

# What's New with NED?

The NED Team/Rick Ebert

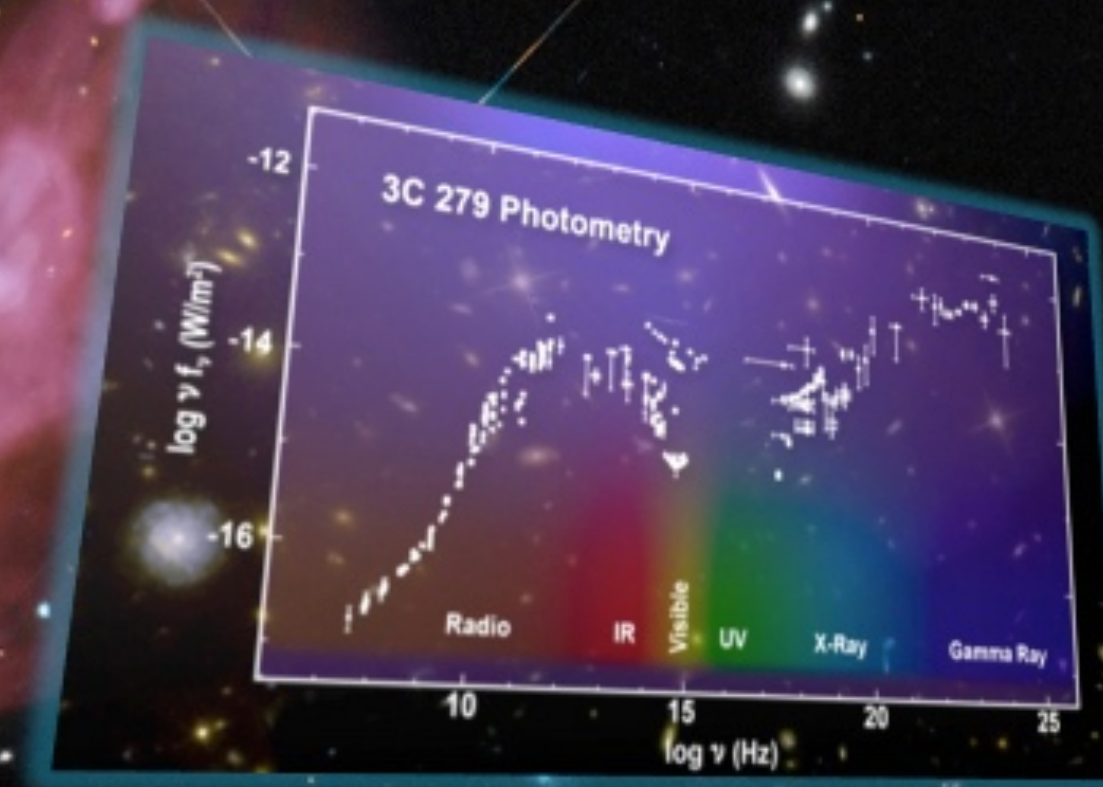
IVOA-Nov2018

# Scope

## Maintaining the Census of the Extragalactic Universe

- ❖ NED is the world's largest database of cross-correlated multi-wavelength data for extragalactic objects;
- ❖ NED is a synthesis of extragalactic data spanning all NASA astrophysics missions, large sky surveys and the peer-reviewed literature;
- ❖ NED is the first port-of-call for most extragalactic researchers; and is integral to all phases of research;
- ❖ NED plays a key role in IVOA, US-NASA, US-NSF and international efforts to construct a collaborative and distributed Virtual Observatory

NASA/IPAC Extragalactic Database



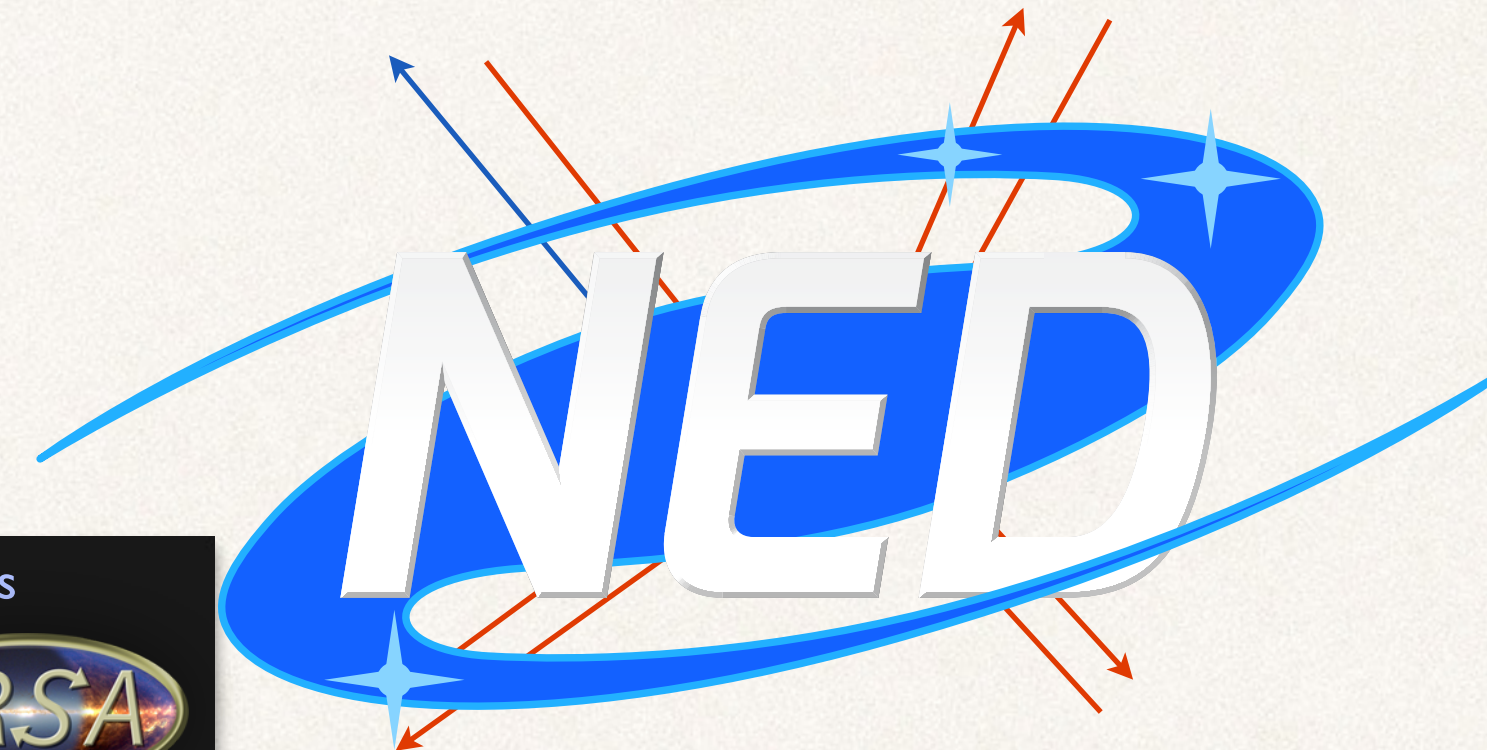
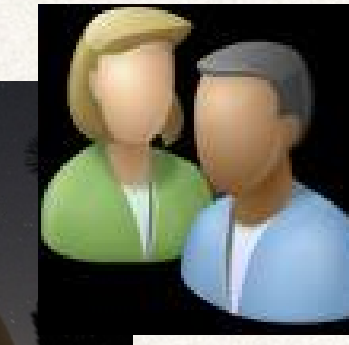
# Community Ecosystem

