Content Integration Networks: Turning Search Upside Down

KMWorld/Intranets 2006

San Jose, CA

November 1, 2006

Jayne Dutra

Jet Propulsion Laboratory

California Institute of Technology

JPL

Who Am I?

- Core Capabilities:
  - Search, web design, portal technology, web content management, usability, knowledge management, enterprise information architecture, taxonomy development, Semantic Web
- Past Credits:
  - Inside JPL Portal (internal home page), Daily Planet Web Site (employee news), many outreach sites for JPL flight projects, Co-Chair of NASA Web Managers, Manager for NASA Taxonomy and Core Metadata Specification, JPL Engineering Taxonomy and Metadata Core Specification Task Manager, team member on NASA Problem Failure Reporting Taxonomy development for CEV data mining, Lead on JPL Project Semantic Search Pilot
- Current:
  - JPL Knowledge Management Process Owner, mainly I work for the CIO's Office on Enterprise Architecture tasks, with an emphasis on information architecture problems

Presentation Agenda

- Overview and Problem Statement
- Vision and Technologies (high level)
  - Semantic Web, Metadata, SOA's
- Strategy
  - Designing Semantic Relationships
  - Targeting Content Using Taxonomy Terms
- Knowledge Management Benefits
  - Connecting People and Data
- Summary and Challenge

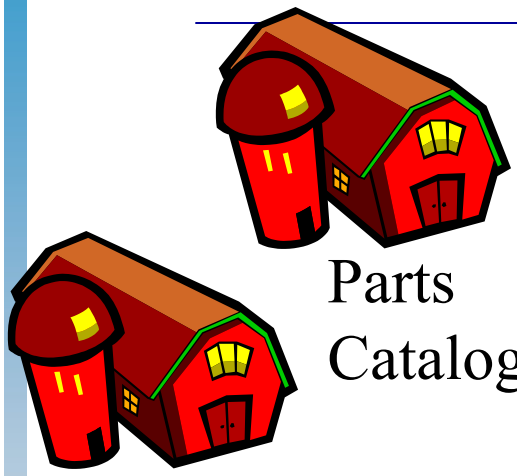
Overview

- More data
- More sources and repositories
- More silos (how many passwords do you have?)
- More fragmented information space

Result: Search getting harder than ever!

JPL

JPL Today



Parts Catalogues

Electronic Libraries

Engineering Repositories

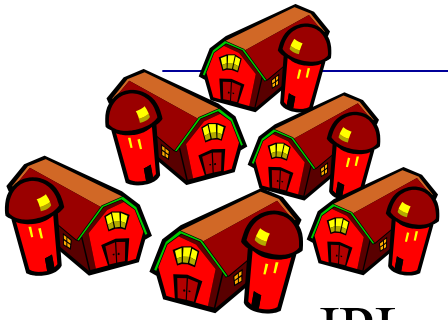
What did I call it?
Where did I put it?
How do I find it?

