As-yet-unnamed LTER Network Data deCenter

LTER Information Management Committee

DRAFT 2015-07-07

The LTER Network Information Management Organization (NIMO) envisions a scientific community in which information management contributes to long-term data stewardship through development and implementation of data and design practices that support scientific research across LTER sites and among LTER and other environmental observing programs. To enable new scientific discoveries in ecology by ensuring access to high-quality data, NIMO will adopt, design and implement a sustainable yet nimble and responsive infrastructure that supports long-term data curation through the development of decentralized information systems, protocols and tools for managing ecological information, cyberinfrastructure and workforce training.

Introduction

The LTER Information Management Committee (IMC) envisions a flexible organization for LTER Network-level data management that can respond to community-identified needs with transparent technology, and which facilitates usability and adoption of appropriate tools by data management personnel at LTER sites. In this model, intellectual and fiscal management are separated, and decisions are made and negotiations are carried out by a group of network representatives, rather than by a single Principal Investigator. Divisions between decision-making (Governance Committee), financial responsibilities (Fiscal Entity), and execution (Project Manager) are designed to maximize transparency, optimize return from expenditures, streamline the process of information management, and provide the best support for science that can be achieved within budget constraints.

The planned services encompass major categories, and within each, prioritization will be carried out by the IMC according to Network requirements and in preparation for budgeting. Of highest priority are core technical services essential for continued data availability and for Network identity. Ancillary targeted databases and scientific working group support, along with assistance to individual sites for their local needs, and activities that foster connections between the LTER and the broader community of environmental sciences will be prioritized on a case by case basis (‘community requirements’ or ‘ideas bucket’ in Figure 2). A transition period is planned, during which highest priority services will be maintained while the new structure for decision making is put in place.

The [NIMO] described here does not replace site-based information management; rather it relies upon the distributed talent within LTER IMC while offering central services and support. NIMO will coordinate with the new LTER Communications Office in regards to responsibilities for maintaining certain network databases and data management-related support for science synthesis and working groups. Furthermore, detailed contracts need to be negotiated with NSF

and with the Fiscal Entity, including a chain of responsibility and reporting that satisfies NSF requirements and LTER data management needs.

Overall organization

A fully functional data management system requires expertise in many areas: project management, system and database administration, website design and implementation, content strategy, programming, technical writing, and data curation and stewardship, and others. Experience has taught us that the breadth and depth required to adequately carry out all tasks is rare in one person, or even a small group of individuals. Therefore, we envision a system where the strengths of the diverse group of LTER information managers plus a few targeted hires can be tapped and compensated.

Although there are centralized (or singular) elements, the overall model proposed is decentralized with services provided by various vendors, locations, or LTER sites (Fig 1). A Governance Board with membership composed of LTER information managers and scientists takes input from the IMC and LTER scientific community. A single Project Manager supports and implements its decisions. Fiscal responsibility is carried out by a single entity, acting mainly as an escrow organization without a technological mission. This organizational strategy will help ensure that no one person's (or institution's) technological agenda is prerequisite. Data management tasks will be accomplished by a variety of individuals. Some are already at sites (e.g., site information managers or other site staff), and the choice to retain dedicated people working in academic environments recognizes the contributions of the LTER data management community. However that choice may limit the available skill sets, and so targeted hires (project specific and possibly short term) from outside the current pool is also likely. Multiple adequate solutions already exist for distribution of computational hardware. Specific requirements, tasks and responsibilities for each entity are below. Figure 2 illustrates the management process and interactions between these entities.

1. Governance Board

Requirements

- Rotating chair who is not the project PI, but is an IM (term length and rotation TBD)
- Rotating membership (staggered terms, length TBD)
- Membership: 7? site information managers (one serves as chair), two scientists (one an EB rep), fiscal entity rep (ex-officio, non-voting)
- Service would be compensated (suggest one month of compensation for members and two months for chair? TBD? exclusive of ex-officio members)
- Activities must be transparent with good outbound communication and full IMC involvement in critical decisions.

