



THE UNIVERSITY
OF QUEENSLAND
AUSTRALIA

Putting the Shock back into the Future

Prof. Jane Hunter
jane@itee.uq.edu.au

Agenda

- PANIC, AONS – Semi-automated Preservation
- FUSION
 - Provenance Model cf. PREMIS
 - Secure Provenance Visualization
- DART
 - Metadata Schema Registry
 - SRB + Fedora
 - Annotation tools
- Future Issues/Research
 - Interoperability of Preservation Metadata – across heterogeneous archiving systems
 - Collaborative Preservation Tools
 - Trusted repositories – Social networks
 - Preservation selection metrics

PANIC Project

Objectives

- Address the long term preservation and accessibility of (composite) digital objects

Partners

- DSTC, UQ

Challenges

Within digital libraries/archives:

- Wide range of file formats - different platforms, different authoring/display software
- Composite mixed-media objects – web pages, images, video, audio, Flash, SMIL, SVG
- Highly proprietary – software & hardware dependent
- Dynamic and interactive
- Difficult to capture – boundary problem
- Large scale
- Little guidance

Existing Tools

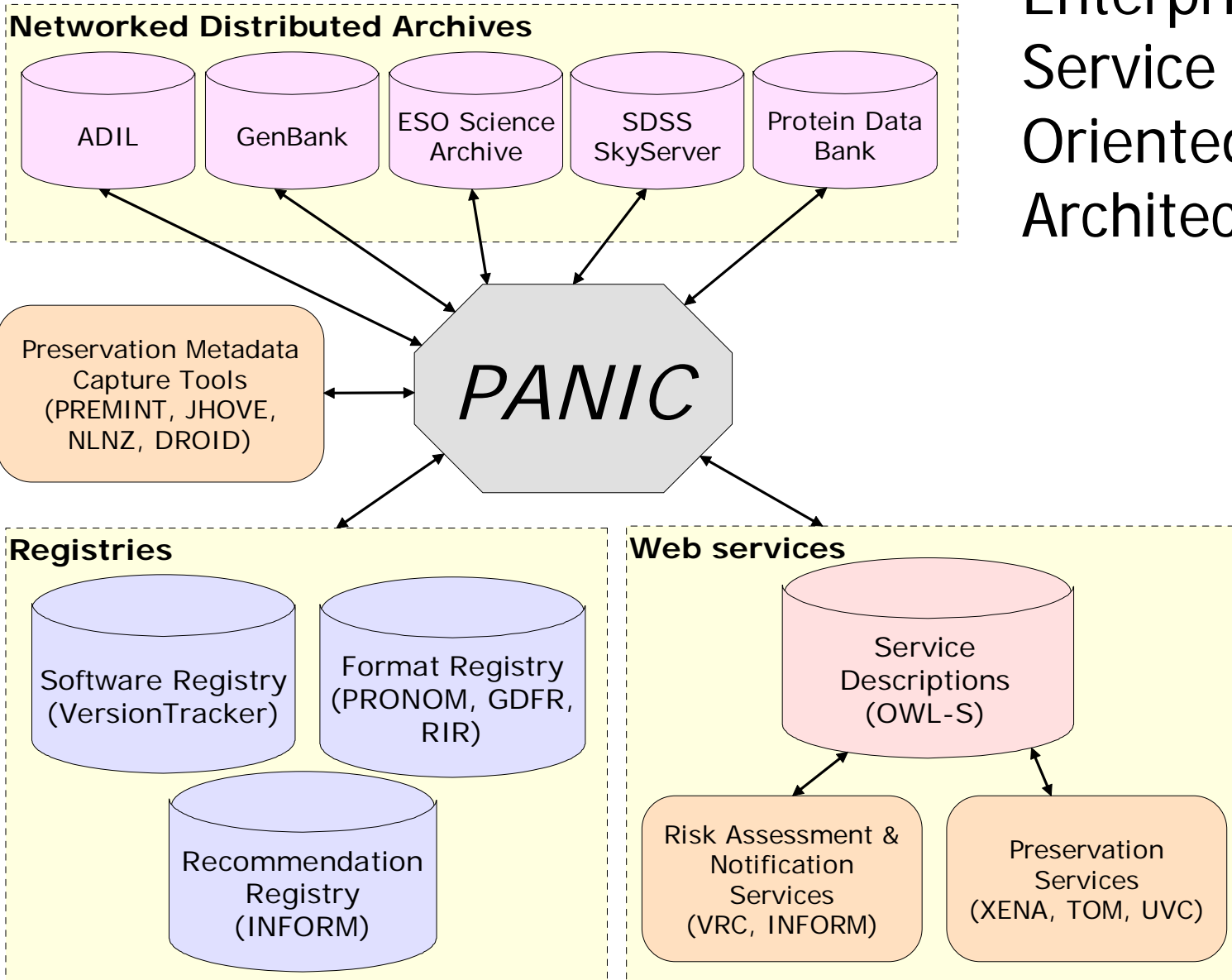
- JHOVE, DROID - Metadata extraction tools
- OCLC's INFORM, Cornell's VRC – risk assessment -> notification services
- GDFR, PRONOM, DCC-RIR – Format registries
- VersionTracker, IIPC – Software Registries
- XENA, TOM – Conversion/migration services
- IBM's UVC (Universal Virtual Computer)
- Koninklijke Bibliotheek - Emulation services

Objectives

Provide an Integrated Preservation Framework which supports:

- Large, heterogeneous, distributed collections
- Multiple formats and composite digital objects
- Changing organizational needs
 - Range of solutions
- Flexible, Dynamic, Scalable, Extensible
- New emerging formats, software, recommendations
- New migration, emulation services
- Recommender services/decision support
- Sustainable - cost-effective, semi-automated
- Collaborative effort!

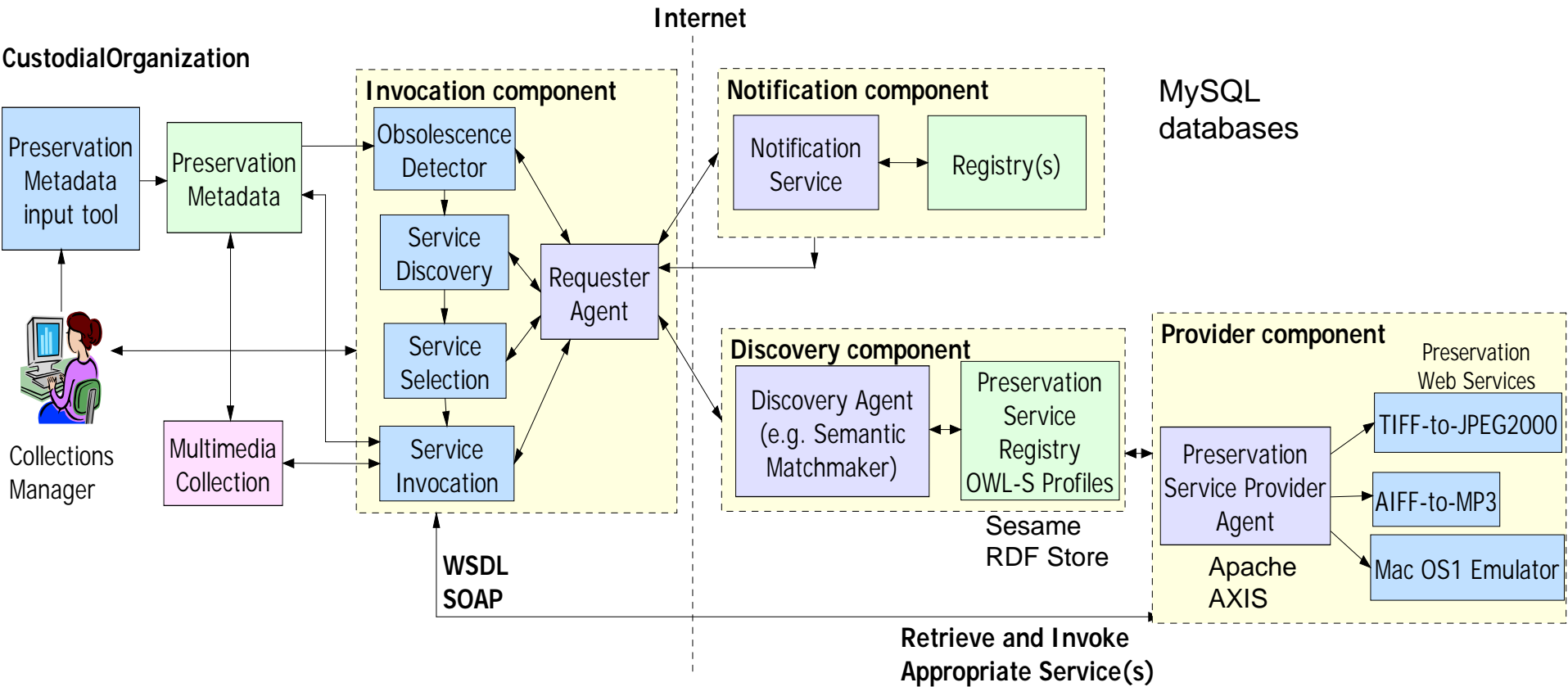
Enterprise Service Oriented Architecture (SOA)



