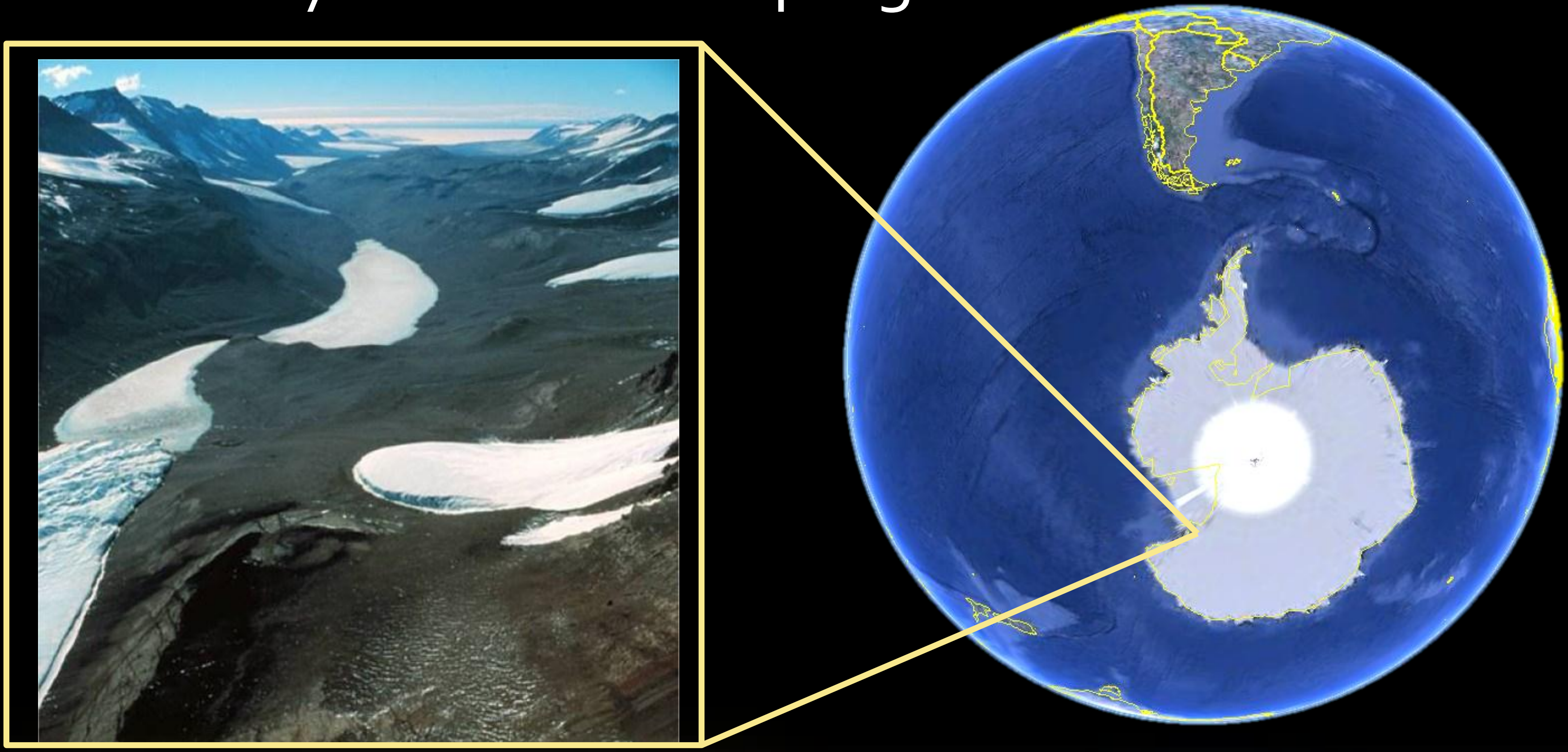
Alessandro Febretti

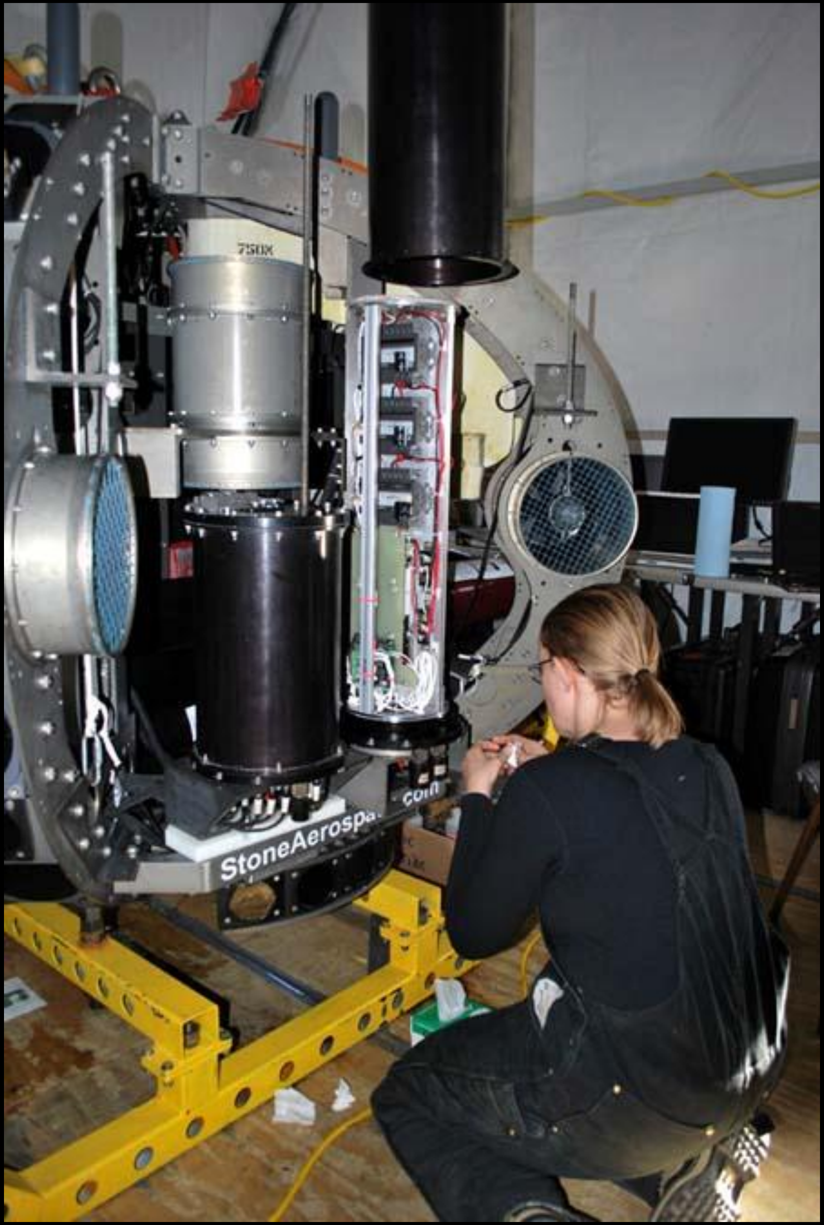
…FOR SCIENCE!

Building visualizations, managing dynamic workflows and coding without swearing too much.