Individual and Team Researchers

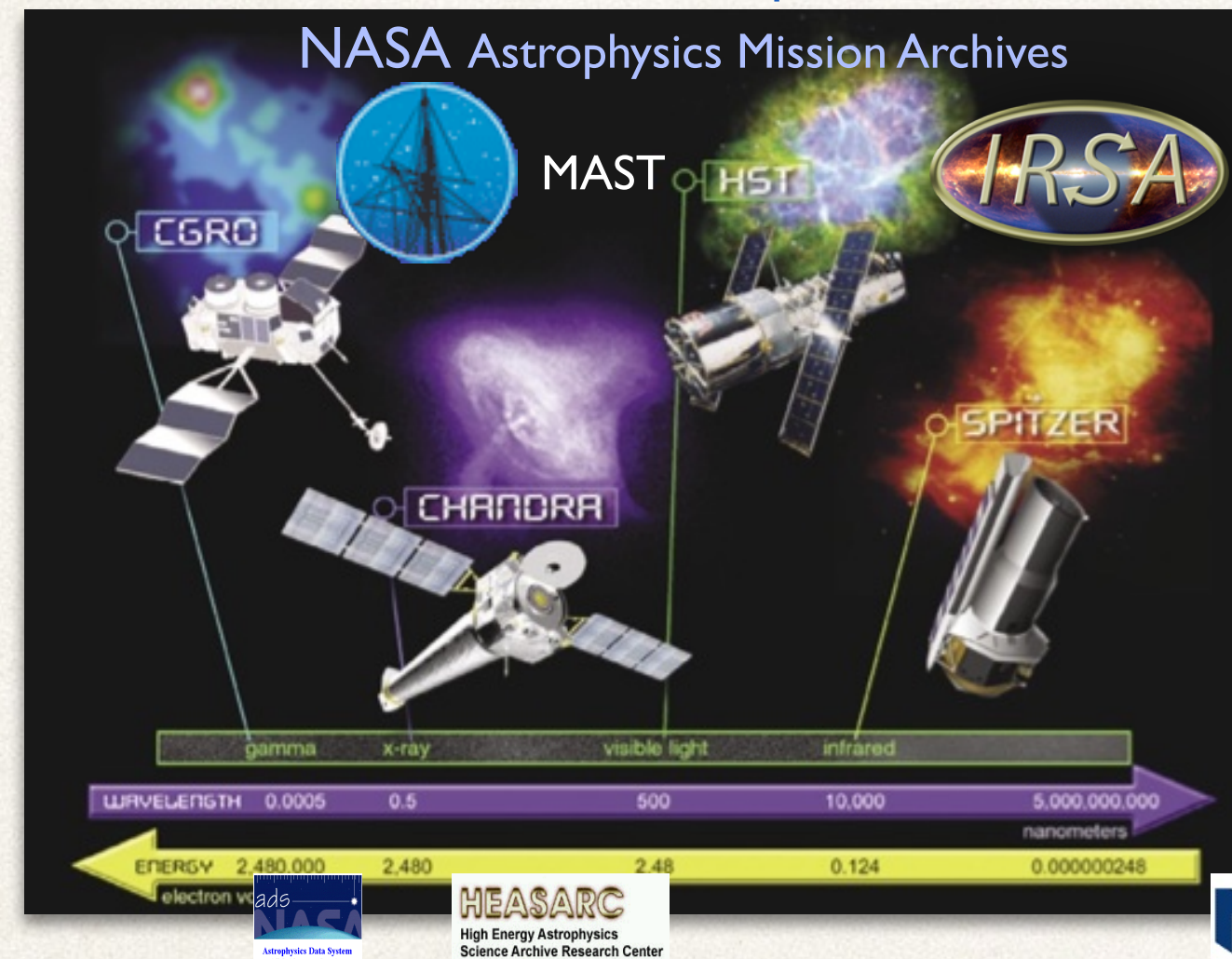
Astrophysics Literature



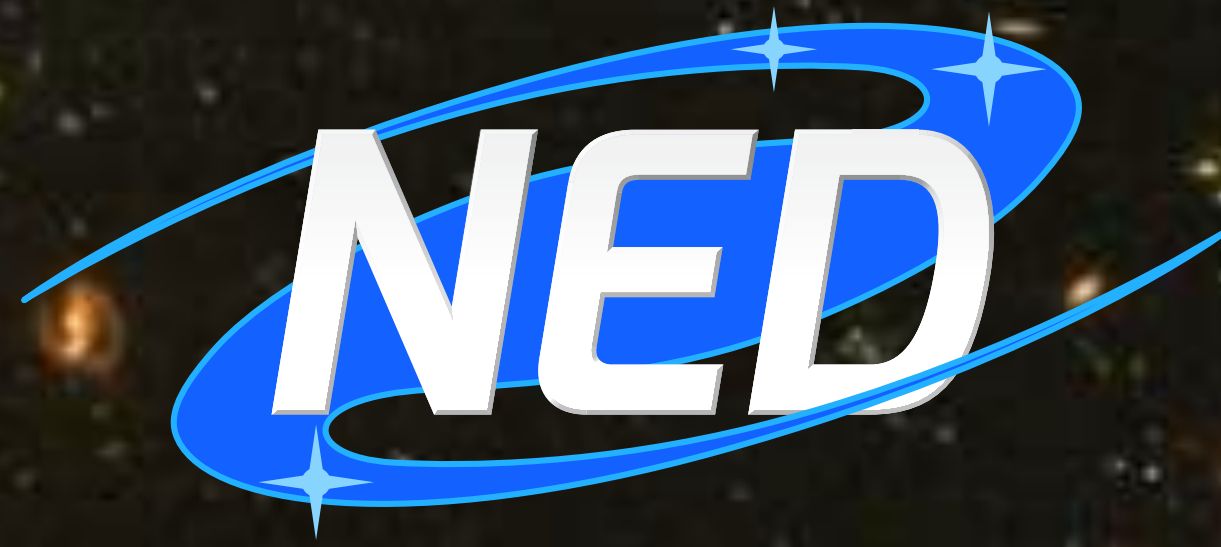
Objects,  
Measurements,  
Attributes, Images,  
Spectra



- ❖ Individual and Team Researchers
- ❖ Interactive & Automated Queries
- ❖ Data Analysis & Discovery
- ❖ Publication



Sky surveys, present & future



# What's in it?

## Published:

- Names
- $(\alpha, \delta)$
- Redshifts
- $D_{\text{Mpc}}$
- Fluxes
- Sizes
- Attributes
- References
- Notes

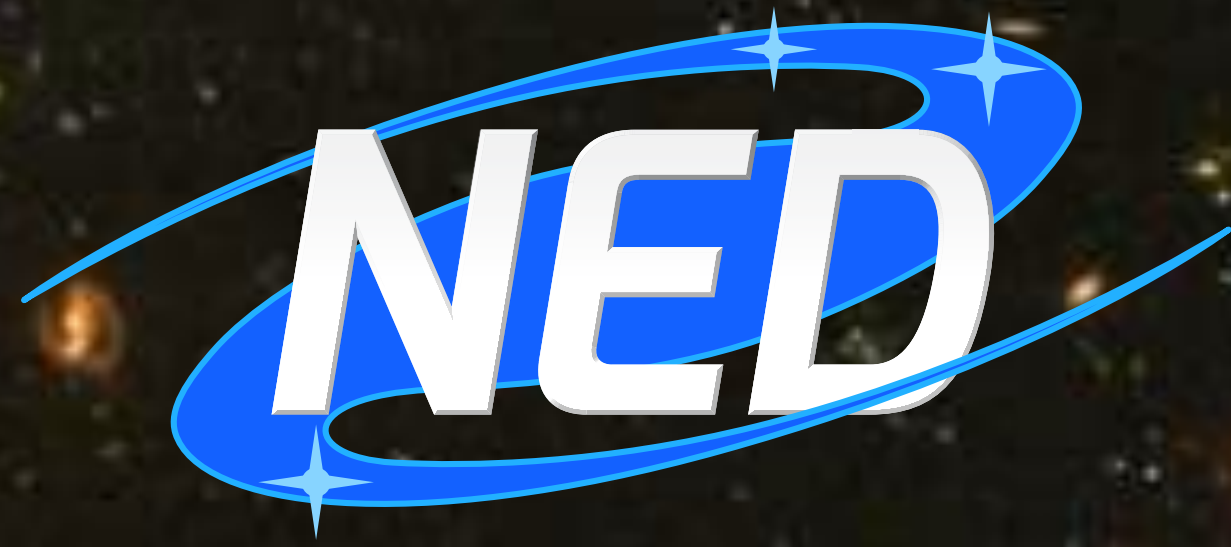
## Contributed:

- Images
- Spectra

## Derived:

- Distances
- Metric sizes
- Luminosities
- Velocity corrections
- Cosmological corrections
- SEDs
- $A_{\lambda}$





# The Team



*NED Team: Ben Chan, Tracy Chen, Cren Frayer, George Helou, Scott Terek, Rick Ebert, Tak Lo, Barry Madore, Jeff Jacobson, Joe Mazzarella, Olga Pevunova, Marion Schmitz, Ian Steer*

Prior members: contributors: Shubash Atreya, Kay Baker, Judy Bennett, Michael Bicay, Marianne Brouty, Harold Corwin, Dario Fadda, Jimmy Jia, Anne Kelly, Dong-Chan Kim, Sarita Kuhn, Cheryl LaGue, Patrick Ogle, Diana Schettini, Hongmin Sun, Xiuqin Wu