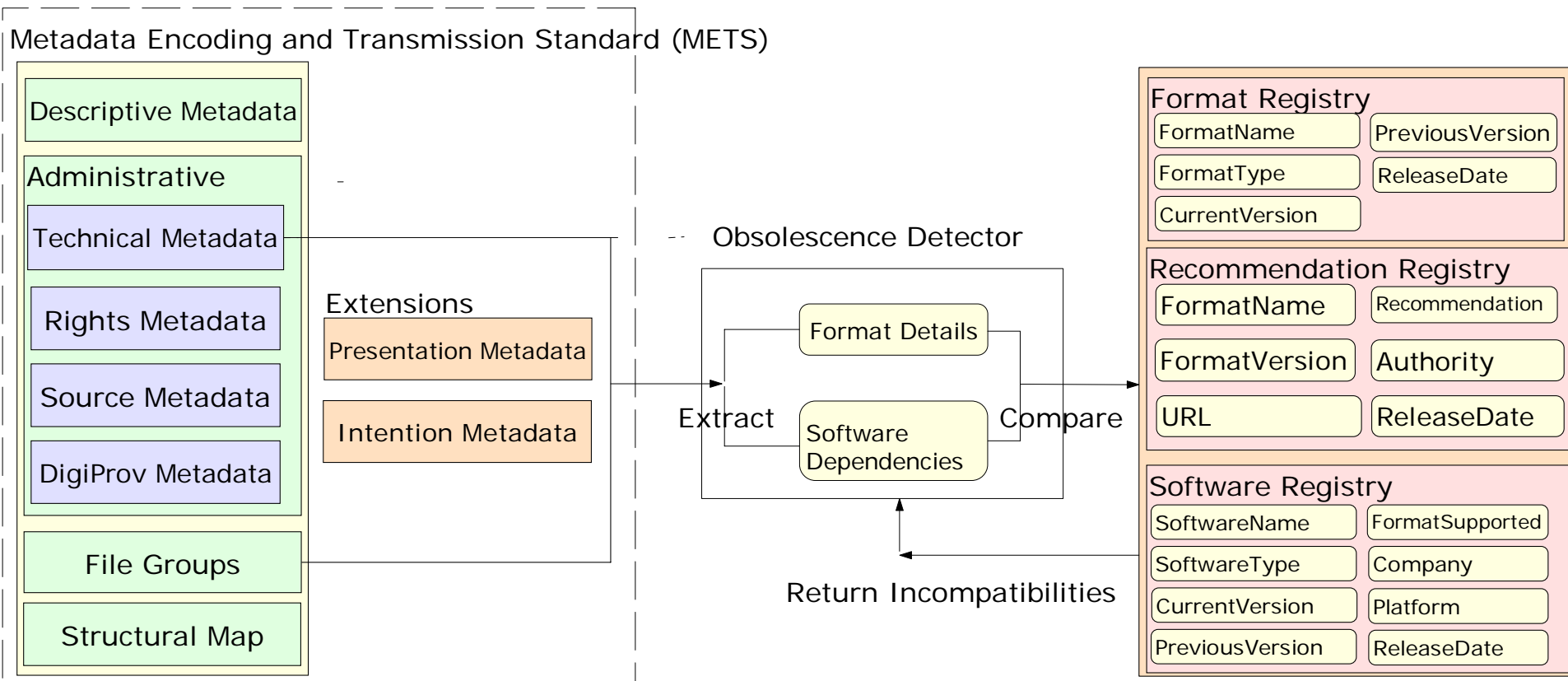
Steps

- Archival – selection and capture of digital object(s) + preservation metadata
- Risk assessment and notification of potential obsolescence
 - New recommendations, formats, software versions
- Service Specification and Request
 - Emulation or Migration
 - Inputs/Outputs
 - Cost
 - Speed
 - Remote/Distributed/Local
 - Reliability
 - Lossiness
- Select, Compose, Invoke Preservation Service
- Record preservation events

PANIC Architecture



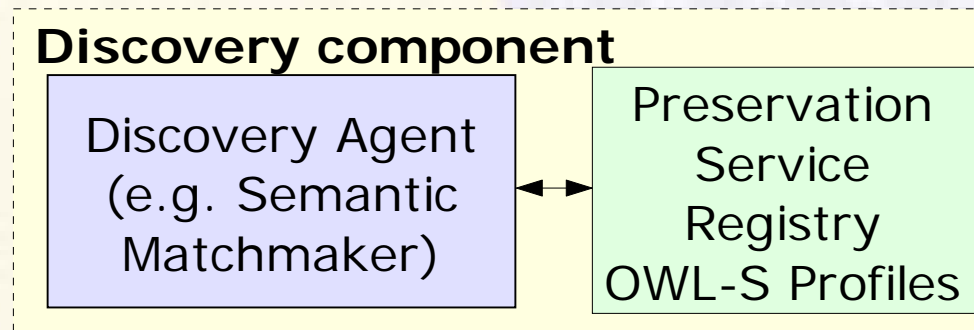
Notification component



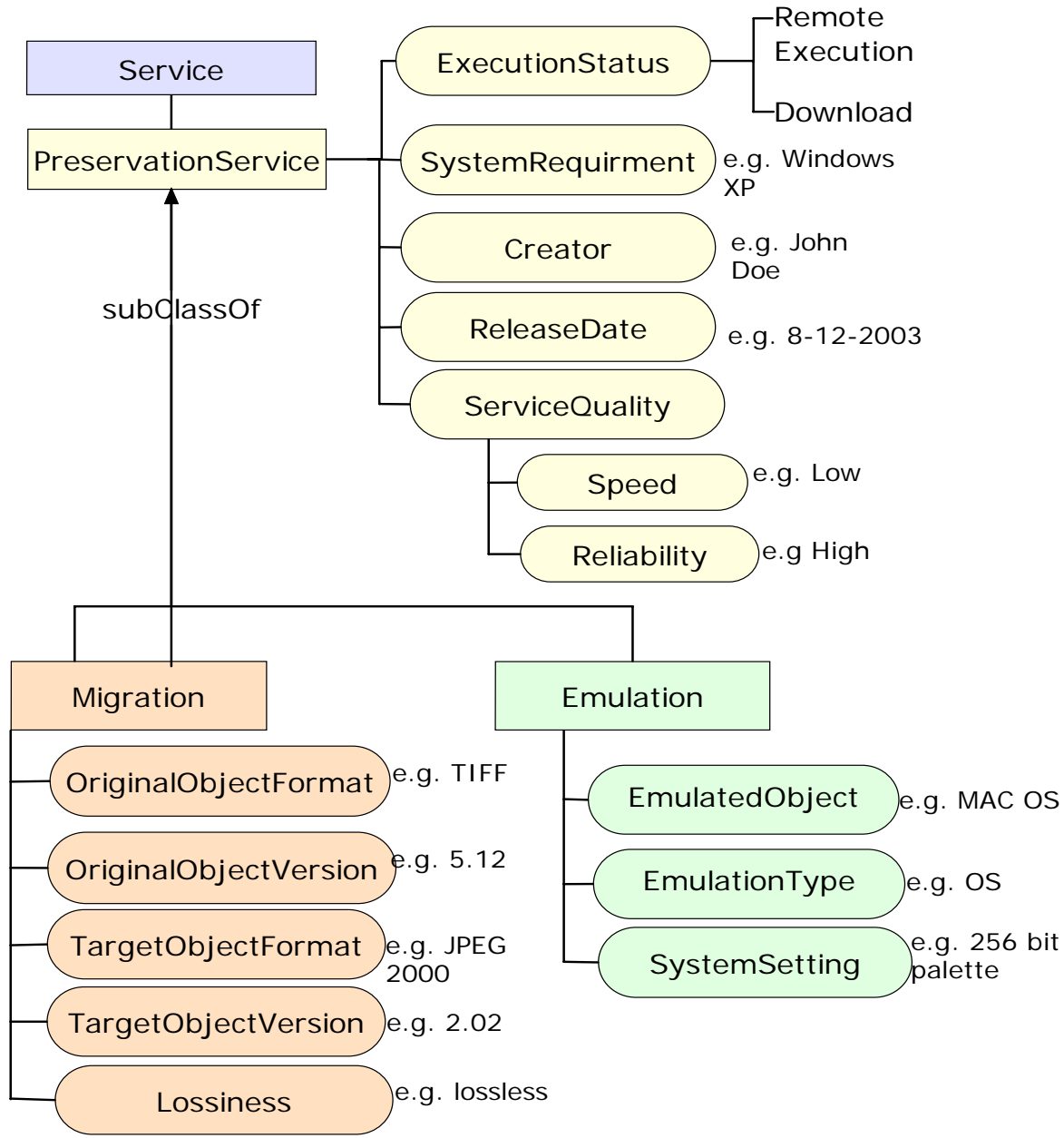
Obsolescence detector – periodically compares the preservation metadata for each object with registries to determine when object is at risk of obsolescence

Discovery component

- Discovery Agent - matches service request against OWL-S descriptions of Preservation Web services
- Returns a ranked list of Preservation Web services that match the request



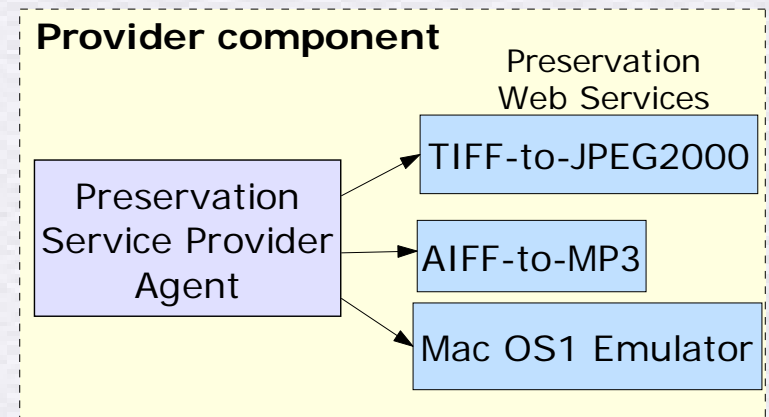
OWL-S Preservation Extensions



Provider component

Provider Agent either:

- retrieves and invokes preservation service locally or;
- Invokes preservation service remotely

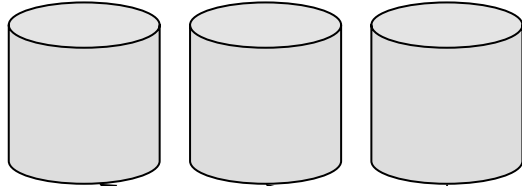


AONS

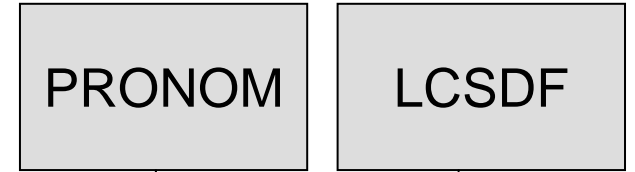
- Automated Obsolescence Notification Service
- APSR (Aust. Partnership for Sustainable Repositories) funded Project
- Collaboration between
 - University of Qld
 - ANU (Peter Raftos, Joseph Curtis)
 - NLA

AONS Architecture

Digital Collections

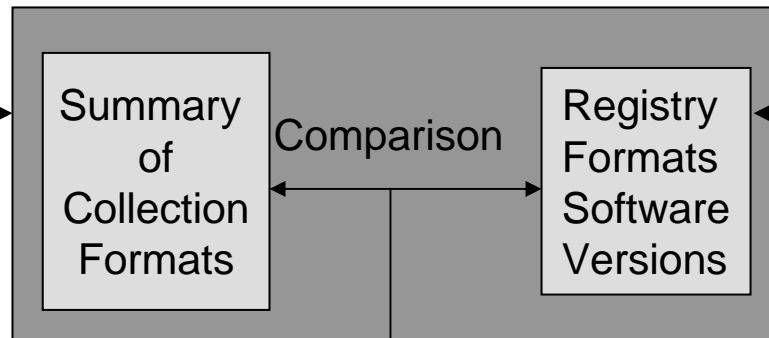


DSpace ANU
Fedora UQ
Pandora



(GDFR)

AONS



Registered Collections Manager,
Owner or Consumers of data?

