

ATER Cyberinfrastructure Site Assessment Survey - Raw Analysis

This survey is designed to assess the current state, and cyberinfrastructure (CI) needs for LTER Network level science being developed as part of the LTER Planning Process. Components of the survey were developed by a CI Team which met in June to initiate the CI needs assessment. Further information can be found in the Cyberinfrastructure section of the LTER Planning Grant Wiki at: <http://intranet.lternet.edu/planning/>

This assessment covers the LTER sites only. The LTER Network Office was assessed separately from the sites.

This survey was written and processed by [John Vande Castle](#), LTER Network Office, using survey tools of "SurveyMonkey.com" (This is neither an endorsement or recommendation by the LTER Network Office). Special thanks are due to the members of the LTER Technology Committee and LTER Information Managers who reviewed drafts of the survey, and especially to ALL the LTER Information managers who took the time to fill out the survey. This survey represents a response from sites of the entire LTER Network, again thanks to the LTER Information Managers.









This survey is meant to cover the following topic areas identified by the LTER CI team, although some questions were added at the request of site information managers or members of the LTER Technology Committee:

- General overview of data collection, QA/QC, archive, delivery
- EML/metadata completeness
- GIS/map services
- Instrumentation - sensors
- Remote sensing infrastructure, data storage
- Analytic tools
- Models
- Visualization
- Number of IM FTE's – break down by type, expertise
- Institutional/site support – email, admin, licensing
- Network bandwidth – internal-external
- Storage capacity, backup solutions
- Database formats/ organization/systems
- Partnerships
- Server architecture
- Computational capacity – models, stats, database
- Collaboration tools – Wikis, vtc, email, file sharing

NOTES: Some sites entered data from multiple sessions, as reflected in some skipped questions, but the total response was from all 26 LTER sites.

NOTES: Some skipped questions indicate the question doesn't apply to the site, or the site does not use or have any of the information requested.

Please enter YOUR LTER roles, duties or committee affiliations (i.e. of the person filling out this survey):

	Response Percent	Response Total
LTER Site Principal Investigator 	6.9%	2
LTER Research Scientist 	17.2%	5
LTER Site Manager	0%	0
LTER Site Administrator	0%	0
LTER Information Manager 	89.7%	26
LTER Climate Committee 	3.4%	1
LTER Technology Committee Representative 	17.2%	5
LTER Network Information System Advisory Committee 	10.3%	3
LTER Education Committee Representative 	3.4%	1
Other (please specify) 	10.3%	2
Total Respondents		29
(skipped this question)		0

OTHER:

- 1. LTER Information Management Executive Committee
- 2. Senior Research Assistant

3. About how many total FTE's does your site allocate for information management from the LTER base funding? In other words, of the annual funding the site receives from NSF, how many FTEs are funded from this amount.

	Response Percent	Response Total
0	0%	0
0.25	0%	0
0.5	19.2%	5
0.75	11.5%	3
1.0	23.1%	6
1.25	3.8%	1
1.5	3.8%	1
1.75	11.5%	3
2.0	11.5%	3
2.5	7.7%	2
3.0	3.8%	1
3.5	0%	0
4.0	0%	0
4.5	0%	0
5.0	0%	0
5.5	0%	0
6.0	0%	0
6.5	0%	0
7.0	0%	0
7.5	0%	0
8.0	0%	0
8.5	0%	0
9.0	0%	0
9.5	0%	0
10.0	0%	0
Great than 10.0	0%	0
Total Respondents		25
(skipped this question)		5

4. About how many total FTE's does your site allocate for information management from non- LTER base funding? In other words, about how many FTE's are there at the site that are not directly funded from LTER funds. This would include cost-share, volunteers and what have you. For example, if in an average year, 0.75 FTE worth of an IM is funded from NSF LTER core funds, but there are really 5 FTEs working at the site from some sort of funds or "personal time donations" then you would want to put "4" here (when id doubt, round to the nearest value).

	Response Percent	Response Total
0	36%	9
0.25	12%	3
0.5	8%	2
0.75	8%	2
1.0	8%	2
1.25	0%	0
1.5	4%	1

1.75	4%	1
2.0	12%	3
2.5	8%	2
3.0	0%	0
3.5	0%	0
4.0	0%	0
4.5	0%	0
5.0	0%	0
5.5	0%	0
6.0	0%	0
6.5	0%	0
7.0	0%	0
7.5	0%	0
8.0	0%	0
8.5	0%	0
9.0	0%	0
9.5	0%	0
10.0	0%	0
Great than 10.0	0%	0
Total Respondents		25
(skipped this question)		5

5. Of the site FTE's for information management, rate from low to high, the level of expertise in the following areas:

	High	Medium	Low	Response Total
Formal training or education in computer science.	8% (2)	46% (12)	46% (12)	26
Acquired (on the job training) computer science training.	46% (12)	50% (13)	4% (1)	26
Formal training or education in data management/database software.	12% (3)	42% (11)	46% (12)	26
Acquired (on the job training) in data management/database software.	62% (16)	35% (9)	4% (1)	26
Formal Ecological/Biological Science training or education.	73% (19)	23% (6)	4% (1)	26
Formal GIS/Remote Sensing training or education.	35% (9)	27% (7)	38% (10)	26
Social Science training or education.	0% (0)	19% (5)	81% (21)	26
Total Respondents				25
(skipped this question)				4

6. What is the GENERAL information management task allocation at your site ? (Please have the total allocation = 100% - the survey software is too dumb to do this...). This question will take some thought to answer since it covers the effort of any and all IM FTEs at an LTER site. It is meant to assess what type of effort goes in to various IM "tasks" at an LTER site. For example, if you select 50% for General site data management, the rest of the tasks can only add up to the remaining 50%...

