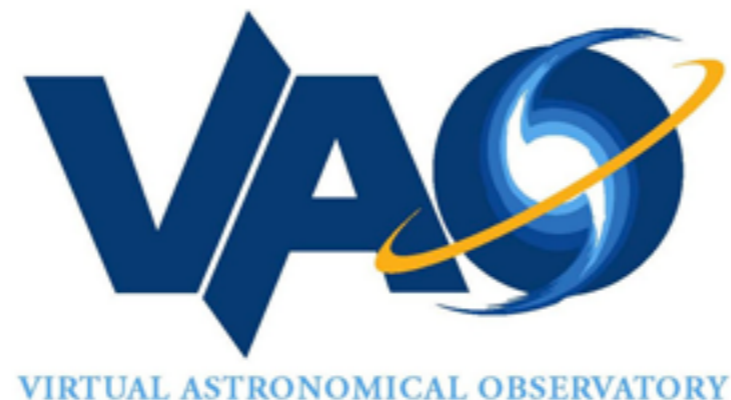


Semantic Interlinking of Resources in the Virtual Observatory Era

Alberto Accomazzi
Rahul Dave
NASA Astrophysics Data System

9 November 2010
Astronomical Data Analysis Software and Systems 2010, Boston, MA



Seamless Search & Discovery

ADS Labs VAO Search

Papers

Objects

Proposals

Datasets

Images

Observations

Journal

Apj

Aj

A&A

PASP

Wavelength

Xray

Optical

Ir

Uv

Search

Cart

xray emission in globular clusters

Sort by HasData HasObjects

[2001A&A...368.1021L](#) **A catalogue of low-mass X-ray binaries**
Liu, Q. Z.; van Paradijs, J.; van den Heuvel, E. P. J.
Astronomy and Astrophysics, v.368, p.1021-1054 (2001) Mar 2001

[2007A&G...48e..12M](#) **Compact objects in globular clusters**
Maccarone, Tom; Knigge, Christian
Astronomy & Geophysics, Volume 48, Issue 5, pp. 5.12-5.20. Oct 2007

[2005ApJ...625..796H](#) **A Deep Chandra Survey of the Globular Cluster 47 Tuc: Catalog of Point Sources**
Heinke, C. O.; Grindlay, J. E.; Edmonds, P. D.; Cohn, H. N.; Lugger, P. M.; Camilo Bogdanov, S.; Freire, P. C.
The Astrophysical Journal, Volume 625, Issue 2, pp. 796-824. Jun 2005

[2008ApJS..179..360G](#) **Thermonuclear (Type I) X-Ray Bursts Observed by the X-Ray Timing Explorer**
Galloway, Duncan K.; Muno, Michael P.; Hartman, Jacob M.; Psaltis, Dimitrios; Chakrabarty, Deepto
The Astrophysical Journal Supplement Series, Volume 179, Issue 2, pp. 360-422. Dec 2008

[2003ApJ...598..501H](#) **Analysis of the Quiescent Low-Mass X-Ray Binary Population in Galactic Globular Clusters**
Heinke, C. O.; Grindlay, J. E.; Lugger, P. M.; Cohn, H. N.; Edmonds, P. D.; Lloyd, D. Cool, A. M.
The Astrophysical Journal, Volume 598, Issue 1, pp. 501-515. Nov 2003

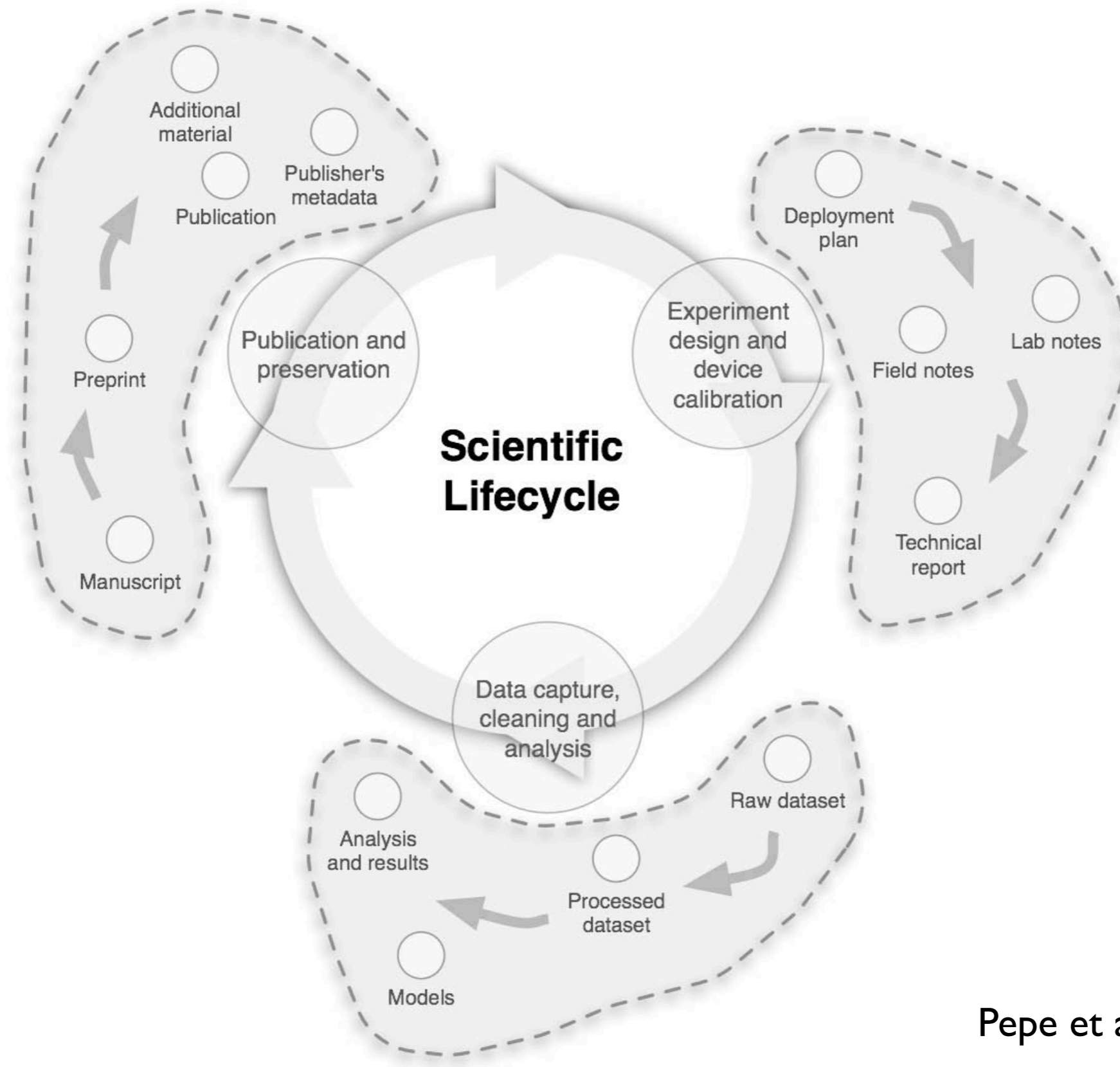
[1992PASP..104..981H](#) **Binaries in globular clusters**
Hut, Piet; McMillan, Steve; Goodman, Jeremy; Mateo, Mario; Phinney, E. S.; Pryor Carlton; Richer, Harvey B.; Verbunt, Frank; Weinberg, Martin
Astronomical Society of the Pacific, Publications (ISSN 0004-6280), vol. 104, no. 681, p. 98 Nov 1992

Overview

- Data Curation & Preservation in the Virtual Observatory
- Resource Metadata in Astronomy
- The Semantic Web
- VAO efforts for Resource Interlinking
- Applications

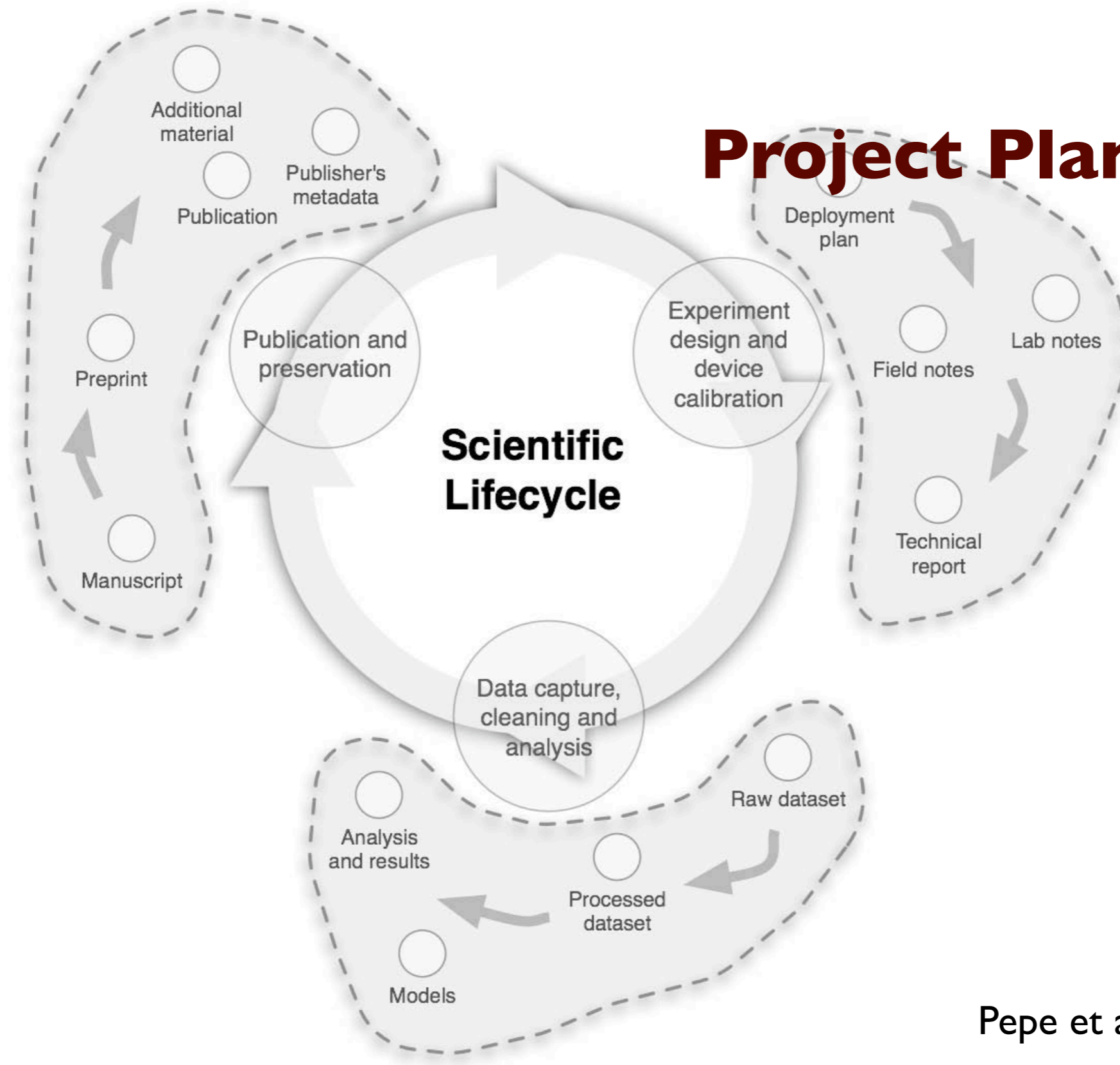
Data Curation & Preservation Principles

- Scientific research requires *repeatability*
- The lifecycle of a research project should be documented by *capturing all artifacts*
- Data, Processes, Results need to be properly described, accessible, and *linked* together
- *Provenance* information should be attached to curated metadata throughout the process

