



A Semantically Structured Information Repository for Distributed NASA Project Teams



Richard M. Keller, Ph.D.

Principal Investigator

*Semantic Technologies Applied Research Team
Information Sharing and Integration Group
Intelligent Systems Division*

NASA Ames Research Center
Supported by NASA Programmatic R&D Funds



Contributors



Core Team

- Dan Berrios
- Robert Carvalho
- David Hall
- Rich Keller
- Ian Sturken
- Shawn Wolfe

Former Teammates

- Jim Chen
- Mike Compton
- Jon Guice
- Dennis Heher
- Kim Hubbard
- Larry Kiser
- Deepak Kulkarni
- David Nishikawa
- Steve Rich
- Keith Swanson
- Sergey Yentus

Key Engineering and Science Collaborators

- Brad Bebout
- Dave Blake
- Dave Des Marais
- David McKay
- Tina Panontin
- Maarten Sierhuis
- James Williams



Outline



- **Introduction & Motivation**
- **SemanticOrganizer system**
- **Applications**
- **Lessons learned**



Motivation:

Information Sharing across Collaborating Teams



- Large distributed NASA project teams must work together to address complex science and engineering questions
- Examples of distributed “virtual” teams:
 - NASA Astrobiology Institute
 - NASA Accident Investigation Teams
 - NASA Mission Design Teams
- Paramount need: to share complex science & engineering data/information across teams

Solut **S** **What solutions are employed?** **Server!**

(e.g., use FTP Server, WebDAV, Document Server, Content Management System)



Our Contention: Conventional file-sharing solutions are insufficient!



- **Little discipline or organization to files; no “librarian”**
- **Lack of metadata**
- **Lack of shared terminology/vocabulary for naming files, folders, or metadata**
- **Many files not inherently indexable/searchable via text-based approaches (lots of data & images – not much text)**
- **Complete reliance on hierarchical file structure for indexing – no multidimensional cross-indexing**



“Islands of Information”



Co-locating information is not enough!

- Disconnected “islands” of information
- File-shared information is:

- not integrated
- not interconnected
- not contextualized

*difficult to ‘connect the dots’
and see the big picture*

- Difficult to see the connections between data/information gathered by different people in different places



Desiderata for an Improved Knowledge Sharing System



- Integrate/interrelate information across the “islands”
- Permit rapid, intuitive access and search of the information space
- Provide high degree of customizability to accommodate each project’s unique work products, processes, and goals
- Enhance consistency and completeness of shared information
- Capture contextual information associated with shared information & data products



1. **Centralize**: Establish centralized access to all project information within a common repository
2. **Standardize**: For a given distributed project team, *engineer* a common information model and vocabulary for describing and interrelating all project information
3. **Contextualize**: Annotate the data with ample semantic context to provide a basis for improved access and search

**Develop a Semantically-structured
Information Sharing Repository**



What is a Semantic Information Repository?



An information management system

(i.e., database / web server / file server / document mgmt. system)

+

an ontology!

Repository users (people *and* intelligent agents):

Upload / enter heterogeneous content:

1. datafiles/documents/images

2. *records* documenting people, places, things, activities
that provide *context* to data

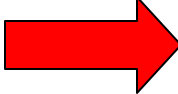
Interlink / interrelate content within a
semantic information network

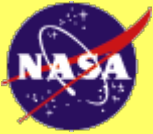
Search, access, and download content



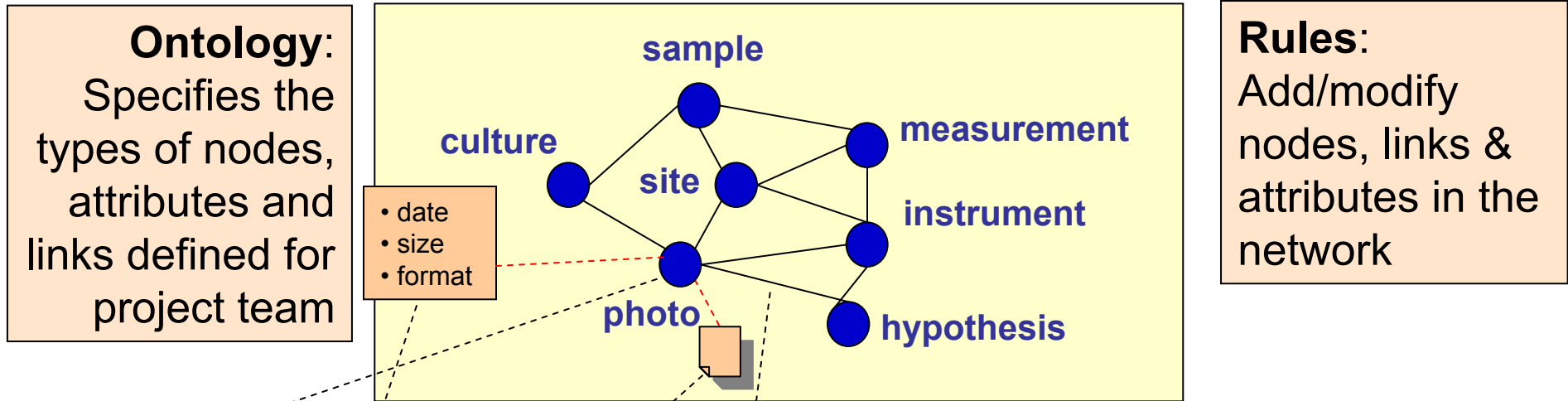
Outline



-  • **Introduction & Motivation**
- **SemanticOrganizer system**
- **Applications**
- **Lessons learned**



SemanticOrganizer Core Representation: Semantic Network Structure



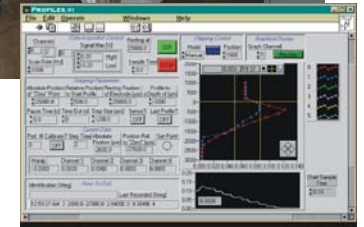
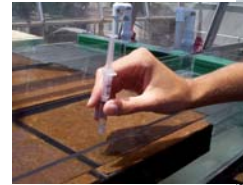
- **Nodes:** key project resources (e.g., people, places, artifacts, activities, processes)
 - **Attributes:** properties of resources (metadata)
 - **Links:** relationships among resources
- **Attached files:** electronic products associated with resources (in almost any type of format)



SemanticOrganizer Sample Application: Astrobiology Research Investigations



**Microbial mat
(algae)**



- Detailed studies of
mat biogeochemistry**
- monitoring
 - analysis
 - experimentation