Future of AONS

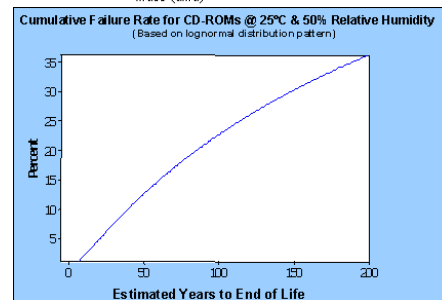
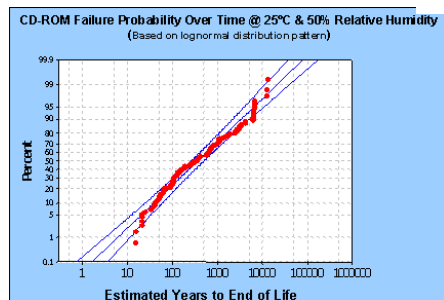
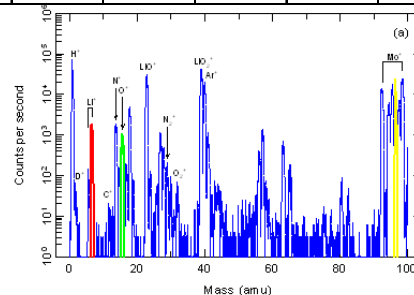
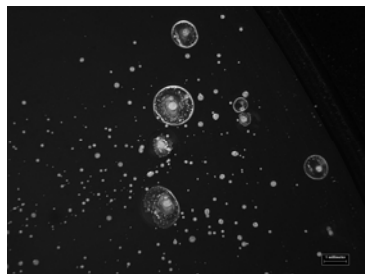
- Implement over Fez at UQ... subset of PANDORA
- Build a GUI
- Test, evaluate and refine
- Investigate release as open source middleware
- Integrate - GDFR, risk assessments/rankings
- Provide access to trusted services - quality ratings
- OWL-S versus WSMO
- Grid Services - Web Services Resource Framework (WSRF)
- Composite services, composite objects

Example – Scientific Models

Slattery, O., Lu, R., Zheng, J., Byers, F., Tang, X.

"Stability Comparison of Recordable Optical Discs- A study of error rates in harsh conditions," Journal of Research of the NIST, 109, 517-524, 2004

	Area	Mean	S.D.	X	Y	Model	Length	Major	Minor	Angle	Int.Der	Back.	Min	Max
1	0.01	208.2	88.14	0.34	0.06	253	0.34	0.11	0.08	102.7	0	0	35	253
2	0.01	206.8	89.14	0.17	0.07	253	0.34	0.1	0.08	17.57	0	0	35	253
3	0.01	212.9	84.54	0.26	0.11	253	0.37	0.11	0.1	158	0	0	35	253
4	0	190.4	98.85	0.07	0.1	253	0.21	0.07	0.05	76.53	0	0	35	253
5	0.03	228.8	68.54	0.67	0.38	253	0.75	0.24	0.15	154.8	0	0	35	253
6	0.09	240.7	50.36	0.34	0.48	253	1.24	0.38	0.3	95.89	0	0	35	253
7	0.08	240.1	51.46	0.59	0.59	253	1.18	0.35	0.28	81.38	0	0	35	253



Derived_from

Each component
Has software,
OS,
hardware
dependencies

$$\text{Average LE} = 1/T \exp -(A - B/T)$$

Preservation of Composite Objects

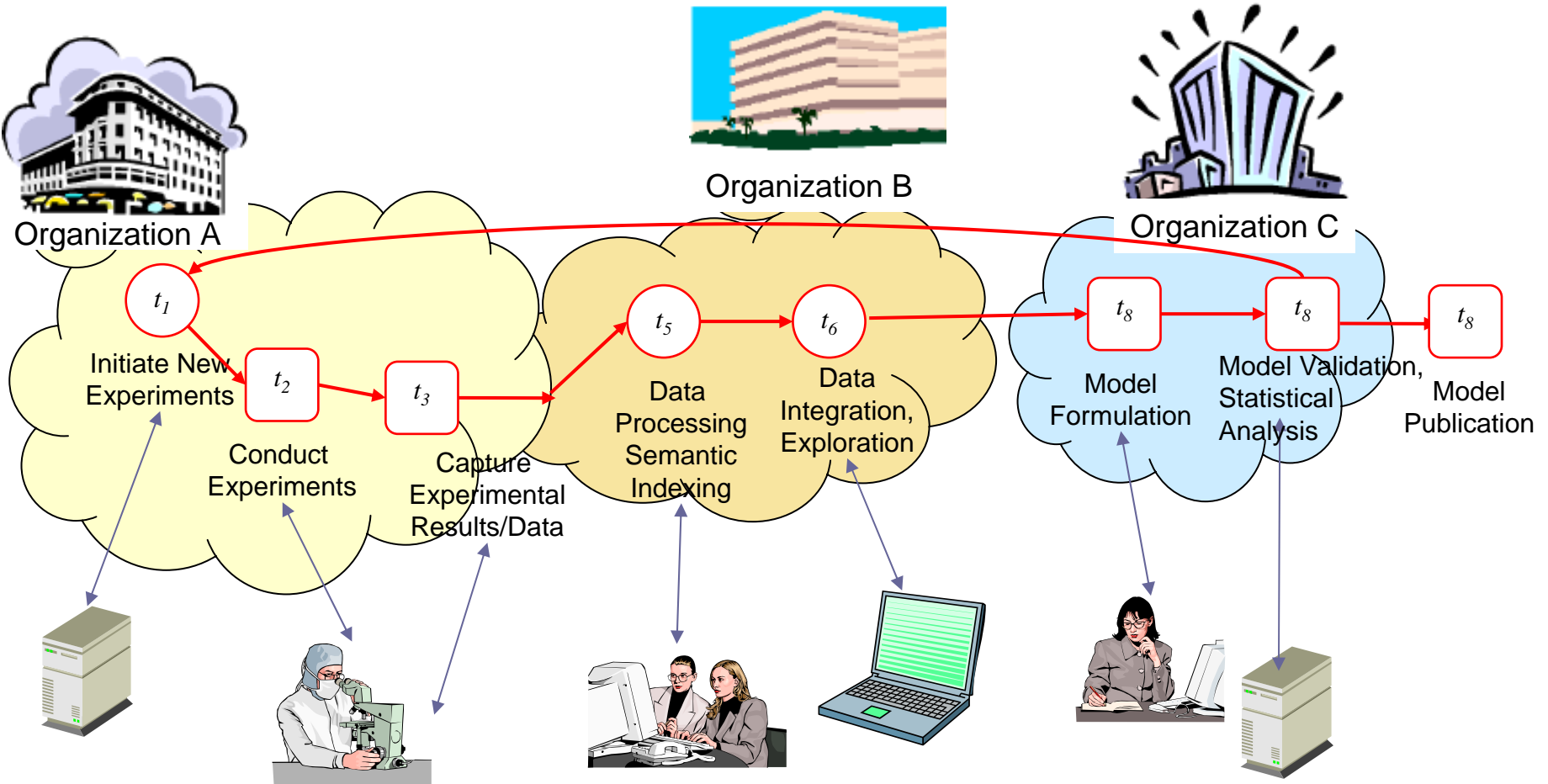
- Use XML to package metadata, component objects and relationships
 - METS, MPEG-21 DIDL, XFDU, IMS-CP
- Maintain preservation metadata for both
 - Composite object
 - Atomic components
- Maintain index of file formats
- Monitor atomic objects first
 - JPEG -> JPEG-2000
- Then check currency of composite objects
 - SMIL 1.0 -> SMIL 2.0

Scientific Publishing

Increasing pressure to:

- publish raw and derivative data
- document precise provenance
- share data and analytical, modelling services
- enable duplication and validation
- protect IP

