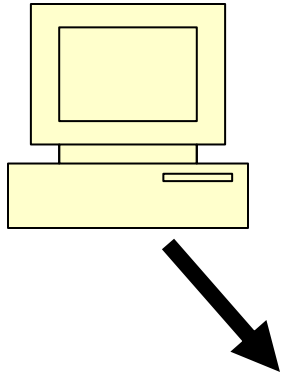
Resulting in **integrated** responses like this:

**Requires Semantic Web technology and Service Oriented Architectures To automate**

- Project System Requirements v 2.0 (DOORS)
- Hardware Drawing - Parts v 1.0 – PDMS
- Hardware Drawing – Assemblies v 3.0 - PDMS
- ECR – PDMS
- ECR waivers – Team Center Community
- Problem Failure Report - PRS
- Associated Documentation - Docushare



# Unified Search for JPL Projects: Goal

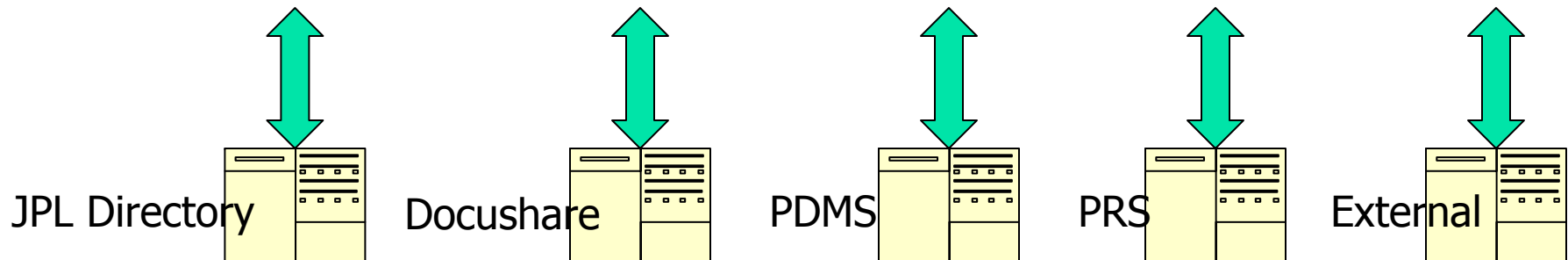


Users would like to be able to find **related sets** of data. For example: “I want to see all of the released documents for my project that are required for system certification for flight”

Combinations of metadata...

- Project = Project XYZ
- Document State = Preliminary or Official
- Content Class = Controlled
- Business Purpose = Product Data
- Content Type = Appropriate Values
- Document Status = Lifecycle values

Query can be applied across multiple repositories...







## Taxonomy Packages Equivalencies for Vocabularies That Change Over Time

Acronyms and Synonyms can be easily defined using Equivalent Terms in RDF (a new language from the semantic web)

- MER A= Spirit, MER B = Opportunity
- MER A and MER B are types of Rovers
- Rovers fall under the subject heading of Robotics/Cybernetics

Built into the back end  
and transparent to users

*Users don't have to have special contextual knowledge to find information items*

# JPL Taxonomy

An Information Product of the JPL Enterprise Architecture Group



Taxonomy Facets

Taxonomy FAQs

Taxonomy XML

**Metadata Specification**

NASA Taxonomy

**JPL Metadata Core Specification v 1.0**

Core Metadata  
Specification  
Online

This specification was developed with a focus on JPL flight project documentation. [Excel Spreadsheet](#)

Field Name	Element Name and Namespace	Definition	Data Type or Source	Comment
<b>The following fields are mandatory for JPL flight project documentation.</b>				
<b>Author</b>	dc:creator	Preparer or responsible individual	String	
<b>Date</b>	dc:date	Date the resource was 'published'.	Cover Date, use format yyyy-mm-dd	For browsing by date, a simple publication date is all that is needed.
<b>Class</b>	jpl:class	Whether or not the content is under configuration management.	Yes/No	If material is controlled, other conditional tags may apply
<b>Content Type</b>	dc:type	The nature or genre of the content of the resource.	Values come from JPL Taxonomy: Content Types	
<b>File Name</b>	jpl:file	Label given to a file that contains the content	String	

*Expressed with  
metadata and  
its values  
(taxonomy)*

<http://jpltaxonomy/metadata.htm>

# Taxonomy Packages Equivalencies for Vocabularies That Change Over Time

```
- <pcv:Descriptor rdf:about="subj:63">
  <pcv:label>Cybernetics</pcv:label>
  <pcv:synonym>artificial intelligence</pcv:synonym>
  <pcv:synonym>control theory</pcv:synonym>
  <pcv:synonym>expert systems</pcv:synonym>
  <pcv:synonym>feedback theory</pcv:synonym>
  <pcv:synonym>robotics</pcv:synonym>
  <pcv:broaderTerm rdf:resource="subj:59" />
  <pcv:relatedTerm rdf:resource="subj:54" />
  <pcv:definition>Includes feedback and control theory, artificial intelligence, robotics and expert systems. For related information see also Man/System Technology and Life Support.</pcv:definition>
  <pcv:code>67</pcv:code>
</pcv:Descriptor>
```

Acronyms can be easily defined using Equivalent Terms or Synonyms in RDF

MER A= Spirit, MER B = Opportunity,  
MER A and MER B are types of Rovers  
and fall under the subject heading of  
Robotics/Cybernetics

Built into the back end  
and transparent to users



# Workforce Planning and Forecasting

---

Answering important business questions for forecasting:

- What work are we doing? What type of business is it?
- How did we acquire it?
- What skill sets does it take to complete?
- Who is funding it? Who is doing it? Who are we delivering it to?
- How much does it cost? How many hours to completion?
- What phase is it in? Where are we doing it?

We want to improve how accurately and how quickly we can answer these questions for management