ASM Workshop Report on "Centralized LTER Information Management?" 24 January 2013

John F. Chamblee[1], Paul Hanson[2], Emery Boose[3], Brian J. Burke[4], Corinna Gries[5], Nik Heynen[6], Margaret O'Brien[7], John H. Porter[8], Phil Robertson[9], Wade M. Sheldon, Jr.[10]

[1] Information Manager, CWT, IMC Co-Chair
[2] Principal Investigator, NTL, NISAC Co-Chair
[3] Information Manager, HFR, LNO Visioning Committee Co-Chair
[4] Post-Doctoral Research Associate, CWT
[5] Information Manager, NTL
[6] Principal Investigator, CWT
[7] Information Manager, SBC, IMC Co-Chair
[8] Information Manager, VCR, NISAC Co-Chair
[9] Lead Principal Investigator, KBS, LNO Visioning Committee Co-Chair
[10] Information Manager, GCE

I. Executive Summary[11]

The LTER Network has seen several calls from outside observers to change network-wide information management practices to improve data availability. One specific suggestion raised revolved around the idea of centralized information management. The ASM Workshop "Centralized LTER Information Management?" was a facilitated discussion exploring the opportunities and challenges associated with centralizing LTER Information Management activities. Site information managers and lead principal investigators were encouraged to fill out a survey on current practice and to attend ASM workshop discussions focused on the actual opportunities and impediments to centralizing some or all aspects of information management within a 5-10 year time horizon.

The survey had a 48% response rate. Results show agreement on general information management practice, but also that some divergent perceptions exist when it comes to allocation of time and resources. Information Managers report involvement in tasks that some principal investigators do not recognize as part of the information management responsibilities list. Responses also vary in terms of their estimates of time allocated to specific tasks, especially in the case of long-term data management, to which information managers allocate more time, when compared to principal investigators.

In workshop discussions, participants classified tasks as strong candidates for centralization or as tasks best done locally. Once several tasks or types of tasks had been classified, we looked for general patterns in the classification. Those tasks most associated with providing support for local science were almost always seen as best done at the site level. Similarly, those tasks where economies of scale are important, such as large-scale software development, were often classified as candidates for centralization.

In discussing challenges, many information managers pointed out that, under the current funding model, they would likely be held responsible for the quality of all IT services used by LTER scientists from their site, whether or not they were centralized. This led to strong concerns that any effort to centralize information management activity, regardless of the potential efficiencies to be gained, would have to be accompanied by a carefully designed plan for governance and accountability. Furthermore, it was noted that LTER research activities were often highly localized. Any effort to build standard solutions to LTER information management challenges will have to involve a philosophical commitment across the LTER community to developing standard methods for solving common problems.

Based on analysis of group responses, the workshop organizers reached six basic conclusions[12]:

1) Models for centralization must be based on consensus understanding of current cost-benefit structures for information management practice. At present, such an understanding does not exist.
2) Some information management tasks can be effectively centralized, others cannot.

[11] In 2012, Chamblee and Boose co-authored a LTER News newsletter article with Susan G. Stafford and Robert J. Robbins that presents an abbreviated version of this summary. See Robbins et al. 2012 for the full text.

3) The principal drivers distinguishing the two are: tasks directly related to supporting local science must stay local; tasks primarily related to supporting third-party data use are candidates for centralization
4) Improved cost-effectiveness can be an argument for centralization, but this must be traded off against local responsiveness. Any effective, on-going movement of some local information management activities into a long-term centralized model must include new approaches to governance, reporting, and dispute resolution that ensure that local responsiveness and effectiveness are maintained while striving for centralized efficiency.
5) Cost-benefit models for existing practices are a starting point for considering centralization, but must be contextualized by an understanding of network needs.
6) The LTER Network would be best served by internally developed, service-oriented information management that respect the goal of serving site science.

II. Introduction

In the past year, the LTER Network has seen several calls from outside observers to change network-wide information management practices to improve data availability. A specific suggestion raised by several commentators revolved around the idea of centralized services. The LTER Thirty Year Review (NSF 2011a: 15), the NSF response to the Thirty-Year Review (NSF 2011b: 3), and the Robbins addendum to the Thirty-Year Review (Robbins 2011: 27 and throughout) all emphasize the importance of centralization to one degree or another.

There is certainly merit to arguments that LTER information managers might be in a better position to deal with the rapidly increasing volume and complexity of ecological data by using an informatics hub and a set of shared tools and approaches to take advantage of economies of scale (cf. Robbins 2011: 28-29). Nevertheless, the implementation of such a hub would represent a major departure from more than twenty-five years of LTER information management practice. Given the size and complexity of both the network and its data, LTER Network leaders would have to carefully and skillfully manage such change.

Assuming that such changes are desirable, the first step in considering a shift toward more centralized information management services is consultation with the people who would be most directly affected by proposed changes. The goal in this case is to learn about the conditions under which more centralized practices would be both practically possible in the short term and broadly beneficial in the long term. In such a dialog, it is also crucial to both outline the challenges involved with such a significant alteration of current practice and understand the scope of approaches to centralizing information management. These activities should be carried out with the understanding that differing options could either benefit or harm the LTER Network's goals of increasing both data availability and the potential of LTER data sets to contribute to cross-site synthesis (see NSF 2011a: 3 and NSF 2011b: 2).