**Collection of
microbial mats
in the field**

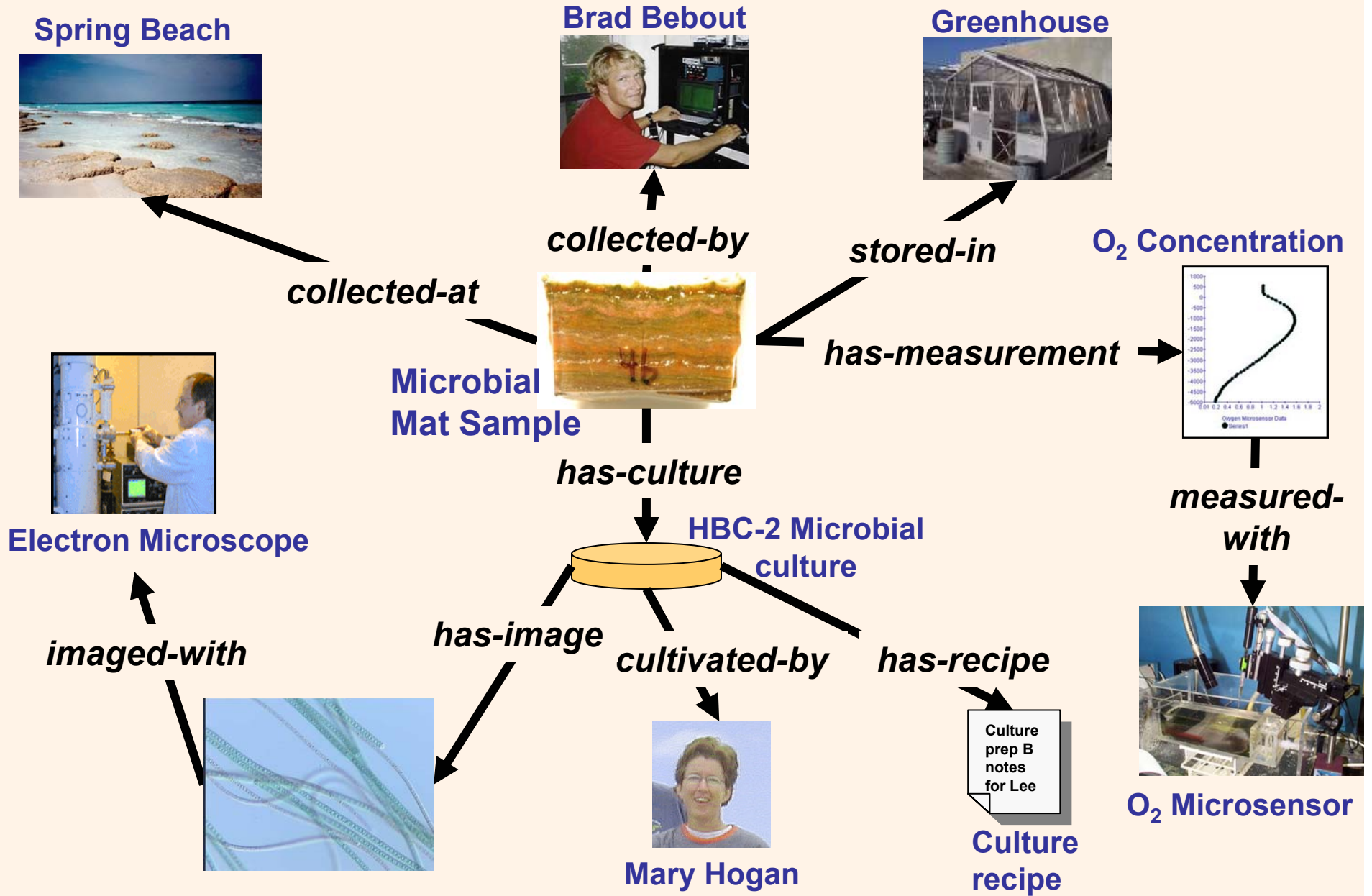
Greenhouse Incubator

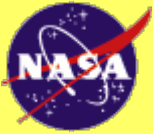


geographically-disbursed
team of collaborators

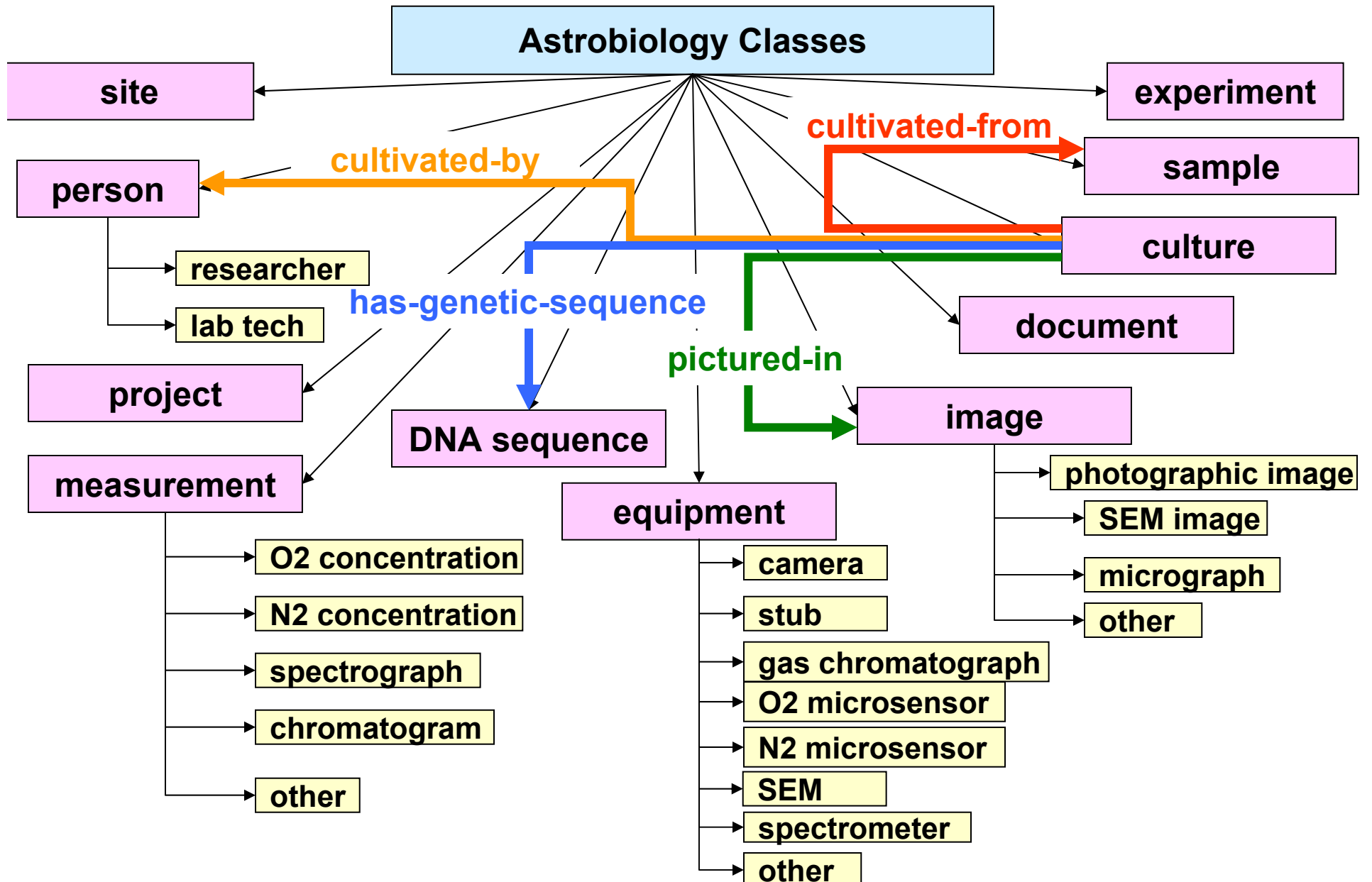


Semantic Network Fragment from Astrobiology Application



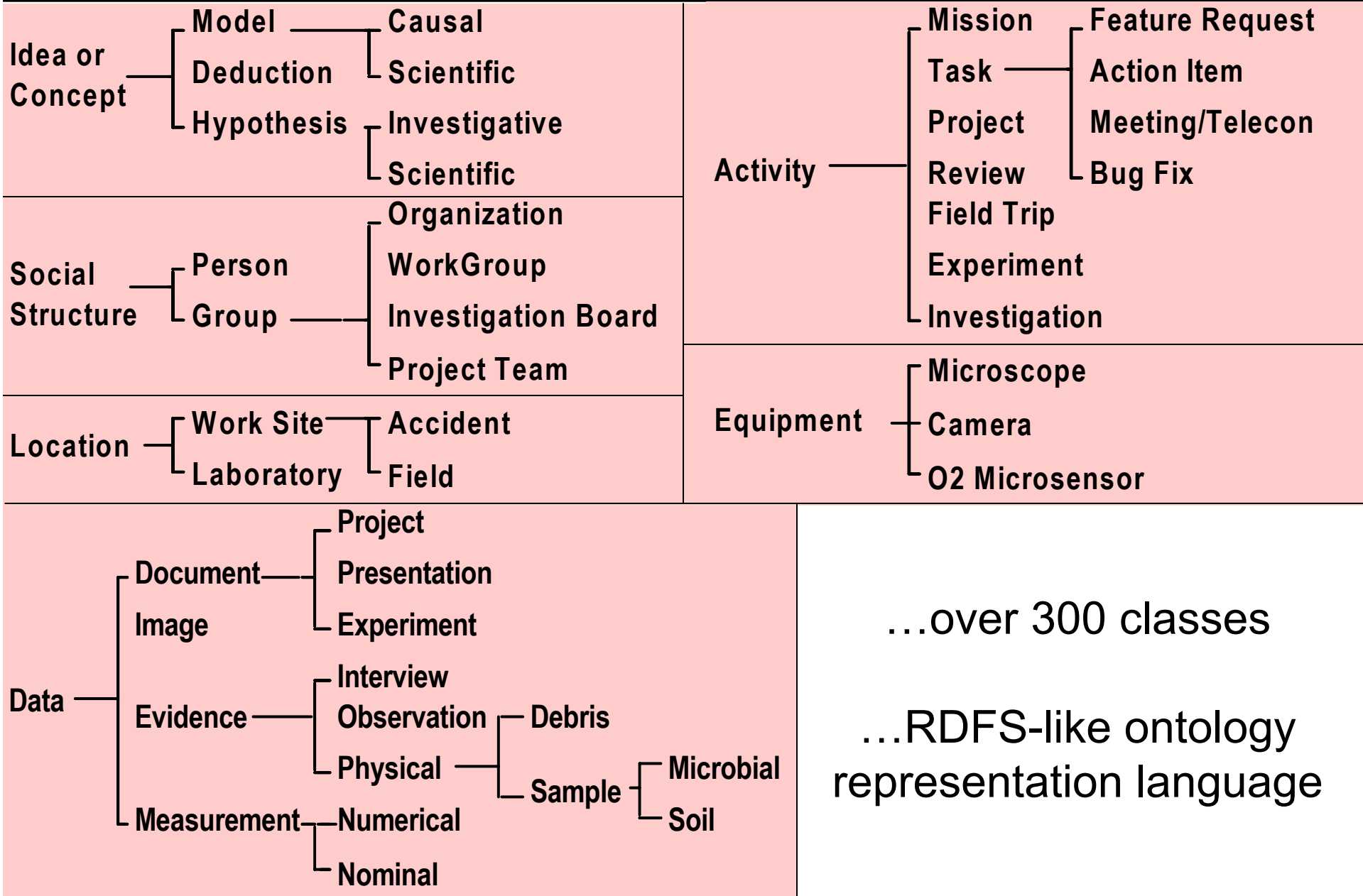


Astrobiology Ontology (partial)



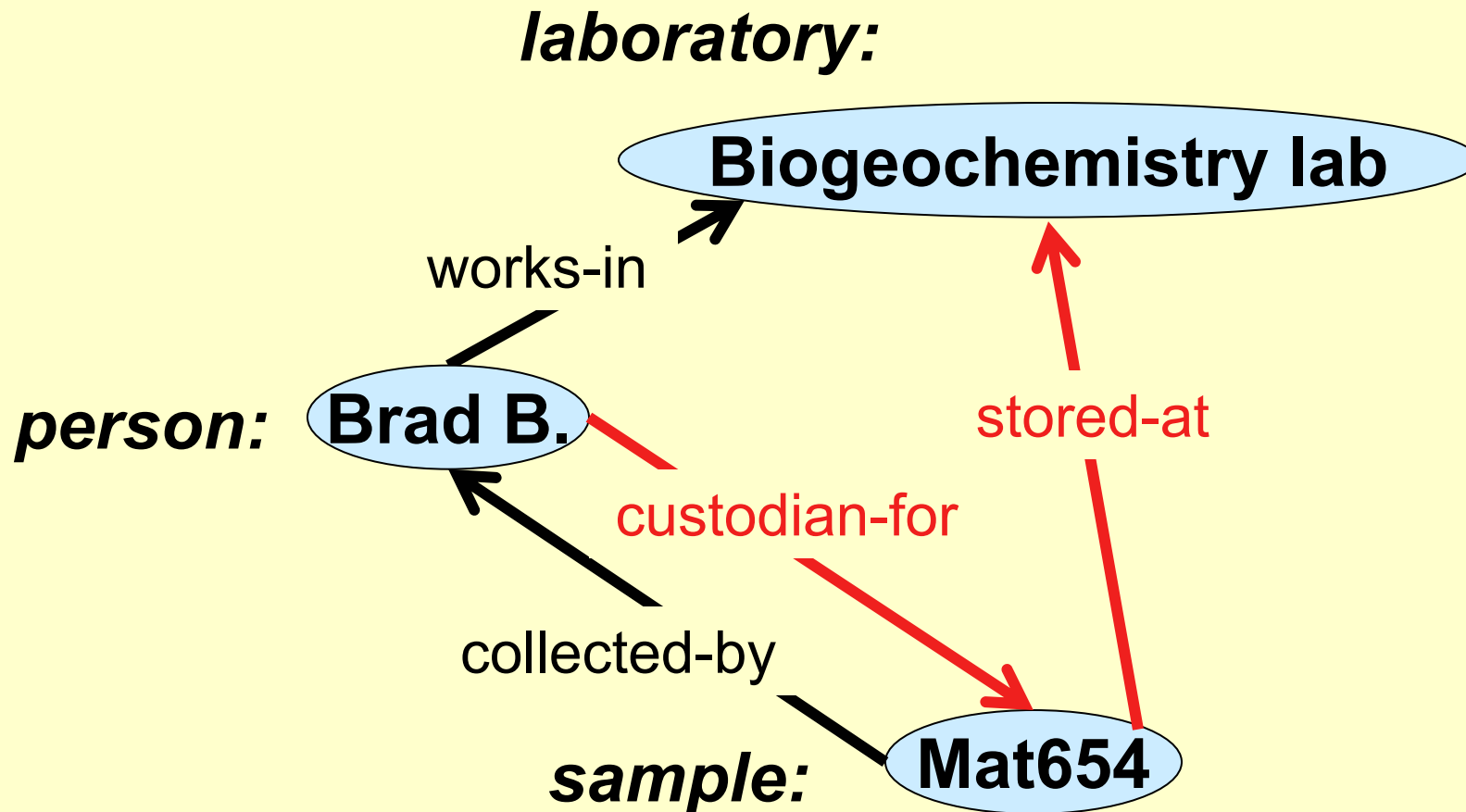


Unified Master Ontology



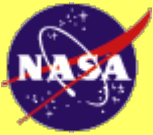


Simple Example of Rule-based Linking



If person P is *custodian-for* sample S, and P *works-in* laboratory L,
Then S is *stored-at* L

