

## Virtual Update Notes December 7 & 8, 2009 - The LNO Operational Plan



Published on *LTER Information Management* (<http://im.lternet.edu>)

Home > Virtual Update Notes December 7 & 8, 2009 - The LNO Operational Plan

---

## Virtual Update Notes December 7 & 8, 2009 - The LNO Operational Plan

Wed, 12/23/2009 - 4:27pm — administrator

Discussion and comments on the operational plan.

---

December 7 notes

Facilitator: Don Henshaw

Present: Don Henshaw, John Chamblee, Karen Baker, Margaret O'Brien, Inigo San Gil, Jason Downing, Suzanne Remillard

John - we need a sense of what activities are covered by ARRA and by core funding.

Karen - NISAC response: PIs cannot respond to a long doc on such a short timeline. they will discuss in their call on Wednesday.

Overall: Doc is quite organized and advanced, but is very traditional and isolated from the sites. Should be reporting to IMC at every operational step of each activity. Milestones should be expanded: small efforts like databits articles and large, like documentation.

John: also missing from milestones: deliverables.

General discussion on reporting bodies: what is the place for design planning - is it NISAC or IMC/IMExec. Many aspects of this plan will benefit from sites input first, and then to be brought to NISAC. Also, IMExec is not mentioned specifically in this plan, as a reviewer. By calling on the IMC as a whole, the doc seems to be requesting input from individuals instead. Clarify: anywhere it says "IMC" do they mean a coordinated response? in which case it should say IMExec. If it from any individual, it should be "IMC members" or IMs. Relates to governance: the IMC is historically defined as "26 reps" (someone correct me if I'm wrong) and that 'IMs' includes IM-teams at sites.

General disc on standardization of data values: not clear where (site or NIS) this will happen. Do they plan for data to reside at sites at all? What is the role of the site? By the way - we dont have any "standards" yet. how are these to be developed?

2 extremes: 1) sites describe all data perfectly (per local standard) and users will know what to do or 2) sites format data for each use (per use standard). Could also be that sites provide their 'workflows' for the workflow manager to generate these product for each use. Neither extreme is perfect, or will work for all synthesis projects. The impact on the sites of either is very high - and this is assuming that we already have PASTA-ready data.

Synthesis and the role of the NIM: must be clarified - this person should also support synthesis groups and their data needs (reformat, etc). Is

scope and detail: there could be considerably more technical info in the CI section. This could be expanded and put into an addendum. The main plan could be more general, and the tech specs could be designed to be changeable

Training: there are several typologies we could make use of, and some are discussed in this plan. We should consider all of these as training opportunities, and perhaps funding/scheduling should be more flexible. To be effective, training needs to be timed correctly and at the right knowledge level.

1. traditional, network-wide (or network-specific) for all IMs
2. traditional, network-wide (or network-specific) for scientists
3. traditional, domain-wide for IMs, for example a GIS management or web development class at a local college
4. traditional, domain-wide for scientists, a GIS class at a local college
5. site-network co-design, e.g., for a NIS a module or a common tool (this might also be a 'production workshop')
6. site-network during NIS-module use and assessment, ie, for a beta or recently implemented product. IMs learn while they give feedback.
7. site-site: co-design, e.g., projectDB, or 'tech-transfer' visits

IMs need better understanding of GUIDs (global identifier for data). Who assigns, who is publisher? (site? network?). Can data be co-owned? e.g., CCE and CalCOFI. Will existing structures work? (e.g, in EML).

---

December 8 notes

Facilitating: Margaret O'Brien.

In attendance: Corinna Gries, Donald Henshaw, Emery Boose, John Chamblee, Wade Sheldon, Jim Laundre, Hap Garritt, Eda Melendez-Colom, Teresa Valentine, John Porter, Ken Ramsey, Linda Powell

Margaret: We want everyone to have an opportunity to comment on the plan. Notes will be collated by Margaret, Corinna, & Don. Also, please get written comments to us.

