# LTER IMC Monthly VWC (February 12, 2018)

# Planning for IMC meeting at ASM 2018 Part 2 (Part 1 was on December 11, 2017)

In attendance: Gastil-Buhl, Suzanne Remillard (notes), Tim Whiteaker, Dan Bahauddin, John Porter, Renee Brown, Jason Downing, Don Henshaw, Johnathan Walsh, Stevan Earl, Wade Sheldon, Yang Xia, Kris Hall, Corinna Gries, Marty Downs, James Conners, Jim Laundre.

Meeting starting with the context from previous meeting presented by Wade and John:

Open discussion throwing ideas out. Some topics IMC focused (moving ball forward) vocab, taxonomies, etc. How can we leverage more broadly. ClimHy should be informed with discussion by domain scientists. What is the next generation product? What are the science use-cases?

New NCO features may accommodate some of the ideas presented for some of the topics (bibliography, personnel, etc.).

Shared approaches, prebuilt technologies for new sites or sites revamping.

Brainstorming specifically on IMC portion of the meeting.

John and Corinna working on vocab.

Gastil: Should we post these topics and get feedback? The topics may need to be grouped.

EDI has made a lot of progress formulating what they would like to achieve at the ASM and are applying for resources. How can IMC interact or participate?

Corinna: most of these are LTER projects, so there should be good opportunity for participation. Working on EML 2.0 with annotation field. Need to learn how to use ontology terms in annotations and connecting our attribute names to something more comparable.

EcoComDP (A dataset design pattern for ecological community data) is still a work in progress. Poplar is another project that is an LTER group that is doing very similar things to EcoComDP. These products should be introduced to domain scientists. This working group product has been very successful. Can we use this for other domains, i.e., climate? Yes, this will be useful for other types of data. We need scientists to tell us what form they would like the data in. Perhaps, could use ODM (observation data model) from Consortium of Universities for the Advancement of Hydrologic Science, Inc. (CUAHSI).

Code generation and analysis tools – EDI is making progress on this. It will be introduced at ASM. There will be a hackathon before the ASM.

Gastil: strategies for a multi-repository data catalog

John P: new reporting requirements; session on meeting NSFs reporting requirements for publications. NSF requires a copy of the publication that they can legally share on their website (no longer citation). Peer reviewed publication goes through new system. But still Bibtxt for other publications. For peer reviewed journals, need DOI and link to site. Need PDF/A version to upload. For articles that journals do not allow posted, can post author copy.

Jonathan W: Mary Martin has been dealing with this and we may want to learn her experience from her directly.

Marty: ASM timeline submission form in about a month for working sessions. Another month to review and organize. There will be categories on the form to link sessions. Can apply for a double session. There will be 7 slots of concurrent sessions (10 tracks).

Stevan: Corinna/EDI ideas are the most fleshed out ideas and are really interesting and related to LTER. What is your vision? What is the status? What do you envision doing to move the annotation forward? EcoComDP? Climate, nutrients, etc.

Corinna: EcoComDP is pretty advanced, but now needs to integrate into other communities. Eric Sokal will get other scientists together. Julian will hopefully work with Corinna to get ways to integrate. ClimDB is only idea right now and needs more brainstorming.

Suzanne/Don: HydroDB was stongly pushed by USFS (Doug Ryan) and other scientists to integrate hydro into climate. The USFS funds really initiated the push and was followed with additional NSF (Henry Gholz) funds for sites to get their data reformatted and submitted.

Could climate or hydrology data go into NCEI (National Center for Environnemental Information <https://www.ncei.noaa.gov/>) or CUAHSI. Corinna has been doing a lot of thinking about all of this, but needs more ideas from others.

Don: Is the climate group planning on meeting this year? If so, we should approach them.

EcoComDP could reside in any data repository, but still be available to work with the data. It is a robust model and would be advantageous to use this approach for other types of data. The idea is that anyone can convert a dataset into the EcoComDP and then it would available for analysis no matter which repository it resides in. So, there is no central infrastructure to maintain the data, just outside code for discovering and accessing (e.g., an R package).

John: ASM is a meeting about getting things started (thinking and discussion), but not completed products. It introduces people working on similar things and is a platform for creating product oriented working groups.

Gastil/Marty/Wade/John: Tagging of datasets with core research areas. What does this mean to tag a dataset with a core research area? This came up in the exec discussion last month. They talked about signature datasets. Tagging datasets with core research areas is more of a bean counting approach than for functionality. Most people have concluded that this is a pretty useless metric. It is more of an accounting exercise.

Are there recommendations or guidelines of what are signature datasets?

It’s a concept of something that comes and go. Our efforts should be how to make our datasets more discoverable rather than tagging to core research areas.

Gastil: IMEXEC should work on a spreadsheet with all of these topic ideas and then they will be presented to the community for comment.

There is a new EML Best Practices (V3) that was completed in October 2017 and is on the EDI webpage.

Don: We’ve gotten away from doing reports at our annual meeting. It would be worthwhile to do a quick overview of the accomplishments over the year from EDI, IMEXEC, VWC, etc. As an editor, DataBits is revealing a lot of interesting information about what has been happening in various arenas. By suspending DataBits for a couple of years, we’ve missed this important sharing mechanism.

Yang: It would also be good to have updates from active working groups too.

Gastil: IMEXEC gets reports from active working groups monthly and these are reported in the monthly notes.

Marty is trying to work kinks out of bibliography and would like to work with one site to work out kinks. She prefers to work with someone who works in EndNote. John P. volunteered after March 31. Yang volunteered now.

EndNote relies on its own identifier and it may be hard to reconcile the same publication.

Next topic: Not yet formalized, but may be practicalities of using BCO-DMO and ADC as alternative or parallel repositories for LTER data (multi-repository landscape).