If sample S is *collected-by* person P, and S has no custodian,
then P is *custodian-for* S



SemanticOrganizer's Browser-based Interface

(Algae Mat Sample Node: Spring-M4-b)



create new nodes create new links search for nodes icon identifies node type modify node metadata

Links to Related Nodes

semantic links

related nodes (click to navigate)

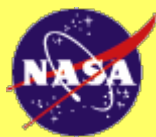
Focal Node

Node Metadata

New Item	Search	Home	Go To	Logout	Help
View Links	Edit Links	Modify Item	Modify Permissions	Delete Item	Duplicate Item
				Put Item In Folder	

- ☆ Spring-M4-b (open all | close all)
 - Associated Project (1 Projects)
 - EMERG
 - Collected At (1 Field Sites)
 - Spring Beach (tidal)
 - Collected By (1 Persons)
 - Bebout, Brad
 - Collected During (1 Field Trips)
 - Baja June 1999
 - Has Custodian (1 Persons)
 - Hogan, Mary
 - Has Measurement (1 Measurements)
 - ExpID#131791-3
 - Has Sequence Info (1 DNA Sequences)
 - 16S3 rRNA sequence
 - Pictured In (1 Images)
 - SprM4b excised**
 - Source Of (1 Cultures)
 - HBC-X2

☆ Mat Sample: Spring-M4-b	
Item ID# 132888 updated 2001/05/23 10:52AM PDT	
Send this Item's web address via Email	
Description	Vibrant yellow-orange slimy mat, approx. 10" square 1cm thick.
Collection Date	03/22/2001
Collection Time	14:30:00
Sample Type	benthic streamer mat
Microorganisms	Synechococcus ("lividus"-type)
Laboratory Preservative	alcohol
Preservative Field Dilution	.2 %
Preservative Buffer	RB-5 ChlorAC
Sample Container Type	tupperware
Storage	greenhouse co
Water Temperature	27 °C
Water pH	7.0
Notes	Excised from a sandy site.
Write Permission	● Bebout, Brad
Read Permission	● EMERG



Sample is *pictured-in* Photo



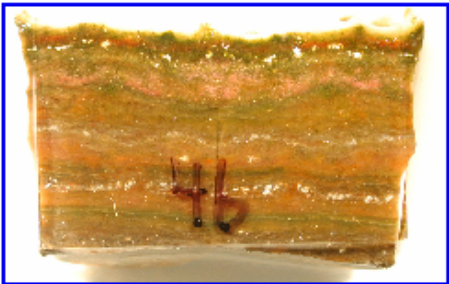
New Item	Search	Home	Go To	Logout	Help
View Links	Edit Links		Modify Item	Modify Permissions	Delete Item
			Duplicate Item	Put Item In Folder	


- [-] SprM4b excised (open all | [close all](#))
 - [-] **Camera Used** (1 Cameras)
 - [-] [Brad's Digital Olympus](#)
 - [-] **Image Of** (1 Samples)
 - [-] [Spring-M4-b](#)
 - [-] **Imaged By** (1 Persons)
 - [-] [Bebout, Brad](#)

Photo: SprM4b excised

Item ID# 132919 updated 2003/06/14 04:28PM PDT

Send this Item's web address via [Email](#)



Click image to view at ACTUAL size.
 Right Click (or click and hold on Macintosh) to download image.
 [Click here to download associated jpg file](#)
[Help with downloading](#)
[Annotate Image](#)

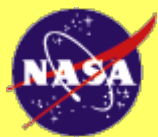
Caption Sample exhibits classic layering associated with mats from the Baja region.

Image Date 03/22/2001

Notes Original photo stored in N239 room 433.

Write Permission [Bebout, Brad](#)

Read Permission [EMERG](#)



Edit Links



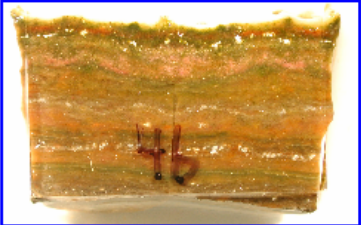
ScienceOrganizer: An Information-Sharing Tool For Scientific Project Teams [NASA Ames](#)

[New Item](#) [Search](#) [Home](#) [Go To](#) [Logout](#) [Help](#)
[View Links](#) **Edit Links** [Modify](#) [Permissions](#) [Delete](#) [Duplicate](#) [Put in a Folder](#)

SprM4b excised (open all | [close all](#))
Camera Used (1 Cameras)
 Link to [existing](#) / [new](#) Cameras
 [Brad's Digital Olympus](#)
Image Of (1 Mat Samples/0 Sediment Samples/0 Study Areas)
 Link to [existing](#) / [new](#) Sediment Samples
 Link to [existing](#) / [new](#) Study Areas
 Link to [existing](#) / [new](#) Mat Samples
 Link to [existing](#) / [new](#) Stromatolite Samples
 Link to [existing](#) / [new](#) Field Sites
 Link to [existing](#) / [new](#) Water Samples
 Link to [existing](#) / [new](#) Gas Samples
 [Spring-M4-b](#)
Imaged By (1 People)
 Link to [existing](#) / [new](#) People
 [Bebout, Brad](#)
 Other Permissible Links...
Contained By (0 Photo Folders/0 Image (other) Folders)
 Link to [existing](#) / [new](#) Photo Folders
 Link to [existing](#) / [new](#) Image (other) Folders
 Link to [existing](#) / [new](#) Item (other) Folders
Figure For Paper (0 Publications)
 Link to [existing](#) / [new](#) Publication
Illustration For (0 Models)
 Link to [existing](#) / [new](#) Models
Is Discussed In (0 Email Messages)
 Link to [existing](#) / [new](#) Email Messages
Map For (0 Study Areas/0 Field Sites)
 Link to [existing](#) / [new](#) Study Areas

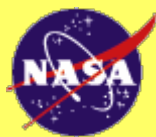
Photo: SprM4b excised

Item ID# 132919 updated 2003/06/14 04:28PM PDT by [Demo, EMERG](#)
 Send this Item's web address via [Email](#)



Click image to view at ACTUAL size.
 Right Click (or click and hold on Macintosh) to download image.
[Click here to download associated jpg file \(389214 bytes\)](#)
[Help with downloading](#)
[Annotate Image](#)

Camera Used [Brad's Digital Olympus](#)
Caption Sample exhibits classic layering associated with mats from the Baja region.
Image Date 03/22/2001
Imaged By [Bebout, Brad](#)
Sample [Spring-M4-b](#)
Notes Original photo stored in N239 room 433.
Write Permission [Demo, EMERG](#)
Read Permission [Demo, EMERG](#)



Linking to an Existing Item



ScienceOrganizer: An Information-Sharing Tool For Scientific Project Teams NASA Ames

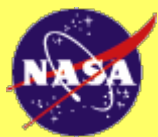
[New Item](#) [Search](#) [Home](#) [Go To](#) [Logout](#) [Help](#)

Select Items to Link via 'Figure For Paper'

Add 'Figure For Paper' Link from [SprM4b excised](#) to either selected Publications below or a [new Publication](#).

Viewing Items 1-18 of 18 Didn't find what you want? [Revise Search](#)

<input type="checkbox"/> type / <input checked="" type="checkbox"/> name	updated	by
<input type="checkbox"/> Abstracts for ABSciCon	2002/03/01 11:20AM PST	Des Marais, Davi
<input type="checkbox"/> BaumgartnerAbSciCon02	2002/03/01 11:31AM PST	Des Marais, Davi
<input type="checkbox"/> Bebout et al 2004	2006/03/14 07:07AM PST	Bebout, Brad
<input checked="" type="checkbox"/> Brad Bebout et al. - Sulfate Experiment	2002/01/31 03:39PM PST	Bebout, Brad
<input type="checkbox"/> Brad's GH sulfate draft 4 + Tori's	2003/07/07 10:35AM PDT	Hoehler, Tori
<input type="checkbox"/> Castenholz AbSciCon2002 Baja Abstract	2002/01/31 04:39PM PST	Bebout, Brad
<input type="checkbox"/> Castenholz AbSciCon2002 YNP Abstract	2002/01/31 02:18PM PST	Fleming, Erich
<input type="checkbox"/> Decker and Potter 2001	2002/05/10 12:02PM PDT	
<input type="checkbox"/> DesMaraAbSciConMcoleus&Lyng	2002/02/18 04:52PM PST	Des Marais, Davi
<input type="checkbox"/> Dillon AbSciCon Abstract	2002/01/31 08:15AM PST	Dillon, Jesse
<input type="checkbox"/> Draft IV	2003/06/29 07:10AM PDT	Bebout, Brad
<input type="checkbox"/> Greenhouse Salinity Manuscript	2002/05/07 03:43PM PDT	Bebout, Brad
<input type="checkbox"/> HabichtAbSciCon02	2002/03/01 12:15PM PST	Des Marais, Davi
<input type="checkbox"/> Jahnke et al.-Lipid Biomarkers	2002/02/04 01:29PM PST	Jahnke, Linda
<input type="checkbox"/> RothrockAbSciCon02	2002/03/01 11:34AM PST	Des Marais, Davi
<input type="checkbox"/> Sulfate Draft III + Tori's	2003/05/07 11:05AM PDT	Hoehler, Tori
<input type="checkbox"/> Sulfate Experiment Manuscript #1	2003/04/10 09:08AM PDT	Bebout, Brad
<input type="checkbox"/> Tori's AbSciCon02 Abstract	2002/01/30 10:17AM PST	Hoehler, Tori



Edit Links



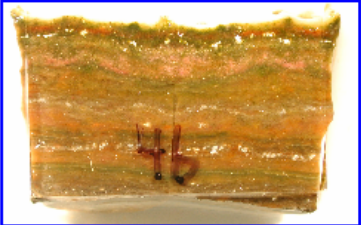
ScienceOrganizer: An Information-Sharing Tool For Scientific Project Teams NASA Ames


[New Item](#) [Search](#) [Home](#) [Go To](#) [Logout](#) [Help](#)
[View Links](#) **Edit Links** [Modify](#) [Permissions](#) [Delete](#) [Duplicate](#) [Put in a Folder](#)

