

Spectroscopic Survey Pipelines for the GBT

Erik Rosolowsky

Use-Cases

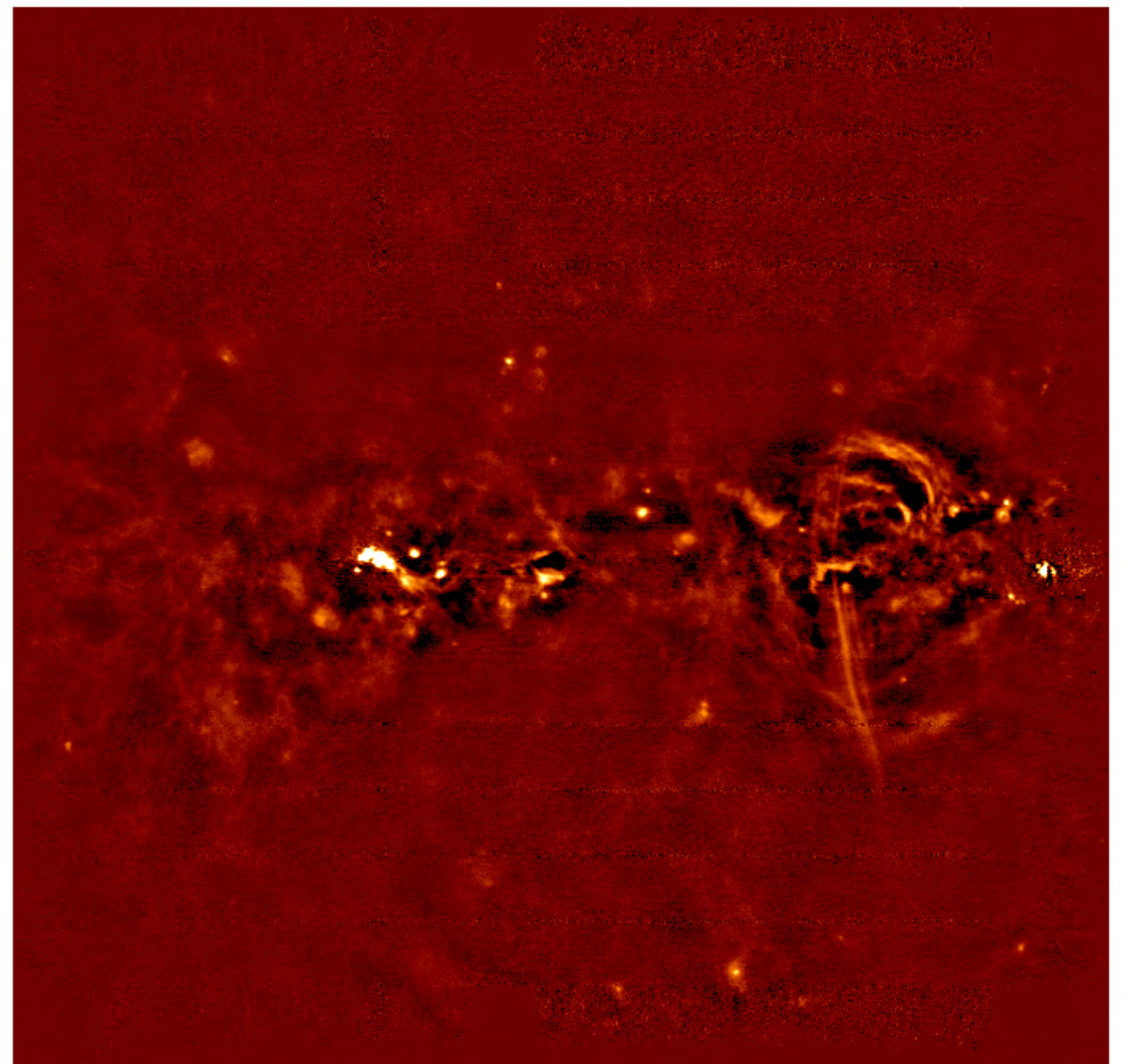
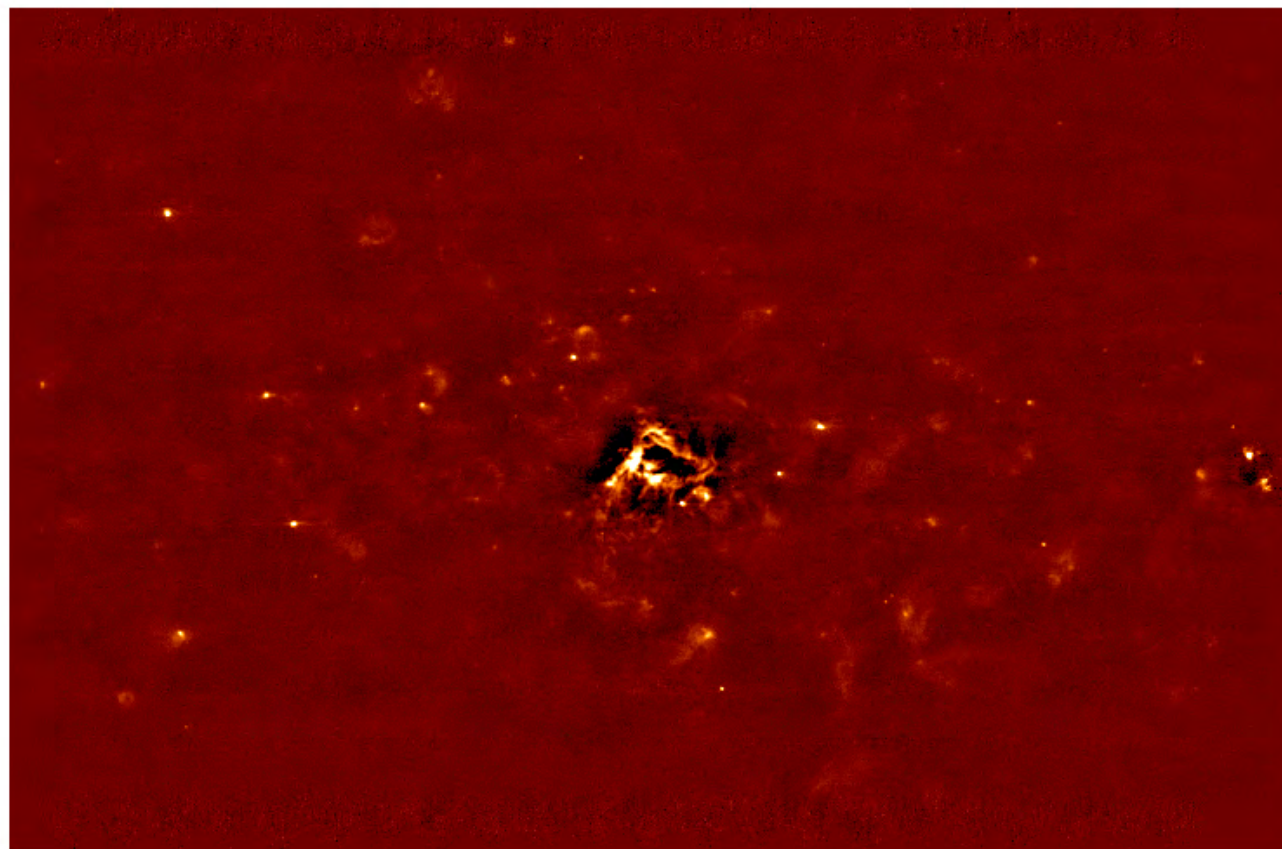
- Operation on sdfits data (after filling)
- OTF spectral line mapping
- Multi-pixel receivers, multi-lines