E-Mail Archives

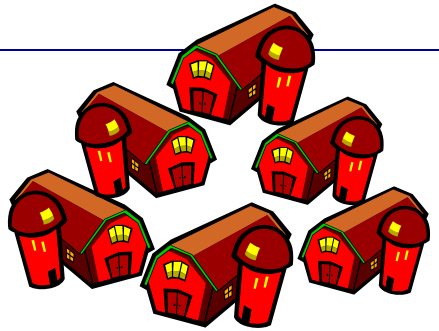
Problem Reporting

Financial Data

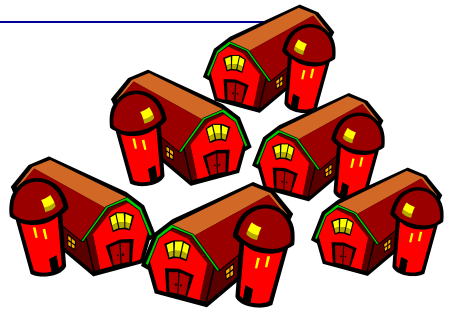
NASA Today



JPL



Kennedy



Johnson

Langley

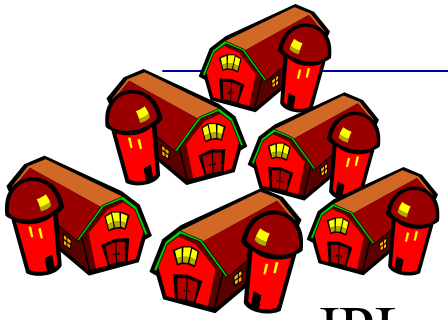
Ames

Goddard

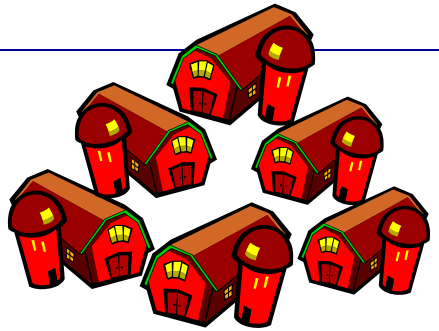
Marshall

JPL

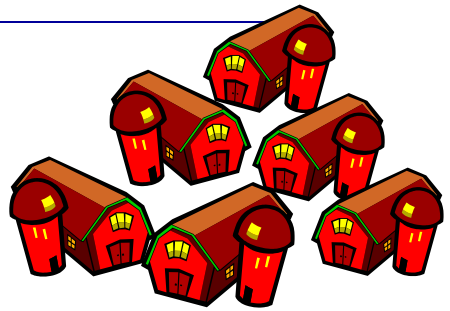
NASA and Other Agencies



JPL



Kennedy



Johnson

Langley

Ames

Goddard

USA

Marshall

Lockheed

ESA

JPL

Knowledge Retrieval

Chances of Finding Needed Information in a Timely Fashion

Knowledge Retrieval

0%

A Different Paradigm

But
what if content
came to you?

Turning Search Upside Down

- Just in time information delivery based on
  - Engineering Lifecycle
  - Task Analysis
  - Associations and relationships
  - Agents and electronic subscriptions
    - **Persistent queries and syndicated content**

New Technology

From Tim Berners-Lee and the W3C

"The Semantic Web is a vision: the idea of having data on the web defined and linked in a way that it can be used by machines not just for display purposes, but for automation, integration and reuse of data across various applications."

http://www.w3.org/2001/sw/

So, What is the Semantic Web?

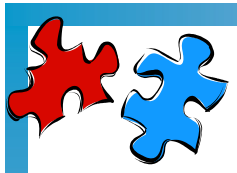
- Today's Web is made for **people** to read and understand
- Tomorrow's Web will be made for computers to read and understand
  - Systems will be able to perform transactions across applications without human help
  - Leverages the vast amount of data accessible on the Web for machine processing
  - Integration of data sets that are currently unlinked using the Web

JPL

So, What is the Semantic Web?

- Today's Web is made for people to read and understand
- Tomorrow's Web will be made for **computers** to read and understand
  - Systems will be able to perform transactions across applications without human help
  - Leverages the vast amount of data accessible on the Web for machine processing
  - Integration of data sets that are currently unlinked using the Web

JPL



Information Building Blocks

An integrated information architecture made up of several components:

- **Common Metadata Specification**
  - Core Metadata Specification for JPL Project Documentation
- **Common language or controlled vocabularies**
  - By discipline, product, and process, etc. - taxonomies
  - Knowledge representations including relationships
  - Intersecting ontology hubs
- **Business Rules for data reconciliation**
  - You say “tomato”……
- *Use new technologies developed for the Semantic Web to enable enhanced capability*

JPL

How Does It Work?

- Focused on encoding metadata about Web resources into Web pages
  - Start with a basic taxonomy of terms and agreed upon definitions
  - Add relationships and associations, ie *ontologies*
- Based on knowledge representation languages
  - RDF, RDFS (Resource Description Framework and Schema)
  - SPARQL (RDF Query Language)
  - OWL (Web Ontology Language)

