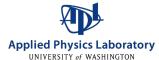


QIceRadar

A tool for Finding, Downloading and Visualizing
Georeferenced Radargrams within the QGIS Ecosystem

Laura Lindzey

12 December 2022
AGU Fall Meeting C15B-03



Applied Physics Laboratory
UNIVERSITY of WASHINGTON



Funded by NSF-CSSI #2209726
(Cyberinfrastructure for Sustained Scientific Innovation)

Goals & Outline

- Motivation
- Planned Functionality
- Data Availability
- Infrastructure

Hi, I'm Laura Lindzey and I'm presenting QIceRadar, a tool for finding, downloading, and visualizing georeferenced radargrams within the QGIS Ecosystem.

This project only just started, but I wanted to give a talk announcing it with a couple goals in mind:

- * First, I want to start discussions with data providers and managers about the issues around making these data easily available and reusable in time to have an impact over the course of the project.
- * I also hope to find additional beta testers outside of my academic network, particularly researchers who aren't already immersed in radar data but are interested in using it. If this describes you, please get in touch.

FAIR data practices

- **FINDABLE**

- How can somebody search for and discover the dataset?

- **ACCESSIBLE**

- Can the dataset be downloaded?
- Can this download be automated?

- **INTEROPERABLE**

- **REUSABLE**

- Is there sufficient metadata and documentation for reuse?
- Was the data released under a license allowing reuse?

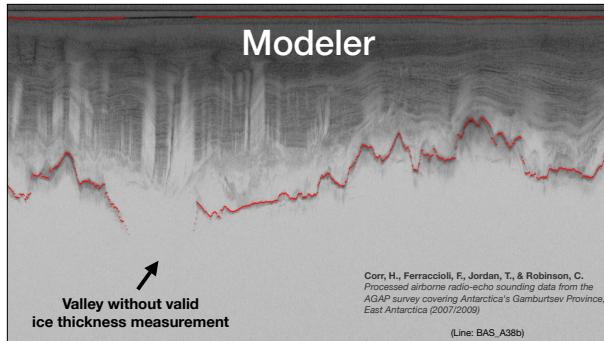
go-fair.org/fair-principles

The scientific community has recently been moving towards adopting FAIR data principles, with an emphasis on enabling automated access and analysis of data. These mesh well with the goals of QIceRadar.

This project will help make radar datasets findable, by providing an index to radar data across many international data centers. We can help make diverse data formats interoperable, by providing parsers that read these formats into a standardized representation. However, we rely on the original datasets being provided in an accessible and reusable way.

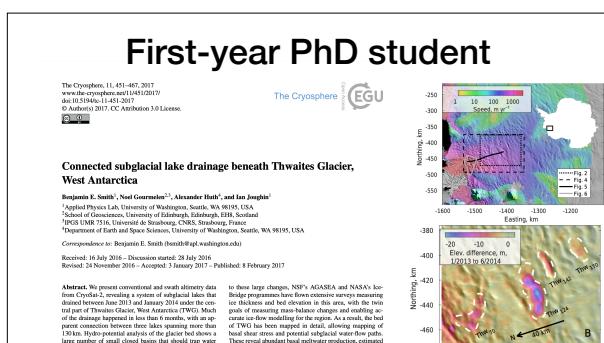
Motivation

QIceRadar is primarily targeted at non-expert users. I want to make it easier for researchers outside of the groups that originally collected these data to use them.



For example, I've spoken with a number of modelers who primarily use gridded topographic data, but when a model doesn't converge, they want the ability to look at the underlying profile data used to create that grid.

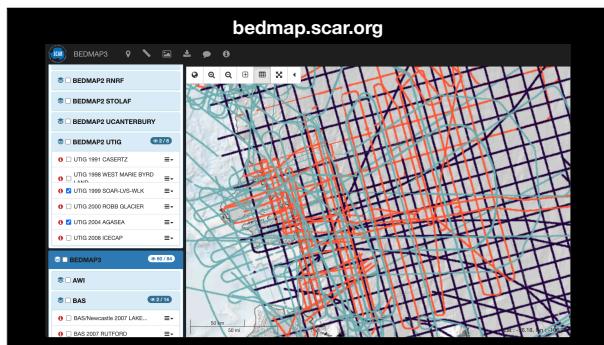
For example, in this radargram, we can see a valley where the gridded data probably underestimates ice thickness because the ice/bed interface wasn't identified in the radargram. And by plotting the surface and bed picks on top of the radargram (shown here as red dots), the researcher can see the interpretation that provided ice thicknesses for the grid.