# Changes

---

- ❖ Platform Solaris86, Linux & Postgresql (from Solaris sparc, Informix)
- ❖ Very Large Catalog Pipeline (>10M entries... 100M)
- ❖ User Interface : HTML1 -> HTML 5.0 & Content Delivery System; IPAC Firefly integration
- ❖ “Metric” Distances
- ❖ VO IVOA services: SIA\_2.0, TAP
- ❖ Other services: Classic “VO” services; SED (SSA-like), ObjectLookup (new name resolver)
- ❖ Challenges:
  - ❖ New Tech & Big(ger) Data



# The User Interface

Grab File Edit Capture Window Help

NASA/IPAC Extragalactic Database - NED

http://ned.ipac.caltech.edu/

**NASA/IPAC EXTRAGALACTIC DATABASE**

**News & Featured Updates (April, 2012)**

- Over 1 million objects and 1.6 million cross-IDs from COSMOS, SUMSS, 2XMM, etc.
- Over 490,000 photometric measurements integrated into SEDs
- Over 127,000 spectra from the 6dF Galaxy Survey (DR3)
- Additions to Level 5, including *Tides in Colliding Galaxies* (Duc & Renaud 2011)
- Server Mode (nedsrv) to be replaced in 2012 with web services

OBJECTS	DATA	LITERATURE	TOOLS	INFO
By Name	Images By Object Name or By Region	References by Object Name	Coordinate Transformation & Extinction Calculator	Introduction Latest News/Updates
Near Name	Photometry & SEDs	References by Author Name	Velocity Calculator	Features FAQ
Near Position	Spectra	Text Search	Cosmology Calculators	Overview (pdf)
IAU Format	Redshifts	Knowledgebase	Extinction-Law Calculators	Source List
By Parameters (All-Sky)	Redshift-Independent Distances	Galaxy Distance Tabulations (NED-D)	Skymplot	Web Links
By Classifications Types, Attributes	Classifications by Object Name	Abstracts	X/Y offset to RA/DEC	Glossary & Lexicon
By Refcode	Positions	Thesis Abstracts	Batch Job Submission Pick Up Results	Team
Object Notes	Diameters		NEW Build Data Table from Input List By Name Near Name/Position (Cross-Matching)	Contact Us or Comment

If your research benefits from the use of NED, we would appreciate the following acknowledgement in your paper: *This research has made use of the NASA/IPAC Extragalactic Database (NED) which is operated by the Jet Propulsion Laboratory, California Institute of Technology, under contract with*

ipac

**NED** NASA/IPAC Extragalactic Database

Home Search Objects » Tools » Information » Classic Services »

Object Name, coordinates with search radius, etc.

**July 2018 Release News**

- The all-sky *2MASS Point Source Catalog* has been cross-matched with NED. The result of cross-matching all 471 million sources is now available; 57 million sources (12%) match with prior NED objects, and their J, H, and  $K_s$  photometric measurements have been merged into SEDs.
- 47 million objects from additional catalogs and journal articles have been cross-identified.
- 1.2 million object links to 3,248 new references added
- 1,064 spectra and 84 images contributed by authors
- For more information see [Release Notes](#) and [Holdings](#).

**The new face of NED!**

We continue to update this modern, streamlined interface that simplifies and consolidates query forms and provides interactive tables with linked graphics. Key interface features are described here: [June 2018 release](#).

For additional information and known issues, see [Information](#) → [Overview](#) → [News](#). We welcome your feedback.

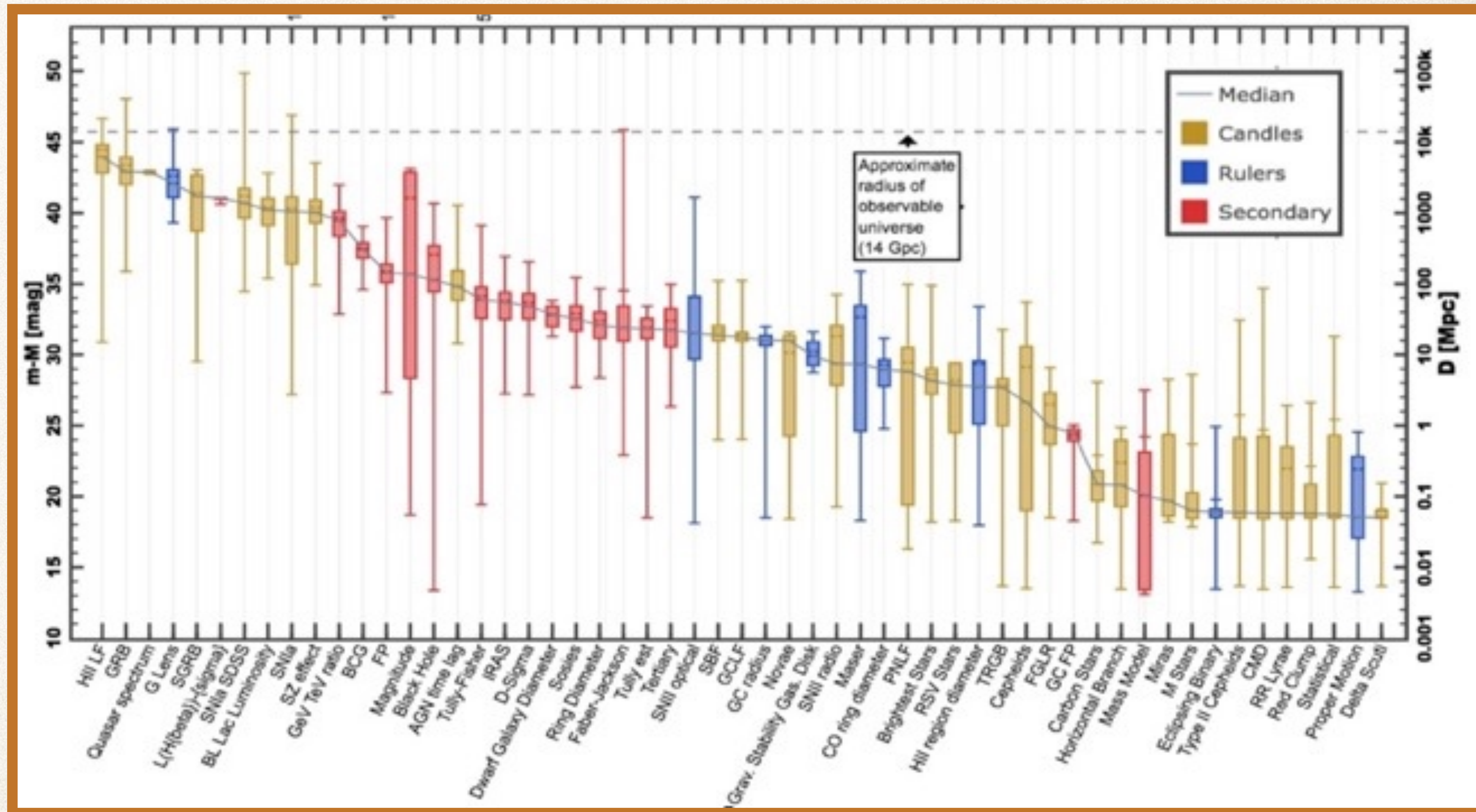
**Phasing out the classic web interface**

While we are completing the transition to the new user interface, quick access to the [original panel menu](#) remains available under the [Classic Services](#) menu option above.

Please see [Information](#) → [Overview](#) → [News](#) for more information.

*“Redshift-Independent Distances in the  
NASA/IPAC Extragalactic Database:  
Methodology, Content, and Use of NED-D”*

Steer, I, et al., *Astronomical Journal*, Volume  
153, Issue 1, article id. 37, 20 pp. (2017).