Not These Use Cases

MUSTANG2 Galactic Plane Survey

PI: Adam Ginsburg

Images courtesy of Brian Mason



Survey Software Roles

Data and Observing Management

Mapping connecting log data to output products

Calibration

Raw telescope data into astronomer units

Mapping

Individual spectra converted into data cubes

Analysis

Derived science projects.

The Landscape

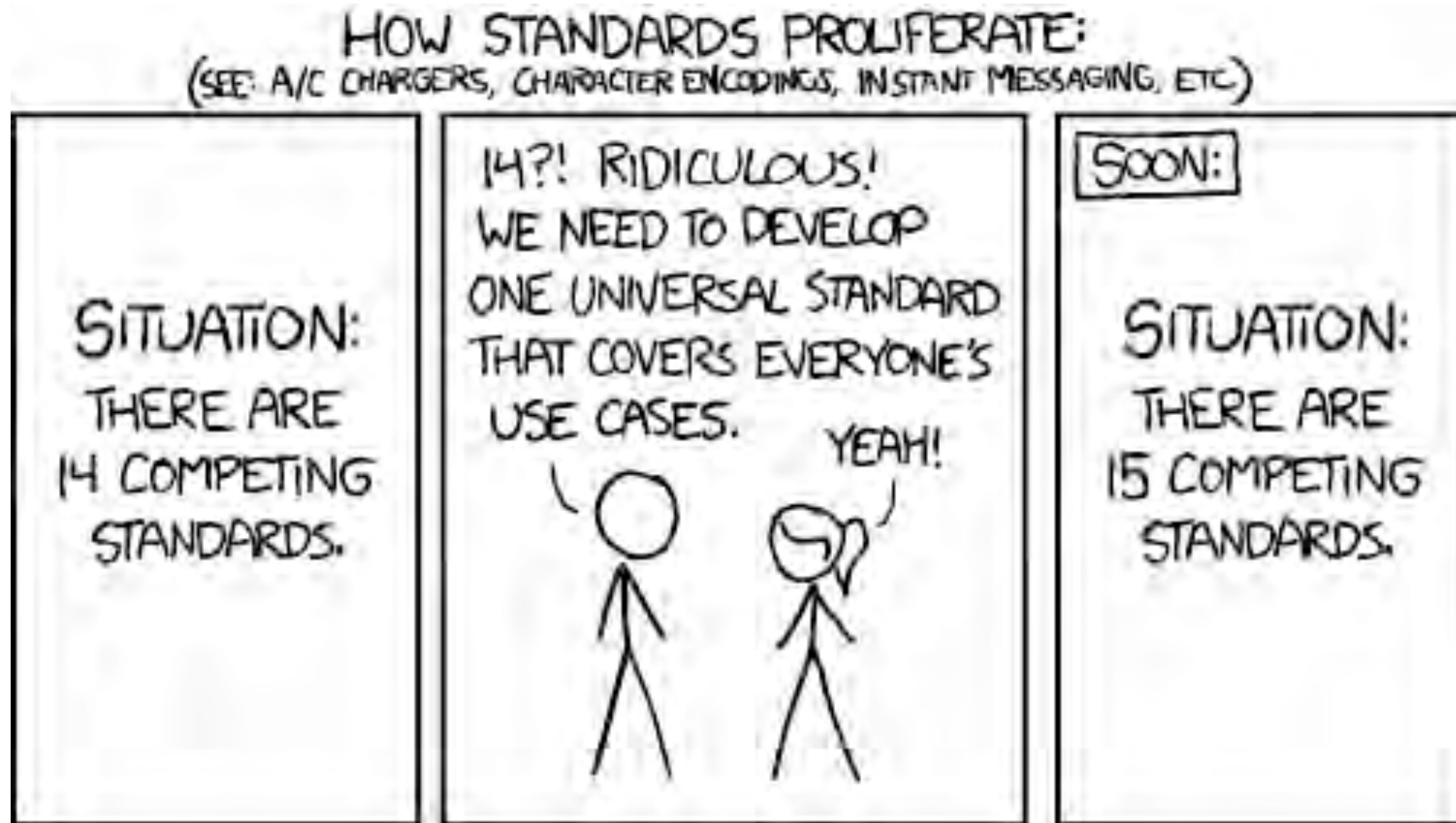
Calibration GBTIDL, gbt-pipeline, user software
(TMBIDL, Robishaw and Heiles, etc.)