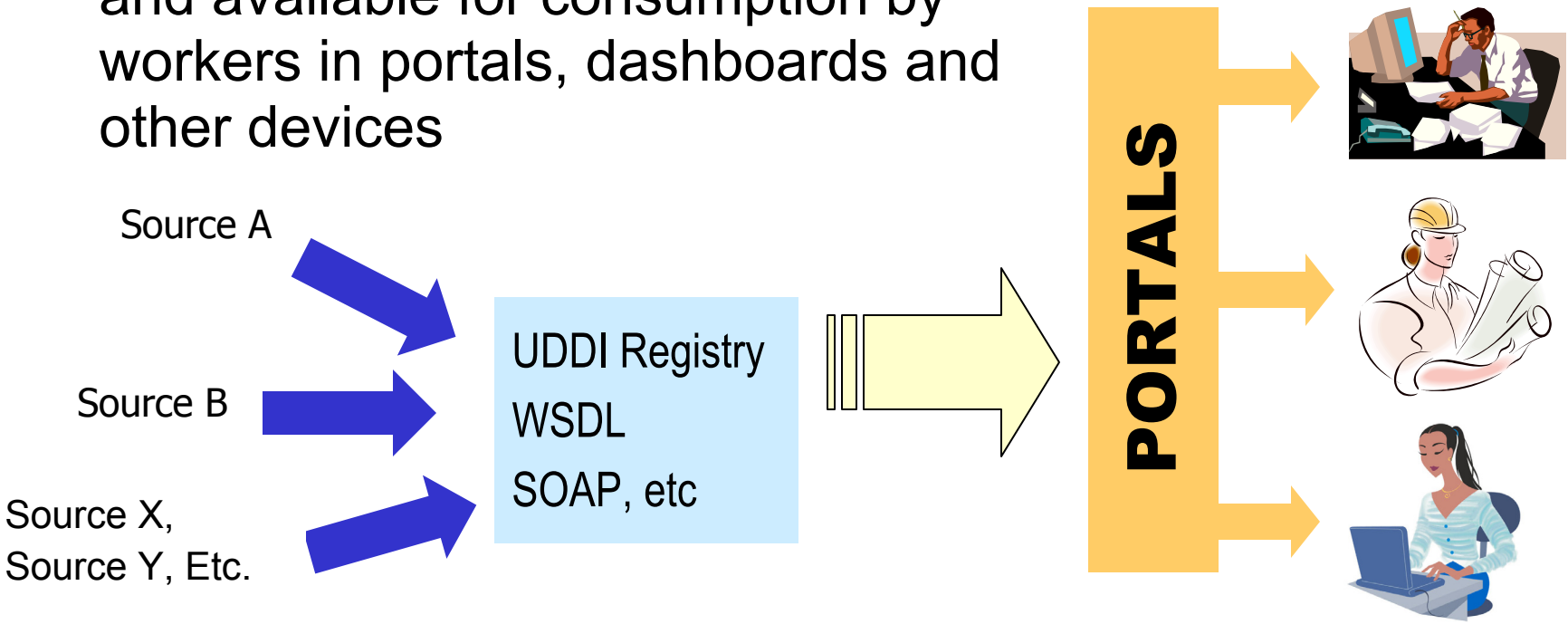
JPL

Then What?

- 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

Then What?

- 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

- Discipline

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

- Discipline
- Competency

But What Goes Where?

Attributes That Describe People	Corresponding Taxonomy Facet
An Engineer	Discipline
Specialty is Electrical Engineering	Competency
Works on propulsion systems	Topic or Subject Matter
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	

But What Goes Where?

Attributes That Describe People	Corresponding Taxonomy Facet
An Engineer	Discipline
Specialty is Electrical Engineering	Competency
Works on propulsion systems	Topic or Subject Matter
Worked on projects X, Y, Z	Past Assignments
Currently working on A	
As a Cog E	
On propulsion subsystem	
Project is in Phase C	
Has published papers on propulsion systems	

But What Goes Where?

Attributes That Describe People	Corresponding Taxonomy Facet
An Engineer	Discipline
Specialty is Electrical Engineering	Competency
Works on propulsion systems	Topic or Subject Matter
Worked on projects X, Y, Z	Past Assignments
Currently working on A	Current Assignment
As a Cog E	
On propulsion subsystem	
Project is in Phase C	
Has published papers on propulsion systems	

But What Goes Where?

Attributes That Describe People	Corresponding Taxonomy Facet
An Engineer	Discipline
Specialty is Electrical Engineering	Competency
Works on propulsion systems	Topic or Subject Matter
Worked on projects X, Y, Z	Past Assignments
Currently working on A	Current Assignment
As a Cog E	Role
On propulsion subsystem	
Project is in Phase C	
Has published papers on propulsion systems	

But What Goes Where?

Attributes That Describe People	Corresponding Taxonomy Facet
An Engineer	Discipline
Specialty is Electrical Engineering	Competency
Works on propulsion systems	Topic or Subject Matter
Worked on projects X, Y, Z	Past Assignments
Currently working on A	Current Assignment
As a Cog E	Role
On propulsion subsystem	System/Subsystem
Project is in Phase C	
Has published papers on propulsion systems	

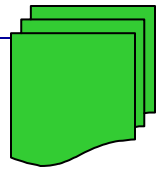
But What Goes Where?

Attributes That Describe People	Corresponding Taxonomy Facet
An Engineer	Discipline
Specialty is Electrical Engineering	Competency
Works on propulsion systems	Topic or Subject Matter
Worked on projects X, Y, Z	Past Assignments
Currently working on A	Current Assignment
As a Cog E	Role
On propulsion subsystem	System/Subsystem
Project is in Phase C	Project Phase
Has published papers on propulsion systems	

But What Goes Where?

