Network Office

Operational Plan (2009 – 2014)

25 February 2010

LTER Network Office
Department of Biology, MSC03 2020
1 University of New Mexico
Albuquerque, NM 87131-0001

TABLE OF CONTENTS

Operational Plan for Activities of the LTER Network Office

1. Introduction and goals

The Long Term Ecological Research (LTER) Network Office (LNO) is responsible for planning, implementing, and supporting activities that advance the goals of the LTER Network. With funds from the National Science Foundation (NSF), the LNO carries out these activities as defined in Cooperative Agreements (CA) with NSF with additional guidance from the Executive Board (EB) of the LTER Network. Specifics of each activity are described in proposals to the NSF that are examined by external reviewers and evaluated by a panel of experts selected by NSF.

In 2008, the LNO submitted a renewal proposal to NSF to continue core operations and add new tasks in synthesis and cyberinfrastructure development, based on the goals of the LTER Decadal Plan (http://www.lternet.edu/decadalplan/). This renewal proposal described 17 activities organized around four thematic areas: Synthesis, Cyberinfrastructure, Core Services, and Development and Outreach. Selected activities described under the Cyberinfrastructure theme are designed to support the research activities under Synthesis through the development of a network information system. The partnerships formed and the training to take place under the Development and Outreach theme inform the development and support the use of the network information system. The Core Services theme describes the underlying role of the LNO in supporting all LTER activities. The NSF funded the renewal proposal in 2009 through two separate Cooperative Agreements.

One of these Cooperative Agreements with the NSF incorporates funds from the American Recovery and Reinvestment Act (ARRA). As part of cost-accounting measures associated with awards using ARRA funds, the NSF requires specific details on the planned expenditure of funds. In the case of the LNO Cooperative Agreement, the NSF requested the development of a detailed implementation plan describing the ten activities to be carried out with ARRA funding and the costs associated with those activities. The NSF further requested that the LTER Executive Board be closely engaged in the development of this operational plan by the LNO and that external reviewers inform the development of that plan.
The LNO has undertaken to develop an operational plan that provides detailed information on each activity funded by ARRA. With guidance from NSF and the Executive Board, the LNO has prepared standard descriptions of each activity, linking these with effort (Table 1) and cost (Table 2) budgets to provide clear links between activities, personnel, and expenditures. This operational plan has been reviewed by the Executive Board, relevant LTER committees, and a panel of outside experts chosen for their knowledge of the technical aspects of cyberinfrastructure and information management. The final operational plan will form the basis of future evaluations of the performance of the LNO with regard to work supported by ARRA funds.

2. Steps in the development of the operational plan

The terms and conditions of the LNO award to the University of New Mexico (UNM) specify the preparation of an operational plan and describe in general detail the elements of such a plan.

Based on these instructions, and with additional guidance from NSF and the Executive Board, the LNO developed a process to draft an operational plan to fulfill the requirements set forth by NSF. This process consists of the following steps.

A. The LNO prepared a general outline of the operational plan with key elements for review and comment by the Executive Board.
B. Once approved by the Executive Board, the outline provided a framework for further discussions of groups of activities clustered into four thematic areas (Synthesis, Cyberinfrastructure, Core Services, and Development/Outreach. Sub-committees consisting of members of the Executive Board were assigned responsibility for detailed review for each of the thematic areas.
C. The sub-committees met in video-teleconference to review LNO activities in each thematic area. The LNO Executive Director presented the details of each activity to the respective sub-committees, who provided comment on the appropriateness of each activity and the allocation of resources.
D. Once the Executive Board sub-committees approved activities proposed in each thematic area, LNO staff undertook to prepare standard descriptions of each activity that included information required by the NSF and the Executive Board.
E. When necessary, LNO staff consulted with the relevant sub-committees to resolve questions on content and presentation of the descriptions of each activity.
F. The LNO proposed members for the external advisory group to the Executive Board, which had final say on the composition of the advisory group.
G. A first draft of the operational plan was prepared by LNO staff under the supervision of the Executive Director. The Executive Board, the LTER Information Management Committee, and the Network Information System Advisory Committee (NISAC) reviewed the draft plan.
H. Based on comments from this review, LNO staff prepared a second draft of the operational plan for review by the external advisory group, the Executive Board, the Information Management Committee, and the Network Information System Advisory Committee.
I. The external advisory group, the Executive Board, and the co-chairs of the Network Information System Advisory Committee met with LNO staff for three days in January 2010 to deliver detailed comments on the second draft of the operational plan.
J. Based on these comments, LNO staff prepared a third draft of the operational plan for review by the Executive Board.
K. After receiving comments from the Executive Board, the LNO prepared the final draft of the operational plan, which was submitted to the NSF along with the comments of the external advisory group.

3. The LNO Operational Plan

The operational plan contains the following information for each of 10 activities in four thematic areas: 1) description of activity, 2) outcomes of activity, 3) operational steps to carry out activity, 4) milestones, 5) process for reviewing progress and evaluating success, 6) benefits to the LTER community, its external partners, and the larger U.S. scientific and education community, and 7) mechanisms to garner feedback from the LTER community and external communities of science

and education. The descriptions of each activity are designed to provide sufficient information to understand purpose, outcomes, and impacts.

3.1 Synthesis

3.1.1 Science Council. Support annual meetings of the Science Council as well as planning meetings in support of its activities.

Background: As part of the new governance structure for the LTER Network, the LTER Bylaws have been amended to create a new entity, the Science Council (SC), comprised of two members from each of the 26 LTER sites and the chairs of standing committees. The SC "has the responsibility to provide leadership and planning for cross-site research and education, to develop proposals for the conduct of Network-level science, to interact with existing and emerging networks, to develop products that synthesize Network-level data and information, and to otherwise manage the science affairs of the LTER Network". The Bylaws require an annual meeting of the SC, and the LNO is responsible for supporting this meeting. The intended long-term impact of SC activities is an increase in operational coordination of science, synthesis, and education across the LTER Network.

Description of Activities: The LNO funds participant travel and lodging for the annual meeting of the SC as well as any additional required expenses (e. g., meeting rooms, audiovisual). The LNO also supports planning and visioning meetings associated with SC activities. The budget associated with the SC is $271048 for five years to cover participant travel of the SC and 0.10 FTE of personnel time to support the meeting. The Executive Director (ED) and the Director for Synthesis Support (DSS) help develop the agenda, supervise meeting preparations and logistics, prepare materials and online databases, participate in the meetings, and record the accomplishments of the meetings.

Outcomes: Outcomes of the operations of the SC include enhanced Network-level synthesis, increased scientific coordination across and beyond the LTER Network, increased funding for synthesis, evolving strategic plans, and effective governance. Synthesis activities of the SC produce Network-level science publications and synthesis volumes, lay the groundwork for proposals to conduct Network-level science and education, and result in advances to the Network research and education objectives. In addition, the SC performs several administrative functions, including resolution of issues from sites or the Executive Board, approval of Bylaw changes, and election of the Chair. The LNO also supports the SC by producing a persistent record of its activities and decisions.

Operational steps:

1. The ED works with the SC to determine a rotating schedule of host sites.
2. Twelve months prior to meeting, the DSS contacts the host LTER site to prepare overall logistics and venue.
3. The DSS works with EB in determining specific science content for the meeting.
4. The DSS publicizes meeting plans to LTER scientists and establishes a web presence for the meeting.
5. The ED and DSS coordinate efforts of LNO and host site staff for logistic support and overall meeting planning.

6. The ED ensures that all aspects of the meeting related to logistics, including participant travel, lodging, food, and meeting rooms are provided.
7. The ED and the DSS attend the SC meeting and provide logistical support in support of science efforts of attendees.
8. The DSS records and preserves all meeting activities and science content such as presentations, working group products and other materials as a permanent record of the meeting.
9. The DSS creates, administers, and analyzes a simple meeting survey to obtain feedback from the meeting from participants.
10. The DSS makes results of the meeting available to LTER scientists.

Milestones:

- The Chair, the ED, and the host site set meeting dates approximately 12 months before the meeting.
- The ED contacts the host site to select venues approximately 9 months before the meeting.
- The ED and DSS work with LNO staff on a monthly basis to coordinate overall meeting logistics including travel, housing and food for attendees.
- The ED and the DSS attend meeting for coordination, and to provide onsite support for science presentations and working group meetings.
- The DSS records activities and archives records of meeting including results of working group meetings and other products resulting from the meeting within 30 days of the meeting.
- The DSS collects and publishes feedback from participants through comments and survey results within 30 days of the meeting.
- The DSS receives, documents, distributes, and archives subsequent products for on-going and new Network level science efforts.

Evaluation: The EB will annually evaluate the LNO's support of SC science efforts through direct experience and feedback from participants. The overall success of the meeting and science products from the meeting will be evaluated by the SC.

Benefits: SC meetings will increase the pace and scope of synthesis in the LTER Network and contribute to attaining the goals of the Decadal Plan. LTER scientists and educators benefit from coordination of scientific endeavors, strategic planning, and governance activities conducted at SC meetings. Increased funding for synthesis resulting from proposals developed by the SC benefits LTER scientists and educators as well as their external partners. The larger scientific and education community benefit directly from science products (e.g., new theory, insights, knowledge) resulting from SC meetings as well as from new scientific initiatives arising from these meetings.

Feedback: Feedback regarding support of Science Council meetings comes from comments from participants of the meeting and site principal investigators as well as survey results.

3.1.2 Research Working Groups. Provide support for Network research goals by funding working groups and intensive research visits for project scientists.

3.1.3 Activities of the Decadal Research Plan. Provide support for Network research goals by facilitating planning and visioning meetings to address the objectives of the Decadal Plan.

Background: The activities that produced the Decadal Plan created new research collaborations among LTER scientists and with scientists from other disciplines, especially the social sciences. These new collaborations need additional resources to encourage their further self-organization. Following guidance from the EB, the LNO provides and coordinates funding to individuals and research groups for this purpose. The working groups that receive LNO support are expected to produce specific products directed at new Network-level research activities. The LNO ensures that proposed synthesis activities are carried out and preserves and documents products from research working groups.

The Decadal Plan also focuses on long-term, social-ecological questions in specific thematic areas that will be targeted for follow-on ASM activities. The Decadal Plan presents the broad outlines of a research program in each thematic area, but additional specifics need to be developed. Planning workshops that ask participants to identify individual questions, observations, experiments, and modeling activities within thematic areas and to specify corresponding education and cyberinfrastructure needs can help develop the details required to advance the Decadal Plan, and the Executive Board has requested the LNO to support these workshops.

In accordance with the guidance of the EB, we have combined research working group opportunities and planning/visioning opportunities into one activity. The total budget for this combined activity is $1,470,783for five years to support participation in research and planning meetings as well as 3.30 FTE of personnel time annually to manage the disbursement of funds, coordinate meeting logistics, and maintain a database of meeting products.

Description of Activity: Following recommendations of the EB, the LNO supports self-organizing groups of site scientists by funding working groups to produce specific research products such as publications, experimental designs, new databases, or common standards. The LNO also provides support for Decadal Plan visioning and planning activities resulting from Science Council activities related to specific focus areas of the Decadal Plan. The LNO may also provide support to allow LTER researchers to focus on short-term synthesis projects that support the Decadal Plan research objectives. These could be short-period support at a specific institution, LTER site or at the LNO to produce a specific product related to Decadal Plan research efforts. The LNO may also provide funds to support a Site Information Manager in selected working groups to communicate research requirements to Network Information System developers.

The Executive Director and the Director for Synthesis Support ensure logistical support in coordination with other LNO staff. This includes soliciting, and collecting proposals to establish these efforts and coordinating their evaluation with the EB. LNO also organizes and maintains records of annual competitions to disburse funds related to these efforts, with proposals evaluated by the EB or their delegated representatives. The DSS, with support from LNO staff, works closely with all funded research working groups to record, preserve, and publicize all products resulting from the research working group efforts.

The Chief Information Officer (CIO) and the Network Information Manager (NIM) will work closely with research working groups to facilitate discovery and acquisition of data through the Network Information System (NIS). They will also provide working groups with guidelines and best practices established by the Information Management Committee (IMC) to manage derived data products. In some cases, the LNO may provide funding for a Site Information Manager (SIM) to carry out this function. Where applicable, the process for creating derived data products will be captured and used to inform use case scenarios that will inform the development of the Network Information System (NIS).

Outcomes: Research working groups establish new partnerships and collaborations that advance a new body of theory in social ecology as one element of a developing strategy to meet Decadal Plan objectives over a range of funding scenarios. The working groups enhance the breadth of LTER research and increase the pace of synthesis by producing multidisciplinary, multi-site publications such as databases, models, books, monographs, and special issues of journals.

The proposed planning groups will produce detailed steps for achieving the science goals of the Decadal Plan; a timeline for the staged implementation of those steps; and interim objectives for LTER sites and scientists. The proposed workshops will lead to a better balance between support for site-based research and Network-level collaborative synthesis. The eventual outcome will be achievement of the goals of the Decadal Plan.

Operational steps:

1. The LNO solicits proposals annually for research working groups and follow-on collaborative discussions held during ASMs or other interactions.
2. Proposals are collected and organized for review by the EB.
3. Logistics for successful proposals are coordinated by the DSS with LNO staff, including on-site support and collaboration as required by each of the activities.
4. The DSS facilitates identification of pertinent data sets from the NIS.
5. The CIO and NIM (or a SIM) provide guidance on best practices for creating and managing derived data products as requested.
6. The CIO and NIM (or a SIM) collect information on the creation of the derived data products and transmit to the NIS developers.
7. Products of working groups are collected, documented and publicized.
8. Solicitations for support of Network-level planning and visioning projects are made following directions from the EB
9. Logistics and science support of planning and visioning projects are coordinated by the DSS with support from LNO staff

Milestones:

- The DSS prepares and distributes annual solicitation for research working groups.
- The DSS arranges for the review of proposed working groups with the EB
- The DSS provides any required support for funded planning and working group efforts.
- The DSS acquires, archives, and distributes research products from working groups and planning meetings as they become available.

Evaluation: The EB evaluates performance of the LNO in supporting working groups as part of its annual evaluation.

Benefits: LTER scientists and educators and their external partners are the primary beneficiaries of working groups and planning meetings. Working groups provide opportunities for the LTER community to mature and advance collective ideas and to create products such as publications, reports, proposals, and databases. Planning meetings allow LTER scientists and their partners to construct new initiatives and partnerships that advance the common good. Support, documentation, and preservation of products from research working groups and visioning meetings advance Network-level research at a pace beyond what would be achieved within the previous framework of activities of the LTER Network.

Feedback: Feedback to advance these activities comes from individual working groups, direct comments by LTER scientists, biennial surveys of the LTER community, and review from the EB.

3.2 Cyberinfrastructure

Background: Developments in cyberinfrastructure (CI) that promote connectivity allow the distributed members of the LTER network to act and feel as part of a close community. Transparent email addresses, a common web presence for public access to LTER science, video teleconferencing, an intranet for building the LTER knowledge base, a shared data catalog that provides metadata for all LTER data sets, systems for the hosting and curation of databases, special-interest websites, and more help to create a sense of community that fosters intersite and network-wide synthesis. The LNO supports Network CI by offering services that both meet the needs of the community (e.g., specialized maillists) and keep pace with evolving technology (e.g., virtual machine hosting). The CI activities described below are supportive of and in harmony with the implementation of the LTER Network CI Strategic Plan, and will provide critical services to the LTER community through four activities:

- improving the operation of the LNO computing, communication, and collaboration infrastructure;
- supporting the development, communication, implementation, and standardization of information management procedures, policies, and practices;
- completing the development and implementation of the LTER Network Information System framework; and
- offering on-site and web-based group and individualized consulting services for improving and maturing site information management practices.

3.2.1 Basic Cyberinfrastructure Support. Provide basic cyberinfrastructure support to the LTER network to enable collaboration, communication, and security.

Description of Activities: The CIO, System Administrator, and Systems Analyst maintain and operate the hardware and software that provide required computing and communication services for LNO and the LTER Network. These include support and management of application servers, support for local and network file services, support for the software development process, and

support for training. The supported hardware and software includes five multiprocessor production servers for network mail lists and local mail handling, database management systems, a web content server, the Metacat data catalog, and a virtual machine server hosting 8 lightweight, fail-over, and development machines. Two file servers support administrative applications including home directories and archives, backups, and shared-access storage. LNO staff daily conduct system health checks, security screening, backups, routine maintenance, repair and replacement of equipment, update verification, system upgrades, and interactions with sites on hosting and email issues. In addition, the CIO, Systems Administrator, and Systems Analyst plan for, implement, and operate collaboration and communication technologies to support video teleconferences. Finally, they administer desktop and laptop computers for all LNO staff and for the LNO training laboratory.

The budget for this activity is $847,054 for five years. In addition to 2.10 FTE of personnel time to conduct the activities described above, additional resources are budgeted for costs of server, desktop computer, and peripheral replacement.

3.2.1.1 Operation of video-teleconferencing services for LTER committees and working groups. LNO supports high-quality, low-cost, standards-based video-teleconferencing (VTC) for the Network, including regular monthly VTCs for the LTER EB and Information Management Executive Committee (IMExec). This service is made possible by a Polycom teleconferencing bridge and gatekeeper software that is owned and operated by the LNO. The bridge supports up to 48 simultaneous connections.

Outcomes: Video-teleconferencing services increase the level of interaction, by extending access to internet-based video-teleconferencing for synthetic interactions to more LTER scientists. An eventual goal is to find and implement technology that will permit virtual meetings of representatives of the 26 LTER sites, thus increasing the frequency of Network interactions.

Operational Steps:

1. The CIO ensures the communication of accessibility of VTC to LTER members on an annual basis and maintains a web-based document detailing VTC options, pricing, protocol, etiquette, and scheduling information.
2. The CIO maintains communication with UNM Media Technology Services (MTS), where the bridge and server are physically housed, to schedule maintenance and upgrades, and to assure a good working relationship with MTS.
3. The CIO meets with the Systems Analyst twice weekly to discuss the VTC schedule and any scheduling or technology issues.
4. As needed, the Systems Analyst verifies site/researcher access to VTC technologies and facilitates the delivery of hardware and software to new members of specific groups (e.g., EB) along with installation and configuration support.
5. Groups wishing to use the VTC conferencing capability poll their members and contact the Systems Analyst to verify availability of the bridge space and provide a list of conference participants.
6. The Systems Analyst verifies the details of each participant,

7. The Systems Analyst confirms participation of each person and connection accuracy for each participant via phone or email.
8. The Systems Analyst programs reservations for meetings into the bridge management software so that participants are automatically contacted at the correct time.
9. The Systems Analyst establishes the call and monitors the call for quality and connection interruptions.
10. When requested, the Systems Analyst produces and archives a digital recording of the call.