This workshop was an attempt to begin that frank and open dialog about the prospects and perils of centralized information management within the LTER network. As organizers, we saw that we could only seriously consider the question "Why Centralize?" if we were able to accomplish three goals: 1) to provide an opportunity to understand current practice, 2) to talk openly about ways to improve those practices, and 3) to link those talks to specific recommendations directed at decision-makers tasked with managing the change necessary to ensure that any potential centralization plans meet the needs of both site and network science.

Given NSF recommendations to think broadly about the structure of the network (see NSF 2011a: 5) and the recent commissioning by the LTER Executive Board of the LTER Network Office (LNO) Visioning Committee, we also sought to provide feedback to the Visioning Committee on how centralized plans might fit into the long-term objectives that the Network leadership may choose to recommend for LNO. In addition to focusing on conceptual aspects of the problem, we also felt that our feedback would be most useful if we put a strong emphasis on functional aspects of the problem (personnel, governance, workflow, computer systems, resource limitations, etc.).

In the next two sections, we define the institutional context in which the workshop took place and then discuss the methods we used to organize the discussion and collect feedback. The

workshop was preceded by an on-line survey and organized as a facilitated discussion in which we recorded results on posters and note cards. We then presented the results of the survey and discussions, highlighted key findings, and concluded with some recommendations. Our results indicate that, while it is clear that some information management tasks are amenable to centralization, it is equally clear that the centralization process is one that will require careful thought, attention to site needs, and a philosophy of service to site science.

III. Scope

In formulating plans for this workshop, we sought to define the boundaries within which our discussions should take place. This basically meant establishing the baseline facts of our existence as participants within the LTER Network. We presented these baseline facts to our participants and they are repeated here as a way of framing the conclusions of the report.

The most basic fact of all for the workshop participants, and one that bears repeating many times, is that, as members of the LTER Network, we are committed to the Network's mission as stated in the LTER Strategic and Implementation Plan (LTER Network 2011:5) "to provide the scientific community, policy makers, and society with the knowledge and predictive understanding necessary to conserve, protect, and manage the nation's ecosystems, their biodiversity, and the services they provide." When we consider centralized information management solutions, we should be doing so in such a way that our proposed solutions would improve our ability to fulfill that mission.

Beyond that basic commitment, the next most important dimension of our roles in the LTER pertains to the current institutional and funding structures within LTER. Figure 1 is adapted from the LTER Strategic and Implementation Plan (LTER Network 2011:5). In the diagram, we highlight the organizations likely to be most directly impacted by plans to centralize information management practice within the LTER. Those most directly affected are likely to be the LNO, the Information Management Committee (IMC), and the Network Information Management Advisory Committee (NISAC), since they would be the ones responsible for implementing such plans. Additionally, individual sites will be dramatically affected as well, as their local information management practices will have to change to adapt to any new requirements.

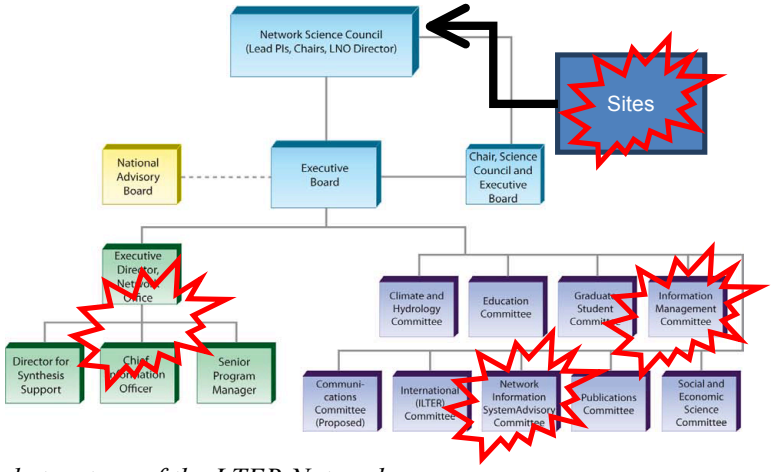


Figure 1. Institutional structure of the LTER Network.

Moving from the specific to the more conceptual, we assert the rather basic, but again crucial point that LTER information management systems are like any other IT system in that they consist of multiple components that are both human and technical (see Figure 2). In developing centralization plans, we must account for all of the components – cultural, technological, process-oriented and structural, that comprise the system.

