

Astronomy 3.0 Style

Alberto Accomazzi
NASA Astrophysics Data System

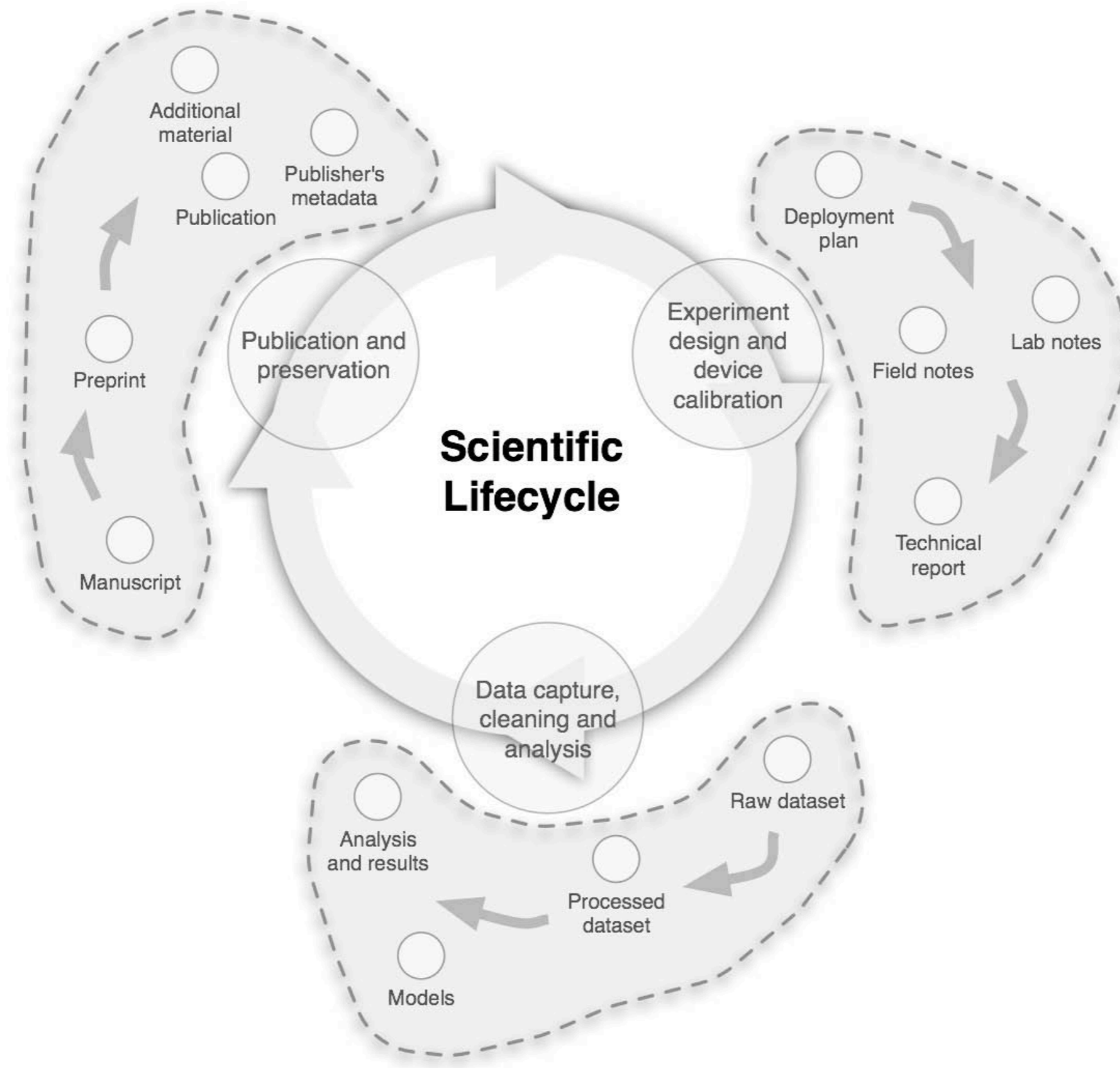
17 Feb 2010
Library and Information Services VI, Pune, India

Overview

- The Scientific Lifecycle model
- Resource metadata in Astronomy
- What is Web 3.0 and how it can help
- A proposal for interlinking Astronomy resources
- Ongoing work and future plans

The Big Picture

- 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 metadata throughout the process



Pepe et al, JASIST (2009)

**High-level
Data Products**

Paper

Report

Mission Design

**Observing
Proposals**

Simulations

**Scientific
Lifecycle**

Analysis

Observations

Data Reduction

Pepe et al, JASIST (2009)

Why do we care?

