8.0 The LBA–Ecology Data and Information System

8.1 Preface

LBA is an international research initiative led by Brazil; as such, responsible Brazilian authorities must obtain, maintain, organize, and archive all data resulting from LBA activities in Brazil. To comply with this requirement, the Brazilian counterpart for each investigation will have a copy of all data and notes collected before an investigator leaves the country with the data. The LBA–Ecology DIS is a component of the overall LBA DIS (Figure 8.0), and, although the LBA DIS has not been fully described, the LBA Science Steering Committee (SSC) recommended the adoption of a distributed data search and retrieval system.[2] LBA–Ecology DIS will be compatible with the overall LBA DIS.

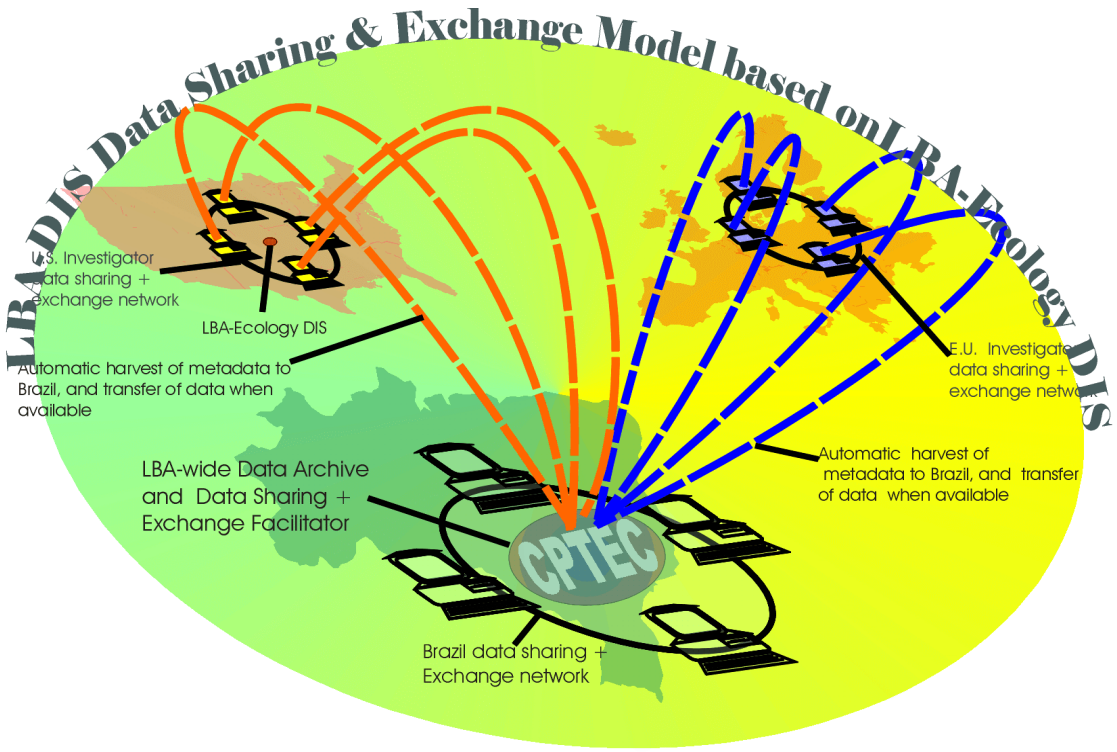

Figure 8.0

---

[2] *"In order to expedite the exchange of information among LBA investigators, the SSC recommends the adoption of concepts* for data searching, indexing, and retrieval presented by the DIS working group (WG). The SSC recognizes that the system described by the DIS WG is applicable to individual PI data sets as well as to larger products (e.g., satellite remote sensing and 4 DDA meteorological analyses). The SSC also recognizes that no single LBA module should support the entire cost of the implementing this data system, and, therefore the SSC recommends that the OIC consider the allocation of resources for implementation of this system across LBA modules. The SSC recognizes that the system presented by the DIS WG does not necessarily provide for all future functions of the LBA DIS.*
**[as footnote] The concepts presented include distributed data hosted by investigators, an Internet–based web–crawler to perform indexing, a Web site containing pointers to data that search functions, and software tools that will facilitate the tasks of the investigators."*

8.2 LBA–Ecology DIS

The LBA–Ecology DIS is a distributed system with a high level of investigator control and responsibility. The LBA–Ecology DIS is comprised of an Internet–based metadata search and data retrieval or delivery component and a general support component.

Five guiding principles used successfully in past efforts will be applied to the LBA–Ecology DIS:

- Minimize the time for data sets to become available to all investigators.
- Minimize requirements (time, hardware and software, human resources) for both data providers and data users.
- Maximize the ease of searching for data sets and retrieving them by users.
- Facilitate the preparation of value–added and integrated data sets.
- Develop and operate at a minimal cost to the NASA and INPE.

Additional considerations included a desire to develop a system that is fully compatible with general LBA Project data requirements and data policies and Brazilian law. Also, it was desired to provide a model or framework around which other funded modules could develop fully interoperable data and information systems if they so desired.

User support of the LBA–Ecology DIS is provided through Web sites at GSFC, ORNL, and INPE. GSFC maintains the dynamic operational project information and science implementation support functions (other than data distribution). The ORNL site mirrors both the Central LBA Web site and the GSFC LBA–Ecology Web site and provides all other information and necessary linkages to the overall LBA data and information system. The ORNL site also maintains some of the pre–LBA data sets and bibliographic information, static project documents, etc. INPE archives the original data.