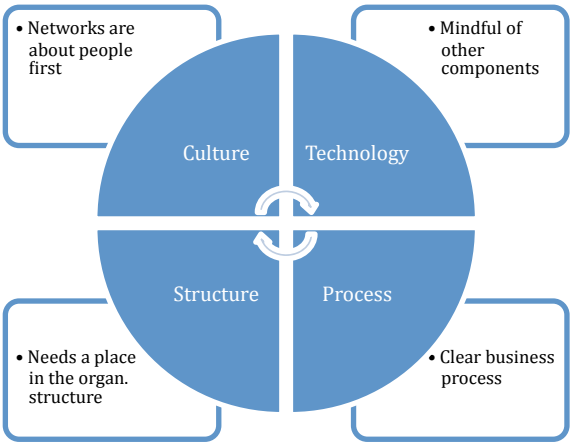


Figure 2. The socio-technical components of an information management system.

Although both the human and technological aspects must be considered during centralization, we must be mindful of the technical requirements necessary for any LTER information management system to operate. Figure 3 presents a view of all of the technological pieces that comprise any site-level system. In considering centralized plans, we must consider not only which components could be centralized and which could not, but also how the various components will have to fit together if any one of them is switched from a distributed to centralized management regime.

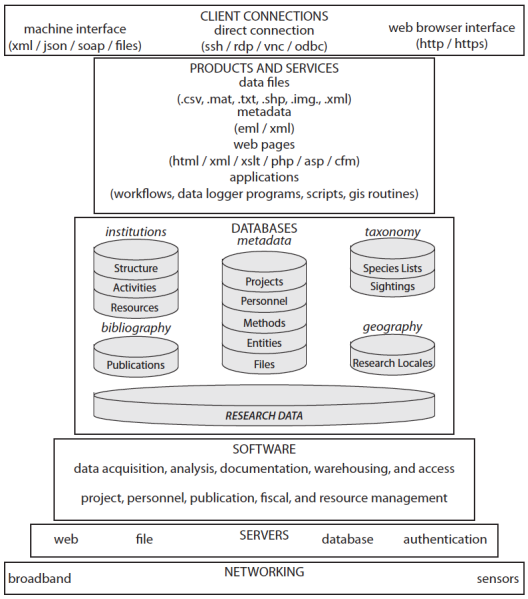


Figure 3. An abstract view of the standard information technology stack required by each individual LTER Site.

A final facet of the workshop's scope related to the social dynamics we needed to foster in order to have a successful discussion. Our goal was to capture the "zeitgeist" of the LTER Network with regard to our problem area, discuss controversial topics openly as needed, and thereby generate a report that would prove useful to our mission of serving the LNO Visioning Committee. We recognized three key challenges:

1. Perceptions of information management and information manager responsibilities vary widely, both within individual sites and across the network. These diverging perceptions should be revealed, even if they cannot be reconciled.
2. Information Management is often a controversial topic within the Network. In addition, the Network itself has inherent social inequalities related to the roles that different professionals play. In order to maximize participation and information availability, discussions would need to be mediated in a way that all concerns could be addressed without creating undue conflict during the event and without long-term social repercussions.
3. In addition to being controversial at times, LTER information management is also complex, involving multifarious connections to scientific practice, institutional and personnel management, basic information technology, and even finance. As was the case with varieties of perception, this complexity would have to be recognized, even if it could not be fully addressed in a single workshop.

We addressed these challenges with methods, outlined below, that were designed around a facilitated discussion to be overseen by trained social scientists within the LTER Network.

IV. Methods and Approach

In order to address the challenges listed above, we had to understand some of the variation in perceptions regarding LTER information management, provide "safe spaces" for discussion of controversial topics, and limit the scope of discussions to a set of topics that could fruitfully be addressed in a single LTER workshop. We chose to accomplish theses tasks by expecting some level of participation in the workshop prior to the ASM itself, by providing participants with the means for anonymous feedback, and by developing a detailed script of the workshop discussion, even if some or all of that script was scrapped during the event.

We elicited pre-workshop participation by developing a survey designed to evaluate variation in the perceptions of current practice within sites and across the network. The survey, included in the report as Appendix 1, was sent to the primary site information manager and the lead principal investigator for each site prior to the ASM. Respondents were asked to complete the survey two days before the workshop, so that results could by analyzed. The survey was designed to be anonymous and no personally identifiable data were collected.

In addition to making the survey anonymous, we increased opportunities for the discussion of controversial topics in two ways – we collected only aggregate responses on the actual results of discussions, and, for instances in which a participant wanted to provide a comment that might be controversial, we provided note cards to all participants so that they could write down comments during the course of the discussion and then turn them in at the workshop's end. Comments were recorded on posters and tacked up along the walls as the discussions progressed. Participants

were also allowed to provide feedback on note cards if they did not wish to address a point publicly. Some participants did so. Addressing the same issue through a number of methods and a number of conceptual angles is a form of "triangulation" that increases the validity of both our qualitative and quantitative data.

We used two avenues to limit discussion to a manageable set of topics. First, we provided two formal presentations at the beginning of the workshop. Paul Hanson (NTL, NISAC Co-Chair) gave a short presentation fully detailing the scope of the workshop and outlining some key aspects of the complexity surrounding LTER IM. The first three paragraphs of the Scope section of this report briefly summarize Paul's presentation. John Chamblee (CWT, Incoming IMC Co-Chair) gave a presentation on the results of the survey. We then presented participants with a fixed set of questions that were presented along with a general workshop outline.