- attribution: links allow one to follow the evolution of research activities
- aggregation: links allow one to find related resources, ancillary data, versioning
- preservation: knowledge about the origin of metadata as important as the data
- discovery: the presence (or absence) of links carries semantic meaning

The case of Astronomy

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

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
- Links across domains (papers, data, objects)
- Provides many of the connections required by the science life-cycle model

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

Some numbers for ADS

- citations: 40M (across databases)
- readership: 18M (90-day window)
- fulltext: 5M (journals, arXiv, ADS)
- astronomical objects: 240K (SIMBAD + NED)
- data products: 130K
- bibliographic groups: 200K

Bibliographic groups

- CfA
- ESO
- ROSAT
- HST
- IUE
- CFHT
- Chandra
- ISO
- PhysEd
- XMM
- USNO
- NRAO
- Gemini
- Keck
- Spitzer

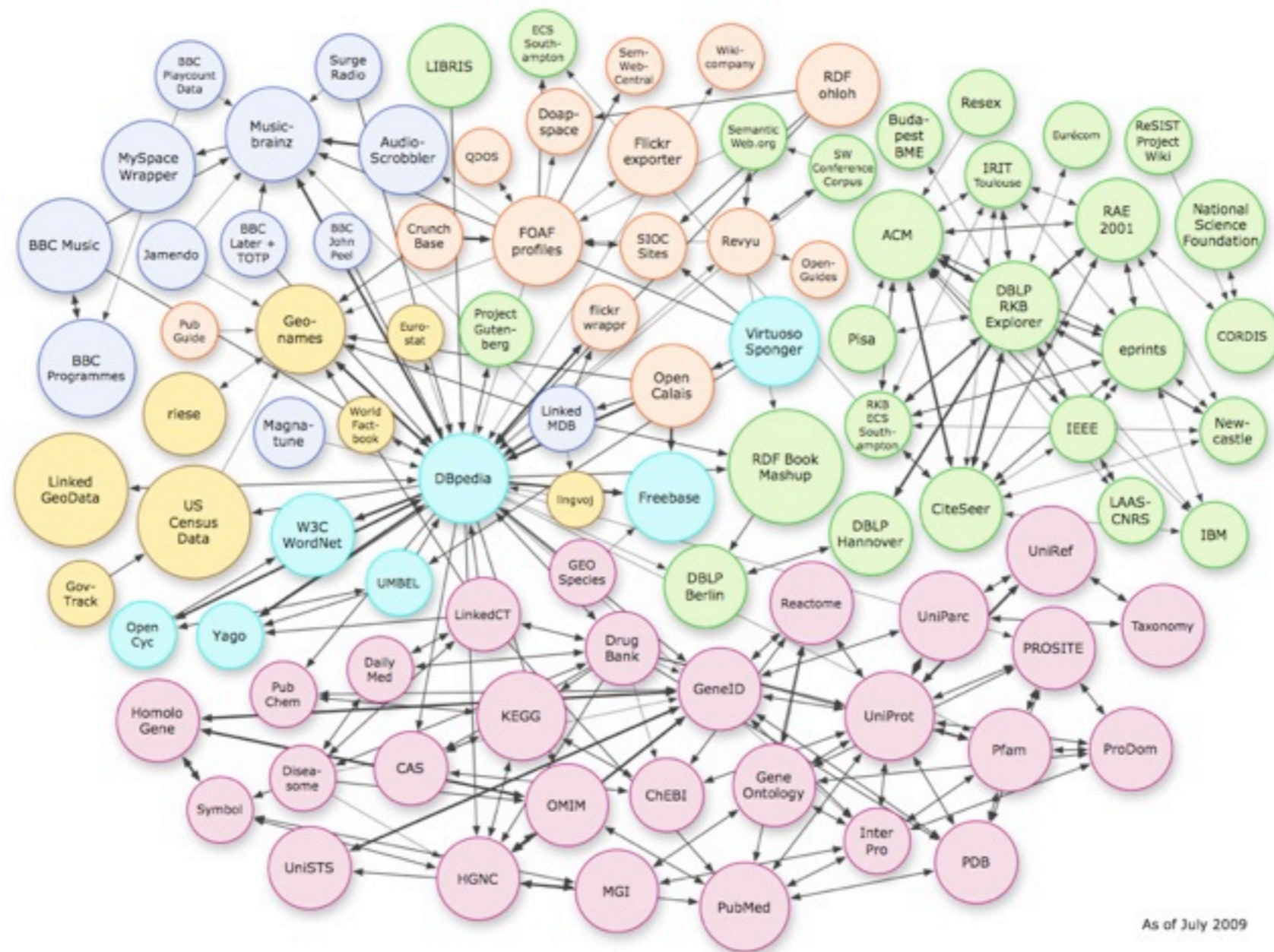
Web 3.0

- The “Web of Data” or “Semantic Web”
- Describes resources in a formal way
- Machine-readable content throughout
- Based on Linked Open Data principles
- “Read, Write, Execute”

Linked Data principles

- resources are named via HTTP *URIs*
- metadata is *open* and in a *standard format*
- links between resources are *typed*
- built on the architecture of the web, *no APIs*

Linked Data Cloud

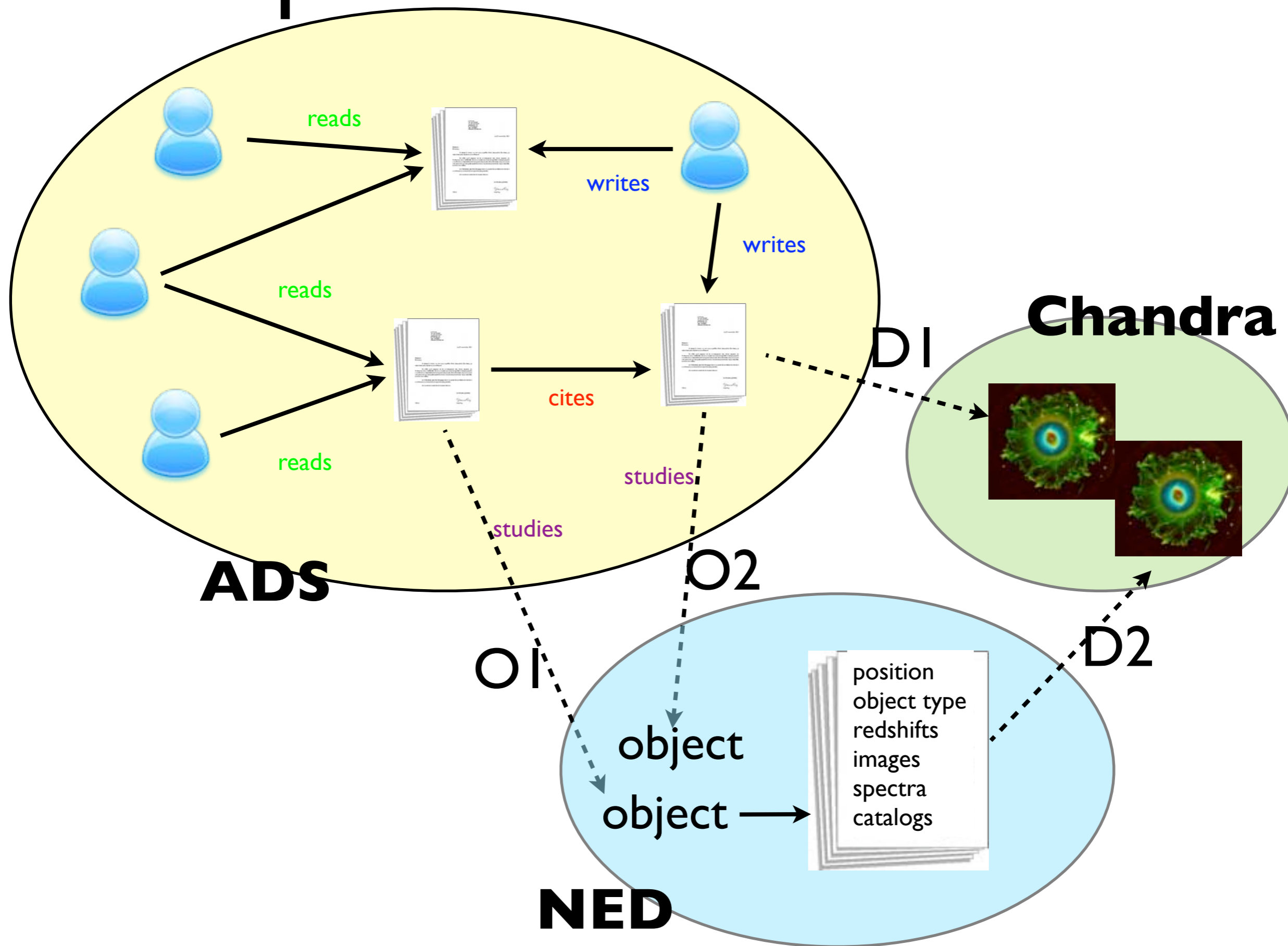


as of jan 2010: 142M links, 13B triples

Links to Datasets

- have been based on reciprocal links between URLs 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