# Advances: cross-matching sources

Algorithm-based probabilistic cross-matching process codifying two decades of experience: *MatchEx*

Local density of objects is used to estimate the background contamination rate

Poisson statistics are used to balance completeness vs. reliability of matches

Scientific vetting is applied to refine algorithms

Ongoing enhancements will go beyond proximity:  $z$ , object classifications, sizes, fluxes

**Search radius:**  $r_s > 3.5\sigma$  (99.8% enclosed Gaussian probability) = 7.5" for GASC

**Background radius:**  $r_b \geq 6.2r_s$  (radius of 90% Poisson probability for  $N=1$ ) = 46.5" for GASC

**Density of background NED objects:**  $n = N/(\pi r_b^2)$

**Expected number:**  $\langle N_s \rangle = n\pi s^2 = N(s/r_b)^2$ ,  $s$  = source-object separation

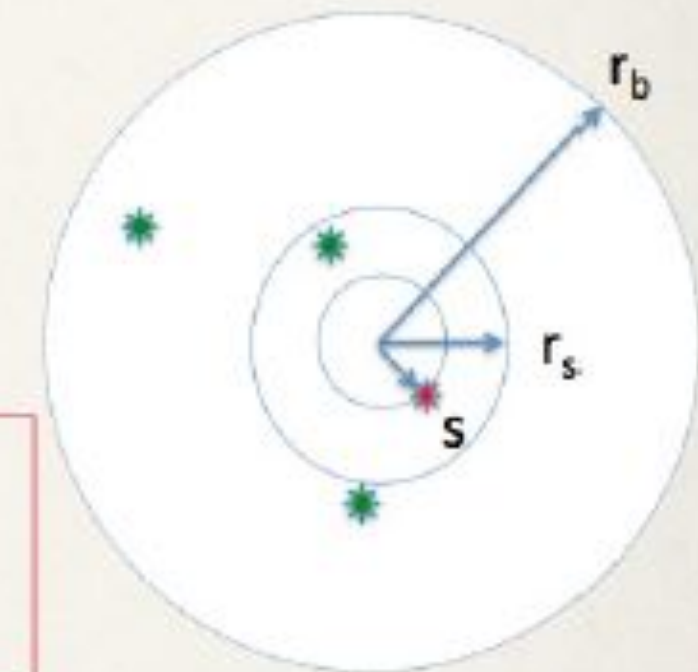
**Poisson probability** of  $x = 0$  NED objects closer to source than  $s$ :

$$P_s(x=k) = \langle N_s \rangle^k \exp(-\langle N_s \rangle) / k!$$

Simplifies to:

$$P_s(x=0) = \exp(-\langle N_s \rangle) = \exp(-N(s/r_b)^2)$$

**False-match probability:**  $P_f = 1 - P_s(0)$



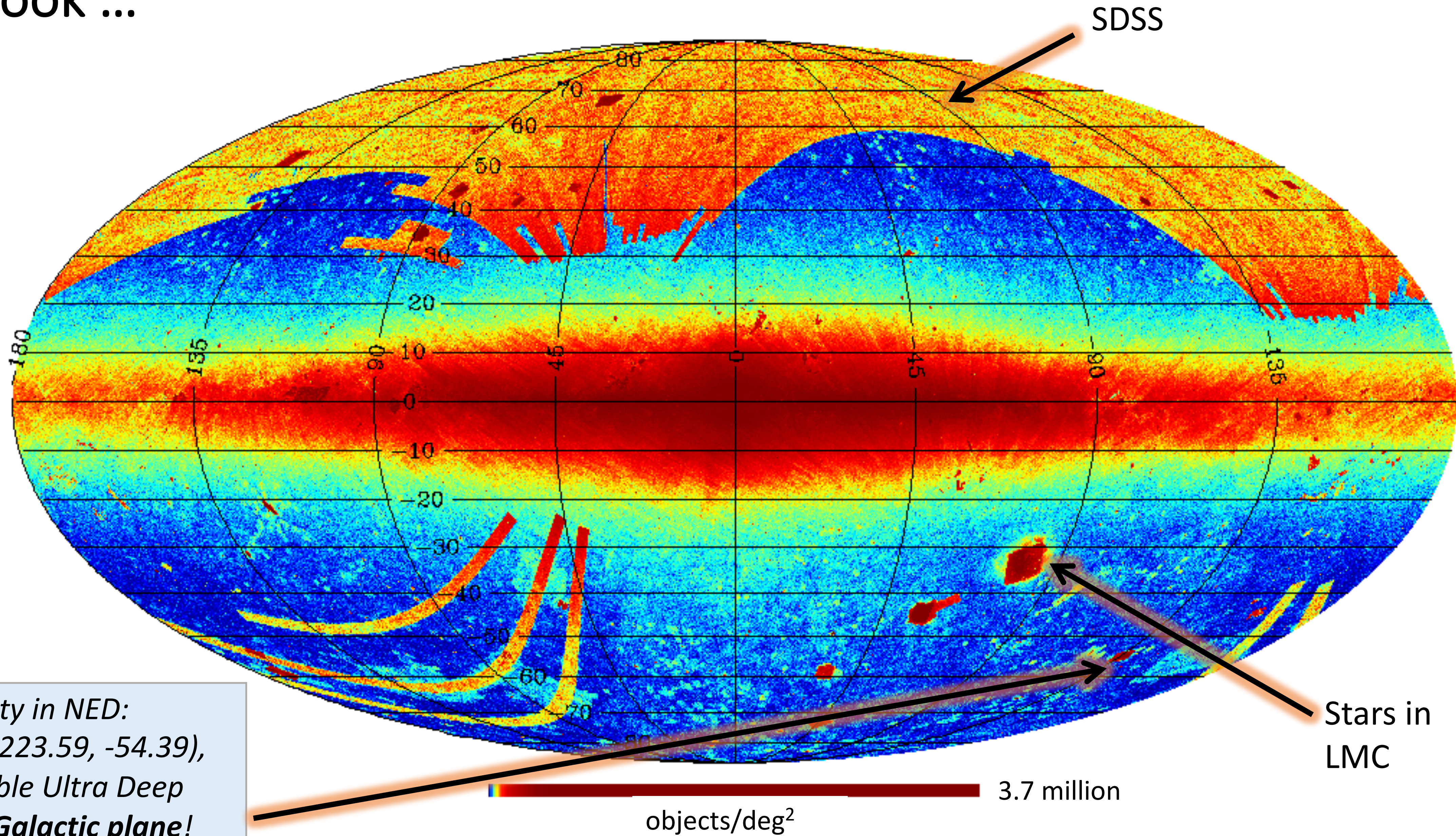
Example:  
 $N = 4$ ,  $s/r_b = 0.08$   
 $P_s(0) = 0.975$   
 $P_f = 0.025$

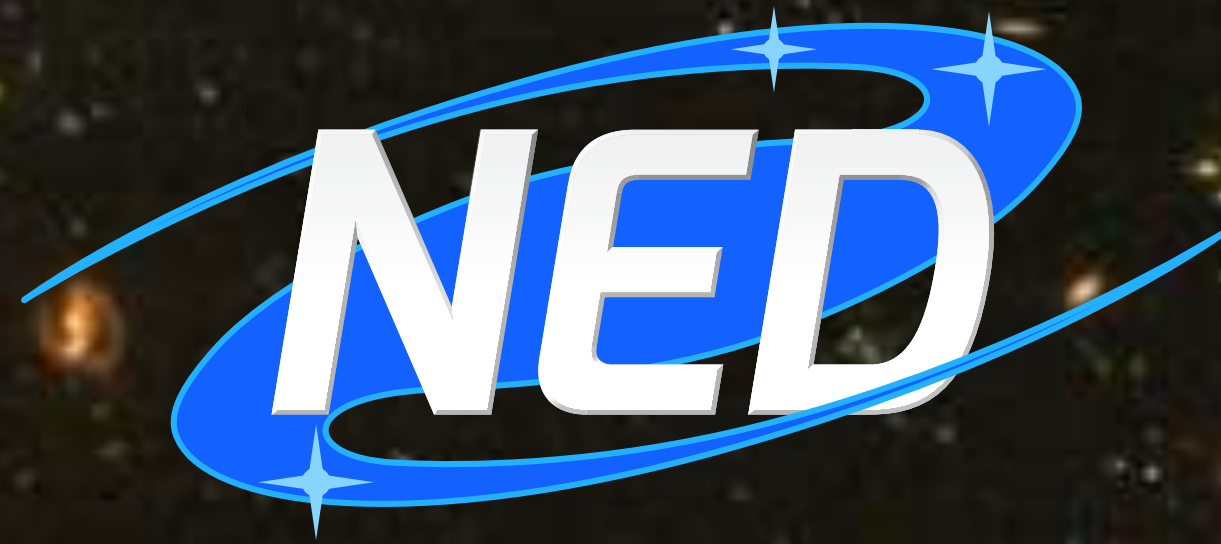
*Example: 2 candidate NED objects were found within  $r_s = 7.5''$  search radius for GALEX ASC, and 2 objects are in the background annulus ( $r_b$ ). Probability of finding no NED objects closer to the GALEX source is 0.98 and probability of a false positive match is only 0.02.*

*Ogle et al. 2015ASPC..495...250 (ADASS)*

*A recent re-implementation of MatchEx with parallel processing is enabling effective cross-matching of surveys with  $> 10^8$  sources.*