At the beginning of the workshop, we tied all these approaches together by handing out an agenda for the workshop that included our framing questions. The agenda is included as Appendix 2. As the workshop progressed, we ended up modifying both the agenda and the questions in response to on-the-spot feedback we received from some participants indicating more consensus on and foreknowledge of the topics than we anticipated. The final survey results, the final questions discussed, and the group's answers to these questions are provided in the next section.

V. Survey and Workshop Results

Sixteen information management site representatives and eight lead principal investigators responded to the survey, providing an overall response rate of 48%. During the workshop presentation of survey results, the organizers presented the results below. In the ensuing discussion, participants and organizers agreed that the survey was characterized by several limitations that readers should bear in mind when reviewing the quantitative results.

Question 3 listed tasks that were assumed to be part of regular information management duties. Some information managers felt that these tasks overlapped, while some scientists felt that the task list was not comprehensive and that some of the tasks might be described differently. The problem of overlap was more significant in Question 4 and the large number of respondents picking one of two sets of two-answer pairs indicates that this question failed to meet its goal of capturing variation regarding potential differences in information management philosophy across sites. Results from Question 4 are therefore not reported here, but included at the end of Appendix 1. A final limitation noted in the workshop the fact that the survey results are not tied to sites – an unfortunate limitation but one that was necessary to maintain respondent anonymity.

While the survey did gloss over a good deal of complexity with regard to the tasks information managers perform, we believe that the limitations of the survey do not invalidate the results. In addition, we show below that the survey fulfilled its intended purposes, which were to ground discussions of centralization in an empirical view of current practice, stimulate discussion, and point out directions for future thought as we consider long-term plans for meeting LTER needs.

Table 1. Percentage of respondents selecting FTE categories for information management funding.

FTE	IM Site Rep	Lead PI	Total
< 0.5	0.0%	0.0%	0.0%
0.5 - 0.75	0.0%	0.0%	0.0%
0.75 - 0.99	18.8%	12.5%	16.7%
1	18.8%	0.0%	12.5%
1.01 - 1.50	25.0%	12.5%	20.8%
2	12.5%	37.5%	20.8%
> 2	18.8%	37.5%	25.0%
Don't know	6.3%	0.0%	4.2%

Table 1 shows results from Question 2, concerning the number of FTE positions committed to information management at each site. We conclude from the divergence in these responses the surveys were in many cases completed by lead principal investigators and information managers from different sites. Taken as a whole, the responses suggest that most sites are dedicating at least 1.5 FTE to information management, and in many cases are dedicating more.

Table 2. Task categories included in LTER IM duties and responses concerning the time taken to do them

	% Selecting Task			*Avg. % Reported*		
Information Management Tasks	**IM**	**LPI**	**All**	**IM**	**LPI**	**All**
Managing long-term monitoring data (data entry, quality control, archiving, and metadata development)	87.5	57.1	75.0	21.6	10.8	16.2
Developing and maintaining the site web presence	81.3	42.9	66.7	14.6	10.0	12.3
Application development in support of data management (e.g. Matlab scripts, data workflows, etc.)	87.5	71.4	79.2	10.6	11.7	11.2
Preparing data products for network distribution (e.g. PASTA and climDB/hydroDB)	81.3	42.9	66.7	6.5	12.8	9.7
System administration	87.5	71.4	79.2	7.1	11.7	9.4
Unknown	87.5	57.1	75.0	0.3	16.7	8.5
Other Network tasks (e.g. IM Exec, working groups, etc.)	81.3	71.4	75.0	9.1	4.7	6.9
Other site tasks (e.g. field work, sensor deployments, instrument maintenance, publishing etc.)	81.3	71.4	75.0	5.6	6.7	6.2
Assisting researchers with archiving and publication of specific, short-term data sets	87.5	28.6	66.7	8.5	3.7	6.1
Populating network databases (e.g. PersonnelDB, BiblioDB, and SiteDB)	87.5	71.4	79.2	3.8	5.7	4.7
Assisting researchers with data analysis (data mining, preparation, quality control, statistical analysis, or model development) prior to the publication process	81.3	42.9	66.7	4.8	2.7	3.7
Site LTER Administrative support (annual report prep, etc.)	75.0	57.1	66.7	3.9	1.3	2.6
End user technical support	12.5	57.1	25.0	3.5	1.7	2.6

Table 2 summarizes the responses to Question 3, in which respondents were asked to estimate the percentage of time that information management staff dedicated to predefined tasks. Column 1 lists these tasks. Several patterns are evident, despite overlap in some tasks and omission of others.

The section entitled "% Selecting Task" (columns 2 – 4) shows, by role, the percentage of respondents providing any non-zero value for a given task. Among information managers, the mode for each task is 87.5% or 14 out of 16 respondents assigning some value, with next two most common being 81.3%. In all analyses of the lead principal investigator respondents for Question 3, we used a total response rate of seven, because one lead principal investigator entered zeros for all but the "unknown" task and, in the comments field replied "It's unclear to me at present how our site organizes this work." Even eliminating that entry, the mode for lead principal investigators is 71% (five out of seven), with next most common value being 57%, and the remaining tasks being selected by less than half of the respondents.