eScience Workflow



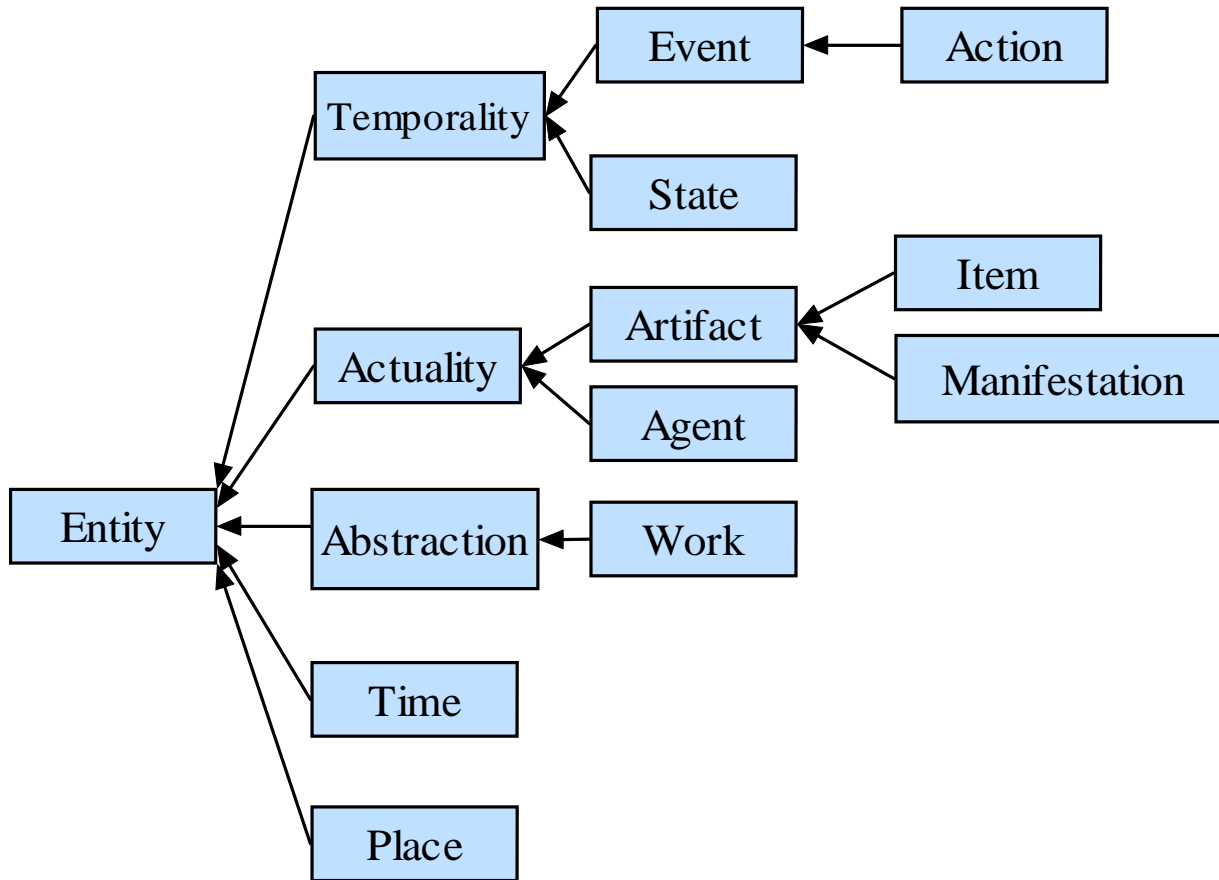
Kepler

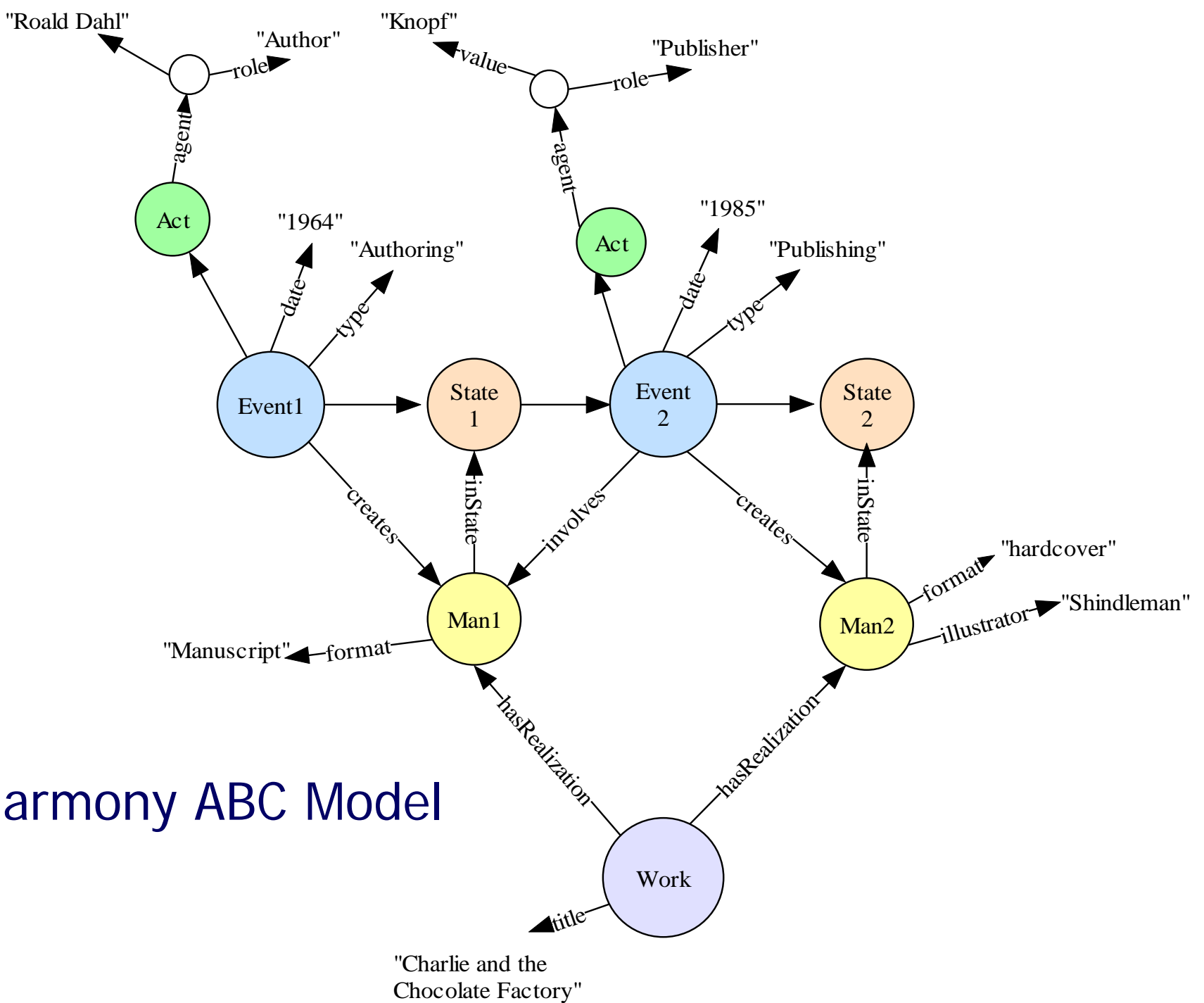
BPEL4WS – workflow based on web services

Components

- Prior work - pre-existing data, models, publications;
- Experimental, observational data
 - numerical data, survey data, images, video, audio, maps, spectral data, real-time sensor data;
 - instrumental conditions, settings and parametric ranges
- Formulae, rules, hypotheses;
- Conceptual models - axioms, models and metaphors;
- Numerical models – mathematical functions;
- Software - source code, executables, applets, web services
 - Analysis, processing, transformation services
 - Computational models – simulation software
- Hardware – instruments and computers;
- Visualizations – 2D, 3D imagery, graphs, tables, charts, diagrams, animations;
- Textual - publications, reports, documentation, annotations, bibliographies, reviews