Example: ADS/NED/Chandra



How to get there

- grow the astronomy data cloud from the bottom-up
- identify our resources (data), name them uniquely, expose their metadata
- expose relationships between resources (links)
- what should be included: observing proposals, observational metadata, instrumental metadata, papers, catalogs, objects

Publication Metadata

- Authors
- Keywords
- Affiliations
- Citations
- Links to data, objects
- Readership

Observational Metadata

- Unique Dataset Identifiers
- Position, footprint and date
- Objects observed
- Wavelengths
- Instruments
- People, Proposals, Collaborations

A VAO Proposal

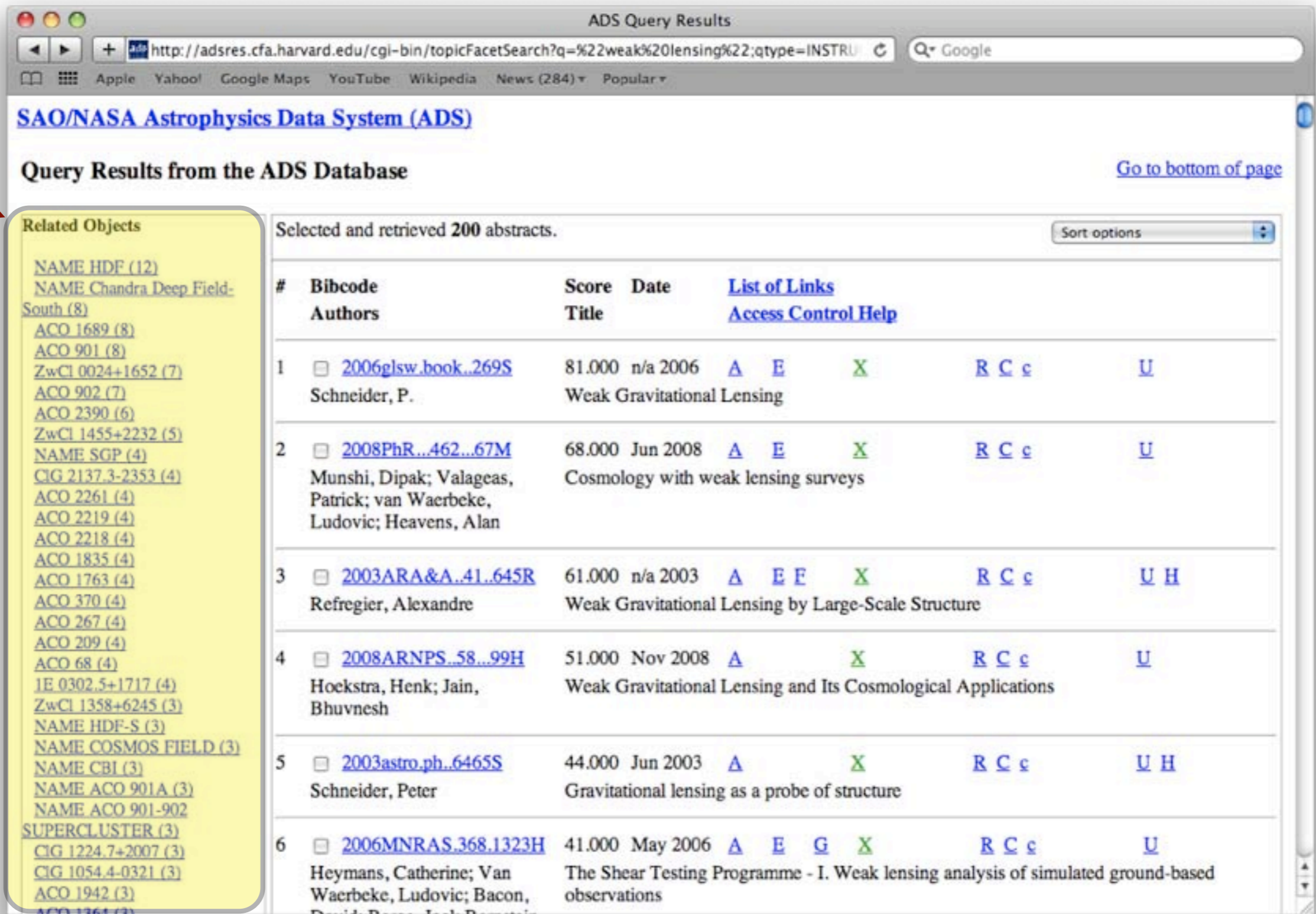
- Identify, collect and expose metadata for Datasets (Chandra, MAST, TSC, etc.)
- Create typed links between Data products, Bibliographic and Object databases
- In the process, create knowledge base about Instruments and Services
- Incorporate all of the above in a “metadata store,” exposing resources using semantics