The last section of Table 2 (columns 5 -7) reports the percentage of total time respondents estimated that information managers at their site spent on the listed tasks. Unknown and other categories were provided so that the time could sum to 100%. Information manager and lead principal investigator responses correspond closely in areas in which the average estimates by all respondents is less than seven percent, Values indicating a greater average time commitment exhibit greater discrepancies. The most notable differences are between estimates for long-term data management. The average information manager estimate is twice that of lead principal investigator.[13]. Information managers also spend more time on web site maintenance and network activities than lead principal investigators think they do, but less time on system administration and network data products. As noted in the methods section, Section 4 results are reported for the record and without comment in Appendix 2. Question 5 asks whether or not information managers regularly attend Site Executive Team meetings. Sixteen of 24 respondents, or two-thirds, said yes.

When the above results were discussed in the workshop, we learned that one reason that a discrepancy exists between lead principal investigator and information managers perceptions concerning the scope of professional duties is that lead principal investigator and site scientists recognize a series of other information management tasks that were left out of the survey. While this is likely not the only reason for the discrepancies, one easy way to move forward is to find out what the missing tasks were. Therefore, the first discussion topic was to enumerate the information management tasks that were left out of the list provided in the survey.

The next step in the discussion was to list all of the tasks that workshop participants thought should be part of LTER information management staff duties and solicit opinions on whether these tasks are best performed in a centralized or decentralized manner. Table 3 summarizes this discussion with a categorized list of tasks accompanied by columns in which an "X" indicates that people provided arguments in favor of either centralized approaches, distributed approaches, or both.

When we organized specific tasks into generalized task types, we concluded that some task groups, such as application development, are clearly amenable to centralization, while others, such as content development, are clearly more suitable for management by and for sites. System provisioning could be handled centrally in some cases, but such options may be complicated by

[13] Though the values are smaller and discrepancies lesser in terms of total percentage, the same pattern is true in terms of the magnitude of the discrepancy for assistance of researchers with short-term dataset publication preparation as well.

the variation in institutional histories and arrangements across sites. End-user support is a mixed category, likely requiring effort at both central and local levels.

Table 3. Task list showing presence of supporting arguments for centralized and distributed approaches.

Task	***Centralized***	***Distributed***
System Provisioning		
Web Hosting	X	
System Administration	X	X
Backups	X	X
Application Development		
Metadata authoring tools	X	
Data Entry Tools & Services	X	
Searching and Keywording engines	X	
Data Publication portals	X	
Data Quality Control tools	X	X
Network Database Development	X	
Instrument Management and Tracking Architecture	X	
End-user support		
Training	X	X
Making User Entry easier	X	X
Assisting Researchers with Data Discovery	X	X
Content Development		
Site Reports	X	X
Metadata Content		X
Schoolyard and Citizen Science Projects		X
Data set Orientation and Consulting about data set availability and appropriateness		X
Science of CyberInfrastructure and Management (Policy and Management)	X	X
Production of Standard Synthesis Products	X	X
Consultation with Site Scientists on Technology and Data Availability for Future Research		X

During discussion of the individual tasks, participants provided a variety of arguments for and against centralization on a task-by-task basis. After the group finished the task listing exercise, we went back and started summarizing these arguments into a more succinct form. Tables 4a and 4b list, respectively, the arguments for distributed (Table 4a) and centralized (Table 4b) information management services. Overall, these arguments revolve around balancing tension between dealing with LTER's inherent and necessary diversity and the need to build better tools and increase data availability and quality by leveraging economies of scale.

Table 4a. Arguments in Favor of Distributed Information Management

It is necessary to have IMs who are familiar with not only data and technology, but also the people and cultures of practice at a site.
Collaboration on IT and data intensive projects must be built on personal relationships, communication, and trust. These are most efficiently maintained at the site level.
The institutional structure of the LTER is inherently diverse and includes legal and political contexts that have policy implications for the development of IT solutions. These frameworks will impact what is possible at a given site.
Many of the field stations in the LTER network are very remote. The high network latency and low accessibility in these situations require local solutions with local support.
Almost all sites have some specialized IM tasks that require the expertise of an on-site information manager, and information managers are held responsible for the quality of IT services sites receive.

Table 4b. Arguments in Favor Centralized Information Management

Sites have a number of generic tasks that are common to all LTER participants. Provisioning of the tools to complete these tasks could be centralized if the tools serve sites needs. This is especially true of tasks that could be automated or require little human intervention.
The establishment of centralized information management services could facilitate the development of some common data standards by providing a common technological framework with which to work.
The development of centralized software frameworks on centralized systems could facilitate greater community integration by allowing information management specialists to focus their diverse skill sets on a common set of tools.
Some of the broader network challenges for information management require a concentration of highly specialized skills that, given the generalist-orientation necessary for successful site information management, are not always practical at the site level.