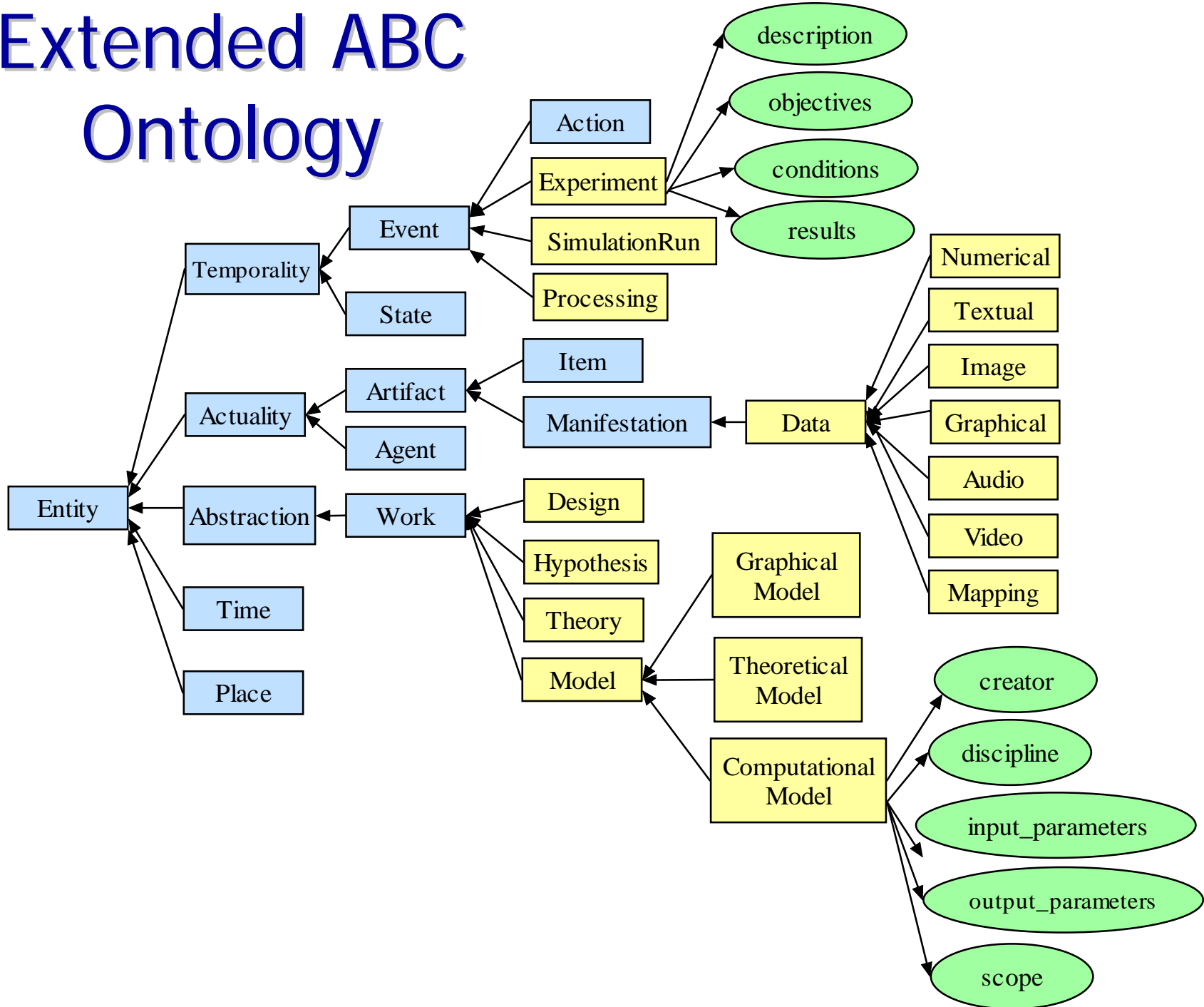
Harmony ABC Model



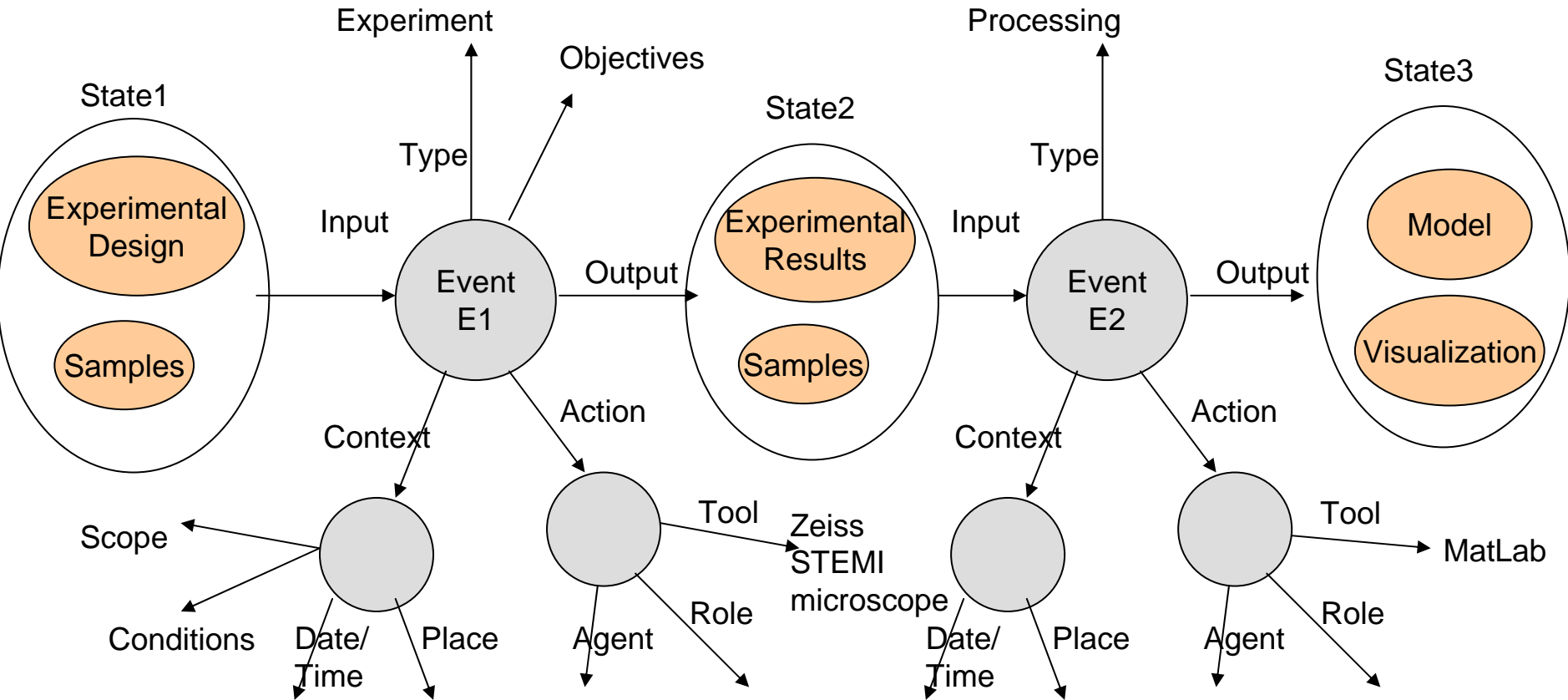


Harmony ABC Model

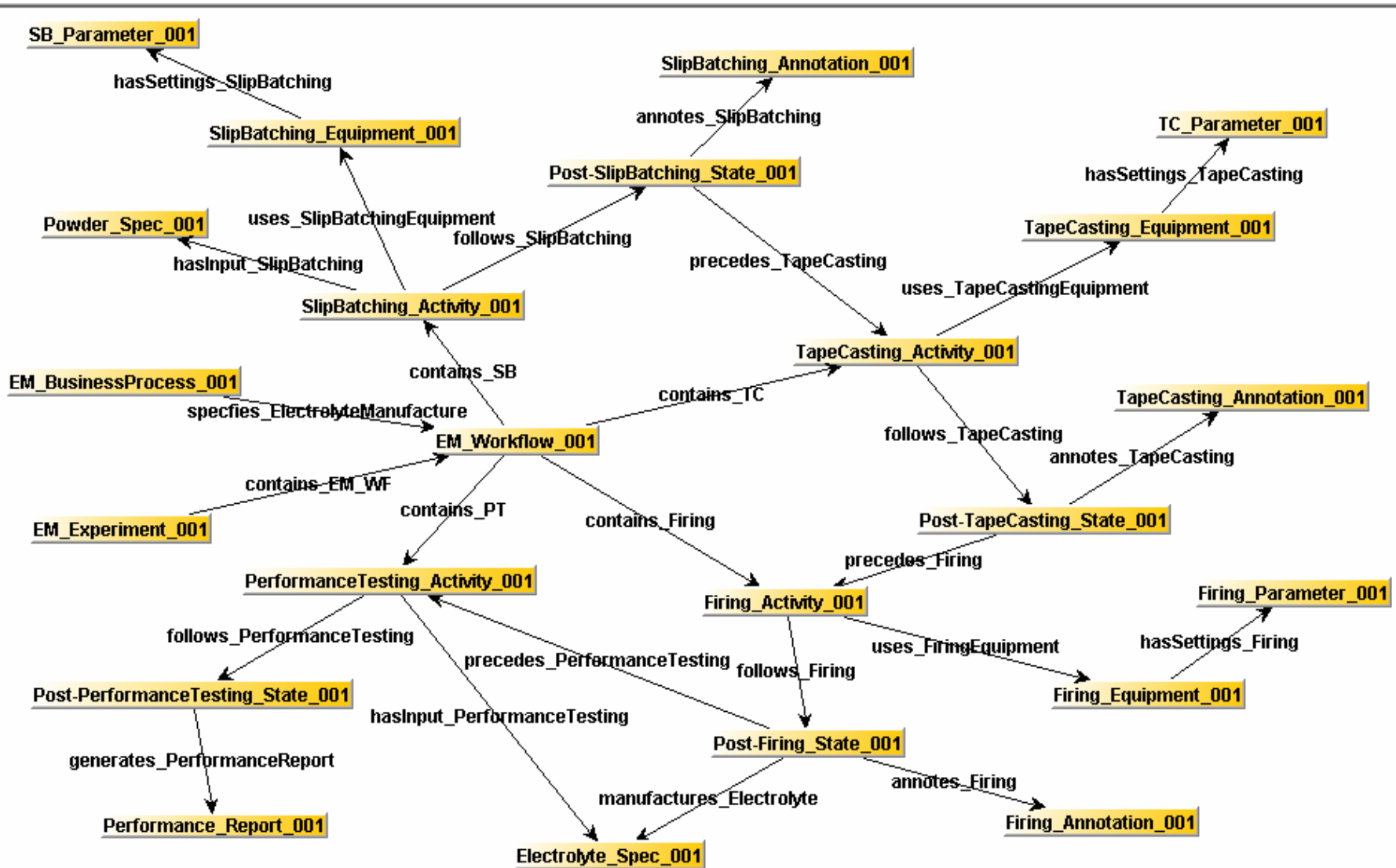
Extended ABC Ontology



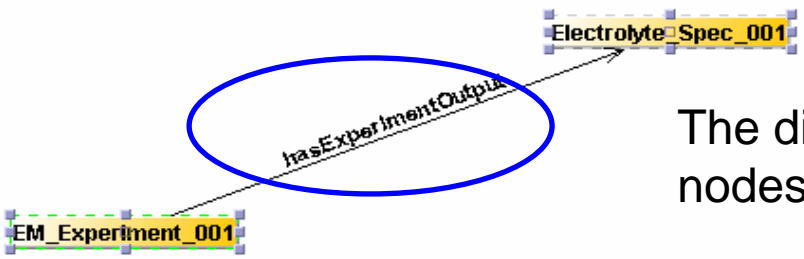
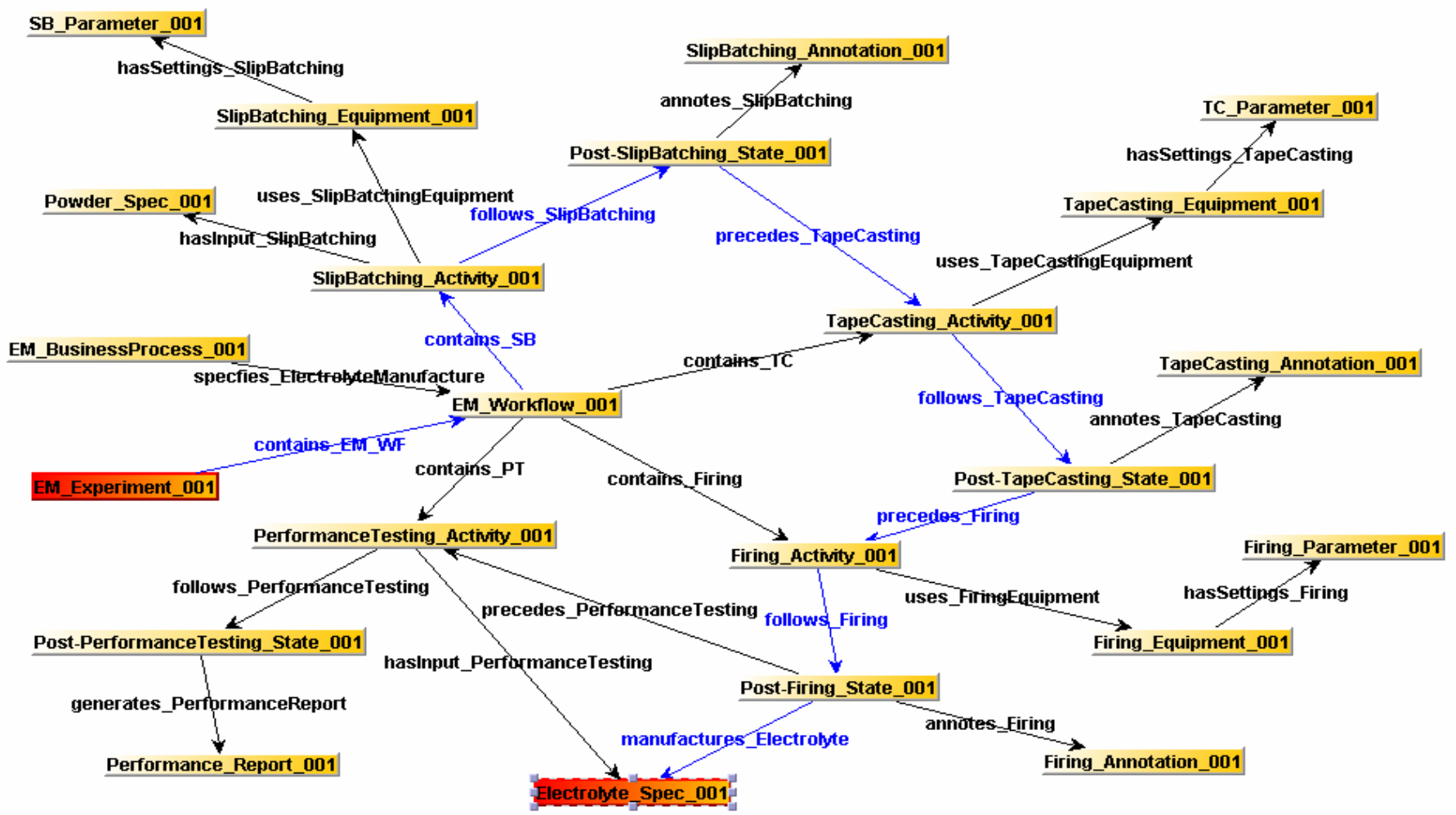
Modelling eScience Provenance