The data archive function for NASA is primarily resident with the ORNL DAAC for biogeochemical Earth science data; however, LBA–Ecology will coordinate closely with the DAAC, primarily through the Project Office's ORNL DIS support activities. The ORNL LBA–Ecology DIS support will include a temporary data archive, as discussed below.

Data sets to be submitted to the ORNL DAAC for permanent archive will be selected by the Science Team, NASA Program Management, or other established LBA review panels. The Project Scientist will lead the activity to review data sets for final archive. The existing DAAC standards for data archival are currently under discussion, and it is anticipated that some modification of the standards may result prior to the archival of LBA–Ecology data sets. All efforts will be made by the Project to minimize the effort required to assimilate the LBA–Ecology data into the ORNL DAAC through the DIS activities described below.

All data sets will be available through the DIS after a period of time set forth by LBA–Ecology data distribution policies. Prior to this period, the investigators control

access to their data, but are expected to share data as freely as the data quality and scientific practice warrant. The Project Office staff at GSFC will assist with quality assurance for metadata and data with guidance from the Science Team, but this process will not limit data distribution.

The following sections discuss how the LBA–Ecology DIS plans to accomplish these goals for Science Team data. The first section presents the approach from the perspective of the users of the LBA–Ecology DIS by elaborating on the five guiding principles. The second section provides a descriptive overview of six fundamental technical components. This technical overview describes the components and their respective interactions that are necessary in the functioning of the overall LBA DIS to be able to handle the Science Team metadata and data appropriately, in accordance with LBA and LBA–Ecology international participation requirements. The technical overview section highlights Project Office, ORNL support, and science investigator roles associated with each LBA–Ecology DIS component. This section also drafts a time line that indicates LBA–Ecology DIS component activities relative to the data life cycle, from data collection planning and acquisition to data archival.

8.3 The User's Overview of the LBA–Ecology DIS

8.3.1. Minimizing the Time for Data to Become Available to All Investigators

In a distributed data system, as soon as data are ready to be released to fellow researchers, the data and metadata are placed on an Internet–accessible computer by the investigator. The investigator has to run a Web server on his/her Mac, PC, or UNIX platform (or transfers the data to ORNL or GSFC where Project Office staff coordinates the technical details). The investigator then sends ORNL the address of the node. This is the first and only contact the investigator has to make to ORNL during the entire project life cycle. Metadata is harvested from the investigator's machine by the system and maintained in an index at ORNL to allow for searching, but the metadata files and the data files remain on the investigator's machine. Users are able to search for data and documentation files through the ORNL index and download them directly from the investigator's machine (or the machine at ORNL, if they were transferred there).

The user interface is based on a search engine interface. There is a simple format established for metadata files, and the variable pick lists feature of the Web–based data entry tool will aid the investigator in creating the metadata. In addition to the metadata inputs, a readme file, ASCII manuscript, or HTML can be included in the metadata. It is hoped that investigators will register documentation describing the data before even collecting data identifying themselves to the system.

The authorization (password protection) capability allows investigators to select a group of collaborators to examine their data. The investigator will remove the password protection when he/she considers the data are ready for general distribution. If subsequent investigation indicates that the data should not be shared, the investigator

removes it from the viewable area on his or her machine; at the next harvest, reference to the data in the central index at ORNL will be deleted.

8.3.1.1 Web–Based Data Search and Retrieval Requirements

The following requirements are associated with *Minimizing the Time for Data To Become Available to All Investigators:*

- The system MUST automatically collect (harvest) metadata without data provider or DIS intervention after data are placed on the data provider's system.

The system MUST give data providers control over which preliminary data the system may access, when it may access the preliminary data, and who may access the data within an agreed–upon time period that is compliant with LBA data policies.

The system MUST allow data and metadata to reside on, and be accessed from, the data provider's machine until it is ready for transfer to long–term archive.

- Both metadata and documentation SHOULD be searchable by the system.
- The system SHOULD allow the data provider to supply only metadata as a means of "registering" his/her data set before the data are ready for distribution. Whatever level of detail the data provider wishes to give should be accommodated by the system. The system SHOULD also allow the data provider to make subsequent updates to registration information.
- The metadata editing tool SHOULD provide assistance in entering valid metadata values, including entry of a value from a list of controlled terms.
- The system MUST allow data suppliers to select the hardware platform (e.g., Macs, PCs or workstation) and web or FTP server software of their choice at their distributed location.
- After an investigator (data provider) indicates a data set location on his/her computer (e.g., URL with appropriate file directory) to the DIS, the system MUST have the capability to make a new data set visible at a regular time interval (e.g., nightly), without the data provider having to notify the DIS of each data set individually.
- The system MUST have a Web–based search interface, and be accessible from standard browsers with Java and JavaScript capabilities (e.g., Netscape 3 and above and Internet Explorer 3 and above).