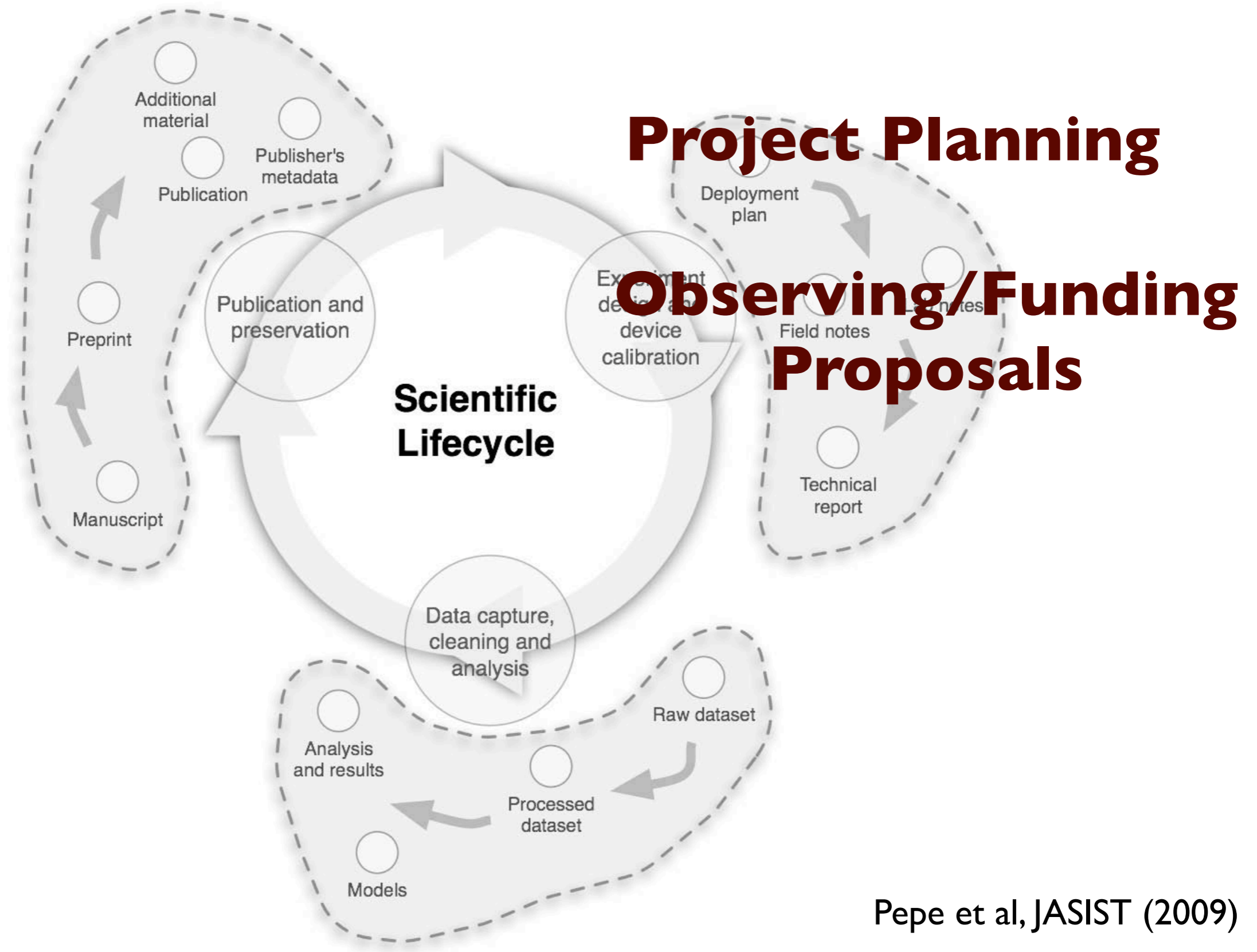


Pepe et al, JASIST (2009)

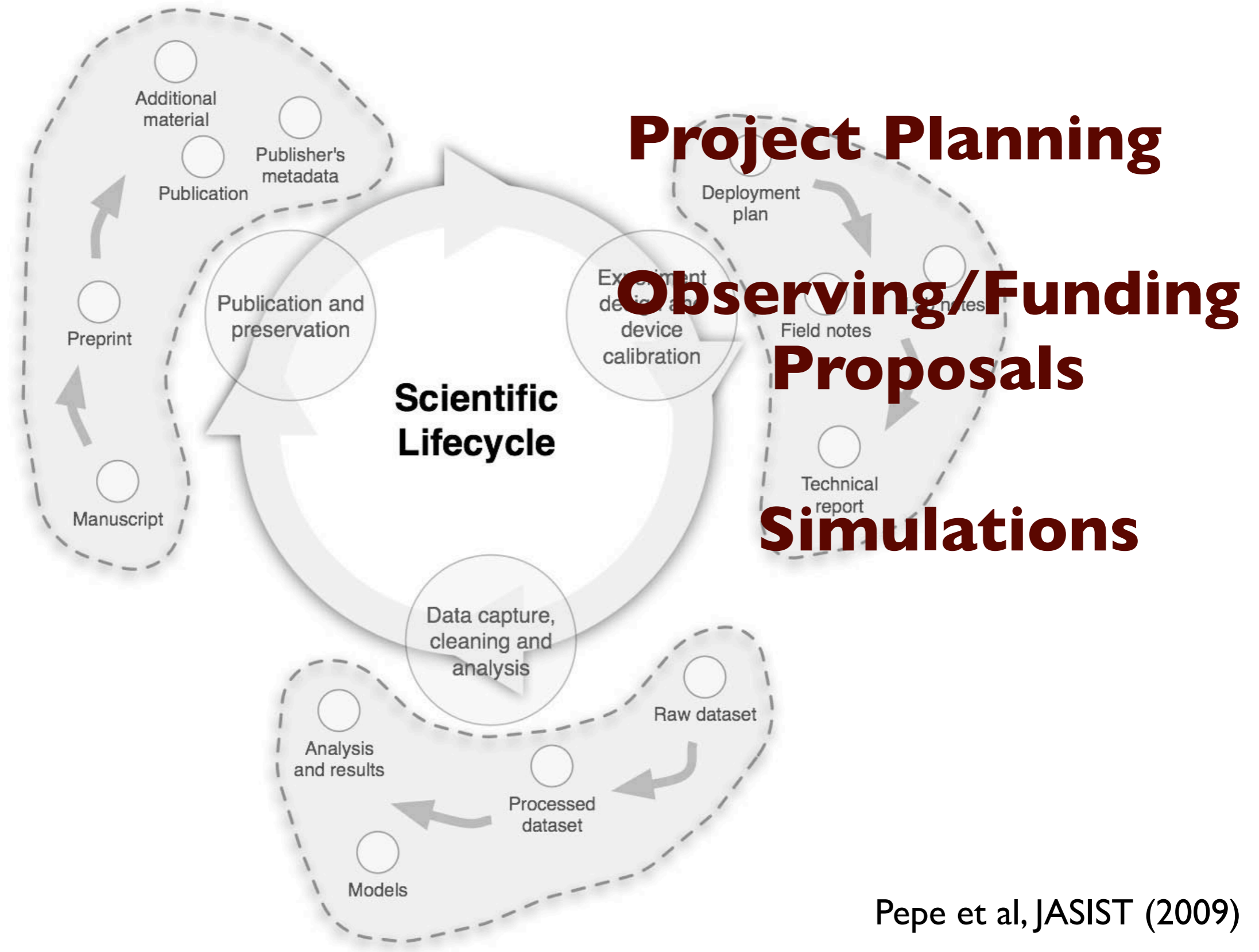
Project Planning



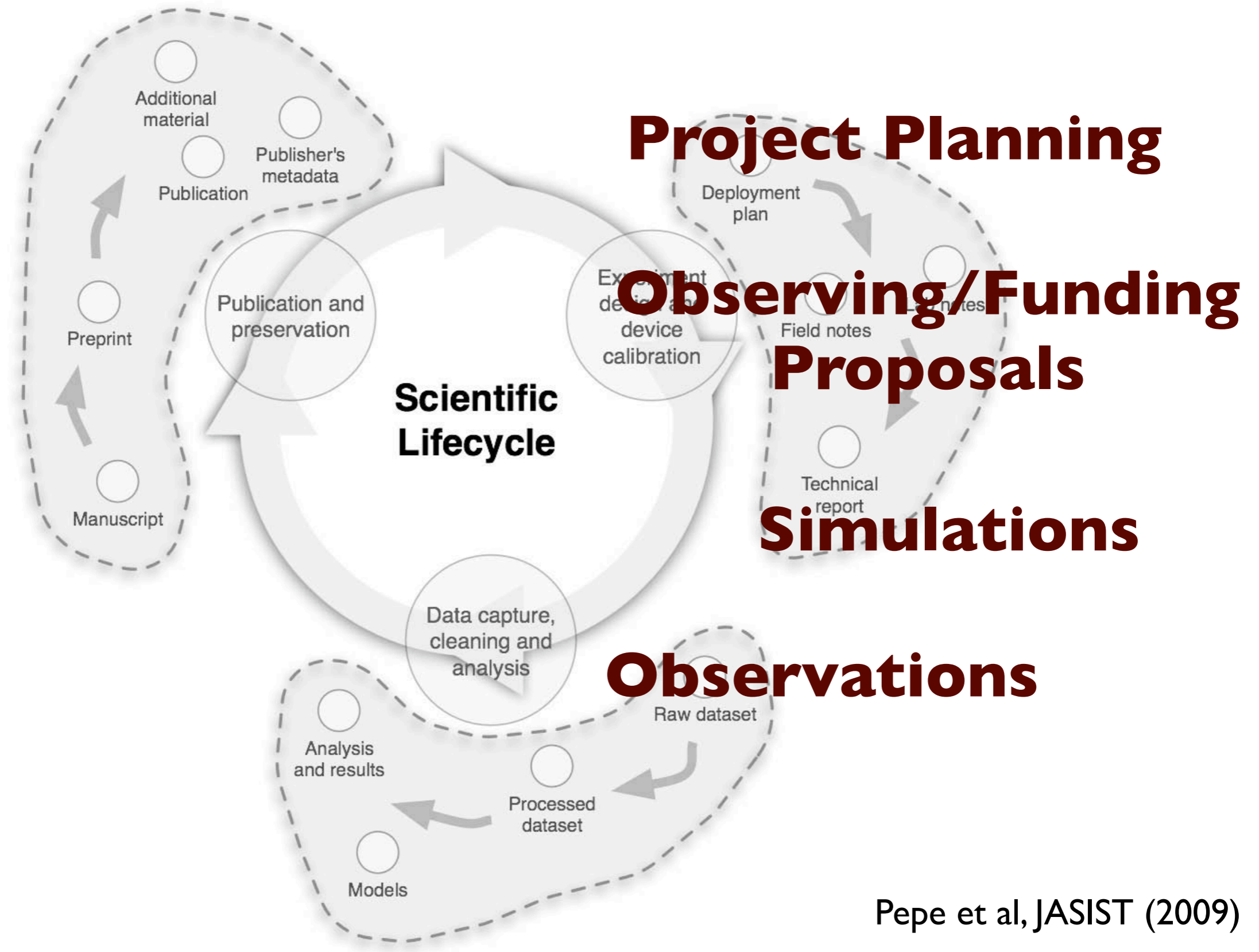
Pepe et al, JASIST (2009)