Milestones: Each scheduled VTC becomes a milestone, and the successful completion of each VTC is recorded with pertinent information about purpose and participants.

Evaluation: A summary of annual effort, including number and kind of VTCs supported and the number of participants, is included as part of the LNO annual report to EB and NSF, which informs the annual performance evaluation of LNO.

Benefits: This activity benefits the LTER Network and its partners by increasing the quality and quantity of productive interactions among LTER scientists and between the LTER Network and other networks, centers, institutes, and agencies.

Feedback: Feedback on this service will be based on questions included in the biennial survey of LNO performance conducted by the EB.

3.2.1.2 Web presence improvements for collaboration and presentation. The LNO maintains web sites targeted at constituencies including the general scientific community, the LTER Network, and specific focal groups. To increase the quality of and access to web-based information content, the Application Support Analyst (ASA) will augment the current web applications to support Network committees, working groups, and special interests with basic web publishing and collaboration tools. To provide sufficient time from regular duties for the Application Support Analyst to evaluate and implement new technologies that meet user needs, the LNO will recruit a new Professional Intern to conduct maintenance activities on the LNO web sites. The professional intern will, in addition to other duties, review content for out of date information, perform regular and time critical updates to web content, conduct tests of web site operation, and respond to user requests for assistance. In addition, the professional intern will make changes to web content as directed by the LTER Network communication plan when it is completed.

Outcomes: This activity will increase the quality of users' web experience and the delivery of web-based content, and provide a better response to user requests.

Operational Steps:

1. The CIO and ASA will annually recruit a web-savvy intern or will evaluate the existing intern for renewal.
2. The ASA will request targeted feedback from committees on their web-based collaboration needs.

3. The CIO, ED, and ASA will discuss and plan annual priorities for web improvements informed by feedback from biennial surveys and unsolicited recommendations from users and the impending communication plan.
4. The CIO and ASA will develop a review schedule for the professional intern to follow providing for content evaluation, link checking, date checking, presentation consistency, regular reviews of committees membership, chairs, etc.

Milestones:

- Annual priorities for web improvements set in January.
- Quarterly reviews of progress for interns.
- Annual web priorities achieved.

Evaluation: Using information from annual reports, outcomes from biennial surveys of users, and performance guidelines established annually, the EB will recommend changes to procedures and priorities.

Benefits: This activity will benefit all visitors to LTER network web sites including LTER scientists, external partners, the larger U.S. scientific and education community, and the general public by providing improved information content and accessibility.

Feedback: Feedback on this service will be based on questions included in the biennial survey of LTER Network Scientists. Feedback from external partners and the larger scientific and education community will be solicited through a feedback form placed on all websites to encourage comments and recommendations.

3.2.1.3 Technology improvements. This activity provides for needed replacement, upgrade, and improvement of LNO computing infrastructure to support network information services, local desktop computing, and training. The LNO CIO and System Administrator will upgrade server hardware and peripherals twice during the CA period to include storage consolidation and a greater reliance on virtual machine systems for non-intensive and fail-over applications. The Systems Analyst will upgrade hardware and peripherals in the Informatics Training Laboratory (see Section 3.4.2 below) twice during the CA period. The Systems Analyst will also perform annual updates of desktop computing equipment on a phased schedule that provides replacement or rotation of CPU units roughly every three years.

Outcomes: LTER Network activities serving science and education are conducted in a secure, modern, and efficient computing environment with adequate resources for growth and development.

Operational steps:

1. The CIO, Systems Administrator, and Systems Analyst will develop a replacement/upgrade schedule for servers, desktop computers, and training lab computers.
2. The CIO will discuss configuration and quotes with vendor.

3. The CIO will present a plan during year 2 to the ED for approval and hand-off to administrative staff for procurement.
4. The Systems Analyst will receive and check-in new equipment upon arrival.
5. The Systems Analyst and Systems Administrator will schedule equipment installation, stage the equipment, and do basic system checks prior to software installation and configuration.
6. The CIO, Systems Administrator, and ASA will discuss any configuration issues and schedule the bringing on line of new equipment and services.
7. The CIO will revise the schedule as needed based on requirements and technology changes. Steps 2-6 will be repeated for each year's procurement.

Milestones:

- By the end of 2010, develop a technology enhancement/replacement schedule.
- By the end of each calendar year, deprecated desktop computing equipment has been replaced, refurbished, or recycled and appropriate adjustments made to inventory according to the schedule.
- In year 2011, new virtual machine hardware and consolidated storage are purchased and configured.
- In year 2012, existing servers are retooled to use consolidated storage system.
- In year 2014, new hardware to replace aging production servers purchased according to revised schedule.

Evaluation: Equipment is placed into use/production according to schedule.

Benefits: This activity will benefit the LTER Network, its external partners, and the external science and education community by allowing uninterrupted computing and web-based services in an efficient and cost-effective architecture.

Feedback: Service performance will be evaluated at monthly meetings of LNO staff and through biennial surveys of the LTER community. Comments and recommendations from external communities will be solicited through a feedback form placed on all websites to encourage comments and recommendations.

3.2.1.4 Communication. The CIO and Systems Analyst will acquire, deploy, and maintain "rich media" recorders to capture and stream presentations via the Internet. Through an investment by UNM in this technology, the LNO will be able to leverage an additional recorder for loan to working groups and committees as needed. The annual LTER mini-symposium at NSF, critical science council presentations, and the triennial All Scientist Meeting plenaries are but a few examples of events that should be preserved through recordings. In addition, LNO will use this technology to capture and stream training and security sessions.

Outcomes: The new capacity to record and stream LTER presentations will improve outreach to the LTER and external scientific, education, and information management communities.

Operational Steps:

1. The CIO and Systems Analyst will purchase and set up rich media recorders for operation in coordination with the vendor and UNM MTS.
2. The vendor will train necessary LNO staff in the transport, setup, and operation of the recorders.
3. The Systems Analyst will accompany equipment to NSF to capture and stream annual LTER Mini-symposium.
4. The CIO will communicate the capability to the network and solicit requests for use while developing a plan for prioritization and use of the equipment in coordination with the LTER Network communication plan.

Milestones:

- In 2009, three rich media recorders will be purchased and configured for operation with the UNM MTS servers.
- In 2010, necessary LNO staff including the Systems Analyst will be trained in setup and use of the equipment.
- In March of 2010 and each year subsequent, the Systems Analyst will successfully capture and stream the LTER Mini-symposium at NSF for the broader LTER and external science, education, and information management communities.
- In 2011, the CIO will develop a plan based on 2010 demand, use, and feedback for operation and prioritization of the technology.

Evaluation: Annual milestones are met and use and demand are included in annual report to EB and NSF. The CIO reports results and plan of operation as part of LNO annual report to EB and NSF.

Benefits: This activity will benefit the LTER network and the external scientific, education, and information management communities by enabling communication of LTER science and technology to a broader audience.

Feedback: Feedback will be sought on the effectiveness of this effort by including questions on the biennial survey and providing a feedback form directly from the web application.

3.2.1.5 Cyber-security and Data Management Web-based information events. The CIO, NIM, Systems Administrator, and Systems Analyst will develop and support communications about data management, cybersecurity, and other timely and related computer and data security issues to LTER scientists, staff, and students through webcasts and associated web-articles developed for this purpose. Each year, the LNO will offer four technology transfer web events including "how to" and "best practice" guidelines on network and computer security. This activity is dependent on completion of 3.2.1.4 above.

Outcomes: The activities described above will increase awareness and understanding of data management and security issues for LTER scientists, staff, and students.

Operational Steps:

1. The CIO will garner feedback from a variety of sources on topics of interest and priority security items to be covered and will establish a production schedule that will be made available online.
2. The CIO will work with appropriate LNO staff to script the featured webcast.
3. The Systems Analyst will schedule production, secure space, see to operational readiness of the equipment, and capture and publish each webcast according to the schedule.
4. The Systems Analyst will announce the availability and URL of the next installment.

Milestones:

- By the end of calendar year 2009, production schedule is established and initial security topics identified.
- Annually, beginning in 2010 at least four security topics are produced, widely disseminated, and published to the UNM Mediasite server.
- Late in calendar year 2010, feedback on course effectiveness is sought via short web-based survey.

Evaluation: Annual milestones are met, and report and results of survey are included as part of annual report to EB and NSF.

Benefits: This activity benefits the LTER community by encouraging up-to-date knowledge of cyber and data security issues. Although the primary beneficiaries of this activity will be LTER scientists and information managers, the external scientific, education, and information management communities may also benefit.

Feedback: Feedback on this activity will be gathered via questions on the biennial survey as well as through a feedback form directly on the web application.

3.2.2 Information Management. Improve information management (IM) for the Network by supporting communication and coordination among SIMs, strategic data integration, data stewardship, curated data storage, and other data operations that promote Network synthesis and the creation of data legacies.

Description of Activities: The CIO and NIM regularly exchange information with LTER SIMs regarding new technologies and developing information management practices. In addition, the LNO supports short working visits to the LNO by SIMs. The NIM manages the LTER personnel directory, all-site bibliography, site directory, data catalog and several synthetic databases. The budget for these activities totals $636,118 over five years and supports 1.00 FTE of personnel time to manage existing database and develop new data modules. In addition, resources are provided for production meetings for information managers and compensation for sites whose information managers dedicate time to NIS activities.

Broader Impacts: New information management approaches will be embodied in new standard procedures and protocols that will be published as “best practices” documents. By improving information management at a Network level, the LTER Network will become a high quality data

provider for scientists and educators and a repository of scientifically rigorous, richly documented data available for synthesis.

3.2.2.1 Information Management working groups and Site Information Manager travel. The LNO will support working visits to the LNO or other sites by one or two SIMs per year and will support two product oriented workshops each year to address critical IM issues related to participating in the NIS.

Outcomes: Visiting information managers will develop and implement solutions for specific IM challenges affecting the Network including addressing the need for standard specifications for data format, metadata content, quality assurance/quality control (QA/QC), and exchange protocols. Product oriented working groups will each year address critical IM issues related to the NIS. For example, site database personnel will meet to determine functional requirements necessary to implement the water chemistry database requested by the Science Council. Product-oriented working groups may contribute to the development of IM approaches or best practices documents.

Operational Steps:

1. Annually NISAC will make recommendations to the IMC and LNO on priority areas of the NIS development where SIMs could make a valuable contribution.
2. Every year information managers will develop working group proposals and travel proposals.
3. Information managers may discuss proposal ideas with CIO to determine relevance to NIS.
4. CIO will receive proposals from individuals or coordinated through IMExec.
5. CIO will evaluate proposals for alignment with NISAC requests and may request NISAC feedback.
6. CIO will notify working group leaders and IMExec about funding decisions.
7. Working group leaders will complete LNO meeting request form and arrange logistics with LNO core services.
8. Working groups will provide a progress report to LNO by Spring of the following year. They will complete work prior to the IMC meeting of the following year and provide a final report within 1 month of completion.

Milestones:

- NISAC makes recommendations to IMC and LNO in Q1.
- Proposals received from information managers at LNO by CIO in Q2.
- CIO evaluates proposals and notifies organizers in Q3.
- Meeting request form complete and logistics arranged in Q3.
- Work completed before the next IMC meeting.
- Meeting organizers provide report to LNO within one month of completion.

Evaluation: LNO performance in logistical arrangements will be evaluated in annual surveys; IMC working group performance in contributing to the network information system will be evaluated by EB.

Benefits: This activity will benefit the LTER Network by contributing effort to the completion of the NIS. The activity will benefit the broader scientific information management community by providing contributions to a community 'standard of practice' for IM.

Feedback: LNO will receive ad hoc comments from site management, working groups, and NISAC on the value of the activity.

3.2.2.2 Maintain Network databases. LNO staff manage the personnel, site characteristics, and bibliographic databases for the network. The Network Information Manager will have primary responsibility for the management of these Network databases. Network databases will be upgraded and migrated to a new framework through activities in 3.2.3.11. The new framework will include standardized data models based on IMC working group collaborations on this activity and will also include specific web service definitions based on use cases from IMC and others.

Outcomes: The outcome of this activity will be the consistent and up-to-date database of network and site information that can be used to support Network communication and synthesis activities.

Operational Steps:

1. The NIM will begin using existing web-based forms and email-based support system to maintain personnel, site characteristics, and bibliographic data sets.
2. NIM will work with NIS developers and IMC to chart an integration path for personnel, site characteristics, and bibliographic data to the Network Information System framework.
3. NIM will migrate existing data to the new framework.
4. NIM will assist sites in converting to the new framework for managing these data.

Milestones:

- The LNO sponsored working group on network database integration meets in 2010.
- Plan will be developed by the end of year 2 for the future curation of these datasets.
- Migrations and site accommodations will be made by the end of year 4.

Evaluation: Progress on this activity will be reviewed as part of the LNO annual report to EB/NSF.

Benefits: By improving the management of network databases, a greater level of integration can be achieved that provides connectivity between datasets, people, and publications. This interconnected knowledge base will benefit the broader scientific community by increasing the quality and integrity of the LTER data repository.

Feedback: LNO will seek feedback as part of regular surveys and will vet the network database curation plan as a formal request for comment (RFC). In addition, there will be web-based comment opportunities associated with database interfaces.

3.2.2.3 Migrate existing and create new Network databases. The Network Information Manager will work with scientists, the IMC, and sites to migrate existing synthetic data to the LNO and create new synthetic datasets. The NIM will also work with scientists and the IMC to develop new synthetic data products resulting from SC and Decadal Plan activities and integrate them into the NIS framework. The CIO, NIM, and Network Developer (ND) will work with IMC and SC leadership to revise the process for inclusion of new data products.

Outcomes: This activity will increase the accessibility of existing Network data, the number of synthetic data sets, and the quality and number of Network data sets accepted for management and curation by the LNO.

Operational Steps:

1. The NIM will work with staff from the H.J. Andrews LTER (AND) to transition the management and curation of climate and hydrologic databases (ClimDB/HydroDB) to the LNO.
2. The NIM will manage ClimDB/HydroDB in this format until an integration strategy to the NIS has been developed.
3. The NIM will also work with scientists and the IMC to develop new synthetic databases resulting from SC and Decadal Plan activities, along with interfaces to such databases, and integration into the NIS framework.
4. The ED, CIO, ND, DSS and NIM will develop a plan, process, and schedule for integrating new synthetic databases.

Milestones:

- LNO recruits and hires a new information manager by 1 February 2010.
- The ClimDB/HydroDB curation is assumed by LNO 1 May 2010.
- LNO assumes complete responsibility for ClimDB/HydroDB from AND by the end of 2010.
- Plan for new synthetic database integration developed by July 2011.

Evaluation: Progress on this activity will be reviewed as part of the LNO annual report to EB/NSF.

Benefits: Improving information management at the Network level will benefit LTER scientists and the broader community of scientists and educators by creating a repository of scientifically rigorous, richly documented data available for synthesis.

Feedback: LNO will seek feedback as part of regular surveys and will vet the synthetic database development plan as a formal RFC to an appropriately targeted group.

3.2.2.4 Development of NIS derived data products that extend the LNO EcoTrends data and application work. The Network Information Manager will support the management and curation of data error-checked for EcoTrends and harvested into the NIS. In addition the NIM will collaborate with NIS developers, site scientists, and information managers to implement a dynamic process for the creation of EcoTrends and other derived data products. These products

will be selected first as 'low hanging fruit' for proof of concept and later as scientific priorities from datasets that have been error-checked and returned to sites by the EcoTrends Information Manager. Consequently, this activity will start slowly and will not be fully operational until Year 3.

Outcomes: Increase the number of synthetic data sets and increase the quality and number of Network data sets accepted for management and curation by the LNO.

Operational Steps:

1. The NIM in coordination with the EcoTrends Information Manager and CIO will identify data sets ready for harvesting.
2. The NIM will validate these datasets for quality of metadata and completeness of data using NIS data validation tools and other methods.
3. The NIM will coordinate harvesting with the SIM and confirm loading of data from site repositories into the NIS Data Cache.
4. The NIM will coordinate the adoption and/or development of program scripts with the Site Information Manager and create workflows to transform data from native formats to the EcoTrends time-series format,
5. The dataset will be reviewed by the CIO and accepted for LNO curation.
6. The process will be repeated for new datasets as identified.

Milestones:

- Datasets deemed 'low hanging fruit' are accepted (2011 Q3).
- Production begins on datasets deemed scientific priorities (2013 Q4).

Evaluation: Target datasets will be established annually as part of the development of LNO performance criteria for review. Progress on this activity will be reviewed as part of the LNO annual report to EB/NSF.

Benefits: This activity will benefit the LTER Network and the greater ecological community by providing new derived data products for synthesis.

Feedback: Feedback will be included as part of biennial LNO surveys to the LTER community. In addition, priorities for new derived products will be set with regard to the plan developed as part of 3.2.2.4 above.

3.2.3 Network Information System. Complete the design, development, and implementation of the LTER Network Information System in collaboration with NISAC, IMC, and strategic community partners.

Background: The mission of the LTER Network Information System is to promote advances in collaborative and synthetic ecological science at multiple temporal and spatial scales by providing the information management and technology infrastructure to increase:

- *availability and quality of data from LTER sites* – by the use and support of standardized approaches to metadata management and access to data;
- *timeliness and number of LTER derived data products* – by creating a suite of middleware programs and workflows that make it easy to create and maintain integrated data sets derived from LTER data; and
- *knowledge generated from the synthesis of LTER data* – by creating standardized access and easy to use applications to discover, access, and use LTER data.

Resources dedicated to the development of the NIS include 2.00 FTE of personnel time and travel to NISAC meetings at a total cost of $727,792 over five years. Partial funding for these efforts comes from the NSF Office of Cyberinfrastructure (OCI).

Roles associated with the development of the LTER Network Information System at the LNO are:

- LTER Chief Information Officer (CIO) – responsible for LTER Network Information Infrastructure.
- NIS Developer (ND) – lead developer of LTER Network Information System.
- NIS Analyst/Programmer III (NAP3) – lead programmer of the LTER Network Information System.
- NIS Analyst/Programmer II (NAP2) – secondary programmer of the LTER Network Information System.
- NIS Programmer Intern (NPI) – tertiary programmer of the LTER Network Information System.
- Application Support Analyst (ASA) – LTER Network Office lead web-application developer.
- Network Information Manager (NIM) – LTER Network Office lead information manager.