	0	5	10	15	20	25	30	35	40	45	50	55	60	65	70	75	80	85	90	95	100	Response Total
General site data management including data entry and access.	0% (0)	0% (0)	15% (4)	8% (2)	4% (1)	12% (3)	19% (5)	0% (0)	4% (1)	0% (0)	19% (5)	0% (0)	12% (3)	4% (1)	4% (1)	0% (0)	0% (0)	0% (0)	0% (0)	0% (0)	0% (0)	26

Data archive and backup.	4% (1)	69% (18)	23% (6)	0% (0)	0% (0)	0% (0)	4% (1)	0% (0)	0% (0)	0% (0)	0% (0)	0% (0)	0% (0)	0% (0)	0% (0)	0% (0)	0% (0)	0% (0)	0% (0)	0% (0)	0% (0)	26
Site System Administration (site-based hardware and network support).	4% (1)	23% (6)	46% (12)	8% (2)	4% (1)	12% (3)	4% (1)	0% (0)	0% (0)	0% (0)	0% (0)	0% (0)	0% (0)	0% (0)	0% (0)	0% (0)	0% (0)	0% (0)	0% (0)	0% (0)	0% (0)	26
User System Support (hardware support for site personnel).	15% (4)	54% (14)	23% (6)	0% (0)	8% (2)	0% (0)	0% (0)	0% (0)	0% (0)	0% (0)	0% (0)	0% (0)	0% (0)	0% (0)	0% (0)	0% (0)	0% (0)	0% (0)	0% (0)	0% (0)	0% (0)	26
User Support (help for site personnel such as data retrieval, study design, statistics, etc.).	0% (0)	31% (8)	35% (9)	27% (7)	8% (2)	0% (0)	0% (0)	0% (0)	0% (0)	0% (0)	0% (0)	0% (0)	0% (0)	0% (0)	0% (0)	0% (0)	0% (0)	0% (0)	0% (0)	0% (0)	0% (0)	26
Network and cross site IM support.	8% (2)	50% (13)	35% (9)	4% (1)	4% (1)	0% (0)	0% (0)	0% (0)	0% (0)	0% (0)	0% (0)	0% (0)	0% (0)	0% (0)	0% (0)	0% (0)	0% (0)	0% (0)	0% (0)	0% (0)	0% (0)	26
Software development.	27% (7)	19% (5)	15% (4)	8% (2)	8% (2)	15% (4)	8% (2)	0% (0)	0% (0)	0% (0)	0% (0)	0% (0)	0% (0)	0% (0)	0% (0)	0% (0)	0% (0)	0% (0)	0% (0)	0% (0)	0% (0)	26
Site administrative tasks (filling out paperwork, and doing things like answering surveys).	12% (3)	73% (19)	12% (3)	4% (1)	0% (0)	0% (0)	0% (0)	0% (0)	0% (0)	0% (0)	0% (0)	0% (0)	0% (0)	0% (0)	0% (0)	0% (0)	0% (0)	0% (0)	0% (0)	0% (0)	0% (0)	26
Other tasks.	29% (5)	35% (6)	18% (3)	18% (3)	0% (0)	0% (0)	0% (0)	0% (0)	0% (0)	0% (0)	0% (0)	0% (0)	0% (0)	0% (0)	0% (0)	0% (0)	0% (0)	0% (0)	0% (0)	0% (0)	0% (0)	17
Total Respondents																						25
(skipped this question)																						4

7. What major information management support, and level is provided by the home/host institution(s) of your site? Select any that apply.







	None	Some	All	Response Average
Email (use the institution's email system)	4% (1)	19% (5)	77% (20)	2.73
Database (use the institution's database system)	85% (22)	15% (4)	0% (0)	1.15
Computational Infrastructure (use the institution's computational infrastructure for analysis, statistics, modeling, etc.)	58% (15)	42% (11)	0% (0)	1.42
System administration is provided by the institution	35% (9)	62% (16)	4% (1)	1.69
Technical support is provided by the institution	19% (5)	77% (20)	4% (1)	1.85
Institutional (site) software licensing or educational discount	0% (0)	85% (22)	15% (4)	2.15
Total Respondents				25
(skipped this question)				4

8. In contrast to the direct support provided by the home/host institution(s) of your site, what does the site support by itself?

	None	Some	All	Response Average
Email (the site uses its own email system)	81% (21)	15% (4)	4% (1)	1.23
Database (the site uses its own database system)	0% (0)	23% (6)	77% (20)	2.77
Computational Infrastructure (use the site's computational infrastructure for analysis,	4% (1)	46% (12)	50% (13)	2.46

statistics, modeling, etc.)				
System administration is handled by the site itself.	0% (0)	58% (15)	42% (11)	2.42
Technical support is handled by the site itself.	0% (0)	69% (18)	31% (8)	2.31
The site itself has arranged special software licensing or educational discounts.	38% (10)	58% (15)	4% (1)	1.65
			Total Respondents	25
			(skipped this question)	4

9. What type of collaboration tools are used at your site (select all that apply, and enter any other not listed here)?

		Response Percent	Response Total
Email is used extensively.		104%	26
Common filesharing is used (such as a shared file system for users at the site).		92%	23
Telephone and video conferencing.		60%	15
Web Calendars.		52%	13
Collaborative web tools such as wikis are used at the site.		32%	8
Other (please describe)		12%	3
		Total Respondents	25
		(skipped this question)	4

Other:

1. scheduled meetings; web applications for requesting or reserving resources and equipment
2. Intranet, Online collaboration, IM "chat"
3. Monthly meetings

10. What type of collaborative cyberinfrastructure / information management partnerships (outside of LTER) is your site engaged in?