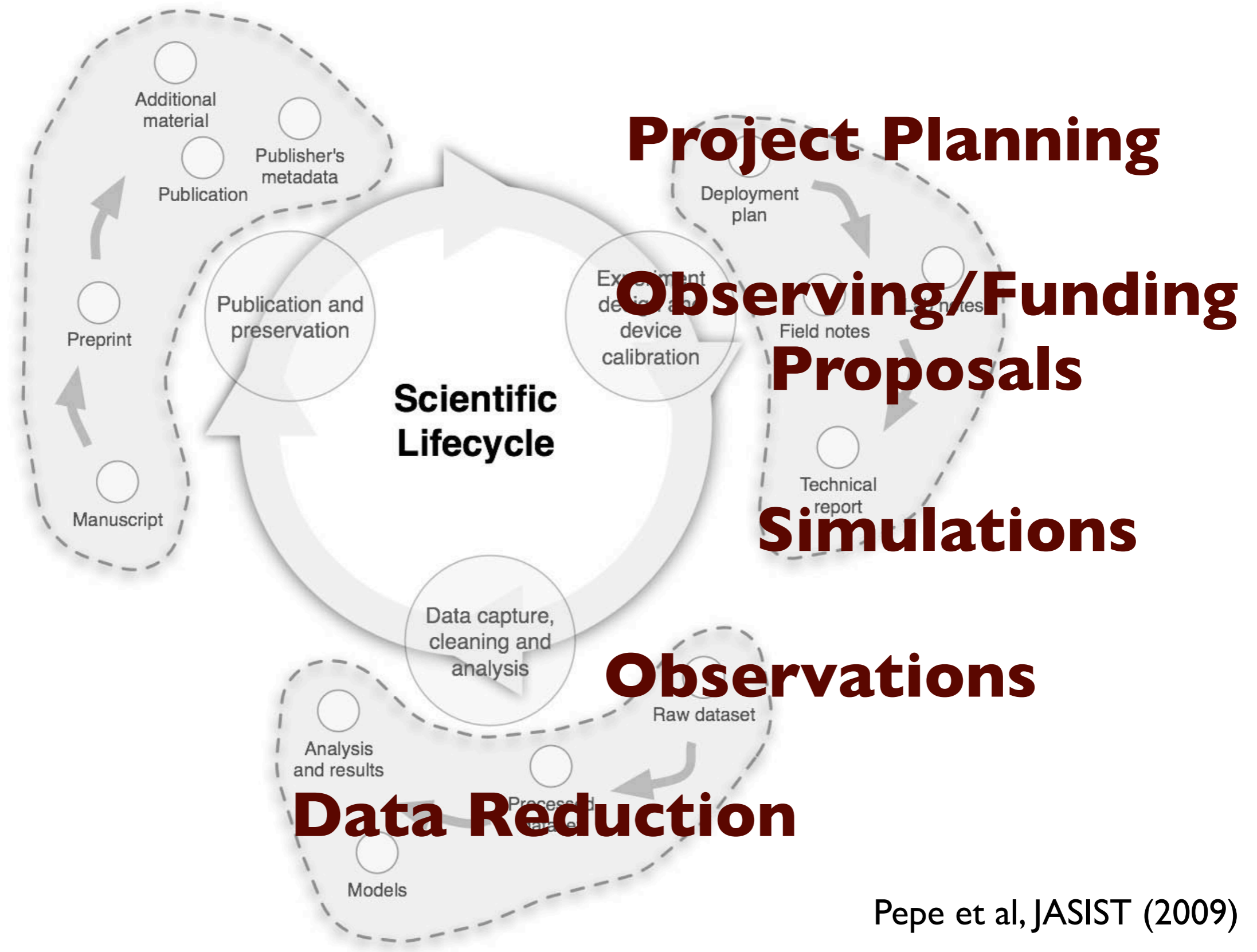
Pepe et al, JASIST (2009)



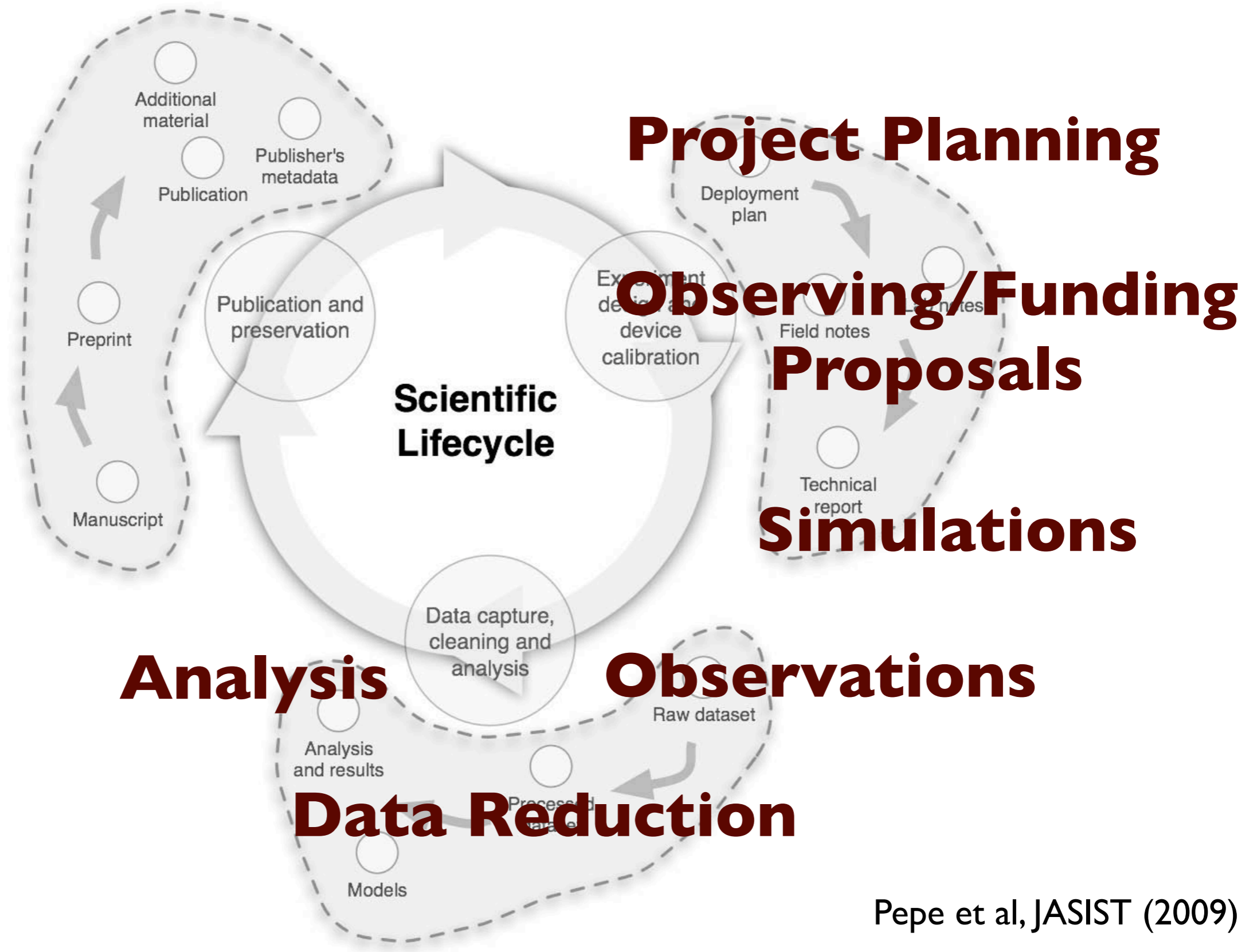
Pepe et al, JASIST (2009)



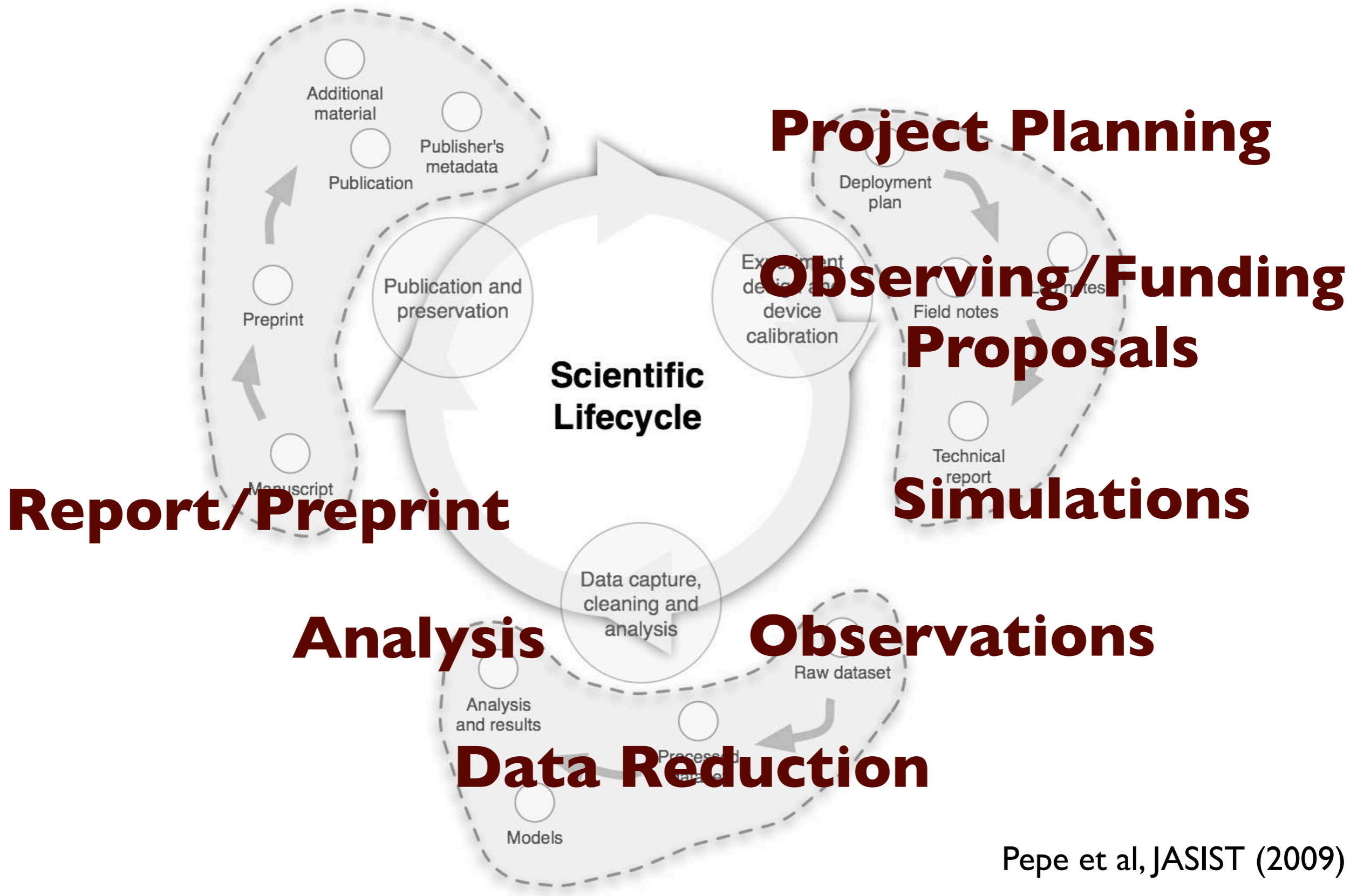
Pepe et al, JASIST (2009)