- ☒ SprM4b excised (open all | [close all](#))
- ☒ **Camera Used** (1 Cameras)
 - Link to [existing](#) / [new](#) Cameras
 - ☒ [Brad's Digital Olympus](#)
- ☒ **Image Of** (1 Mat Samples/0 Sediment Samples/0 Study Areas)
 - Link to [existing](#) / [new](#) Sediment Samples
 - Link to [existing](#) / [new](#) Study Areas
 - Link to [existing](#) / [new](#) Mat Samples
 - Link to [existing](#) / [new](#) Stromatolite Samples
 - Link to [existing](#) / [new](#) Field Sites
 - Link to [existing](#) / [new](#) Water Samples
 - Link to [existing](#) / [new](#) Gas Samples
 - ☒ [Spring-M4-b](#)
- ☒ **Imaged By** (1 People)
 - Link to [existing](#) / [new](#) People
 - ☒ [Bebout, Brad](#)
- Other Permissible Links...
 - ☒ **Contained By** (0 Photo Folders/0 Image (other) Folders)
 - Link to [existing](#) / [new](#) Photo Folders
 - Link to [existing](#) / [new](#) Image (other) Folders
 - Link to [existing](#) / [new](#) Item (other) Folders
 - ☒ **Figure For Paper** (0 Publications)
 - Link to [existing](#) / [new](#) Publication
 - ☒ **Illustration For** (0 Models)
 - Link to [existing](#) / [new](#) Models
 - ☒ **Is Discussed In** (0 Email Messages)
 - Link to [existing](#) / [new](#) Email Messages
 - ☒ **Map For** (0 Study Areas/0 Field Sites)
 - Link to [existing](#) / [new](#) Study Areas

Photo: SprM4b excised

Item ID# 132919 updated 2003/06/14 04:28PM PDT by [Demo, EMERG](#)
 Send this Item's web address via [Email](#)



Click image to view at ACTUAL size.
 Right Click (or click and hold on Macintosh) to download image.
 [Click here to download associated jpg file \(389214 bytes\)](#)
[Help with downloading](#)
[Annotate Image](#)

Camera Used [Brad's Digital Olympus](#)

Caption Sample exhibits classic layering associated with mats from the Baja region.

Image Date 03/22/2001

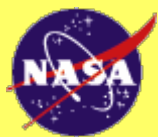
Imaged By [Bebout, Brad](#)

Sample [Spring-M4-b](#)

Notes Original photo stored in N239 room 433.

Write Permission [Demo, EMERG](#)

Read Permission [Demo, EMERG](#)



Linking to New Item



ScienceOrganizer: An Information-Sharing Tool For Scientific Project Teams

NASA Ames

[New Item](#) [Search](#) [Home](#) [Go To](#) [Logout](#) [Help](#)

[View Links](#) [Edit Links](#)

SprM4b excised (open all | [close all](#))

- Camera Used (1 Cameras)
 - Link to [existing](#) / [new](#) Cameras
 - Brad's Digital Olympus
- Image Of (1 Mat Samples/0 Sediment Samp)
 - Link to [existing](#) / [new](#) Sediment Samp
 - Link to [existing](#) / [new](#) Study Areas
 - Link to [existing](#) / [new](#) Mat Samples
 - Link to [existing](#) / [new](#) Stromatolite Sa
 - Link to [existing](#) / [new](#) Field Sites
 - Link to [existing](#) / [new](#) Water Samples
 - Link to [existing](#) / [new](#) Gas Samples
- Spring-M4-b
- Imaged By (1 People)
 - Link to [existing](#) / [new](#) People
 - Bebout, Brad
- Other Permissible Links...
 - Contained By (0 Photo Folders/0 Im)
 - Link to [existing](#) / [new](#) Photo Fol
 - Link to [existing](#) / [new](#) Image (oth
 - Link to [existing](#) / [new](#) Item (othe
 - Figure For Paper (0 Publications)
 - Link to [existing](#) / [new](#) Publicatio
 - Illustration For (0 Models)
 - Link to [existing](#) / [new](#) Models
 - Is Discussed In (0 Email Messages)
 - Link to [existing](#) / [new](#) Email Mes
 - Map For (0 Study Areas/0 Field Sites)
 - Link to [existing](#) / [new](#) Study Arc
 - Link to [existing](#) / [new](#) Field Sites
 - Organizer Home Page For (0 People)

Create either a new Publication Folder or Publication:

Publication Folder **Publication**

Publication Name:

Locate the Publication file to upload from your local systems:

or type the URL that points to the remote file.

[Help with remote files](#)

Please select the file type for this file

Check here to unpack this file after unloading (for zip files only!):

Name is optional for zip files and will be prepended to internal file names.

To upload multiple files at once, [follow these directions](#)

Publication Title

Publication Type

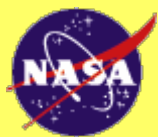
Publication Status draft final

Acceptance Status

Author
[Help with selecting multiple entries](#)

Project
[Help with selecting multiple entries](#)

Journal



Resulting Publication with Links Established




ScienceOrganizer: An Information-Sharing Tool For Scientific Project Teams NASA Ames

[New Item](#) [Search](#) [Home](#) [Go To](#) [Logout](#) [Help](#)

[View Links](#) [Edit Links](#)

BioSim Journal Article (open all | [close all](#))

- Authoring By (1 People)
 - [Bebout, Brad](#)
- Document For Project (1 Groups)
 - ◇ [EMERG](#)
- Has Figure (1 Images/0 Figures)
 - [SprM4b excised](#)
- Publication For Project (1 Groups)
 - ◇ [EMERG](#)

Publication: BioSim Journal Article
Item ID# 201297 updated 2006/10/30 09:42AM PST by ● [Demo, EMERG](#)
Send this Item's web address via [Email](#)
 [Click here to download associated pdf file \(2092906 bytes\)](#)
[Help with downloading](#)

Publication Title	Stramatolite Populations in Baja Mexico under Longerm Environmental Stress Conditions
Publication Type	journal article
Publication Status	draft
Acceptance Status	in review
Author	● Bebout, Brad
Project	◇ EMERG
Journal	Biosimulation



Search



Search within Names Files Fields

or

Select type(s) to search:

- All Items
- Culture
- DNA Sequence
- Email Message
- Experiment Log
- Field Site
- Field Trip
- Gas Sample
- Global Control
- GreenHouse Table Pass
- Institution
- Laboratory
- Mailing List
- Mat Sample
- Model
- Note
- Person
- Sediment Sample
- Stromatolite Sample
- Study Area
- Test Point
- URL
- Water Sample
- Item (other)

- All Groups
- Project
- Workgroup

- All Experiments
- GreenHouse Image Experiment

- All Images
- Micrograph
- Photo
- Phylogenetic Tree
- SEM Image
- Image (other)

- All Measurements
- Acetylene Reduction
- Ammonium
- Biomarker Analysis
- Carbon Isotope
- Carbon Monoxide
- Dissolved Inorganic Carbon
- Dissolved Organic Carbon
- Fluorescence
- Hydrogen Sulfide
- Hydrogen
- Light
- Mat Image
- Methane
- Nitrate
- Orthophosphate
- Oxygen
- pH
- Salinity
- Temperature
- Volatile Sulfur
- Measurement (other)

- All Documents
- Culture Recipe
- Equipment Document
- Experiment Document
- Figure
- Presentation/Poster
- Project Document
- Publication
- Trip Document
- Document (other)

- All Equipment
- Camera
- Flow Box
- Fluorometer
- Gas Chromatograph
- H Microsensor
- H2S Microsensor
- Ion Chromatograph
- Light Microsensor
- Mass Spectrometer
- O2 Microsensor
- pH Meter
- pH Microsensor
- Refractometer
- SEM
- Spectrophotometer
- Thermometer
- Tray
- Equipment (other)

Enter criteria for Field-Restricted Search. [Help with Restricting Fields](#)

Fill in any Fields you want to restrict during search.
You can specify a range for numbers, dates, and time.
Leave blank any fields you do not care about or are unsure about.

Publication Name:

Publication Title:

Publication Type:
journal article
conference paper
magazine article

Publication Status:
final

Acceptance Status:
in review
accepted
rejected

Author:
Albert, Dan
Baumgartner, Laura Kathleen
Bebout, Brad [Help with selecting multiple entries](#)

Project:
Ames Exobiology Culture Collection
EMERG [Help with selecting multiple entries](#)

Journal:

Issue:

Pages:

Conference/Workshop Proceedings Title:

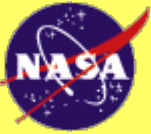
Edited Collection Title:



Additional SemanticOrganizer Features



- Semantic Search
- Email discussion list integration
- Automated semantic hyperthreading
- Microsoft office integration
- Collaborative image annotation
- Interoperation w/external agents



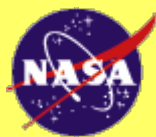
Semantic Search



- Search by semantic pattern (patterns involving extended chains of nodes, links, and values)
- Executed by matching pattern against the instance nodes and links in the semantic network

“Find all DNA Sequences associated with Stromatolite Samples collected at the Field Site named “Stromatolite Beach””

The screenshot shows a web interface for semantic search. At the top, there are navigation buttons: Home, New Item, Search, Logout, and Help. Below these are buttons for 'Perform Search' and 'Start Over'. The main area is titled 'Your query so far...' and contains a search form. The form has a dropdown menu for 'Set item to:' with the value 'HBC-1: 16S rRNA sequence'. Below this is a table with columns for 'Sequence For' and 'Name'. The 'Sequence For' column contains 'Stromatolite Sample 1' and an 'Add New Link' button. The 'Name' column contains various fields: 'Name', 'Sequence Origin' (no search value), 'Gene' (no search value), 'Sequencing Laboratory', 'Sequencing Method' (radio buttons for 'Both strands sequenced' and 'One strand sequenced'), 'Sequencing Date' (mm/dd/yyyy), 'Reading Direction' (radio buttons for 'bidirectional', '5 --> 3', and '3 --> 5'), 'Best BLAST Hit Data:' (highlighted in yellow), '% Identity' (%), 'Genus', 'Species', and 'Taxa'. There is also a 'Notes' field and a 'Set Field(s)' button. To the right of the main form is a sidebar with a 'Set item to:' dropdown set to 'Core T-1B'. Below this is a section for 'Collected At' with 'Field Site 1' and an 'Edit' button. Below that is 'Has Sequence Info' with 'DNA Sequence 1' and an 'Unset' button. At the bottom of the sidebar is a button for 'Field Site 1: Stromatolite Beach'. The search results area at the bottom shows '0 Search Results' and a timestamp 'Search started on Apr 16, 2004 4:00:33 PI' with a 'Name This Search' button.



Email Discussion List Integration



ScienceOrganizer: An Information-Sharing Tool For Scientific Project Teams NASA Ames/Computational Sciences

[New Item](#) [Find Items](#) [Home](#) [Go To](#) [Logout](#) [Help](#)

[View Links](#) [Edit Links](#) [Modify Item](#) [Modify Permissions](#) [Delete Item](#) [Duplicate Item](#) [Put Item In Folder](#)

March Baja trip logistics (open all | [close](#))

Followed By (1 Email Messages)

- [Greenhouse Mat Experiment and T](#)

Preceded By (1 Email Messages)

- [Re: Baja, again](#)

Sent By (1 Persons)

- [Des Marais, David](#)

Sent To (1 Persons/Mailing Lists)

- [EMERG General Mailing List](#)

Other Permissible Links...