8.3.2 Maximizing the Ease of Searching for and Retrieving Data

Users of the system are able to search on fields such as KEYWORDS or PRINCIPAL INVESTIGATOR for focused searches, or conduct a free–text search on the metadata and or data documentation files. Also, users are able to conduct navigational searches when they are unclear about the contents of the data base. The user can move down through a hierarchy of classifications to find the data that he/she needs. For example, the different levels might be "Earth science," "atmospheric science," "atmospheric chemistry," "carbon dioxide," and "atmospheric concentrations of carbon dioxide." Reports can also be created to list the projects and data sets by "status parameter," indicating data availability and level of data documentation and QA. Investigators will

also have the ability to search for data of interest within a particular region and/or time frame.

The system will be fully compliant with national and international data sharing protocol standards, and will include options to search dozens of other data bases.

8.3.2.1 Web–Based Data Search and Retrieval Requirements

The following requirements are associated with *Maximizing the Ease of Searching for and Retrieving Data:*

- The system MUST support queries of metadata.
- The system MUST support pick lists for fielded queries.
- The system MUST support free text searching anywhere in the complete metadata and documentation.
- The system should assist the user in narrowing search criteria through the use of search browse trees.
- The system MUST support geospatial queries.
- The system MUST support temporal queries.
- The system MUST support automatic transfer of fielded metadata to a data base structure for more rigorous queries and reports. An example would be a report on data set status for the use of the LBA-Ecology Project Office or Headquarters.
- The system MUST be available for use to all LBA–DIS partners (e.g., EU, Brazil).
- The system MUST support international data sharing standards to facilitate cost–effective data exchange among LBA–wide communities and communities external to LBA, and to fulfill U.S. Government mandates for data sharing.

8.3.3 Facilitating the Preparation of Value–Added and Integrated Data

It is expected to be important possible that the Project will develop some integrated data and value–added products derived from data for both current and future investigations.

Project Office staff may work with the Science Team to produce these integrated data sets (Scientists may produce integrated products without Project Office participation).

8.3.3.1 Web–Based Data Search and Retrieval Requirements

The following requirements are associated with *Facilitating the Preparation of Value–added and Integrated Data That May Be Requested by the Science Team:*

- The system should be preconfigured with templates/definitions to support national and international metadata standards and allow for translation and cross–formatting of common metadata formats and standards (e.g., DIF, Dublin Core, GILS, FGDC, and Z39.50).
- The system should have both Web–based and standalone integrated metadata entry tools, which the DIS can give to data providers at no cost to them.

8.3.4 Developing and Operating the LBA–Ecology DIS at a Minimal Cost

The Commercial Off The Shelf (COTS) software that the system is using is produced by Blue Angel Technologies, Inc., and is currently being implemented in Internet search and retrieval systems worldwide. The modules are linked in a design of choice. The Alpha prototype is being tested, and already enables HTML searches, provides for fielded and free–form queries, automatically builds the data entry tool template from the data base definition, and automatically builds the relational data base and the navigation tables. Modifications to enhance the functionality and automation (lower ultimate costs) have been identified and are being costed (time and funding) for implementation. However, the system is already close to an operational system suitable for LBA–Ecology.

8.3.4.1 Web–Based Data Search and Retrieval Requirements

The following requirements are associated with *Developing and Operating the LBA–Ecology DIS at a Minimal Cost:*

- The system MUST have sufficient documentation to allow system maintainers and users to learn about all aspects of the system.
- The system MUST be based on well–documented, supported software to allow flexibility in incorporating evolving technology.
- The system should be based on well–supported (e.g., COTS) software.
- The system MUST be able to access data files in any format (e.g., ASCII, netCDF, GEOTIFF, gif, jpeg, spreadsheets, SAS data sets, etc.).

8.3.5 Summary of Features and Functionality

LBA DIS is designed to support the data and information needs of field and laboratory programs where the critical aspects of a desirable system are 1) early exchange of data among investigators, 2) control of preliminary data visibility in the system to be maintained by investigators, 3) rapid and economical deployment, and 4) high automation and scaleability to reduce maintenance costs.

The system allows both free text and fielded searches, using the Meta tag functionality of HTML. The approach of the system is to allow distributed metadata files to be searched in order to identify desired data sets for retrieval. The system is designed to make maintenance highly automated to reduce costs and deployment time. Figure 8.1 provides a graphic representation of the system component functions.