Does everyone understand the comments: LNO is going to get extra ARRA funds. EB wants a long Operational Plan. This plan has overlap between core funds and ARRA. The two different funding sources have their own reporting scheme. Years 2, 3, and 4 will be hosting NSF. LNO Have to produce additional reports for ARRA to NSF because there is a different reporting scheme.

Emery: Emery has a question about the scope of NIS – what is role of individual site information systems after the NIS is up and running?

Wade: This is going to be on a site volunteer basis. In the manner of the data access server. Have the capabilities and encourage its use. No emphasis and orientation for a centralized architecture. Big reluctance at LNO to pursue that gain. At the EB at NISAC level, we need to try to get critical mass. We need to have a discussion about the appropriate level buy in and push to get these prototyped. Leaving population these to strictly voluntary, would be a mistake.

Jim: How will this affect site in terms of unfunded mandates. Also, this is a lot of money and I am not sure it will help individual sites in terms of doing their own work? Where are the sites? How can we accomplish any of our own work? Will this just be more work for us?

Corrina: This obviously our major concerns. We need constructive suggestions as to how we need to be included. How can we put language in there that makes NSF aware and that we expect continuous supplement funding to do these tasks. There is a division between how the IMC and the Information Managers training and there is the LNO with the NIS. We need to rephrase training so that training is used specifically for site-level input into the NIS, to train us in the technology, to help provide feedback on the use cases, and to develop the documentation.

Jim: We need to make constructive suggestions, we need to have tools developed for PASTA that are useful at the Site level. But it isn't training that I need, it is assistance.

Linda: When I look at all different values from different methodologies and all the different approaches, it is not possible to put everything in the same database, because we don't want users using data as if they all come from the same protocols. This is just at the site level, it is much worse at the network level. This is not a decision that LNO needs to be making. Will the EB make

these big decisions about standardization. These decisions belong with the EB, not the LNO.

Corrinna: They are ready to listen to us regarding to standards. But not data collection standards and change protocols at individual sites. Todd Crowell was on IMEXEC and the opinion is that Information Management can handle discrepancies. When Corrinna told them that the data collection was never intended to be comparable. It is not LTER to standardize.

Linda: If we don't do some standardization, then we'll never be able to develop these different databases. It does make a difference in terms of how you collect and process data.

Margaret: At some level, we need to these areas of responsibility. PIs need to be specific about what their needs are. In some cases, the request has been very non-specific.

Hap: "Standardization" is in quotes. It seems that what has happened in the network, there's are initiatives that started and created the standard from the start. And the intent behind EML was to allow some of the conversion, in theory, you should be able to convert unit and can then compare it. Ecotrends has done a version of this as well. Defining this without a question or them is really difficult. If you get into something like PASTA, you will use workflows based on theme. These are standards that will be question driven.

Don: I agree with Hap. The information managers were never able to accomplish this without scientists. This is the kind of a thing a synthesis workgroup, proposed by scientists. So the new information manager for LNO would work with synthesis group. Instead, the production workshop is about producing improvements to LNO. One problem is that IMExec should be able to review the production workshop. ImExec should be able to look at IMC needs and the propose production workshops. We need to prioritize.

Linda: We need to have the scientists more engaged with LNO, ImExec, NISAC. We have to move forward, this question of methodologies protocols, really have me stumped at my own site. When we had people check EcoTrends data, folks didn't like it because of all the assumptions. We might ask scientists how they really feel about the fact that we've never really collected data for comparative purposes.

Emery: The real challenge is the difference in methodology and don't know how to do that. Returning to the scope of the NIS, something like Wade's statement needs to appear in the plan, so that the NSF isn't surprised.

JFC: I wonder if rather than dedicating our training efforts to supporting NIS, we should simply say that feedback, more integrative approach with IMC are responsibility of LNO and those, along with Documentation need to be included in deliverables.