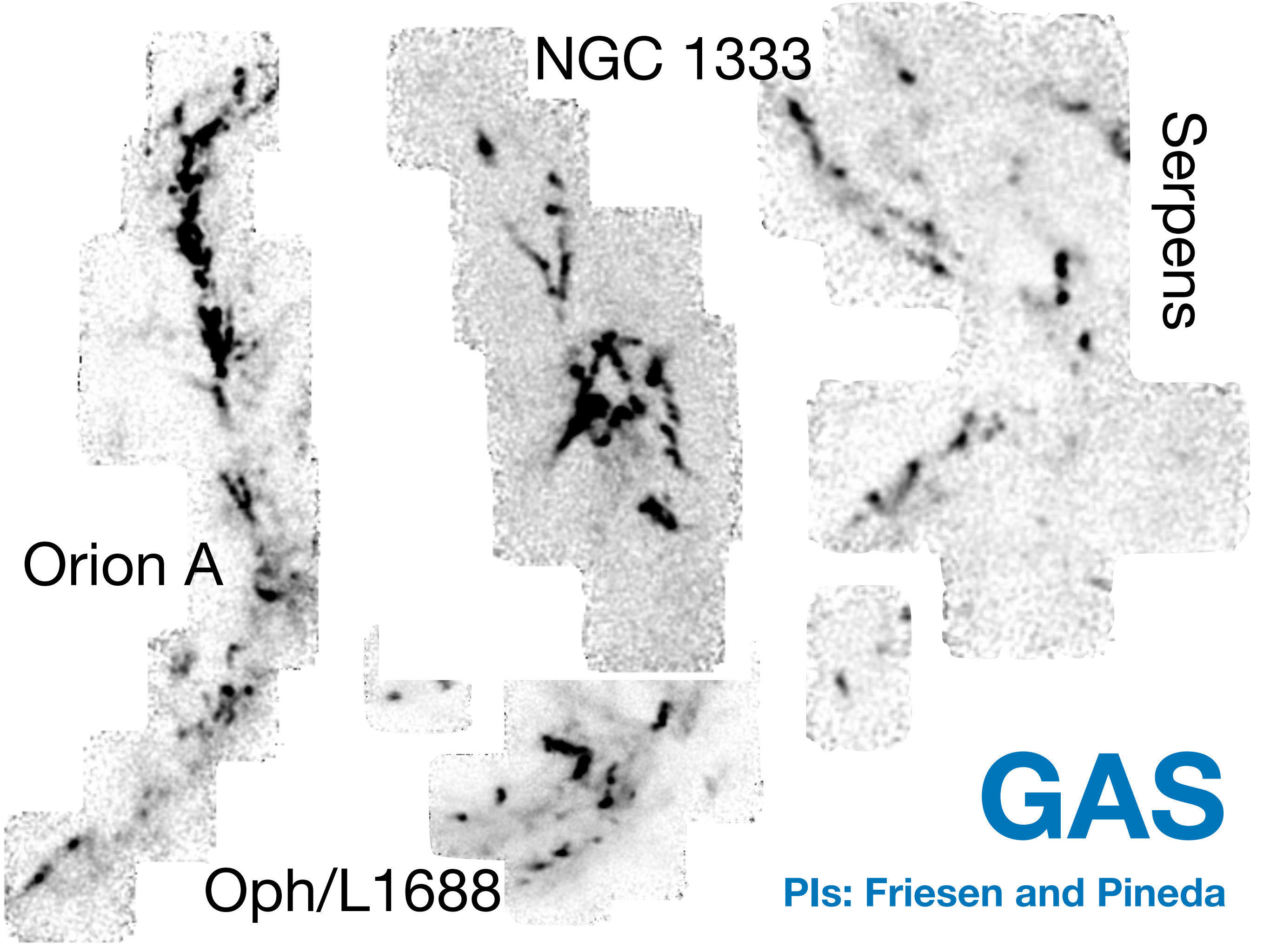
Mapping AIPS gridder, gbtgridder, cygrid,
sdgrid (CASA)