The final topic presented to the group related to the challenges facing any plan that might be developed for centralizing some or all aspects of information management practice within the LTER network. This resulted in an extended discussion that listed not only challenges to centralization plans, but also a list of needs that were not being met by current infrastructure and could be addressed in future long-term plans. Table 5 summarizes the challenges participants provided and Table 6 summarizes the unmet needs that participants listed.

These results are culled from both the specific discussions of challenges and needs and from the earlier discussions of centralization with respect to specific tasks. The relationships between the specific outcomes and feedback from the survey and the discussion and the themes just listed are addressed in the next section.

Table 5. Challenges involved in creating centralized information management services.

Challenges
Cultural bias toward existing site practice
The current funding and governance model for the LTER Network as a whole, which evaluates sites on an individual basis and does not provide much in terms of frameworks to "scale-up" beyond the site.
The lack of concordance between LNO mission and the potential need for centralized information management services that serve site needs.
The institutional and scientific divisions inherent to the diverse nature of the LTER network.
The complexity of partnerships and relationships that crosscut the LTER Network.
Information management is expensive and current cost-models for LTER information management do not account for the true cost of information management activities.
Mechanisms for accountability and responsibility in providing information management services to sites from a centralized location are not currently defined.
The lack of expertise in managing transitions from a more-or-less fully distributed system to one that is more centralized. Experts in managing such transitions do exist and it would be good to bring such individuals into any centralization project, rather than re-inventing the wheel.
Broader political and economic contexts for scientific research centralization involve top-down standardization models imposed from above and resented and resisted by those forced to implement them.
Lack of attention, both within sites and across the network, to client-centered, customer-service, and performance-oriented models for information management from the world of commercial IT and IM.
Insufficient infrastructure for communication and translating Network information management expertise, both LTER-wide and within LNO, to the wider network.
Lack of priorities for classes of data to be centralized and/or standardized across the network.
The lack of a comprehensive and rigorous analysis of the costs, benefits, and trade-offs involved in centralizing some or all information management services. This should be conducted internally before imposed externally.
A lack of clarity on whether or not timesaving incentives related to centralization will be sufficient to defray the front-end costs of a distributed-to-centralized model.
No mechanism for dispute resolution between centralized service(s) and clients.
No mechanism for ensuring equitable distribution of resources.

Table 6. Unmet needs in current information management practices.

Needs
Integrated data products that would require feedback from scientists on standards and priorities
Consistent access to all LTER data.
Knowledge about data use from across the network.
Consistent messages from priority setters, both within and beyond the network.
Rapid, easy-to-use tools for ingesting, documenting, and publishing data.
An informed community that is ready to participate in the data publication process.

VI. Discussion

The overall workshop goals were to gain an understanding of how information management practice is perceived and understood by information managers and scientists alike and then, based on that understanding, begin discussions about the prospects and difficulties associated with centralizing some or all parts of information management. The purpose of the workshop was to use the survey and discussions to synthesize a broader set of results to present to the LNO Visioning Committee. These broader themes are presented below and include similarities and differences in perceptions of practice, types of tasks most amenable to centralization, common themes in arguments for centralization and distribution, and patterns evident in challenge and needs lists for any information management centralization plans.

Survey results show much consensus concerning the overall roles that information managers play within the site and the network. However, there are some points of difference. While these discrepancies could be linked to the fact that lead principal investigators and information managers from different sites completed the survey, a more likely interpretation is that there are complexities to current site-level information management practices that are not fully understood beyond the information management community. One lead principal investigator explicitly stated s/he did not understand information management practice at his or her site at all. These data, coupled with the fact that the average value for "unknown" (essentially unaccounted for) time is 16%, even without the outlier responses, reinforce the conclusion that the full scope of work for an LTER information manager is not well understood outside the LTER information management community.

Discussions of specific tasks clearly show that some tasks would be amenable to centralization, especially those related to system provisioning and the development of software applications to support both work at both the site and network level. However, some tasks are also clearly best left at the site, especially those requiring detailed knowledge of information management and scientific practice, as well as those reflecting the manifest diversity in scientific practice and institutional structure present in the LTER network. Arguments for centralized and distributed approaches reinforce the task-based results and highlight not only areas where centralization potential is high, but also those areas where site-based practices are and will remain crucial. Discussions regarding challenges and unmet needs covered the topics of governance, cost, accountability, scientific priorities, and the diversity present in the LTER Network. Discussions of cost and accountability focused on three areas:

1) Information managers will ultimately be held responsible for LTER IT solutions.
2) Any centralization plans must be grounded in both a clear understanding of costs and benefits and an operationally detailed outline for transitioning to a centralized model.
3) Centralization plans should be based on both more formal cost-benefit assessments of centralized versus distributed practices and on inclusion of "change-management" experts with experience in helping organizations transition between business models.

Discussants also cautioned against top-down solutions that would impose centralization through standard practices to be adopted without accounting for the costs involved in implementation. Overall, discussants shared the view that the LTER network would be best served if plans for centralizing information management services came from within the LTER community.