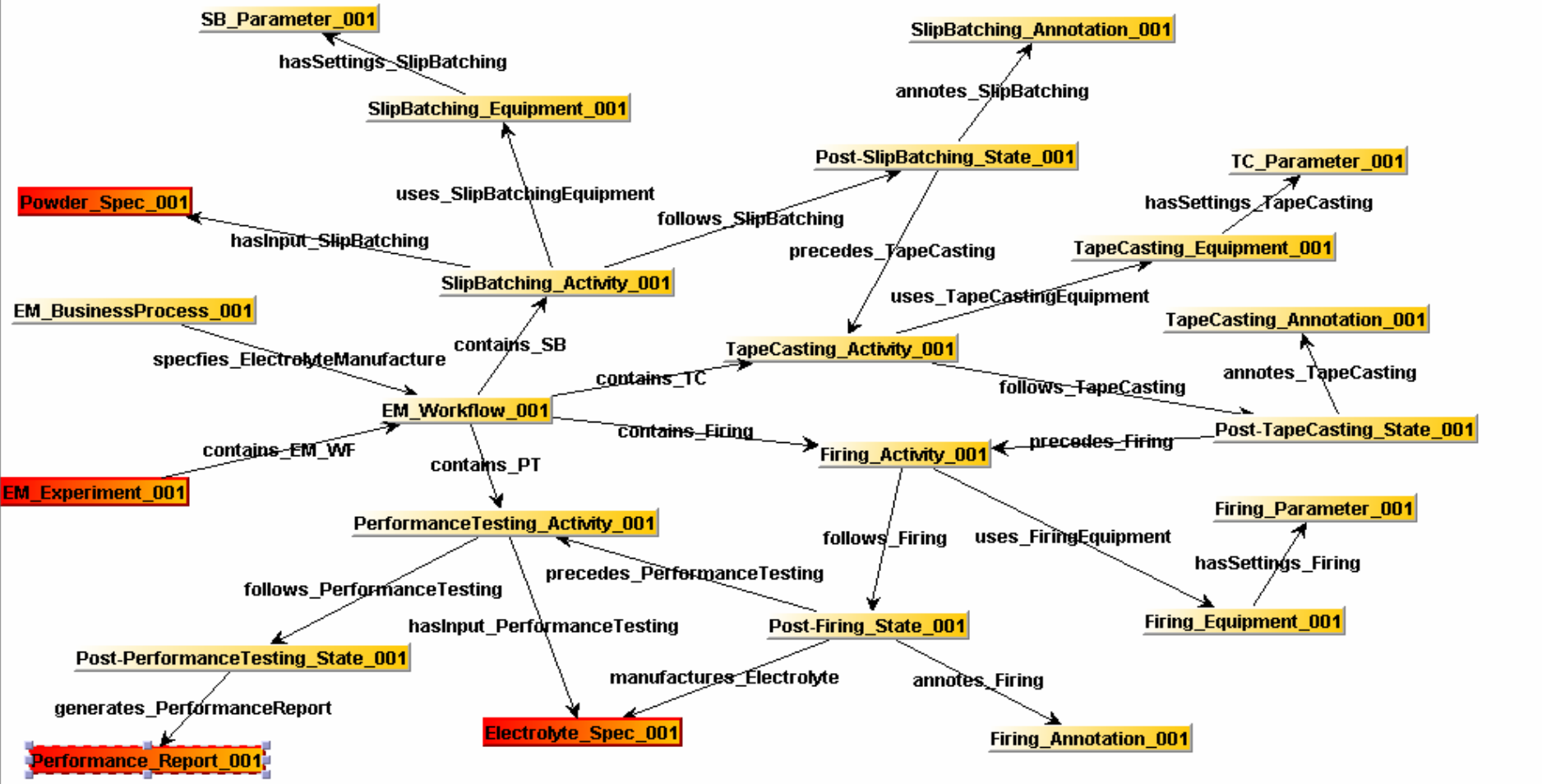
Agents can be people, instruments or software e.g., web services



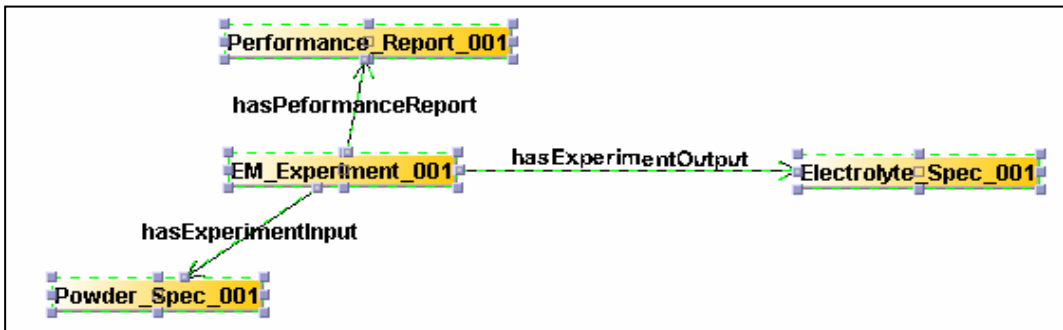
A Visualisation of an Electrolyte Manufacture experiment.



The direct relationship between selected nodes is automatically inferred.

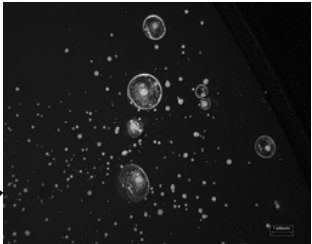


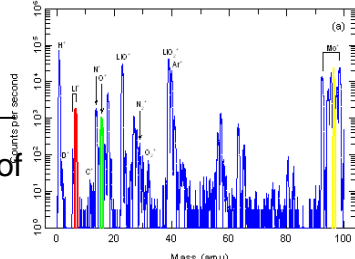
Coarse-grained view of provenance



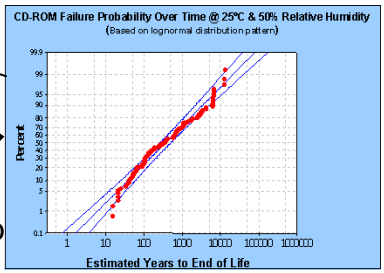
Example of a Scientific Model Package

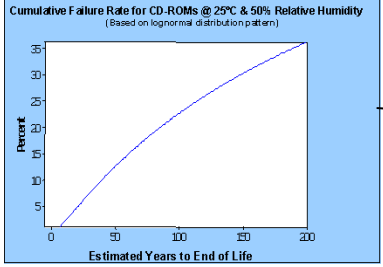
	Area	Mean	S.D.	X	Y	Mode	Length	Major	Minor	Angle	Int.Der	Back.	Min	Max
1	0.01	208.2	88.14	0.34	0.06	253	0.34	0.11	0.08	102.7	0	0	35	253
2	0.01	206.8	89.14	0.17	0.07	253	0.34	0.1	0.08	17.57	0	0	35	253
3	0.01	212.9	84.54	0.26	0.11	253	0.37	0.11	0.1	158	0	0	35	253
4	0	190.4	98.85	0.07	0.1	253	0.21	0.07	0.05	76.53	0	0	35	253
5	0.03	228.8	68.54	0.67	0.38	253	0.75	0.24	0.15	154.8	0	0	35	253
6	0.09	240.7	50.36	0.34	0.48	253	1.24	0.38	0.3	95.89	0	0	35	253
7	0.08	240.1	51.46	0.59	0.59	253	1.18	0.35	0.28	81.38	0	0	35	253





Average LE = $1/T \exp -(A - B/T)$



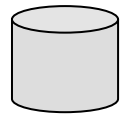


Drennan, J., Knibbe R., Auchterlonie, G., "Effect of porosity and firing Temperature on fuel cell efficiency", Journal of Solid State Ionics, Vol 8, No 2, 517-524, 2004

