An Iterative Approach to **Library Data Services in Astronomy**

NEASIST | C. Erdmann | @libcce

History background

→ Learned how to program

→ Joined a dotcom

→ Data mining, data linking, faceting

→ Became data savvy

Being able to program has opened the door to new possibilities for me in the library world.

http://i.livescience.com/images/i/000/027/364/i02/vlt-brunier-nuit-1600.jpg?1337613157

European Southern Observatory (ESO)

- Multinational organization
- Headquarters in Munich
- Ground-based telescopes

Telescope Bibliography (telbib): workflow

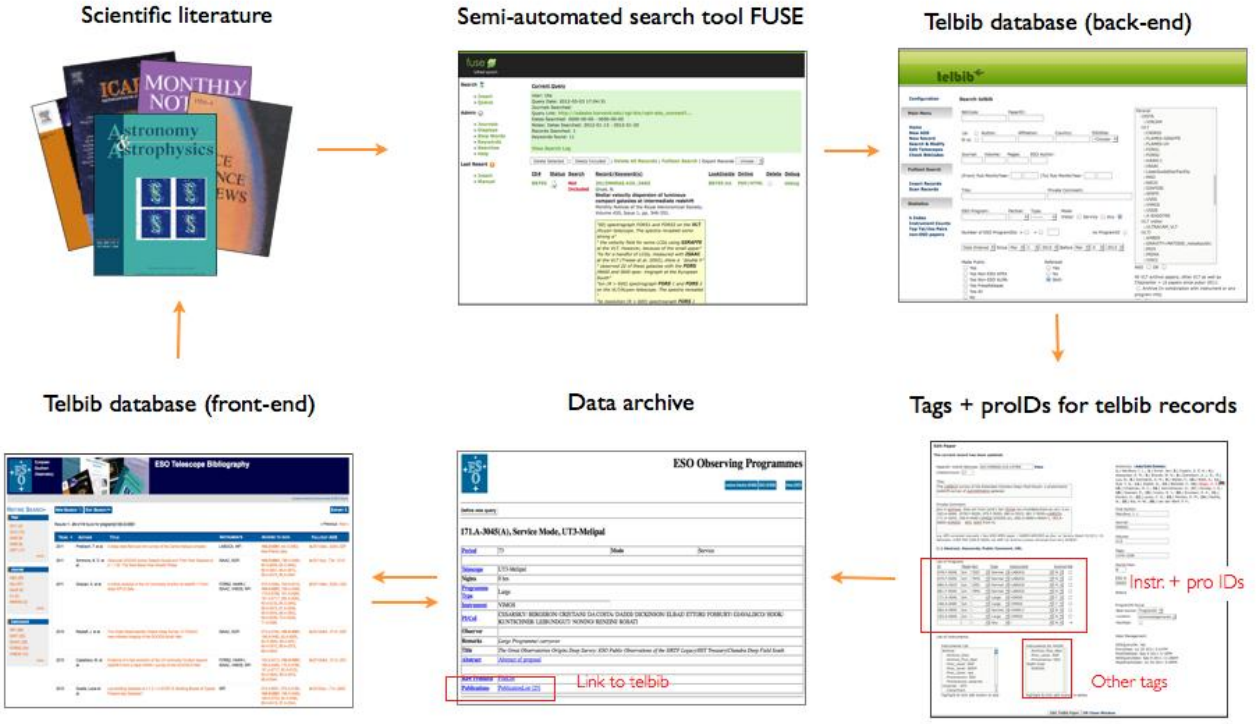


Image courtesy U. Grothkopf, ESO Library

Limit your search

- Top papers
- Authors
- Database
- Keywords
- Publications
- Refereed status
- Bib Groups
- Grants
- Data apply
  - CDS (409)
  - NED (309)
  - MAST (48)
  - HST (44)
  - ESA (43)
  - more...
- Vizier Tables
- Publication Year
  - from 1972 to 2013 apply

Database : astronomy OR physics Clear all

‹ Previous | 1 to 20 of 931 | Next › Analyze View Export Sort

1. 2013arXiv1312.7417W [X R]
Measuring the X-ray luminosities of SDSS DR7 clusters from RASS
Wang, Lei; Yang, Xiaohu; Shen, Shiyin; Mo, H. J. *and 7 coauthors*
Published in Dec 2013
... noticeably the ***CfA redshift survey*** (e.g.Geller & Huchra 1983), the Las Campanas Redshift Survey (e.g.Tucker ...