# A closer look ...



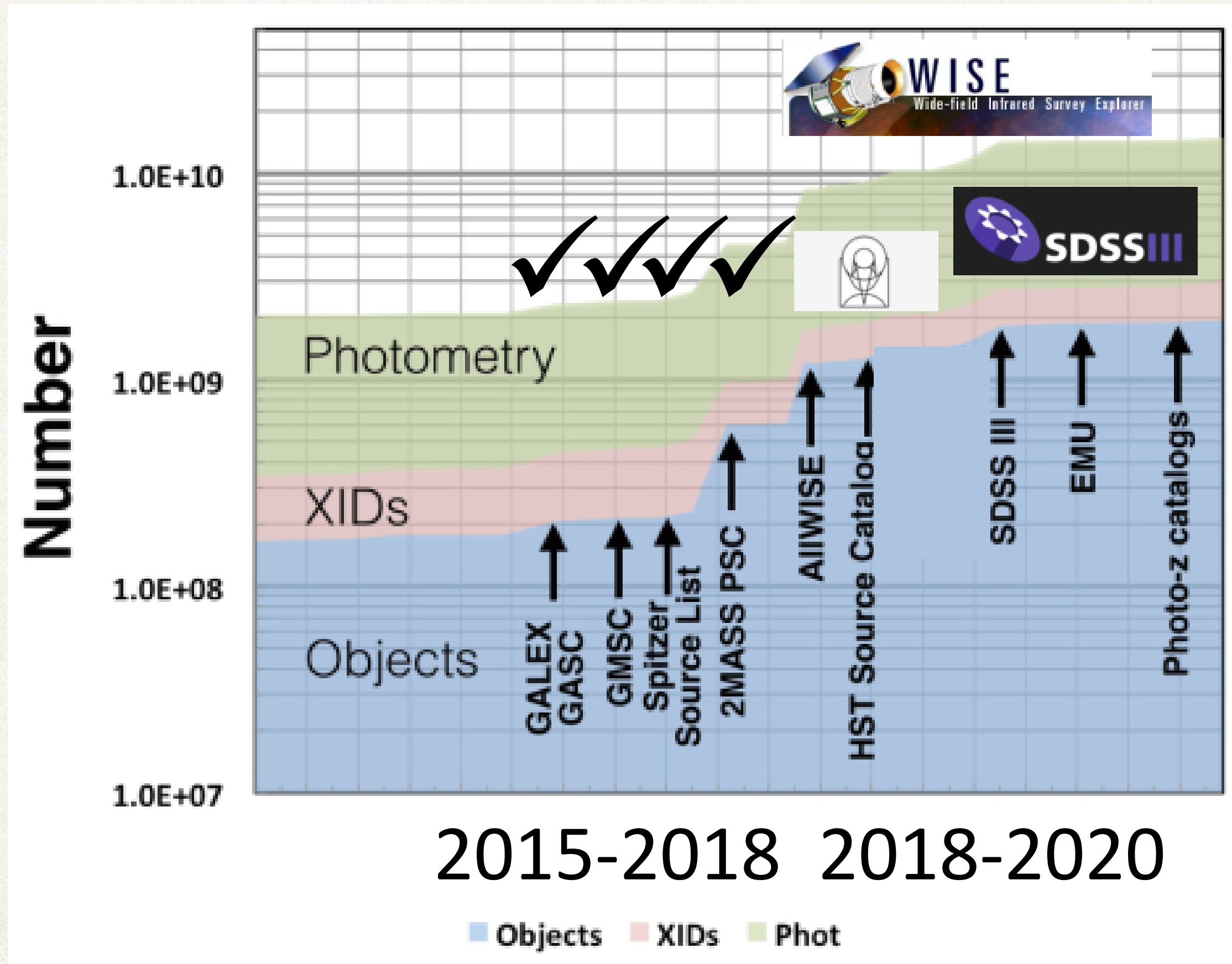


# Counts Then/Now

	April 2009	July 2018
Distinct Objects	163 M	667 M
Cross-IDs	170 M	773 M
Associations	0.19 M	1.4 M
References	68 K	113 K
Literature links	5.0 M	45 M
Redshifts	1.4M	11.3 M
Photometrics	1.7 B	4.98 B
Diameters	609 M	609 M
Spectra	56 K	645 K
Images	2.3 M	2.5 M
Database Records	3.3B	17B



# Into the 2020's



\* Volume, Velocity, Variety, Veracity

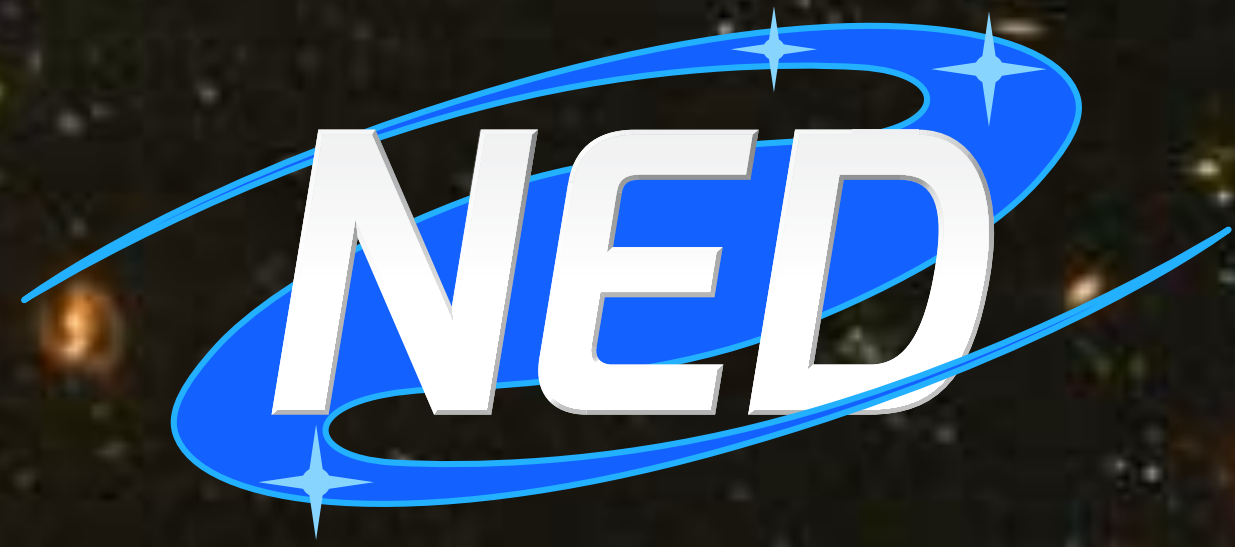


Euclid  
JWST  
WFIRST

*And data from thousands of journal articles*

2020+

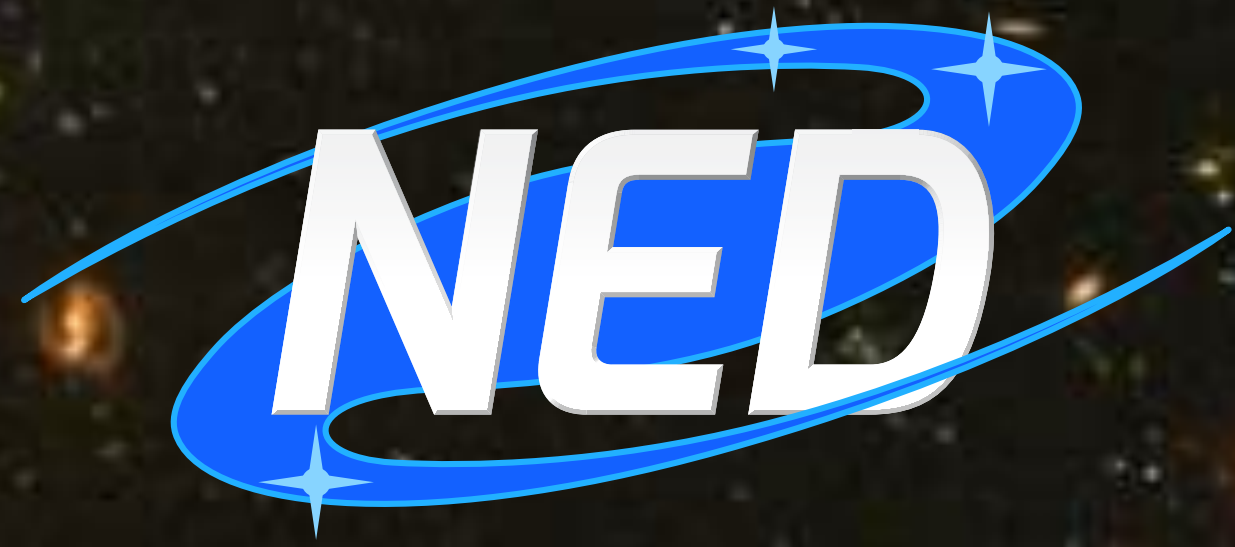
**Surpassing 1 billion X-matched objects with 10s of billions of attributes.**