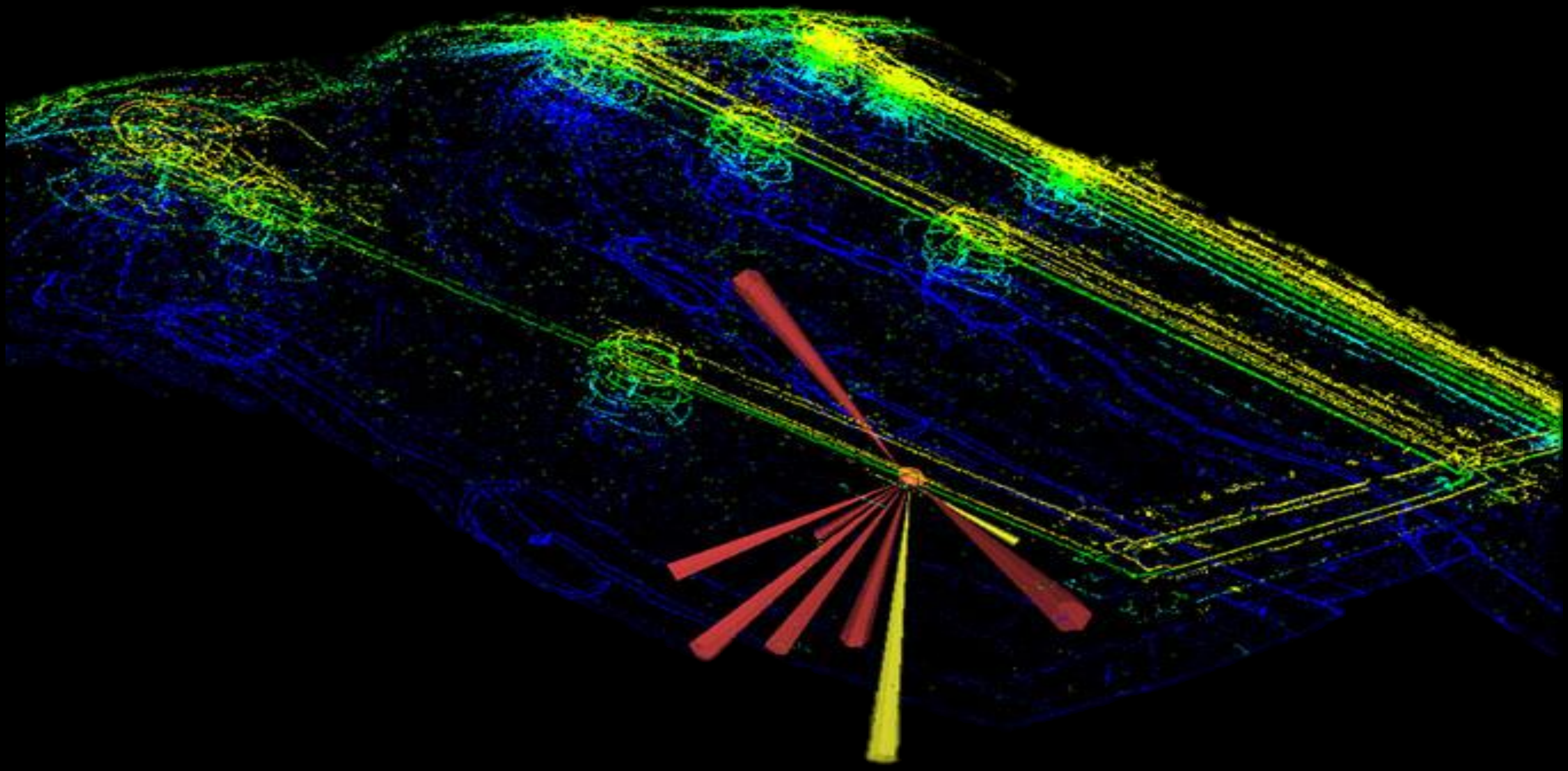
ENDURANCE

- **Environmentally Non - Disturbing Under - ice Robotic ANtarctic Explorer**
- Funded by NASA ASTEP program