Another example might be an early stage grad student—they're reading broadly, in search of a project.

As a concrete example, let's say the student just finished reading this 2017 paper by Ben Smith et al. presenting data about the surface expression of a set of active subglacial lakes beneath Thwaites Glacier.

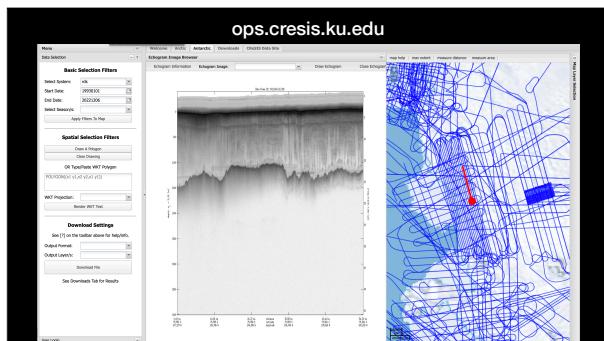
And the student is curious about the basal interface and englacial layers in the vicinity of these features. So ... what would it take to indulge that curiosity?



The SCAR bedmap data portal is good for seeing what campaigns collected data in a given region.

I've zoomed into Thwaites, and there are so many radar lines that they obscure the coastline.

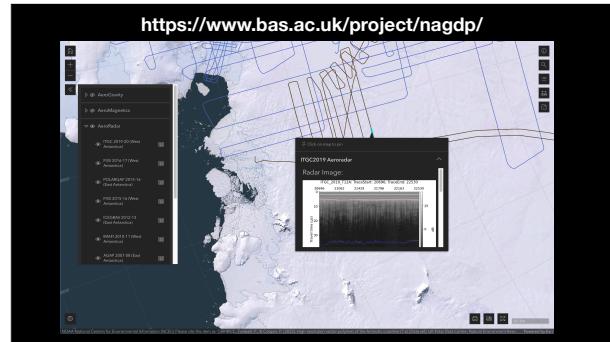
We can see that UTIG, CRESIS, and BAS all have relevant datasets.



To find the CReSIS data, we'll go to Open Polar Server.

Clicking on a line lets you preview the radargram and download it right from the webpage.

However, you have to eyeball where the features of interest are, and which lines cross them.



BAS recently released a very similar portal.

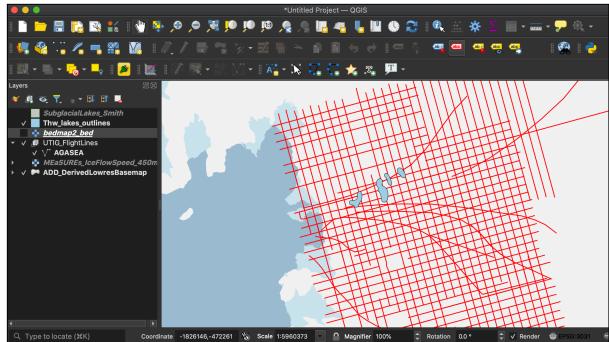
The screenshot shows a dataset information page from the USAP-DC. At the top, there's a navigation bar with links for Home, Contact, Search, Services and Support, Other Resources, Contact, and Log in. Below the navigation is a section titled "Dataset Information" with a sub-section for "Radar Sounding Observations of the Amundsen Sea Embayment, 2004-2005". This section includes a "Data DOI" link (<https://doi.org/10.15783/etds/129464458>), a "DOI" link, and a "Cite" link. The main content area contains a detailed description of the dataset, mentioning the AGASEA project, the data source (Pine Island Glacier and Thwaites Glacier), and the processing steps. It also includes a map showing the study area and a small thumbnail image of a radar cross-section.

And finally, the AGASEA radargrams were released via the USAP Data Center.

It doesn't have a slick web interface...

This screenshot shows a dataset page with a "Supplemental Docs" section containing a "README_601436.txt" file. Below it is a "Data Files" section listing several files: "93.7 GB", "69.5 kB", "24.1 MB", "125.4 kB", and "100.8 MB". A red arrow points to the "100.8 MB" entry, which is labeled "UTIG_FlightLines.kml". The page also includes a note at the bottom stating, "This dataset has been downloaded 19 times since March 2017 (based on unique data/IP combinations)".

... but if you scroll down, you'll find a KML of all the flight lines.