Attributes That Describe People	Corresponding Taxonomy Facet
An Engineer	Discipline
Specialty is Electrical Engineering	Competency
Works on propulsion systems	Topic or Subject Matter
Worked on projects X, Y, Z	Past Assignments
Currently working on A	Current Assignment
As a Cog E	Role
On propulsion subsystem	System/Subsystem
Project is in Phase C	Project Phase
Has published papers on propulsion systems	Topic or Subject Matter

Matching Attributes for People to Attributes for Content



Attributes About People

- Competency/Subject Matter
- Discipline
- Past Task Assignment
- Current Task assignment - Role
- Subsystem
- Task Phase
- Associations to objects as Author

People Ontology (POPS)

Attributes About Info Objects

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

Engineering Taxonomy

Matching Content Attributes

Attributes About Info Objects

- **Products related to a Competency**
- **Objects related to Topic Areas**
- **Objects associated with a Role**
- **Information on a Subsystem**
- **Objects associated with a project phase**
- **Information on project products**
- **Information on technologies**

Attributes About Info Objects

- **LLIS objects about elec engineering**

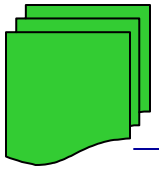
Matching Content Attributes

Attributes About Info Objects

- Products related to a Competency
- Objects related to Topic Areas
- Objects associated with a Role
- Information on a Subsystem
- Objects associated with a project phase
- Information on project products
- Information on technologies

Attributes About Info Objects

- LLIS objects about elec engineering
- LLIS objects about propulsion design



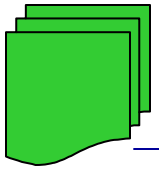
Matching Content Attributes

Attributes About Info Objects

- Products related to a Competency
- Objects related to Topic Areas
- Objects associated with a Role
- Information on a Subsystem
- Objects associated with a project phase
- Information on project products
- Information on technologies

Attributes About Info Objects

- LLIS objects about elec engineering
- LLIS objects about propulsion design
- Designs related to team activity



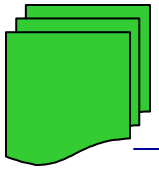
Matching Content Attributes

Attributes About Info Objects

- Products related to a Competency
- Objects related to Topic Areas
- Objects associated with a Role
- Information on a Subsystem
- Objects associated with a project phase
- Information on project products
- Information on technologies

Attributes About Info Objects

- LLIS objects about elec engineering
- LLIS objects about propulsion design
- Designs related to team activity
- Anomalies involving propulsion systems



Matching Content Attributes

Attributes About Info Objects

- Products related to a Competency
- Objects related to Topic Areas
- Objects associated with a Role
- Information on a Subsystem
- Objects associated with a project phase
- Information on project products
- Information on technologies

Attributes About Info Objects

- LLIS objects about elec engineering
- LLIS objects about propulsion design
- Designs related to team activity
- Anomalies involving propulsion systems
- ECRs related to propulsion system

Matching Content Attributes

Attributes About Info Objects

- Products related to a Competency
- Objects related to Topic Areas
- Objects associated with a Role
- Information on a Subsystem
- Objects associated with a project phase
- Information on project products
- Information on technologies

Attributes About Info Objects

- LLIS objects about elec engineering
- LLIS objects about propulsion design
- Designs related to team activity
- Anomalies involving propulsion systems
- ECRs related to propulsion system
- GIDEP alerts about products related to propulsion

Matching Content Attributes

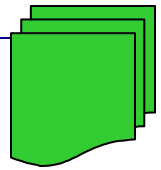
Attributes About Info Objects

- Products related to a Competency
- Objects related to Topic Areas
- Objects associated with a Role
- Information on a Subsystem
- Objects associated with a project phase
- Information on project products
- Information on technologies

Attributes About Info Objects

- LLIS objects about elec engineering
- LLIS objects about propulsion design
- Designs related to team activity
- Anomalies involving propulsion systems
- ECRs related to propulsion system
- GIDEP alerts about products related to propulsion
- Published papers on relevant subjects and technologies

Associations Between Attribute Sets



Attributes About People

- An Engineer
- Discipline is Electrical Engineering
- 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

Attributes About Info Objects

- LLIS objects about propulsion systems
- Published papers on relevant subjects and technologies
- ECRs related to subsystem
- Information about projects
- Anomalies involving propulsion systems
- GIDEP alerts about products related to propulsion

Web of Knowledge

The KM-ness of All This

Making Connections Across Data

People ⟷ People

People ⟷ Content Objects

People ⟷ Process (Engineering Life Cycle)

KM End Products

People ⟷ People

- Social Networks
- Experts Locators
- Team Collaboration Tools - Portals
- Portals for Communities of Practice

Use Cases

Project Element Manager:

"I wonder who else has done this type of work before?"

"I want to hire someone at a different Center to be my team Cog E since the work is being done there. Who has the right skills and experience?"

KM End Products

People ⟷ Content

- Effective Knowledge Discovery
- Robust Knowledge Base
- Targeted Delivery and Transparent Search

Use Cases

Manager:

"I'd like to see all documents needed to complete my Certification for Launch and what state they're in, no matter where they are."