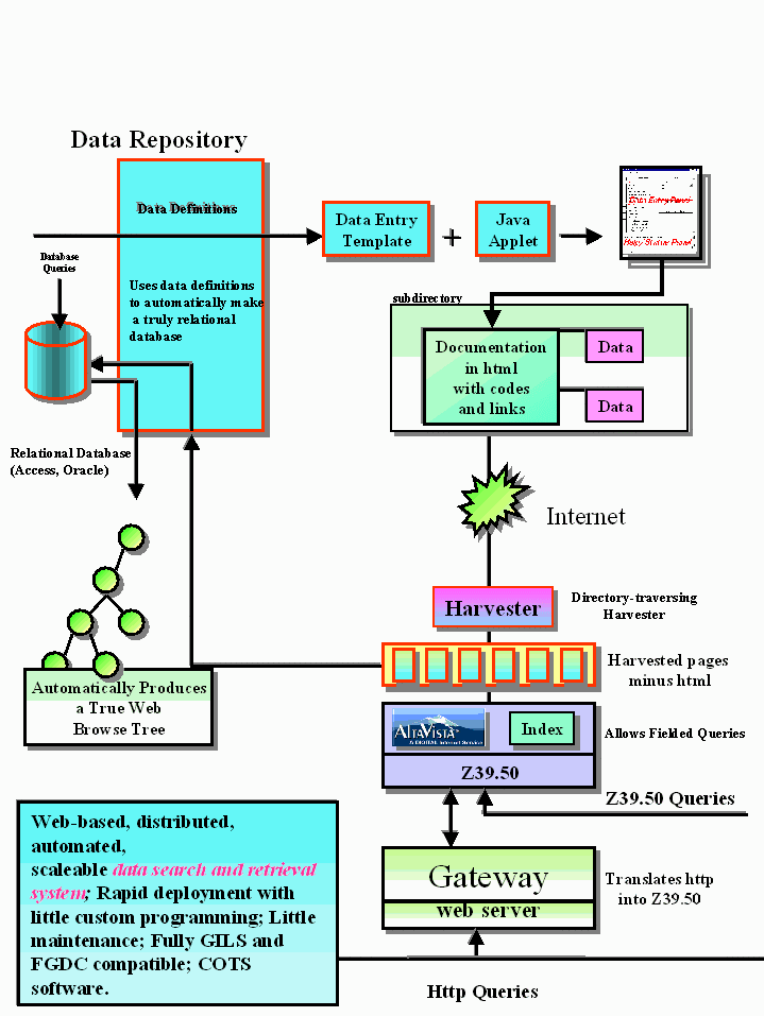

Figure 8.1

The system has the following characteristics:

- Efficient, simple, highly automated architecture.
- Data automatically included at any Internet–connected site.
- Same metadata and data architecture at all sites.
- No traditional data base programming necessary.
- Flexible, comprehensive free text and fielded searching.
- Keyword pick lists for data entry and searching.
- Fast transfer between system components.
- Computer platform independence (Mac, PC, Unix workstations).
- Simplicity and low investigator requirements.
- No special software or programming by investigators.
- PI in full control of preliminary data visibility by manipulating data at his/her site.

The system is an Alpha prototype at ORNL, capable of harvesting environmental data from a number of machines on three continents. The Beta stage of Mercury will more extensively modify the COTS software to achieve: 1) a more flexible data entry tool, capable of free–form as well as constrained vocabularies, 2) a modified harvester that

will employ Web–tree browsing allowing just one Universal Resource Locator (URL) per site to define the search schema, and 3) fully automated updating of controlled vocabularies for data entry templates as well as for free–form and fielded searches.

8.4 A Technical Overview of the LBA–Ecology DIS

The LBA–Ecology DIS represents the people, technology necessary to facilitate data management activities, and efforts focus on identifying and implementing data management services. Technical elements of Science Team data handling include 1) data registration, 2) metadata quality assessment, 3) data search and retrieval, 4) data quality control and assurance, 5) value–added data product generation, and 6) data archival. The following sections summarize each of these elements.

8.4.1 Data Registration

Data registration refers to the acquisition of data descriptions (metadata) and documentation. Such registration information includes who collects data, what types of measurements are collected, and when and where data are collected. Various characteristics about the data, such as specific parameters and units of measurement, may also be described. Metadata and documentation tools will be provided to assist with data registration and documentation. The tools will achieve cost savings to the DIS by assisting investigators with expediting data registration and documentation processing, minimizing data description entry errors, and producing standard metadata file formats required for automated procedures. Figure 8.2 illustrates basic principles of the data registration process.

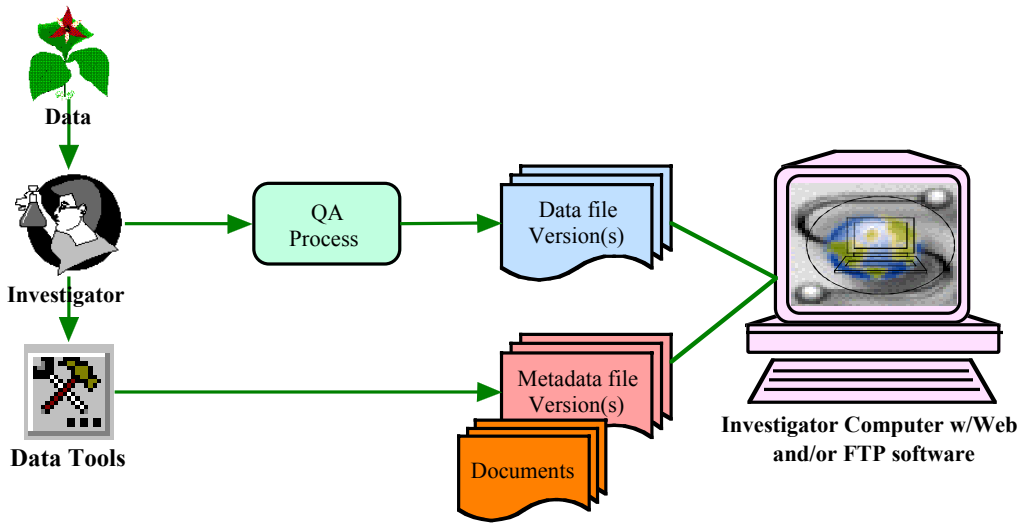

Figure 8.2

Once an investigator has collected data, the data tools can be used to generate applicable metadata and documentation. The investigator describes data via an intuitive tool interface. The descriptions are translated into standardized formats (e.g., DIF), and the formatted document is

saved to a file. In order for the DIS to acquire and process registration information automatically, the investigator is required to run simple web or FTP software on the investigator's computer, where metadata and documentation are stored. Figure 8.2 also indicates that data are stored on the system, which may be processed for quality control/assurance at the investigator's discretion. The investigator can edit existing metadata/documentation and release more complete metadata descriptions and additional versions of data as the project progresses. New versions of metadata, documentation, and data will be automatically registered by the DIS software.