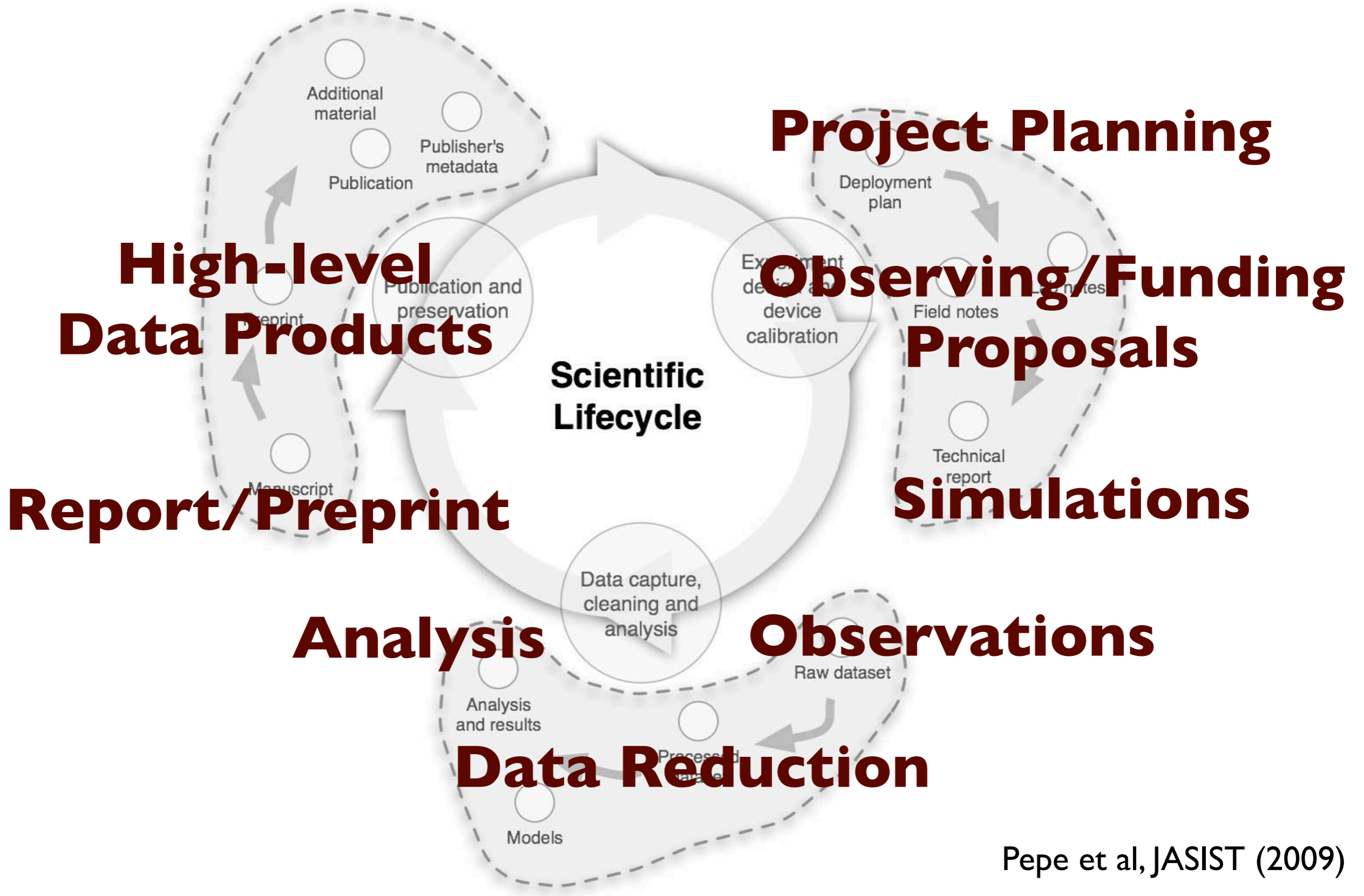
Pepe et al, JASIST (2009)



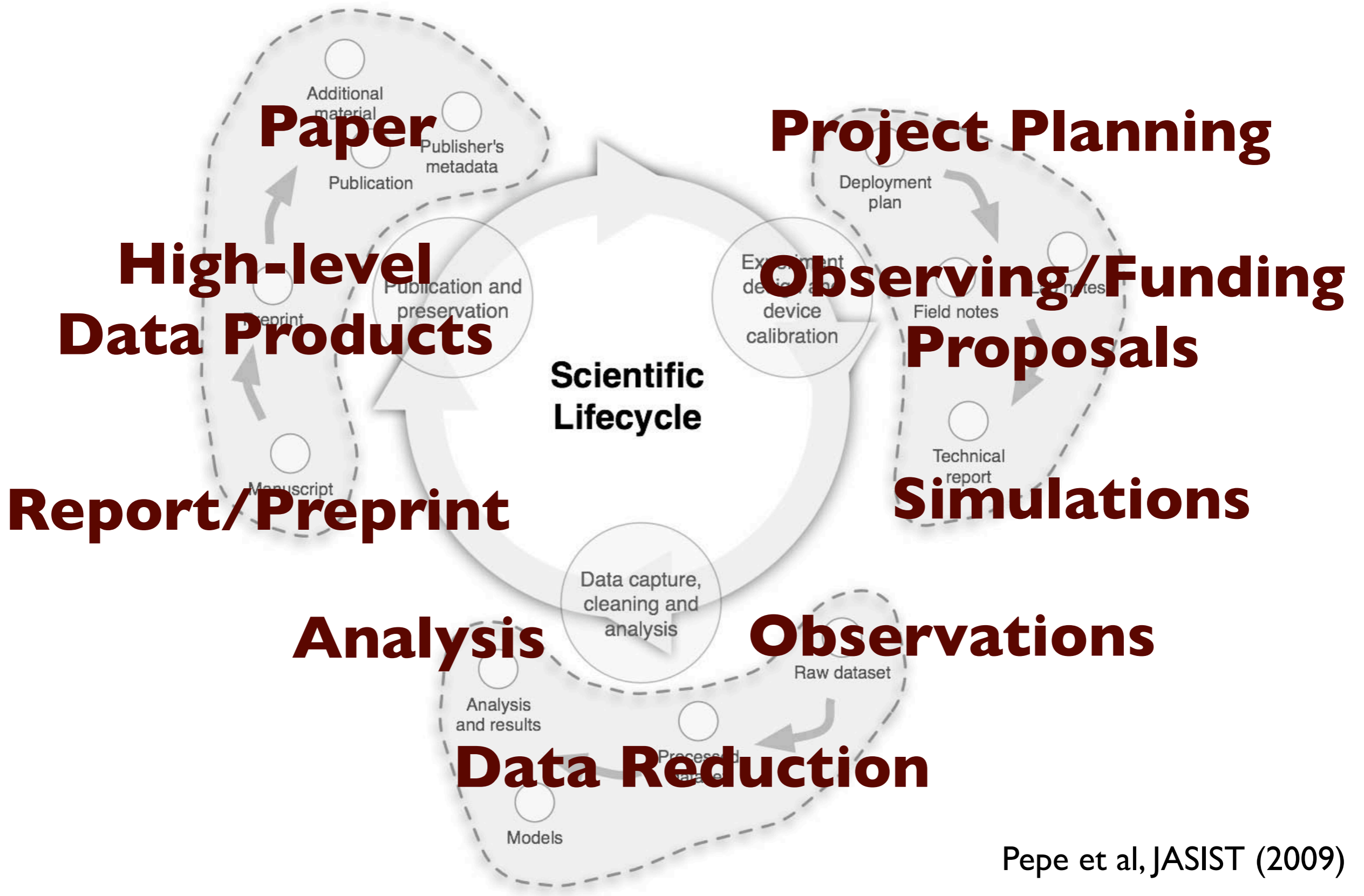
Pepe et al, JASIST (2009)



Pepe et al, JASIST (2009)



Pepe et al, JASIST (2009)



Pepe et al, JASIST (2009)

Curation in Astronomy

- Bibliographic Data & Metadata: Publishers, ADS
- Observational Data & Metadata: Archives, Observatories, Surveys
- Object Metadata: NED, SIMBAD, VizieR
- VO protocols “tie things together”