Analysis GBTIDL, pyspeckit, user software

**Observation
Management** Whatever you like

How we got here





Orion A

NGC 1333

Serpens

Oph/L1688

GAS

PIs: Friesen and Pineda

GAS

<https://github.com/GBTAmmoniaSurvey>

Observation
Management

Google Docs

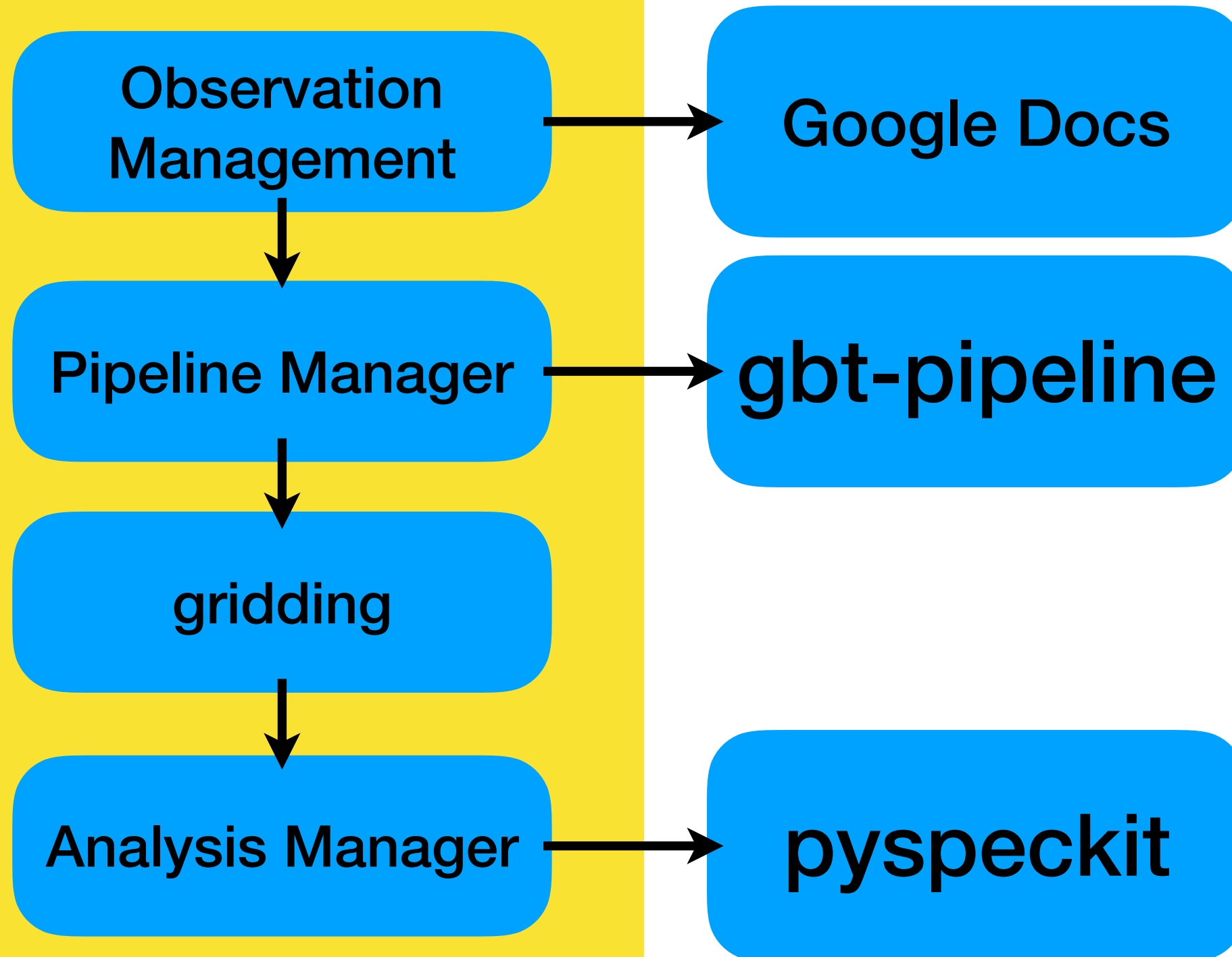
Pipeline Manager

gbt-pipeline

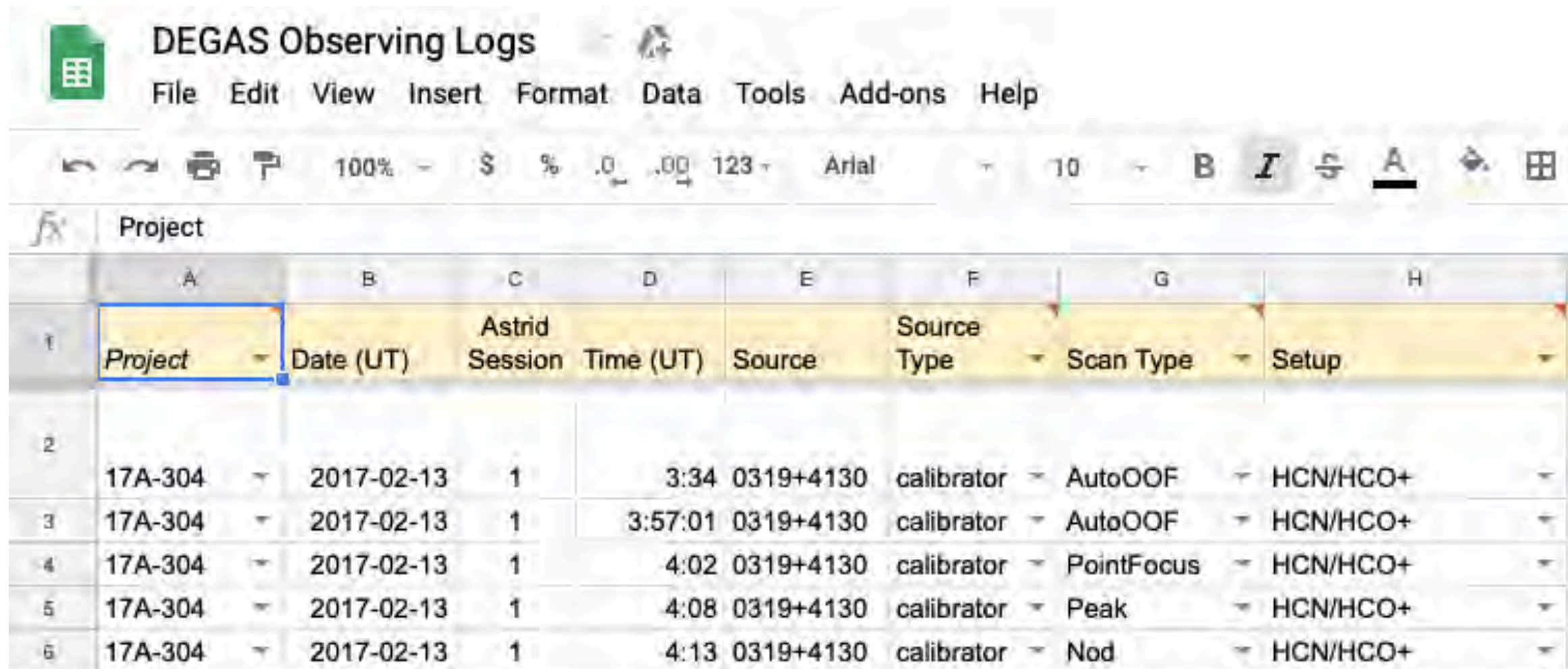
gridding

Analysis Manager

pyspeckit



Observing logs control pipeline



The screenshot shows a Google Sheets interface titled "DE GAS Observing Logs". The menu bar includes File, Edit, View, Insert, Format, Data, Tools, Add-ons, and Help. The toolbar shows various icons for undo, redo, print, and formatting, along with a font face of Arial and a size of 10. The spreadsheet has columns labeled A through H. The first row (row 1) contains headers: "Project", "Date (UT)", "Astrid Session", "Time (UT)", "Source", "Source Type", "Scan Type", and "Setup". The second row (row 2) contains data: "17A-304", "2017-02-13", "1", "3:34", "0319+4130", "calibrator", "AutoOOF", and "HCN/HCO+". The third row (row 3) contains data: "17A-304", "2017-02-13", "1", "3:57:01", "0319+4130", "calibrator", "AutoOOF", and "HCN/HCO+". The fourth row (row 4) contains data: "17A-304", "2017-02-13", "1", "4:02", "0319+4130", "calibrator", "PointFocus", and "HCN/HCO+". The fifth row (row 5) contains data: "17A-304", "2017-02-13", "1", "4:08", "0319+4130", "calibrator", "Peak", and "HCN/HCO+". The sixth row (row 6) contains data: "17A-304", "2017-02-13", "1", "4:13", "0319+4130", "calibrator", "Nod", and "HCN/HCO+".

	A	B	C	D	E	F	G	H
1	Project	Date (UT)	Astrid Session	Time (UT)	Source	Source Type	Scan Type	Setup
2	17A-304	2017-02-13	1	3:34	0319+4130	calibrator	AutoOOF	HCN/HCO+
3	17A-304	2017-02-13	1	3:57:01	0319+4130	calibrator	AutoOOF	HCN/HCO+
4	17A-304	2017-02-13	1	4:02	0319+4130	calibrator	PointFocus	HCN/HCO+
5	17A-304	2017-02-13	1	4:08	0319+4130	calibrator	Peak	HCN/HCO+
6	17A-304	2017-02-13	1	4:13	0319+4130	calibrator	Nod	HCN/HCO+

```
In [1]: import GAS
```

```
In [2]: GAS.updateLogs( )
```