Sample Applications

- A view of bibliographies with facets based on objects, wavelengths, keywords, etc
- A recommender / alert service for data products, objects
- A UI integrating views of bibliographies, objects, and datasets interacting together
- An application computing metrics on data, objects based on citation and use of papers

Views on Papers

object
facets



The screenshot shows the ADS Query Results page for the query "%22weak%20lensing%22;qtype=INSTRU". The page displays a list of 200 abstracts. A sidebar on the left, titled "Related Objects", lists various astronomical objects and their counts, such as "NAME HDF (12)", "NAME Chandra Deep Field-South (8)", "ACO 1689 (8)", "ACO 901 (8)", "ZwCl 0024+1652 (7)", "ACO 902 (7)", "ACO 2390 (6)", "ZwCl 1455+2232 (5)", "NAME SGP (4)", "ClG 2137.3-2353 (4)", "ACO 2261 (4)", "ACO 2219 (4)", "ACO 2218 (4)", "ACO 1835 (4)", "ACO 1763 (4)", "ACO 370 (4)", "ACO 267 (4)", "ACO 209 (4)", "ACO 68 (4)", "1E 0302.5+1717 (4)", "ZwCl 1358+6245 (3)", "NAME HDF-S (3)", "NAME COSMOS FIELD (3)", "NAME CBI (3)", "NAME ACO 901A (3)", "NAME ACO 901-902 (3)", "SUPERCLUSTER (3)", "ClG 1224.7+2007 (3)", "ClG 1054.4-0321 (3)", "ACO 1942 (3)", and "ACO 1361 (3)".

The main table of results is as follows:

#	Bibcode	Score	Date	List of Links	Access Control Help
1	<input type="checkbox"/> 2006glsw.book..269S Schneider, P.	81.000	n/a 2006	A E X	R C c U
2	<input type="checkbox"/> 2008PhR...462...67M Munshi, Dipak; Valageas, Patrick; van Waerbeke, Ludovic; Heavens, Alan	68.000	Jun 2008	A E X	R C c U
3	<input type="checkbox"/> 2003ARA&A..41L..645R Refregier, Alexandre	61.000	n/a 2003	A E E X	R C c U H
4	<input type="checkbox"/> 2008ARNPS..58...99H Hoekstra, Henk; Jain, Bhuvnesh	51.000	Nov 2008	A X	R C c U
5	<input type="checkbox"/> 2003astro.ph..6465S Schneider, Peter	44.000	Jun 2003	A X	R C c U H
6	<input type="checkbox"/> 2006MNRAS.368.1323H Heymans, Catherine; Van Waerbeke, Ludovic; Bacon, David; Besci, Jean-Benoit	41.000	May 2006	A E G X	R C c U

Recommendations?

Collaborations

My data

Recent data

Hottest data

myADS Personal Notification Service
for Alberto Accomazzi
Tue Nov 10 04:57:53 2009
Astronomy database

ADS Main Queries

- [Astronomy](#)
- [Physics](#)
- [arXiv e-prints](#)
- [FAQ](#)
- [What's new](#)

Current Tables of Contents

- [Astronomical Journal](#)
- [Astronomy & Astrophysics](#)
- [Astrophysical Journal](#)
- [Astrophysical Journal Letters](#)
- [Monthly Notices of the Royal Astronomical Society](#)
- [Publications of the Astronomical Society of Japan](#)
- [Recent Astro-PH](#)

Search Recent Papers

Author
Subject

ACCOMAZZI, ALBERTO - Citations: 223 (total 242)

