2015-12-04 Combined Project Brainstorming Meeting notes

Attendees

- Jason Votava, Jong Lee, Rui Liu, Smruti Padhy, Praveen Kumar, Brock Angelo, Luigi Marini

Goals

- A 90 minute meeting to discuss some crucial dependencies developing between the SEAD/Clowder, Browndog, Geosemantics, and IMLCZO projects. We will consider what compelling technologies can be put in place to help us get over some tipping points and scale these projects through synergy.

Discussion items

Item	Who	Notes
Introduction	Praveen	- SEAD - data and science life cycle (data curation, data flows) - uses project spaces, extractors, data publication, data providence - Brown Dog - data software interopterability, deep patterns in data - LIDAR (pulse), LIDAR (wave form), topographic, hyper-spectral, AMERIFLUX, landsat - Geosemantics - model data interoperatibility - semantic ontology services, models as web services, model data, meta data, sensors, observations - IMLCZO - large heterogeneous repository with interested, active users. Critical zone science that can embrace other efforts - Appears to interact with other projects in an opportunistic way more often than a strategic way from the perspective of the researchers - First three projects are tech based, IMLCZO is science based - Purpose of this meeting is to bring the research students and developers together to find a way to integrate these efforts more effectively
How the tech pieces interact and inter-operate	Luigi	- Individually, science does tend to drive development based on need. - It is only recently that project spaces from Clowder were available to be shared across efforts and be part of the same build. - Brown Dog is not a client as much as services. Tools built for BD are available to all users across any data collection that has been granted access to that tool. - Extractors and the data tilling service all unite as part of Clowder - Geosemantics uses data services and modeling services - SEAD contributes work toward Clowder GUI development and a virtual archive/BCPR/long term storage library - Brown Dog contributes DTS (Clowder), DAP (polyglot) and uses DataWolf and specific tools and extractors - Geosemantics contributes to Clowder, and model/data integration service - IMLCZO contributes a result of Geodashboard and Clowder - Tech is heavily interdependent, but not always visible on the user side/front end
Brainstorming	All	- what other things could be integrated across multiple projects? - there is less emphasis on data analytics compared to data extraction - how often will these services be deployed locally (on your own laptop or for your own department or for ...)? These are designed for server deployment, perhaps this isn't made clear - usability of these services requires a certain technical learning curve at this point. this is planned to improve as development continues - there is a gap between what the results or data pieces are and how these data products are used. - this can be bridged in part by closer communication of what is needed or specific capabilities that are available. - informal demos - documentation review - tools catalog - work flow schematics - Not everything can be automated, there is always going to be some level of manual work between tools. This effort increases as the specific needs or specific data becomes more particular and finely defined. How is this work minimized and as painless as possible?

Action items

- [ ]