Tasks/Responsibilities

- Develop by-laws and procedures based on current IMC Terms of Reference (LTER IMC, 2011)

    - decisions by consensus with occasional votes (see IMC ToR sample language)
    - develop procedures to make decisions transparent and ensure input from all IMC members
- Establish System's overall direction and goals
    - community served (e.g., LTER only, single investigators outside LTER, etc)
    - services to be provided, and implications for technological framework
- Manage overall finances
- Agree on a decision making framework based on IMC and advisory input to set priorities for
    - IT infrastructure
    - Software to be supported
    - Architecture (web services, databases, etc.)
    - Workflows, approaches
- Receive proposals for specific projects, and set priorities based on decision making framework above. Examples include, but are not limited to:
    - A site IM develops a certain software package which may have more widespread use. [NIMO] provides programmer time to
        - bring to production stage,
        - maintain code and/or upgrade,
        - provide user training and support
    - A site needs access to commercial software. [NIMO]
        - buys license
        - provides on central server
    - A scientist asks for code or a data product to support a synthesis project
- Communicate with the IMC, EB, LTER CO, NSF, and Fiscal entity

2. Fiscal Entity

Requirements

- Legally able to engage in a cooperative agreement with NSF
- Independent from LTER sites
- No direct technology mission which could supersede decisions of governance board
- Facilitates provision of services and support structure
- Reduced overhead

Tasks/Responsibilities

- Enter into a cooperative agreement with NSF
- Hire personnel as directed by the Governance Committee
- Reimburse LTER IMs for time spent on network projects
- Administer subcontracts with universities
- Administer funds for IT infrastructure and services as directed by the Governance Committee
- Administer travel funds?

3. Project Manager

Requirements

- Technology experience - estimate costs and technical issues (IT Project Management)
- Personnel management experience - hiring, evaluation
- Fiscal experience - formulate and manage budgets
- Writing skills - ability to write compelling proposals and accurate reports
- Verbal communication
    - ability to get responses from project participants
    - with IMC and Governance Board
- Scientific literacy
    - demonstrated ability to work with scientists
    - able to work with LTER culture
- Possibly someone who has been a site IM or worked in a research data management environment

Tasks/Responsibilities

- Make sure the bills get paid (interface with fiscal entity)
    - administer sub-contracts
    - technology support (computing hardware, licenses, etc.)
- Manage other personnel (e.g., programmers, system admins, coordinators, administrators)
    - hiring, evaluation, promotion, termination
- Implement and track progress on the tasks identified by the Governance Board
- Provide periodic updates on status of each project
- Take actions to assure that projects are proceeding as desired and within scope
- Manage logistics for IMC working groups and specialized meetings
- Organize or facilitate training activities
- Adhere to data management best practices

4. LTER at large

Role/Responsibilities

- Provide science drivers for data management goals
- Facilitate collaboration between synthesis projects and IM
- Advise Governance Board

Scope of Services

Currently five major categories of service are identified. Of highest priority are core technical services essential for continued data availability (Table 1, “Core Technical”). Over its lifetime, the Network has adopted several ancillary databases targeted for specific activities, types of data or audiences. The needs for hosting and maintenance of these varies, and will be determined on a case-by-case basis and prioritized within the spectrum of all services (Table 1,

“Ancillary and Targeted Databases”). Data management support for scientific synthesis projects is envisioned to be coordination with the LTER Communications Office, and specific areas of responsibility and mechanisms are yet to be determined (Table 1, “Scientific”).

The framework is flexible enough to accommodate planning for other types of services that may require specific additional funding. To promote a homogeneous network approach to site-level data management, sites have received occasional technical support and technology, which could be continued by [NIMO]. Contracting with highly-skilled personnel to enact network approaches at sites could enhance data management practices locally, and be supported under the same priority conditions as the databases and science support. (Table 1, “Site”). Collaborations and connections between the LTER and the broader community of environmental sciences are possible via sophisticated technological approaches already being developed within broad thematic areas, such as by EarthCube or S2I2 (e.g., Chandler, et al. 2015). Adoption of these products or collaboration with similar activities would expose LTER data broadly, making it accessible far beyond our own Network. And lastly, a broadening of the services to single investigators, other groups (e.g., OBFS) on a fee-basis is envisioned in future years.(Table 1, “Community Presence”).