8.4.1.1 Investigator and Project Support Roles

Investigator Roles:

- Provide metadata and provide data set documentation in standard formats. Create initial version of metadata and data documentation, and make subsequent changes to information as needed.
- Provide feedback and requirements to the LBA-Ecology Project Office regarding metadata and documentation tool development.

GSFC Project Office Roles:

- Assemble requirements and lead design of metadata and documentation tool development. Interact with investigators to ensure that Science Team tool requirements are met.
- Implement the data documentation tool.

ORNL Project Office Support Roles:

- Work with GSFC to implement and/or identify existing metadata/documentation tool(s) that meet Science Team and data registration requirements.

ORNL DAAC Roles:

- Provide feedback on DAAC metadata and documentation requirements relative to tool development activities, to help ensure the smooth transition of archivable data to the DAAC.

8.4.2 Metadata Quality Assessment

The preliminary assessment of metadata quality is a simple yet important step in ensuring the accuracy of data registration information. The DIS depends on accurate metadata, as users will rely on the DIS for rapidly producing quality data and information of interest. Preliminary metadata quality can be assessed by computer–automated procedures; however, human intelligence will be required for comprehensive checks. For example, the data registration tool can automatically check for incomplete mandatory fields, ensure accurate entries through the use of controlled pick lists where applicable, detect mispelled words, and provide other checks based on heuristic semantics. The development of in–depth analysis algorithms that ensure total metadata accuracy is cost–prohibitive, and beyond the scope of LBA–Ecology DIS tool functions. Data providers ultimately ensure quality metadata and documentation for their data, and Project Office staff will be available to assist with metadata assessment and documentation production activities, if requested by the Science Team.

8.4.2.1 Investigator and Project Support Roles

Investigator Roles:
- Provide feedback and requirements to the GSFC Project Office regarding simple metadata validation algorithm development.

GSFC Project Office Roles:
- Receive reports regarding new or updated metadata, and review the metadata to assess their validity.
- Interact with investigators to help clarify and validate any questionable metadata entries.
- Provide ORNL with metadata QA checking algorithms that are incorporated into the metadata entry tool.

ORNL Project Office Support Roles:
- Work with GSFC to implement processes and reporting mechanisms that are needed for GSFC to obtain updates on new and changed metadata.
- Implement QA checking algorithms as part of the metadata entry tool.

ORNL DAAC Roles:
- Provide feedback on GSFC metadata issues in terms of past experience and future expectations.

8.4.3 Data Search and Retrieval

A DIS can be measured by its ability to locate and provide access to data of interest and of high quality rapidly. The data search and retrieval component provides automated and conventional technologies for rapidly discovering available data and making data available to DIS end users, while maintaining investigator control of preliminary data access and distribution.

Web search engines will provide the capability for investigators to locate data based on user–specified search criteria (e.g., spatial, temporal, key words, etc.). Data are planned to be distributed across many locations (i.e., investigator sites). The search and retrieval system provides "web–crawler" mechanisms for automatically collecting data registration/metadata information from each distributed data site. Metadata collections are automatically indexed and loaded into a data base structure to facilitate transparent search engine capabilities. Figure 8.3 briefly illustrates data search and retrieval functionality. Investigators play two different but synergistic roles within the DIS: data providers and data users.