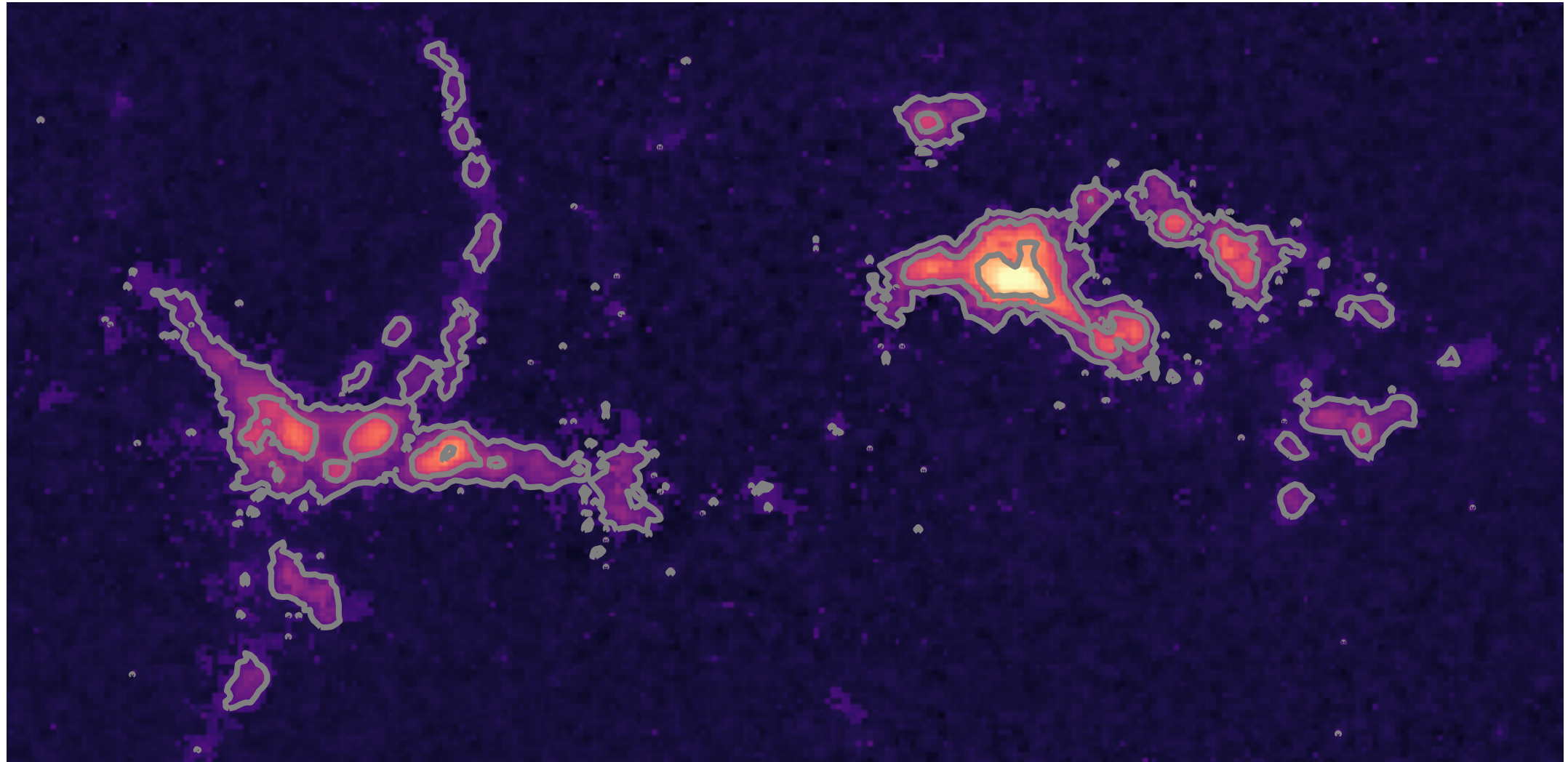
```
In [3]: GAS.reduceAll( )
```

Pulls from Google Docs

Calibrates all new data

KEYSTONE

PI: James DiFrancesco



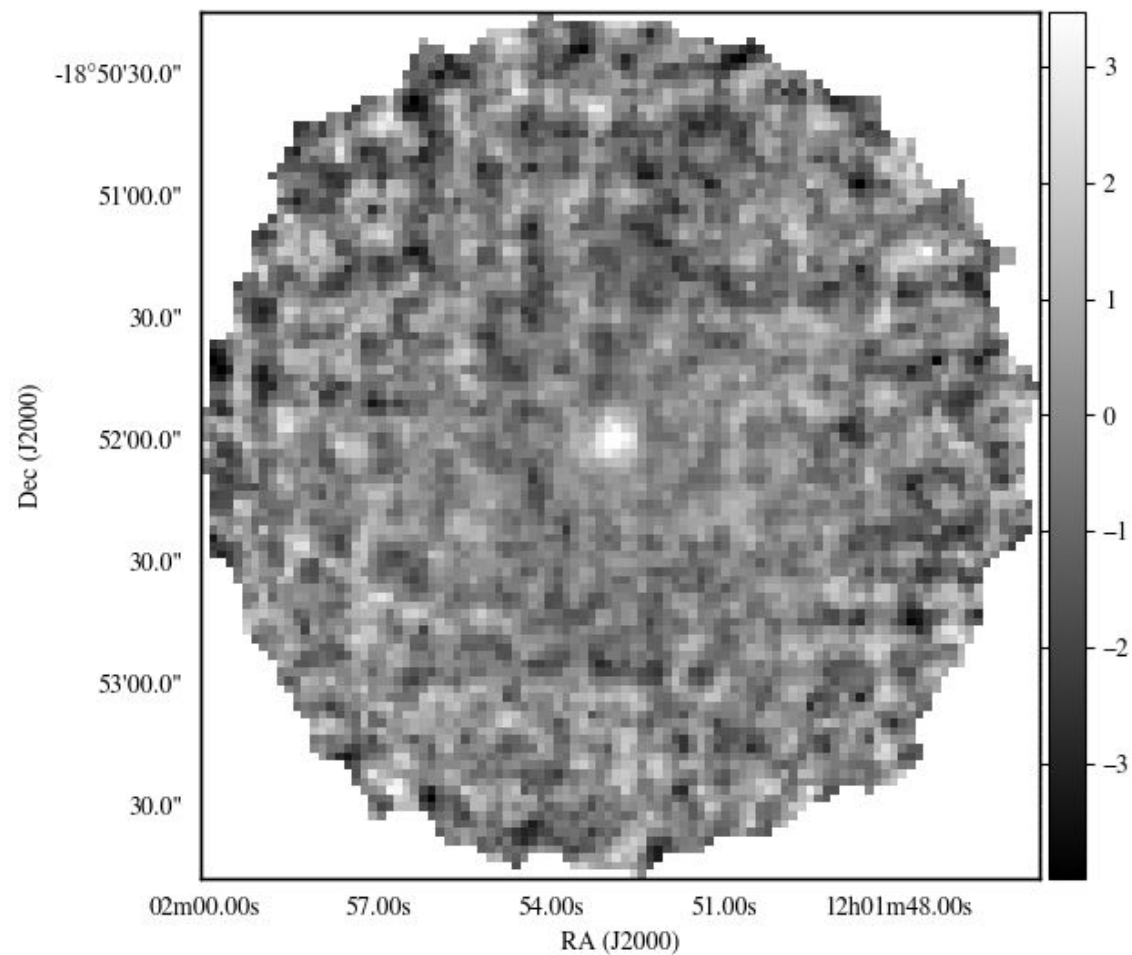
Similar mapping strategy

Different spectral setups (KEYSTONE uses RAMPS setup)