Implementation of the services described below may be accomplished using a variety of mechanisms including short-term “code fests,” projects led by LTER site information managers, full or part-time [NIMO] employees, contracts with external organizations or even (following the astronautics community) prizes for the best suitable system (Porter et al., in prep). It is anticipated that the choice of mechanism to accomplish a specific task will vary. The goal will be to obtain the best result while making maximum use of the creative potential within the LTER Network and beyond, and to build sustainable systems that can provide many years of service. Table 2 shows an anticipated timeline for broad classes of activities during the first few years of the project.

Table 1. Example services. Priorities are still to be determined, and expected to be partly based on relative costs.

Type	Description	Services
Core Technical	Systems essential for data availability and Network identity	● PASTA ● Network data catalog ● DataONE member node ● improvements to core services
Ancillary and Targeted Databases	Services to the entire network, targeted for specific activities (e.g. information management, scientific). Some coordination	● EML support ○ Unit Registry ○ Controlled Vocabulary ○ new databases (e.g. taxonomy)

	with the LTER Communications Office is expected for some databases.	● personnelDB (LTER CO) ● bibliography (LTER CO) ● Legacy data (e.g., discontinued sites, LandSat imagery)
Scientific	Support for working groups per scientific priorities. Some services are likely to require close coordination with LTER Communications Office.	● existing data products ○ climHydroDB ○ GeoNIS ○ EcoTrends ● new activities ○ Assistance finding data ○ Create EML-based data products for Network data catalog ○ Curate analysis code
Site	Assistance to individual sites for their local needs	● Tech support ● training / software adoption assistance ● New IM personnel mentoring ● Maintain/develop DEIMS code ● Host DEIMS for individual sites ● Metabase assistance ● Matlab Data Toolbox assistance
Community Presence	Foster connections between LTER data systems and other communities	● Cross-referencing between LTER collections and domain repositories (eg, ACADIS, BCO-DMO, streamChemDB) ● Collaboration with external groups ● Data management services to single investigators and other groups on a per fee basis (e.g., OBFS)

Table 2. Anticipated Timeline (technology transition in Years One and Two)

Period	**Goals**
Year Zero	● Define, review, and ratify governance and management Terms of Reference for GB and IMC ● Identify fiscal entity ● Negotiate Cooperative Agreement ● Agree on reporting commitments

	● Create MOU (or contract) with Fiscal Entity ● Create MOU with LTER Communication Office
Year One	Operations: ● Initiate Governance Board ● Hire project manager ● Transfer technical personnel, as appropriate, to fiscal entity Services initiation: ● Core: Transition to new platforms or external hosting ● Targeted databases: Develop technical requirements and priorities ● Scientific: Develop planning process, participate in one working group as a prototype ● Site: Characterize site needs and priorities ● Community presence: Identify groups of interest
Year Two	Operations: ● Governance Board turnover (__ seats) Services progress: ● Core: Review, ratify technology road maps and support model ● All: initiate additional services as determined by priorities
Year Three	Operations: ● Governance Board turnover (__ seats) ● mid-term review Services progress: ● All: Review community proposal and set priorities ● Core, Targeted, Scientific: identify additional necessary features ● Site, Community presence: Seek additional funding
Year Four	Operations: ● Governance Board turnover (__ seats) Services progress: ● As in Year Three

Literature cited

Chandler, C., A. Shepherd., P. Wiebe, D. Kinkade. 2015. BCO-DMO and GeoLink: Improving Data Discovery and Access for Ocean Science Research Results. Presented at EarthCube All Hands Meeting, Washington, DC, 27-29 May 2015. http://earthcube.org/document/2015/bco-dmo-geolink