VII. Conclusions and Recommendations

In assessing the workshop as a whole, we feel confident saying there was consensus among the participants in three areas:

1) Some parts of the information management workload are amenable to centralization.
2) Other areas, specifically those most closely tied to site practice and site science, need to remain distributed.
3) Any plan to centralize information management services must be focused on meeting the needs of site information managers and, more importantly, the goals dictated by site science.

However, to achieve greater consensus than this, the wide-ranging nature of the discussions and diversity of opinions we observed lead us to conclude that further agreement would have to be built on a more practical and operationally complete picture of what centralized services might look like. The first impediment to such a plan is the divergence of opinion on current practice. While the survey provided for the workshop was flawed and somewhat limited, it was effective enough to show that there are different viewpoints on what information managers are doing and what they should do. We need to understand and resolve these differences.

The best way to achieve such understandings may be to more carefully study practice. Since any development of centralized plans must be built on a clear understanding of relative costs and benefits, it would be in the LTER Network's best interest to conduct cost-benefit analyses internally and in a way that includes objective measures and a complete understanding of current practices. Cost-benefit analyses that do not account for all the work information managers are relied upon to complete will not serve the long-term interests of the network. Once we better understand current practice and its cost-benefit structure, we can begin thinking about how to develop a plan for centralizing many, but not all, aspects of information management.

As a committee, we are optimistic about the prospects for centralized information management services because, if executed well, they would benefit scientists, the public, and information managers. However, we must be mindful of focusing exclusively on reducing costs by capitalizing on economies of scale and eliminating redundancy as the main reason to centralize. Currently, the LTER lacks both the mechanisms and the strategies to build a path to centralized services that serves the needs of the sites.

Developing the expertise to centralize is likely to have costs of its own and, more importantly, if the transition is not carried out carefully, the result could very well be a centralized approach that does not meet site needs, thereby resulting, paradoxically, in further cost increases. To avoid such scenarios, accountability – specifically accountability to sites – must receive equal emphasis with cost. This should be achieved by the developing the mechanisms for governance, accountability, and effectiveness that ensure responsiveness to local needs.

The other underlying and often unspoken motivator for centralizing information management is to better to serve the scientific objectives of LTER scientists. Though sometimes distant from the complexities of information management, it is this objective that must remain in the forefront as any information management centralization plans develop.

References Cited

Long Term Ecological Research Network. 2011. Strategic and Implementation Plan Long-term Ecological Research Network (LTER): Research and Education. Manuscript on File. LTER Network Office, University of New Mexico, Albuquerque (*http://intranet2.lternet.edu/sites/intranet2.lternet.edu/files/documents/LTER%20History/Planning%20Documents/LTER_SIP_Dec_05_2010.pdf, accessed 1 December 2012)*.

National Science Foundation. 2011a. Long-Term Ecological Research Program A Report of the 30 Year Review Committee. Manuscript on File. National Science Foundation (*http://www.nsf.gov/pubs/2012/bio12001/bio12001.pdf, accessed 1 December 2012)*.

2011b. NSF Responses to LTER Thirty-Year Review Report. Manuscript on File. National Science Foundation (*https://www.nsf.gov/attachments/125204/public/Response_to_LTER_30-year_review.docx, accessed 1 December 2012)*.

Robbins, Robert J. 2011. Data Management for LTER: 1980 – 2010, A Position Paper. Manuscript on File, National Science Foundation (*http://www.nsf.gov/pubs/2012/bio12002/bio12002.pdf, accessed 1 December 2012)*.

Robbins, Robert J., Susan G. Stafford, John F. Chamblee, and Emery R. Boose, 2012. From the Anthropocene to the Applet: A tale of two workshops. *LTER Network News* 25(3), (*http://news.lternet.edu/Article2676.html, accessed 1 December 2012).*

Acknowledgements

The author/organizers wish to thank all of the LTER and non-LTER information managers and scientists who filled out surveys, participated in the workshop, and provided feedback. We also thank Robert J. Robbins for comments on the Executive Summary.

Appendix 1

Survey questions provided to information managers and lead principal investigators prior to the workshop.

What would centralized Information Management look like? A survey of current practice

This survey collects information about current practices and activities for LTER Site Information Management. The survey is designed members of both Information Management and Site Executive teams. The survey is anonymous and designed to elicit differences in understanding of Information Management roles across and within LTER sites. Please answer individually and do not coordinate your answers with anyone else. The survey is investigative and there are no "correct" answers.

1. What are your LTER roles (please select all that apply)?

- a. Information Manager (Site Representative)
- b. Other Information Management team member
- c.Lead Principal Investigator
- d. Site Executive team member

2. What is the FTE (Full Time Equivalent -- the proportion of salary dedicated to one person in a year) provided for IM activities at your site?

<0.5 0.5 - 0.75 0.75 - 0.99 1 1.01 - 1.5 2 >2 Don't know

3. What percentage of time does your site's Information Manager or IM team spend on the following activities (the sum will total to 100%)?