Drag it into QGIS, where we already have the lake outlines from Smith's supplement.

Then it's easy to identify exactly which lines we want.

Supplemental Docs
View Download README_601436.txt

Data Files

File Name	Size	Description
Large dataset. Contact us to arrange download.	93.7 GB	Large dataset. Contact us to arrange download. MD5 Checksum: e47250076c6edab0cb359a5d20242e0001 File Type: XLS, Adobe Portable Document Format (PDF), Readme Text File, Google Earth, NetCDF
G01IG_Radartrack	68.5 kB	G01IG_Radartrack MD5 Checksum: 5303a029039444a4e4e0110c0d95 File Type: Text File
MasterMap.pdf	24.1 MB	MasterMap.pdf MD5 Checksum: ee7250076c6edab0cb359a5d20242e0001 File Type: Adobe Portable Document Format (PDF)
README.pdf	125.4 kB	README.pdf MD5 Checksum: 5d6eb3c2031e10507feab2609321d1 File Type: Readme Text File
UTIG_FlightLines.kml	100.8 MB	UTIG_FlightLines.kml MD5 Checksum: 100bd5d770ba1f31160064e0d0c1 File Type: Google Earth

This dataset has been downloaded 19 times since March 2017 (based on unique date-IP combinations)

Then, back to the portal, where we discover that downloading the data requires sending an email.

A human will respond with a download link to data on Google drive.

That's just finding and downloading the data.

Each file is in a different format, and the student will have to figure out how to:

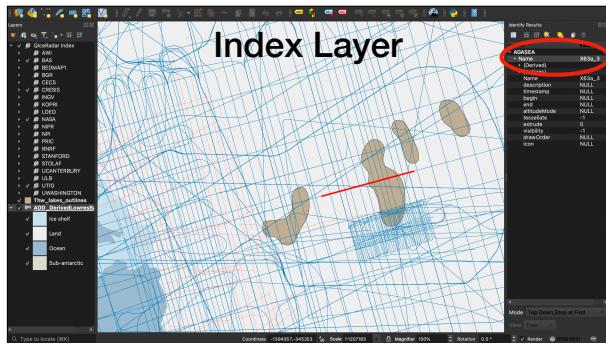
- * plot the radargrams
- * Plot their GPS track
- * Associate traces on the radargram with points on a map

In all, this is hours to days of work for somebody who is somewhat comfortable with Python. It's not insurmountable, but it's also not trivial. I imagine that it would probably be pretty daunting for a non-programmer.

Planned Functionality

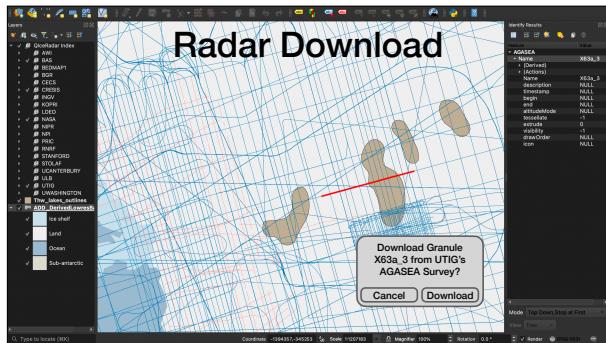
- Index Layer
- Radar Downloader plugin
- Radar Viewer plugin

I think this should be a lot easier, so I proposed QIceRadar. It will have 3 components.

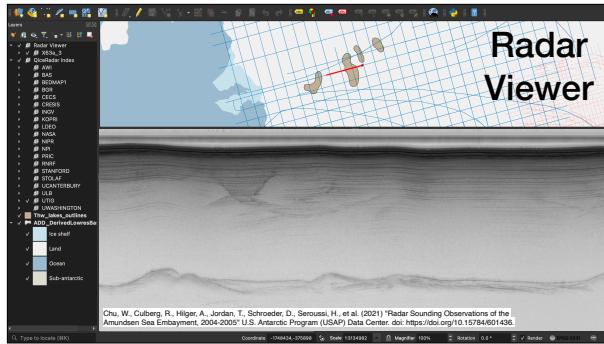


First, there will be an index layer including all known radar tracks, color-coded based on whether the data is available.

We click on a line in the map and can see in the “Identify Results” pane which survey, segment, and granule contains that data.