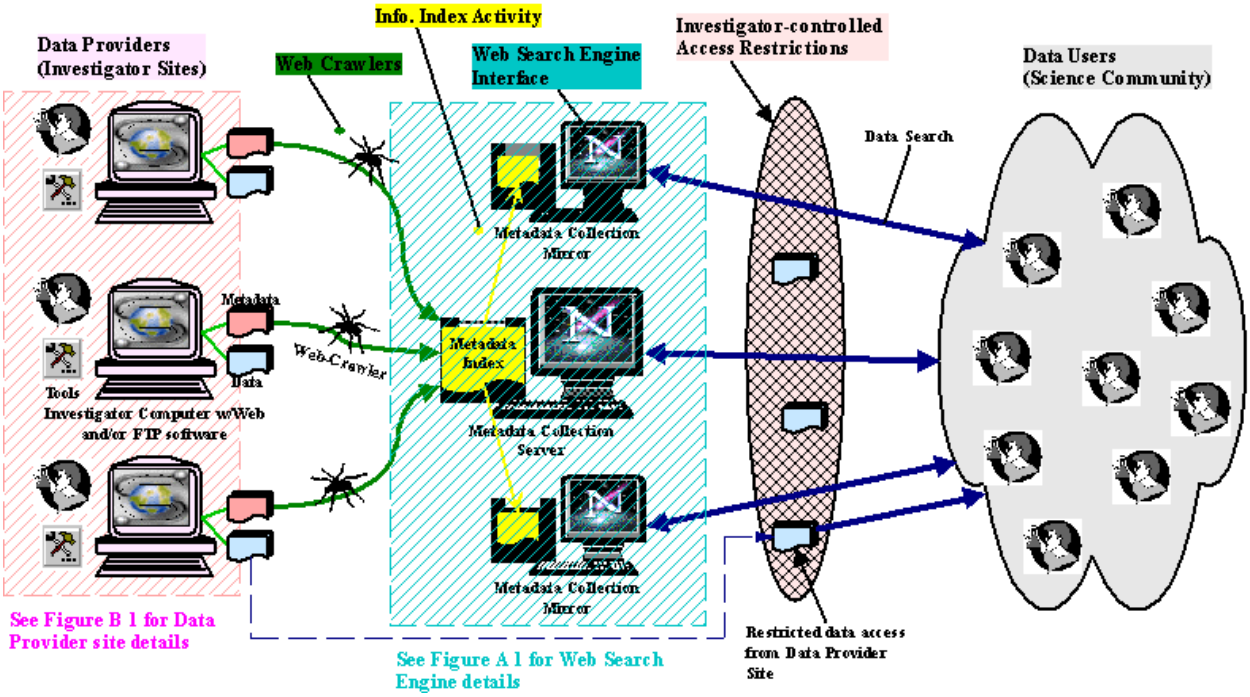

Figure 8.3

In the data provider role, an investigator utilizes data tools to describe data and generate standard metadata/documentation formatted files. The investigator then stages data and metadata files to a computer running Web server software. The data search/retrieval component automatically detects the staged files, transfers the information to search/retrieval server system(s), and indexes the information. Note that there are satellite or mirror search/retrieval servers that essentially receive copies of indexed information (Figure 8.3), increasing search engine performance by distributing the load of data requestors. It is anticipated that Brazil will host the same or similar data search and retrieval software, and serve as the central data harvesting point for all LBA data. LBA–wide original data will be archived in Brazil as well.

Data users locate data of interest through the use of a server system's search engine. The search engine consults metadata indices, and returns a matching list of data sets based on specified search criteria. Data users will retrieve metadata and data files associated with their search criteria from staged online data at investigators' sites preliminary data, however, may be restricted by investigators to a selected group of end users.

8.4.3.1 Investigator and Project Support Roles

Investigator Roles:

- Establish and maintain a "harvestable" data site (i.e., computer with a Web or FTP server), or ask the Project Office to maintain the data site.
- Describe data (with metadata and documentation tools) and place available metadata, documentation, and data file(s) on the Web–accessible site.

- Inform ORNL (System Operator[s]) of location (URL) of metadata/data, and any applicable data access restrictions.
- Contact ORNL for general questions and/or assistance regarding the system.
- Provide feedback to GSFC and ORNL regarding requirements and overall functionality.

GSFC Project Office Roles:
- Provide project–specific functional requirements related to data search, retrieval, and distribution.
- Interact with LBA–Ecology investigators to obtain feedback on the data search and retrieval functions and capabilities, and provide ORNL with regular reports on any feedback obtained from investigators.
- Ensure that any preliminary data access restrictions imposed by investigators are in accord with LBA data sharing policies.

ORNL Project Office Support Roles:
- Implement and maintain data search and retrieval components, including temporary Project data archive of mature data sets.
- Interact with investigators on issues related to configuration and operational aspects of the data search and retrieval system (e.g., user consultation regarding system use, software troubleshooting and testing, etc.).
- Distribute copies of indexed information to applicable data search and retrieval mirror sites.

ORNL DAAC Roles:
- Provide guidance regarding Project temporary archive to facilitate eventual transfer of data to permanent archive.

8.4.4 Data Quality Assurance

Data quality analysis is needed to ensure optimal usefulness of data to current and future data users. Computer–automated procedures can be used for simple quality assurance (QA) such as gross–bound checks or consistent measurement units; however, human intervention will be required to analyze data at various levels. Tools will be made available to investigators for their use in QA. Project Office staff will be available to assist investigators with data QA efforts on a case–by–case basis, and as time and resources permit. Further, as data become available to the Project Office, the data will be automatically checked using the same simple data QA utilities.

8.4.4.1 Investigator and Project Support Roles

Investigator Roles:
- Interact with the Project Office to identify data QA tool requirements.
- Request Project Office resource assistance for data QA tasks, if needed.
- QA Data

GSFC Project Office Roles:
- Lead development and progressively more capable versions of data QA software.

- Develop automated procedures for using data QA tools to check incoming data; and any report, QA anomalies/problems to respective data providers.
- Interact with Science Team to define QA tool requirements and functionality.
- Assist Science Team with data QA tasks if requested, and as resources allow.

ORNL Project Office Support Roles:
- Work with GSFC to ensure that metadata tool file formats are interoperable with the data QA tool.

ORNL DAAC Roles:
- Interact with and provide feedback on LBA-Ecology QA efforts in terms of past experience and future expectations.

8.4.5 Value–Added Product Generation

It is anticipated that investigators will derive various products from existing data sets based on their scientific requirements. For example, meteorological and satellite data can be integrated and overlaid with topography in a GIS. Remote sensing data can also be georeferenced with ground control data, and imagery can be subsetted to complement various GIS map–related data. Figure 8.4 depicts quality assurance and value–added product generation activities.