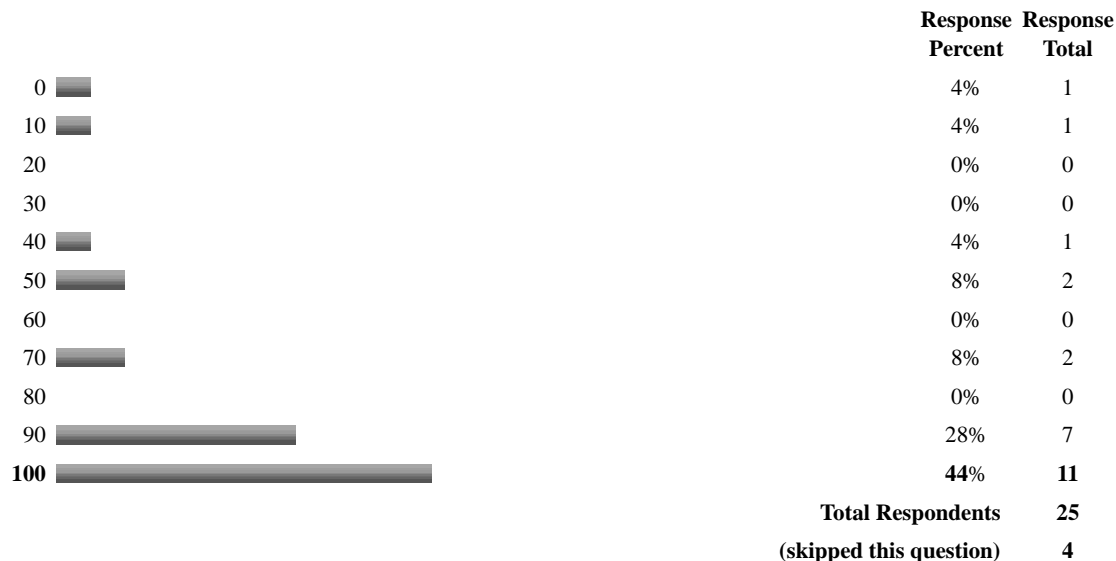
		Response Percent	Response Total
The site maintains an active collaboration with high performance computer centers (SDSC, NCSA, etc).		20%	4
The site maintains an active collaboration with NASA archive centers.		20%	4
The site maintains an active collaboration with USGS centers such as NBII.		30%	6
Other (please describe)		70%	14
		Total Respondents	20
		(skipped this question)	9

Other:

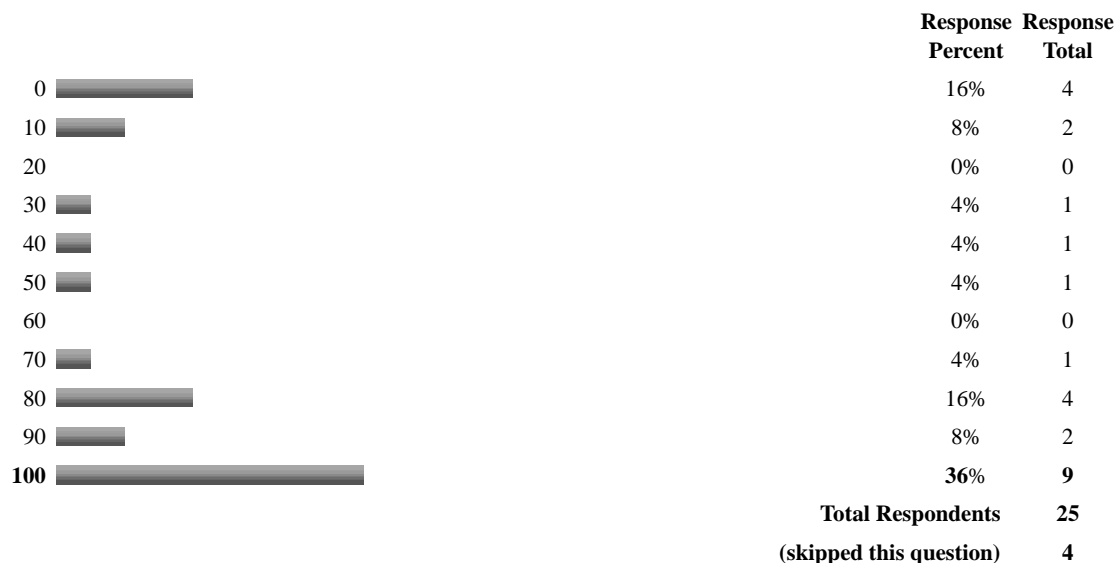
1. Marine Consortiums and Digital Library
2. Marine consortiums, Digital Library
3. The Arctic System Science Data Coordination Center
4. Collaborate with NASA to validate numerous satellite instrumentations for many years. Hyperspectral imagery (aerial-based and now satellite-based) is an example of one of these collaborations. Also, collaborating with the Canopy Project (Evergreen State College) to explore their implementations of database templates with associated analytical and visualization tools.
5. National Center for Ecological Synthesis and Analysis (NCEAS) Partnership for the Interdisciplinary Study of Coastal Oceans (PISCO)