Curated Metadata

Paper

Project Planning

**High-level
Data Products**

**Observing/Funding
Proposals**

Report/Preprint

Simulations

Analysis

Observations

Data Reduction

Scientist's
Desktop
Observatories
Journals
Archives

Scientific
Lifecycle

Data capture,
cleaning and
analysis

and results

Deployment
plan

Publisher's
metadata

Publication

ion and
vation

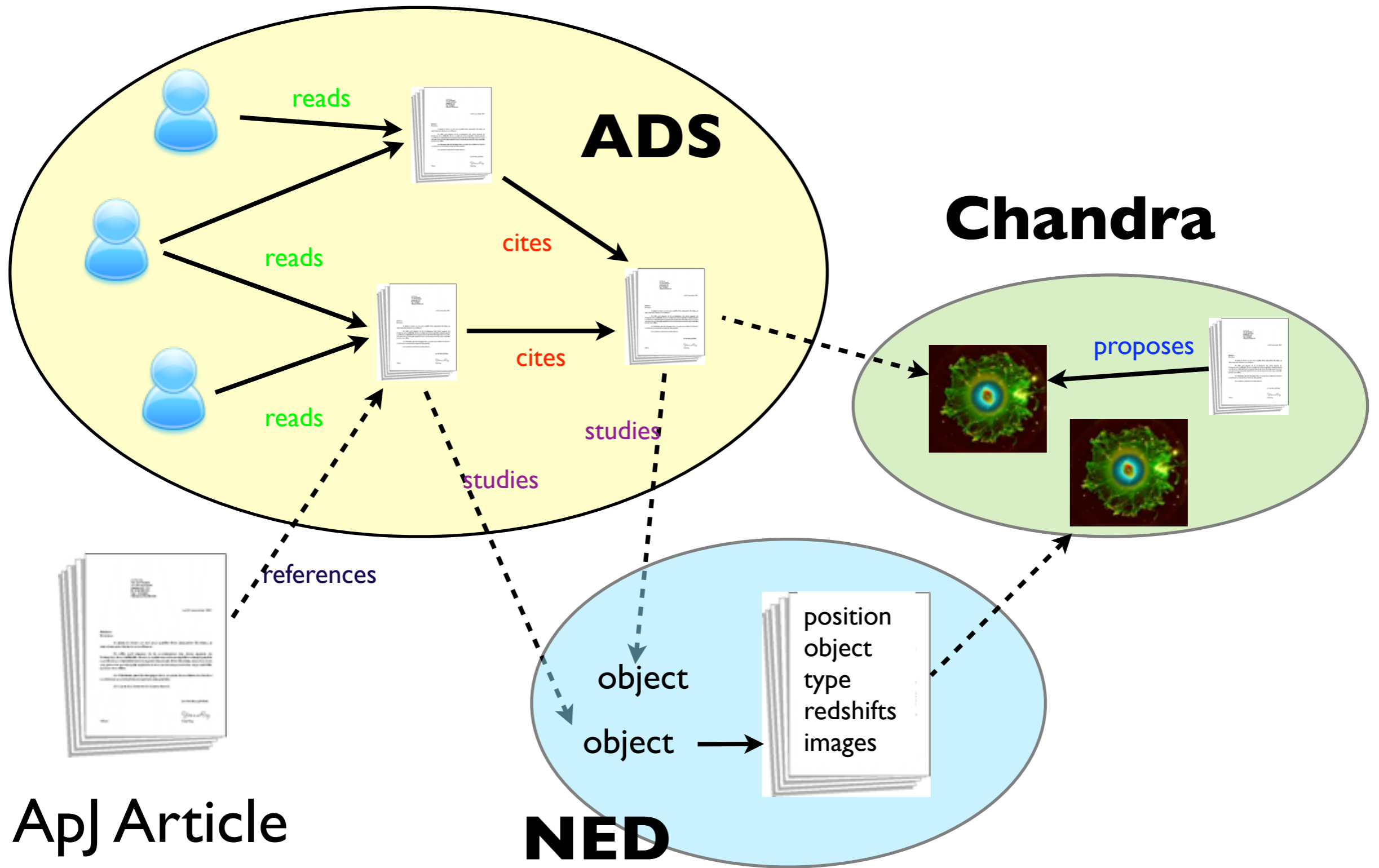
The Web of Astronomy Links

- Established very early in the game thanks to agreement on identifiers (pre-web!)
- Dense, curated by individual projects on behalf of the community
- Distributed and Heterogeneous: links span across domains (papers, data, objects)

What Resources?

- People, organizations, proposals
- Observations, instruments, wavelengths, footprints, observing logs
- Papers, citations, authors, readers
- Objects, catalogs, surveys
- Software, algorithms

ApJ/ADS/NED/Chandra



Optical follow-up of new Small Magellanic Cloud wing Be/X-ray binaries

http://adsabs.harvard.edu/abs/2007MNRAS.381.15615

Most Visited Getting Started Latest Headlines

Optical follow-up of new Small M... +

aaccomazzi@cfa.harvard.edu | my Account | Sign off

SAO/NASA ADS Astronomy Abstract Service

- [Find Similar Abstracts \(with default settings below\)](#)
- [Custom Format](#)
- [Electronic Refereed Journal Article \(HTML\)](#)
- [Full Refereed Journal Article \(PDF/Postscript\)](#)
- [FIND IT @ HARVARD](#)
- [arXiv e-print \(arXiv:0708.0228\)](#)
- [On-line Data](#)
- [References in the article](#)
- [Citations to the Article \(4\) \(Citation History\)](#)
- [Refereed Citations to the Article](#)
- [SIMBAD Objects \(10\)](#)
- [NED Objects \(5\)](#)
- [Also-Read Articles \(Reads History\)](#)

- [Translate This Page](#)

Title: Optical follow-up of new Small Magellanic Cloud wing Be/X-ray binaries

Authors: [Schurch, M. P. E.](#); [Coe, M. J.](#); [McGowan, K. E.](#); [McBride, V. A.](#); [Buckley, D. A. H.](#); [Galache, J. L.](#); [Corbet, R. H. D.](#); [Still, M.](#); [Vaisanen, P.](#); [Kniazev, A.](#); [Nordsieck, K.](#)

Affiliation: AA(School of Physics and Astronomy, Southampton University, Highfield, Southampton SO17 1BJ), AB(School of Physics and Astronomy, Southampton University, Highfield, Southampton SO17 1BJ), AC(School of Physics and Astronomy, Southampton University, Highfield, Southampton SO17 1BJ), AD(School of Physics and Astronomy, Southampton University, Highfield, Southampton SO17 1BJ), AE(South African Astronomical Observatory, PO Box 9, Observatory 7935, South Africa), AF(Harvard-Smithsonian Center for Astrophysics, 60 Garden Street, Cambridge, MA 02138, USA), AG(Universities Space Research Association, X-ray Astrophysics Laboratory, Mail Code 662, NASA Goddard Space Flight Center, Greenbelt, MD 20771, USA), AH(South African Astronomical Observatory, PO Box 9, Observatory 7935, South Africa), AI(South African Astronomical Observatory, PO Box 9, Observatory 7935, South Africa)

Done

Your NED Search Results

http://nedwww.ipac.caltech.edu/cgi-bin/nph-objsearch?objr

Recently Bookmarked DChart -- Ocean, Cli... Airfare Deals - Chea... YouTube - Google S... SKUA project BaseX - Download Pubget: the search e...

Data Related Directly to Object Names	Site/Service
Query SIMBAD by primary NED object name -- MESSIER 031	SIMBAD (CDS, Strasbourg, France)
Revised New General Catalogue -- NGC 0224	VizieR Catalog Query (U.S. mirror, CfA/Harvard)
Uppsala General Catalog -- UGC 00454	VizieR Catalog Query (U.S. mirror, CfA/Harvard)
Original Zwicky Catalog -- CGCG 535-017	VizieR Catalog Query (U.S. mirror, CfA/Harvard)
Query UZC Spectral Archive (60 arcsec search radius)	Updated Zwicky Catalog Data (Harvard/SAO)
Morphological Catalog of Galaxies -- MCG +07-02-016	VizieR Catalog Query (U.S. mirror, CfA/Harvard)
The Third Bologna Survey -- B3 0040+409	VizieR Catalog Query (U.S. mirror, CfA/Harvard)
2MASS Extended Source Images (JHKs) -- 2MASX J00424433+4116074	NASA/IPAC Infrared Science Archive (IRSA)
IRAS Point Source Catalog -- IRAS 00400+4059	VizieR Catalog Query (U.S. mirror, CfA/Harvard)
IRAS Faint Source Catalog -- IRAS F00400+4059	VizieR Catalog Query (U.S. mirror, CfA/Harvard)
Catalogue of Principal Galaxies -- PGC 002557	VizieR Catalog Query (U.S. mirror, CfA/Harvard)
Retrieve mean data from LEDA -- PGC 002557	The Lyon/Meudon Extragalactic Database (LEDA)
Query GALEX (NUV/FUV) Mission Archive (6' search radius) -- MESSIER 031	GALEX Mission Data Archive at MAST
General Archive Resources -- All queries centered at 00h42m44.3s, +41d16m09s (J2000)	Site/Service
Query Optical and UV Mission Archives (Default search radius)	Multimission Archive at STScI (MAST)
Query High Energy Mission Archives (Default search radius)	HEASARC (NASA/GSFC)
Explore resources with DataScope (15' search radius)	HEASARC (NASA/GSFC)
Visualize Coverage Map with IMPReSS Size: 1 deg	Astrophysics Data Facility (NASA/GSFC)
Retrieve 2MASS Atlas Images Band(s): Ks Size: 2'	NASA/IPAC Infrared Science Archive (IRSA)
Retrieve IRAS ISSA Images Band(s): 60um Size: 30'	NASA/IPAC Infrared Science Archive (IRSA)
1-D Coadd of IRAS Scans (ADDSCAN/SCANPI)	NASA/IPAC Infrared Science Archive (IRSA)
Retrieve NVSS Image Size: 15' <input checked="" type="radio"/> Contours (PS) <input type="radio"/> JPEG <input type="radio"/> FITS File	NRAO/VLA Sky Survey (NVSS)