2. 2013arXiv1312.4294S [X R]
AGN surveys to study galaxy evolution along cosmic times
Spinoglio, Luigi
Published in Dec 2013
... selected sample of Seyfertgalaxies (25 Seyfert 1 and 23 Seyfert 2) has been extracted from the ***CfA Redshift*** ...
... ***Surveyof*** ~2400 objects (1/50 or 2% efficiency) (Huchra & Burg 1992). Veilleux & Osterbrock1987 proposed ...

3. 2013PASJ...65..113T [E F X R]
The 9 and 18 Micrometer Luminosity Functions of Various Types of Galaxies with AKARI: Implication for the Dust Torus Structure of AGN
Toba, Yoshiki; Oyabu, Shinki; Matsuhara, Hideo; Malkan, Matthew A. *and 5 coauthors*
Published in Oct 2013
... and the Centerfor Astrophysics ***(CfA) redshift survey*** (ZCAT: Huchra et al.1995). The flow chart used for sample ...

4. 2013MNRAS.434.2556W Cited by 3 [E F X R C]
The impact of systematic uncertainties in N-body simulations on the precision cosmology from galaxy clustering: a halo model approach
Wu, Hao-Yi; Huterer, Dragan
Published in Sep 2013
... Peebles 1977) and the ***CfA Redshift Survey*** (Huchra et al. 1983; Geller Huchra 1989) revealing the cosmic ...

5. 2013arXiv1309.7786S [X R]
Galaxy structures - Groups, clusters and superclusters
Sepp, Tiit; Gramann, Mirt
Published in Sep 2013