RDF Package

- Title
- Creator
- Description
- Type
- Discipline
- Date.Published
- License

ePrints database



refers_to

image_of

graph_of

analysis_of

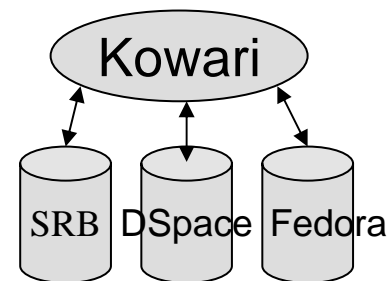
derived_from

refers_to

Each component has software, OS, hardware dependencies + interdependencies

Required Tools

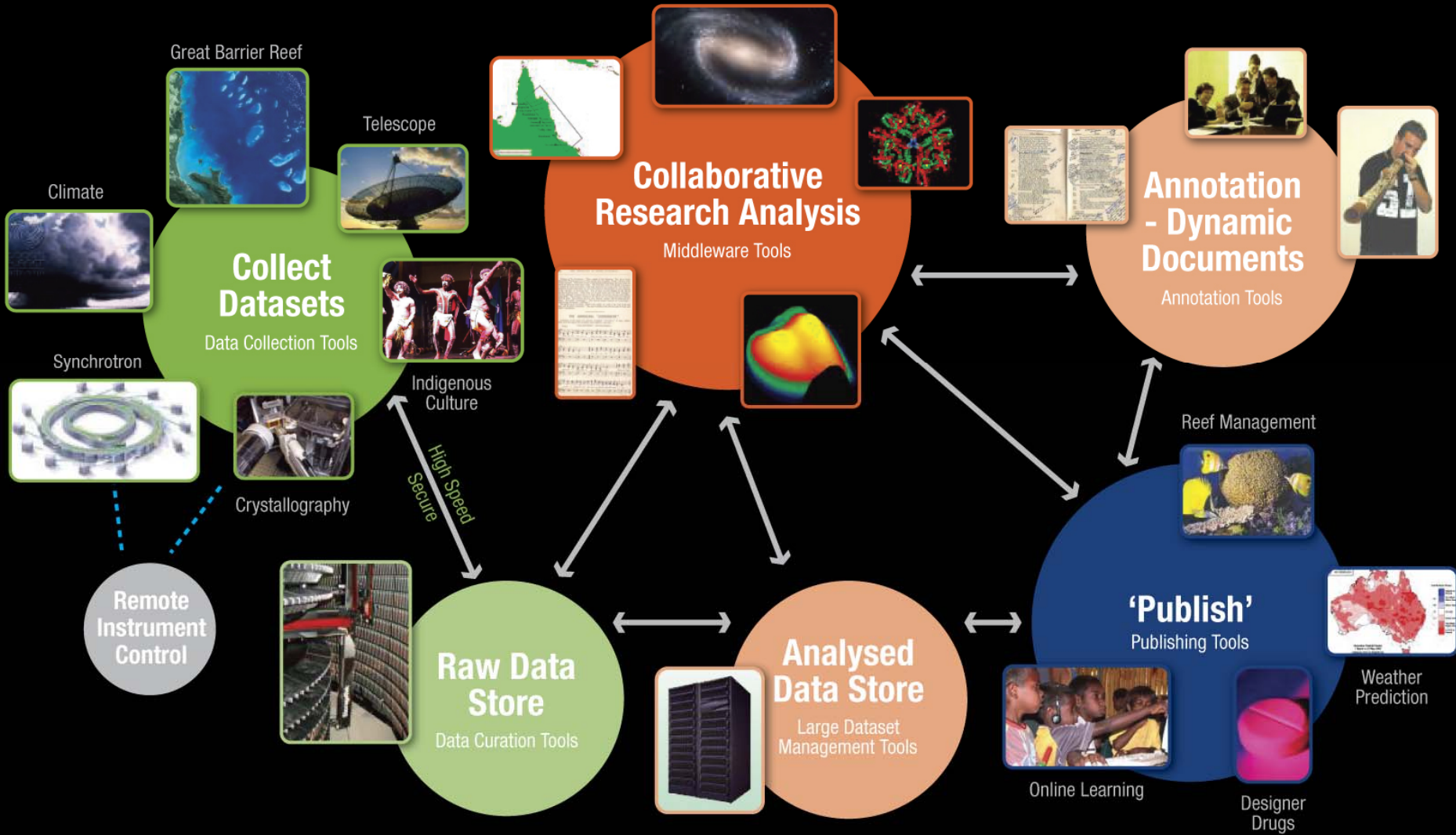
- Scientific Model/Publication construction tools
 - Drag and drop tools – hyperlinks and bitstreams
 - Metadata generation/capture tools
 - Science Commons license attachment
 - Database ingestion – to institutional repository
 - RDF Datastore
 - Kowari + links to SRB, DSpace, Fedora
- Search, Browse and Retrieval
 - RDFQL
 - Jgraph, Haystack – Relationship graphs



DART

Dataset Acquisition, Accessibility & Annotation eResearch Technologies

- **\$3.23M – DEST ARIIC funding**
- **3 partners (Monash (PI), UQ, JCU)**
- **15 months -> Dec 2006**
- **28 Separate work packages**
 - Data Collection, Monitoring and Quality Assurance (DMQ)
 - Storage and Interoperability (SI)
 - Content and Rights (CR)
 - Annotation and Assessment (AA)
 - Discovery and Access (DA)



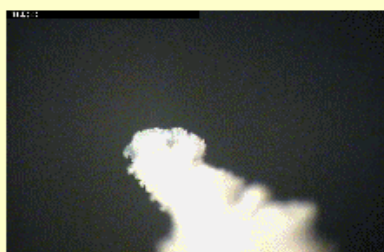
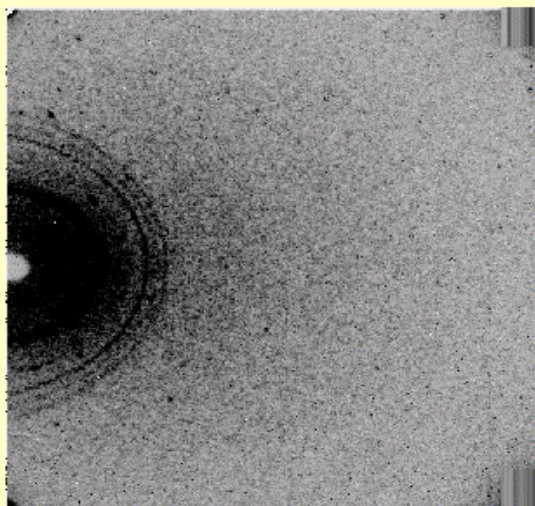
UQ Workpackages

- DMQ4 – Online remote access to instruments/sensors
- AA1 – Annotations of scientific data
- AA2 – Secure annotation server
- AA3 – Collaborative annotations
- SI1 – Integration of Fedora and SRB
- SI3 – Semantic search interface on SRB
- CR2 – Creative Commons – licensing and enhanced search engine
- CR3 – Science Commons tools (SHERPA/ROMEO)
- DA3 – Metadata Schema Registry



Bruker SMART6000 CCD

IUMSC Bay 1 - IUMSC:Bruker:06417



Time	2006-03-06 02:20:24 (UTC)
Current data frame	064175.200
Crystal temperature (C)	-145.5
Instrument enclosure humidity (%)	15.6
Instrument enclosure temperature (C)	25.9
Instrument bay temperature (C)	19.7
Instrument bay humidity	25.9
Crystal image	camera2_2006-03-06_02:20:52Z.jpg
Time for crystal image	(UTC)
Lab image	camera1_2006-03-06_02:20:52Z.jpg
Time for lab image	(UTC)
X-ray coolant water in (C)	17.9
X-ray coolant water out (C)	24.2
CCD Chip Temperature (C)	-54.86
Frame #	200

Start

<<<100

<<10

<1

>1

>>10

>>>100

List Annotations Policies


- [-] Review ISWC
 - ... Evaluation created by jane on 2006-03-06T15:42:56Z
- [-] Review ISWC
 - ... Comment created by imrank on 2006-03-06T15:41:26Z
- [-] Information Regarding Income inequality hypothesis
 - ... Question created by michael on 2006-03-06T15:44:47Z
 - [-] Re: Information Regarding Income inequality hypothesis
 - ... Reply created by suzanne on 2006-03-06T15:46:02Z
 - [-] Re: Information Regarding Income inequality hypothesis
 - ... Reply created by ronalds on 2006-03-06T15:45:24Z
 - [-] Re: Re: Information Regarding Income inequality hypothesis
 - ... Reply created by michael on 2006-03-06T15:46:42Z
- [-] Review ISWC
 - ... Comment created by imrank on 2006-03-06T15:40:07Z
- [-] Comment about ePrints
 - ... Comment created by michael on 2006-03-06T15:47:56Z

Body Policy Other Info

POLICIES

- [-] DART Group (unknown)
 - [-] GROUPS
 - [-] uq_members
 - [-] RULES
 - ... READ is Permitted
 - ... LIST is Permitted
 - ... READPOLICY is Permitted
 - [-] monash_members
 - [-] RULES
 - [-] jcu_members

Attribute	Function	Value	Issuer
eduPersonOrgDN	string-equal	itee	uq.edu.au
eduPersonOrgUnitDN	string-equal	dke	uq.edu.au
eduPersonAffiliation	string-equal	staff	uq.edu.au


THE UNIVERSITY OF QUEENSLAND AUSTRALIA
UQ Library

UQ HOME SEARCH CONTACTS STUDY NEWS EVENTS MAPS LIBRARY

ePrintsUQ


Home About Browse Search Register User Area Help

Income Inequality and Health: A Multi-Country Analysis


Asafu-Adjaye, John (2004) Income Inequality and Health: A Multi-Country Analysis. *International Journal of Social Economics* 31(1/2): 195-207.

Full text available as:
[PDF](#) - Requires [Adobe Acrobat Reader](#) or other PDF viewer.

Abstract

This  paper investigates the effect of income inequality on health status. A model of health status was specified in which the main variables were income level, income inequality, the level of savings and the level of education. The model was estimated using a panel data set for 44 countries covering six time periods. The results indicate that income inequality (measured by the Gini coefficient) has a significant effect on health status when we control for the levels of income, savings and education. The relationship is consistent regardless of the specification of health status and income. Thus, the study results provide some empirical support for the income inequality hypothesis.

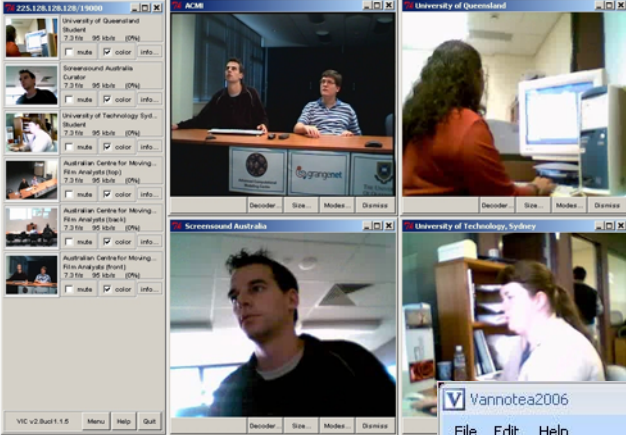
EPrint Type: Journal (Paginated)

Keywords: Income inequality,  Income inequality hypothesis, Health status

Subjects: [340000 Economics](#): [340200 Applied Economics](#): [340204 Health Economics](#)

ID Code: 801

Vannotea – Collaborative Annotation and Discussion of Medical Images/Videos

A screenshot of the Vannotea2006 software interface. The window title is 'Vannotea2006'. The interface is divided into several panels:

- Media Browser:** Shows a search bar, a 'Browse' button, and a list of video files. The selected file is 'SVL.Riess.5.04.mov' with the URL 'http://static.cjp.com/gems/test/SVL.Riess.5.04.mov'.
- Annotations Sidebar:** Contains a list of annotations: 'The reimplantation procedure explained' and 'The Valsava Prothesis'.
- Body Policy Other Info:** Shows a diagram of a heart with an aneurysm. The text next to it reads: 'The a cut, ju above aortic annul the co ostia (open where coron arterie attach the ac root). disea portio aorta remov'.
- Video Player:** Displays a surgical video of a heart procedure. The video is paused at 00:05:07.50. The total duration is 00:11:37.63.
- Annotea Timeline:** Shows a timeline with a 'ronalds' annotation at 00:02:54.
- Bottom Bar:** Contains buttons for 'Annotea Timeline', 'Jabber Client', 'Record and Replay', and 'Jabber Chat'. The status bar at the bottom right reads 'ActiveTaskManager: ManageActiveThreadList: 0 remaining'.

Annotation of Crystallographic Structures (Annotating) - Microsoft Internet Explorer

File Edit View Favorites Tools Help

Address <http://datura.itee.uq.edu.au/mjh/3molAnnotated/index.php?model=models/caffeine.xyz> Go Links

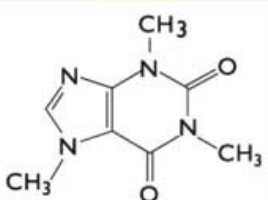

List Annotations Policies

- Supercritical carbon dioxide extraction
- Theophylline
- C8H10N4O2
- Wikipedia Reference
- Molar mass

Body Policy Other Info


(disambiguation).