Roles associated with the development of the LTER Network Information System at the Network are:

- Network Information System Advisory Committee (NISAC) – provides guidance and priorities for NIS development from a domain science and information management perspective.
- Information Management Committee (IMC) – provides planning and reviews input from a site information management perspective; the IMC may filter input through specific individuals or groups within the information management community.
- Site Information Manager (SIM) – site-based information manager.
- Executive Board (EB) – decision making body for the LTER Network; review of overall operational plan.

3.2.3.1 LTER NIS/PASTA framework overview

The LTER Network Information System will utilize the Provenance Aware Synthesis Tracking Architecture (PASTA) framework (Figure 1), which will provide the LTER community a

metadata-driven data-flow architecture for automatically loading data from LTER research sites and making it available through a standard and well defined software interface. We distinguish the PASTA framework from the more generalized Network Information System by classifying framework components as critical and enabling cyberinfrastructure that, collectively, provide the services defined by the above mission. All data products within the NIS will have associated metadata, including provenance information where applicable. We anticipate that sites will make available a majority of their data for harvesting into the NIS, with an initial focus on well documented tabular data, which will be followed by more complex data (i.e., spatial and remote sensing data, and/or video data) at a later date. A key goal of the PASTA framework is to simplify the site's burden of participating in the NIS by only requiring that the data be described with complete metadata using the Ecological Metadata Language standard and having the data accessible through one or more Internet protocols. Adoption of any future LTER data standards at the site is not necessary for harvesting into the NIS, but will facilitate future analysis and/or synthesis of such data by applications internal and external to the NIS. Site data that is harvested into the NIS will be available to the broader community through the NIS Data Portal and through the site's information system (if the site wishes to support such access).

We view participation by the site, and specifically, by Site Information Managers, as critical to the overall success of a fully functional Network Information System by 1) participating in the development of NIS tools and applications that specifically support site information management needs, 2) informing the development of the PASTA framework components through direct and indirect transfer of knowledge and experience to NIS developers, and 3) creating well documented and quality data products for harvesting into the NIS. In addition, the collective or individual components of the PASTA framework may service site needs by providing a centralized and off-site data archive and distribution point, metadata and data congruency checks for quality assurance and correctness, automated metadata generation and management, data access and use audits, and a source of workflow and data transformation algorithms, to mention but a few. We note that site participation will be a function of funding that is available to support Network-level information management goals. Overall, we expect that the completed NIS will reduce the workload (and cost) of sites to participate in Network synthesis and analysis.

The PASTA framework comprises nine functional elements of the LTER Network Information System – 1) the data management suite of components (Data Manager, Data Catalog, Data Loader, and Data Cache), 2) the Workflow Manager, 3) the Metadata Factory, 4) the Data Portal, 5) the Discovery/Access Application Programming Interface (API), 6) the Identity Management Services, 7) the Persistent Identifier Services, 8) the System Monitor, and 9) the Audit Services. Each component will interact to varying levels with an integrated security model. The Data Manager, Data Catalog, Data Loader, Data Portal, and Audit Services (formerly, Data Access Server) have been developed to varying degrees under previous funding, but will require modifications and enhancements necessary for inclusion into the PASTA framework. Identity Management Services and Persistent Identifier Services will be adopted from community and industry standards; PASTA development efforts will consist primarily of integration and deployment. The remaining elements require new development and are addressed as part of this operational plan, as described in detail below.

Site Information Managers will interact directly with the Data Manager component interface to configure and schedule metadata harvests into the Data Catalog and to identify "PASTA-ready" data (i.e., data that are made available to PASTA and conform to the necessary metadata standards). Changes to metadata in the Data Catalog that represent new and/or updated site data will trigger the loading of this data into PASTA by the Data Loader. Once in the Data Cache, the site-based data will be available for synthesis projects to produce value-added derived data by incorporating various workflows into the Workflow Manager. The resulting derived data may become part of a recognized NIS Data Module, which is typically an agreed upon data model resulting from a science-driven goal. Provenance metadata will be captured during processing for derived data products and will be included in the final metadata package generated by the Metadata Factory. Discovery and access of both site-based and NIS Data Module data products will be achieved through a standard interface by value-adding applications, such as the LTER NIS Data Portal. The Audit Services will support and comply with the LTER Data Policy to track LTER data access and usage.

3.2.3.2 LTER NIS data levels

All data products within the NIS will be classified from levels 1 to 4 (level-0 refers only to data located at the site) based on the level of post-processing applied to the original site data. Although there exist many standards for classifying data based on processing (e.g., NASA, CODMAC, EOC, and NEON), we have adopted a modified version of the Ameriflux Network classification scheme because of similarities of definitions for processing level between the two networks. The following classification is suggested and may be modified after further evaluation by NISAC, IMC, or another designated party:

> Level-0 – Site located data (may be raw or modified by the site) that are made available for harvest into the NIS. These data may be accessed directly through the site resource or the Audit Services (if registered). Discovery of Level-0 data may be performed through the LTER Data Portal and/or Discovery/Access API (if the associated metadata are harvested into the Data Catalog and the metadata supports direct data access) or through the site resource.
>
> Level-1 – Data that are harvested into the NIS as archived replicates of the Level-0 site data; the Level-1 storage structure may differ from that of the Level-0 structure. All Level-1 data will have EML metadata documentation identical to the site harvested metadata with the exception to the physical storage description and will be discoverable through the LTER Data Portal and/or Discovery/Access API. Provenance information will include reference to the Level-0 data.
>
> Level-2 – Data are structurally corrected (e.g., realignment or adjustment of columns to match metadata), but retain the same content where possible of the Level-0 data. All Level-2 data will have EML metadata documentation identical to the site harvested metadata with the exception to the physical storage description and/or data description and will be discoverable through the LTER Data Portal and/or Discovery/Access API. Provenance information will include reference to the Level-0 data and any processing correction applied.

Level-2 data may or may not be accepted into the NIS depending on the cause for correction and/or changes in LTER policy.

Level-3 – Data that are qualified and processed into one of the NIS Data Modules. Processing may include unit normalization, new label conventions, adjustment to reporting intervals, and reformatting as necessary to be consistent with the target data product. These data will be available through the LTER Data Portal and/or the Discovery/Access API. Provenance information will include reference(s) to the previous metadata document(s) and all processing steps applied to the data to generate the Level-3 data product.

Level-4 – Data that is gap-filled and semantically adjusted to meet the needs of particular synthetic data products. This data level is a place holder for anticipated future use.

3.2.3.3 Software development methodology

The software life-cycle and development process of the LTER Network Information System and PASTA framework will follow the Rational Unified Process (RUP), which is an industry standard for iterative and incremental software development. The RUP is an adaptable framework in which iterative cycles are combined into four primary time-segments (inception, elaboration, construction, and transition) across a single development goal. Each iteration consists of a series of operational steps that vary in effort according to the current time-segment. These steps, and their associated activities, are:

1. Component Modeling – Identification and refinement of component goals, including defining the business goal of the component (i.e., the benefits and/or high-level achievement of the component); identification of current iteration goals,

2. Requirements – Identification of use-case scenarios that define and drive functional and non-functional requirements of the component (use-case scenarios are critical to the initial software development process for identifying system functionality; utilizing use-case scenarios in planning often results in more precise requirements, which generally translates to more accurate and timely deliverables),

3. Analysis & Design – Analysis of use-case scenarios and refinement of requirements; generation of component design (using the Unified Modeling Language); identify relevant technology; identify component unit tests to meet requirements,

4. Implementation – Implementation of unit tests; implementation of component software and evaluation against unit tests,

5. Test – Review and evaluation of completed component software for iteration or overall component; confirm completeness and correctness against unit tests; sign-off of iteration milestone, and

6. Deployment – Place software in functioning environment.

The completion of each iteration progressively achieves the software development goals of the system by incrementally releasing successively more functionality. This incremental process scales throughout the project beginning with each individual iterative cycle, to the completion of a component, to the integration of components into deliverable and testable packages, and finally, to the fully complete and functioning system. Because it is an iterative process, the RUP can adapt to modifications in both goals and schedules without complete restructuring of the operational plan. Each PASTA framework component will follow its own iterative development trajectory, thereby introducing high-frequency release cycles that will allow timely evaluation and feedback by stakeholders. As such, planning details are better defined for earlier cycles; later cycles will be fleshed out in greater detail as they near the present. The development process will also generate comprehensive documentation for the planning of all framework components, including progress reports, which will be made available through the NIS project website.

All software source code that is developed as part of the LTER Network Information System and the PASTA framework will be considered "open source" and will be licensed as "permissive free software". The project will follow the guiding principles of "open source development" – that is, access to software source code, including planning and reference documents, will be unfettered and freely available to the general public for review, comment, and reuse, and the contribution of software improvements and features from the community may be accepted back into the project under a managed process. As such, all software development artifacts will be stored in a revision control system and will be publicly accessible through the LTER Network software repository. Software collaboration and contribution deemed non-critical to the operational functions of the PASTA framework, but significant to the general Network Information System, will be integrated into the system where and when appropriate. To this end, the specific type of "open source" license and guidelines for the contribution of software from the LTER Network and broader community will be presented to NISAC for evaluation prior to the start of actual development. An "open source" policy statement will be written and clearly presented as part of the LTER strategic communication plan and on the NIS project website.

Software development and framework component integration will occur in two major phases. The first phase will focus on delivery of core functionality of the PASTA framework and result in an operational prototype by the second quarter of 2012 (Figure 2). We will enlist a set of 3-4 "early-adopter" sites to assist in testing and evaluating software of the operational prototype. The second phase will concentrate on extending and enhancing existing components, and include addressing more complex functionality. The complement of remaining sites will be recruited during the second phase. The second phase will end with the production deployment of the PASTA framework in early 2014, which will be followed by an ongoing operational and maintenance period.

Official releases of the Network Information System that represent functional milestones will be designated with a numeric string denoting the *major-minor-revision* version of the release (e.g., NIS 1.3.5). An increment in the *major*, *minor*, and/or *revision* version signifies a forward progression towards the final goal of the system. A *major* version change generally indicates a significant improvement or addition to features/functionality of the software; such changes may not be backward compatible with a previous *major* version. A *minor* version change may also

include new features/functionality, but backward compatibility will be retained. A *revision* change simply includes software "bug" corrections and does not generally provide any new features or functionality. The LNO will plan frequent releases that engage the community to generate awareness and excitement for the LTER Network Information System. These releases will strive to provide a sufficiently rich set of features so less technically inclined stakeholders would be able to use its functionality (see "Package Level" in section 3.2.3.6). Such awareness will be critical in the adoption of the NIS, in addition to providing regular feedback to system developers on the capabilities of the PASTA framework. To this end, the LNO will develop a "milestone release schedule" within the first six-month period of active planning, including a list of potential benefits associated with each release. Releases will be announced to the stakeholder community as defined in the LTER strategic communication plan.

3.2.3.4 Design, technical, and data standards

Acknowledging and adopting standards is critical for interoperability between different communities and for improving efficiencies of the overall project development plan. To help disambiguate the standards addressed in this operational plan, we have categorized standards into three categories:

1. Design and usability standards for user interfaces – having to do with the programming of web user interfaces (e.g.,Web Content Accessibility Guidelines 2.0, W3C). Due to the extensive use of the Internet, and specifically, web-based user interfaces in the NIS and PASTA framework, we will adopt the applicable guidelines suggested by the World Wide Web Consortium (W3C) Web Accessibility Initiative for web interface usability and access as specified in the Web Content Accessibility Guidelines 2.0. This standard addresses the accessibility of web-based presentation, primarily for disabled users, but also for agents with limited device capability, such as mobile phones.

2. Technical standards – published and acknowledged standards that define aspects of software development, the interoperability of software components, the structuring, documentation, and communication of data (e.g., Rational Unified Process, Unified Modeling Language, Ecological Metadata Language). Adoption of technical standards ensures interoperability and leads to faster development time through "reuse" of effort performed by other communities. Development of the PASTA framework, and more generally the NIS, builds upon the Ecological Metadata Language (EML 2.1) as the core metadata standard for describing LTER Network generated data. The EML is hierarchical in design (i.e., specified as an XML schema) and flexible in specificity of the information used to describe data. For example, streaming sensor data is now very common across all 26 LTER Network research sites. A standard for streaming sensor data, approved by the Open Geospatial Consortium for describing sensor configuration and measurement processes, is SensorML. We recommend that data generated by streaming sensors be described with the EML standard, which can "wrap" the pertinent SensorML XML in either the "methods" or "additional metadata" section of the EML. It is the high-quality EML that will enable harvesting the sensor data into the NIS. Other standards that are being evaluated for use in the NIS include: Handles, DOIs, PURLs, and LSIDs for persistent identification of digital objects; SKOS and Z39.19-2005 for

controlled vocabularies; InCommon, OpenID, and LDAP for user/agent authentication. The software development process will adopt the Rational Unified Process standard for life-cycle management, including the Unified Modeling Language for design.

3. Data standards – project specific specifications for the storage, display, and communication of data products (e.g., NetCDF for climate data). For the Network Information System, data standards provide interoperability through the use of common storage formats (e.g., NetCDF for climate data) and, perhaps more important, the ability to integrate disparate data for generating derived and synthetic ("downstream") data products (e.g., ClimDB/HydroDB and EcoTrends). Identification and adoption of data standards is primarily the domain of the science working groups, but will also require interaction with technical personnel for integration into the NIS.

Although there are standards currently identified for use within the NIS/PASTA framework (e.g., EML), additional standards will be evaluated through a vetting process by NISAC, IMC, and other relevant parties (e.g., domain scientists).

3.2.3.5 Security model

The LTER Network Information System, and more specifically the PASTA framework, provides the opportunity for members of the LTER Network and the broader community to interact directly with an Internet-based system to discover and access LTER Network data products. Equally important, LTER Network site information managers will need access for ingesting metadata and data into, and interfacing directly to tools provided by, the NIS. The LNO acknowledges that the nature of interaction proposed between the LTER Network community and the general public to the NIS dictates that a strong, but open, security model be adopted. We recognize three significant classes of security that must be incorporated into the NIS:

1. Identity/Authentication Management – Members of the LTER Network and related communities often access similar web-based services during daily work routines (e.g., Gmail, MySpace, and Amazon.com). Most, if not all, of these services require a user's identification and an associated password to accurately match (or authenticate) a specific user with their service identity before any transaction can occur. The identification and password information provided by users to these service providers is often called their "credentials". It is a goal of many cyberinfrastructure projects to offer their users a single credential that is accepted across different communities for authentication, thereby eliminating the need for multiple identity recognition efforts for the same user. The current authentication mechanism for accessing LTER Network resources is performed through the validation of a user identity within a local LDAP database. To utilize a broader system, the LNO will adopt the preferred community standard for federated identity management (see section 3.2.3.10.5, Identity Management Services). We will work collaboratively with organizations that research and support federated identity management, like the National Center for Supercomputing Applications' CILogon project, to integrate a comprehensive identity management approach for the NIS so that users will be able to authenticate with the same identity across common service providers.

2. Access Control to Metadata/Data – All metadata and data within the LTER Network should be considered publicly accessible due to its origin of funding support and its importance for public knowledge. Understandably, however, certain metadata and data should be embargoed for limited periods of time (e.g., innovative or novel approaches to research that require further validation) or be banned from general access (e.g., sensitive information regarding endangered species). The Ecological Metadata Language, used to describe LTER Network data products and the key metadata standard identified for the NIS, provides access control directives for both metadata and data. As such, all components and services in the NIS/PASTA framework that interact with metadata and/or data will adhere to the access control rules documented within the EML.

3. System-level Security – The LTER Network Information System, and specifically the PASTA framework, is designed to be Internet accessible and supports interactions with many users. The underlying system software and its hardware, like most cyberinfrastructure, are vulnerable to both malicious and non-malicious events. To this end, NIS software development will embrace security "best practices" (e.g., International Information Systems Security Certification Consortium's Best Practices for Secure Software Development) to ensure that all systems include a committed level of secure operation. Design and implementation practices will be reviewed for consistency in techniques for prevention of malicious activity, while procedures for backup and redundancy will be followed for metadata/data integrity. In addition, administrative personnel will be trained in protocols for secure management of computer and Internet resources (e.g., LTER Network Office Cyber-security briefings).

3.2.3.6 Community engagement

Success in the development of an information system depends not only on the cyberinfrastructure underlying the system but also on the engagement of stakeholders whose needs define system functionality. Stakeholders must be engaged in a broad range of activities from documenting data to evaluating applications. The underlying motivations for stakeholder engagement in the NIS are the joint responsibility of the stakeholder and developers to make LTER data accessible and the desire to have improved tools for research synthesis. However, to encourage significant and persistent involvement across a broad cross-section of the LTER community, proactive steps must be in place to remove or lower barriers to participation as partners in system design and development. Mechanisms in this operational plan to promote stakeholder engagement are described below.

The most important element in fomenting successful partnerships with LNO stakeholders is frequent and useful communication of progress and of opportunities for involvement. The LNO is committed to insuring that such communication takes place, and to that end will develop a plan that is focused on the NIS as part of the LTER strategic communication plan (see section 3.4.1 for additional detail). As part of this process, the LNO will seek expert advice on the development of partnerships between domain scientists, information managers, and system developers. In addition, the LNO will consider the use of social networking approaches focused

on the NIS and submitting a proposal to the CI-TEAM program at NSF to disseminate information to LTER scientists on best practices in building effective partnerships.

The LNO will assist leaders of new LTER Network research initiatives in preparing effective data management schemes and proposals that reflect those schemes. The processes by which LTER research teams standardize and analyze data and synthesize results provide important information for system developers, but defining and capturing the steps of those processes require special effort. To that end, the LNO will also provide support to compensate information managers to implement these schemes for each research group and to translate the analytical processes into functional requirements for the NIS. To build partnerships between domain scientists, information managers, and system developers, the LNO will engage an expert to offer advice in the formation of research teams to this end. The LNO will work with domain scientists and information managers to develop best practices for data management, to formulate data standards for targeted data sets, and to identify non-LTER data that need to be incorporated into the NIS.