Margaret: Do we need to make training, etc. We need to align what we are

requesting with what we perceive the needs to be in one or two. Some of these things in terms of training

Wade: What we could propose that put into this document, we need to make sure that we get a process for building this more active partnership between the LNO development and the alignment with site activities. Include the groups that should be included. A more regular series of briefings and a more active series of exchanges and bringing development expertise across the network. This is going to be critical to addressing Jim's concern which I share that the LNO's development has been spectacularly failing to bring products that are useful to the sites. What we need is the right partnership model between the LNO and the sites. There's a possibility that these could be useful, but there's a need for greater partnership. We need cleaner site-to-network interoperability and there's a lot missing from this document in order to make that happen. There's also a real lack of coordination between science working groups and NIS development. Science working groups really need to be involved. The timeline here is so protracted, that if we don't parallelize some of this, and bring some of this forward, so that, four five years from now we have this expectation that it won't meet our needs but then it won't.

Hap: There's a lot of promises. I don't have the technical expertise to say is this do-able? What expectation will NSF have after six years? How do we define milestones and deliverables? How do we see how this performs as it goes along? If this some aspect of this doesn't work well and meet goals it will difficult to get support to do this again.

Corrinna: These workflows could be something that sites develop. I don't see LNO develop any workflows. Only we know the data. To Linda's comment: the only way we will be able to do this is via higher level products. The scientists and synthesis needs to be involved in the synthesis working groups. The IMs do need to be involved if there is to be some higher level product. There's never gonna be raw data than can be compared. These workflows are going to be developed by us. We will have to do this. Where in all this scheme is the work of each single information manager.

Hap: Are you suggesting we do this with a site budget or are who are we going to pay for it? Ecotrends had funding.

Corrinna: There's several project. There's going to be some higher level products we all need to produce. For that, we need continuous supplement funds. Then there's synthesis workgroups that they need to pay for in terms of information management for each synthesis project. This means additional funds outside of the operational plan.

Margaret: We don't have a distinguish between site-required synthesis and project driven synthesis.

Porter: I'm assuming that PASTA harvests our metadata, so all the sites really need to do is have good eml metadata. If we have good documentation, then all

we need is phosphorous measured one way vs. phosphorous measured another. This would be a table with lots of empty columns, but it depends on having good standards for documentation and data description. We will have to take that on and find resources.

Wade: Another thing that's missing from the plan is a statement about how network standards and controlled vocabularies could be layered onto these systems in order to handle the grunt work. If PASTA is correctly configured it would do this. There's this known weakness in the process because NSF knows that ARRA is incomplete. Todd Crowell knows that this is incomplete. What is missing in this document is any language addressing the known unfunded portions and acknowledging them.

Corinna: EB is aware of this and is actively pursuing different possible funding opportunities.

JFC: Will PASTA be able to be used to develop workflows? Is this supposed to be the tool which we use to do the workflows as they fall on us?

Wade: That's the expectation. But this is very immature. That's why we need to have a co-design process. It will start out using Christine's EcoTrends process and using R and Kepler. But we need lots more collaborative process and more network wide expertise.

John P.: A lot of data work goes from data cache to API. You need workflows to create synthesis products for NIS modules. In principal all you need to do this is good metadata.

Jim: I wonder if all this workflow stuff will expose areas where, even with EML, the standardization is insufficient. I see this is a big failure.

Wade: Agreed.

Don: In ClimDB/HydroDB we did standardization at the site. I'm not sure we've figured out how to do these standardization. Does it need to be done at Site, or is centralized? A different workflow for every data set seems unmanageable over the long term. Synthesis work groups – there should be a data manager associated with each of these groups, whether the NIM or a site information manager. There's discussion for each working group about standards. But there's requirement for Information Manager. There's no link between synthesis and how this will really work.

Ken: In research working groups, there's no reference to IMC at all. A little concerned that even if we have well documented EML and can get the data together, and still come up with scientifically inappropriate measures. In my recent experience, successful synthesis projects have been collaborations between scientist, IMs and LNO. We are not a "site perspective" on the CI, but are the network, when we work together. It looks as if LNO is making synthesis decisions. It seems key that synthesis involve NISAC. This document lends a

reading that everything could be done centrally with the NIM and NISAC and then the sites simply have to adopt it.