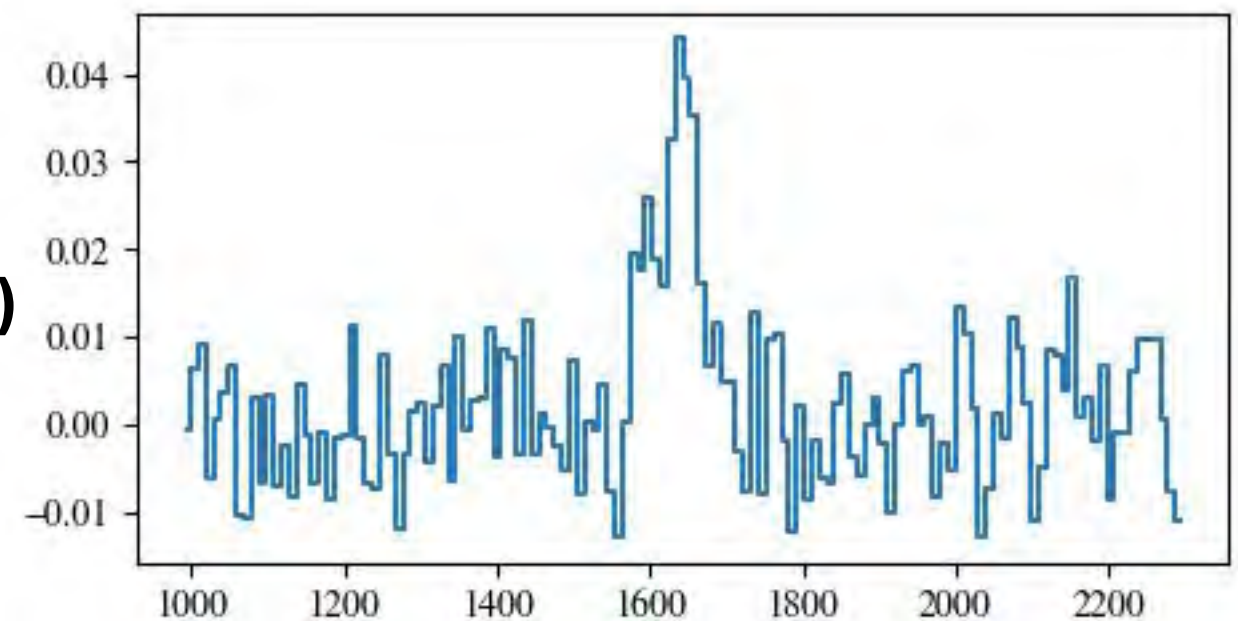
DEGAS

PI: Amanda Kepley

NGC 4038



T_b (K)



V (km/s)

New mapping strategy
New calibration data needed

gbtpipe

<https://github.com/GBTSpectroscopy/gbtpipe>

The screenshot shows the GitHub repository page for **GBTSpectroscopy / gbtpipe**. At the top, there are buttons for **Unwatch** (1), **Star** (1), and **Fork** (3). Below this is a navigation bar with links for **Code**, **Issues** (5), **Pull requests** (0), **Projects** (0), **Wiki**, **Insights**, and **Settings**.

The main content area features the description "Packaged version of the GBT pipeline" with an **Edit** button. Below this is a section for repository statistics: **63 commits**, **1 branch**, **4 releases**, and **2 contributors**.

A horizontal bar indicates the current branch is **master**. Below this bar are buttons for **New pull request**, **Create new file**, **Upload files**, **Find file**, and a green **Clone or download** button.

The commit history section shows the latest commit by **low-sky** with the message "Try/except for corrup scans in FDNUM" on Sep 11. Below this is a list of files and their commit history:

File	Commit Message	Time Ago
astropy_helpers @ 02858df	Initial layout for package	2 years ago
docs	Initial layout for package	2 years ago
gbtpipe	Try/except for corrup scans in FDNUM	2 months ago
.gitignore	Initial layout for package	2 years ago