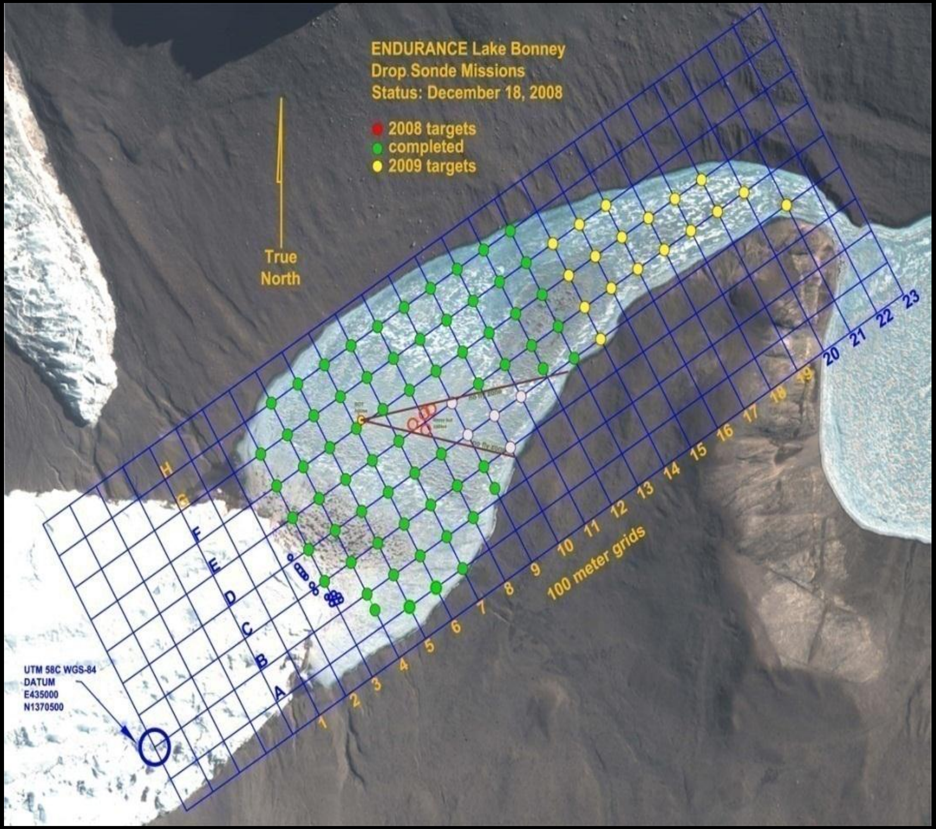


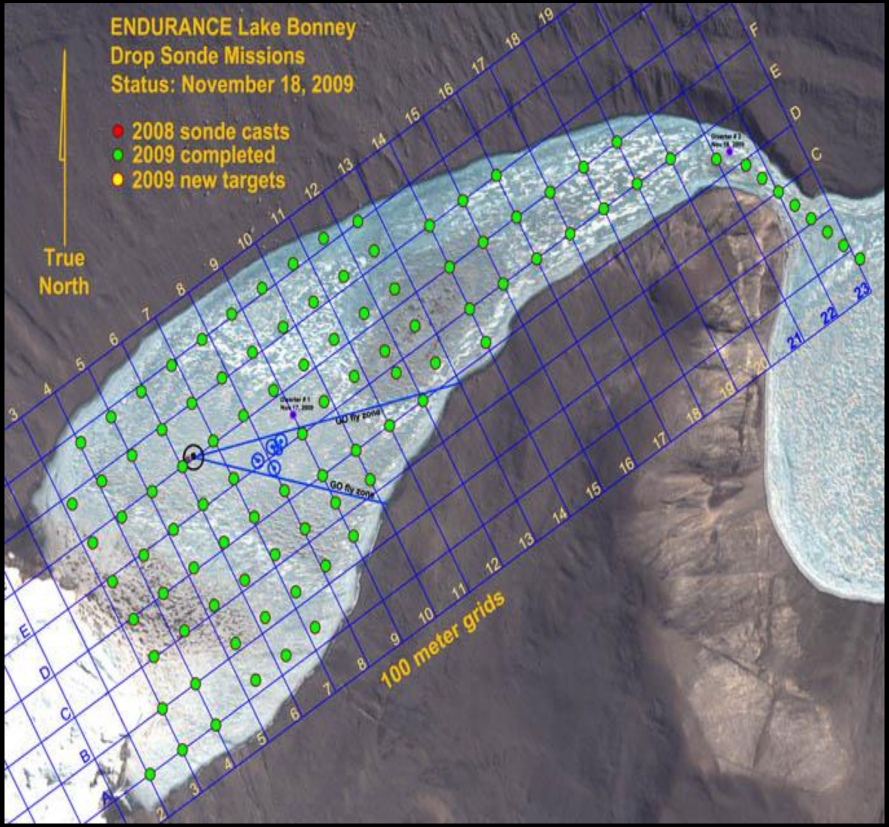

750X
StoneAerospa
com

Data collection

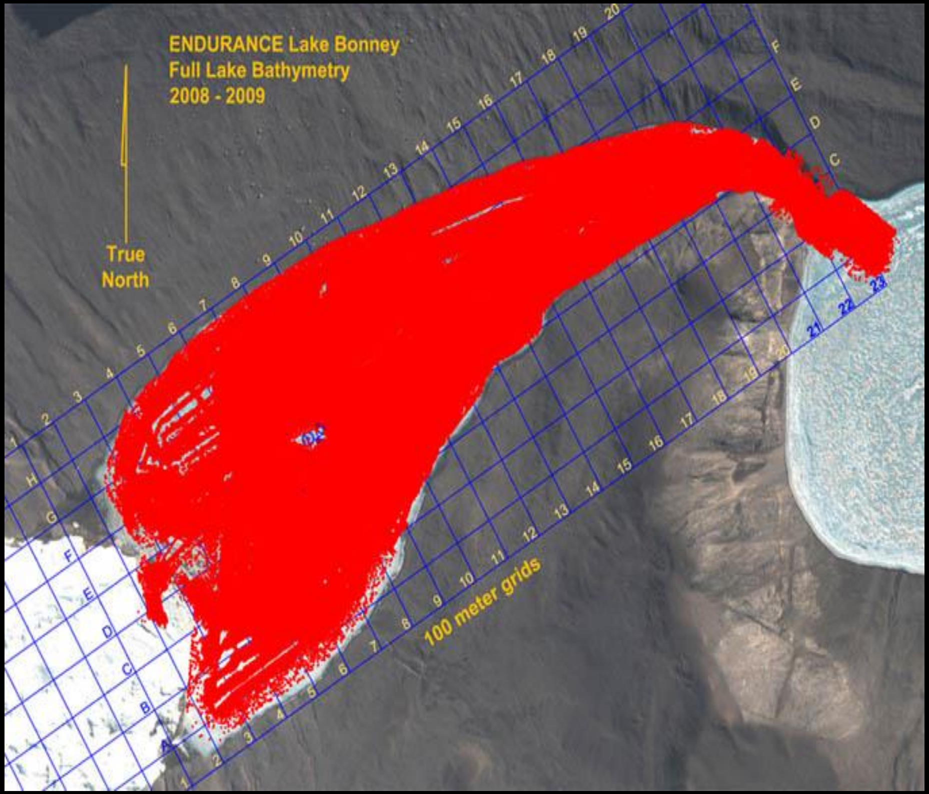


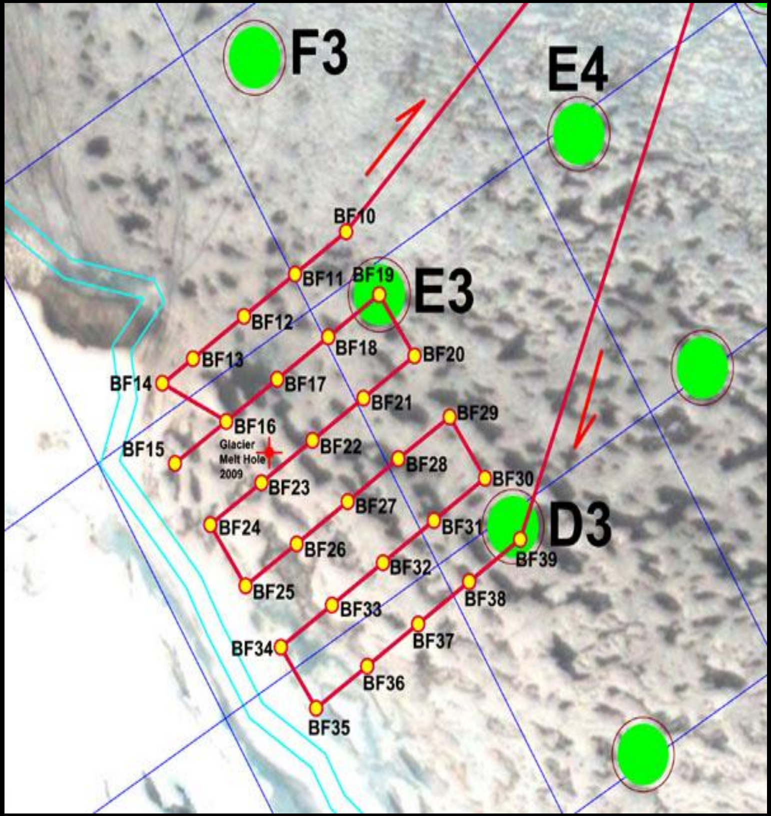
2008, 2009 Missions



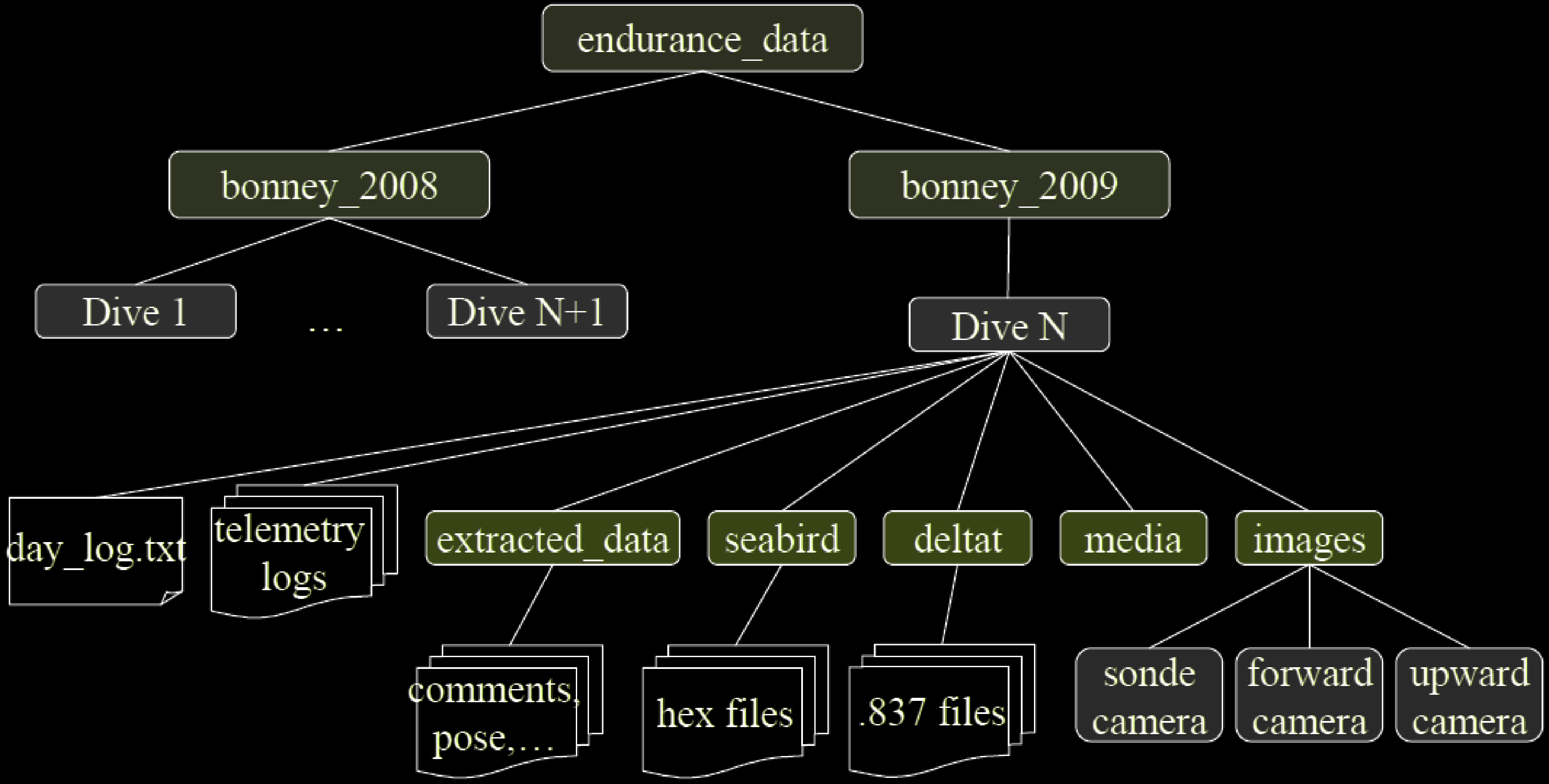


Sonar & Fine Grid Sonde Drops





Dataset structure



ENDURANCE: Main tasks

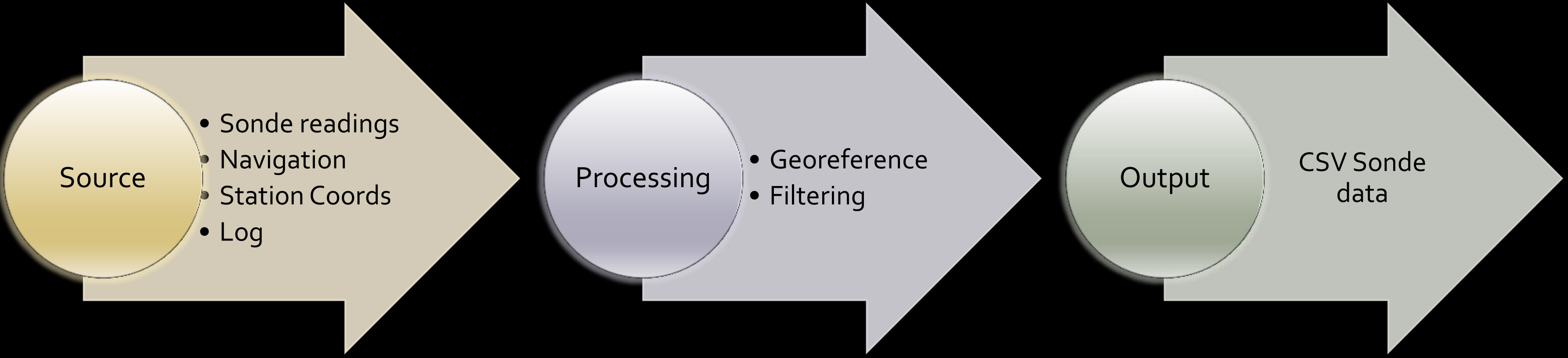
- Correct sonde drop data
- Process AUV images
- Create new lake model
- Develop visualization app(s) that:
  - aggregate and visualize processed datasets
  - Allow intuitively ‘querying’ the data
  - Generate additional data products (plots, csv files, etc)

Dynamic workflows

- Data may go through processing steps before being ready for visualization
  - existing tools
  - Custom tools
  - Manual vs automated steps

- Workflow = data sources + processing blocks + output products
- Static vs dynamic workflows

Sonde data



Source

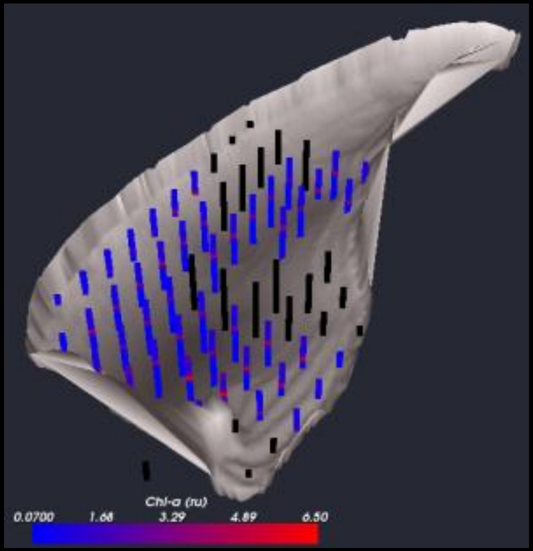
- Sonde readings
- Navigation
- Station Coords
- Log