Engagement with the LTER community during software development will occur at multiple levels, including at the iterative, component, package, and system levels (see below). Each level defines the granularity of the development effort and will provide a "check point" with a community stakeholder to evaluate and sign-off on the level milestone (Figure 3). The following list provides a general description of the engagement level from the finest to coarsest development effort and the community stakeholder responsible for interaction:

Iterative – The "Iterative Level" is defined as the finest level of effort and consists of a single iteration within the development period of an individual PASTA framework component. Points of engagement will be addressed at the "modeling", "requirements", and "testing" steps of each iteration (see section 3.2.3.3 regarding the use of RUP); evaluation and sign-off of the iterative milestone goal will be performed by the community stakeholder at the end of the "testing" step. The community stakeholder at this level is a small number of individuals (e.g., information managers, students, and domain scientists) with a vested interest in the specific component being developed and who agree to provide timely review and feedback to NIS developers throughout the development period. Collectively, this set of individuals is called a "tiger team". The identification of individuals for a "tiger team" will be the joint responsibility of NISAC and IMC.

Component – The "Component Level" is recognized as the deliverable software of a PASTA framework component when all requirements that have been identified are complete (the sum output of all iterations). The component milestone marks the point of engagement with the community stakeholder. We see the stakeholder at this level as a combination of NISAC, IMC, and/or a designated party who is assigned to the task. The community stakeholder will evaluate the component for completeness and correctness (as defined by its list of requirements). A successful milestone will require that the community stakeholder sign-off on the completion of the component.

Package – The "Package Level" is defined by the successful integration of one or more PASTA framework components into a community usable package that delivers higher level

and more complete functionality (e.g., integration of the Data Manager, Data Catalog, Data Loader, and Data Cache will provide a end-to-end metadata and data harvesting package that identifies and loads data from the site and into the NIS). The point of engagement occurs when a fully integrated package is available to the community as a NIS version release (see section 3.2.3.3). The version release will be announced through channels defined by the LTER strategic communication plan and will solicit for "early adopter" sites and individuals who will evaluate and use the package as a prototype of the NIS. Review and feedback of the package will compose an informal "check point", with positive reaction considered to be successful sign-off from the stakeholders.

System – The "System Level" consists of a fully functioning Network Information System, including fulfillment of all requirements identified for components of the PASTA framework. Points of engagement are coarse grained and are represented by all users of the LTER Network Information System, including LTER Network scientists, information managers, students, and the broader scientific community. Feedback will occur through systematic surveys and reviews, as well as informal forms that will be available through NIS user interfaces. Successful completion of all system goals will result in the LTER Executive Board signing off on the development phase of the NIS (at which point, the NIS will enter an operational and maintenance phase).

The PASTA framework will provide direct opportunities for LTER and non-LTER stakeholders to create applications that interface with the NIS through the Discovery/Access Application Programming Interface. The LNO will also create specific opportunities for the information management community to address site functional requirements through support of product-oriented workshops, release time for information managers, visits to sister sites and the LNO, and formal training exercises.

System developers will set schedules so that software will be released in operational packages that incrementally meet functional requirements and thus provide useful tools to stakeholders (see section 3.2.3.3). In addition to the "requests for comments" directed to a broad stakeholder community, smaller focus groups of students will be used to evaluate release functionality in more detail. By partnering with several early-adopter sites, system developers will receive frequent informal feedback on problems with interfaces to site data systems. Finally, metrics of user satisfaction obtained through biennial surveys of the LTER community and metrics from the annual evaluation of software development by NISAC will provide additional input to the development team.

3.2.3.7 Metrics of progress

Development progress on the NIS will be evaluated annually by the Executive Board with input from NISAC on technical elements. Evaluation of progress requires mutually-agreed upon goals and milestones and metrics to assess whether these milestones have been met. Three kinds of metrics will be documented: 1) metrics of user satisfaction, 2) descriptive (use) metrics, and 3) software development metrics. Metrics of user satisfaction measure whether stakeholder expectations are met within the period under consideration. Questions relating to user satisfaction with the NIS are included in the biennial survey of the LTER community, and

changes in user satisfaction will be measured over time. In addition, focus groups for each software package released will provide immediate feedback on user satisfaction. Descriptive statistics measure the how many people have adopted software applications and the intensity of use. The number of data sets in the NIS and the number of people accessing these data are examples of descriptive statistics that will be collected by the LNO. Software development metrics are divided into two kinds: *functional* metrics (e.g., incorporate a search engine) and *non-functional* metrics (e.g., return results from a standard search within five seconds). For each annual milestone set for the NIS, appropriate metrics will be jointly chosen by the EB and NIS developers. NISAC will evaluate annually whether technical milestones have been met using data provided by the LNO.

3.2.3.8 Policies for incorporating external data

The desire to incorporate data external to the LTER Network (e.g., NADP data) in synthesis analyses has been expressed in various contexts. However, because the challenges of incorporating external data are significant and the range of possible data sets large, the incorporation of external data will need to follow a strict policy. LTER data will have the highest priority in the initial stages of NIS development. Synthesis needs as defined by LTER research projects will dictate the scope and priority for eventual inclusion of external data in the NIS. Access to external data will be incorporated into workflows for specific synthesis projects and provided by existing data at sites. In the long run, links to DataONE and other data warehousing projects will provide broader access to data external to LTER.

3.2.3.9 Risk management

Risks are inherent in any cyberinfrastructure/software development project. While it is not efficient use of resources for a project of this size to engage in a comprehensive risk management approach, it is useful to identify the highest known risks and acknowledge that there are unforeseen and unidentified risks. The primary impact of risks in software development projects is schedule slippage, and a general mitigation strategy is to assume the risks through expanded schedule contingency. To address this concern, schedule contingency has been included in our work plan to mitigate the majority of these risks. Other risks might be mitigated through the reallocation of effort and/or of funds identified in salaries, working groups, sub-contracts, and mini-sabbaticals, or project de-scoping. The CIO and Network Developer will monitor risk as part of their regular schedule of project management tasks to make sure that emergent risks are identified early and to account for schedule contingency that has been expended on known risks.

The known high-risks are reported here, along with the mitigation strategy:

1. Analyst/Programmer staffing – The analyst/programmer staffing for the NIS development is small at 2 FTE + 1 FTE annual intern. The risk is that there will be a high level of turnover among these staff leading to schedule slippage, high workload for existing personnel, and potential for low morale as a result. Mitigation of this risk is through schedule contingency. If monitoring reveals that schedules have slipped beyond the allocated contingency, the approach would be to consider reallocation of resources or de-scoping the project.

2. Software integration challenges – The NIS relies on a number of open-source software components that will be integrated as part of the PASTA framework. The risk is that the software products might not integrate as expected, resulting in some levels of software re-engineering, additional workload, and schedule slippage. Mitigation of this risk is through schedule contingency. If monitoring reveals that schedules have slipped beyond the allocated contingency, the approach would be to consider reallocation of resources or de-scoping the project.

3. Workflow product complexity – The Workflow Manager component of the PASTA framework will drive the integration and transformation of data into derived data products. Current planning looks to the maturation of workflow products like Kepler and Taverna to fill this role. The risk is that integration of workflow systems into the PASTA framework may prove more complex than anticipated, thereby resulting in additional integration work. Mitigation of this risk is through schedule contingency. If monitoring reveals that schedules have slipped beyond the allocated contingency, the approach would be to consider reallocation of resources or de-scoping the project.

4. Site metadata quality – The NIS is dependent on high quality EML metadata describing published data products at the sites. The risk is that sites will not be able to participate because of metadata quality issues, lack of preparation, limited resources, or disinterest. This risk is expected to be partially mitigated by the ongoing expenditure of funds to train site information managers in the preparation of metadata for the NIS and in developing QA/QC metrics that can provide quality feedback to the sites via the Data Manager/Data Loader components.

5. Data Portal product complexity – The NIS Data Portal is the primary access point for LTER data. While basic functionality of delivering LTER data products is assured based on prototypes, the risk is that functional requirements developed from use cases by scientists will prove to be more complex than can be accomplished within the scope of this project. The project assumes this risk and will have to engage in feature prioritization as this component develops.

6. Future budget scenarios do not support NIS maintenance and operations – There is a high risk associated with the completion of the NIS project and having inadequate resources to operate and maintain the NIS. The mitigation strategy here is to develop good estimates of operations and maintenance costs and convey this information to LTER management and NSF program managers as the project moves toward completion.

There are a number of known medium and low risks related to computing and data capacity underestimates that are not elaborated here.

3.2.3.10 Completion of the LTER Network Information System.

The following sub-sections describe in detail pertinent development activities related to the PASTA framework as outlined in the proposal and over the 5-year performance period; these

include the data management suite of components (Data Manager, Data Catalog, Data Loader, and Data Cache), the Workflow Manager, Metadata Factory, NIS Data Portal, and Audit Services. Four additional activities, 1) development of a Discovery/Access Application Programming Interface (API), 2) evaluation and adoption of Identity Management services, 3) evaluation and adoption of Persistent Identifier Services, and 4) development of a System Monitor, have since been identified as necessary to complete the LTER NIS.

As described previously, software development of PASTA framework components will follow the Rational Unified Process methodology (see section 3.2.3.3). As such, each component's development life-cycle will follow the same general operational steps (component modeling, requirements, analysis and design, implementation, testing, and deployment). In general, the sequence and description of operational steps are similar for each component. The principal LNO personnel involved in each of the component development efforts include the Chief Information Officer, the NIS Developer, and software developers (Analyst/Programmer 3, Analyst/Programmer 2, and a Programmer Intern). The Chief Information Officer and NIS Developer are responsible for interaction with LTER community stakeholders and developing, along with stakeholders, the overall design and goals of the NIS. The NIS Developer, together with all software developers, will develop detailed design plans and perform the implementation and testing of all system software. Review of completed software will be the responsibility of the Chief Information Officer, the NIS Developer and the designated community stakeholder. Engagement of the LTER community is described in greater detail in section 3.2.3.6.

3.2.3.10.1 Data Manager, Data Catalog, Data Loader, and Data Cache

Description: The Data Manager, Data Catalog, Data Loader, and Data Cache suite of PASTA framework components are tightly coupled services that provide both metadata and data harvesting capability to the NIS. These four components represent the automated extraction and loading process to move Level-0 data into a managed community data cache (Level-1).

Site Information Managers will interact directly with the Data Manager user interface to register metadata for harvest into the Data Catalog on regular intervals. All metadata will conform to the Ecological Metadata Language (EML) standard through established LTER best practices. The same registration process will allow information managers to mark specific data sets identified within the metadata for harvest by the Data Loader. Once a data set is registered, the Data Manager will initiate loading of Level-0 data from the site by the Data Loader only if they are new and/or updated as indicated by changes in the harvested metadata. These Level-0 data will be replicated in the centralized Data Cache using the same general format designated by the site, but in a PASTA-specific structure designed to accommodate the class of data being loaded (e.g., tabular data will be stored as a relational database table, remote sensing data as binary objects, and GIS data in a spatially compatible structure). These replicates will be characterized as Level-1 data. Services of the Data Loader will provide "structural" quality control and error detection of Level-0 data, thus enabling corrections and/or flagging of these data prior to being stored; corrected/flagged data will be characterized as Level-2 data. All errors or flagged data that produce Level-2 data will be reported back to the originating site. The Data Manager service interface will also provide metadata/data quality checking capabilities (i.e., ensuring data formats comply with descriptions in metadata) to site information managers for pre-harvest

evaluation of their data packages (metadata and data). All harvest event information will be logged and provided back to the site as regularly scheduled and/or ad-hoc reports.

The Data Manager component will also provide an integration point for site-developed tools, such as the Controlled Vocabulary and Unit Registry. The process of integration will be decided through collaborative planning between site and NIS developers during initial planning of the Data Manager component.

Outcomes: The Data Manager, Data Catalog, Data Loader, and Data Cache will provide automated extraction and loading of site metadata and data into a community repository for centralized access and preservation. The Data Manager user interface will act as a single point of interaction for scheduling metadata/data harvests, serve as the interface for ad-hoc evaluation of data packages, and provide an integration point for site developed tools. All harvest events and related information will be documented and provided to Site Information Managers as regularly scheduled or ad-hoc reports.

Operational Steps:

1. The ND will compose detailed use-case scenarios through targeted requests and/or formal working groups of stakeholders, including SIMs, IMC, and NISAC, that characterize the Data Manager service interface and operations of the Data Catalog, Data Loader, and Data Cache.
2. The ND will analyze use-case scenarios and will formulate functional requirements and an implementation plan for tabular data only (phase 1); the implementation plan will be reviewed by the IMC and NISAC for completeness and correctness.
3. Execution of the implementation plan will be performed by the software developers; incremental releases will include prototypes of the user interface.
4. Review of operational prototype deliverables will be performed by the IMC via a RFC and NISAC via regular annual review process.
5. Review of use-cases and functional requirements, including new use-cases for additional data classes (phase 2), will be performed by the CIO, ND and NISAC during second year review.
6. ND will generate new/additional functional requirements that will be added to the implementation plan.
7. Execution of implementation plan will continue by the software developers (phase 2); incremental releases will include prototypes of the user interface.
8. A final review of deliverables by the IMC and NISAC will be performed to measure success.

Milestones:

- Use-case scenarios document (phase 1) (2010 Q2)
- Implementation plan (phase 1) developed (2010 Q2)
- Operational prototype (phase 1) deployed (2011 Q2)
- Review of operational prototype (2011 Q3)
- Use-case scenarios document updated (phase 2) (2012 Q2)
- Implementation plan updated (phase 2) (2012 Q2)

- Production deployment of Data Manager, Data Catalog, Data Loader, and Data Cache (2012 Q3)
- Review of final deployment (2012 Q4)

Evaluation: Criteria for evaluation will be included in annual LNO performance review criteria and will cover delivery of all functionality as described by the requirements and implementation plan on the scheduled date(s). Review of performance will be by NISAC during regular annual reviews.

Benefits: The Data Manager service interface will benefit Site Information Managers by streamlining the process of harvesting metadata/data into the NIS. The LTER community will benefit from data being cached in a community repository in an accessible format that enables persistence through managed curation. The broader scientific community will be able to use this service to perform data quality checks on EML described data.

Feedback: Direct feedback will occur from Site Information Managers to the ND for evaluating the efficacy and efficiency of the Data Manager service interface via an RFC process. Internal NIS metrics will measure the number of harvested metadata and data objects, thereby providing indirect feedback on usage.

3.2.3.10.2 Workflow Manager (Transformation Engine)

Description: The Workflow Manager provides management and execution of workflows by authenticated individuals and/or on behalf of the system. A workflow in the PASTA framework is simply a sequence of computational tasks that perform some action on one or more datasets that reside in the Data Cache. Workflows provide a standardized way to document the processing steps to generate derived data products and the ability to modify and retrace those steps. The level of automation possible and the effort required to develop a workflow is highly dependent on the complexity of both source data and the data model. Source data complexity would be reduced because these data will be described using the EML standard and will reside in a common repository. Data models will be defined in research working groups, and algorithms to generate Level-3 data will be documented by SIMs integrated with the working group. Initial development of the Workflow Manager will rely on manually written scripts and/or executable code that access predefined data sets from the Data Cache for generating derived data products found in existing NIS Data Modules. Future work will allow the user to create workflow sequences by integrating external workflow management systems (e.g., Kepler and/or Taverna) into the Workflow Manager. The Workflow Manager will also include a "state monitor" for system administrators and workflow owners to oversee the execution of one or more workflow sequences, in addition to system reporting of all events related to the Workflow Manager.

A first order Workflow Manager scenario would begin with the installation of one or more executable scripts that read data from one or more predefined Level-1 or Level-2 data sets found in the Data Cache, operate on that data, and then write output to the appropriate NIS Data Module (e.g., the ClimDB for derived climate data). The executable scripts and the script-dependent data sets found in the Data Cache would be registered with the Workflow Manager. When new or modified data are added to the Data Cache, the Data Manager will trigger a "data

update notice" to the Workflow Manager. The Workflow Manager will then resolve all data dependencies and execute the appropriate script(s). New derived data will be added to the corresponding NIS Data Module and updated metadata will be added to the Data Catalog.

Outcomes: Completion of the Workflow Manager will provide users with the ability to define a sequence of computational tasks that may be applied to one or more data sets residing in the Data Cache and to have the same tasks executed automatically when new and/or updated data are harvested into the NIS.

Operational Steps:

1. The ND will compose detailed use-case scenarios through targeted requests and/or formal working groups of stakeholders, including site scientists/researchers, SIMs, IMC, and NISAC, that characterize the Workflow Manager services, including levels of interoperability, interface requirements, and workflow library interactions.
2. The ND will analyze use-case scenarios and will formulate functional requirements and an implementation plan for manually generated scripts and executable code (phase 1).
3. The implementation plan will be reviewed by the IMC and NISAC for completeness and correctness. The implementation plan will utilize one or more of the current NIS Data Modules (e.g., ClimDB and/or HydroDB) as a baseline example.
4. Execution of the implementation plan will be performed by the software developers to demonstrate an operational prototype that can execute workflows either manually or on a schedule.
5. The CIO and ND will evaluate off-the-shelf workflow process management solutions (phase 2).
6. Review of operational prototype deliverables will be performed by the IMC and NISAC.
7. Review of use-cases and functional requirements, including new use-cases for integrating more complex workflow applications (e.g., Kepler and/or Taverna) (phase 2), will be performed by the CIO, ND and NISAC during second year review.
8. ND will generate new/additional functional requirements that will be added to the implementation plan.
9. Review of updated implementation plan will be performed by the IMC and NISAC.
10. Execution of implementation plan will continue by the software developers (phase 2); incremental releases will include prototypes of the user interface.
11. A final review of deliverables by the IMC and NISAC will be performed to measure success.

Milestones:

- Use-case scenarios document (phase 1) developed (2011 Q4)
- Implementation plan (phase 1) developed (2011 Q4)
- Operational prototype (phase 1) deployed (2012 Q1)
- Review of operational prototype will be performed by IMC and NISAC (2012 Q2)
- Review of off-the-shelf workflow process management solution by CIO and ND (2013 Q1)
- Use-case scenarios document updated (phase 2) (2013 Q1)
- Implementation plan will be updated (phase 2) by ND (2013 Q2)
- Production deployment of the Workflow Manager performed (2014 Q1)

- Review of final deployment will be performed by IMC and NISAC (2014 Q2)

Evaluation: Criteria for evaluation will include delivery of all functionality as described by the requirements and implementation plan on the scheduled date(s). Additional criteria will include the validation of derived data products (Level-3) as output from the Workflow Manager.

Benefits: The Workflow Manager will benefit Site Information Managers and scientists/researchers by providing a framework for the development, storage, and automated execution of computational tasks against Level-1 and Level-2 data. This work will also benefit communities that are engaged in workflow integration into production data management systems.