- Contained By** (0 Email Message Fo (other) Folders)
- Discusses** (0 Items)
- Is Discussed In** (0 Email Messages)
- Next** (0 Email Message Folders)
- ScienceOrganizer Home Page For** (0 Persons)
- Unspecified** (0 Items)

Email Message: March Baja trip logistics

Item ID# 60371 updated 8/1/02@11:14AM
Send this Item's web address via [Email](#)

'From:' Line David Des Marais <d-desmarais@mail.arc.nasa.gov>
Sender [Des Marais, David](#)

'To:' Line emerg@science.nasa.gov, "David A. Stahl" <d-stahl@nwu.edu>, "John R. Spear" <John.Spear@Colorado.EDU>, SOGIN@evol5.mbl.edu
Recipients [EMERG General Mailing List](#)

Date sent
Date received 2001-02-16 00:00:00

Body

Dear Group,

We are still looking forward to a March 5 to 18 schedule for the Baja trip. Last week, our Mexican research permit was recommended forward from SEMARNAP, Mexico's Department of the Environment and Fisheries, to SRe, their State Department. This week, the State Department reviewed the package, received inputs from other relevant agencies, and made additional requests from me with which I have now complied. I expect that we will receive final approval within days.

Accordingly, if you feel that making plane reservations now will save you more funds than you might lose with a postponement, please go ahead and make those reservations.

TRAVELING TO MEXICO

If you are taking a plane flight to join the trip, you are a member of the "Airplane Group." You should plan to meet in San Diego on March 5 and ride down together in a passenger van rented from Pearson Ford in San Diego. It might not be necessary for the Airplane Group to rendezvous with the Ames Group, unless some of you want to transfer equipment to our vans for customs declaration. We currently are holding a 12 passenger van in my name, but I must transfer that reservation to someone in the "Airplane" group as soon

Links

Email Body

Enhanced email indexing via semantic linkage



Automated Semantic Hyperthreading [Concept Extraction from Email] (experimental)



Email body (extract)

Relevancy
rank

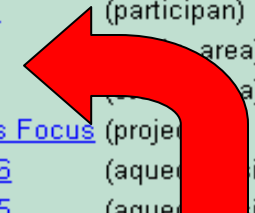
Node

Type

"FIELD OBJECTIVES

This field trip will address a number of objectives that include those of the Ecogenomics project as well as those associated with previously-funded grants. Accordingly, we will be performing biogeochemical measurements and collect samples for various analyses back at our respective laboratories...

1.063660	Spear, John	(participan
1.063660	Stahl, David	(participan
1.063660	Bebout, Brad	(participan
1.036137	Ames	(area)
1.036137	Baja	(a)
1.031293	Ecogenomics Focus	(projec
1.020290	Pond 5 near 6	(aque
1.020290	Pond 4 near 5	(aque



Straw **sampling** Protocols

In situ **measurements** of fluxes:

- ... **number** of diel **periods** : 2 (once in **Pond 4 near 5**, once in **Pond 5 near 6**)
- ... **number** of chambers: 6 (3 **mats**, three **blanks**)
- ... **sampling** times: 0600h, 0900h, 1200h, 1500h, 1800h, 2100h, 0000h, 0300h
- ... **Water volume taken** per **sampling time**
- ... **Gas volume taken** per **sampling interval**

2. Relevancy ranking score computed:

- based on heuristic combination of factors

1.009677	visscher, Pieter	(participan
1.009677	Hoehler, Tori	(participan
1.009677	Hogan, Mary	(participan
1.009677	Garcia-Pichel, Fer	(participan
1.009677	Dillon, Jesse	(participan
1.009677	Turk, Kendra	(participan
1.009677	Miller, Scott	(participan

...SAMPLING

We recognize that the acquisition and preservation of samples for later use will be a major objective of this trip. A discussion of sampling preferences during the trip will be held at Ames.



1. Text marked up to identify:

- nouns matching node names/types
- verbs corresponding to links
- references to attribute values
- WordNet synonyms/hyponyms used to enable inexact semantic matching



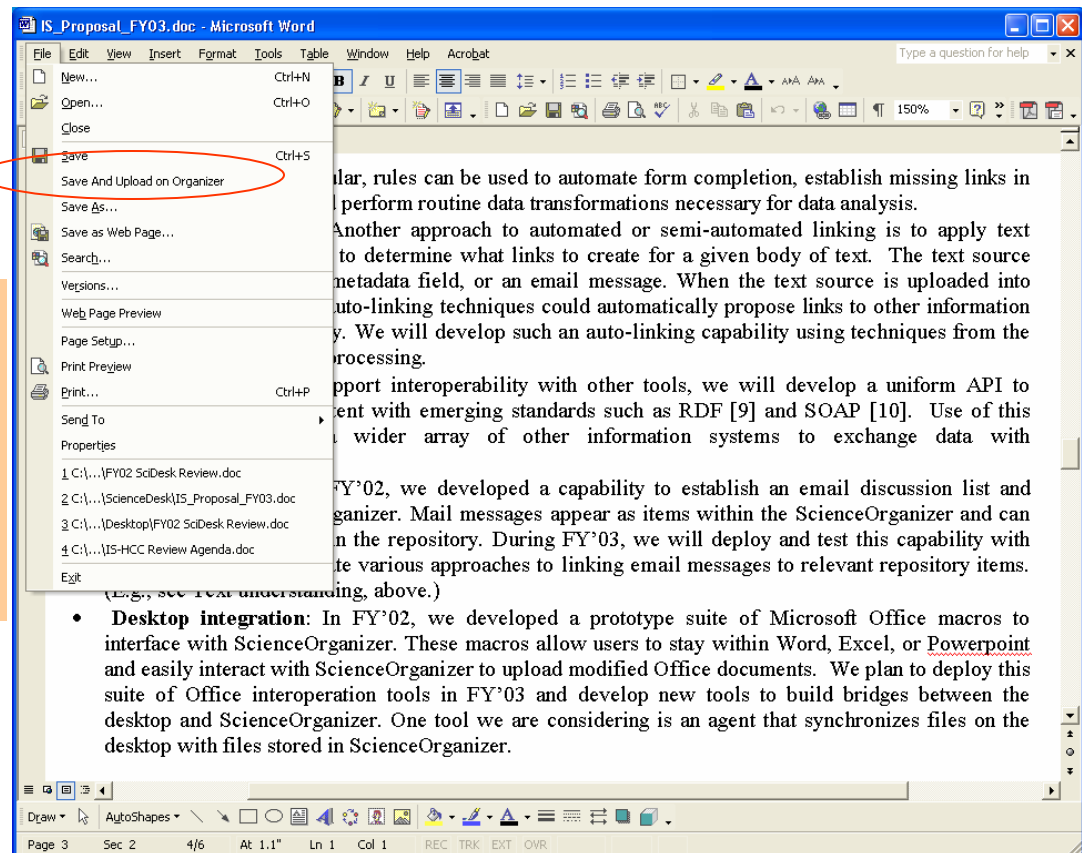
Microsoft Office Interoperation

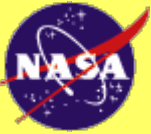


- Developed suite of Microsoft Office macros to enable direct upload and subsequent modification of documents without leaving Office application

“Save and upload to Organizer”

- Macro communicates w/server
- User fills out metadata using standard Organizer form on creation
- Subsequent saves are transparent

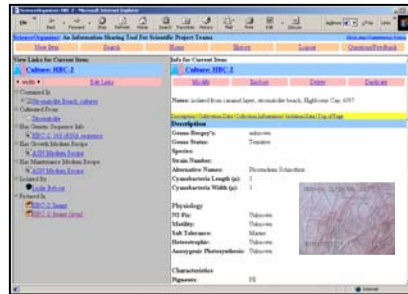




Shared Image Annotator



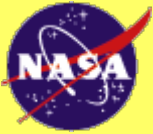
Collaborative Whiteboarding & Annotation Tool



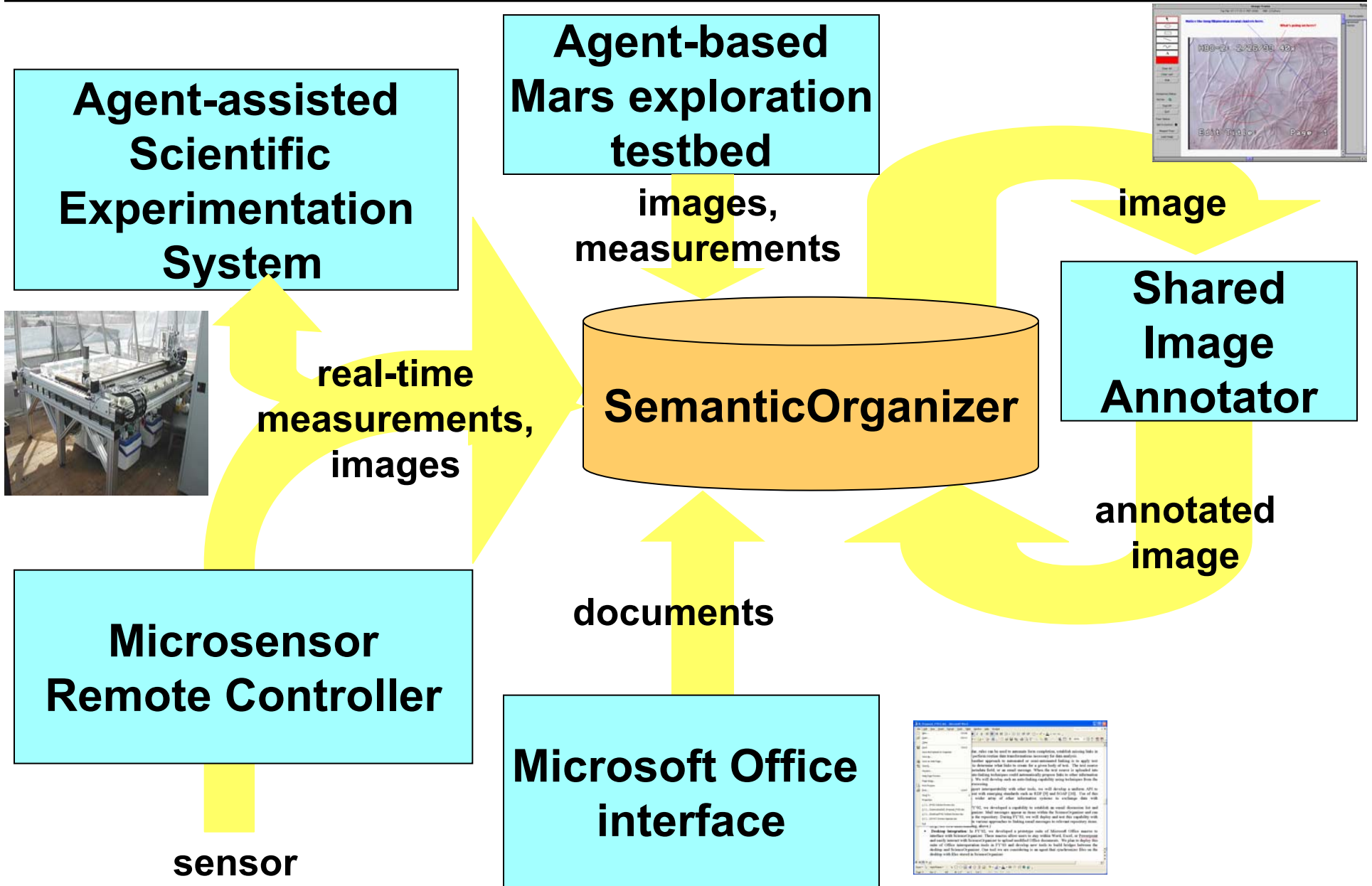
SemanticOrganizer

The interface features a central image frame with the following elements:

- Title Bar:** Image Frame, Tue Mar 07 17:15:11 PST 2000, HBC-2 Culture
- Toolbar:** Includes a mouse cursor icon (highlighted with a red box), an oval, a rectangle, a line, a wavy line, a letter 'A', and a red bar.
- Buttons:** Clear All, Clear Last, Hide, Sign Off, Quit, Request Floor, Load Image.
- Status:** Connection Status: Online (green dot); Floor Status: Not In Control (red dot).
- Image Content:** A microscope image of biological strands with text overlays: "HBC-2: 2/26/99 40x" and "Edit Title: Page 1".
- Annotations:** A blue line points to a cluster of strands with the text "Notice the long filamentous strand clusters here."; a red line points to another cluster with the text "What's going on here?"; a red oval highlights a specific cluster.
- Participants:** A list on the right side showing "hgreenleaf" and "rkeller".



Interoperation w/ External Agents & Systems via API



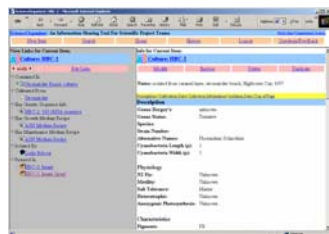


Architecture



Interface Layer

Node Browser Interface

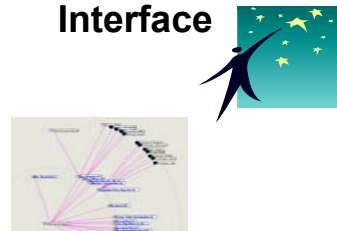


**Structure Editors/
Viewers**

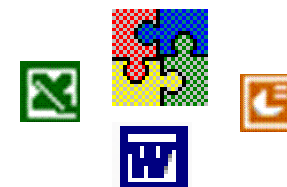


**Radial Network
Browser**

**Programmatic/Agent
Interface**



**Microsoft Office
Macro Integration**



**Representation
& Reasoning
Layer**

Email Ingestor
(email & attachments)

**Semantic
Repository**
(items, links,
attribute values)

Ontology
(item types, link types,
attributes, rules)

RDFS-equivalent

Semantic Annotation

**WordNet
thesaurus**

**Inference
Engine**

**Implementation
Layer**

**Java
servlets**

**Apache-
Tomcat**

**MySQL
Database**

Jess

file system

images
documents
datafiles



Outline



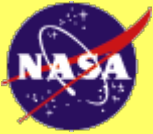
- Introduction & Motivation
- • **SemanticOrganizer system**
- Applications
- Lessons learned



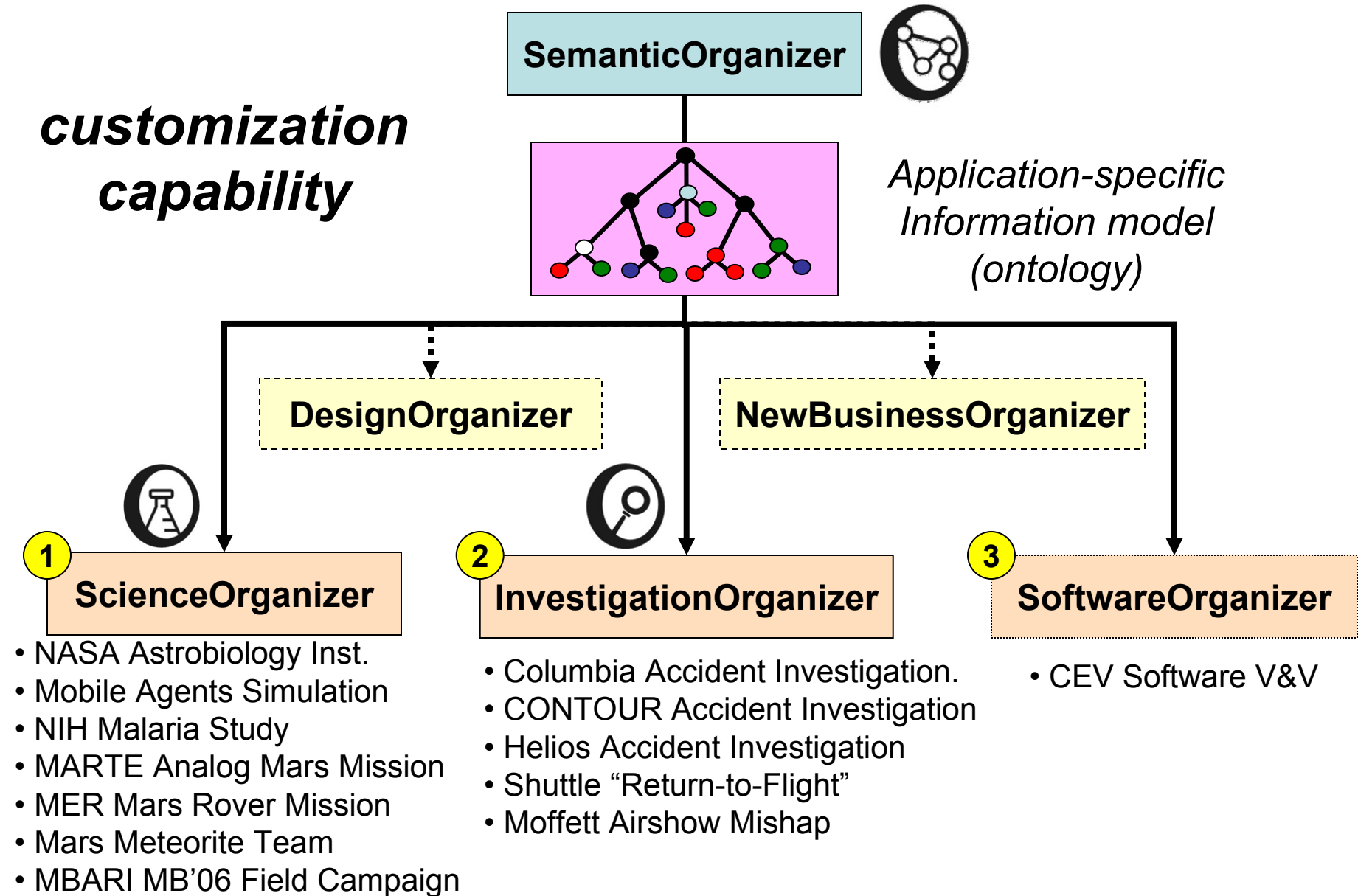
Operational Status



- **System deployed initially in 2001**
- **Over 400 registered individual users from over 50 organizations within NASA at peak**
- **Over 30 projects hosted**
- **Over 45,000 nodes & 160,000 links in repository**
 - **Over 10,000 electronic files stored (documents, image, datasets)**
 - **Over 10,000 archived email messages**



SemanticOrganizer Applications





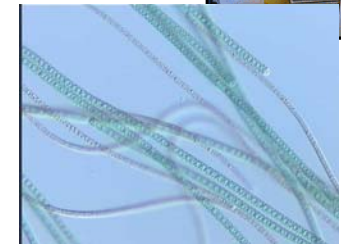
1. Science Organizer Deployments



Astrobiology Field and Lab Science



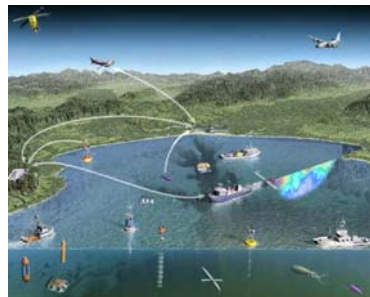
NIH-NASA African Malaria Study



Electron Microscopy Image Archive (Martian Meteorites)



Simulated Mars Surface Exploration



Monterey Bay Oceanographic Campaign



MARTE Mars Analog Drilling Mission



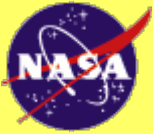
Mars Exploration Rovers Hypothesis Management (prototype)



Activities Supported by ScienceOrganizer



- **Collecting scientific field data**
- **Initiating/monitoring laboratory experiments**
- **Acquiring data from remote instruments**
- **Archiving samples and images**
- **Tracking scientific hypotheses**
- **Writing scientific papers & proposals**
- **Conducting education and outreach**

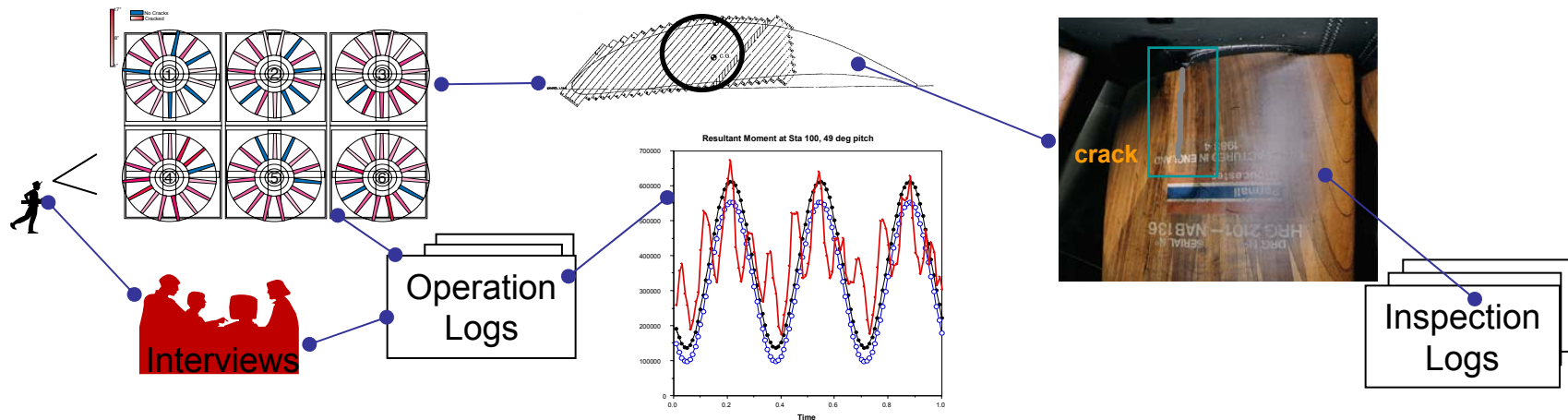


2. InvestigationOrganizer: *Support for Accident Investigation Teams*

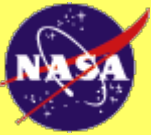


Large amount of data/evidence must be organized/shared rapidly:

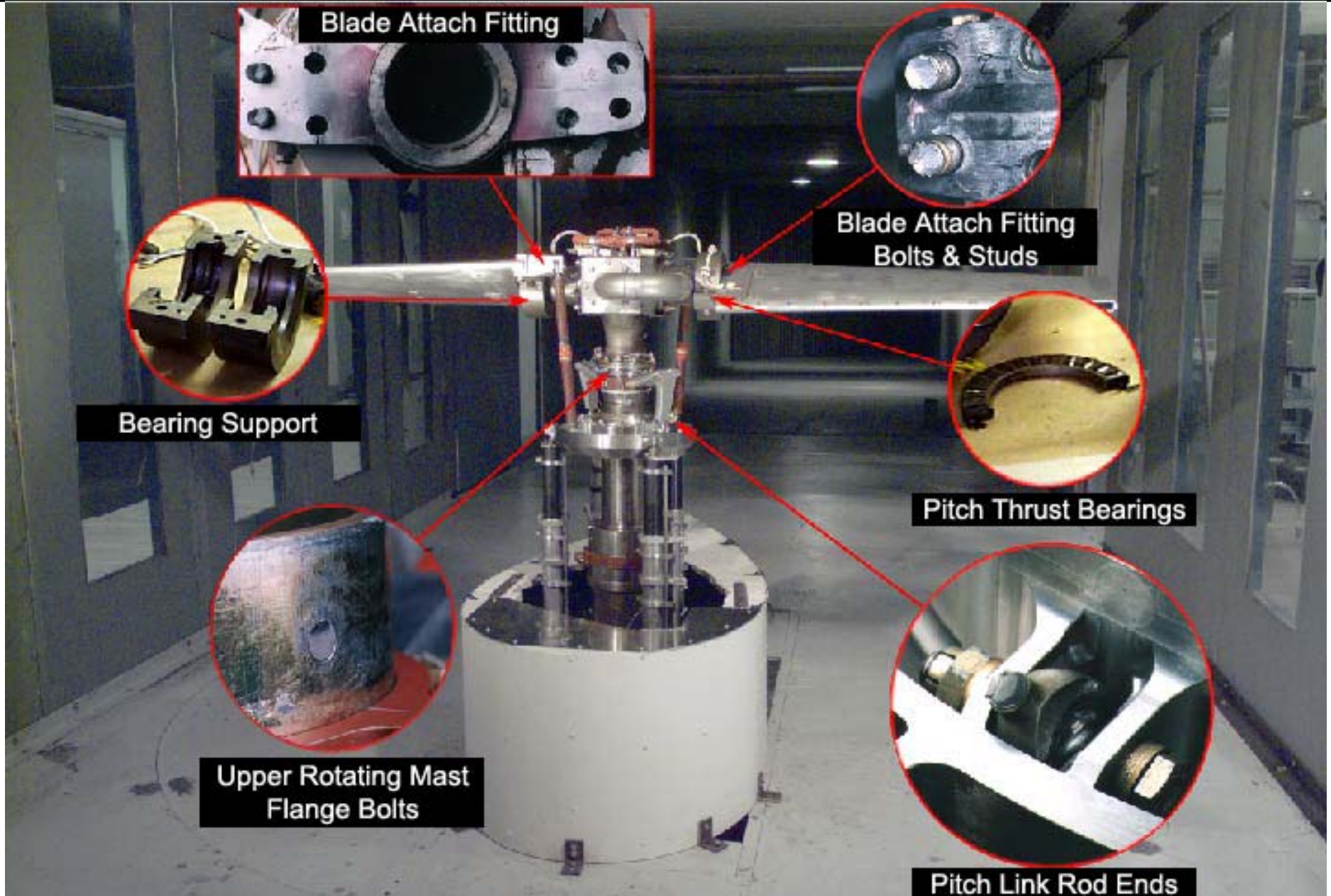
- design specifications
- materials analyses
- procedure manuals
- operations logs
- inspection logs
- defect images
- witness transcripts
- investigation notes
- site maps, images
- lab reports
- engineering drawings
- risk analyses
- test data
- training records



- ***Much of the information is interrelated***
- ***Interrelationships & information context are important!***



Investigation Organizer Example: Helicopter Rotor Wing Mishap Investigation





Semantic Network Fragment from Helicopter Rotor Mishap Investigation



Rotor Wing Mishap Investigation

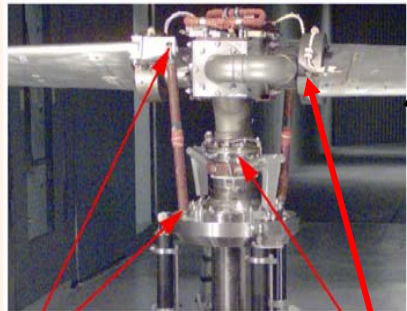
System under investigation:

Has site:

Has hypothesis:

Wind tunnel

Rotor assembly

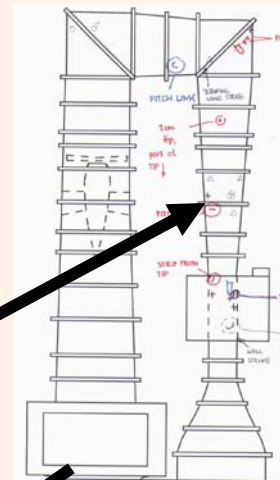


Has design record:

Design schematic

Improper Bolt Tightening

Supported by:



Has subpart:

Rotor pitch link



Found at:

Operated by:



John Smith

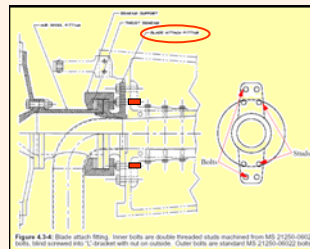
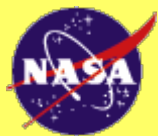
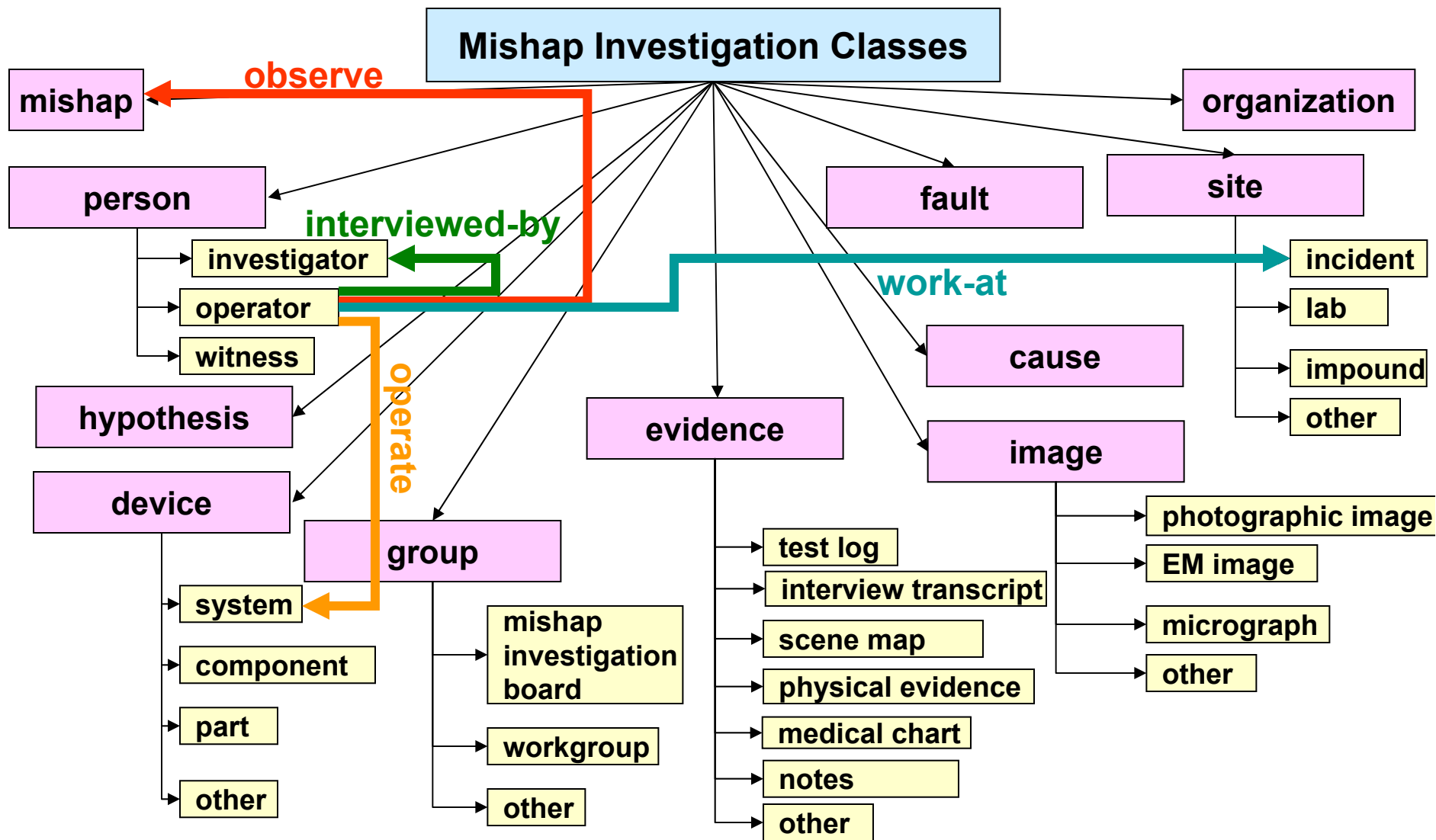
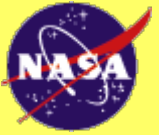


Figure 4.3-4: Blade attach fitting. Inner bolts are double-headed bolts machined from MS 21250-1A022 bolts. Inner bolt is 1/2" diameter with 1/4" hole. Outer bolts are standard MS 21250-1A022 bolts.