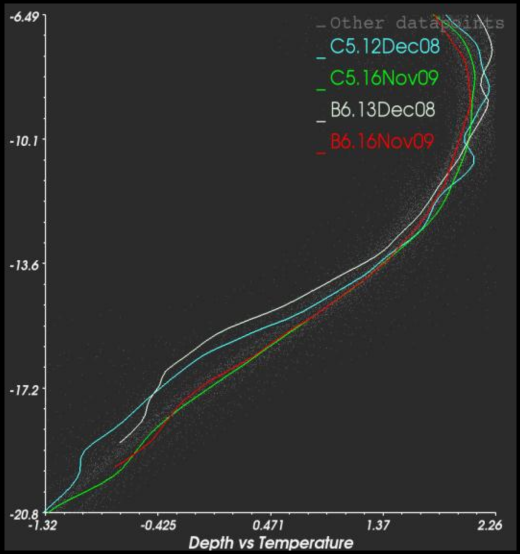
Processing

- Georeference
- Filtering

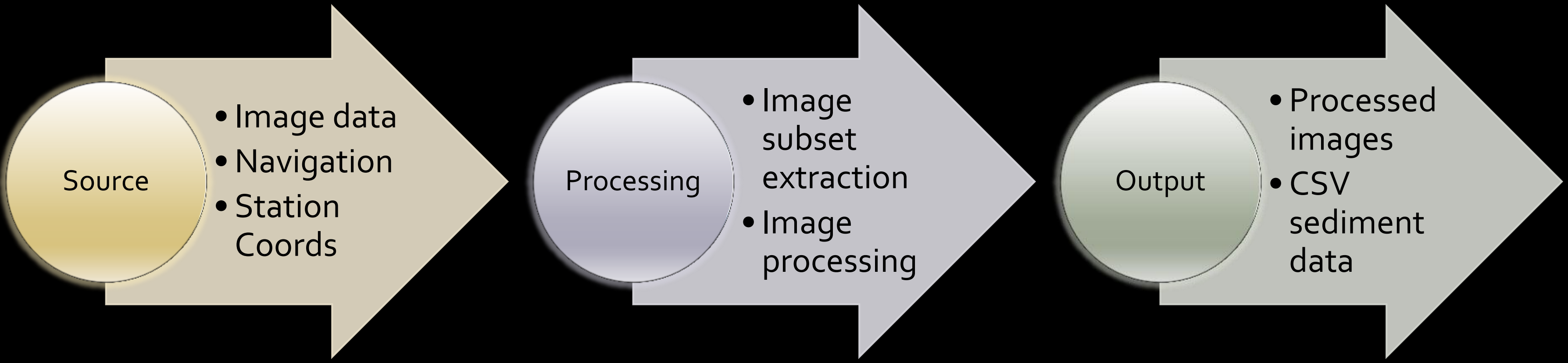
Output

CSV Sonde data





Sediment images



Source

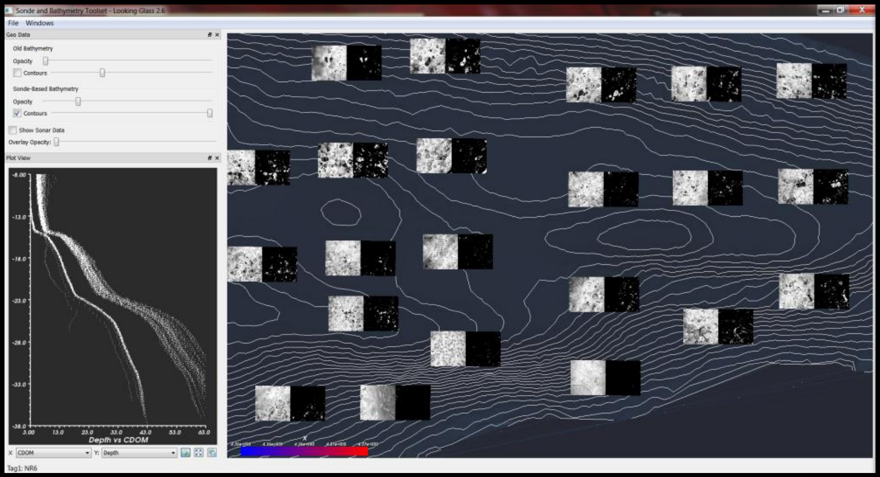
- Image data
- Navigation
- Station Coords

Processing

- Image subset extraction
- Image processing

Output

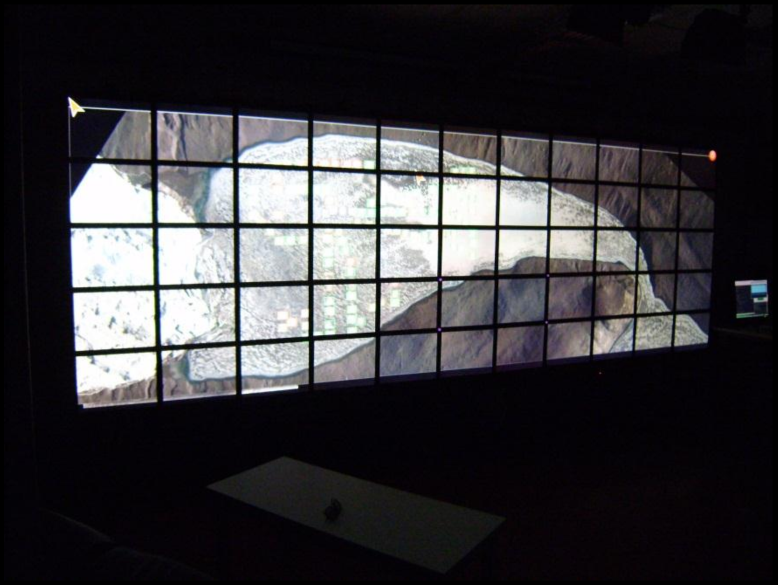
- Processed images
- CSV sediment data



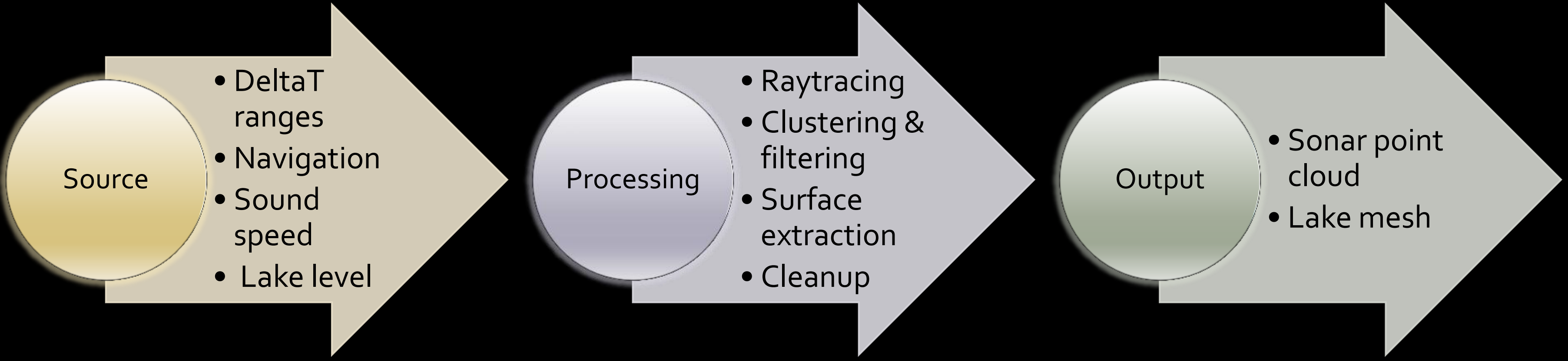
Sonde and Bathymetry Toolset - Looking Glass 2.6

X: CDOM Y: Depth

Tag1: NR6



Bathymetry



Source

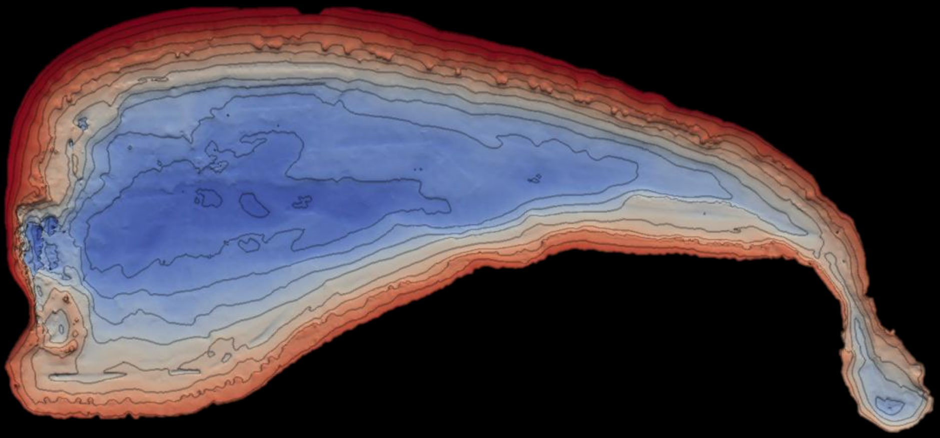
- DeltaT ranges
- Navigation
- Sound speed
- Lake level