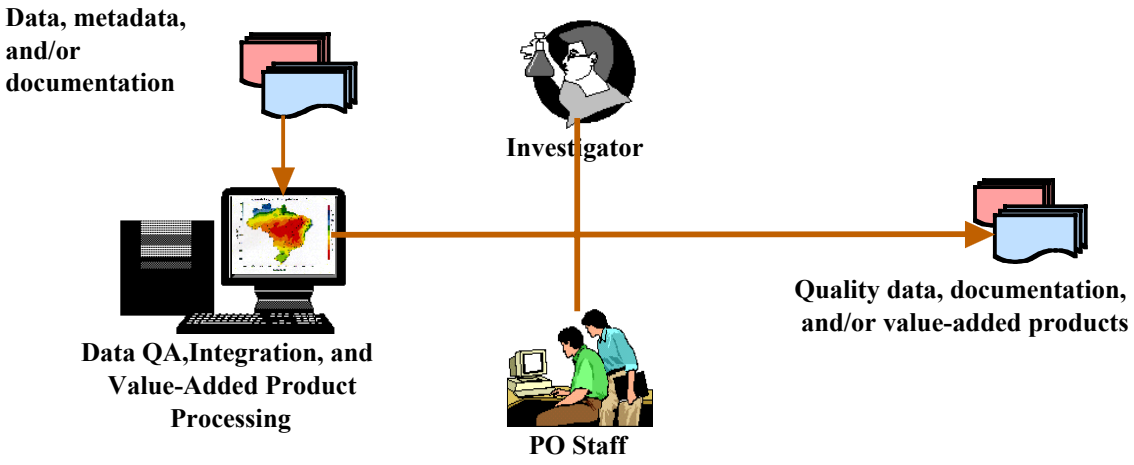

Figure 8.4

If requested by the Science Team, the LBA-Ecology DIS can facilitate the production and sharing of value–added products with other investigators through the data search and retrieval component and/or the production of CD–ROMs. As LBA-Ecology Project resources permit, project Office staff can also assist with Science Team requested value–added derivations through the use of Project Office computing and GIS processing facilities.

8.4.5.1 Investigator and Project Support Roles

Investigator Roles:
- Interact with GSFC to provide requirements and guidance in the derivation of value–added data products (e.g., GIS, image enhancements, merged/modeled data).
- Request assistance, if needed, for generating value–added data products.
- Create value–added data products and register and document them.

GSFC Project Office Roles:
- Interact with investigators to identify value–added products that may be useful to the Science Team and/or LBA in general.
- Provide assistance in value–added data product generation.
- Register and document value–added products with the data search and retrieval system for data products that are created at GSFC.

ORNL Project Office Support Roles:
- Work with GSFC to make value–added data products available through the data search and retrieval system.
- Provide GSFC with information about potential data sets for follow up contact.

ORNL DAAC Roles:
- Work with GSFC as needed to archive selected data sets.

8.4.6 Data Archival

Data archival refers to long term starage of quality assured data and associated metadata and documentation. Data archival for LBA-Ecology will be at the ORNL–DAAC and in Brazil; these data may potentially be published on CD–ROMs or other appropriate media. The level of metadata, documentation, and data quality is driven by EOSDIS and LBA guidelines. Further, archived data will be accessible to the general public. It is currently anticipated that most or all LBA–Ecology data will be publicly accessible within one year of initial data acquisition either through the long term archive or unrestricted access at the Investigators web site.

All project data may not necessarily be archived. Science Team members and Project Management will determine which data should be archived, in order to secure a long–term investment in LBA–Ecology research efforts. All project data, however, will be transferred to Brazil, regardless of whether they are archived for the long term. Figure 8.5 illustrates the data archival component, as data are documented, QA'd and potentially archived at the ORNL DAAC. It is anticipated that all data judged to be of high quality and potential value to other will be archived, but intermediate data sets and data of dubious long term value may not be preserved beyond the duration of the LBA–Ecology Project.

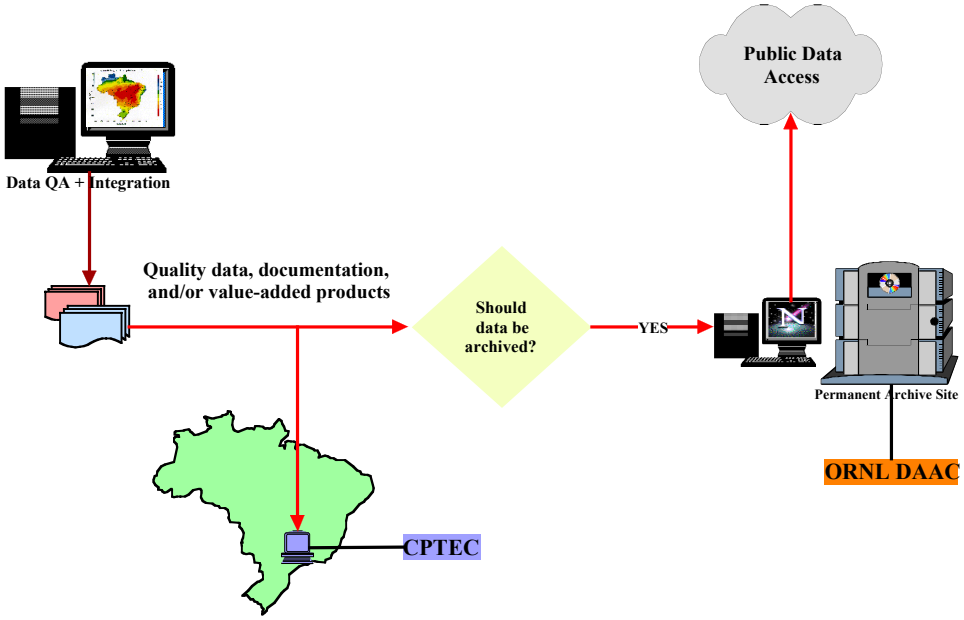

Figure 8.5

8.4.6.1 Investigator and Project Support Roles

Investigator Roles:

- Recommend data sets for long term archival and participate in decision–making process with LBA and LBA–Ecology Project management to determine which data will be archived.