- 6. The site maintains an active collaboration with the Canopy Databank Project at The Evergreen State College; The site maintains an active collaboration with the LTER Network Office by harvesting metadata to their metacat and contributing to the development of the McCube (metadata database system); The site maintains an active collaboration with ARS and USFS to manage spatial data about prairie dogs and other features
- 7. The site collaborates with other coastal oceanography projects (mainly PISCO) and with consults with NCEAS
- 8. Ameriflux North American Pollen Database
- 9. local host institution
- 10. NOne

11. What percentage of all site data has corresponding structured metadata of any type, including EML?



12. What percent of all site metadata has been converted to EML to at least the "identification" (base) level?



13. What percent of all site metadata has been converted to EML to the "discovery" level or beyond?



60	0%	0
70	0%	0
80	16%	4
90	8%	2
100	32%	8
Total Respondents		25
(skipped this question)		4

14. What percent of all site metadata has been converted to EML to the "integration" level or beyond?




	Response Percent	Response Total
0	44%	11
10	8%	2
20	0%	0
30	4%	1
40	8%	2
50	0%	0
60	0%	0
70	8%	2
80	16%	4
90	4%	1
100	8%	2
Total Respondents		25
(skipped this question)		4

15. Of the known site historical/legacy data - i.e. data the site might not consider part of its standard research data, what percentage has corresponding EML metadata (at any level)?

	Response Percent	Response Total
0	34.8%	8
10	13%	3
20	0%	0
30	0%	0
40	0%	0
50	8.7%	2
60	4.3%	1
70	13%	3
80	8.7%	2
90	4.3%	1
100	13%	3
Total Respondents		23
(skipped this question)		6

16. What is the general way research data are managed at your site? Select all that apply.






	Response Percent	Response Total
The site information manager(s) enter and manage most site data.	76%	19
Researchers at the site enter most of their data into the site	44%	11

information management system.			
Student research data (i.e. thesis data) are generally included in the site information management system.		64%	16
Standard procedures are in place for use by researchers to enter and manage their data.		56%	14
Other (please describe)		24%	6
		Total Respondents	25
		(skipped this question)	4

Other:

1. technicians and field crew enter some data
2. Information manager interacts with researchers to complete the metadata. We get metadata fully completed and we also fully complete metadata by communicating with researchers. We are currently developing applications to foster the former to reduce IM involvement.
3. inclusion of student research data is planned for 2006
4. Standard procedures are in place for use by researchers to document their data in the format required by information managers.
5. The data are entered by technicians, workstudy people, RAs and students. They are managed by the information manager.
6. Data entry templates are filled out by researchers and technicians, then the metadata and data are standardized and uploaded to the database by the information manager

17. What type of QA/QC procedure does your site follow for site data? Select all that apply.

		Response Percent	Response Total
The site has documented specific QA/QC procedures of its own.		68%	17
The site follows specific QA/QC guidelines (e.g. EPA, USGS etc).		12%	3
QA/QC guidelines are followed for MOST site data.		60%	15
QA/QC guidelines are not necessarily followed for all site data (such as student thesis data).		40%	10
Other (please describe)		24%	6
		Total Respondents	25
		(skipped this question)	4

Other:

- Researchers do there own QA/QC
2. The quality of data are the responsibility of the responsible investigator. The information management staff facilitates the generation and maintenance of the metadata, and often the data, as well. Data QA/QC is performed by the researcher, research assistants, or information management staff depending on a case-by-case basis. The IM staff performs QA/QC when responsible for actual data entry; otherwise, the information management staff facilitates QA/QC when requested depending on the relative value of the research to site research initiatives and resources available.
 3. The site has a fortran program to perform QA/QC for most of the long-term data sets
 4. Researchers are currently responsible for their own QA/QC
 5. We have implemented a comprehensive rule-based QA/QC framework for all tabular data. This framework supports domain limit checks, sanity checks, complex algorithmic analyses and manual flagging of values. QA/QC criteria can be defined for any attribute and are stored in metadata templates for application to raw data. QA/QC rules are also automatically reapplied whenever data is changed or added during interactive analysis.
 6. Researchers are responsible for the QA/QC for the data they provide

18. What is the general "delivery" procedure for data at your site? Select ALL that apply.

Response Response

	Percent	Total
Most site data are online and freely available.	92%	23
A "data license" or data use agreement is required for data use.	68%	17
Some site data are online although most data requests are filled by an information manager.	12%	3
Most site data are provided to requesters by an information manager.	4%	1
Other (please specify)	12%	3
Total Respondents		25
(skipped this question)		4

Other:

1. Currently, the compliance with the site data policies are voluntary. We are currently developing applications to make it 'mandatory' in the next year. All public datasets are available online. All completed metadata is available online. Access to restricted data is by request to the information manager and responsible investigator. Access to GIS/remote sensing data is currently by request to our GIS/Remote Sensing Specialist. We are developing applications that will make public GIS/remote sensing data available online. Restricted access GIS/remote sensing data will still require a request to the GIS/Remote Sensing Specialist.
2. Some sensitive data are stored offline and are provided to requesters by an information manager with permission from the principal investigator.
3. Some data that are not available online are provided by the IM

