...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
Data collection
2008, 2009 Missions
Sonar & Fine Grid Sonde Drops

ENDURANCE Lake Bonney
Full Lake Bathymetry
2008 - 2009

True North

100 meter grids

F3

E4

E3

D3
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

- 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
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?
Putting It All Together

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