Scientist:

"I'd like to see what types of data were returned on earlier missions using a particular instrument to help with the Science Definition Goals of my new project."

KM End Products

People ⟷ Process

- Effective Knowledge Discovery
- Smart Work Flows
- Just in Time Content Delivery
- No Search Search

Use Cases

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."

"I'd like to see all engineering rationale documents (Technical IOMs) that pertain to a particular trade study outcome on a certain type of mission or subsystem design."

Content 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!

Summary

- Taxonomies and Semantic Web technologies allow us the ability to structure data availability in a way that enables automated targeted content delivery based on:
  - Who you are
  - Which task you're working on
  - The phase of your task
- Content no longer has to stay in one place
  - It can move to where it is needed on a proactive basis
  - Increases efficiency and quality of work

JPL

The Challenge

Something to think about:

What tasks are performed where you work and how does your taxonomy define characteristics that can be used to enable smooth data flow?

JPL

Thanks for your time!

Jayne.E.Dutra@jpl.nasa.gov

JPL

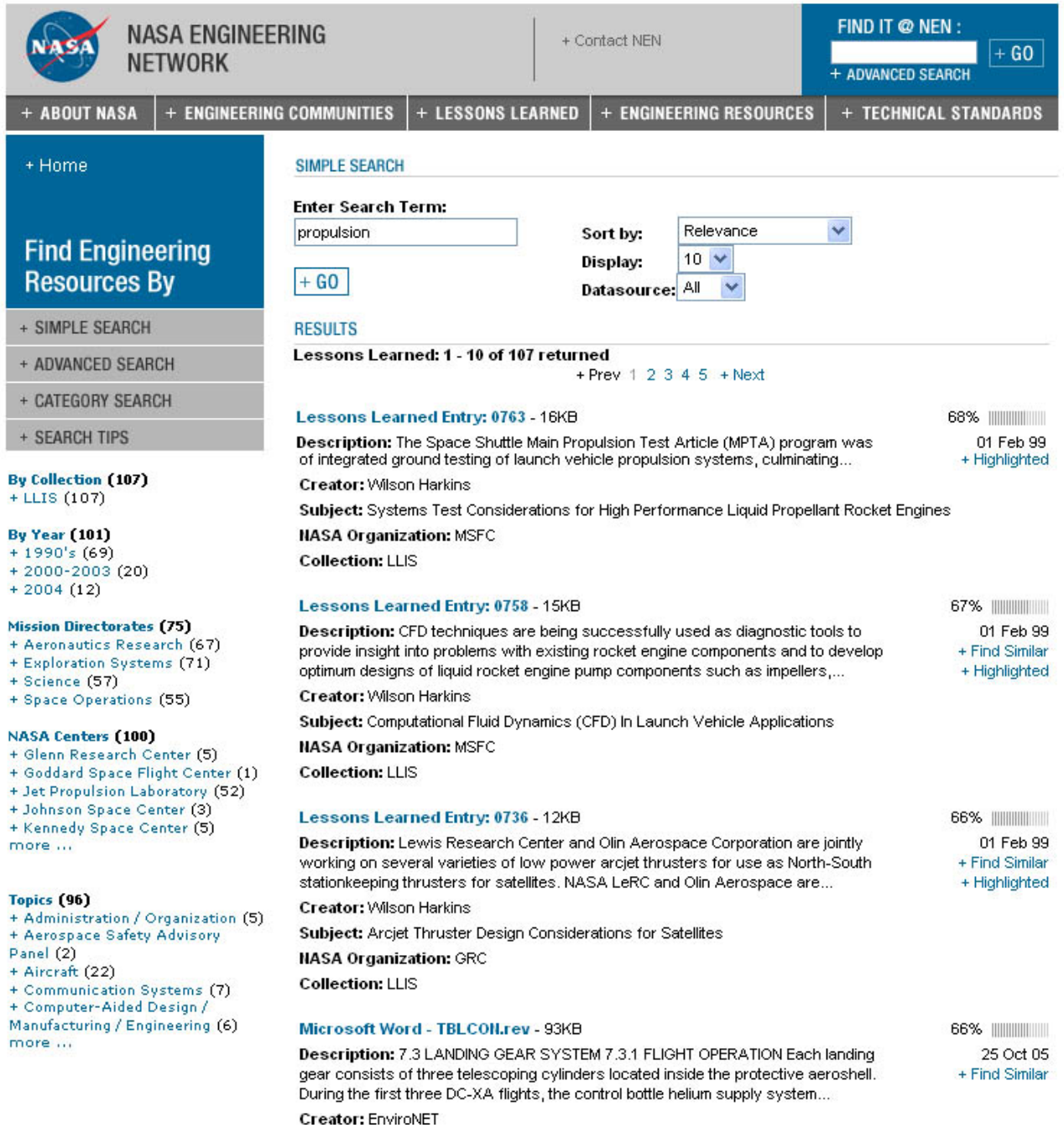
Back Up Slides

(Dry Technical Stuff)