Project Office Roles:

- Assist the Science Team with data and documentation processing efforts in order to meet ORNL DAAC data archival requirements.
- Organize the efforts to publish final LBA–Ecology data collection on CD–ROM or other appropriate media.

Project Scientist Roles:

- Organize the review of data and documentation by project management and investigators to determine those data that should be archived.

ORNL Project Office Support Roles:

- Work with GSFC and Science Team to participate in discussions regarding which data should be archived.

ORNL DAAC Roles:

- Work with GSFC to define exchange processes that allow efficient transfer and version control of data and documentation for public release.

8.4.7 DIS Component Activities Relative to Data Maturation

The DIS components described in the overview can be viewed as basic steps to data maturation. At the data registration level, data are planned to be collected or are currently being acquired. At this stage, investigators may describe or register their data, which in turn makes the data "harvestable," or automatically available for others to discover, through the data search and retrieval component. Data registration actually facilitates data sharing and exchange early on, and the data, at subsequent levels of revisions, normally remain visible throughout the project's life cycle, and possibly beyond. The registered metadata are validated through simple metadata QA tools, and simple data checks can be initiated as well, using simple data QA tools. Once simple metadata and data QA are favorably completed, data maturation may progress to more rigorous metadata and data QA analysis. At this point, metadata and data may progress through various levels of QA, and complementary data integration may take place. Data documentation may also progress, and useful value–added data products may be created. At the final stages of the data QA step, data and documentation will maintain levels that are significant for archival by the DAAC. Figure 8.6 depicts the progressive steps of data maturity relative to DIS components. Input for this section is currently being acquired from Science Team representatives.

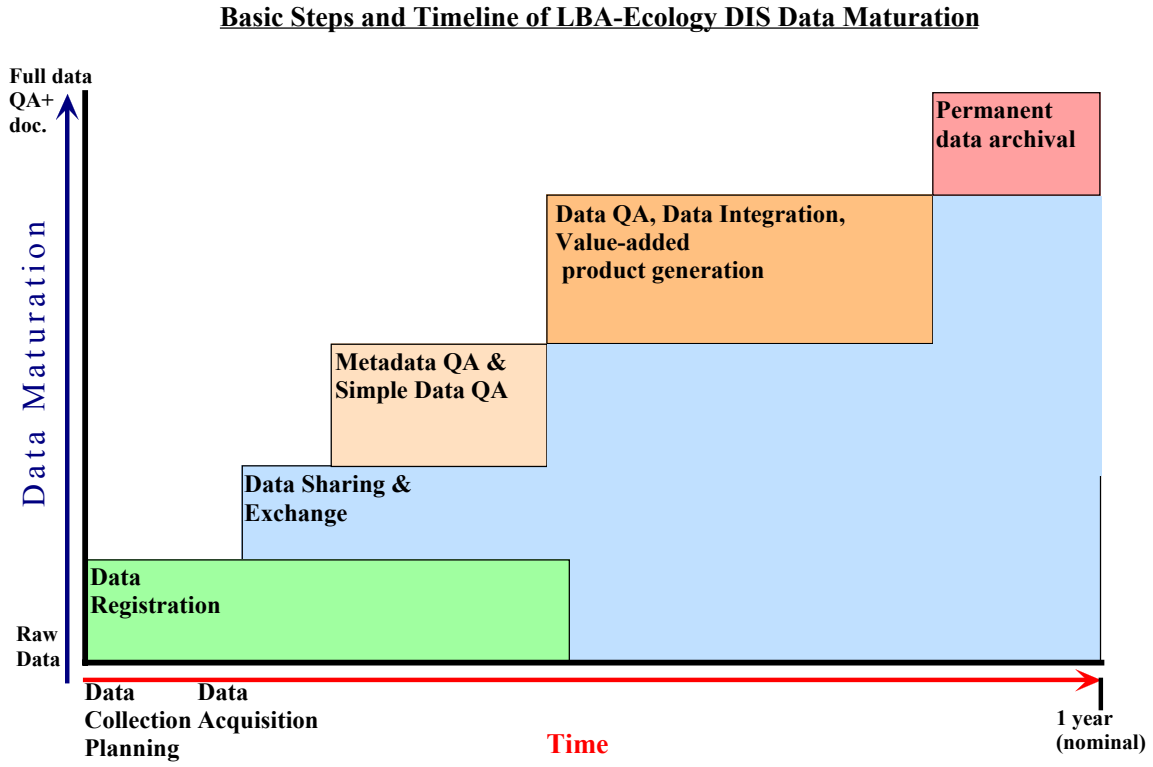

Figure 8.6