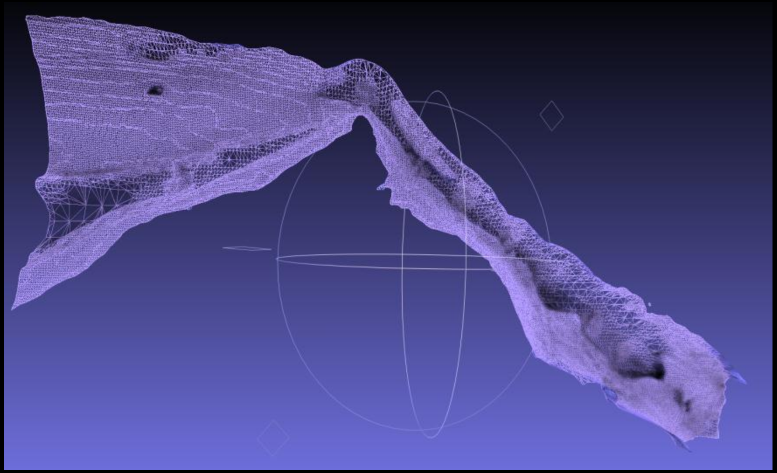
Processing

- Raytracing
- Clustering & filtering
- Surface extraction
- Cleanup

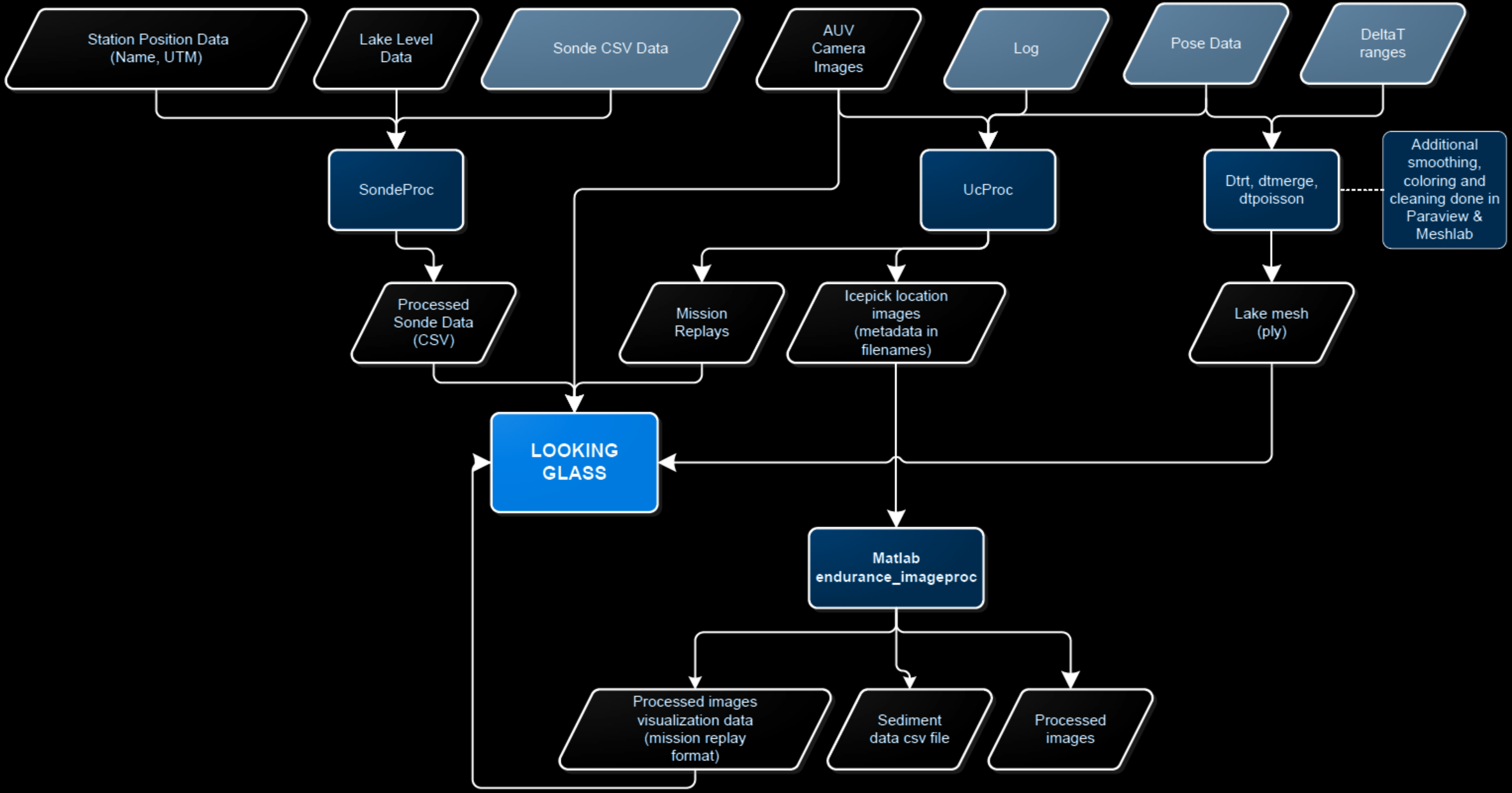
Output

- Sonar point cloud
- Lake mesh





ENDURANCE Workflow



Station Position Data (Name, UTM)

Lake Level Data

Sonde CSV Data

AUV Camera Images

Log

Pose Data

DeltaT ranges

SondeProc

UcProc

Dtrt, dtmerge, dtpoisson

Additional smoothing, coloring and cleaning done in Paraview & Meshlab

Processed Sonde Data (CSV)

Mission Replays

Icepick location images (metadata in filenames)

Lake mesh (ply)

LOOKING GLASS

Matlab endurance_imageproc

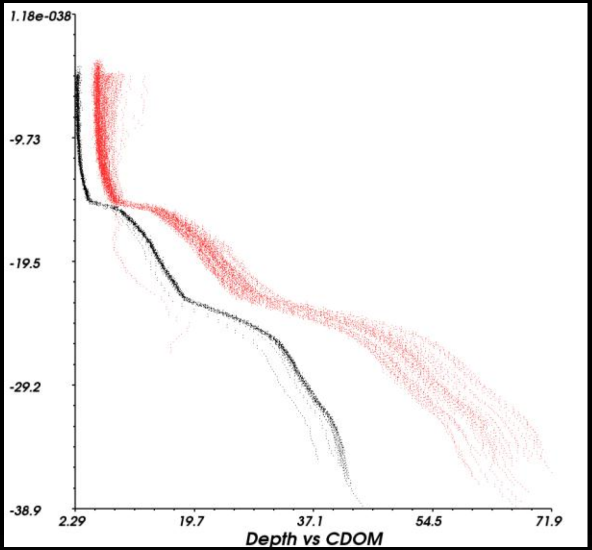
Processed images visualization data (mission replay format)

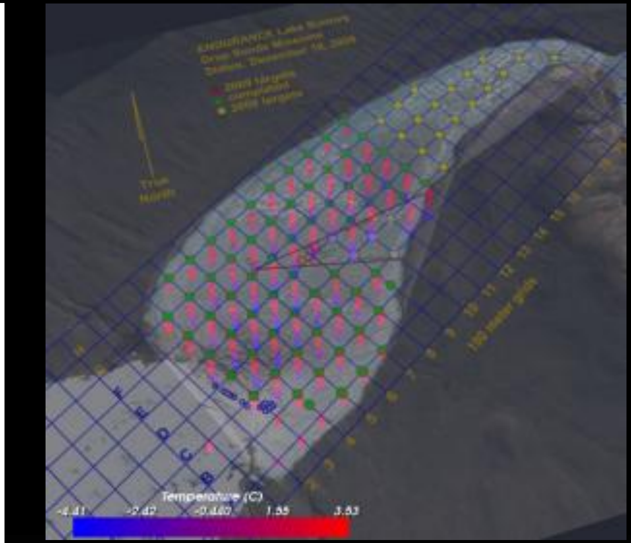
Sediment data csv file

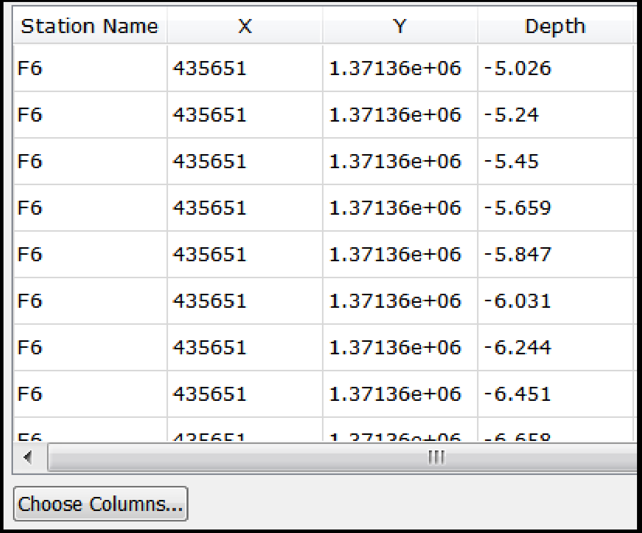
Processed images

Putting It All Together

- What do researchers want to **SEE** in the data?
  - Define Views
    - What data do you need?
    - In which format?
    - What tools?