# ObjectLookup

## What is a Name Resolver?

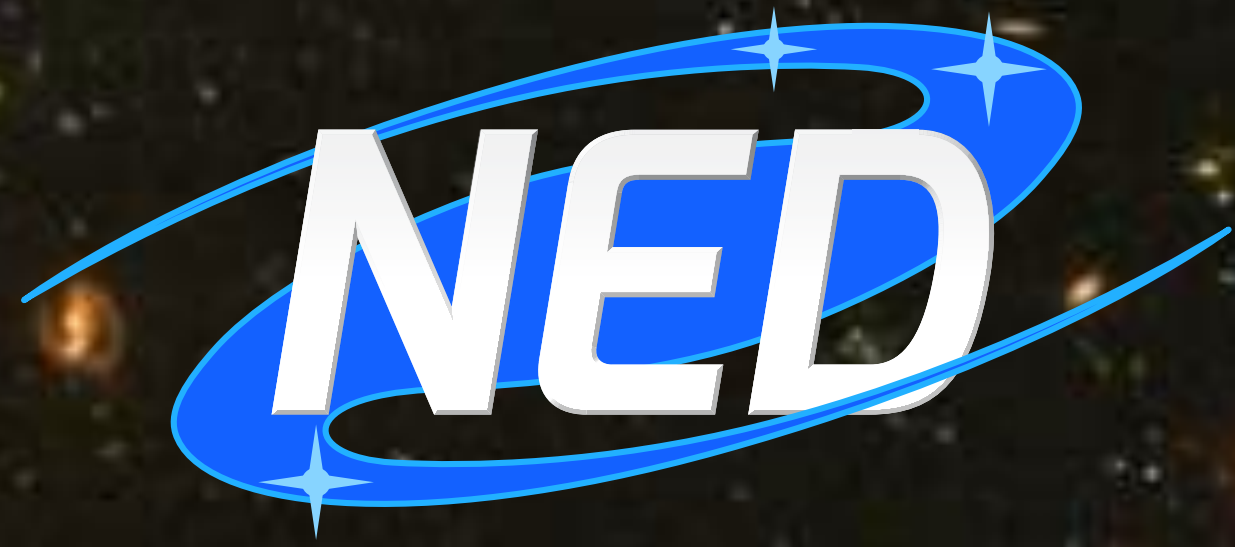
- Accepts any string a user enters
- Attempts to make sense of that string in some rule-set (NED's EGRET)
- Returns the “resolved name” for a specific domain (or error)
- May also return information about the named object for a specific domain



# ObjectLookup

ObjectLookup:  
The Next NED Name Resolver

[http://ned.ipac.caltech.edu/srs/ObjectLookup?  
json={\"name\": \"<STRING>\"}](http://ned.ipac.caltech.edu/srs/ObjectLookup?json={\)

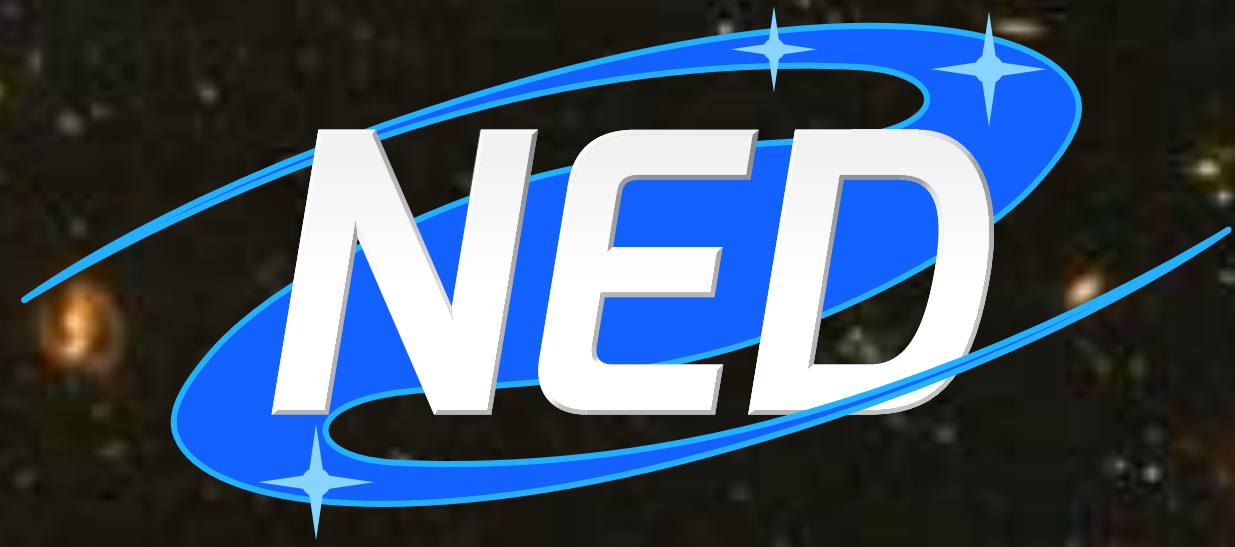


# ObjectLookup

<https://ned.ipac.caltech.edu/Documents/Guides/Interface/ObjectLookup>

```
{
  "Copyright": "(C) 2016 California Institute of
Technology",
  "Version": "1.1",
  "NameResolver": "NED-Egret",
  "QueryTime": "Tue May 17 12:20:37 2016",
  "StatusCode": 100,
  "ResultCode": 3,
  "Supplied": "n1068",
  "Interpreted": "NGC 1068",
```

```
  "Preferred": {
    "Name": "MESSIER 077",
    "ObjType": {
      "Value": "G",
      "RefCode": null
    },
    "Position": {
      "RA": 40.66962917,
      "Dec": -0.01328055556,
      "UncSemiMajor": 5.555555556e-5,
      "UncSemiMinor": 5.555555556e-5,
      "PosAngle": 0.0,
      "RefCode": "1997ApJ...476L..67C"
    },
    "Redshift": {
      "Value": 0.00379300001,
      "Uncertainty": 9.99999975e-6,
      "QualityFlag": null,
      "RefCode": "1999ApJS..121..287H"
    }
  }
  .....
}
```



# ObjectLookup

<https://ned.ipac.caltech.edu/Documents/Guides/Interface/ObjectLookup>

```
"aliases": {"v": true}
```

```
"Aliases": ["MESSIER 077", "NGC 1068", "UGC 02188", "ARP 037", "KUG 0240-002", "CGCG 388-098",  
"CGCG 0240.1-0013", "MCG +00-07-083", "3C 071", "4C -00.13", "PKS 0240-00", "2MASX  
J02424077-0000478", "2MASXi J0242407-000047", "IRAS 02401-0013", "IRAS F02401-0013", "AKARI  
J0242406-000039", "ISOSS J02426-0000", "CGS 161", "LDCE 0174 NED006", "HDCE 0163 NED002",  
"USGC U143 NED02", "LQAC 040-000 029", "HIPASS J0242+00", "NSA 024009", "NSA 023964", "PGC  
010266", "RBS 0348", "UZC J024240.7-000048", "PKS B0240-002", "PKS J0242-0000", "PMN  
J0242-0000", "MRC 0240-002", "MG1 J024240-0000", "87GB[BWE91] 0240-0013", "[WB92] 0240-0013",  
"FIRST J024240.7-000046", "NVSS J024240-000047", "VLSS J0242.6-0000", "UITBOC 0219", "S3  
0240-00", "OD -067", "AT20G J024240-000046", "DA 082", "NRAO 0112", "TXS 0240-002", "Cul  
0240-002", "GC 0240-00", "LHE 067", "MSH 02-014", "PAPER J040.87+00.15", "NEWPS_5yr_5s 063",  
"NEWPS_5yr_5s_15 062", "AFGL 4220S", "CXO J024240.8-000047", "RX J0242.6+0000", "RX  
J0242.6-0000", "1RXS J024240.9-000046", "1RXP J024240.9-000042", "2PBC J0242.7-0000", "PBC  
J0242.7-0000", "2XMM J024240.7-000046", "2XMMp J024240.7-000046", "1XMM J024240.7-000046",  
"1AXG J024241-0000", "1H 0244+001", "1ES 0240-002", "XSS J02445-0000", "CXO  
J024240.79-000046.4", "3FGL J0242.7-0001", "SWIFT J0242.9-0000", "SWIFT J0242.6+0000", "2FGL  
J0242.5+0006", "[KWP81] 0240-00", "[dML87] 057", "LGG 073:[G93] 002", "[MHH96]  
J024239+000001", "RX J0242.6-0000:[BEV98] 003", "[M98j] 041 NED07", "[VCV2001]  
J024240.7-000047", "[CRK2003] J0242.6-0000", "[SLK2004] 0333", "NGC 1068:[LB2005] X01",  
"[VCV2006] J024240.7-000047", "[KRL2007] 025", "NGC 1068:[L2011a] X0001", "[MSC2011]  
J024240.71-000047.7", "[BTM2013] 0144", "[AHG2014] B126", "0244+001", "LEDA 010266",  
"0240-002", "0240-0013", "0240-00"]}]
```



# Challenges ...

- ❖ Volume, Velocity, Variety, Veracity
- 

- ❖ Virtualization,
- ❖ Horizontal Scaling,
- ❖ Clustered VM,
- ❖ Containers,
- ❖ Platforms
- ❖ and then ...