Done

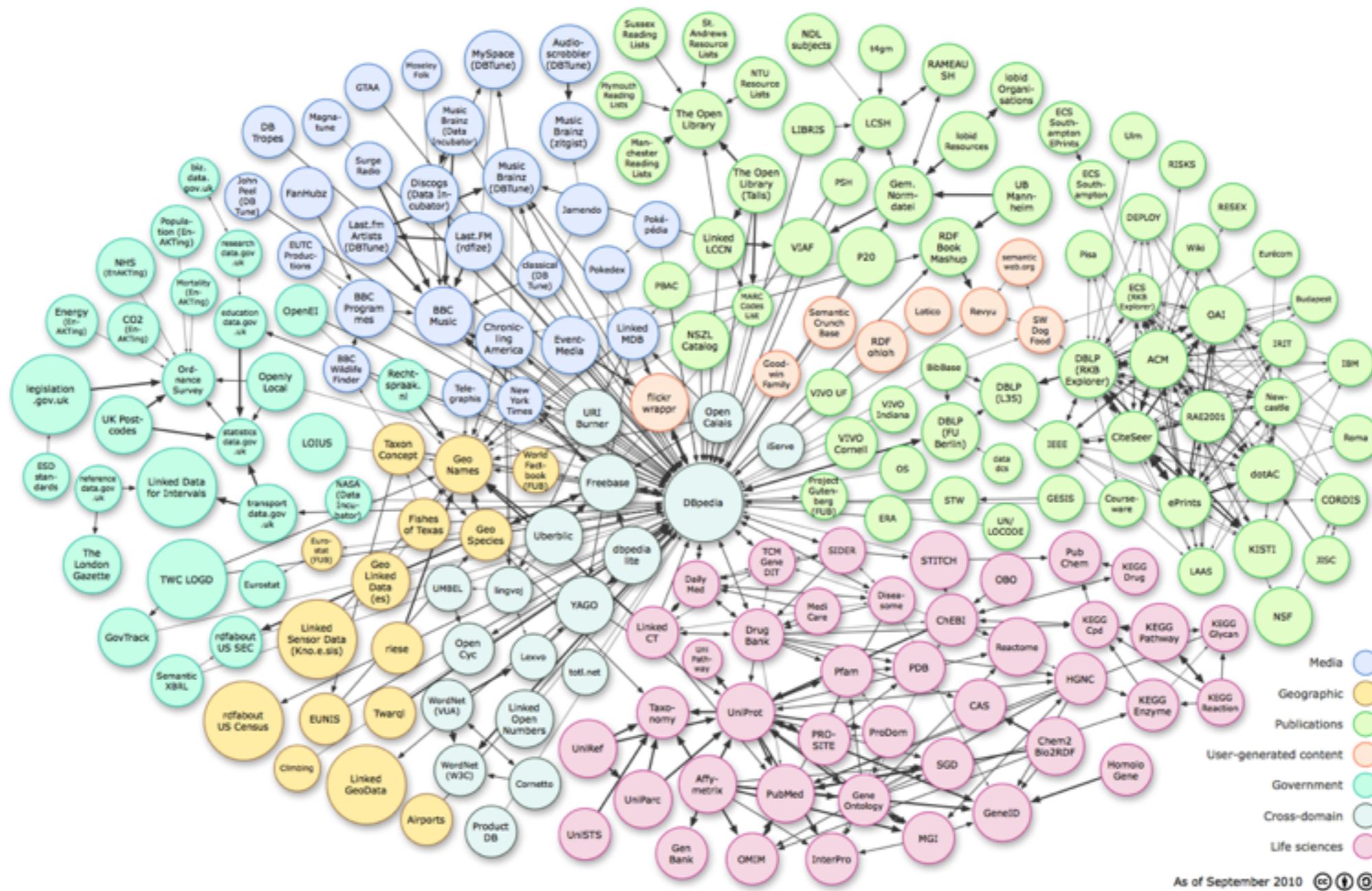
Why Semantics

- Built on the architecture of the web (Linked Data)
- A common language for resources and links, expressed in a machine-readable format (RDF)
- Knowledge representation based on concept schemes (Vocabularies, Ontologies)
- Support for Heterogeneous Knowledge Bases (triple/quad-stores), Rules, Reasoning, Inferencing, Annotation, Provenance
- Able to deal with incomplete/inconsistent information

Linked Data principles

- Resources are uniquely named via *URIs*
- Metadata is *open* and in a *standard format*
- Links between resources are *typed*
- Resource metadata includes *links* to other resources
- Built on the architecture of the web, *no APIs*

Linked Data Cloud



Linking Open Data cloud diagram, by Richard Cyganiak and Anja Jentzsch. <http://lod-cloud.net/>

Resource Description Framework (RDF)

- Reduce all knowledge to simple statements in the format of subject-predicate-object
- Entities and relationships are formally defined
- Provenance is preserved throughout

ESO> **paperA** **fundedBy** **grantG**

ADS> **paperA** **cites** **paperB**

ADS> **paperB** **describes** **obsX**

CXO> **obsX** **isIn** **ChandraArchive**

CXO> **obsX** **covers** **FootprintF**

NED> **objectY** **in** **FootprintF**

NED> **objectY** **isA** **QSO**

Reasoning

SIMBAD> paperA describes obsX

SIMBAD> paperA describes obsY

CXO> obsX covers FootprintA

CXO> obsX in ChandraArchive

MAST> obsY covers FootprintB

MAST> obsY in HLArchive

AGENT> FootprintA overlaps FootprintB

AGENT> paperA isA multiWavelengthStudy

Links to Astronomy Datasets

- Have been based on reciprocal links between *HTML pages* curated by collaborators
- Do not identify resources as *unique URIs*
- Do not make use *standard vocabularies* to describe data or the *types of links* between them
- Are *not actionable* by applications

Semantic Interlinking in the VO

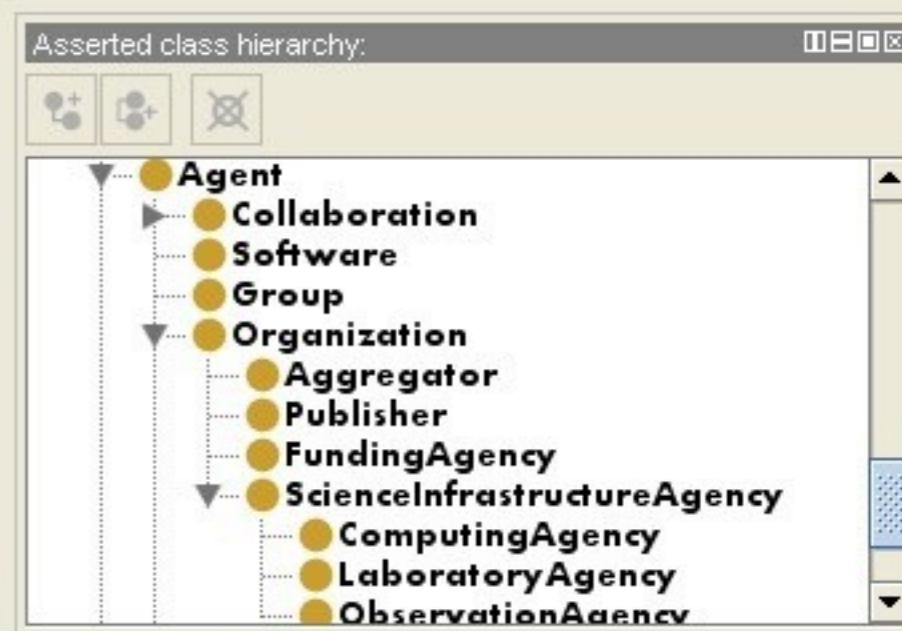
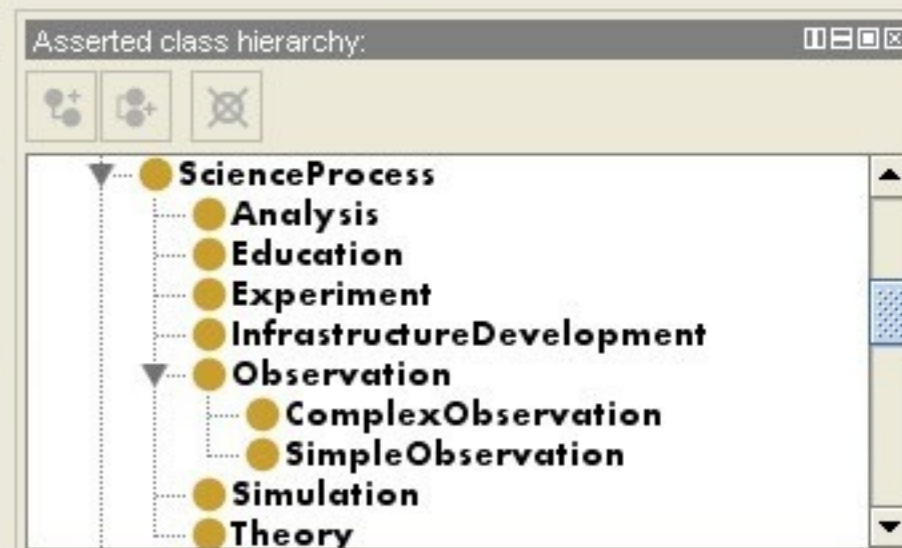
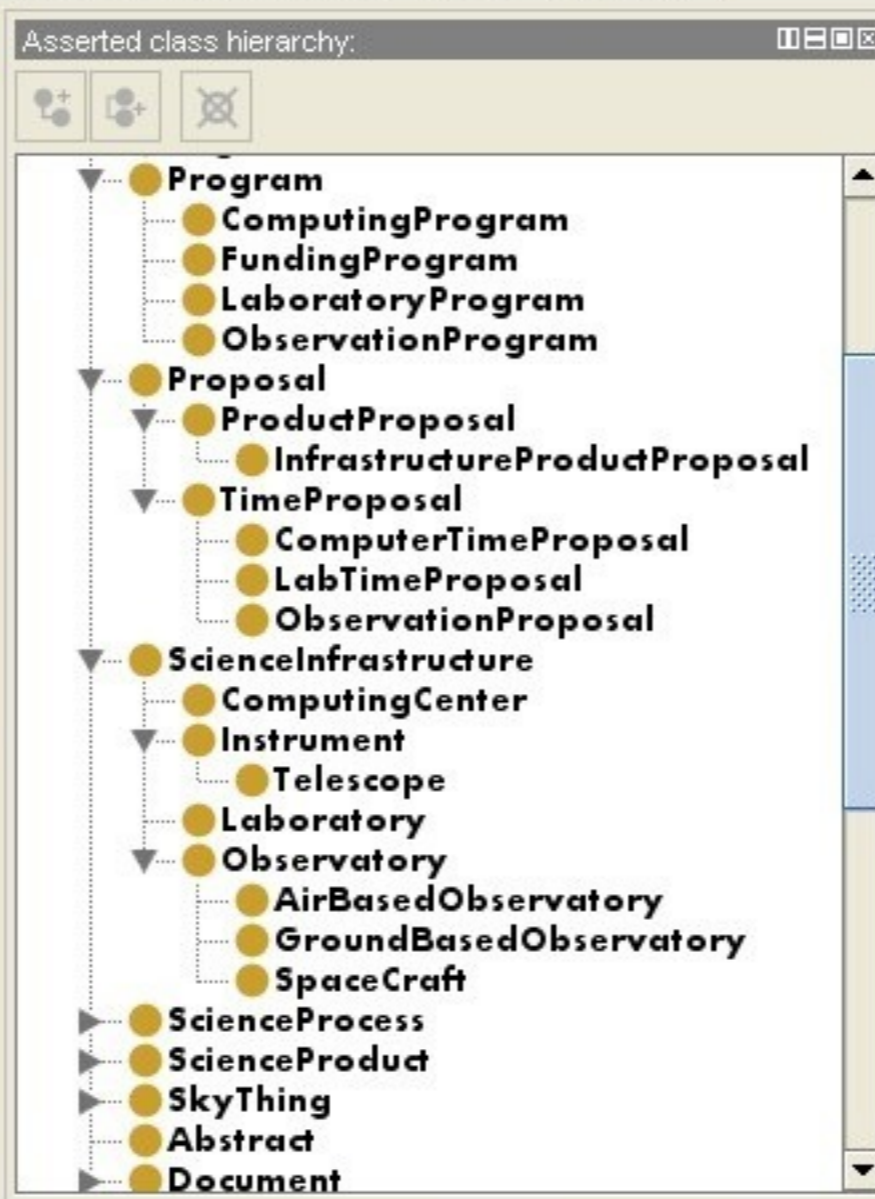
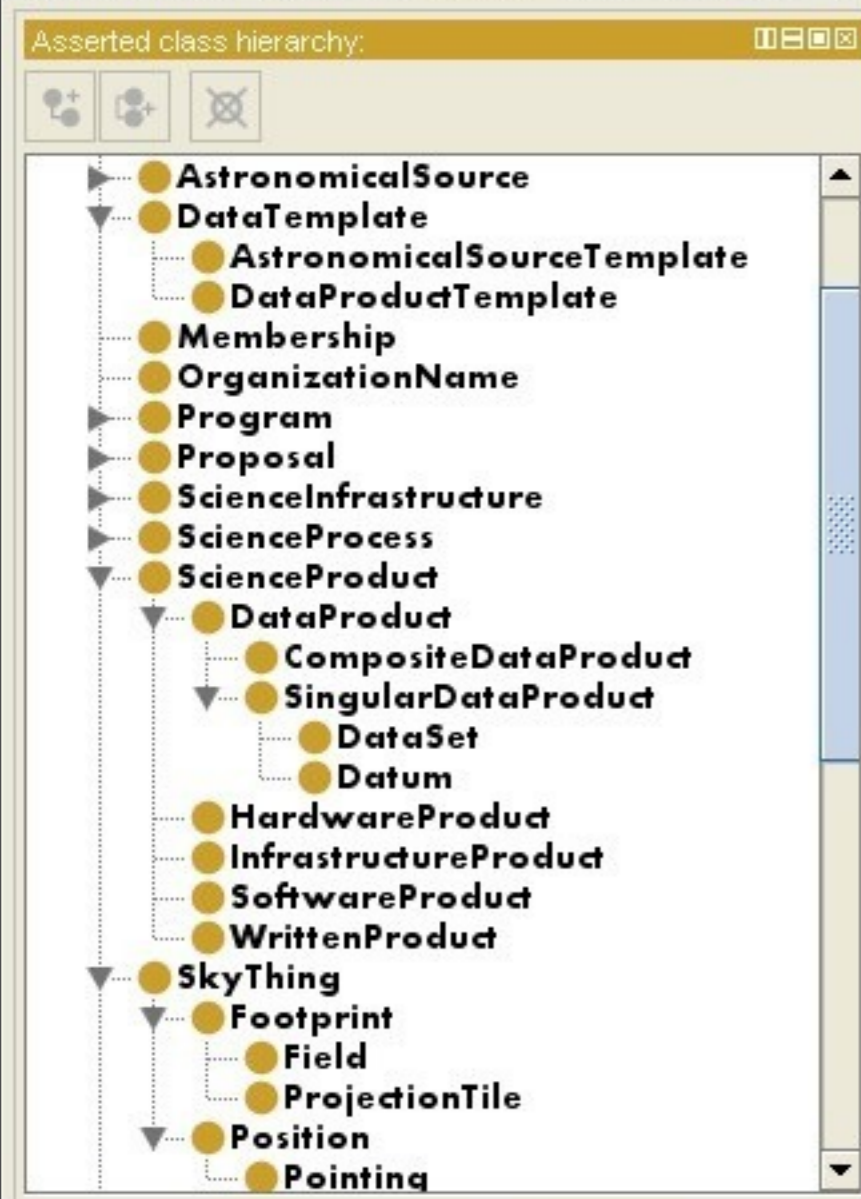
- Exposing and linking metadata from separate archives allows computing over *heterogeneous VO resources*
- Tackles *outstanding issues* with dataset IDs, relates to ongoing IVOA DC&P activities
- Provides semantic layer over existing resources, *crossing boundaries* between archives
- Now finally possible(?) thanks to ObsCore, DM, IVOA Semantics, Portal development efforts