19. On-line site data are provided through the following mechanisms: Select all that apply.




	Response Percent	Response Total
The site website provides direct access to data.	100%	25
Site data are available though an "ftp" mechanism.	16%	4
Site data are generally managed as "flat files" in an organized file system.	56%	14
Site data are managed through a database system such as MySQL or Oracle.	64%	16
Site data are registered in the Metacat server.	56%	14
Other (please specify)	12%	3
Total Respondents		25
(skipped this question)		4

Other:


1. Site data will be registered on the LNO Metacat server VERY SOON.
2. The site has made initial steps to create EML and register it in the metacat. 10% of the datasets are registered and the site is working on the rest.
3. Data files can be remotely accessed via HTTP and analyzed or integrated using MATLAB client software (freely available on the GCE web site)

20. For GIS data maintained at the site (select all that apply):

	Response Percent	Response Total
Most site GIS data are provided online.	65%	13

Most site GIS data have corresponding EML metadata.		25%	5
Internet map services are used at the site for DISPLAY of GIS data.		70%	14
Internet map services are used at the site for ACCESS to GIS data.		25%	5
		Total Respondents	20
		(skipped this question)	9







21. For site GIS data, how are remote sensing data managed ? Select all that apply.

		Response Percent	Response Total
Remote Sensing Data are primarily managed by the site Information Manager.		31.8%	7
Remote sensing data are generally managed by researchers rather than site information managers.		50%	11
Most remote sensing data (for instance, original Landsat data) are provided online.		18.2%	4
Products generated from remote sensing data are provided online.		36.4%	8
Other (please specify)		27.3%	6
		Total Respondents	22
		(skipped this question)	7

Other:

1. Remote sensing data are primarily managed by the LTER GIS Specialist.
2. Managed by field station personnel.
3. Most remote sensing data are available by request from our GIS/Remote Sensing Specialist. Some RS data are available from researchers. None of our RS holdings are available online, although we plan on distributing them online in the next year.
4. A half-time GIS specialist manages GIS data.
5. Some remote sensing data are provided on-line
6. Remote sensing data are managed by the remote sensing lab on campus

22. What archive and backup procedures are used at your site (select all that apply) ?








		Response Percent	Response Total
Active site data are maintained on fault-tolerant (i.e. RAID) systems.		48%	12
The site uses a documented archive and backup plan.		72%	18
Site data are archived using tape backup systems.		76%	19
Site data are archived using CD or DVD backup systems.		32%	8
Site data are archived using mirrored disk systems.		32%	8
Site data are archived using remote data archive facilities.		12%	3

Data backup includes off-site/secure storage.		88%	22
Other (please describe 		12%	3
		Total Respondents	25
		(skipped this question)	4

Other:

1. Online remote sensing data are maintained on a RAID system; offline remote sensing data are on CD.
2. Data is also archived on external USB/FireWire hard drives, magneto-optical (MO) drives, and CD/DVD media which are stored on-site and off-site at multiple locations on and off-campus. If LNO offers the off-site storage of boxes of media discussed a couple of years ago, we would take advantage of that service.
3. Site data are archived using remote external hard-drives.

23. How are routine meteorological data collected/managed at the site (select all that apply)?









		Response Percent	Response Total
Meteorological data are manually collected by observing static instrument readings.		29.2%	7
Meteorological data are collected by interpreting paper, strip-chart or other information.		20.8%	5
Meteorological data are automatically collected by a digital or digitized hard-wired system.		25%	6
Meteorological data are collected by automated data logger systems, and later downloaded.		75%	18
Meteorological data are collected by automated wireless phone (i.e. cell phone) system.		4.2%	1
Meteorological data are collected by automated radio or wireless transmission and collected automatically.		45.8%	11
Other (please specify) 		20.8%	5
		Total Respondents	24
		(skipped this question)	5

Other:

1. Automatic from Buoy and shipboard
2. Our meteorological data are downloaded from the National Climatic Data Center (NOAA) and formatted to the Climdb standard.
3. Data collected from datalogger via modem and posted online every 15 minutes
4. We are in the process of installing wireless transmission stations.
5. Meteorological data are collected by automated microwave transmission (USGS) and collected automatically

24. What type of GPS location information is maintained for the LTER site? Please select all that apply, and add any information not listed here.






		Response Percent	Response Total
The primary research site locations are maintained in a file or database.		87.5%	21

The primary research site locations are publicly available on the site webpage.		70.8%	17
High precision control points are established for reference at the LTER site.		45.8%	11
The LTER site is mapped with a consistent grid of GPS locations.		33.3%	8
Most research data are not generally stored with GPS location information.		16.7%	4
Most research data are collected with GPS location information.		66.7%	16
Most research data are collected with GPS location information accurate to better than 15m.		37.5%	9
GPS location information is required for all research data.		29.2%	7
Other (please describe)		16.7%	4
		Total Respondents	24
		(skipped this question)	5

Other:

1. Uses shipboard GPS
2. Our site requests that all GPS data used in the information system be collected at sub-meter accuracy using a standard datum and coordinate system.
3. Historical research data site information may or may not be collected with GPS.
4. Not all research locations are publicly available because of human subject research restrictions, but are available upon request.