Feedback: Direct feedback will occur from site scientists/researchers to the ND for evaluating the efficacy and efficiency of the workflow integration process, including the steps required for loading a workflow and monitoring its activity. Indirect feedback will be obtained by reviewing system reports. The system will be reviewed by NISAC during annual reviews. Communities that are involved in similar workflow manager integration applications, such as the National Center for Supercomputer Applications, may provide additional feedback on efficacy of the Workflow Manager.

3.2.3.10.3 Metadata Factory (Provenance Tracking)

Description: The Metadata Factory will be responsible for generating metadata using the Ecological Metadata Language (EML) standard format, including provenance information, for all Level-1 and Level-2 data found in the Data Cache and, more importantly, all derived data products (Level-3) that are produced by the Workflow Manager. Each EML document produced by the Metadata Factory will be harvested into the Data Catalog, thereby making it (and associated data) available for future discovery.

Provenance metadata for each derived data product will include one or more references to metadata documents (also as EML found in the Data Catalog) that describe the original site-based data set(s) (Level-1 or Level-2) used to generate Level-3 derived data products. A natural language description of the program used to generate the derived product will be included in the metadata, as well as the source code of any script or executable code used to generate the derived data, including configuration and steps that are part of an external workflow package. The Metadata Factory will generate metadata sufficient to recreate a derived data product based on its Level-1 or Level-2 inputs and the computational tasks used in the process.

The metadata structure that will be used to store provenance information is the “methods” sub-tree of EML. This sub-tree will contain core elements of the original EML that describes the Level-0 data and Level-1 and/or the Level-2 data that are used to produce any derived data product, including a direct reference to these documents within Data Catalog – we refer to this reference as a “metadata chain”. The “methods” sub-tree will also contain available source code and/or a description of the executable used in the workflow sequence.

Outcomes: Completion of the Metadata Factory will result in ongoing and automated metadata documentation of all Level-1 or Level-2 data in the Data Cache and derived data products (Level-3), including provenance information. The Metadata Factory will enable researchers to determine the data and workflow versions that lead to any derived data product.

Operational Steps:

1. The ND will compose detailed use-case scenarios by closely reviewing the operations of the Workflow Manager and assessing the steps required by the Metadata Factory; these use-case scenarios will be reviewed by the IMC and NISAC.
2. The ND will analyze use-case scenarios and will formulate functional requirements and an implementation plan for generating metadata as EML for Level-3 data; the implementation plan will be reviewed by the IMC and NISAC for completeness and correctness. The implementation plan will utilize one or more of the current NIS Data Modules (e.g., ClimDB /HydroDB) as a baseline (phase 1).
3. Execution of the implementation plan will be performed by the software developers to demonstrate an operational prototype.
4. Review of operational prototype deliverables will be performed by the IMC and NISAC.
5. Review of use-cases and functional requirements, including new use-cases for integrating more complex metadata associated with workflow applications (e.g., Kepler and/or Taverna) (phase 2), will be performed by the CIO, ND and NISAC during second year review.
6. ND will generate new/additional functional requirements that will be added to the implementation plan.
7. Execution of implementation plan will continue by the software developers (phase 2).
8. A final review of deliverables by the IMC and NISAC will be performed to measure success.

Milestones:

- Use-case scenarios document (phase 1) developed (2011 Q4)
- Implementation plan (phase 1) developed (2011 Q4)
- Operational prototype (phase 1) deployed (2012 Q1)
- Review of operational prototype performed by IMC and NISAC (2012 Q2)
- Use-case scenarios document updated (phase 2) (2012 Q2)
- Implementation plan updated (phase 2) (2012 Q1)
- Production deployment of the Metadata Factory (2013 Q2)
- Review of final deployment will be performed by IMC and NISAC (2013 Q3)

Evaluation: Criteria for evaluation will include delivery of all functionality as described by the requirements and implementation plan on the scheduled date(s). Additional criteria will include the validation of generated metadata as output from the Metadata Factory.

Benefits: A key issue facing scientific data delivery systems today is the lack of effective traceability from derived products to source data. Scientific data repositories (e.g., DataONE) will benefit from a standard approach to provenance tracking that is metadata-based and system independent. The broader scientific community will benefit from the metadata generated by the Metadata Factory by 1) making such derived data discoverable and 2) having adequate

provenance information to be able to recreate the derived data. In addition, communities that require provenance tracking may utilize techniques pioneered by the Metadata Factory in their own projects.

Feedback: Direct feedback from the user community to the ND for evaluating the correctness and efficacy of the metadata which describes derived data products. Indirect feedback will be generated based on the number of discoveries and access events to derived data that occur through the Data Catalog.

3.2.3.10.4 NIS Data Portal

Description: The NIS Data Portal (herein, Data Portal) will be the LTER community's web presence for discovering, accessing, and exploring LTER site-based and derived data products. The Data Portal will provide interfaces for user authentication, data discovery, data access (download), data exploration, personal data management (e.g., saving search results), and user feedback. In addition, the Data Portal will be customizable by the user by setting preferences for the "look and feel" of their browser view of the Data Portal and criteria for data discovery and access. The Data Portal will set the standard for key "portlets" that may be reused by other web-based portals (in support of the LTER community "skin" adoption of the current LTER Data Catalog web presence). The Data Portal will not be tightly coupled to any other PASTA framework component.

A typical use-case scenario will begin with a user authenticating to the system by entering their community credentials into the authentication interface. Once authenticated, the user may be greeted with news of another month of data being added to a data set that they are watching. The user may then navigate to the catalog webpage that contains hierarchical lists separated by different thematic or geographic categories (e.g., Climate-Temperature-Surface Temperature or North America-United States-California). Following the catalog link in the geographic category to California, the user may search for all LTER data sets that originate in the state of California. Finding a particular tabular data set of sea-level height, the user may plot this data set against other data sets that were saved in their personal data management space. Being satisfied with the quality and usability of the data set, the user may then download the data set and its associated metadata to their local work station for further analysis.

Outcomes: Completion of the Data Portal will secure a web presence for LTER site-based and derived data products, and provide a single point of access for LTER scientists/researchers and the broader community to such data. The Data Portal will become a local research home page for many scientists/researchers who regularly access ecological data of the LTER Network.

Operational Steps:

1. The ND will compose detailed use-case scenarios through a series of concept working groups involving SIMs, site scientists/researchers, IMC, and NISAC, that characterize the Data Portal services.
2. The ND will analyze use-case scenarios and will formulate functional requirements and an implementation plan (phase 1) for creating the Data Portal web application.

3. The implementation plan will be reviewed by the IMC and NISAC for completeness and correctness. The implementation plan will build upon past experience learned from the current LTER Data Catalog and the EcoTrends web portal.
4. Execution of the implementation plan will be performed by the software developers to demonstrate an operational prototype; incremental releases will include prototypes of the user interface.
5. Review of operational prototype Data Portal will be performed by the IMC and NISAC.
6. Review of use-cases and functional requirements, including new use-cases for integrating more complex features (e.g., data exploration and advanced discovery) (phase 2), will be performed by the NISAC during second year review.
7. ND will generate new/additional functional requirements that will be added to the implementation plan.
8. Execution of the implementation plan will continue by the software developers (phase 2); incremental releases will include prototypes of the user interface.
9. A review of deliverables by the NISAC and site scientists/researchers will be performed to measure success.

Milestones:

- Concept working group (2013 Q4)
- Use-case scenarios document (phase 1) developed (2011 Q1)
- Implementation plan (phase 1) developed (2011 Q1)
- Operational prototype (phase 1) deployed (2011 Q2)
- Review of operational prototype performed by NISAC and site scientists/researchers (2011 Q3)
- Use-case scenarios document updated (phase 2) (2013 Q4)
- Implementation plan updated (phase 2) (2013 Q4)
- Production deployment of the Data Portal (phase 2) (2014 Q1)
- Review of final deployment performed by NISAC and site scientists/researchers (2014 Q2)

Evaluation: Criteria for evaluation of the Data Portal will include delivery of all functionality as described by the requirements and implementation plan on the scheduled date(s). Additional criteria will include meeting web accessibility standards as specified by the Web Content Accessibility Guidelines.

Benefits: The Data Portal will provide a single point of access to LTER site-based and derived data for LTER scientists/researchers and the broader community. Users will be able to discover such data through multiple interfaces and download the data and associated metadata to their local workstation. In addition, the Data Portal will serve as a local home page to many scientists/researchers who regularly access LTER data; the Data Portal will support many personalized features related to managing discovered data sets.

Feedback: Direct feedback from the scientist/researcher user community to the ND for evaluating the efficacy of the Data Portal, including accuracy of searching for data sets and ease of managing data sets. The Data Portal will include a “feedback” interface for ad-hoc interaction.

3.2.3.10.5 Discovery/Access Application Programming Interface

Description: The Discovery/Access Application Programming Interface (API) will be the standard interface by which external web-based applications will interact with the LTER Network Information System and, more specifically, PASTA framework services. The Discovery/Access API will provide a web-services interface that follows a Service Oriented Architecture (SOA) design, thereby supporting a neutral programming layer that may be used by multiple applications regardless of their hardware and software preference. The Discovery/Access API will support a core set of services to provide discovery and accessibility to LTER site-based (Level-1 and Level-2) and derived (Level-3) data products. In addition to the core services, the API will also support user authentication for single sign-on through the Identity Management Services, system information (including, state of health), site-developed applications and tools (e.g., Controlled Vocabulary and Unit Registry), and Project Specific Data Services (section 3.2.3.11.4) (e.g., CUAHSI Hydrologic Information System (HIS), DataONE Member Node, National Ecological Observatory Network (NEON) functions). The Discovery/Access API will be an abstraction layer that removes (or minimizes, at the very least) tightly-coupled dependencies between underlying PASTA framework components from external applications; as such, framework components may evolve as necessary without adversely affecting applications that are built on top of the Discovery/Access API.

An example use-case scenario would be that of the NIS LTER Data Portal. The Data Portal would configure its discovery engine to query the Discovery/Access API discovery service using a standard query syntax (*in lieu* of connecting directly to the Data Catalog). The discovery service would then translate the standard query syntax into the specific syntax used by the Data Catalog. Upon completing the query, a Data Catalog formatted query result would be returned to the Discovery/Access API discovery service. The discovery service would then translate the Data Catalog query result format into a standard format, and return it to the Data Portal for final processing.

Outcomes: Completion of the Discovery/Access API module will provide a standardized application programming interface by which external web-based applications will discover and access LTER site-based (Level-1 and Level-2) and derived (Level-3) data products. The Discovery/Access API will support industry standard web-services using SOA principles to provide access for external applications. The Discovery/Access API abstraction will insulate external applications from evolving PASTA framework components.

Operational Steps:

1. The ND will compose detailed use-case scenarios through targeted requests and/or formal working groups of stakeholders, including Site Information Managers (and programmers/developers), community programmers/developers, and NISAC, that characterize the Discovery/Access API; additional review of existing systems (e.g., Virtual Data Center) will inform use-case scenario development.
2. The ND will analyze use-case scenarios and will formulate functional requirements and an implementation plan (phase 1) for creating the Discovery/Access API and associated

services. The implementation plan will build upon past experience learned from similar projects (e.g., EarthGrid).
3. The implementation plan will be reviewed by the CIO, ND, and NISAC for completeness and correctness.
4. Execution of the implementation plan will be performed by the software developers to demonstrate an operational prototype.
5. Review of operational prototype deliverables will be performed by the IMC and NISAC.
6. ND will publish Discovery/Access API specification to the broader community.
7. Review of use-cases and functional requirements, including new use-cases generated during phase 1 from community input for integrating more complex features (e.g., user authentication, system state of health, or external project support) (phase 2), will be performed by the CIO, ND and NISAC during second year review.
8. ND will generate additional functional requirements from new use-case scenarios that will be added to the implementation plan.
9. Execution of implementation plan will continue by the software developers (phase 2).
10. A final review of deliverables by the IMC and NISAC will be performed to measure success.

Milestones:

- Use-case scenarios document (phase 1) developed (2010 Q2)
- Implementation plan (phase 1) developed (2010 Q2)
- Operational prototype (phase 1) deployed (2011 Q1)
- Discovery/Access API specification (phase 1) published (2011 Q1))
- Review of operational prototype performed by IMC and NISAC (2011 Q2)
- Use-case scenarios document updated (phase 2) (2012 Q2)
- Implementation plan updated (phase 2) (2012 Q2)
- Production deployment of the Discovery/Access API (phase 2) (2013 Q1)
- Discovery/Access API specification (phase 2) published (2013 Q1))
- Review of final deployment will be performed by IMC and NISAC (2013 Q2)

Evaluation: Criteria for evaluation of the Discovery/Access API will include delivery of all functionality as described by the requirements and implementation plan on the scheduled date(s). Additional criteria will include service up-time as defined by non-functional requirements.

Benefits: Separation of underlying PASTA framework components from external applications through a published interface will benefit LNO Analysts/Programmers, Site Information Managers, and community developers/programmers by eliminating direct dependencies that may otherwise affect such external applications. Community developers will benefit from having programmatic access to LTER data and metadata through an SOA supported interface for development of external web-based applications.

Feedback: Direct feedback from the developer community to the ND for evaluating the efficacy of the Discovery/Access API interface specifications and services; indirect feedback will include measuring the number and correctness of service calls performed in a given period.

3.2.3.10.6 Identity Management Services

Description: The Identity Management Services module will provide support for user single sign-on authentication to access the LTER NIS services. Single sign-on implies that a user may authenticate at one, of perhaps many, authentication interface points associated with the LTER Network (including the broader community of service providers) and would not have to re-issue the same authentication credentials at each service that requires user authentication; the credential or token representing the user provides short-lived authentication on behalf of the user to services requiring authentication. Because of the complexity and widespread need of such services, the design and implementation of the Identity Management Services will be developed collaboratively through multiple working groups within the community, including the CILogon project at the University of Illinois Urbana Champaign (http://www.cilogon.org) and DataONE (http://dataone.org) project at the University of New Mexico.

One use-case scenario for Identity Management Services would include the authentication of an LTER Site Information Manager, who may have dual roles when interacting with the LTER NIS. In one role, the Site Information Manager may be a user of the Data Manager interface for scheduling new metadata harvesting; and in a second role, the Site Information Manager may be a user of the Data Portal when searching for LTER data products. To schedule a new harvest, the Site Information Manager would authenticate with the Data Manager authentication service by entering credential information. After configuring the new harvest, and while still logged into the NIS web-based interface, the Site Information Manager loads the Data Portal interface into the local web browser. Using a session token, the Data Portal recognizes the Site Information Manager and performs an automatic authentication using the token, at which point the Information Manager has full access to the Data Portal.

Outcomes: The Identity Management Services module will support session-based single sign-on of users for authentication to access services of the LTER NIS.

Operational Steps:

1. The ND will compose detailed use-case scenarios through targeted requests of stakeholders, including Site Information Managers (programmers/developers), site scientists/researchers, community programmers/developers (e.g., CILogon and DataONE), and NISAC, that characterize the Identity Management Services.
2. The ND will analyze use-case scenarios and will formulate functional requirements of the Identity Management Services.
3. The ND will work collaboratively with leaders of community projects working toward similar single sign-on goals and jointly develop an implementation plan.
4. The implementation plan will be reviewed by the IMC and NISAC for completeness and correctness.
5. Execution of the implementation plan will be performed by the software developers to demonstrate an operational prototype.
6. A final review of deliverables by the CIO, ND, and NISAC will be performed to measure success.

Milestones:

- Use-case scenarios document (phase 1) developed (2010 Q4)
- Implementation plan (phase 1) developed (2010 Q4)
- Operational prototype (phase 1) deployed (2010 Q4)
- Review of operational prototype performed by IMC and NISAC (2011 Q2)
- Collaborative implementation plan (phase 2) developed (2013 Q1)
- Production deployment of the Identity Management Service (2013 Q4)
- Review of final deployment performed (2014 Q1)

Evaluation: Evaluation will be performed by NISAC during annual reviews of LNO performance. Criteria for evaluation of the Identity Management Services will include delivery of all functionality as described by the requirements and implementation plan on the scheduled date(s). External projects (e.g., CILogon and DataONE) will be asked to evaluate deliverables for correctness and to ensure interoperability.

Benefits: The Identity Management Services will simplify access to NIS services and resources for all users, including the broader scientific community, by supporting a single sign-on authentication service, thereby eliminating the need for repeated authentication by users. The implementation of the Identity Management Services will act as a guide and model to community developers of other cyberinfrastructure projects requiring similar authentication infrastructure. For example, the DataONE (http://dataone.org) project at the University of New Mexico is now reviewing existing authentication protocols and procedures for implementing identity management. The Identity Management Services of the LTER NIS may demonstrate techniques that can be reused within DataONE.

Feedback: Direct feedback from the NIS user community to the ND will occur for evaluating the correctness of the Identity Management Services; indirect feedback will include measuring successful access to LTER NIS services and resources.

3.2.3.10.7 Persistent Identifier Services

Description: A problem that often occurs when digital objects are identified with absolute locations, such as a web URL to a data set, is that the direct link to the object "breaks" when the object is moved to a new location or the domain name of the web server is changed. To overcome the limitation of absolute identifiers, persistent identifiers use a relative identification scheme that maps the identifier to the current location of the object. As such, the user of a persistent identifier can be assured that the identifier will always resolve to the object, regardless of its physical location. The Persistent Identifier Services that will be adopted by the PASTA framework will provide persistent identification of (and resolvability to) digital objects (including metadata documents and data sets). In addition to object persistence, the Persistent Identifier Services will ensure that once an object is identified by a unique identifier, no other object within the operational domain may be identified with the same identifier.

Persistent and unique identification systems are now under review by many cyberinfrastructure projects, including the biodiversity Taxonomic Data Working Group project and the DataONE

(http://dataone.org) project at the University of New Mexico. The Persistent Identifier Services will utilize the method for assigning and resolving persistent identifiers that is adopted by the broader community, thereby maintaining interoperability between organizations. An alternative system will also be identified if the community does not adopt one method within the time-scope necessary for the NIS development.

The use of persistent identifiers will be optional for LTER Network data that resides in a site-based information system and is referenced through a site-generated EML document found in the Data Catalog. All metadata and data found in the PASTA framework, however, will be identified with a persistent identifier.

Outcomes: Completion of the Persistent Identifier Services component will ensure persistent identification and access to LTER data for users and services from within the PASTA framework and for external users and services. In addition, by adopting a community standard, interoperability between organizations will be maintained.