- Managing long-term monitoring data (data entry, quality control, archiving, and metadata development)
- Assisting researchers with archiving and publication of specific, short-term data sets
- Application development in support of data management (e.g. Matlab scripts, data workflows, etc.)
- Preparing data products for network distribution (e.g. PASTA and climDB/hydroDB)
- Populating network databases (e.g. PersonnelDB, BiblioDB, and SiteDB)
- Developing and maintaining the site web presence
- Site LTER Administrative support (annual report prep, etc.)
- Assisting researchers with data analysis (data mining, preparation, quality control, statistical analysis, or model development) prior to the publication process
- System administration
- End user technical support
- Other site tasks (e.g. field work, sensor deployments, instrument maintenance, publishing etc.)
- Other Network tasks (e.g. IM Exec, working groups, etc.)
- Unknown

4. Which statements describe your site's approaches to information management (choose as many as apply)

- a. The Information Manager or IM team determines priority of tasks within the framework of project needs, data generation, and other demands.
- b. IM tasks follow a strict schedule determined by field work in which the Information Manager or IM team is intimately involved.
- c. The Site Executive team and Information Manager or IM team work closely to determine long- and short-term information management priorities.
- d. The Site Executive team looks to the Information Manager or IM team for guidance in terms of best practices and implementation of new technologies for site data management.
- e.I would like to provide additional comments regarding our site's approach to information management.

4a. Please provide any additional comments relevant to your site's approach to information management below

5. Does the Site Information Manager regularly attend meetings of the site's Executive Committee or equivalent body?

- YES
- NO

Appendix 2
Survey results.

What would centralized Information Management look like? A survey of current practice.

1. 1. What are your LTER roles (please select all that apply)?

#	Answer		Response	%
1	a. Information Manager (Site Representative)		19	70%
2	b. Other Information Management team member		1	4%
3	c.Lead Principal Investigator		10	37%
4	d. Site Executive team member		10	37%

2. 2. What is the FTE (Full Time Equivalent -- the proportion of salary dedicated to one person in a year) provided for IM activities at your site?

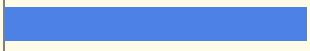
#	Answer		Response	%
1			1	4%
2	0.5 - 0.75		1	4%
3	0.75 - 0.99		4	15%
4	1		4	15%
5	1.01 - 1.5		5	19%
6	2		5	19%
7	>2		6	23%
	Total		26	100%

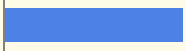
3. 3. What percentage of time does your site's Information Manager or IM team spend on the following activities (the sum will total to 100%)?

#	Answer	Min Value	Max Value	Average Value	Standard Deviation
1	Managing long-term monitoring data (data entry, quality control, archiving, and metadata development)	0.00	100.00	22.67	19.92
2	Assisting researchers with archiving and publication of specific, short-term data sets	0.00	20.00	6.33	4.62
3	Application development in support of data management (e.g. Matlab scripts, data workflows, etc.)	0.00	30.00	9.48	8.72
4	Preparing data products for network distribution (e.g. PASTA and climDB/hydroDB)	0.00	30.00	7.78	8.31
5	Populating network databases (e.g. PersonnelDB, BiblioDB, and SiteDB)	0.00	15.00	3.74	4.01
6	Developing and maintaining the site web presence	0.00	45.00	11.52	9.37
7	Site LTER Administrative support (annual report prep, etc.)	0.00	10.00	2.74	2.58
8	Assisting researchers with data analysis (data mining,	0.00	20.00	4.85	5.10

	preparation, quality control, statistical analysis, or model development) prior to the publication process				
9	System administration	0.00	30.00	7.44	7.02
10	End user technical support	0.00	10.00	2.89	3.14
11	Other site tasks (e.g. field work, sensor deployments, instrument maintenance, publishing etc.)	0.00	30.00	6.22	6.82
12	Other Network tasks (e.g. IM Exec, working groups, etc.)	0.00	20.00	6.78	5.71
13	Unknown	0.00	100.00	7.56	26.66

4. 4. Which statements describe your site's approaches to information management (choose as many as apply)

#	Answer		Response	%
1	a. The Information Manager or IM team determines priority of tasks within the framework of project needs, data generation, and other demands.		22	81%
2	b. IM tasks follow a strict schedule determined by field work in which the Information Manager or IM		3	11%

	team is intimately involved.			
3	c. The Site Executive team and Information Manager or IM team work closely to determine long- and short-term information management priorities.		13	48%
4	d. The Site Executive team looks to the Information Manager or IM team for guidance in terms of best practices and implementation of new technologies for site data management.		18	67%
5	e.I would like to provide additional comments regarding our site's approach to information management.		2	7%

5. 4a. Please provide any additional comments relevant to your site's approach to information management below

Text Response
Many routine tasks related to network databases are automated so they require little attention
It's unclear to me at present how our site organizes this work.

6. 5. Does the Site Information Manager regularly attend meetings of the site's Executive Committee or equivalent body?

#	Answer		Response	%
1	YES		19	70%
2	NO		8	30%
	Total		27	100%