JPL

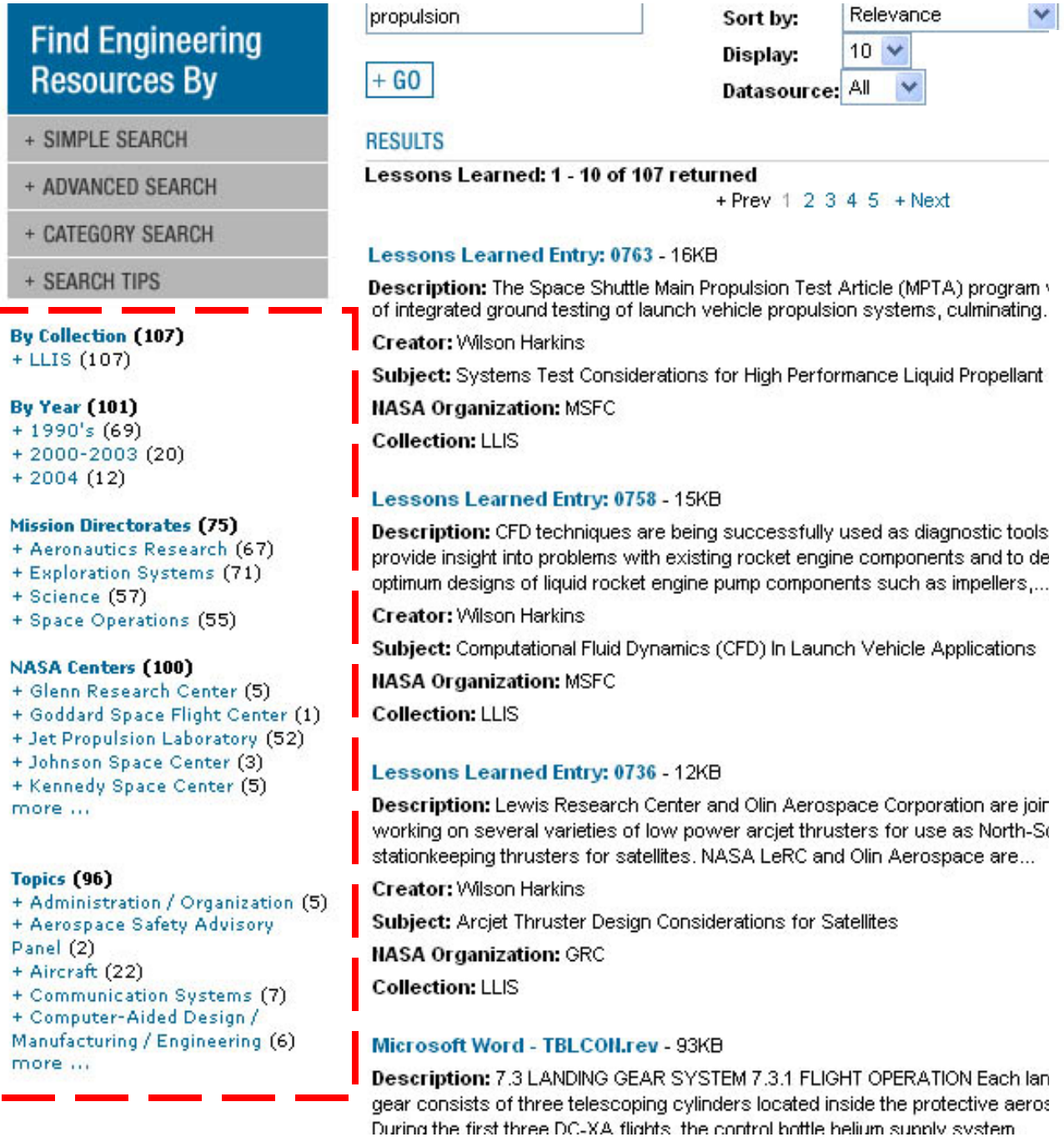
NASA Engineering Network

- Leverage the vast knowledge resources of NASA and its partners *across* domains
  - Making resources more accessible and useful by proactively integrating capabilities
- Rearchitect the way in which lessons learned are captured, stored, shared, and learned at NASA
  - Embed lessons in processes and tools using workflow
  - Provide semantic search
  - Connect engineers to expertise
  - Capture tacit knowledge
  - Manage communities of practice
  - Provide customizable portals



NEN Search Results Screenshot

The Beginning!