25. What type of GPS equipment is available for use at the site (select all that apply).

		Response Percent	Response Total
A high precision GPS base station (for differential correction) is maintained or is available for use at the LTER site.		25%	6
Differential GPS (DGPS) equipment (> 3m accuracy) is available for use at the site.		50%	12
Wide Area Augmentation System (WAAS) enabled (> 5m accuracy) GPS receivers are available for use at the site.		20.8%	5
Conventional GPS receivers (15m accuracy) are generally available for use at the site.		79.2%	19
The site does not maintain GPS receivers for research use.		0%	0
Other (please specify)		16.7%	4
		Total Respondents	24
		(skipped this question)	5

Other:

1. A high precision GPS base station is available in the Madison area
2. The host institute maintains the base station on-campus. Most GPS data are collected at sub-meter accuracy after post-processing correction.
3. We have borrowed UNAVCO high precision Trimble GPS through LTER/UNAVCO agreement
4. DGPS equipment is installed on all research vessels, but the site does not maintain receivers for land-based studies (some investigators supply their own)



26. What type of sensor systems are routinely used for data collection at the site (Please describe)?

Total Respondents 19
(skipped this question) 10

Other:


1. Thermistor chains for water temperature profiles, ADCP, Dissolved oxygen, meteorological measurements (air temperature, precip, relative humidity, wind speed and direction, PAR, long and short wave radiation)
2. Automated shipboard systems
3. Profiling instruments deployed over-the-side; ship lab flow through systems
4. Data Loggers for meteorological data, stream and lake depth, light, conductivity and temperature.
5. Not clear what is meant by 'sensor system.'
6. Standard oceanographic instrument packages, ADP, ADCP, CTD, thermistors, wave, tidal height recorders, all with internal data recorders
7. Eddy Co-Variance Tower; Bowen ratio Tower; TDR Probes; Radiometer
8. JPL Sensor Web 3.2 -- sensor network in which all sensors are talking to each other simultaneously. This system is being used to monitor microclimates under three shrub species.
9. None to date
10. Automated meteorological station Eddy flux towers (3) Portable PAR, CO2, etc stations

27. Wireless internet is available for researchers at the site.

	Response Percent	Response Total
Yes 	73.9%	17
No 	26.1%	6
	Total Respondents	23
	(skipped this question)	6




28. The site has installed wireless internet for automated data collection at the site (this would include radio data transmission that is eventually linked directly to the internet). An example is the wireless network that was installed at the VCR LTER site. If, so please BRIEFLY append a description to question #34

	Response Percent	Response Total
Yes 	41.7%	10
No 	54.2%	13

If present, please describe in "other" of question #34 

Total Respondents 24
(skipped this question) 5

29. What primary SERVER architecture is used for SITE data management (select all that apply) ?

	Response Percent	Response Total
MS Windows-based systems 	70.8%	17
Mac-based systems (OS X Tiger, etc) 	8.3%	2
Linux-based systems 	33.3%	8
Other Unix (SunOS etc) based systems 	33.3%	8
Other (please describe) 	4.2%	1

Total Respondents 24
(skipped this question) 5

Other:

1. Netware

30. What general online data storage capacity does your site maintain for site data (round to nearest power of ten - i.e. 4tb would select 1tb, 6tb would select 10tb)? Please ALSO describe more complex storage systems.

	Response Percent	Response Total
1gb	0%	0
10gb	8%	2
100gb	36%	9
1tb	48%	12
10tb	0%	0
100tb	0%	0
>100tb	0%	0
Other (please describe)	8%	2
Total Respondents		25
(skipped this question)		4

Other:

1. 4 TB
 2. 600gb

31. What is the primary link speed from the Home Institution(s) to the Internet? In other words, what kind of Internet connection does your primary site institution, university etc. have? For multiple institution sites, select what most have, or what the primary site institution has.

	Response Percent	Response Total
Internet link is greater than 1 GB/s	28%	7
Internet link is 1Gb/s.	32%	8
Internet link is 100mb/s	32%	8
Internet link is 10mb/s	0%	0
Internet link is less than 10mb/s	8%	2
There is no internet connection.	0%	0
Total Respondents		25
(skipped this question)		4

32. What is the primary Internet Link at the Research SITE ?

	Response Percent	Response Total
Internet link is greater than 1 GB/s.	4.2%	1
Internet link is 1Gb/s.	8.3%	2
Internet link is 100mb/s	20.8%	5
Internet link is 10mb/s	20.8%	5
Internet link is less than 10mb/s	41.7%	10
There is no internet connection.	4.2%	1

Total Respondents 24
 (skipped this question) 5

33. What type of internet capability is available for researchers at the Home Institution(s)? Select any that apply.

	Response Percent	Response Total
Internal (local area) connections are 1gb/s or better	32%	8
Internal (local area) connections are 100mb/s	64%	16
Internal (local area) connections are 10mb/s or slower	16%	4
External (wide area) internet connections are BETTER 10mb/s	20%	5
External (wide area) internet connections are 10mb/s or less	12%	3
Wireless internet network is available	52%	13
Other (please specify)	8%	2
Total Respondents	25	
(skipped this question)	4	

Other:

1. Servers and some researchers are interconnected at 1 GB/s
2. home institution = field site

34. What type of internet bandwidth is available at the research SITE (select all that apply)?

	Response Percent	Response Total
Internal (local area) connections are 100mbs or better.	45.8%	11
Internal (local area) connections are 10mbs or slower.	33.3%	8
Wide area internet connections are better than T1 speeds.	4.2%	1
Wide area internet connections are T1 speeds or less.	50%	12
Wide area internet connections are generally not available on site.	0%	0
Wireless internet network is available.	54.2%	13
Wireless internet network is used to collect site data.	25%	6
Other (please describe)	33.3%	8
Total Respondents	24	
(skipped this question)	5	

Other:

1. RF radio modem to internet for weather data
2. A Moxa DE-302 is used as serial-to-Ethernet, providing the connectivity between the base station (RF400) and T1 line. There are 6 remote sites, 1 weather station and 5 hydrology, up to 8 miles range.
3. wireless is available at the field station main building only via WIFI
4. Wireless backbone consists of two Tringolink10 EXT Bridge Radios. These radios connect to 802.11b access points, and the clients communicate with these access points.
5. None
6. home institution = field site
7. dial-up speeds available at permanent field camps
8. nearly complete: the wireless system connects a ctd mooring to a windows xp computer in the laboratory. process data is mapped to the web server.
- 9.A 900 Mhz IP radio extends 22 km to Broadwater Tower. An amplified 802.11b access point there provides a link to Machipongo Station 8 km to the north that has its own amplified AP. Jointly they provide communications to sensors and web cameras.