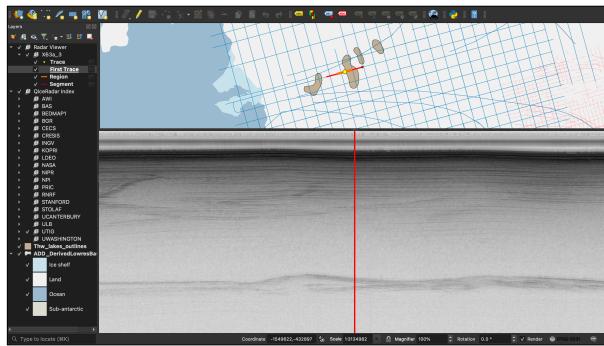


Then, clicking on that line with the radar downloader plugin will open a dialog offering to download that data.

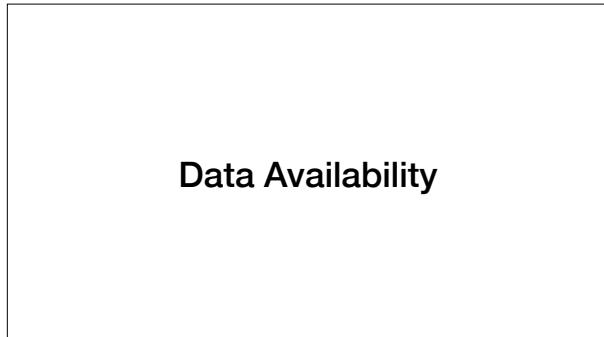


Finally, clicking on the line with the radar viewer plugin will pop up a radargram in a separate panel.

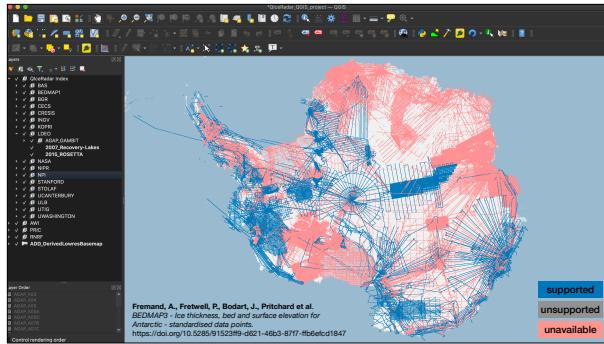
The extent of this granule is shown on the map view in red, with a dot indicating the first trace.



As the user zooms in and out, the currently-displayed section of the radargram will be highlighted (here, in bold orange), and the trace under the user's mouse will appear as a dot.



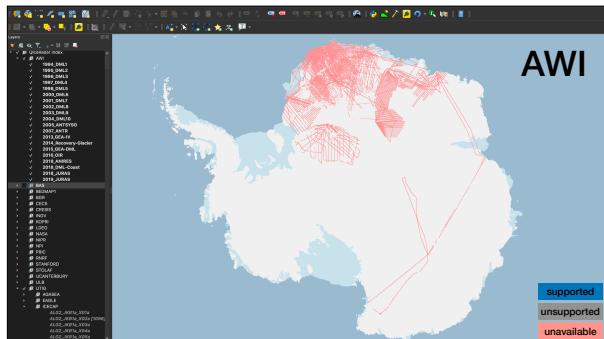
So, that's what it will do.
But what data will you be able to access with this tool?



This map is based on the dataset of ice thickness measurements released alongside BEDMAP3.

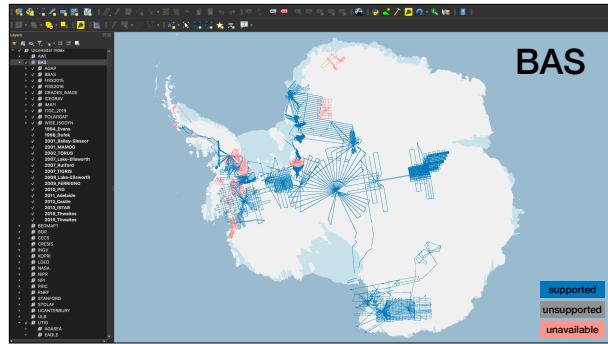
- * GPS tracks for radargrams that I have found and downloaded are shown in blue. All of these will be viewable in QIceRadar.
- * Tracks for data that I haven't yet found are in salmon. If you see any lines that shouldn't be pink, please let me know where to find the data!

This is a very busy map, so I'm going to break it out by data provider. For all of the large providers I'm about to show, QIceRadar will index their data and support their format in the radar viewer, IF they release all of their data in a self-consistent format.