What does that mean to put an “archive” in the cloud?

# ... to the Cloud?

---

What does that mean to  
put “VO” in the cloud?



What does that mean to put anything in the cloud?

---

Consider the alternatives:

Archivable & Reproducible

vis-a-vis

Interactive & Collaborative



# “Velocity problem”

---

A scientist writes a script ...

# Velocity problem

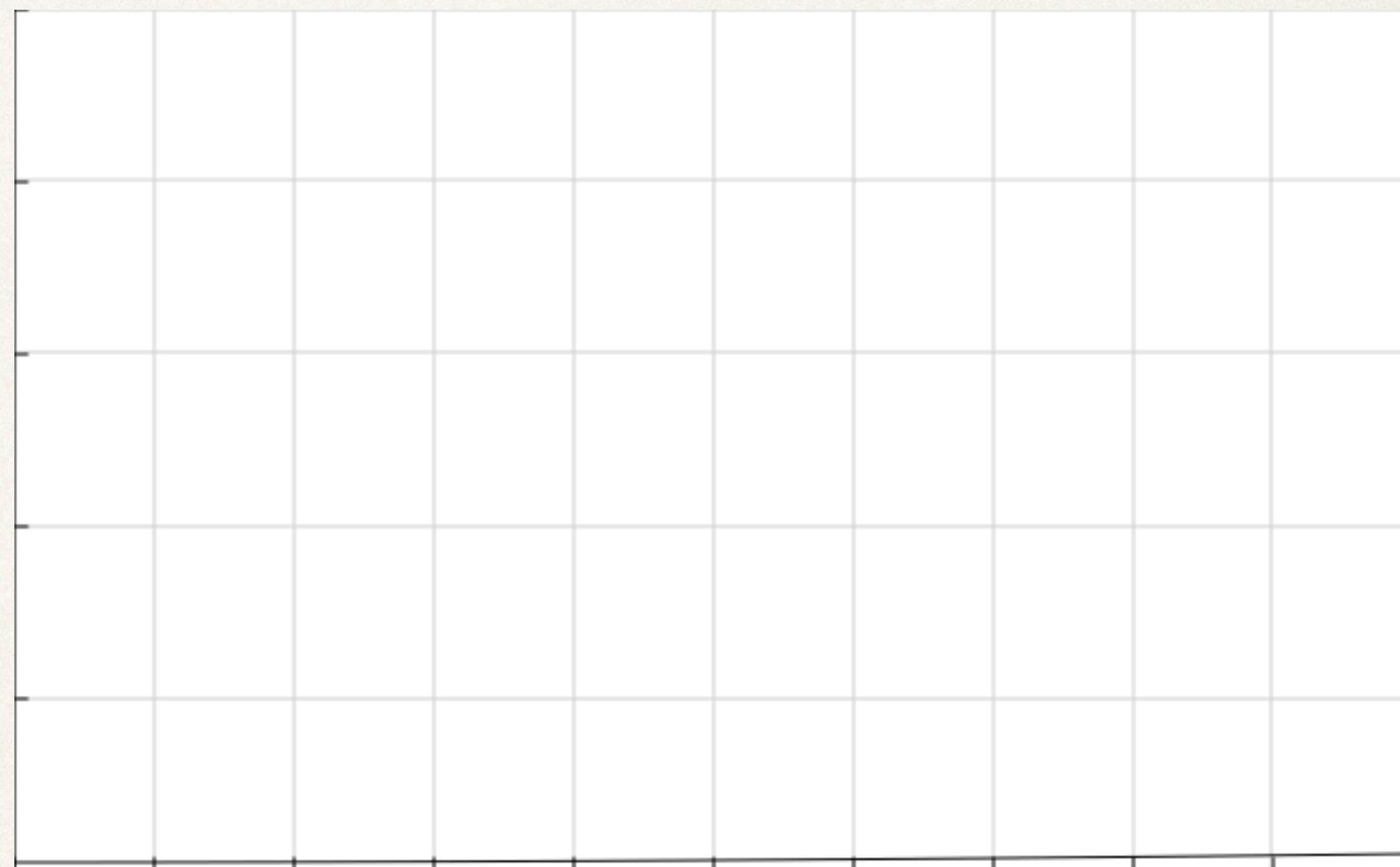
---

It uses a VO API library,  
which is “cool”

It runs on her laptop - and it does  
a small ‘cone search’ on a list of objects...

A thousand positions in,  
a couple of minutes later,  
a thousand result lists out.

‘Really Cool’



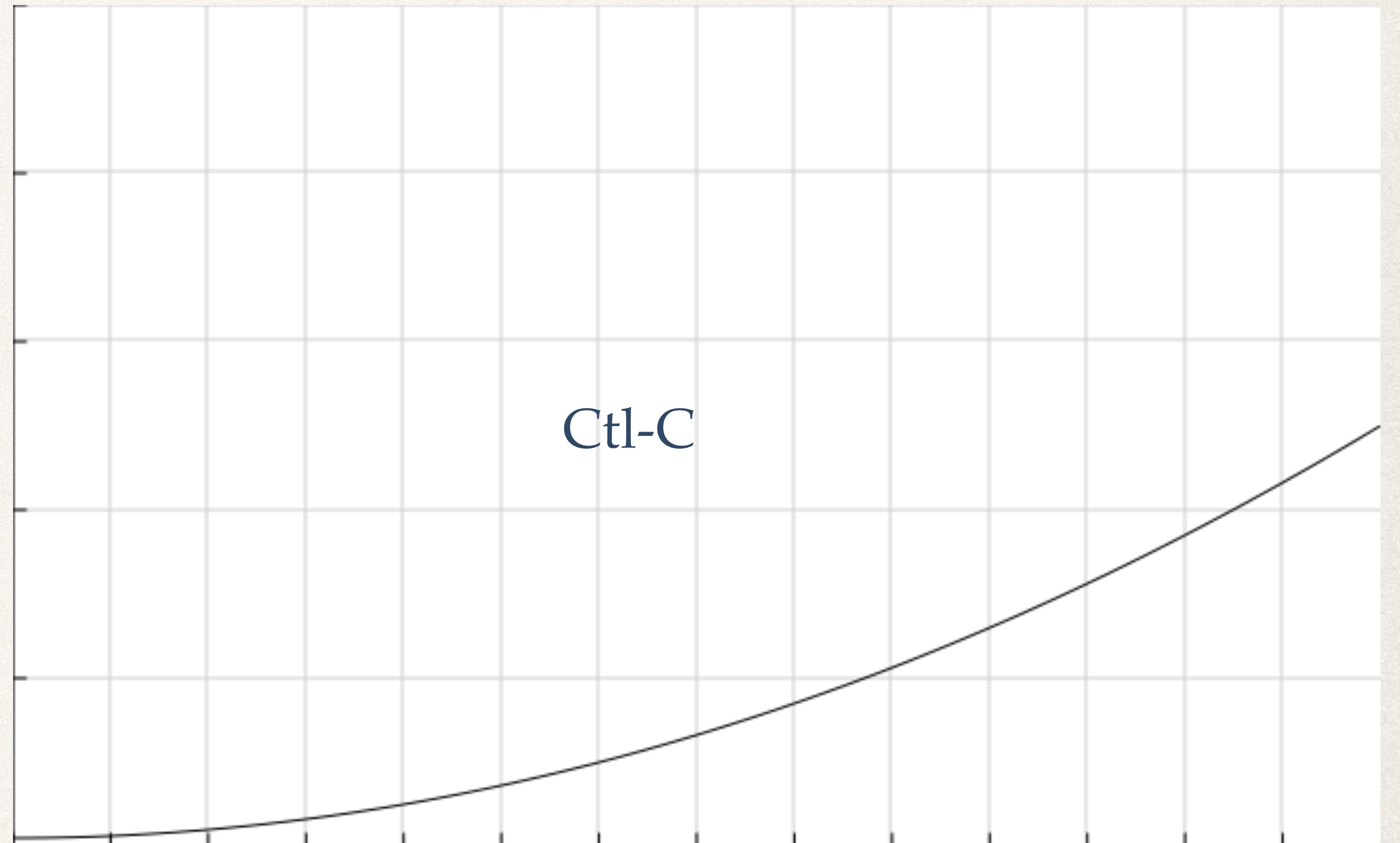
She has 10 million positions ...