- [2010ARIST..44....3K](#): Kurtz,+; Usage Bibliometrics
- [2009MNRAS.399.2165R](#): Ramirez,+; The nature of the near-infrared core source in 3C 433
- [2009MNRAS.399.1622G](#): Goyal,+; Unusual optical quiescence of the classical BL Lac object PKS 0735+178 on intranight time-scale
- [2009ASPC..411..384H](#): Henneken,+; Exploring the Astronomy Literature Landscape
- [2009IISys..24....74C](#): Chen,+; Mapping the Sloan Digital Sky Survey's Global Impact

Favorite Authors - Recent Papers

- [2009Ap&SS.322..101R](#): Rosen,+; Laboratory experiments to study supersonic astrophysical flows interacting with clumpy environments

"VIRTUAL OBSERVATORY", etc - Recent Papers

- [2009Obs...129..397P](#): Prema; Galaxy formation and evolution using the Virtual Observatory
- [2009LNP...791...81W](#): Walton,+; AstroGrid and the Virtual Observatory
- [2009A&A...506..455S](#): Solano,+; The LAEX and NASA portals for CoRoT public data
- [2009MNRAS.399.2231S](#): Scaringi,+; Classifying broad absorption line quasars: metrics, issues and a new catalogue constructed from SDSS DR5
- [2009SoPh...155U](#): Ulrich,+; Modeling Total Solar Irradiance Variations Using Automated Classification Software on Mount Wilson Data

"VIRTUAL OBSERVATORY", etc - Most Popular

- [2009ASPC..411..101S](#): Seaman,+; Optimal DN Encoding

(ASTROPHYSICS DATA SYSTEM, etc - Recent Papers)

- [2009amos.confE..87S](#): Schmalz,+; Comparison of Neural Networks and Tabular Nearest Neighbor Encoding for Hyperspectral Signature Classification in Unresolved Object Detection
- [2009MNRAS.399.2231S](#): Scaringi,+; Classifying broad absorption line quasars: metrics, issues and a new catalogue constructed from SDSS DR5

(ASTROPHYSICS DATA SYSTEM, etc - Most Popular)

- [2008ApJ...674..768O](#): Oyaizu,+; A Galaxy Photometric Redshift Catalog for the Sloan Digital Sky Survey Data Release 6
- [2009NewA...14..649B](#): Bora,+; Automated star-galaxy segregation using spectral and integrated band data for TAUVE/ASTROSAT satellite data pipeline
- [2006MNRAS.373..845B](#): Ball,+; Bivariate galaxy luminosity functions in the Sloan Digital Sky Survey
- [2009MNRAS.392..233Z](#): Zhang,+; Morphology classification and photometric redshift measurement of galaxies

[Update myADS Settings](#)

APOD Application

The screenshot shows a web browser window with the URL `http://dabrowser.semantic.ads.org`. The search bar contains the text "lmc" and a "Go" button. The breadcrumb trail is "All > Query1 > Query2 > Query3: lmc".

The interface is divided into three main panels, each with a "List" and "SpaceTime" view toggle and an "ON" status indicator:

- Datasets:** Displays a 2D map with a yellow line and a green shaded area. Labels include "Macho 1367", "OGLE lmc_scl", "Object One", "Object Two", and "Object Three". A "Time" axis is shown at the bottom left with a "Download?" link.
- Publications:** Lists search results with links for "Pivot, Details, More":
 - Macho LMC catalog etc etc
 - Microlensing in the LMC
 - Ogle Variables in the LMC
 - ...
- Astronomical Objects:** Lists search results with links for "Pivot, Details, More":
 - Object One
 - Object Two
 - Object Three
 - ...A "FILTER by facet" menu is open, showing options: Variability, Wavelength (highlighted), Optical, Xray, Shape, Lightcurves?, and Spectra?.

created with Batsamig Montage - www.batsamig.com

Ongoing Efforts

- IVOA / ADEC dataset identifiers
- IVOA data models, vocabularies for keywords and UCDs, ontology of object types
- MSR / WWT work on mashups
- ADS work with VAO, text mining, faceted views

The Role of Libraries

- Astronomy librarians have been at the intersection of archives and journals
- They are experts in metadata curation and search, maintain bibliography-data links
- They have been collecting/generating much of this metadata already
- They understand the need for (and limitations of) metrics

A call to action

- Tell us what bibliographic and observational metadata you are collecting and curating
- Use our upcoming fulltext search on current literature to collect/annotate papers
- Drink the Web 3.0 coolaid
- We have freed our data, now let's free your metadata!