Current Effort

- Develop & Adapt Ontologies to describe Astronomical Resources
- Identify, collect, and expose metadata for Datasets
- Create typed links between Data products, Bibliographic and Object databases
- In the process, create knowledge base about Instruments and Services
- Incorporate all of the above in a “metadata store,” exposing Resources using Semantic Web standards

Ontology Development

- Define URIs for all Resources
- Stay within OWL-DL
- Build on skos, pav, agents, foaf, cito, fabio
- Skos based keywords from IVOA
- Object Type Ontology from SIMBAD
- Develop VAObase, VAObib, VAOobsv



Sample Applications

- A view of bibliographies with facets based on objects, wavelengths, keywords, etc.
- VO-powered portals (e.g. VAO, CDS)
- A UI integrating views of bibliographies, objects, and datasets interacting together
- An view of the “Enhanced Paper,” incorporating links to datasets and objects
- Applications computing metrics on data, objects based on citation and use of papers



The Smithsonian/NASA Astrophysics Data System ADS-Labs project



Home

ADS Classic

Search

Search

xray "globular clusters"

Disable Topic Search

Keyword Search:	Subject Area Search:
<input type="radio"/> Most relevant	<input type="radio"/> Most popular
<input type="radio"/> Most recent	<input type="radio"/> Most useful
<input type="radio"/> Most important	<input checked="" type="radio"/> Most instructive



The ADS is Operated by the Smithsonian Astrophysical Observatory under NASA Grant NNX09AB39G



The Smithsonian/NASA Astrophysics Data System ADS-Labs project



Home

ADS Classic

Search

Query Results from the ADS Database

Object: Galaxy [X]AND Keyword: galaxy globular clusters [X]AND Author: Grindlay, J [X]

FILTER BY:

SIMBAD Objects

 X-Ray Source Other object Galaxy NGC 104 (86) M 15 (61) NGC 6397 (59) NGC 6624 (50) NGC 6752 (47) NGC 5139 (43) NGC 6440 Star Nebula[ADS Publications](#)[SIMBAD Info](#)[World Wide Telescope](#)[Aladin applet](#)

7. [2005ApJ...625..796H](#) **A Deep Chandra Survey of the Globular Cluster 47 Tucanae: Catalog of Point Sources**
Heinke, C. O.; Grindlay, J. E.; Edmonds, P. D.; Cohn, H. N.; Lugger, P. M.; Camilo, F.; Bogdanov, S.; Freire, P. C.
The Astrophysical Journal, Volume 625, Issue 2, pp. 796-824. Jun 2005
10. [2003ApJ...598..501H](#) **Analysis of the Quiescent Low-Mass X-Ray Binary Population in Galactic Globular Clusters**
Heinke, C. O.; Grindlay, J. E.; Lugger, P. M.; Cohn, H. N.; Edmonds, P. D.; Lloyd, D. A.; Cool, A. M.
The Astrophysical Journal, Volume 598, Issue 1, pp. 501-515. Nov 2003
14. [2006ApJ...651.1098H](#) **Faint X-Ray Sources in the Globular Cluster Terzan 5**
Heinke, C. O.; Wijnands, R.; Cohn, H. N.; Lugger, P. M.; Grindlay, J. E.; Pooley, D.; Lewin, W. H. G.
The Astrophysical Journal, Volume 651, Issue 2, pp. 1098-1111. Nov 2006
- [2007ApJ...657..286L](#) **Chandra X-Ray Sources in the Collapsed-Core Globular Cluster M30 (NGC 7099)**
Lugger, Phyllis M.; Cohn, Haldan N.; Heinke, Craig O.; Grindlay, Jonathan E.; Edmonds, Peter D.

ads The Smithsonian/NASA Astro

Home Help Sitemap

On searches for pulsed emission with application to four globular cluster X-ray sources - NGC 1851, 6441, 6624, and 6712

Leahy, D. A.; Darbro, W.; Elsner, R. F.; Weisskopf, M. C.; Kahn, S.; Sutherland, P. G.; Grindlay, J. E.

Astrophysical Journal, Part 1, vol. 266, Mar. 1, 1983, p. 160-170.

Periodic pulsations have been sought in the NGC 1851, 6441, 6624, and 6712 using the Einstein Observatory's Monitor Proportional Counter. Special methods of analysis, which correctly account for several effects that have been overlooked by other researchers. No pulsed emission was detected in the periods searched. The 90% confidence upper limits for the pulsed fraction are given.

Keywords: GLOBULAR CLUSTERS, PULSARS, STELLAR RADIATION, X RAY SOURCES, CONFIDENCE LIMITS, FAST FOURIER TRANSFORMATIONS, HEAO 2, MONTE CARLO SIMULATIONS

DOI: [10.1086/160766](https://doi.org/10.1086/160766)