NEN Results Clustered By:

- Collection
- Year
- Directorate
- Organization
- Topic

What Makes a Technology Semantic?

Makes the Web understandable to computer systems

Has the ability to:

- **Represent knowledge**
  - More than just data element definitions
  - Expresses data relationships and process
  - Richness in statements about a specific knowledge domain
- **Reason over knowledge to create new knowledge**
- **Make connections between data that are non-explicit**
- **Deploy a knowledge model for run time consideration**
- **Support disparate, distributed resources**
  - Ask questions across repositories for integrated results

JPL

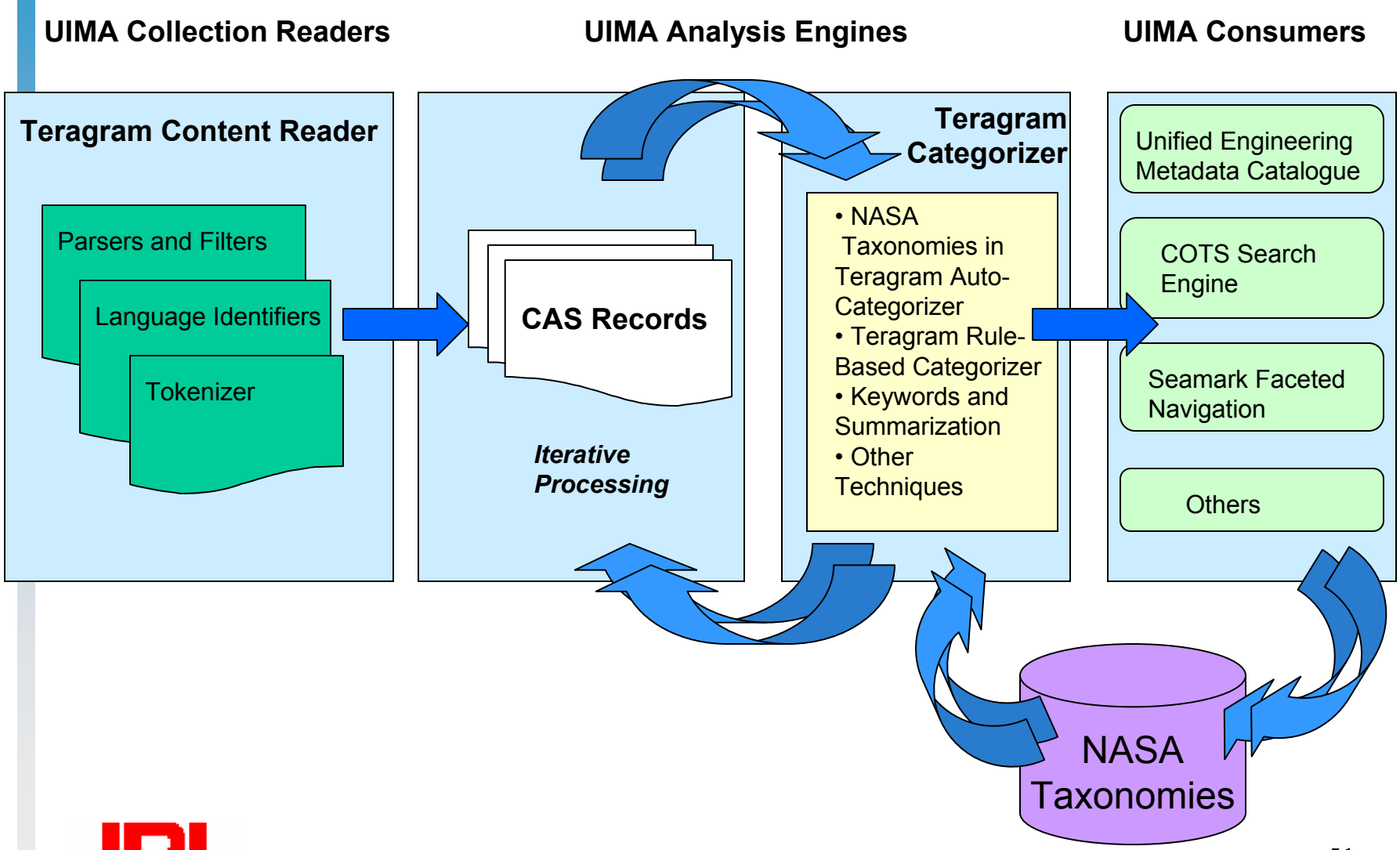
But What About the Legacy Data?

Tools, frameworks, architectures now available to deal with this problem.

