

The LTER Network Information System - a data cooperative for ecological science

The 2013 release of the Network Information System (NIS) makes 30+ years of data from the Long Term Ecological Research (LTER) Network accessible and usable for synthesis by the broader scientific community. Data producers and data consumers can now reliably interact within a single framework when working with LTER data. The NIS provides a secure repository, provenance chaining, DOI registration, and a pipeline to DataONE. The NIS is a key contribution to the development of a robust and efficient national cyberinfrastructure for ecological research. The LTER Network Office is exploring making this infrastructure available to other communities of data producers as part of a data cooperative (co-op).

Why did the LTER Network create an information system?

Sites in the LTER Network collect thousands of diverse data sets from a range of ecosystems. Access to these data advances collaborative and synthetic ecological science at multiple temporal and spatial scales. The goals of the LTER Network in expanding access to the NIS are to:

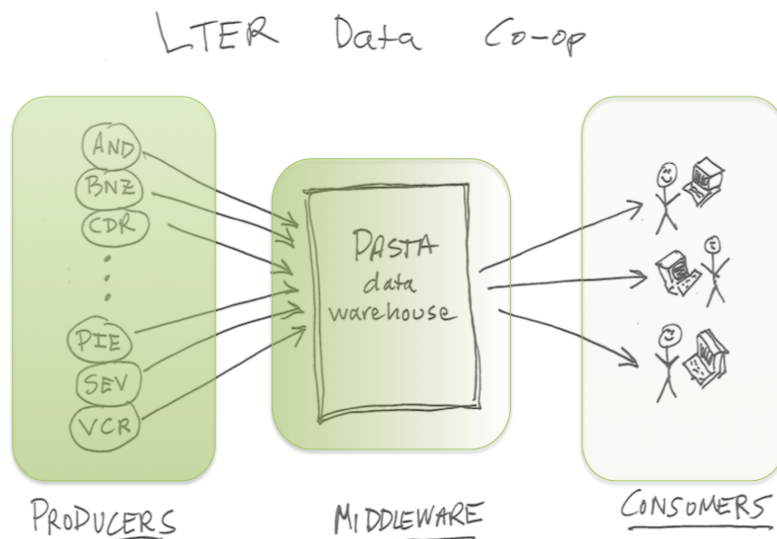
- provide access to data and metadata to a broad community of data consumers,
- improve the quality of ecological data,
- provide a persistent archive of ecological data,
- facilitate synthesis of ecological data into new, derived data, and
- provide access to ecological data through other community portals such as DataONE (<http://www.dataone.org/>).

How was the NIS developed?

The LTER community engaged in an intensive, long-term planning effort to identify requirements for managing ecological data. Professional information managers from each of the LTER sites defined the scope of the NIS in collaboration with LTER researchers. Software engineers at the LTER Network Office developed the cyberinfrastructure underlying the NIS with support from the National Science Foundation.

What does the NIS comprise?

The NIS comprises data producers at field sites that submit data packages (consisting of data and associated metadata); a community of data consumers including scientists, students, and educators; and software that connects data producers and data consumers (the Provenance Aware Synthesis Tracking Architecture or PASTA).



What is the PASTA software?

Based on a service-oriented architecture, PASTA is a scalable and extensible software framework that may be used for a broad spectrum of data-centric scientific projects. PASTA components interact across the Internet through an application programming interface (API) that uses RESTful web-services. The API also allows consumers to develop their own imaginative tools to manage and access data packages. Interfaces to R, Matlab, SAS, and SPSS are already available. The NIS Data Portal (<https://portal.lternet.edu>) uses a powerful and intuitive web-based interface for managing, discovering, and accessing data content in PASTA.

What does the NIS do?

- Uses the Ecological Metadata Language to document and describe data and to track the provenance of derived data,
- Assigns each submitted data package a *DataCite* registered Digital Object Identifier to facilitate data citation,
- Performs a comprehensive quality review of each submitted data package to ensure consistency between data and metadata,
- Integrates the LTER controlled vocabulary of terms and units to facilitate search and discovery,
- Creates linked versions of data packages as they are updated,
- Provides for automatic or manual submission of data packages,
- Meets requirements for open data access,
- Replicates data in DataONE via the LTER Member Node to ensure data persistence
- Makes data discoverable through the LTER, DataONE, and ONEMercury data portals.

Who can use the NIS?

The LTER Network makes data collected by LTER sites broadly accessible and open to other investigators, students, educators, and the general public. To date, there are 39,486 data packages in the NIS from LTER sites, the EcoTrends project, and independent projects funded by NSF's Long Term Research in Environmental Biology (LTREB) and Macrosystems Biology programs.

The LTER NIS and PASTA have all the necessary ingredients to support the data repository needs of other scientific data producers. LTER is creating a new "Data Co-op" through which other investigators can share their own data. Individuals or projects that provide data packages in appropriate formats can participate in the Data Co-op. Participation in the Data Co-op provides all of the advantages of the NIS.

Please contact James Brunt (jbrunt@lternet.edu) at the LTER Network Office for more details.