Operational Steps:

1. The ND will compose detailed use-case scenarios involving the use of persistent identifiers in the PASTA framework through the evaluation of community adopted solutions.
2. The ND will analyze use-case scenarios and will formulate functional requirements of the Persistent Identifier Services.
3. The ND will work collaboratively with leaders of community projects working toward similar persistent identifier goals and jointly develop an implementation plan.
4. The implementation plan will be reviewed by the NISAC for completeness and correctness.
5. Execution of the implementation plan will be performed by the software developers to demonstrate an operational prototype.
6. A final review of deliverables by the NISAC will be performed to measure success.

Milestones:

- Use-case scenarios document developed (2011 Q2)
- Collaborative implementation plan will be developed by ND and other community project leaders (2011 Q2)
- Review of implementation plan by NISAC (2011 Q2)
- Production deployment of the Persistent Identifier Services will be performed by ND and NAP3/NAP2 (2011 Q3)
- Review of final deployment will be performed by CIO, ND, and NISAC (2011 Q4)

Evaluation: Persistent Identifier Services will be evaluated based on the delivery of all functionality as described by the requirements and implementation plan on the scheduled date(s). This evaluation will be performed as part of the NISAC review of LNO performance.

Benefits: The Persistent Identifier Services will benefit all end users and services of the LTER NIS by ensuring that identifiers used by metadata and data objects will accurately obtain requested objects regardless of their physical location.

Feedback: Direct feedback from the NIS user community to the ND will occur for evaluating the accuracy and persistence of requested metadata and data objects as identified through the Persistent Identifier Services. Communities that are involved in similar persistent identified-based applications, such as the Global Biodiversity Information Facility, may provide additional feedback on efficacy of the Persistent Identifier Services.

3.2.3.10.8 System Monitor

Description: The System Monitor provides system state-of-health information for all modules/services. The System Monitor will consist of a number of component tests that evaluate the state of specific services and/or entire components; each test will be measured from a known performance base-line; therefore, changes in performance (either positive or negative) may be noted. The System Monitor will operate as a background process from which state-of-health alerts and reports may be continually generated. External applications will be able to access a subset of System Monitor functions through the Discovery/Access API (section 3.2.3.5).

A typical use-case scenario would be for the System Monitor to track performance of the Data Catalog at 10-minute intervals to determine overall responsiveness based on a specific query. The System Monitor would issue the query and measure the response time of the Data Catalog. If the response time increased above some threshold, the System Monitor would send an alert to the LTER NIS system administrators.

Outcomes: Completion of the System Monitor will ensure that the LTER NIS continues to operate a measured level of service. In addition, the System Monitor will provide near real-time alerts in the event of any service and/or module failure. Regular reports will be generated for evaluating the overall performance of new or modified services and/or components.

Operational Steps:

1. The ND will compose detailed use-case scenarios through targeted individuals, including NIS Analysts/Programmers and NISAC, which characterize the System Monitor.
2. The ND will analyze use-case scenarios and will formulate functional requirements.
3. The ND will evaluate existing system monitoring applications (e.g., Cacti, Hobbit, or Nagios). Based on such evaluations, an implementation plan will be developed for developing the System Monitor.
4. The implementation plan will be reviewed by the NISAC for completeness and correctness.
5. Execution of implementation plan will be performed by the software developers; incremental releases will include prototypes of the user interface.
6. A final review of deliverables by the NISAC will be performed to measure success.

Milestones:

- Use-case scenarios document developed (2013 Q4)
- Implementation plan developed (2013 Q4)
- Production deployment of the System Monitor (phase 2) (2014 Q1)

- Review of final deployment performed NISAC (2014 Q2)

Evaluation: The System Monitor development will be evaluated as part of the regular NISAC review of LNO performance. Criteria for evaluation of the System Monitor will include delivery of all functionality as described by the requirements and implementation plan on the scheduled date(s).

Benefits: The System Monitor will benefit all users of the LTER NIS by ensuring that the system is operating at an optimal level. In addition, the System Monitor will benefit the NIS system administrators by providing up-to-date and timely information on the state-of-health of PASTA framework services and/or components.

Feedback: Direct feedback from the System Monitor, in the form of alerts and reports to the ND, will occur for evaluating the overall performance of the PASTA framework services and/or components.

3.2.3.10.9 Audit Services (formerly, Data Access Server)

Description: The Audit Services will record all metadata/data access events and provide key services in compliance of the LTER Data Policy, including authentication, reporting, and notification. Specifically, the Audit Services will support the replacement of direct access URLs to all LTER metadata and data products within the PASTA framework with the use of a proxy URL that will route such requests through its logging, authentication, reporting, and notification services. The Audit Services will support the use of persistent identifiers for object identification and resolution as part of the proxy URL.

The Audit Services will also be available for use by LTER information managers to replace direct access URLs with proxy URLs for any network accessible location outside of the PASTA framework (e.g., site web pages), thus continuing support provided by the original Data Access Server.

A typical use-case scenario would begin with a user requesting access to a specific data set. The user will be required to authenticate by using their community credentials, unless previously authenticated at another interface point. Once authenticated, the access event will result in the generation of an audit record that documents the event and a pair of notifications (e.g., email, SMS text message, or instant message) that are sent to 1) the data set owner/manager alerting them of the access event and 2) the user who is accessing the data with suggested citation information. The data are then passed through the proxy server and back to the requesting user.

Outcomes: The Audit Services will support compliance to the LTER Data Policy without separate and additional effort on behalf of LTER Site Information Managers, and with the added benefit of data use statistics and reporting.

Operational Steps:

1. ND will review use-case scenarios and functional requirements for the Audit Services, including new use-cases for integrating standard persistent identifier protocols.
2. ND will generate an implementation plan based on functional requirements.
3. The implementation plan will be reviewed by IMC and NISAC for completeness and correctness.
4. Execution of implementation plan will be performed by the software developers; incremental releases will include prototypes of the user interface.
5. A final review of deliverables by the IMC and NISAC will be performed to measure success.

Milestones:

- Use-case scenarios document developed (2011 Q4)
- Implementation plan developed (2011 Q4)
- Production deployment of the Audit Services (2012 Q3)
- Review of final deployment performed by the IMC and NISAC (2012 Q4)

Evaluation: The Audit Services will be reviewed as part of the regular NISAC reviews of LNO and NIS performance. Criteria for evaluation of the Audit Services will include delivery of all functionality as described by the requirements and implementation plan on the scheduled date(s).

Benefits: The Audit Services will benefit the LTER scientists/researchers and their funding agency(s) by tracking, in a consistent manner, the access to LTER site-based and NIS derived metadata/data and provide standard use statistics. Site Information Managers will also benefit from the Audit Services since they will be able to use an LTER-wide application for complying with the LTER Data Policy and will not be required to develop a site-based solution.

Feedback: Direct feedback from the LTER scientists/researchers to the ND for evaluating notifications and their impact on user citation of LTER data use. Additional feedback will include response from Site Information Managers concerning ease-of-use of the Audit Services service interface and reporting. Indirect feedback on the Audit Services will include data access statistics.

3.2.3.11 Integration of Network databases into the LTER NIS.

The LTER Network supports a number of existing synthesis database applications, including EcoTrends and ClimDB/HydroDB, which are maintained as isolated projects and without direct coordination between one another. As part of the development process of the LTER NIS, these separate database applications will be integrated into the PASTA framework, thereby taking advantage of the structure and coordination of PASTA framework services and/or components. In addition, derived products from the ClimDB/HydroDB NIS Data Module will be exposed through a Consortium of Universities for the Advancement of Hydrologic Science, Incorporated (CUAHSI) Hydrologic Information System (HIS) web-service for subsequent discovery and access to users from CUAHSI. The following sections describe in detail activities of the integration tasks.

3.2.3.11.1 Integration path for EcoTrends and ClimDB/HydroDB

Description: The existing synthesis database applications (EcoTrends and ClimDB/HydroDB) of the LTER Network all follow similar processing steps: 1) collect site data and organize it into a standard format, 2) store restructured data into a common relational database schema, and 3) provide discovery and access tools via a web-based interface, including basic analytical applications. Each one of these applications, however, requires separate management and maintenance resources (this is especially true between EcoTrends and ClimDB/HydroDB because of their very different architectures). In some cases, applications require additional resources at the site-level to reformat data into a standard structure. As such, integrating these independent database applications into the LTER NIS will provide scalability and ensure data integrity by utilizing a common application framework (PASTA) that supports automated data harvesting and quality assurance checking. New features of the PASTA framework will become directly available to the integrated database applications without additional effort at the sites.

This task will focus on development of an integration trajectory for each of the existing database applications (EcoTrends and ClimDB/HydroDB). It will require understanding the underlying processes of the current systems, including stakeholder (both data providers and consumers) requirements and dependencies, data transformation issues, and the methods employed for discovery, access, and analysis. The process of integration will require that end products be similar, if not identical, as provided by the original applications, while at the same time introduce efficiencies brought about by the common PASTA framework.

Outcomes: Completion of the integration path for the existing synthesis database applications (EcoTrends, ClimDB/HydroDB) will provide a structured and well managed step-wise sequence for integrating existing synthesis database applications into the common PASTA framework.

Operational Steps:

1. The ND will perform a detailed analysis of the operations and processes of the existing synthesis database applications, looking particularly at data origination issues, transformation issues, and end-user tools (discovery, access, and analysis); this analysis will include direct communications with SIMs, site scientists/researchers, NISAC, and the broader community that utilizes the current applications.
2. The ND will compose detailed use-case scenarios based on the analysis in 1 above.
3. The ND will analyze use-case scenarios and will formulate functional requirements and an implementation plan for performing the integration
4. The implementation plan will be reviewed by the CIO, the IMC, and NISAC for completeness and correctness.
5. Execution of the implementation plan will be performed by the NIM, NAP3, and NAP2.
6. A final review of deliverables by NISAC will be performed to measure success.

Milestones:

- Use-case scenarios document developed (2010 Q4)
- Implementation plan developed (2010 Q4)

- Production deployment of the database application performed (2011 Q3)
- Review of final deployment will be performed by NISAC (2011 Q4)

Evaluation: Integration of Network databases will be reviewed as part of the regular annual review of NIS carried out by NISAC. Criteria for evaluation of the integration path for existing synthesis database applications will include: 1) comparison and validation of end products (derived data) to that of the data produced by the original application, 2) availability of data products as compared to the original application, and 3) ease of data harvesting at the site.

Benefits: All end users of the existing applications will benefit by having the applications integrated into a common framework that supports automated metadata and data harvesting and quality assurance checking. Site Information Managers will benefit by having additional data reformatting requirements removed from their responsibility.

Feedback: Direct feedback from end users of the existing synthesis database applications to the ND will occur for evaluating accuracy and efficiencies of the integrated applications. This feedback will be gathered as part of the biennial survey process.

3.2.3.11.2 Online analysis and processing of LTER NIS data

Description: Site-based data (Level 1 and Level 2) and derived data (Level 3) that will be accessible from the LTER NIS through the Discovery/Access API (section 3.2.3.10.5) will have great value. Tools for discovering and exploiting such data will also benefit the community. A certain number of these tools exist today, but many of them do not meet the full needs of the user community. This task will identify, plan, and develop online analysis and processing tools for data within the LTER NIS through a comprehensive RFC process that will solicit input from the LTER and the broader ecological community. Such tools may include enhanced data discovery through knowledge-based and semantic approaches, sub-setting and integration of data, and descriptive analysis of data.

Outcomes: Completion of this task will result in a set of online analysis and processing tools that may be applied to data within the LTER NIS. These tools will be at the request of the LTER Network and broader ecological community as specified in a formal RFC process.

Operational Steps:

1. The ND will compose a set of RFC documents for specific online analysis and processing tools, including enhanced data discovery, data sub-setting and integration, and descriptive analysis. These RFC documents, as informed by the research working groups, will be sent to stakeholders, including LTER scientists/researchers and other end users of the ecological community, for soliciting inputs on the benefits, design, and function of such tools.
2. The ND will perform an analysis of RFC responses and generate use-case scenarios.
3. The ND will analyze use-case scenarios and will formulate functional requirements and an implementation plan for each tool; the implementation plan will be reviewed by the IMC and NISAC for completeness and correctness.
4. Execution of the implementation plan will be performed by the software developers.

5. A final review of delivered tools by the NISAC will be performed to measure success.

Milestones:

- RFC documents developed (2013 Q1) for each tool
- Use-case scenarios developed for each tool (2013 Q1)
- Implementation plan developed (2013 Q1) for each tool
- Production deployment of each tool performed (2013 Q4)
- Review of final deployment of each tool performed by (2014 Q1)

Evaluation: The on time and on target completion of these tools will be reviewed as part of the regular review of the NIS carried out annually by NISAC. Criteria for evaluation of each online analysis and processing tool will include delivery of all functionality as described by the requirements and implementation plan on the scheduled date(s). Additional criteria will include evaluation of tool's end product for correctness and accuracy. The usefulness of these tools to the scientific community will be assessed via the biennial survey of the LTER Network.

Benefits: More flexible and interoperable access to LTER data provided by the NIS will stimulate and increase LTER science, improve the quality of LTER data and metadata, foster the use of data for synthetic projects, and thereby enhance both national and global research. Online analysis and processing tools for LTER site-based data (Level 1 and Level 2) and derived data (Level 3) found within the LTER NIS will benefit LTER scientists/researchers and others involved in research from the broader ecological community by providing enhanced data discovery through knowledge-base and semantic tools, data sub-setting and integration tools, and descriptive data analysis tools.

Feedback: Feedback will be requested from LTER scientists/researchers and end users from the broader community of the efficacy of the online analysis and processing tool. Additional feedback will come from the metered use of such tools.

3.2.3.11.3 Site-based data integration into LTER NIS Data Modules

Description: The existing network-supporting database applications (site characteristics, personnel, bibliography) of the LTER Network will be integrated into the LTER NIS. The existing network-supporting databases will be redesigned and integrated as NIS Data Modules and their content will be made available through the Discovery/Access API. These network database applications currently follow site practices for generating NIS specific data, which are then imported into existing data modules for use by LNO, Site Information Managers and LTER scientists/researchers. A goal of the LTER NIS is to support the integration of these applications into the NIS without increased effort. As such, the Network Information Manager will support Site Information Managers to transition their data preparation effort from a localized process to an automated process as part of the PASTA framework.

Outcomes: The network-supporting databases will be accessed directly through the Discovery/Access API (as web-services). The Network Information Manager will assist in the transition of site-based data preparation from local SIMs to that of an automated PASTA

framework process, increasing the overall data quality and reducing the overall effort of the Site Information Managers.

Operational Steps:

1. The associated product oriented working group will meet to determine needs.
2. The ND will compose detailed use-case scenarios based on output from the product oriented working group.
3. The ND will analyze use-case scenarios and will formulate functional requirements.
4. The implementation plan will be reviewed by the IMC and NISAC for completeness and correctness.
5. Execution of implementation plan will be performed by the software developers and the Network Information Manager.
6. A final review of deliverables by the NISAC will be performed to measure success.

Milestones:

- Working group product(s) delivered (2011 Q2)
- Use-case scenarios documents developed for network-supporting databases (2011 Q3)
- Implementation plan developed (2011 Q4)
- Production deployment of the integration process performed (2012 Q3)
- Review of final deployment will be performed by NISAC (2012 Q4)

Evaluation: This effort will be evaluated as part of the regular review of NIS activities by NISAC. Criteria for evaluation of site-based data integration of databases and applications will include delivery of all functionality as described by the requirements and implementation plan on the scheduled date(s). Additional criteria will include the validation and verification of correct data harvesting into the PASTA framework and availability of network-supporting data through the Discovery/Access API.

Benefits: Completion of site-base data integration will benefit SIMs by reducing overall effort applied to manually preparing data for ingestion in the current NIS Data Modules and make accessing network-supporting data more efficient.

Feedback: Feedback will be gathered from Site Information Managers regarding the efficacy of automated processing and preparing of site-based data and accessing network-supporting data. Additional feedback will include automated data quality and validation testing as part of the Data Loading module (section 3.2.3.10.1).

3.2.3.11.4 Project Specific Data Services

Description: Project Specific Data Services will be a set of custom web-services of the Discovery/Access API that may be developed in support of projects that are synergistic or closely related to the mission of the LTER Network Information System. Examples of projects include the Consortium of Universities for the Advancement of Hydrologic Science, Incorporated (CUAHSI) Hydrologic Information System (HIS) web-services, DataONE

(Observation Network for Earth) Member Node services, and/or the National Ecological Observatory Network (NEON) data interface. Any services development will be based on available support and required effort and its importance as decided by the LTER Executive Board. The CUAHSI HIS web-service has been previously identified by NISAC as a potential candidate as a Project Specific Service. The following sections assume acceptance of the CUAHSI HIS web-services as a Project Specific Data Service.

Outcomes: Completion of the CUAHSI HIS web-services integration into the Discovery/Access API will expose ClimDB and HydroDB data from their respective NIS Data Modules to CUAHSI and broader communities who are performing hydrologic research.

Operational Steps:

1. The ND will review and analyze the set of CUAHSI HIS web-services.
2. The ND will formulate an implementation plan for the CUAHSI HIS web-services; the implementation plan will be reviewed by the CIO, ND, NISAC, and representatives of the CUAHSI HIS technical team for completeness and correctness. The implementation plan will build upon past experience learned from the current LTER Data Access Server.
3. Execution of the implementation plan will be performed by the software developers.
4. A final review of deliverables by the CIO, ND, NISAC, and representatives of the CUAHSI HIS technical team will be performed to measure success.

Milestones:

- Implementation plan developed by ND (2013 Q4)
- Production deployment of the CUAHSI HIS web-services performed (2014 Q1)
- Review of final deployment performed by NISAC, and representatives of the CUAHSI HIS technical team (2014 Q2)

Evaluation: This activity will be evaluated as part of the regular annual review of NIS activities by NISAC. Criteria for evaluation of the CUAHSI HIS web-services will include delivery of all functionality as described by the requirements and implementation plan on the scheduled date(s). Additional criteria will include testing and validating output from the web-services by both the ClimDB and HydroDB data providers and CUAHSI scientists/researchers.

Benefits: The integration of CUAHSI HIS web-services into the Discovery/Access API will expose LTER ClimDB/HydroDB data through a public and standardized interface, thereby benefiting CUAHSI and broader communities who are performing hydrologic research.

Feedback: The ND will receive direct feedback from both scientists/researchers from CUAHSI and broader communities who are performing hydrologic research on the efficacy, correctness, and ease-of-use of the data from the LTER ClimDB/HydroDB. Additional feedback will include measuring the number of access events to ClimDB/HydroDB through the Discovery/Access API CUAHSI HIS web-services.