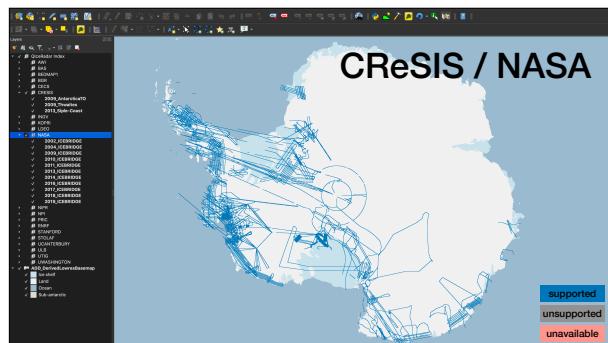


AWI hasn't yet released their data yet, but Olaf says they're working towards a portal like BAS, though they can't commit to a delivery date yet.

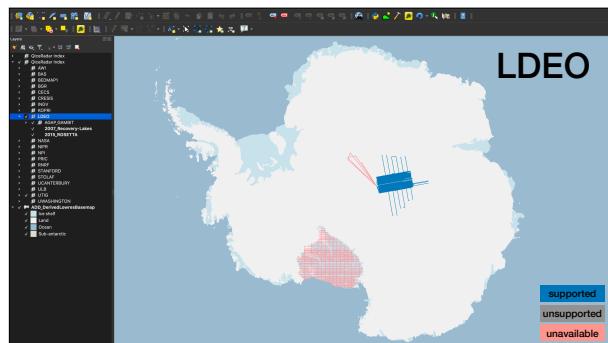
If they release their data by the end of 2023, it will be included in the initial QIceRadar index.



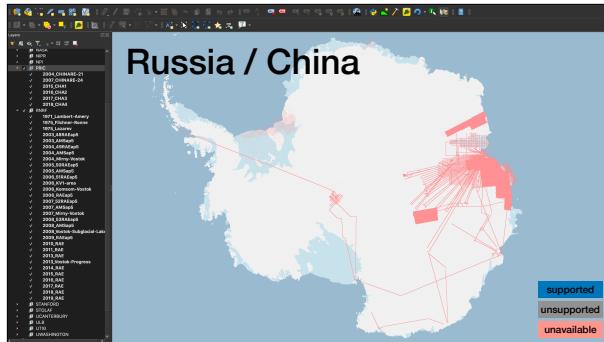
BAS has released all of their more recent data, though the older stuff is still a work in progress



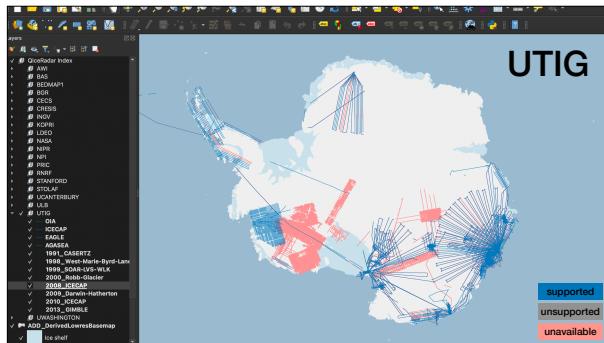
CReSIS of course posts everything to their own server and/or NSIDC



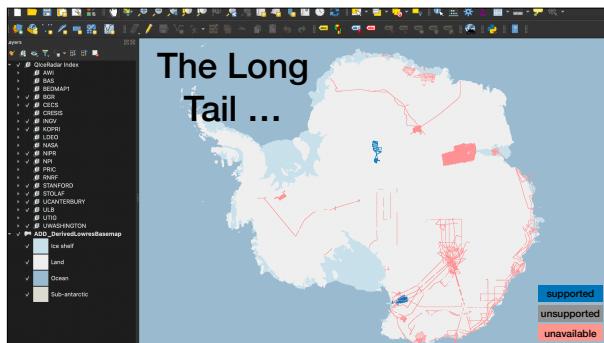
LDEO has released the AGAP-GAMBIT survey, but I haven't found radargrams for ROSETTA.



I haven't found any data from either Russia or China.



And UTIG has released most of their data since 2004, with the exception of the GIMBLE survey in West Antarctica. The old SOAR surveys are not available, with the exception of Lake Vostok, which was released by LDEO.