Mishap Investigation Ontology (partial)





Investigation Organizer Deployments



CONTOUR Comet-Follower



Shuttle Columbia



Helios Solar-powered UAV



Helicopter Rotor Wing

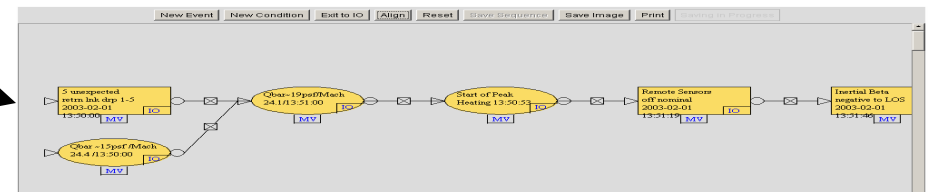
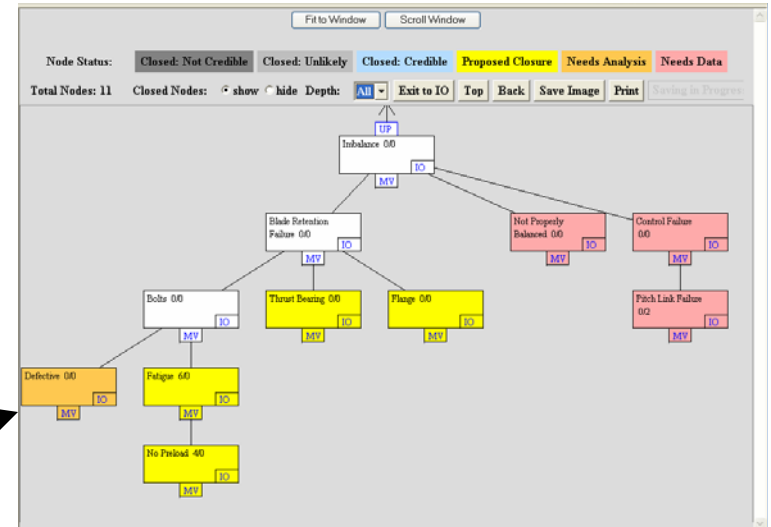




Activities Supported by InvestigationOrganizer



- **Collecting evidence**
- **Tracking hypotheses**
- **Modeling accidents**
 - **fault tree development**
 - **event timelines**
- **Analyzing data**
- **Supporting investigative action-tracking/workflow**
- **Writing reports**



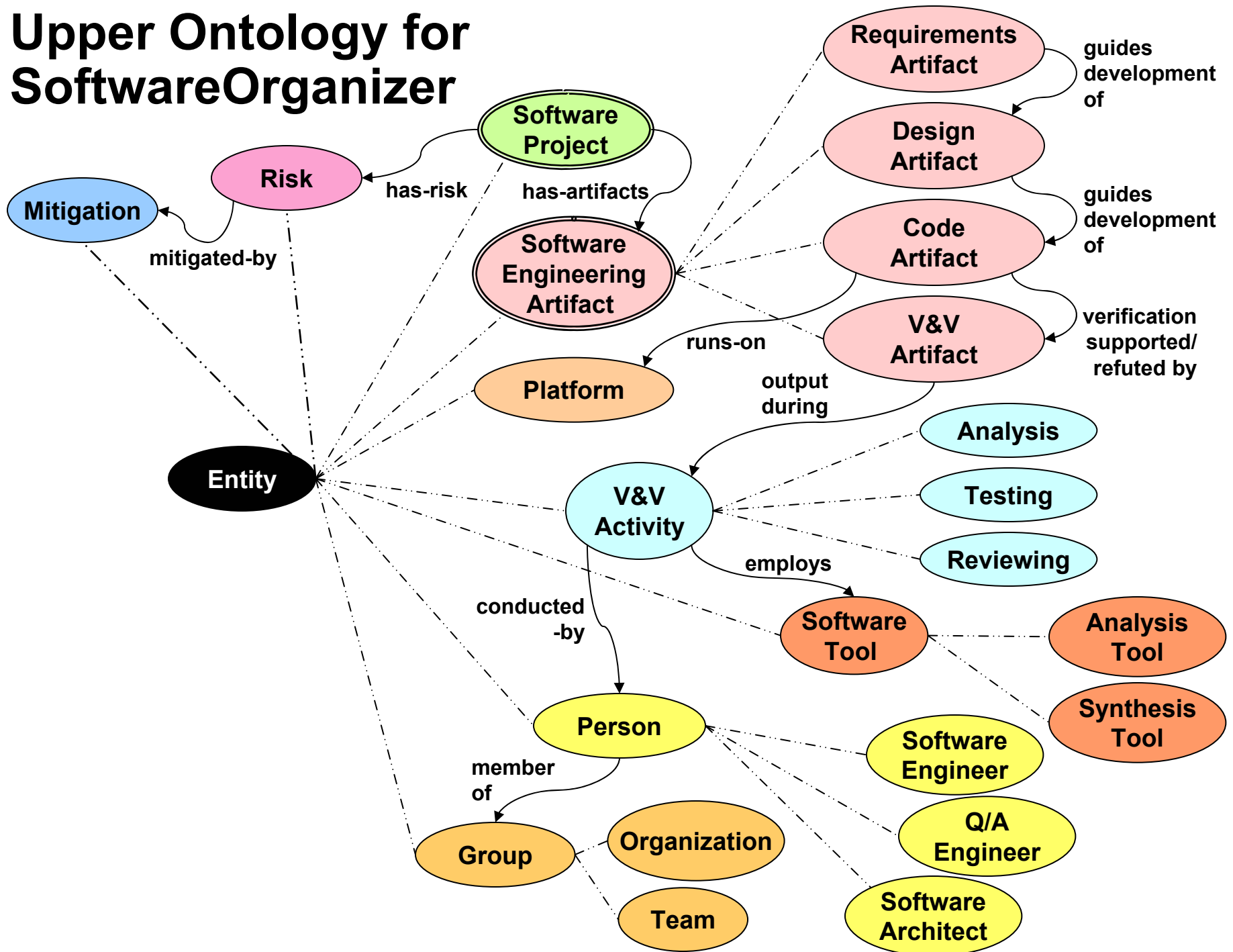


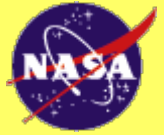
3. SoftwareOrganizer Application (under development)



- **Target Users:** NASA software acceptance, V&V, QA personnel
- **Goals:**
 - Store and interrelate artifacts from throughout the software lifecycle: requirements, design, coding, V&V
 - Support end-to-end requirements traceability
 - Provide an historical record of activities conducted and artifacts produced as part of a NASA software delivery
- **Motivation:**
 - Due to large number of contractors, no single place to examine all software engineering artifacts for a project
 - Existing vendor tool suites capture only a subset of the desired artifacts and are not customizable for NASA needs
 - Closed, vendor-specific solutions not suitable for diverse NASA contracting environment

Upper Ontology for SoftwareOrganizer





Outline



- **Introduction & Motivation**
- **SemanticOrganizer system**
- • **Applications**
- **Lessons learned**



Lessons Learned



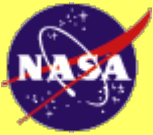
- **Lesson #1:** Network-structured Storage Models Confound Users
- **Lesson #2:** Need to “Loosen” Semantics to Accommodate Different Organizing Styles
- **Lesson #3:** Navigating the Information Space can be Difficult
- **Lesson #4:** Automated Knowledge Extraction is Key



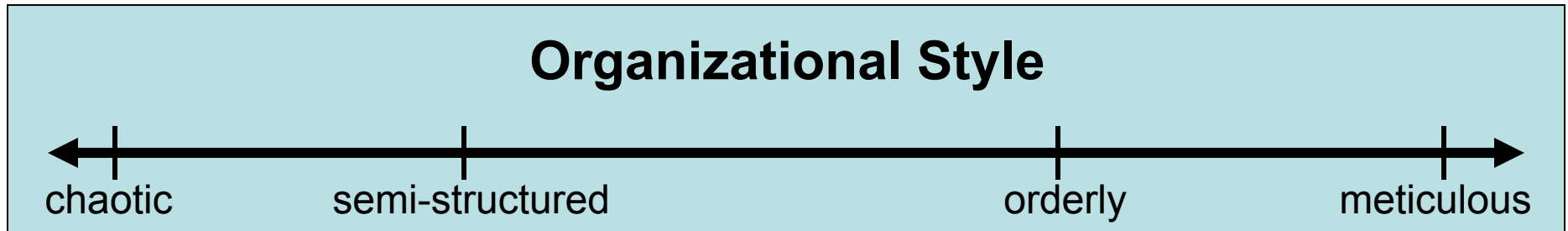
Lesson #1: Network-structured Storage Models Confound Users



- **Despite ubiquity of the Web, users' comfort level with networked storage schemes is low**
- **Networks can be disorienting**
- **Hierarchical storage schemes rule! (e.g., folders, directory structures)**
- **Mitigation strategy: Introduction of typed folders and subfolders into network**
 - **Largely unsuccessful: semantics and type restrictions are unintuitive and counterproductive**



Lesson #2: Need to “Loosen” Semantics to Accommodate Different Organizing Styles



- **Personal Organizational style depends on:**
 - Individual differences and preferences
 - Task demands
- **SemanticOrganizer frustrated:**
 - Casual (and chaotic) organizers
 - Users whose task demands or time available didn't justify the overhead required to provide semantic linkages: cost-benefit proposition wasn't clear
- **Compromises that serve all users are difficult to achieve**



Lesson #3: Navigating the Information Space can be Difficult



- **Information space is very large**
 - Over 45,000 nodes
 - Over 150,000 links
- **Can't display entire information space**
- **Difficult to comprehend the scope and extent of the information space**
- **Difficult to overview and navigate the space**



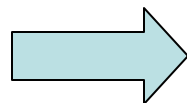
Semantic web visualization research



Lesson #4: Automated Knowledge Extraction is Key



- Relying on purely manual entry and linkage of information is not tenable
- Automated methods of creating and linking information are essential:
 - Inference rules
 - Object and link extraction from text
 - API for communication with agents/processes
 - Integration with document management software (MS Office)
 - Interoperation with stand-alone sources (databases, Web sites)



Semantic annotation research



Key Idea: Model-based Information Integration



- **Use a common information model to “glue” together disparate data and information into a coherent, integrated whole**
- **Use model semantics to check validity of information and improve consistency and completeness of information**
- **Use semantic context to facilitate information access, retrieval, and search**




Questions?



Web Site <http://sciencedesk.arc.nasa.gov>

Paper Keller et al.: “SemanticOrganizer: A Customizable Semantic Repository for Distributed NASA Project Teams”, International Semantic Web Conference, Hiroshima, Japan, November 2004.

Honors  Winner, Best paper award, 2004 International Semantic Web Conference



Finalist in the 2003 NASA Software of the Year Competition and recipient of a Space Act Award for significant technical contribution to NASA.



Winner of a 2004 NASA Turning Goals Into Reality Award for outstanding accomplishment in supporting the Columbia Accident Investigation Board



Finalist in the 2004 Semantic Web Challenge Competition