Provides interface to GBT calibration framework in python

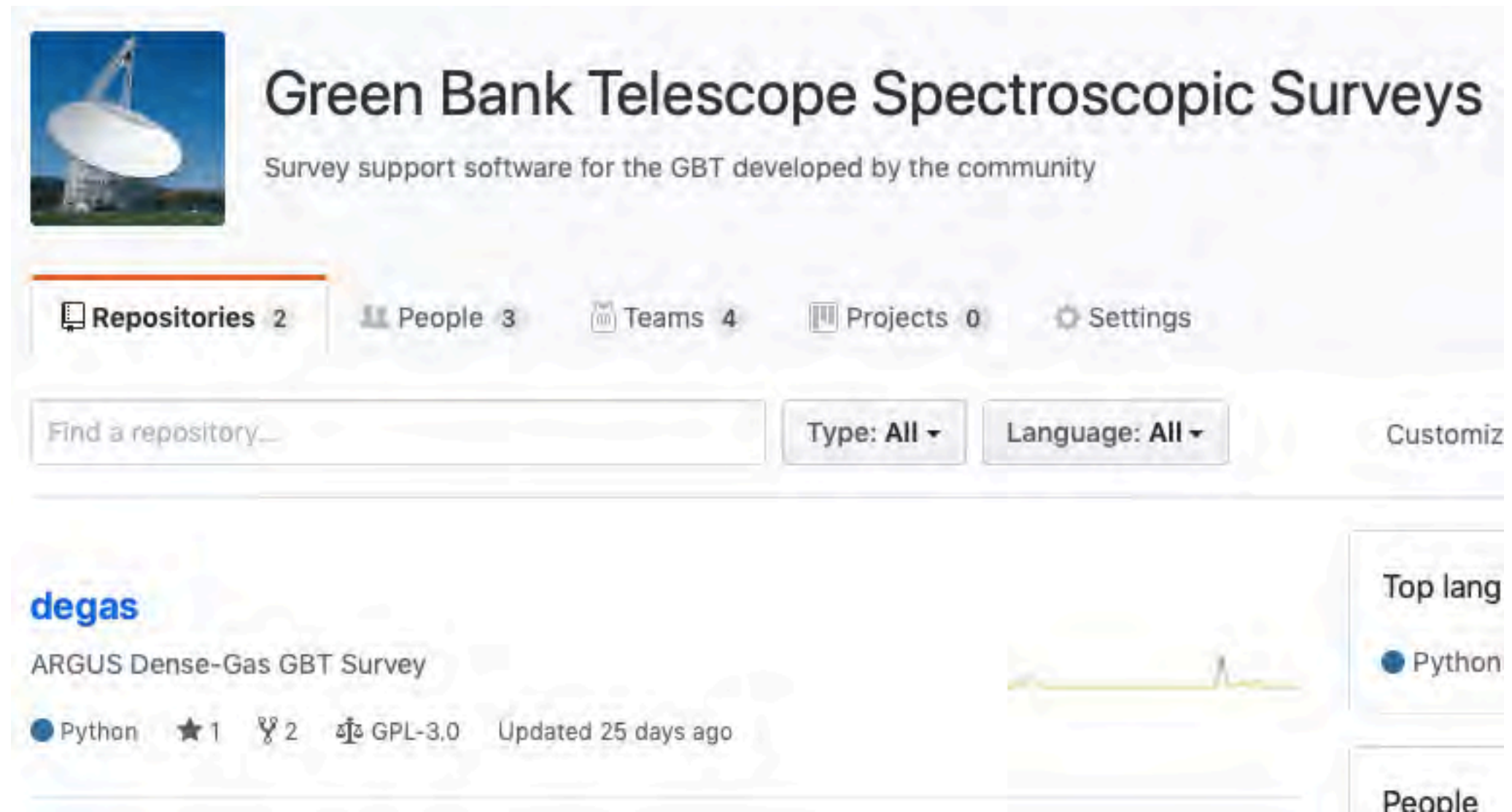
Github model for pipeline management

- Allow multiple contributors to the pipeline development via pull request model
- Data release
 - Build end-to-end script for making data products
 - Create tag in github
- Documentation and public release
- Possibility for regression testing

Areas for Partnerships

- **Web-based access to GBT weather data**
- Connections between software teams on different surveys
- Local storage and faster access to on site machines (X2GO rather than VNC)?