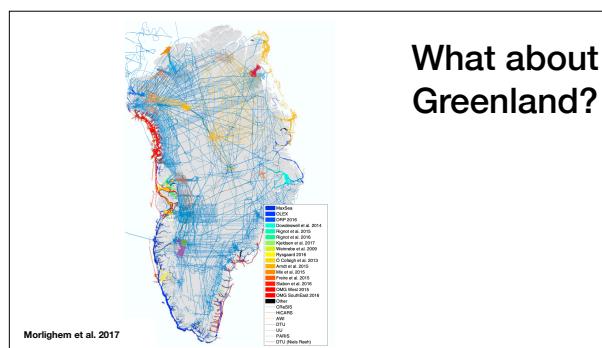


And there's a long tail of smaller data providers, where I don't have the budget to write a custom importer for each of their formats.

Instead, we will provide the ability to include their data in the index if it is released in a supported format and available for programmatic download on a per-line basis.

If this describes your data, or you'd like it to describe your data, and

you'd like to be an early tester of this pipeline, get in touch!



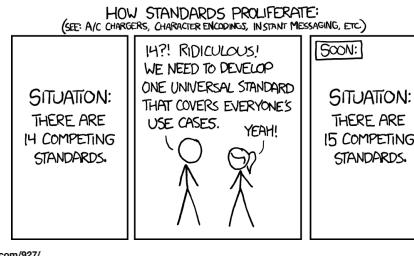
What about Greenland?

Greenland is also in scope, and I'm working with QGreenland to make an index layer that's compatible with their project.

Challenges

Unfortunately, if somebody came up to me today and said “Your plans for QIceRadar sound awesome! How do I get my data included?” I don’t have fantastic answers. There are two big technical challenges that the community is in the process of working on.

Data Format



First, what format should you release your data in?

Right now, every data provider releases data in their own format, and unfortunately, each is too specialized for me to recommend it as a general standard.

The Open Polar Radar project is working on this problem, and I plan to support reading data in their standard. However, I'm already fielding questions from researchers who need to release data now and can't wait for that standard to be developed.

Data Hosting

Data Center	Provenance?	Programmatic access?	Per-line downloads?
NSIDC	Y	Y	Y
UK PDC (BAS)	Y	Y	Y
AADC (AAD)	Y	Y	Varies
USAP-DC	Y	N	Varies
Zenodo	Y	Y	N
Self-Hosted	N	Y?	Y?

Next, where should you put your data?

To start with, just post it somewhere! Anything is better than nothing.

In order for tools like the one I'm creating to work, the features that matter most are:

- 1) What I'm calling provenance. The data needs to be citable, and there needs to be an expectation that the data will be available at

the same link for decades.

- 2) Next, there's enabling programmatic download. Tools like this can assist users to download data, but if it requires sending an email or solving a captcha, that's going to be a bottleneck. Instead, I think that requiring login credentials is generally a better solution.
- 3) Finally, make the data available in small chunks. Neither the end user nor the data center is happy if the scientist has to download 10GB of data in order to access the 100MB transect that they care about.

I'm not super familiar with it, but it appears that the marine seismic community has figured out an OK data portal: <https://www.marine-geo.org/index.php> offering direct download of individual lines.

Admin/Management

Timeline

- Fall 2022: Download & index all available data
- Winter/Spring 2023: Develop v0
- Summer 2023: Share with beta testers
- Summer 2024: Release v1 at SCAR-OSC
- June 2025: Funding ends

Advisory Committee

- | | |
|---|---|
| • Kenichi Matsuoka
(NPI Quantarctica, RINGS) | • Nick Holschuh
(Amherst) |
| • Twila Moon
(CIRES QGreenland) | • Lenneke Jong
(AAD) |
| • Duncan Young
(UTIG Open Polar Radar) | • Joseph MacGregor
(NASA) |
| • John Paden
(Kansas Open Polar Radar) | • Felicity McCormack
(Monash) |
| • Alice Fremantle
(BAS Bedmap3) | • Ted Scambos
(CIRES) |
| • Christine Dow
(U. Waterloo) | • Martin Siegert
(Imperial College London) |
| • Jamin Greenbaum
(Scripps) | |

I am the sole PI and software developer for this project, but it also includes an advisory committee representing related projects (including Open Polar Radar, Quantarctica, QGreenland and bedmap3) as well as end users in a variety of disciplines.

They will provide a pool of beta testers, as well as feedback/discussion at yearly meetings.



Photo: Laura Lindzey