NASA Arctic-Boreal Vulnerability Experiment (ABoVE)
User Stories for the ESIP Earth Science Collaboratory

0.0 Preamble

- Authors:
  - Aulenbach, Steve (saulenbach@usgcrp.gov)
  - Griffith, Peter (peter.c.griffith@nasa.gov)
    http://scholar.google.com/citations?user=7Ra4418AAAAJ
  - McWhirter, Jeff (jeff.mcwhirter@gmail.com)
  - Wee, Brian (bwee@neoninc.org, http://orcid.org/0000-0002-0038-9381)

- Sources of information in this document include:
  - Ref 1: Report of the NASA Arctic Boreal Vulnerability Experiment (ABoVE) Workshop – 13 to 15 June 2012, Boulder, Colorado
  - Ref 2: Peter Griffith's data management doc titled "ABoVE Data Management"

1.0 Background

ABoVE is a NASA sponsored field campaign studying impacts of climate change on the Arctic and Boreal region. Climate change in the Arctic and Boreal Region (ABR) is unfolding faster than anywhere else on Earth, resulting in a longer ice-free Arctic Ocean during summer, warming and thawing of permafrost, increases in the frequency and severity of climate driven disturbances, and widespread changes to surface water extent, soil moisture, and vegetation structure and function.

To more fully understand the evolving ABR environment and provide the information required to develop options for societal responses to the impacts of ABR climate change, the Arctic-Boreal Vulnerability Experiment (ABoVE) has been proposed as a NASA-sponsored field campaign. The research conducted to address these questions would emphasize observations, analyses, syntheses, and modeling. (*Ref 1, Pages 3 and 4*)

The ABoVE Information System is not currently implemented.

2.0 Goal

The development of a data and information system would be another important component for ABoVE. The ABoVE Information System would serve the field campaign as a short-term repository and clearinghouse for all data sets collected and data products produced as a result of ABoVE research. It would provide access to other datasets that would be used during ABoVE that were generated from other ABR research projects, products from land management agencies, and from long-term monitoring efforts. (*Ref 1, Page 5*)

Needs include:

- Serve the field campaign as a short-term repository and clearinghouse for all data sets collected and data products generated as a result of ABoVE research. (*Ref 1, Page 5*)
- Provide access to other datasets that would be used during ABoVE that were generated from other ABR research projects, products from land management agencies, and from long-term monitoring efforts. (*Ref 1, Page 5*)
- Provide access to the results and assessments being produced through modeling and other analyses to a wide range of end users. (*Ref 1, Page 5*)
- Provide support for experiment planning during ABoVE. (*Ref 1, Page 5*)
- Researchers and scientists sponsored through ABoVE work from a unified suite of base maps/datasets (e.g., digital elevation maps, climate downscaling, surficial geology in common projection, resolution, and format). (*Ref 1, Page 28*)
- Data management plan and policies to facilitate data preparation, sharing, discovery, integration, and analysis. (*Ref 1, Page 28*)
- Database of ABoVE funded projects. (*Ref 2*)
- Information artifacts related to ABoVE project, such as those for field notes (hopefully structured in some way!), meetings. (*Ref 2*)
- Daily log of activities: where / when / what measurements had been made by whom. (*Ref 2*)
- Publications, reports, assessments to be shared among the team of scientists, and they require access rights to be associated with them.
- Managing the modeling life cycle.
- Documenting derived data products and their provenance.

3.0 User Stories

3.1 Problem: Who's who?

The list of participants is growing like kudzu. At any given moment, it consists of many people, each person filling one or more roles, representing one or more operational groups, from one or more organizations. On bad days, this information is captured in several spreadsheets and managed by several people. On good days, the information is harmonized and collected into a single, centralized spreadsheet. All of the spreadsheets have different structures.

Discussion:
The project team and the participants are spending a lot of time managing basic information about the people who make up ABoVE. Mistakes are being made and there is an impressive amount of duplicated information. A valid participant list with all the appropriate organizational information is foundational to many of the ABoVE activities and, eventually the ABoVE products. We need a simple way to manage this.

3.2 Problem: Sharing documents

The ABoVE Science Definition Team (SDT) will hold their first meeting in May. Before the meeting participants will want to share their presentations and documents with other participants. Many of the files are too large to mail to the ABoVE project management NASA team.

Discussion:
This is a common problem across the sciences. Meeting participants want to be able to share their presentations with others. Many ad-hoc solutions are used - email, ftp upload, shared drives (e.g., Dropbox, GDrive), and custom web forms.

3.3 Problem: Science Bibliography

The ABoVE project team has assembled a bibliograpy of articles that are relevant to the science mission of ABoVE. They have both the bibliographic information as well as electronic copies of the articles. Access to these documents needs to be restricted to members of the ABoVE SDT but other users should be able to see the bibliographic information.

Discussion:
In the geoinformatics world we often focus on the data and do not provide much support for the broad range of digital stuff that makes up the science endeavor.

3.4 Problem: Site and Project Information

The ABoVE project team has put together (in a spreadsheet) a detailed survey of data centers, projects and instrument sites. The intent is to provide information about what kinds of data and products are currently available in the project area to help the ABoVE SDT in determining the science project plans. This spreadsheet contains information regarding the location of sites, the project leaders, sponsors, types of instruments, etc.

Discussion:
This information has gotten rather broad and deep and is pushing the limits of what a spreadsheet can offer in terms of access and organization. We need to support querying, browse and mapping of this information space.

3.5 Problem: Capturing Real-Time Data

It's late August 2013. Almost 8 years to the day after Katrina struck tropical storm Imogene is strengthening and becoming a hurricane as it approaches New Orleans. A number of University researchers want to compare the approach Imogene is taking with the tracks and weather data (models, radar, soundings, satellite) from Katrina. After some considerable effort they found text ATCF track data from a NOAA repository and some level 2 radar from NCDC. The software package they used could not read that version of the radar data and they had to write a custom script to convert analyze the track data. As to the other data from 8 years prior? Not much is there.

Now fast forward 3 days…
Imogene strengthened and hammered into New Orleans. Luckily the levees held but there was much damage along the coast. The same group of researchers wanted to create a case study of all of the relevant meteorology data. Luckily they have a Unidata IDD feed so most of the data was local – model, radar, metar and satellite. This consisted of about 20 gigabytes of data from 3000 files. They want to capture this data and make it available for the future. Hopefully they can use something more advanced than FTP and HTTP serving of files.

Now fast forward 1 year…
An undergraduate severe storms lab is studying hurricane Imogene. They are generating time series of images of satellite and model data. They want to generate a QuickTime movie and publish it to an external repository like YouTube. But, their movie has spatial and temporal metadata and they want to publish it along with the other data that makes up the Imogene case study.