Margaret: Imexec was left out of early documents. Sorry Margaret, Ted G. walked in and I lost this.

Hap: Some of these are about data we haven't collected yet. Going forward, especially with SS coming on board, there's going to be social science data. Is there a way to short cut site level data for Social Science or new data in general so that we don't have to go site by site to harvest.

Wade: this should be the NIM's role. I think that too much of this document was too insular – there are too few points of collaboration between LNO and IMC/IMexec/Sites. Also there's not enough attention to governance. Leaving out governance is something that needs to be addressed. NISAC needs to be re-evaluated in terms of what's mission. What role is NISAC playing? Being a deciding body was really part of NISAC mission and it has been assumed by several groups that this is so. We need a better description and see how CI decision making is supposed to be done.

Ken: Assuming that this is annual thing. Going forward it would be nice to have more details in this document. We need details of what's going on and how its evolved and how what 's being produced. We need to have more up-front lead time to provide input. Going forward, if you have more opportunities, to make more comments.

Ada: Who's going to officially respond to Operational Plan?

Margaret: EB asked for collated comments from Corinna. Margaret and Don will be helping.

Ada: You need to pull in additional people to do this. Two people is too few.

Margaret: These meetings are to get feedback. If anything is left out.

Wade: Time frame is a problem.

Don: Are we filtering comments? Are we reporting everything?

Margaret: In general, we are getting broad comments, then specific comments. We could organize our feedback similarly. Summarize comments ,then include specifics.

Corinna: We have to summarize, but we can post it on the web and have people comment.

Wade: this is not the end of the process. The 15th it will go out an external advisory committee and they'll go back to NISAC and presumably IMC meeting. In late January Corinna and Wade will be involved in an EB/LNO/NISAC meeting on comments. There will be another comment period. We can then bring concerns that will inform a final draft to go to the EB. While we won't

have additional input into the draft, Corinna and Wade will have additional opportunity to re-address things not dealt with in the first round.

Margaret: Stuff that came up yesterday. Put CI technical stuff in a separate document.

Ken: There's a hazard in separating CI. They need to be tightly integrated.

Don: The idea was to leave in operational step and the thinking was that there is not enough information for a technical document. More technical specs could be added, but you would still have CI components at the level of detail that everything else is in the plan.

Wade: At EBs request, the general LNO operations were included for EB. This will be pulled back out for Final plan. So this is compounded by the ARRA. This is why details are scater on Operational Agreement because they are covered elsewhere.

Hap: Missed this comment. Sorry.

Wade: There's a lot detail lacking to bring databases inter-operability. This is our big opportunity to move forward on things that can't be funded using grants because they are operations. These missing details will be addressed in our meeting tomorrow.

Don: Will you be addressing governance?

Wade: we have a lot of questions about governance ourselves. We are going to be looking into NIS and CI input without solving broader problems of governance.

Don: There are places where IMExec could be doing more or be better integrated.

Wade: Exactly. Especially in the area of training and where site IMs need and what they can contribute to overall process. We need to use IMexec to make sure site im's needs are being met and I've already got that down as something to discuss.

Hap: One last thing I haven't sent comments in. Where do comments go?

Corinna: Send to all three.

Meeting adjourned around 12:30.

- Virtual Updates [1]

- Copyright © 2012 Long Term Ecological Research Network, Albuquerque, NM -  
This material is based upon work supported by the National Science Foundation under Cooperative Agreement #DEB-0236154. Any opinions, findings, conclusions, or recommendations expressed in the material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.  
Please contact us with questions, comments, or for technical assistance regarding this web site.



**Source URL:** <http://im.lternet.edu/node/548>

**Links:**

[1] <http://im.lternet.edu/taxonomy/term/169>