Caffeine

General

Systematic name	1,3,7-trimethylxanthine
Other names	trimethylxanthine theine mateine guaranine methyltheobromine

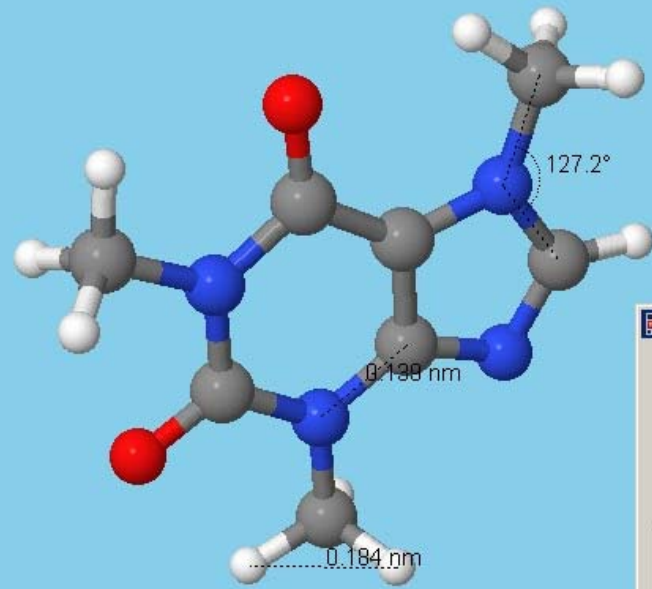


dataset acquisition
accessibility & annotation
e-research technologies

Annotation of Crystallographic Structures

[View model](#)
[Search publication databases](#),

Viewing models/caffeine.xyz



0.138 nm
127.2°
0.184 nm

Selected Atom: { 10 } (O 11 #1)

[Annotate](#)

Show bounding box
 Spin the model
[Rotate once about the x axis](#)

Custom Jmol scripting:
[\(what is this?\)](#)

[Execute](#)

Create New Annotation

Title:

Type:

Rank:

Language:

Body Policy

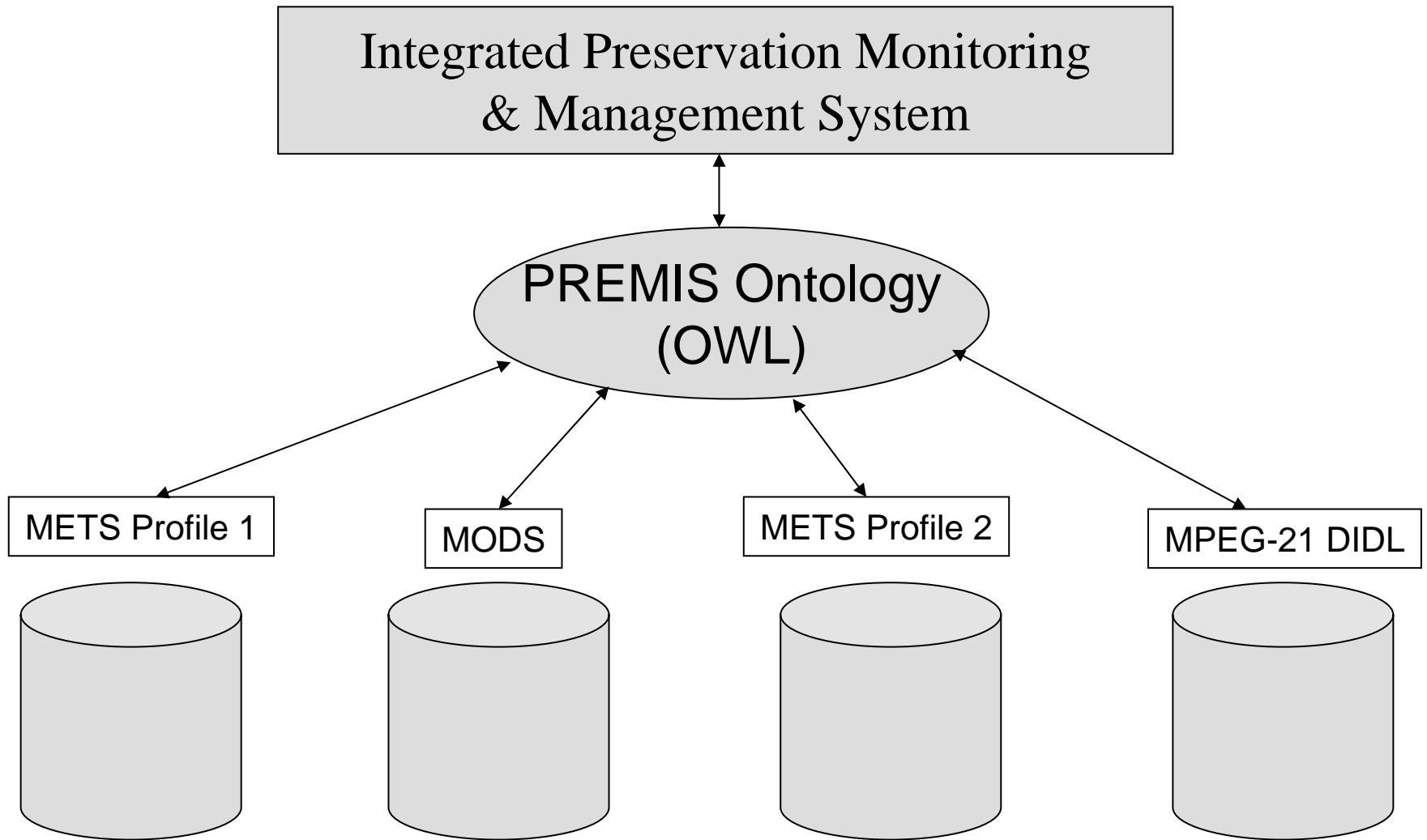
Plain Text URL File

O=C1C2=C(N=CN2C)N(C(=O)N1)C

Future Research

- Interoperability of Preservation Metadata – across heterogeneous archiving systems
- Collaborative Preservation – Decision support
 - Jabber instant messaging, chat, skype
 - Producer, Consumers, Repository Manager
- Trusted repositories -> trusted data/files/annotations
 - Social networks, FOAF with ratings
- Preservation metrics

Preservation Metadata Interoperability



Network of heterogeneous digital archiving systems

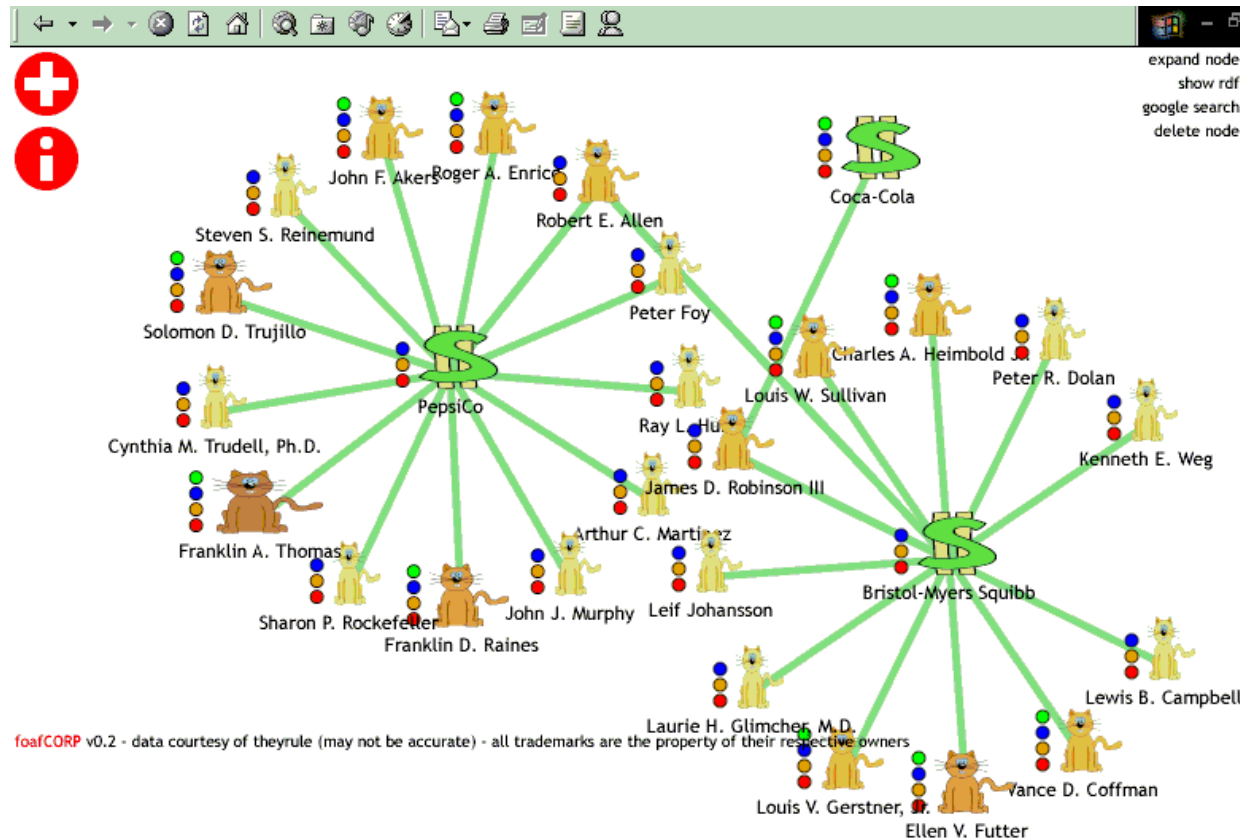
Trusted Repositories

Objects

- Data/files - genres
- Services/software
- Methodologies
- Annotations/reviews

Agents

- Organizations
- Research Groups
- Individuals



Preservation Selection Metrics

- **What is “high” quality data?** - what are the significant attributes
 - Accessibility
 - Accuracy, Completeness, Consistency
 - Reliability, Trustworthiness – reputation of source
 - Authenticity, certified – not tampered with
 - Provenance, metadata
 - Repeatable, reproducible, validation
 - Value-add – annotations, metadata
 - Re-use - citations
 - Uniqueness
 - Objectivity – unbiased
 - Relevance
 - Concise representation
 - Standards compliance
 - Currency, timeliness
 - Positive peer reviews, citations
- **How to measure/assess quality**
 - Which attributes are measurable? How to measure them?

References

<http://www.itee.uq.edu.au/~eResearch>
contact: jane@itee.uq.edu.au