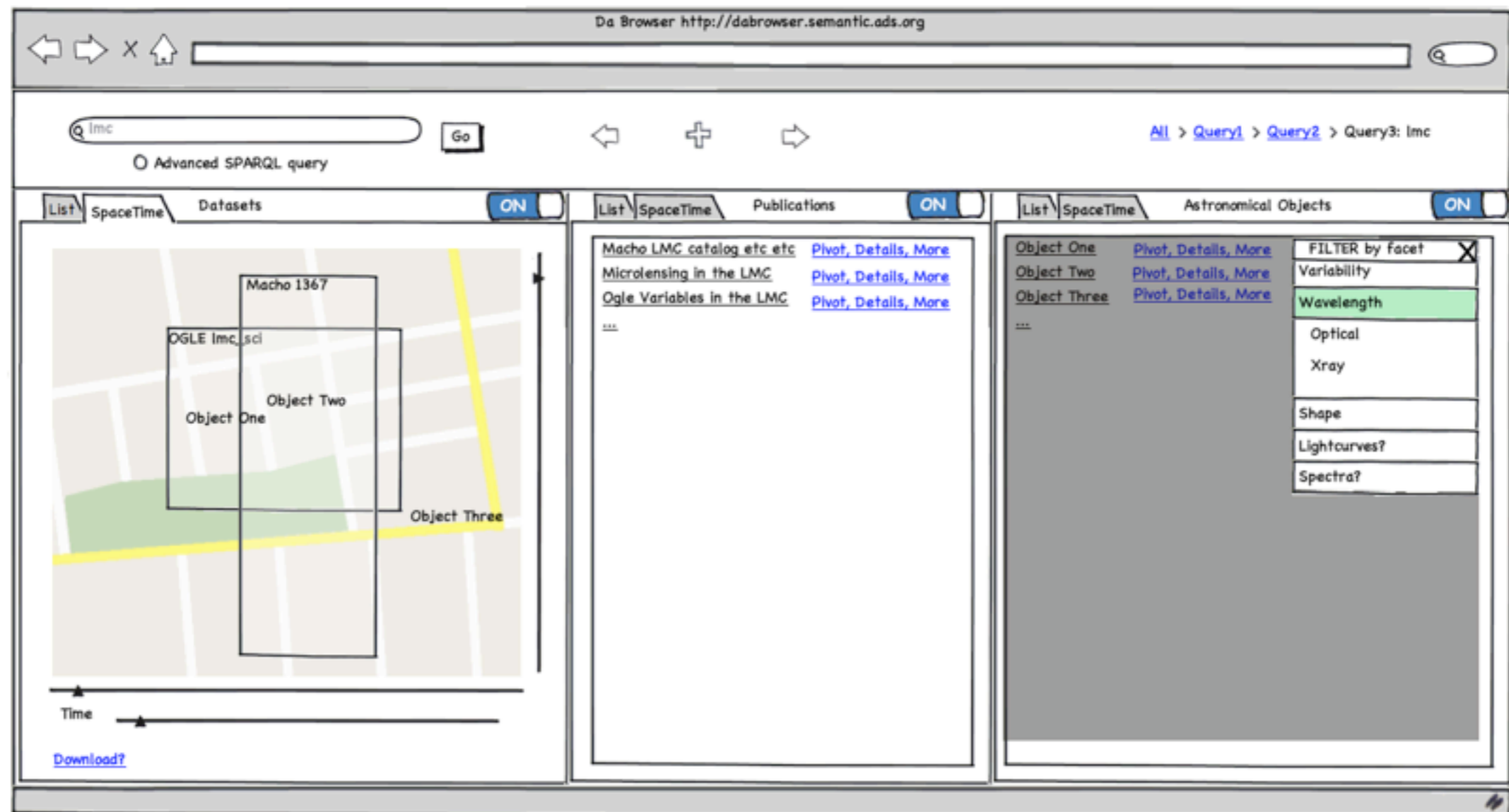
The ADS is Operated by the Smithsonian Astrophysical Observatory under NASA Grant

```

1 <!DOCTYPE HTML PUBLIC "-//W3C//DTD HTML 4.0 Transitional//EN" "http://www.w3.org/TR/REC-html40" >
2 <html>
3 <head>
4 <title>NASA ADS: On searches for pulsed emission with application to four globular cluster
  NGC 1851, 6441, 6624, and 6712</title>
5 <link rel="stylesheet" type="text/css" media="screen" href="/abs_doc/ads.css" />
6 <meta name="citation_language" content="en" />
7 <meta name="citation_doi" content="10.1086/160766" />
8 <meta name="citation_abstract_html_url" content="http://adsabs.harvard.edu/abs/1983ApJ...266...160L?version=2" />
9 <meta name="citation_title" content="On searches for pulsed emission with application to four
  cluster X-ray sources - NGC 1851, 6441, 6624, and 6712" />
10 <meta name="citation_authors" content="Leahy, D. A.; Darbro, W.; Elsner, R. F.; Weisskopf,
  Sutherland, P. G.; Grindlay, J. E." />
11 <meta name="citation_keywords" content="GLOBULAR CLUSTERS, PULSARS, STELLAR RADIATION, X RAY
  STARS, COMPUTERIZED SIMULATION, CONFIDENCE LIMITS, FAST FOURIER TRANSFORMATIONS, HEAO 2, MONTE
  CARLO SIMULATIONS" />
12 <meta name="citation_issn" content="0004-637X" />
13 <meta name="citation_date" content="03/1983" />
14 <meta name="citation_journal_title" content="The Astrophysical Journal" />
15 <meta name="citation_volume" content="266" />
16 <meta name="citation_firstpage" content="160" />
17 <meta name="citation_lastpage" content="170" />
18 <link title="schema(PRISM)" rel="schema.prism" href="http://prismstandard.org/namespaces/1.1/" />
19 <meta name="prism:publicationName" content="The Astrophysical Journal" />
20 <meta name="prism:issn" content="0004-637X" />
21 <meta name="prism:volume" content="266" />
22 <meta name="prism:startingPage" content="160" />
23 <meta name="prism:endingPage" content="170" />
24 <link title="schema(DC)" rel="schema.dc" href="http://purl.org/dc/elements/1.1/" />
25 <meta name="dc:language" content="en" />
26 <meta name="dc:identifier" content="doi:10.1086/160766" />
27 <meta name="dc:date" content="1983-03" />
28 <meta name="dc:source" content="Astrophysical Journal, Part 1, vol. 266, Mar. 1, 1983, p. 160-170" />
29 <meta name="dc:title" content="On searches for pulsed emission with application to four globular
  cluster X-ray sources - NGC 1851, 6441, 6624, and 6712" />
30 <meta name="dc:creator" content="Leahy, D. A." />
31 <meta name="dc:creator" content="Darbro, W." />
32 <meta name="dc:creator" content="Elsner, R. F." />
33 <meta name="dc:creator" content="Weisskopf, M. C." />
34 <meta name="dc:creator" content="Kahn, S." />
35 <meta name="dc:creator" content="Sutherland, P. G." />
36 <meta name="dc:creator" content="Grindlay, J. E." />
37 </head>
38 <body>
39 <div id="wrap">
40 <div id="header">
41 <table>
42 <tr><td>
43 <div id="logos">
44 <table>
45 <tr><td>
46 
  
```

APOD Application

(Astronomy Papers, Objects & Datasets)






created with Balsamiq Mockups - www.balsamiq.com

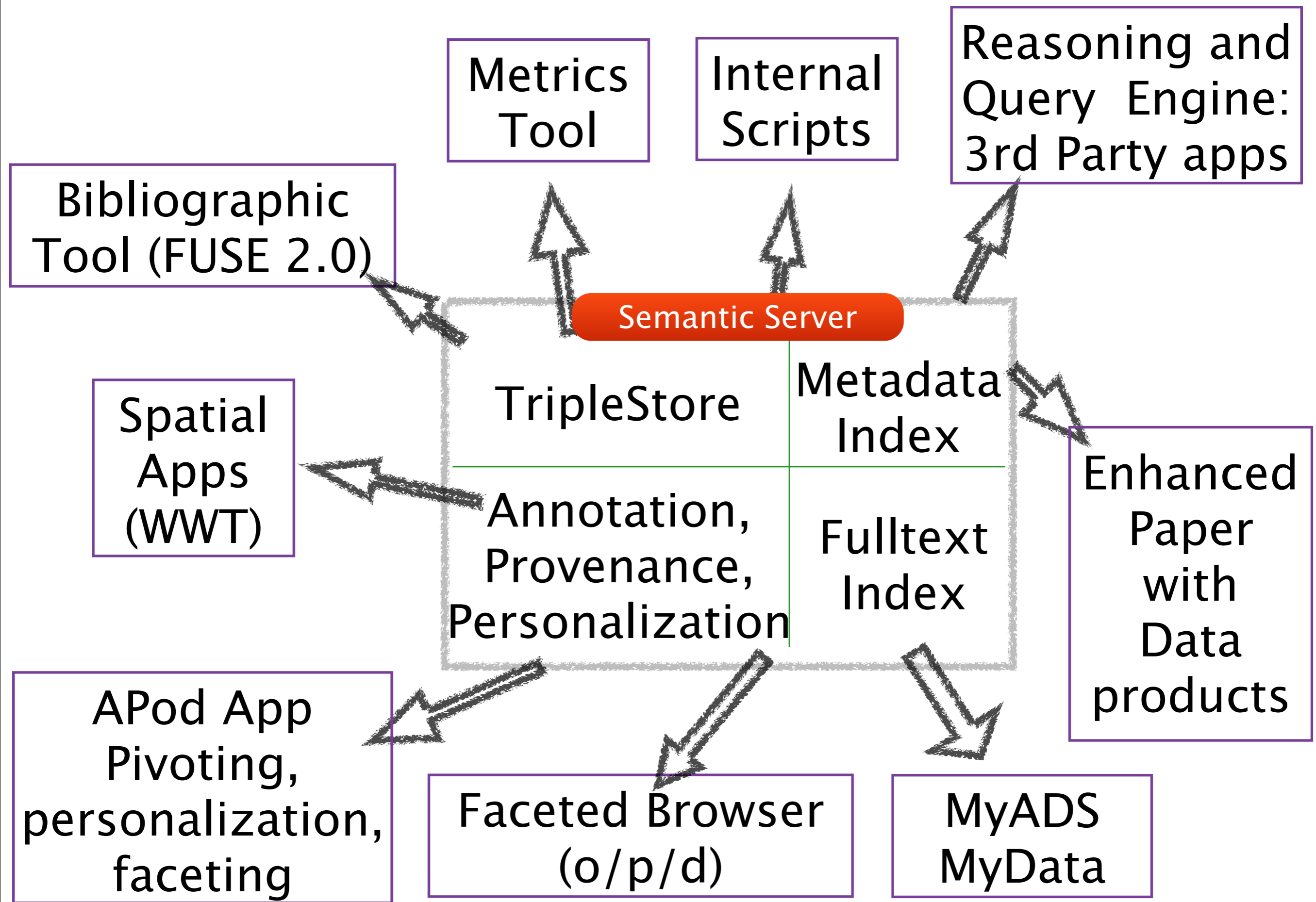
Datasets

Papers

Objects

Paper, Data “Mashup”

paper	associations			
<p data-bbox="441 583 812 625">Quasars in the LMC</p> <p data-bbox="455 699 776 768">Authors: P. Protopapas, D.W. Kim, et al</p> <p data-bbox="455 810 798 1283">Abstract: We identify new EBs in the MACHO Dataset in the LMC using supervised learning techniques. After establishing our algorithms on the OGLE2 dataset we let it loose on the MACHO LMC dataset, discovering some hitherto unknown eclipsing binaries.</p> <table border="1" data-bbox="524 1388 743 1640"><tr><td data-bbox="524 1388 743 1472">Text</td></tr><tr><td data-bbox="524 1472 743 1556">Figures</td></tr><tr><td data-bbox="524 1556 743 1640">Datasets</td></tr></table>	Text	Figures	Datasets	<p data-bbox="1100 583 1264 625">Datasets</p> <div data-bbox="1059 674 2189 1304"><p data-bbox="1141 716 1476 747">Macho lmc_lc1 1234.43.9</p><p data-bbox="1141 789 1421 821">Ra, dec, mag, etc etc</p><p data-bbox="1950 705 2115 737"><u>From TSC</u></p></div> <div data-bbox="1059 1346 2189 1818"><p data-bbox="1141 1430 1347 1461">Ogle2 aaa+bbb</p><p data-bbox="1141 1503 1421 1535">Ra, dec, mag, etc etc</p><p data-bbox="1909 1377 2101 1409"><u>From OGLE2</u></p></div>
Text				
Figures				
Datasets				



Conclusions

- Effort will create useful VO infrastructure
- Tackles outstanding issues with dataset IDs, DC&P activities
- Provides semantic layer over existing resources, leverages existing services
- Eases computation of metrics, supports text-mining, portal development
- Follow ongoing development in IVOA Semantics and at <http://github.com/rahuldave/ontoads>

Thanks to...

- The “Seamless Astronomy” group at CfA:
Alyssa Goodman, Gus Muench, Alberto Pepe
- ADS and Chandra projects at SAO:
Michael Kurtz, Doug Burke, Sherry Winkelman
- Collaborators at MAST, NED, NRAO, CDS,
IVOA
- Support from Microsoft Research and the
Virtual Astronomical Observatory