3.2.4 IT, Data base, and Web Consulting. Provide IT, database, and web consulting to LTER sites and synthesis working groups.

Background: Resources available to LTER sites to participate in Network data standards and applications are limited. The LNO has successfully gained site participation in Network standards and synthesis activities through the use of individualized consultation and 'hands on' technical help for issues ranging from web site design to EML implementation. LNO will now fill the majority of these generalized needs with web-based and targeted in-house training. However, one-on-one support provides an opportunity to help sites over minor hurdles that may be delaying their participation. This activity will develop a process to identify and prioritize these opportunities. The budget for this activity is $76,920 over five years to support 0.20 FTE of personnel time.

3.2.4.1 Technical support – The CIO will evaluate and prioritize needs and requests to provide a limited amount of information technology, database, and web application consulting services from existing labor pools to sites and working groups.

Outcomes: Individualized technical help provided to LTER sites at critical junctures increases their capability and desire to participate in Network science and education goals defined in the Decadal Plan. Synthesis working groups will leverage their analytical work by producing published data products.

Operational Steps:

1. CIO develops a process for evaluation in coordination with ED and ND.
2. CIO vets process with NISAC and IMEXEC.
3. CIO identifies sites/groups in need of assistance based on technical support requests or by other means during establishment of annual performance criteria.
4. CIO and ED evaluate requests for impending calendar year and rank based on the likelihood of increasing site participation in Network standards.
5. CIO develops a scope of work and a schedule based on resource availability.

Milestones:

- CIO develops and vets process for evaluating requests by Q4 2010.
- Sites and groups are identified during establishment of annual LNO performance criteria.
- Scope of work established and scheduled.
- Work completed.

Evaluation: Performance will be evaluated as part of LNO annual performance review. Participants will be asked to evaluate the effectiveness of the help.

Benefits: The broader scientific community will benefit from a greater number of site and synthesis data products being available online.

Feedback: Feedback will be requested as part of regular LNO surveys of sites.

3.2.4.2 Site visits – The CIO and selected LNO staff may travel to LTER sites during the performance period to consult and evaluate site-specific issues related to participation in the Network Information System and meeting Network standards for information management. The intent is to be able to visit all sites twice during the performance period if needed with the understanding that there are economies of traveling to multiple sites, opportunities when traveling on other business, and not all sites will desire the visitation.

Outcomes: Face-to-face interactions with site personnel are the most productive way to understand site needs and plan optimally crafted technical solutions. The outcome of LNO site visits will be improvements in information management plans and techniques across the LTER Network that increase the overall quality of Network data products.

Operational Steps:

1. The CIO identifies sites wanting assistance in preparing for review, working on renewal proposals, or improving their information system to meet Network standards.
2. The CIO fields ad hoc requests during the performance period and evaluates the requests based on priority and resource availability.
3. LNO personnel visit the site to develop an understanding of the site's requirements and help evaluate the level of effort required to make needed improvements.
4. LNO personnel provide a written evaluation within 2 weeks following the visits.

Milestones:

- Sites are identified during establishment of annual LNO performance criteria.
- CIO and ED evaluate requests.
- LNO visit scheduled and completed for each site.
- LNO written evaluation delivered to each site.

Evaluation: Performance will be evaluated as part of LNO annual performance review.

Benefits: Some sites will benefit from having additional expertise and perspective on specific components of their site IM and CI. The LTER Network will benefit by LNO facilitating broader site participation and accessibility. The LNO external CI partners and broader scientific community will benefit from having better overall access to quality data products.

Feedback: Feedback will be requested as part of regular LNO surveys of sites.

3.3 Core Services

The LNO provides many services to LTER scientists, sites, and the Network, most of which are conducted with funding from a separate Cooperative Agreement. One of these services, the acquisition of data, hardware, and software, directly contributes to LTER cyberinfrastructure, and thus was included in the proposal upon which this operational plan is based.

3.3.1 Acquisition of Data, Hardware, and Software. Facilitate the acquisition of commercial and public data, hardware, and software products for site-based and Network synthesis activities.

Background: Communication and data sharing are facilitated by common hardware and software, which can often be acquired at reduced cost through joint purchases. On several occasions in the past, the LNO has served as the Network's agent in acquiring common cyberinfrastructure, and it will act in this capacity more frequently in the future to fulfill the goals of the Decadal Plan. The Network research and education goals developed in the Decadal Plan emphasize synthesis using comparable datasets over broad spatial scales. To address issues of land use, climate change, and invasive species, LTER scientists will require time series of remotely-sensed data. Specific examples of such data sets include: satellite data collected by mission oriented agencies such as NASA, NOAA, or DOD; field data collected through networks of monitoring stations; and photographic images. The LNO facilitates collection of such data by interacting with LTER sites and the data providers to ensure accurate and complete geographic coverage of LTER sites.

Description of activity: This activity has a budget of $164,462 over five years that supports 0.20 FTE of personnel effort so the Network can respond opportunistically to improve hardware, software, and data capabilities. Because these opportunities are not known at this time, the plan describes the modes of response without identifying specific hardware, software, or data acquisitions.

Hardware/software: The LNO works to fulfill needs for cyberinfrastructure and communication tools as identified by sites and the goals of the Decadal Plan. By evaluating the utility and cost-effectiveness of different alternatives, the LNO assists the Network in making decisions about hardware and software acquisitions. Through proposals and partnerships, the LNO obtains resources to acquire common cyberinfrastructure and communication tools. Network-level acquisition and implementation includes communication and coordination with site personnel for site-specific needs, negotiations of costs for hardware and software, and distribution of resources across the LTER Network.

Data: The LNO works with private vendors and with partners at UNM centers, NASA, USGS, DOD, and other agencies to identify and acquire data that contribute to LTER research and education goals. The Director for Synthesis Support oversees the collection of information on site needs and the communication of this information to the data provider. The Chief Information Officer is responsible for ensuring delivery of the acquired data to LTER users through a web portal. When required, the LNO provides or requests funding for data acquisition.

Outcomes: Central acquisition of cyberinfrastructure, communication tools, and data reduces cost and improves interoperability while reducing redundant effort at each LTER site, as measured by the following outcomes:

- The consistent acquisition of standard and new data products across the LTER Network provides opportunities for cross-site synthesis.

- Direct collaborations with private companies and government agencies reduces redundancy of effort by individual LTER sites and enhances access to a variety of data including remote sensing data products.
- Centralized common acquisition of hardware, software, and data translates into savings in time and money by individual LTER sites.

Operational steps:

1. By the end of 2009, the LNO will complete a draft operational plan to update LTER Network spatial data holdings in conjunction with the LTER Spatial Data and Analysis Committee directives and Executive Board review.
2. On an annual basis thereafter, the LNO will engage in communication and coordination with LTER committees such as the Information Management Committee and the Spatial Data Analysis Committee as well as site personnel to identify needed new research resources.
3. The LNO will consult with the EB to prioritize requests and to identify sources of funds to acquire, hardware, software, or data.
4. When necessary, the LNO will prepare proposals and supplement requests for acquisitions.
5. The LNO will provide central coordination for acquisition of various hardware, software, and data.
6. The LNO will engage in negotiations for best costs with vendors of hardware and software.
7. The LNO will purchase and distribute new hardware, software, and data.

Pursuant to the internal LNO Business and procedures manual the processes for purchase include the following steps:

1. Initial request for the acquisition of hardware/software or data,
2. Check for availability of funds,
3. Identify vendor(s) and negotiate for specific costs,
4. Comply with approved and audited internal UNM purchasing processes: purchase requisition, purchase order, tracking and receipt of goods, inventory placement and tracking,
5. Implementation of hardware/software or data, and
6. Maintenance of hardware/software or data.

Milestones:

- By the end of the 2009, the LNO will complete a draft operational plan to update LTER Network spatial data holdings.
- The LNO will perform an annual evaluation of Network needs for hardware, software, and data.

Evaluation: Annual prioritized requests for hardware, software, and data will be incorporated into performance criteria defined by the EB. The annual report by the LNO to NSF and the EB will describe progress under each milestone of the performance criteria to inform the annual review process by the EB.

Benefits: Overall, this activity will enhance access to a variety of hardware, software, and data including remote sensing data products. In the long term (7-10 years), this will increase levels of network coordination and standardization through the consistent use of hardware and software. Ultimately, there will be an enhanced ease of cross-site data exchange and use, particularly for LTER scientists and educators and their external partners.

Feedback: Feedback will be gathered from the biennial survey and directly from users of the acquired data, hardware, and software.

Hardware/software: The LTER community will provide feedback through the survey administered every other year on the types of hardware/software requested, received, and used, along with support.

Data: As databases are accessed within the website, numbers of users will be monitored. Feedback may be indirectly parsed out of the volume and diversity of data requested by the number of users or through direct feedback on the website from a form that may be filled out in response to users' satisfaction.

3.4 Development and Outreach

3.4.1 Communication and Outreach. Effectively communicate information about the LTER Network—purpose, activities, and achievements—through an array of proven, high-visibility mechanisms.

Background: The LNO budget allocates $449,833 to support 1.00 FTE of personnel time in years 3 through 5 as well as participant support costs, publication costs, and travel to disseminate information about the accomplishments of the LTER Network. Efforts to communicate information about the LTER program to its various constituencies have grown over the years and adopted new technologies as they have come in to common use. The LNO has gradually shifted from an early reliance on printed documents to a mix of print, multimedia, and web-based communication approaches. In the absence of clearly-defined Network goals for communication and outreach, the LNO has used simple and cost-effective media approaches to reach perceived key constituencies. Increases in the size and scope of the Network and interest in the Network from new audiences have complicated communication and outreach efforts and led to the initiation of a strategic communication plan.

Description of Activity: The LNO Public Information Office, working with other members of the LNO staff and Network governance committees, communicates information about LTER to the LTER research and education community and its partners, the external scientific community, funding agencies, policy makers, and the public. Primary means of communication include web portals, video and presentations, a traveling multimedia exhibit, and a variety of print materials. Core activities that presently form the foundation for the LNO communication and outreach program include: (1) publication of LTER Network Newsletters, site and Network brochures, and flyers and occasional publications; (2) maintenance and continued upgrading of the LTER Network and LNO web sites; (3) periodic revision of the LTER traveling multimedia exhibit, presentations, and posters; and (4) representation of the LTER Network at scientific meetings

such as the American Association for the Advancement of Science (AAAS), ESA, and SACNAS. A Strategic Communication Plan presently under development likely will lead to the modification of some existing communication and outreach activities as well as to the addition of new activities that are approved by the Executive Board.

Outcomes: LNO communication and outreach activities will increase understanding of LTER capabilities, achievements, and contributions to science and society, especially with respect to understanding complex, long-term phenomena and to informing decision-making on ecological issues.

Operational steps:

1. Until the Strategic Communication Plan is completed and approved, the Public Information Officer, supervised by the Executive Director, will continue ongoing communication and outreach activities, including:
   a) Publication of the Spring and Fall LTER Network Newsletter
   b) Production of two revised site brochures annually
   c) One or more revisions of the LTER Network brochure
   d) Maintenance and regular updating of the LTER and LNO web sites
   e) Re-design of the traveling multimedia exhibit in Year 1 and the addition of one or two revised poster modules annually
   f) Production of flyers, posters, and occasional publications as requested by the Executive Board
   g) Outreach to annual meetings of ESA and SACNAS
2. Upon acceptance of the communication plan, the Executive Director and the CIO will review new tasks in the plan and match tasks with existing expertise and effort levels.
3. The ED and CIO will re-allocate effort and other resources to address new tasks.
4. If necessary, the ED and CIO will revise position descriptions of the Public Information Officer.
5. If necessary, the Public Information Officer will undergo training to meet new responsibilities.
6. If necessary, the ED and the CIO will recruit new staff to complete the skill set needed for new tasks.

Milestones:

- By 2010, the Public Information Officer will complete a re-design of the traveling multimedia exhibit; one or two revised display modules will be added annually in subsequent years.
- Beginning in 2009, the LNO will publish annual Spring and Fall LTER Network Newsletters unless contraindicated by the Strategic Communication Plan.
- Beginning in 2009, the LNO will assist two sites in revising their brochures annually unless contraindicated by the Strategic Communication Plan.
- Beginning in 2009, the Public Information Officer will create or acquire updated content for LTER and LNO web sites at least quarterly.

- Beginning in 2009, the Public Information Officer will deploy the multimedia exhibit at the annual meeting of the ESA.
- By 2012, LNO staff will publish a revised LTER Network brochure unless contraindicated by the Strategic Communication Plan.
- By 2012, the LNO will modify or enhance staff to address new skills required by the Strategic Communication Plan.

Evaluation: The EB will evaluate the successful completion of the suite of communication and outreach activities by 1) reviewing progress against approved milestones, and 2) reviewing the quality of communication and outreach products, 3) inspecting the results of the biennial survey of LNO performance, and 4) analyzing performance metrics defined in the Strategic Communication Plan.

Benefits: The LTER Network and its external partners will benefit from an increased understanding and appreciation of their work by targeted audiences. The larger scientific and education community will benefit from easily accessible information about LTER activities, including research results, learning tools, opportunities for engagement, information management standards and practices, training opportunities, and future plans.

Feedback: The EB will receive formal feedback on communication and outreach activities from the biennial survey of LTER scientists and educators and informal feedback from the communities reached by LNO communication and outreach activities.

3.4.2 Training. Provide or coordinate training for LTER scientists and information specialists in support of Network science and CI development.

Background: The dissemination of technical information throughout the LTER Network raises understanding to a common level, facilitating productive interactions and promoting the adoption of best practices. As new analytical techniques and technical tools come into use, a regular program of training opportunities presents the most efficient approach to disseminating standard practices across the LTER Network. In the past, LTER scientists and information managers have participated in various LNO training exercises funded through external partnerships with the Organization of Biological Field Stations and the SEEK project. The success of these training exercises and the repeated requests for additional learning opportunities prompted the LNO to initiate a regular program of training.

Description: A budget of $326,105 over five years will provide support for 0.20 FTE of personnel time to organize training activities for LTER sites and scientists as well as materials and supplies for meetings and participant support costs for trainees. Each year, the LNO will organize and host one or two three-day training workshops that reach scientists or information managers at the 26 LTER sites. As part of this exercise, the LNO will coordinate workshop planning and training activities, cover participant support costs, capture lectures and live demos on a rich media recorder, and produce formative and summative evaluation tools. Technology training for scientists could alternate with information technology training. LTER information managers have proposed several possible IT training topics, including cybersecurity and advanced GIS analysis. Possible science technology training topics include: creating and

running scientific workflows for complex analyses (e.g., Kepler) and establishing sensor networks for *in situ* measurement programs (e.g., communications, QA/QC, analysis).

Outcomes: Training sessions will increase the use and proficiency of use of advanced technologies across the LTER Network as well as leading to increases in standardization of methods among sites. Increased proficiency and standardization will lead to an increase in cross-disciplinary and cross-site publications based on use of new technologies or scientific approaches. Recorded lectures and demonstrations will provide the means to reach a larger audience as well as to prepare newcomers to the LTER Network.

Operational steps:

1. The Executive Director will prepare an annual call for suggestions for training workshops from scientists, information managers, technical staff, and students, to be submitted simultaneously with proposals for research working groups.
2. Interested groups will submit short descriptions of proposed training activities to the LNO, where they will be evaluated for practicality, breadth of interest, availability of appropriate trainers, relevance to LTER goals, and cost.
3. The Executive Director will present the proposed training activities to the EB along with results of the LNO evaluation and recommendations on funding.
4. The EB will determine which proposed activities will be funded.
5. The Executive Director will inform the LTER Network of the decision of the EB and open enrollment for the highest priority training activity. Sites will be allowed to nominate one trainee per year. If space remains after enrollment is closed, sites will be allowed to nominate a second trainee up to a total enrollment of 26. If the number of nominations exceeds the spaces available, trainees will be chosen by lot.
6. The Executive Director will recruit trainers and set the dates and location of the training. In most cases, the training will take place in Albuquerque, New Mexico, except when an alternate training facility specializing in the training topic is available.
7. Trainers will be responsible for the content of the training.
8. The LNO will organize and pay for travel and lodging for the workshop.
9. The LNO Systems Analyst will prepare the training laboratory for the workshop and be responsible for assisting the trainers during the workshop.
10. The LNO Systems Analyst will make video recordings of the training workshop and archive those recordings along with any audiovisual or written materials used in the training. The Systems Analyst will make these materials available to LTER scientists and information managers.
11. Formative and summative evaluations of the training will be completed by workshop attendees, and reviewed by the Executive Director.
12. Should training funds accumulate over time because of cancellations or other reasons, the Executive Director may schedule a second training activity in a single year. In that case, he will select the proposal ranked second in priority by the Executive Board.

Milestones: By 2014, each LTER site should have had the opportunity to participate in five training exercises.

Evaluation: The EB will evaluate each annual training workshop to modify future training priorities and the process for selecting training workshops. The Executive Director will evaluate each training opportunity from the perspective of trainee satisfaction and cost effectiveness. These evaluations will inform future decisions on training topics and venues.

Benefits: LTER scientists and information managers will be the primary beneficiaries of training.

Feedback: Feedback on the training exercises themselves will be acquired through formative and summative evaluations of the training by participants. Feedback on the LNO's role in organizing the training will be received through the biennial survey of LTER sites.

Table 1. Person effort in full-time equivalents (FTE) for LNO staff involved in ARRA funded activities.