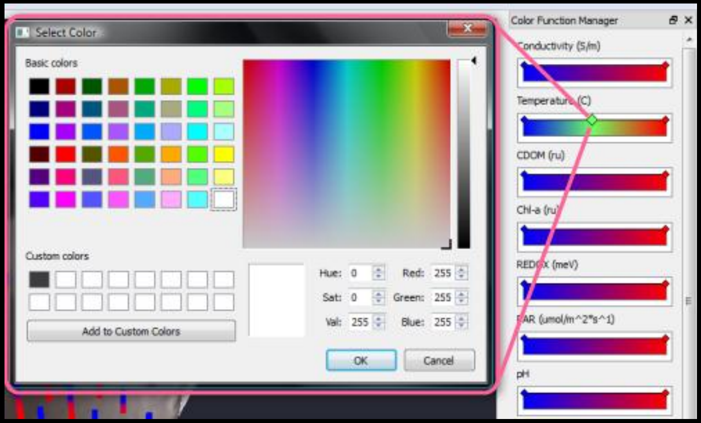




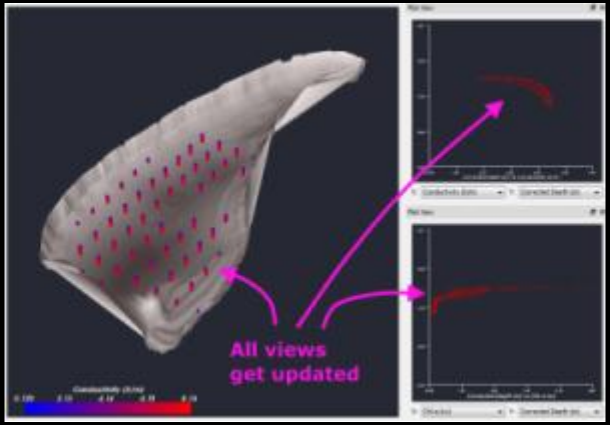
Station Name	X	Y	Depth
F6	435651	1.37136e+06	-5.026
F6	435651	1.37136e+06	-5.24
F6	435651	1.37136e+06	-5.45
F6	435651	1.37136e+06	-5.659
F6	435651	1.37136e+06	-5.847
F6	435651	1.37136e+06	-6.031
F6	435651	1.37136e+06	-6.244
F6	435651	1.37136e+06	-6.451
F6	435651	1.37136e+06	-6.658

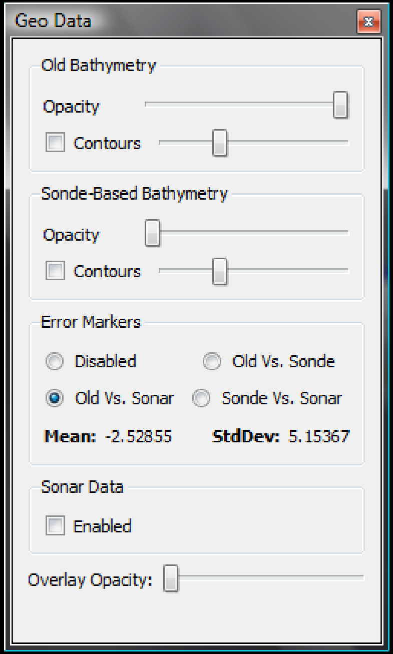
Putting It All Together

- What do researchers want to **DO** with the data?
  - Define Operations
    - Querying vs Manipulation
    - Local vs global operations
    - User interface



Hue: 0 Red: 255
Sat: 0 Green: 255
Val: 255 Blue: 255





Old Bathymetry

Opacity

☐ Contours

Sonde-Based Bathymetry

Opacity

☐ Contours

Error Markers

○ Disabled ○ Old Vs. Sonde

● Old Vs. Sonar ○ Sonde Vs. Sonar

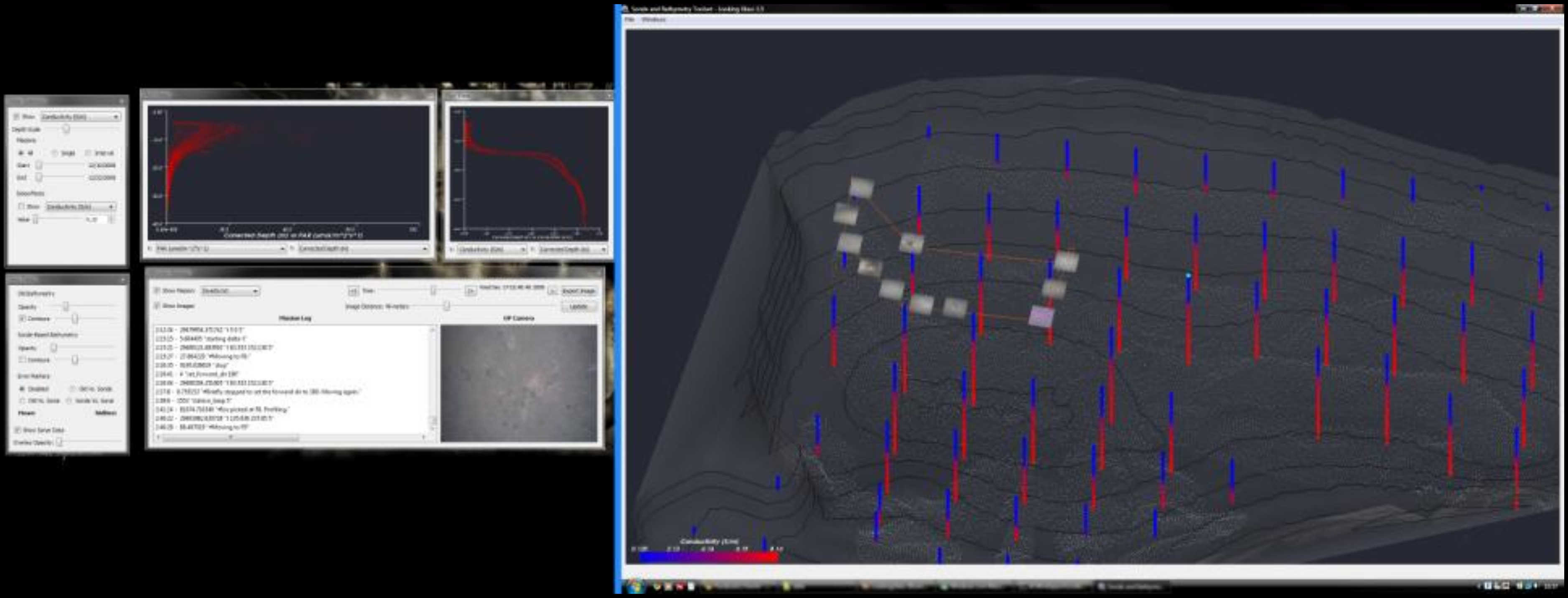
Mean: -2.52855 **StdDev:** 5.15367

Sonar Data

☐ Enabled

Overlay Opacity:

Looking Glass

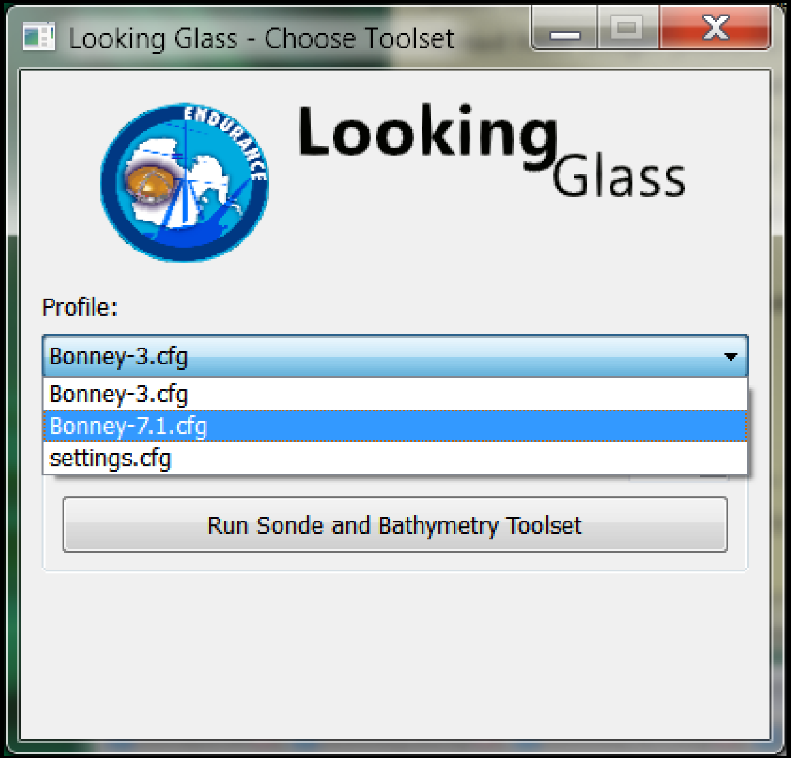


Looking Glass: challenges

- Data is changing all the time
    - Content AND FORMAT
- Different people from different teams modify the data
- Tools in the application change as well
- Old versions of the application may not work with new data

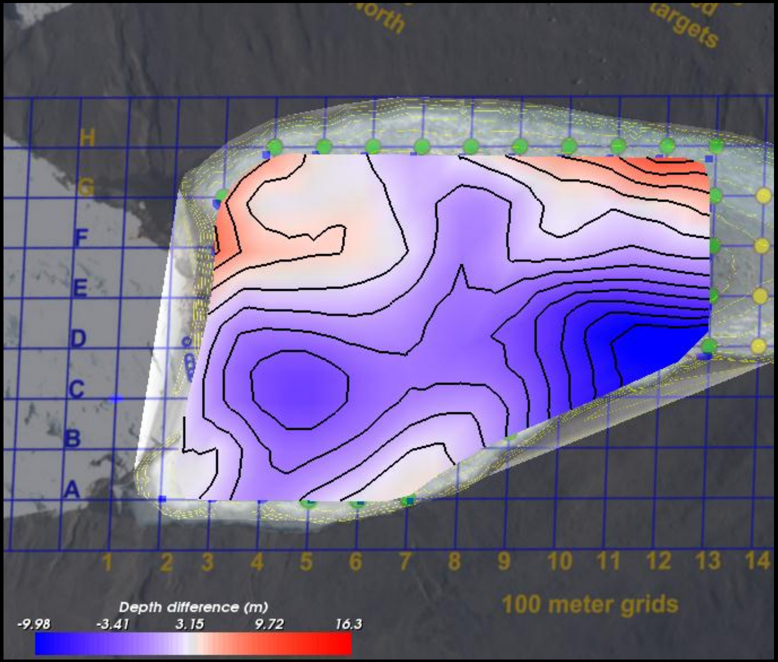
Solution: DATASET PROFILES

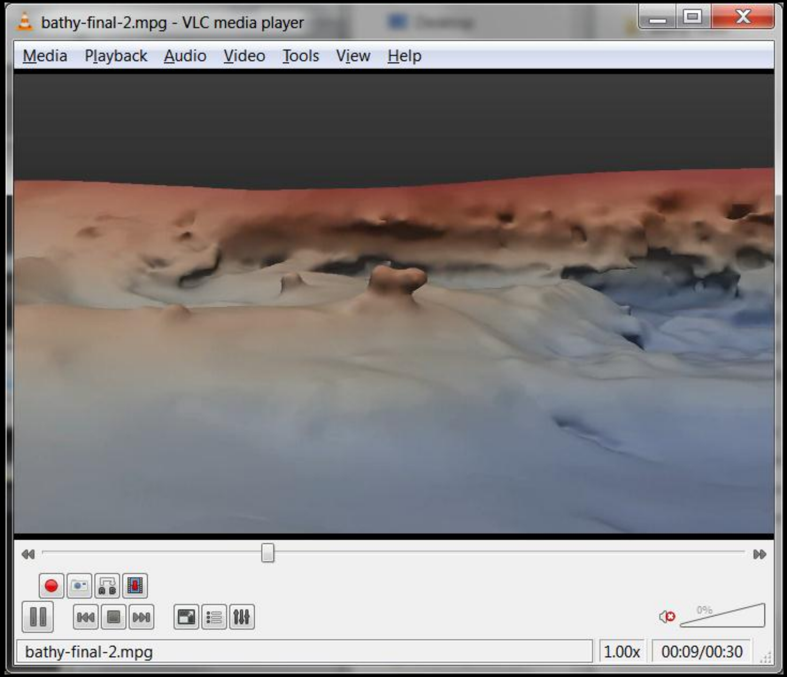
- Specify the content of dataset and version of the viz tools
  - Each version of the data runs it’s own toolset
  - Unavailable data or tool binaries get downloaded at startup



Additional Data Products

- Can be workflow byproducts
- One-time vs high frequency products
- Exploit existing tools





Lessons: Data management

Make every data transformation a REPEATABLE PROCESS

- You may have to re-do it in the future
- OTHERS may have to re-do it
- Tracks provenance of data products

Lessons: Data management

Batch / automated processing is nice
but don't overdo it

- It will need maintenance
- How many data items / processing iterations you have?

Lessons: Programming

Don’t Reinvent The Wheel!

- Use standard libraries / APIs
- Customize existing code to fit it into your project (beware of licensing)
- Glue + script

Example: dttools

- Library code (139’000 LOCs, 98%)
  - Raytracing (mbsystem)
  - Linear algebra: (vmmlib)
  - Octree building / traversal: (ravec)
  - Poisson surface reconstruction (tool source)
  - Surface normal estimator (vcg)
  - Mesh cleaner (vcg)
- My code (2’400 LOCs, 2%)
  - Tool configuration
  - Data conversion
  - Input / output

Lessons: Programming

Don’t Fall In Love With Your Code!

- Version & throw away freely.
- Form is nice, but function wins.
- Prototype a lot.

That’s all, Folks!

Alessandro Febretti

…FOR SCIENCE!

Building visualizations, managing dynamic workflows and coding without swearing too much.