It runs on her laptop - and it does  
a 'cone search' on her big list of objects...

It will take awhile ... but ok...



Wait... hold on, 10M positions,  
at 1000 minute ... that's ...  
10,000 minutes ! Gotta do this faster...



... to the Cloud!

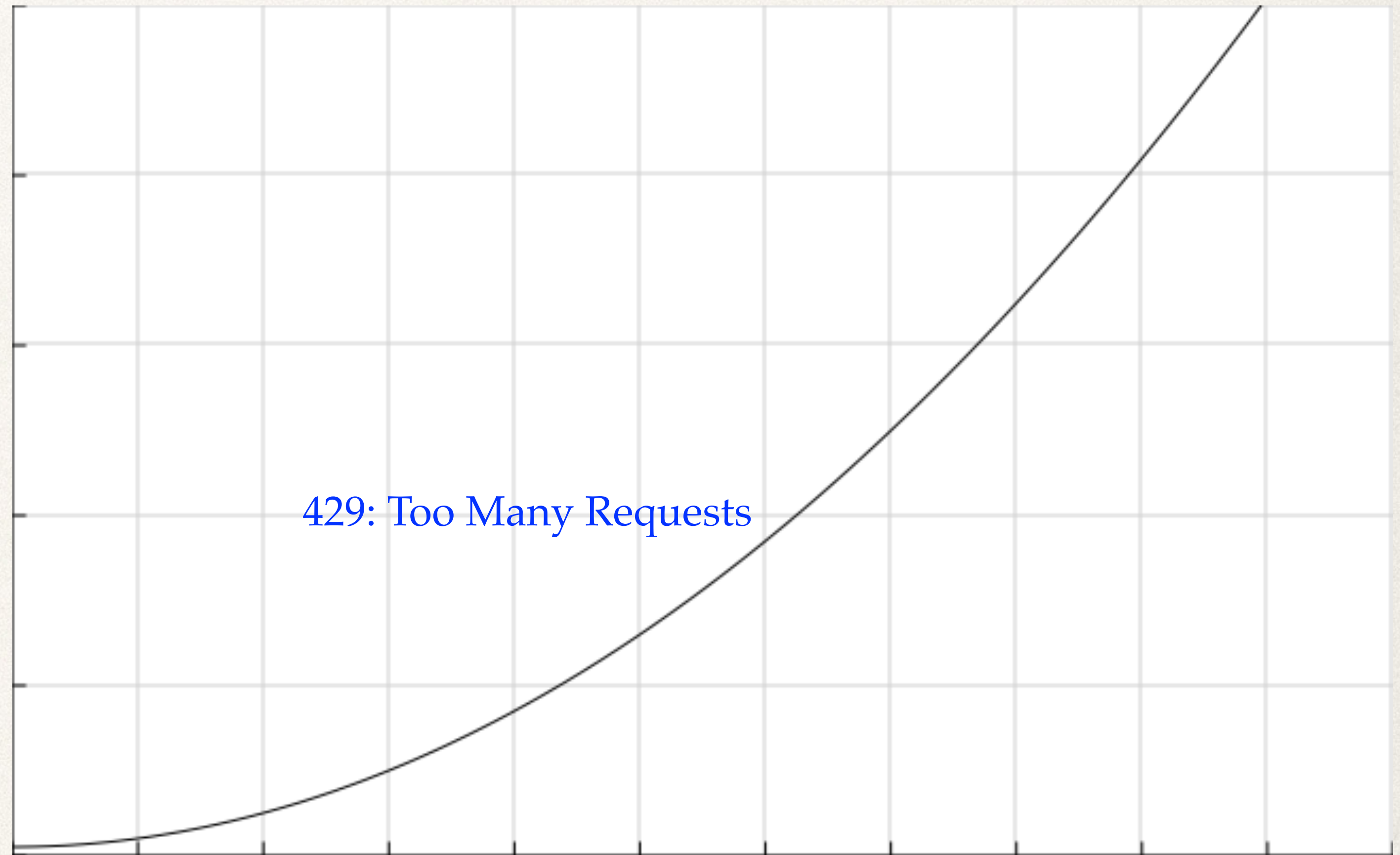
---

Uh oh...





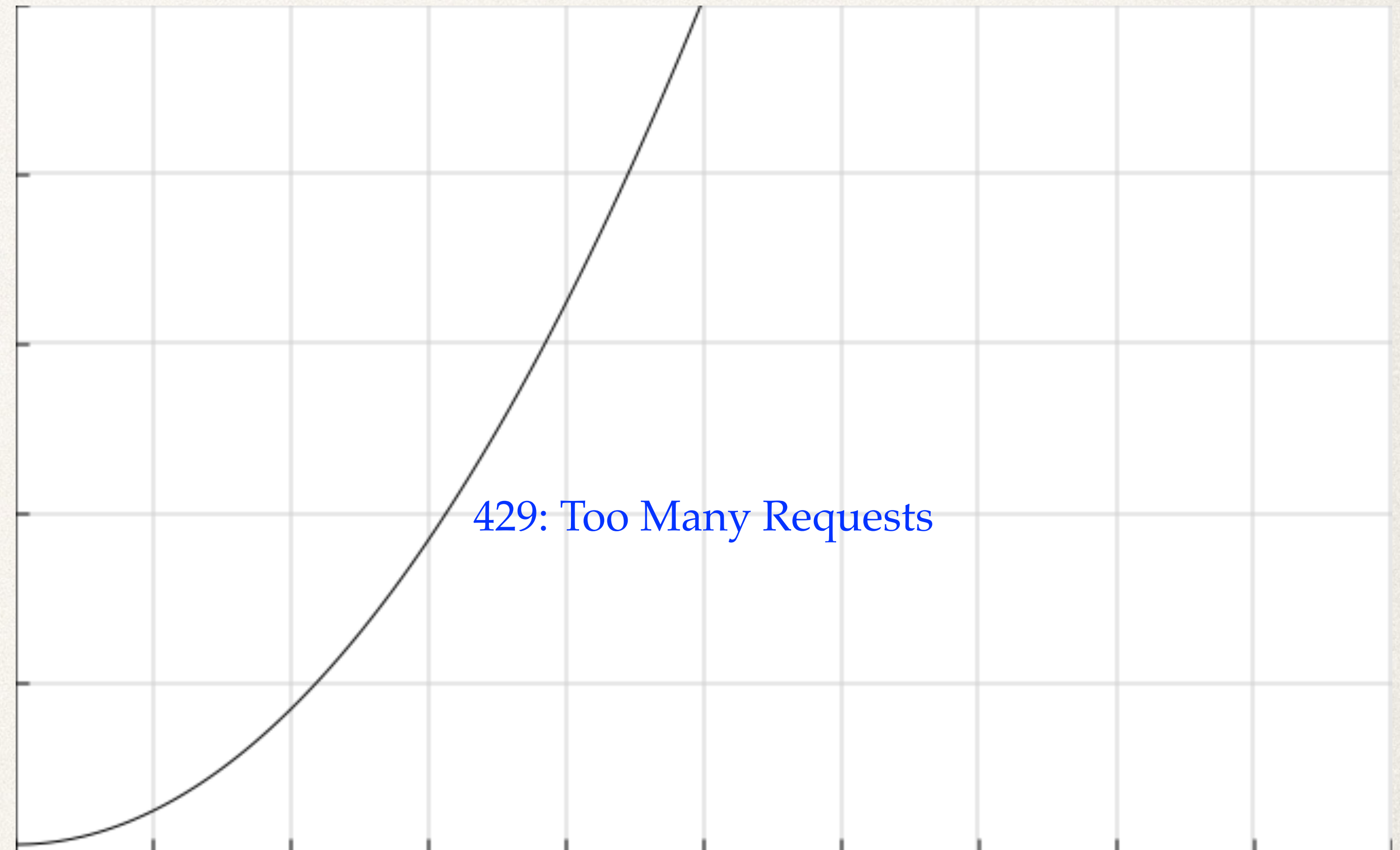
Hmmm, I can run this in parallel  
on my department compute cluster!



---

---

# 503: Service Unavailable



# What to Do?

---

- ❖ Let's practice and encourage
  - ❖ "status checking",
  - ❖ Reasonable response to
    - ❖ 429
    - ❖ 503
  - ❖ Retry-After:
- ❖ Do we need:
  - ❖ X-RateLimit-\*
- ❖ Community / IVOA review / endorsement of CLIENT APIs?



Thanks!

The NED Team/Rick Ebert

IVOA-Nov2018