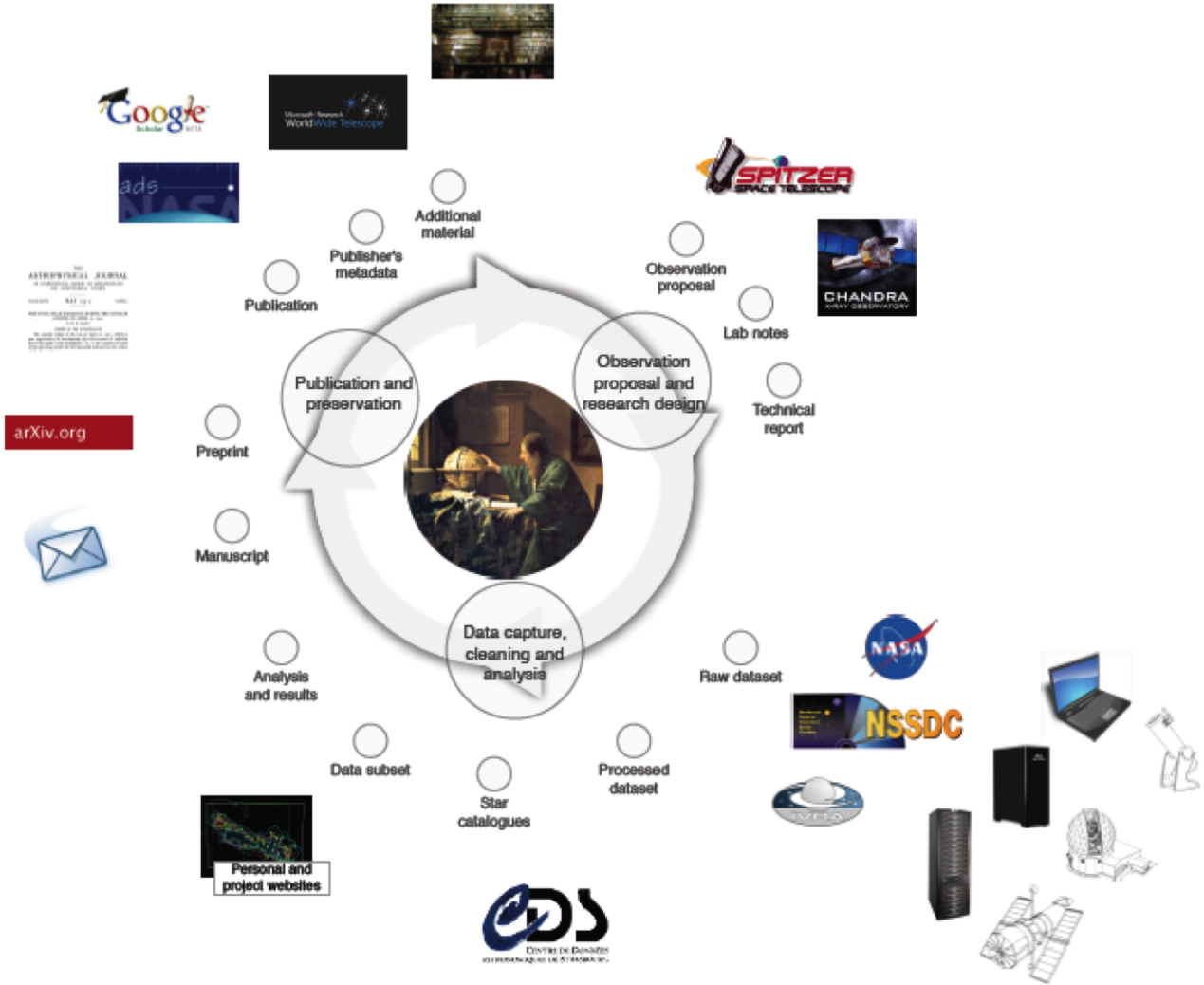


Image courtesy A. Pepe, Harvard-Smithsonian Center for Astrophysics

http://commons.wikimedia.org/wiki/File:Center_for_Astrophysics_at_Harvard.jpg

Harvard-Smithsonian Center for Astrophysics (CfA)

- Collaboration between institutions
- Based in Cambridge, MA
- Multinational projects
- Global & space-based facilities

Administrative Services

Policies & Copyright Advice

DMPs & DMPTool

DM Training Programs

DM @ Harvard Site

Research Data Collaborative

DataCite

E-Science Institute

WH OSTP Response

Survey: Story of Your Data?

Publishing Research Material

- Zenodo
- Dataverse
- Figshare
- Datahub

Lessons Learned

Low Barrier
User Experience
Author Workflow Integration
Social & Sharing Aspect
Citation Ease
Permissions & Copyright
Paper <-> Data Linking
Selling the Service

Objective Advice
Visual Customization (APIs)
Metadata Customization
Traditional vs Data Savvy Approach
Repository-centric
Tracking Work
Curation Advice
Learning Curve
Versioning

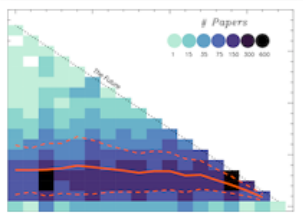
What If We Could Be More?

If We Assume

[About] [Contact] [Press] [Mock Twain]

The Pace of NSF Funded Research

Topics: academia, Astronomy, costs, statistics



Recently on Facebook I came across a note by Chris Erdmann that some handy folks at Harvard put together statistics on (nearly) every astronomy paper from 1995 to present that was funded through an NSF AST grant. This seemed like a really interesting dataset, especially for a young (read: financially uncertain) research such as myself.

So parsing through all 29,042 papers listed, here are two interesting things I've learned...