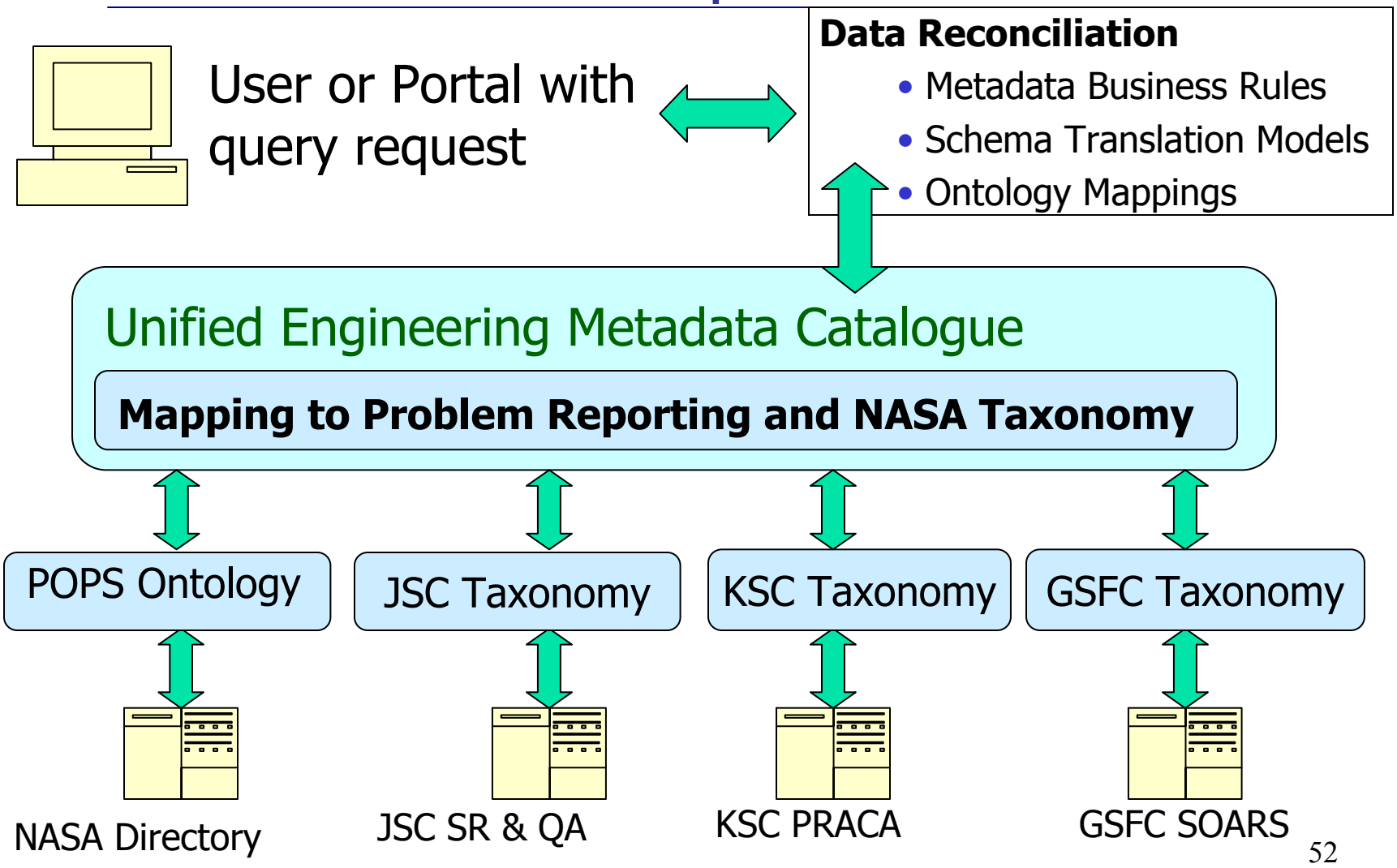
- Architecture:
  - UIMA from IBM
- Tools:
  - Teragram
  - Metatagger
  - Inxight
  - Many more

Requirements and market survey done as part of this paper

UIMA Framework Applied to Teragram Capability



Next Steps: Problem Reporting Systems Example



NEN Benefits

- Semantic Search capability will index all NASA engineering knowledge
- Helps capture NASA expert's tacit knowledge
- Cross-organizational structure and processes of tools that break down NASA's silos
- Expertise Location through POPS Semantic Search
- Communities of practice across all key NASA engineering disciplines
- Processes that encourage the sharing of lessons learned, expertise and experiences
- Tools that support the individualized processes and needs of each NASA engineering discipline
- A gold-source reference for all the tools and resources available to the NASA engineer

About UIMA

"IBM's Unstructured Information Management Architecture (UIMA) is an architecture and software framework for creating, discovering, composing and deploying a broad range of multi-modal analysis capabilities and integrating them with search technologies."

-UIMA SDK User's Guide and Reference (August 2005), p. 13

JPL