LTER IMC. 2011. Terms of Reference. http://im.lternet.edu/home/tor

Porter, J. H., M. Gastil-Buhl, K. Vanderbilt. in prep. Draft: Types of Projects & Products https://docs.google.com/document/d/1o3aqtxxHVcqJeQqRIZ6TWJ6OqLf3piHPNAmhoVU_SA8/edit?usp=sharing

Possible Organization Names

NIMO - Network Information Management Organization (conflicts with National Incident Management Organization, National Internet Math Olympiad and a few other organizations, 9M hits on google). Or
NEIMO - Network Ecological Information Management Organization (conflicts with French Rock band, 83K hits).
NEMO - Network Ecoinformatics Management Organization (conflicts with Disney Movie, 83M hits).
L-NIMO - LTER Network Information Management Organization (14K hits). climate change refs possible!
NLTIMO - Network Long-Term Information Management Organization (37K hits),
LTIMO Long-Term Information Management Organization (700K hits)
LIMO (then we'd have Limo services)
LIM (or LIMN, the N is silent, conflicts with Lab Information System)
L-RIMO - LTER Research Information Management Organization
LIMP - LTER info mgt Partnership
LINC - LTER info mgt Consortium

Figures

Figure __. High Level Activity Management Process

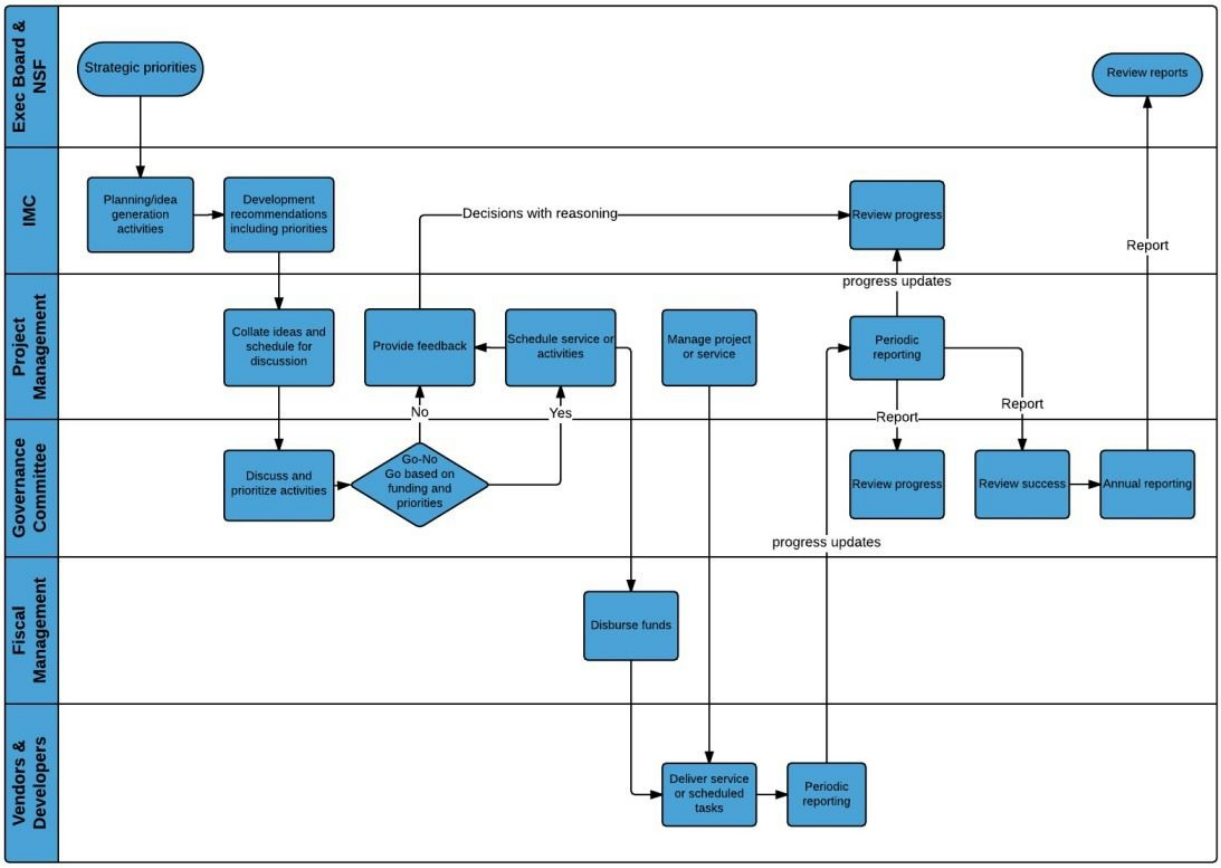

Figure __. Organizational Relationships

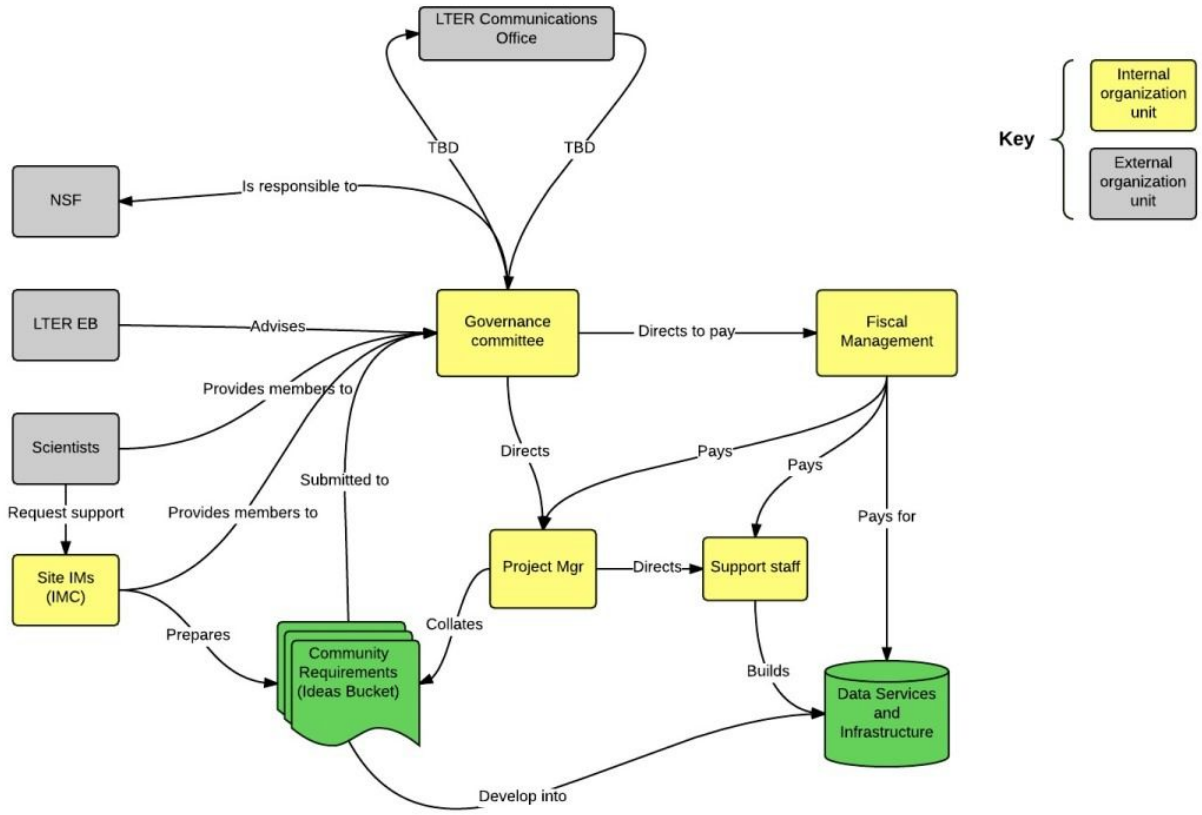

Figure __. Decentralized Services and Structure