35. What type of conferencing capability is available at the site's home institution(s)? Select all that apply.

	Response Percent	Response Total
Local (on-site) phone conferencing is available.	100%	25
Shared phone conferencing capabilities are available at the institution.	68%	17
Local or shared ISDN video teleconferencing.	32%	8
Voice over Internet (such as Skype).	36%	9
Local internet video conferencing (such as Polycom video).	52%	13
Shared internet video conferencing (such as Polycom video) is available at the institution.	36%	9
Local satellite video conferencing.	8%	2
Shared satellite video conferencing is available at the institution.	16%	4
Local advanced video conferencing (such as ACCESS Grid) is available i.e. the site has direct access or maintains an ACCESS Grid system.	0%	0
Shared advance video conferencing (such as ACCESS Grid) is available somewhere at the institution.	4%	1
Other (please describe)	0%	0
	Total Respondents	25
	(skipped this question)	4

36. What type of conferencing capability is available and used at the research SITE (select all that apply)?







	Response Percent	Response Total
Phone conferencing.	95.8%	23
Internet video conferencing.	50%	12
Satellite video conferencing.	8.3%	2

Advanced video conferencing (such as ACCESS Grid).	0%	0
Other (please specify) 	4.2%	1
Total Respondents		24
(skipped this question)		5

Other:

- 1. None

37. What type of computational capabilities does your site provide to researchers for data analysis - modeling, statistical analysis and data synthesis? Select all that apply.

	Response Percent	Response Total
Investigators use their personal systems for analysis. 	92%	23
The university/home institution provides most computational support for data analysis. 	36%	9
A dedicated system is supported by the site for data analysis. 	24%	6
A local computer cluster is available to researchers for data analysis. 	20%	5
A remote computer cluster is available to researchers for data analysis. 	8%	2
A direct link to high performance computer center is available to researchers for data analysis. 	8%	2
Other (please describe)	0%	0
Total Respondents		25
(skipped this question)		4

38. What PRIMARY database system is used at your site for METADATA? Select any that apply.

	Response Percent	Response Total
SQL Server 	32%	8
Oracle 	8%	2
MySQL 	24%	6
Ingress	0%	0
Postgres 	8%	2
Paradox	0%	0
Other (please specify) 	48%	12
Total Respondents		25
(skipped this question)		4

Other:

- 1. None
- 2. Microsoft Access
- 3. Text files
- 4. My EML Metadata is stored as a static EML file in a folder. Some small amount of metadata is extracted from EML and input into my Oracle database.

- 5. Metadata currently managed in EML. Plan to migrate to MySQL next year.
- 6. Visual Foxpro
- 7. No database system, flat files
- 8. Metadata are not currently in a database. We plan to use MySQL in the future
- 9. Metadata is only in html format
- 10. db2
- 11. with metacat
- 12. MiniSQL, but we are moving towards MySQL

39. What PRIMARY database system is used at your site for DATA? Select any that apply.

	Response Percent	Response Total
SQL Server	28%	7
Oracle	12%	3
MySQL	36%	9
Ingress	0%	0
Postgres	4%	1
Paradox	0%	0
Other (please describe)	56%	14
Total Respondents		25
(skipped this question)		4

Other:

- 1. none
- 2. dBase. Microsoft SQL Server is the main repository for research data and dBase files are used for GIS data analysis. The dBase databases are being migrated to SQL Server.
- 3. flat ASCII files
- 4. Microsoft Access
- 5. IBM DB2
- 6. Data is available as a flat file and can be queried in an Oracle database.
- 7. Data currently managed as flat files. Plan to migrate selected datasets to MySQL over next 1-2 years.
- 8. Visual Foxpro
- 9. No database system, flat files
- 10. flat files are pointed to by metadata
- 11. db2
- 12. MATLAB data structures (support database functionality via programmatic interface)
- 13. ascii text
- 14. For archival purposes, data is stored in ASCII text files.