Could we be of more use to scientists?

Sympathize with their data needs and offer assistance?

Become data savvy?

Clear Need

"By 2018, the United States alone could face a **shortage** of 140,000 to 190,000 people with deep analytical skills as well as 1.5 million managers and analysts with the know-how to use the analysis of big data to make effective decisions."

"Online help-wanted ads for data analysis mavens have **shot up** 46% since April 2011, and 246% since April 2009, to over 31,000 openings now."

[1] http://www.mckinsey.com/insights/business_technology/big_data_the_next_frontier_for_innovation

[2] http://management.fortune.cnn.com/2013/05/10/big-data-jobs/

Data Scientist Training for Librarians (DST4L)

Major goals:

- Hands on experience w/ research data lifecycle
- Inform new forms of library (data) services
- Upgrade librarian knowledge & skills
- Create a community of continuous learning by doing
- Experience new style of working
- Change library mindset
- Explore other learning moments

Course Outline

Extract

Wrangle

Analyze

Visualize

Technologies

Command Line
Git (GitHub)
Excel
OpenRefine
Python
Software Carpentry
R (RStudio)
Tableau
SQL (SQLShare)
MongoDB
Data Repositories
D3

More about the Course

1 class/week 3-4 months

    * 1 hands on session (projects)

Lead organizer (context) w/ expert instructors

Open notes, participants blog (experience)

In-person, not virtual or recorded

Some homework, final presentations

Response

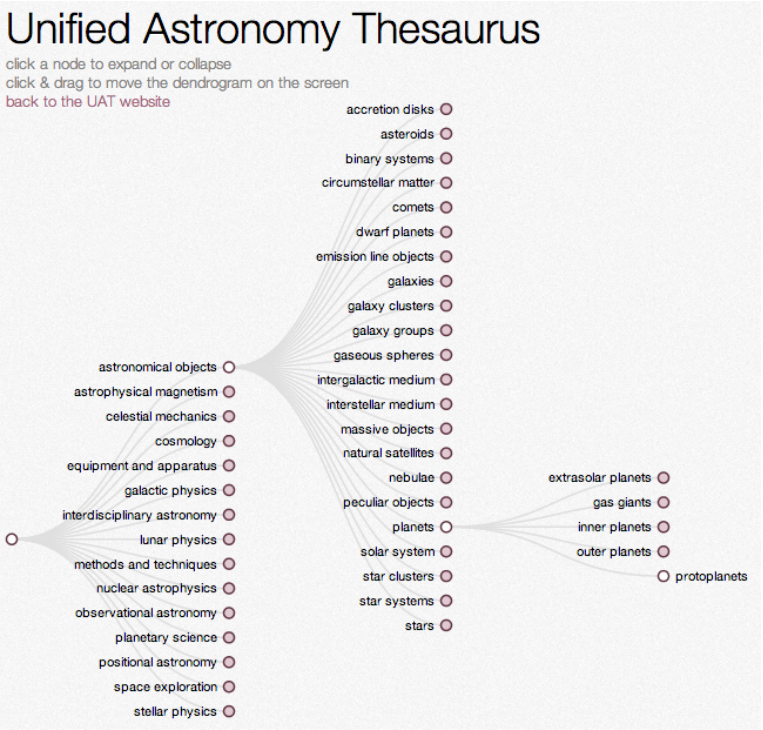
Highlighted student comment:

http://www.youtube.com/watch?v=U5ZYM085bNo&t=1m21s

...very helpful, preparing for long career, going to see it more and more, will keep using skills set, like doing it, fun problem solving...

It’s Hard Work!

But worth it!



- NASA ADS Visualization
- Unified Astronomy Thesaurus
- CfA Bibliography
- DOE Grants

How Do I Start?

Software Carpentry Workshop

Invite Local Experts

Tap Your Own Community

A Challenge

If you dislike change, you're going to dislike irrelevance even more.

-- Eric Shinseki

Questions?