	FTE	3.1.1 Science Council	3.1.2 Research Working Groups	3.1.3 Decadal Plan Research	3.2.1 Basic CI Support	3.2.2 Information Management	3.2.3 Network Information System	3.2.4 IT, Database, and Web Consulting	3.3.1 Acquisition of Data	3.4.1 Communication/Outreach	3.4.2 Training
Director Synthesis Support	0.60	0.10	0.20	0.10					0.20		
Public Information Officer	1.00									1.00	
Application Support Analyst	0.50				0.30			0.20			
NIS Programmer/Analyst	1.00						1.00				
Systems/Analyst	1.00				0.80						0.20
Information Manager	1.00					1.00					
Professional Intern 1	1.00				1.00						
Professional Intern 2	1.00						1.00				
Work study student 1	1.00		1.00								
Work study student 2	1.00		1.00								
Work study student 3	1.00		1.00								

Table 2. ARRA funding among 10 tasks	3.1.1 Science Council	3.1.2 Research Working Groups	3.1.3 Decadal Research Plan	3.2.1 Basic CI Support	3.2.2 Information Management
Salaries/Wages/FB	$ 82,231	$ 329,612	$ 82,231	$ 657,918	$ 381,120
Equipment > $5,000	$ -	$ -	$ -	$ 90,000	$ -
Materials and Supplies	$ -	$ 300,000	$ -	$ 99,136	$ 90,000
Participant Support Costs	$ 188,817	$ 588,471	$ 170,469	$ -	$ 138,598
Publication Costs	$ -	$ -	$ -	$ -	$ -
Sub-awards	$ -	$ -	$ -	$ -	$ -
Travel	$ -	$ -	$ -	$ -	$ 26,400
Total	**$ 271,048**	**$ 1,218,083**	**$ 252,700**	**$ 847,054**	**$ 636,118**
	3.2.3 Network Information System	3.2.4 IT, Database, Web Consulting	3.3.1 Acquisition of Data, Hardware, Software	3.4.1 Communication and Outreach	3.4.2 Training
Salaries/Wages/FB	$ 635,692	$ 76,920	$ 164,462	$ 268,513	$ 68,105
Equipment > $5,000	$ -	$ -	$ -	$ -	$ -
Materials and Supplies	$ -	$ -	$ -	$ -	$ -
Participant Support Costs	$ 59,100	$ -	$ -	$ -	$ 258,000
Publication Costs	$ -	$ -	$ -	$ 16,320	$ -
Sub-awards	$ -	$ -	$ -	$ -	$ -
Travel	$ 33,000	$ -	$ -	$ 165,000	$ -
Total	**$ 727,792**	**$ 76,920**	**$ 164,462**	**$ 449,833**	**$ 326,105**
Total direct costs	$ 4,970,115				
Equipment totals (F&A Excludable)	$ 90,000				
Participant totals (F&A Excludable)	$ 1,403,455				
F&A (TDC - Equip and Part Costs)	$ 1,766,628				
Grand Total	$ 6,736,743				

Figure 1. Provenance Aware Synthesis Tracking Architecture (PASTA) framework

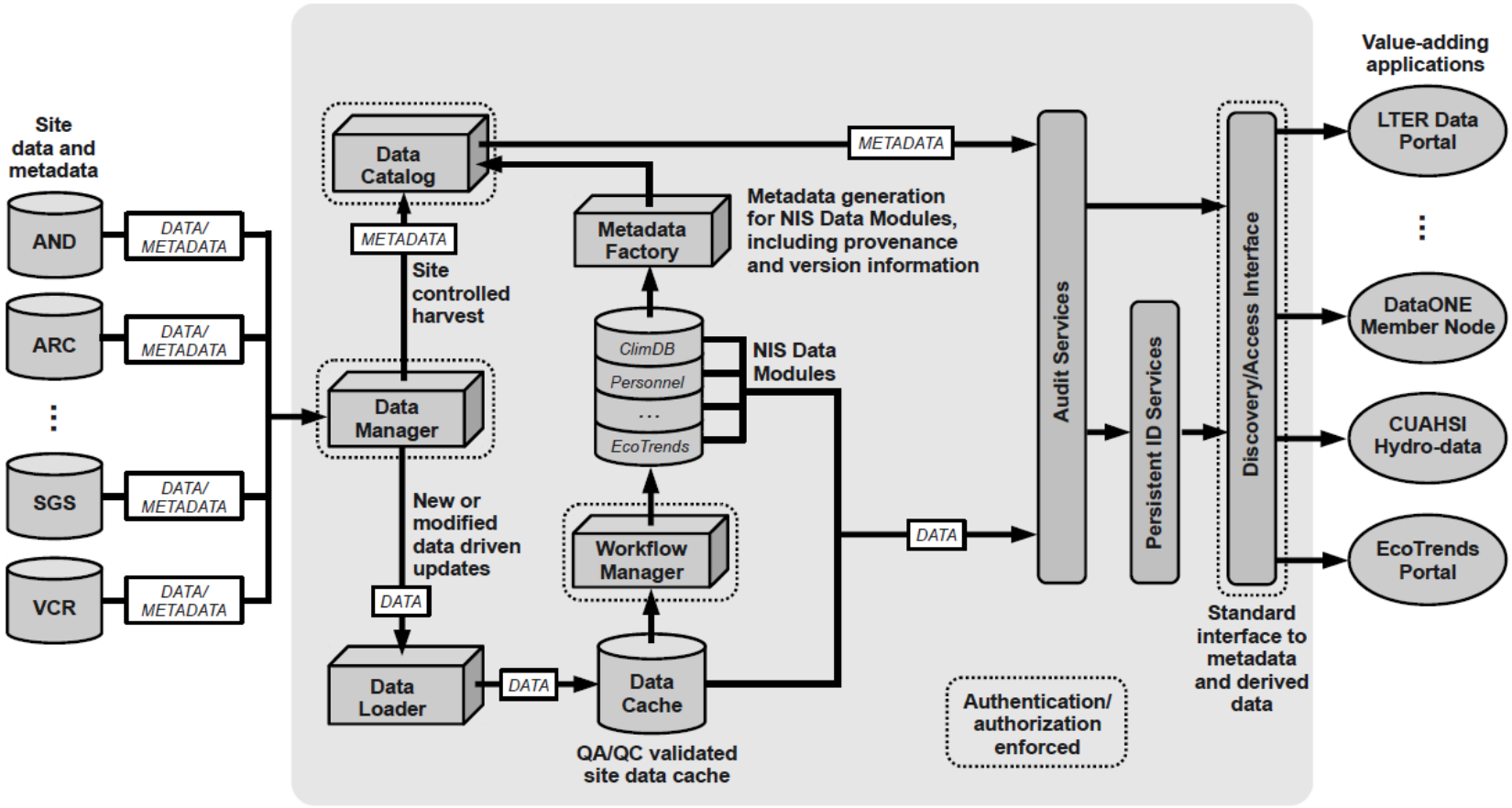


Table 2. Timeline of major development tasks

	WBS	Name
1		End of Project
2		Project Life
3		Operational Plan
4		Review Planning
5		Operational Prototype
6		Production Deployment
7		Transition/Maintenance
8	**3.2.3.10**	**Network Information System**
9	**3.2.3.10.1**	**Data management suite**
10		**Phase 1**
11	3.2.3.10.1.1	Data Manager
12	3.2.3.10.1.2	Data Catalog
13	3.2.3.10.1.3	Data Loader
14	3.2.3.10.1.4	Data Cache
15		**Phase 2**
16	3.2.3.10.1.1	Data Manager
17	3.2.3.10.1.2	Data Catalog
18	3.2.3.10.1.3	Data Loader
19	3.2.3.10.1.4	Data Cache
20	**3.2.3.10.2**	**Workflow Manager**
21		Phase 1
22		Phase 2
23	**3.2.3.10.3**	**Metadata Factory**
24		Phase 1
25		Phase 2
26	**3.2.3.10.4**	**NIS Data Portal**
27		Phase 1
28		Phase 2
29	**3.2.3.10.5**	**Discovery/Access API**
30		Phase 1
31		Phase 2
32	**3.2.3.10.6**	**Id Management Services**
33		Phase 1
34		Phase 2
35	3.2.3.10.7	Persistent ID Services
36	3.2.3.10.8	System Monitor
37	3.2.3.10.9	Audit Services
38	**3.2.3.11**	**NIS Database Integration**
39	3.2.3.11.1	Integration Path
40	3.2.3.11.2	OLAP NIS Data
41	3.2.3.11.3	Site-based Data Integration
42	3.2.3.11.4	Project Specific Data Services

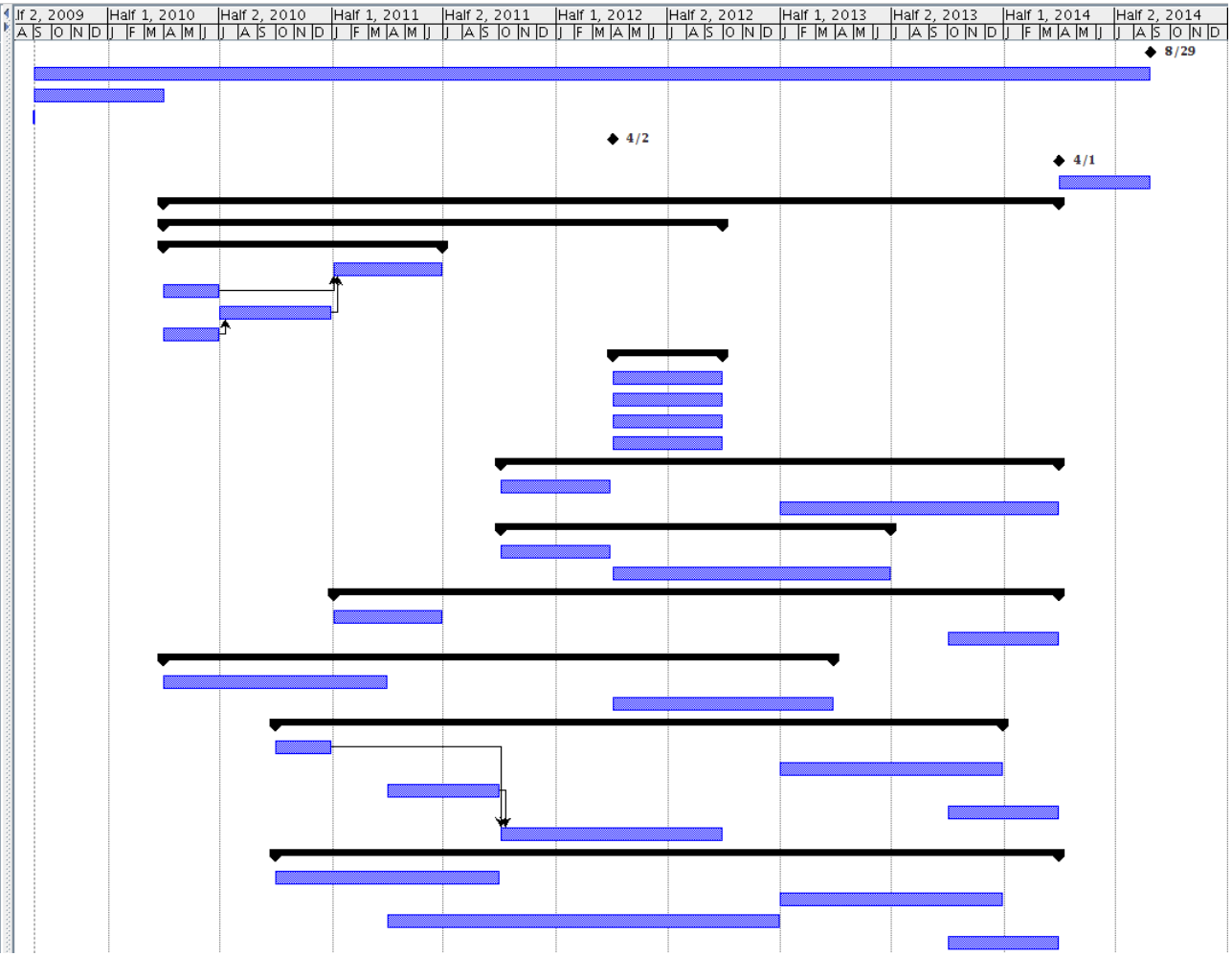
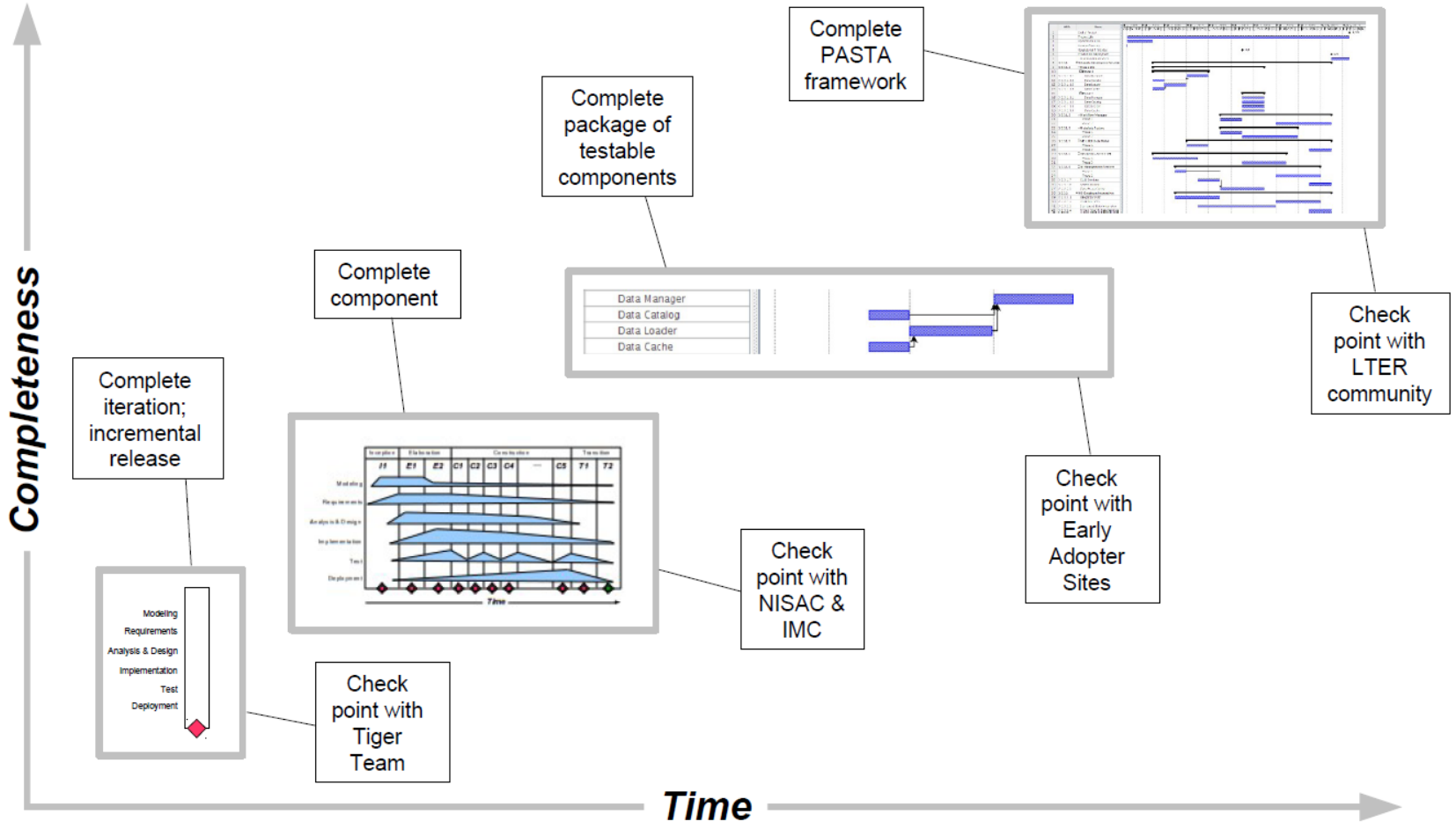


Figure 3. Levels of engagement with the LTER community



Appendix I. List of Acronyms used

AAAS – American Association for the Advancement of Science
AIBS – American Institute of Biological Sciences
ARRA – American Recovery and Reinvestment Act
ASA – LNO Application Support Analyst
ASM – LTER All Scientists Meeting
CA – NSF Cooperative Agreement
CI – Cyberinfrastructure
CILogon – NCSA supported to facilitate secure access to CI
CIO – LTER Network Chief Information Officer
ClimDB/HydroDB – LTER Network climate and hydrology databases
CPU – Central Processing Unit
CUASHI – Consortium of Universities for Advancement of Hydrologic Science, Inc.
CUASHI HIS – CUASHI Hydrologic Information System
DataONE – Data Observation Network for Earth
DOD – US Department of Defense
DOE – US Department of Energy
DSS – LNO Director of Synthesis Support
EB – LTER Network Executive Board
ED – LNO Executive Director
EML – Ecological Metadata Language
ESA – Ecological Society of America
FTE – Full-time Equivalent
GBIF – Global Biodiversity Information Facility
GIS – Geographic Information System
IM – Information Management
IMC – LTER Information Management Committee
IMExec – LTER Information Management Executive Committee
LNO – LTER Network Office
LTAR – Long Term Agricultural Research Network
LTER – Long Term Ecological Research
MTS – UNM Media Technology Services
NAB – LTER National Advisory Board
NAP2 – LNO NIS Programmer Analyst II
NAP3 – LNO NIS Programmer Analyst III
NPI – LNO NIS Programmer Intern
NASA – National Aeronautics and Space Administration
NBII – National Biological Information Infrastructure
NCAR – National Center for Atmospheric Research
NCEAS – National Center for Ecological Analysis and Synthesis
NCSA – National Center for Supercomputing Applications
ND – LTER Network Information System Developer
NEON – National Ecological Observatory Network
NESCent – National Evolutionary Synthesis Center
NIM – LTER Network Information Manager

NIS – LTER Network Information System
NISAC – LTER Network Information System Advisory Committee
NOAA – National Oceanographic and Atmospheric Administration
NSF – National Science Foundation
NSF-BIO – NSF Biology Directorate
OBFS – Organization of Biological Field Stations
OCI – NSF Office of Cyberinfrastructure
OOI – Ocean Observatory Initiative
ORNL – Oak Ridge National Laboratory
ORNL DAAC – ORNL Distributed Active Archive Center
PASTA – Provenance Aware Synthesis Tracking Architecture
QA/QC – Quality Assurance / Quality Control
RFC – Request for Comments
RUP – Rational Unified Process
SACNAS – Society for Advancing Hispanics/Chicanos and Native Americans in Science
SC – LTER Network Science Council
SDSC – San Diego Supercomputer Center
SEEK – Science Environment for Ecological Knowledge
SEV – Sevilleta National Wildlife Refuge LTER
SIM – LTER Site Information Manager
ULTRA – Urban Long Term Research Areas
UNM – University of New Mexico
URL – Uniform Resource Locator
USDA – US Department of Agriculture
USDA FS – USDA Forest Service
USGS – US Geological Survey
VDC – Virtual Data Center INTEROP Project
VTC – Video-Teleconference

Other terms:
Controlled Vocabulary Working Group – IMC working group developing an LTER specific controlled vocabulary.
Unit Registry Working Group – IMC working group developing a database and interface for vetting EML custom units.
NIS Developers – LNO NIS personnel
Chair – The chairperson of the LTER Network Science Council and Executive Board