40. Does your site use any CASE tools (database design tools) for database management? Select any that apply.

	Response Percent	Response Total
None	75%	18
xCase	4.2%	1
ERStudio	0%	0
Rational Rose	0%	0
Oracle Designer	0%	0
ERwin	0%	0
Microsoft Visio	8.3%	2
Other (please specify)	16.7%	4
Total Respondents		24
(skipped this question)		5

Other:

1. php admin, dbdesigner
2. phpadmin; dbdesigner
3. SQLServer
4. MySQL suite

41. Does your site currently use web services ?

	Response Percent	Response Total
Yes	32%	8
No	68%	17
Total Respondents		25
(skipped this question)		4

42. What basic analytic tools are used at your site (i.e. software)? Select any that apply, unless used rarely.

	Response Percent	Response Total
Microsoft Office Tools (Excel, etc.)	100%	25
Matlab	52%	13
Splus	16%	4
SAS	76%	19
SYSTAT	36%	9
SigmaStat	16%	4
SPSS	36%	9
R	24%	6
Other (please specify)	12%	3
Total Respondents		25
(skipped this question)		4

Other:

1. stata
2. mathematica
3. sigmaplot

43. Please select or enter MAJOR Analytic MODELS or MODEL TOOLS in use at the site (there is a good model registry at <http://eco.wiz.uni-kassel.de/ecobas.html>)

	Response Percent	Response Total
BASIN	5.9%	1
CENTURY	29.4%	5
COVER	0%	0
GARP	5.9%	1
GEM	5.9%	1
PNET	23.5%	4
RAMS	11.8%	2
REMM	0%	0
RIVMOD	0%	0
SAGE	0%	0
SAVANNA	0%	0
Other (please specify)	64.7%	11

Total Respondents 17
(skipped this question) 12

Other:

1. DYRESM
2. Marine Biological Laboratory's Multiple Element Limitation Model (MBL-MEL) Marine Biological Laboratory General Ecosystem Model (MBL-GEM) TOPMODEL-based hydrology model
3. ECOTONE SOILWAT
4. DISTANCE; SOILWAT; GEMRAMS; GDAY; DAYCENT; STEPPE
5. ELM -Everglades Landscape Model NSM -Natural Systems Model NYMAN -mangrove models HYMAN -mangrove models Simulation Model of Seagrass Communities in Florida Bay
6. Hydrologic Models
7. none
8. PALS, UGM-PHX, XPPAUT, MM5, urbanSim
9. ALFRESCO,TEM, STMTEM
10. HEC-RAS rating curves oc4v4
11. Individual investigators use a variety of modeling tools and approaches, most of them home grown.

44. What type of data visualization software tools (separate from GIS and statistical tools) are used at your site? Select any that apply, unless used rarely

	Response Percent	Response Total
OpenDX	0%	0
VisDB	0%	0
XmdvTool	0%	0
Spotfire	0%	0
Visionary	0%	0
NONE are used (site uses statistical and GIS packages)	95.8%	23
Other (please specify)	8.3%	2
Total Respondents	24	5
(skipped this question)		

Other:

1. 3D Nature Studio
2. Tecplot, Matlab

45. What type of GIS software tools are used at your site ? Select any that apply, unless used rarely.

	Response Percent	Response Total
ArcInfo	91.7%	22
ArcView	91.7%	22
ArcGIS	83.3%	20
Erdas Imagine	54.2%	13
IDRISI	8.3%	2
ENVI	12.5%	3
IDL	4.2%	1
GRASS	4.2%	1
ERMMapper	4.2%	1
None	0%	0
Other (please specify)	20.8%	5
Total Respondents	24	5
(skipped this question)		

Other:

1. custom developed software
2. eCognition Professional
3. ArcIMS ArcGIS Server
4. GHAAS-RGIS, University of New Hampshire river GIS system
5. Home grown program named CCGIS

46. What type of project management tools are used at your site ?

	Response Percent	Response Total
Microsoft Project	8.7%	2
Open Workbench	0%	0
MindManager	4.3%	1
None	82.6%	19
Other (please describe)	4.3%	1
Total Respondents (skipped this question)		23 6

Other:






1. Project management is accomplished with a database. Access to this database is through an intranet in which annual reports, publications, project descriptions, personnel, and proposals for new projects are entered.

47. If your site were to participate in a cross-site study that involved installation of instruments generating 10 MB (ten MEGAbytes) of streaming data per day from the field, what parts of your network infrastructure would REQUIRE upgrading in order to for you to automatically collect and store that data (select any that apply):







	Response Percent	Response Total
Network from Field to Laboratory (wireless).	79.2%	19
Network from Laboratory to Internet.	33.3%	8
Network at home institution(s).	8.3%	2
More data storage capacity at the field laboratory.	62.5%	15
More data storage capacity at the home institution(s).	54.2%	13
File servers at field laboratory.	33.3%	8
File servers at home institution(s).	20.8%	5
Nothing.	8.3%	2
Total Respondents (skipped this question)		24 5

48. NOW (read carefully, this is different!), if your site were to participate in a cross-site study that involved installation of instruments generating 10 GB (ten GIGAbytes) of streaming data per day from the field, what parts of your network infrastructure would REQUIRE upgrading for you to automatically collect and store that data (select any that apply):

	Response Percent	Response Total
Network from Field to Laboratory (wireless).	87.5%	21
Network from Laboratory to Internet.	62.5%	15

Network at home institution(s).		25%	6
More data storage capacity at the field laboratory.		83.3%	20
More data storage capacity at the home institution(s).		75%	18
File servers at field laboratory.		66.7%	16
File servers at home institution(s).		62.5%	15
Nothing.		0%	0
		Total Respondents	24
		(skipped this question)	5

49. If new cross-site experiments were to increase the volume of data or the number of datasets you are managing by a factor of 10, which of the following would MOST need to be increased at the site. (check ONLY up to 3)

		Response Percent	Response Total
Faster internet connection speed.		36%	9
Better, faster software for managing metadata.		40%	10
Better, faster software for managing data.		28%	7
A faster server(s).		28%	7
Disk space for data storage.		76%	19
More personnel.		76%	19
Nothing.		0%	0
		Total Respondents	25
		(skipped this question)	4