

October 14, 2022 Joint Methods & CI Working Group Call

Attendees: Matt Brousil, Jessica Burnett, Jody Peters, Hassan Moustahfid

Regrets: Quinn Thomas, Jake Zwart

Agenda/Notes:

1. [Data Ingest, Cleaning, and Management](#) Task View
 - a. Matt is moving to a new position. Still needs to figure out to what extent he can lead this. Will talk to his new supervisor
 - b. Have a lot in the Task View so far.
 - c. Problem, solutions section is nice. This is something Alexey came up with so we can ping him again if need be when we get there
 - d. From Hassan: something that could be added on machine learning, deep learning - quality control, data ingestion and cleaning
 - i. Cloud resources and cloud sandboxes
 - ii. Hassan can contribute to common data standard - Darwin core and cf convention standards
 - e. Something to consider for the future (another task view, or how it fits in with this one or other ones) is long term management of data
 - i. This could also be a point in the cloud computing section saying "this is something to consider"
 - ii. Could go along with web-based and commercial storage options
 - f. Cloud computing might need to be a section itself or under data storage access
 - i. From Jessica, just thinking about \$\$\$ associated with data volume and length of time its stored in various components of the workflow and where, not necessarily documentation/archival
 1. I think saying at least, "you should prob think about this" is good too :)
 - g. Best practices mentioned in multiple places - data ingest and data cleaning and harmonization
 - i. Should we have a separate best practices. Is there one best practice over all or are they different for the different steps?
 - ii. Are there different best practices for ingest vs cleaning and harmonization?
 - iii. Wait to see what kind of info go into these sections and then see if it goes within other sections or becomes its own section
 - h. Next steps
 - i. Matt to check
 - ii. Ask people in the group to add resources
 - iii. Next call

2. Forecasting Standards and [Challenge CI](#) Update
 - a. Biodiversity & Ecological Forecasting meeting is May 8-12 in College Park, MD
 - i. Have flexibility for a morning meeting - so good to check with Mike and Quinn about options
 - ii. From Jessica: I am not too familiar with either the NEON/EFI challenge nor the following, but might consider reaching out to NASA EF group re: NASA Space Apps Challenges. The hackathon had something like 20k participants this past round, and Ketih Gaddis is heavily involved with the program. <https://www.spaceappschallenge.org/blog/your-challenge-awaits/>
3. Forecast visualization challenge pitch (this came up on the Sept call, think about if this is something the group wants to move forward with)
 - a. Leaving this as a placeholder, but not quite ready to move forward this, right now.
4. Workshop Proposal
 - a. [Google doc](#) to collectively develop ideas
 - b. The following are notes from the Sept call that Jody is leaving in for reference.
 - c. Identify gaps in CI/methods for forecasting and put forth NSF proposal for a workshop/series of workshops for \$50-100K. Want something modular and interoperable. Want to pull in people with CI and computing resources
 - i. NSF doesn't need to send out these types of proposals for review
 - ii. Mike has pitched a roadmap white paper that the community agrees to what community CI should look like - how it is designed, what are the pieces it should include.
 - iii. Then could ask across agencies to pay to support that CI. No one agency wants to pay the bill for this. NSF can't operationalize
 - iv. Jake has been talking with the interagency working group. This is an issue - there aren't forecasting centers or offices not tied to the National Weather Service. So it is hard to have sustained funding for operational products
 - v. At USGS there are more projects that want to do short term forecasting. Over the next decade there will be more need for this, but don't have a plan to scale up
 - vi. People at NASA already implement but it is boutique since they don't have a forecast center
 - vii. There is a WG within ICAMS that works on CI - this is the big issue in that community as well
 - viii. US NPN used to get funding. Don't anymore because of changeover in leadership.
 - ix. Key word = boutique. If you build community tools, it will cost less since everyone will not need to recreate forecasting workflows

- x. We need a plan for the CI - that is the pitch for the workshop.
 - 1. Get \$50-100K from NSF and can get NASA/USGS to put in that much as well. Want to get agency buy in so that they will pay attention to the recommendations and put effort behind proposed solutions
 - a. Suggest reaching out (or Jessica can) to USGS Mark Wiltermuth for support.
 - b. re: USGS -- check out the USGS Comm for Data Integration funding opportunities. Requires at least 1 USGS PI, smaller grants, typically projects occur over a few to six months (not including writing). Best POC is Leslie Hsu (lhsu@usgs.gov). I will try to post a link to the annual call, but I think it might require individual signup and login (but doesn't require to be USGS employee).
 - a. Also consider ESIP Lab small grants. <https://www.esipfed.org/lab>
 - ii. Need a champion - don't think it would be hard. Think it would take a couple of weeks to pull together
 - 1. Think about the timing for this adopted
 - a. This upcoming year could be good from USGS
 - b. There will be a meeting with USGS and NOAA around modeling efforts that include forecasting
 - c. There will be a forecasting strategic plan written in 2023 for USGS
 - d. NSF and NOAA are having their own pair wise conversations
 - iii. Next steps - open a shared doc on another call to collaboratively work on an outline of the steps and the pitch for the proposal
 - iv. Then have a champion(s) shop it around to agencies
 - v. Hassan is working on a work plan that goes along with this and may have funding available. Still working through this, so no guarantee. But good to stay in touch with Hassan about this
 - 1. Big interest in moving from research to operations
 - 2. If forecast are within 6-7 readiness level then think it is possible to application (not quite operations yet).
 - vi. This concept of providing a framework for researchers to avoid boutique apps is a priority of NASA Earth science data systems (ESDS) group, and increasingly of NASA Ecoforecasting. Jessica is happy to help with helping field agency partnerships and socializing the concept. Alexey is also a good contact for ESDS. share the 1-pager with NASA BDEF (biodiv and ef) when developed.
 - vii. [National Partnership of the Agencies](#) (NOP) that come together and call for proposal -

1. There was a recent call for biodiversity that just called, but think there are lots of calls that would be
 2. <https://nopp.org/category/funding-announcements/open/>
 3. <https://nopp.org/2019/ioos-ocean-technology-call-for-proposals/>
- viii. USGS Community for Data Integration (CDI) has seed grants for topics around data management, assimilation, UI/UX dev, HPC, etc. Heavily focused around using or incorporating USGS assets. SOIs are due September annually, with funding periods over the following summer/fall.