Community Development Partnerships

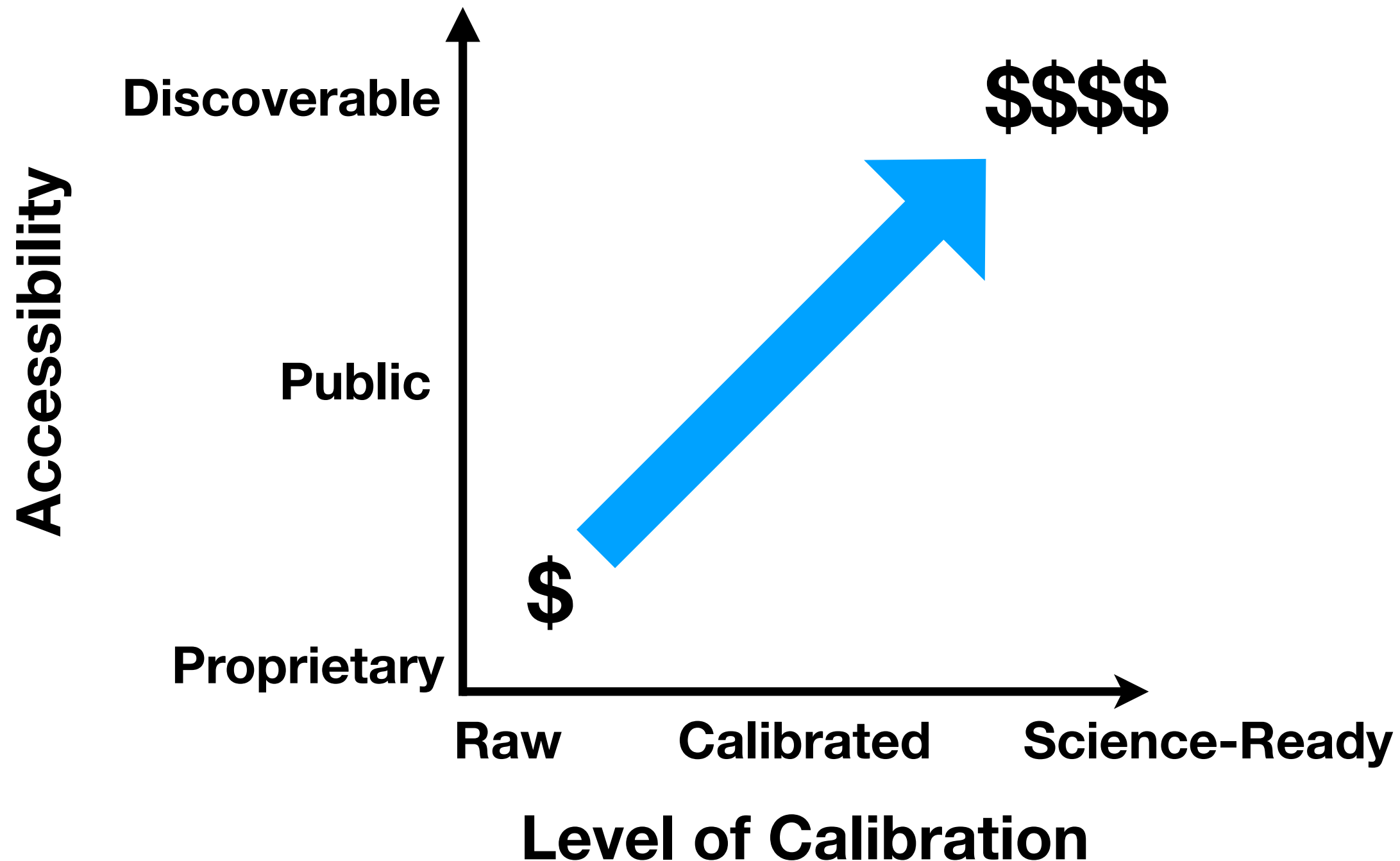


<https://github.com/GBTSpectroscopy>

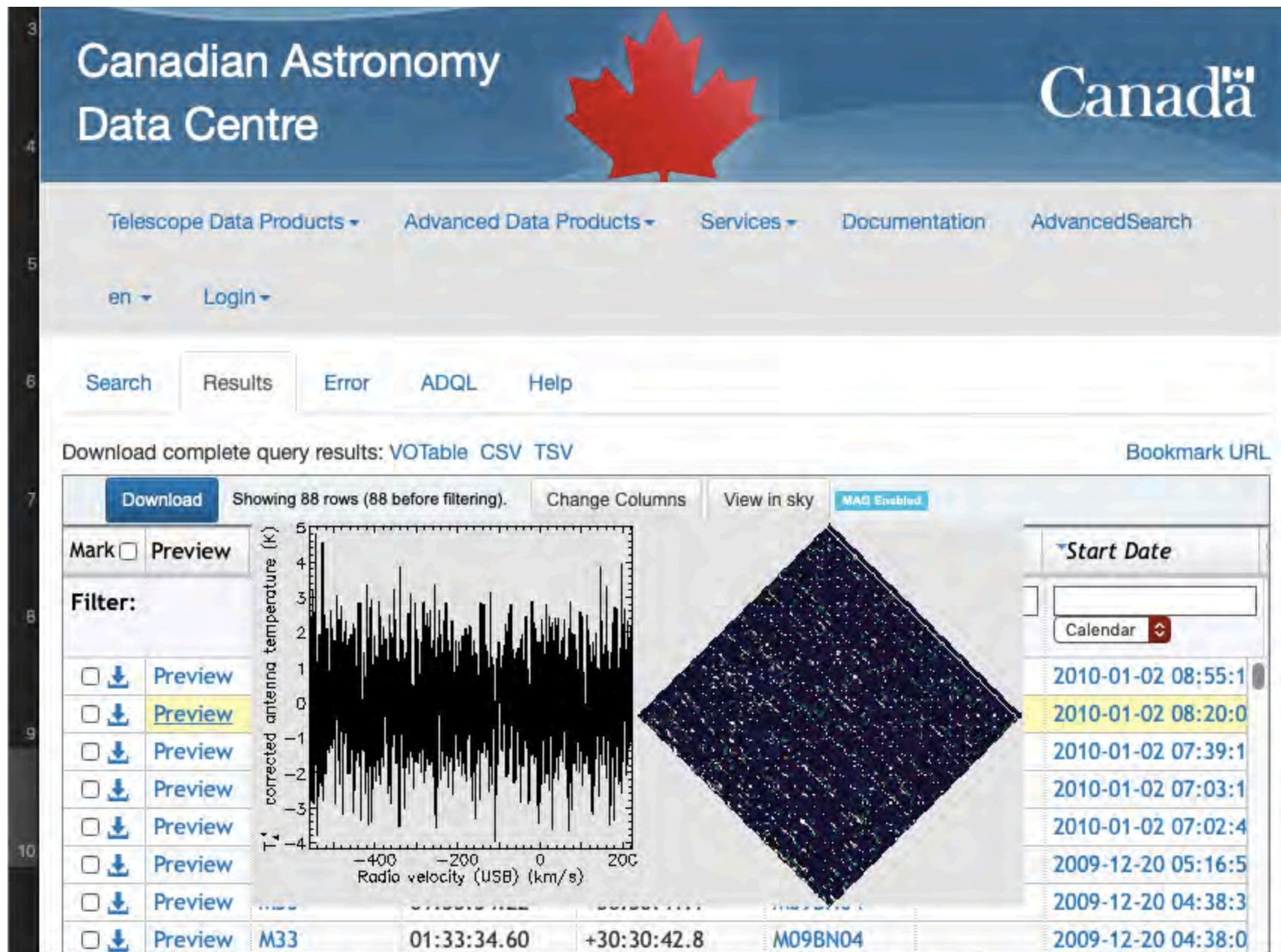
Stupid ideas: software bounties. Hack meetings.
Community software manager

Data Access and SRDP

SRDP = Science-Ready Data Products



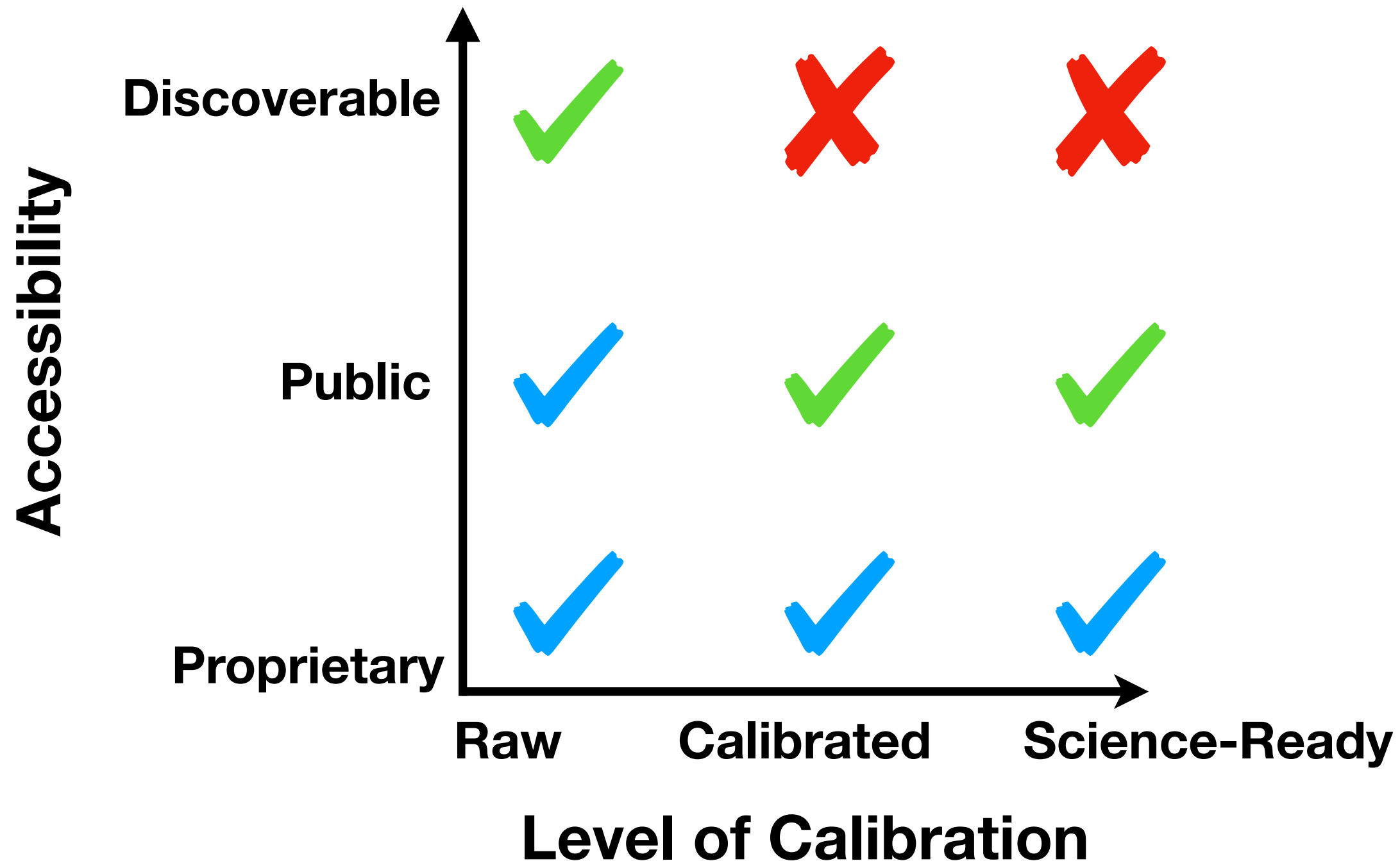
Discoverability — Ability to find a product without knowing that the survey exists.



This is hard. Useful but difficult.

Data Access and SRDP

SRDP = Science-Ready Data Products



Lessons Learned

- Efficiencies gained from parallel development across different surveys
- Excellent software exists in GBT legacy
- Access to individual software pieces allows rapid development of pipelines:
 - ON/OFF pairs
 - Logging
 - Weather
 - Spectroscopic transforms