**ASSESSMENT OF NEEDS AND REQUIREMENTS FOR
A GEOGRAPHIC INFORMATION SYSTEM FOR
BROOKHAVEN NATIONAL LABORATORY**

Sharon L. Baxter
and
Mary L. Daum
Biomedical and Environmental Assessment Group
Analytical Sciences Division
Department of Applied Science

and

Allan Scott
Office of Environmental Restoration

November 19, 1992

MASTER

This research was performed under the auspices of the United States Department of Energy under
Contract No. DE-AC02-76CH00016

DISTRIBUTION OF THIS DOCUMENT IS UNLIMITED

fr

DISCLAIMER

This report was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor any agency thereof, nor any of their employees, nor any of their contractors, sub- contractors, or their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise, does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency, contractor, or subcontractor thereof. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency, contractor or subcontractor thereof.

ABSTRACT

At the request of the Brookhaven National Laboratory Office of Environmental Restoration, and with the cooperation of the Safety and Environmental Protection and Plant Engineering Divisions, a Geographic Information System (GIS) Task Group has undertaken an effort to identify the data requirements, potential applications, and computing environment within which an integrated GIS could be implemented to meet the environmental and facilities management needs of the Laboratory.

The Task Group held discussions with the three participating groups and with Laboratory officials. Software and hardware vendors were invited to demonstrate their products to all participants using BNL data, and private consultants discussed their GIS and database applications.

After considering the needs of the three groups and other general implementation concerns, a list of ten requirements for a BNL site-wide GIS has been determined. On the basis of these considerations, three conceptual models for an integrated GIS are defined. These include a highly centralized system, a fully distributed system, and a composite model. The models differ in their allocation of responsibility for data management, applications development and GIS analysis, and local user support. Several issues remain to be resolved before an integrated GIS for the three operational groups can be fully implemented.

TABLE OF CONTENTS

ABSTRACT	ii
1 Introduction	1
2 Approach	2
3 Assessment Results	3
3.1 Office of Environmental Restoration	3
3.1.1 Mission	3
3.1.2 Computing environment	3
3.1.3 Needs	4
3.2 Safety and Environmental Protection	4
3.2.1 Mission	4
3.2.2 Computing environment	4
3.2.3 Needs	4
3.3 Plant Engineering	5
3.3.1 Mission	5
3.3.2 Computing environment	5
3.3.3 Needs	5
3.4 Summary	5
4 Conceptual Models	8
4.1 Implementation Considerations	8
4.2 Model Options.....	9
4.2.1 Centralized model	9
4.2.2 Distributed model.....	9
4.2.3 Composite model.....	10
5 Issues to be Resolved	11
ACRONYMS.....	13

TABLES

Table 1. Identified GIS Applications at BNL	14
Table 2. Projected GIS Data Requirements at BNL	15
A. Potential Maintenance Responsibility	15
B. Current Data Availability	17
Table 3. Computing Environment at BNL	19
APPENDIX A. Questionnaires	20
APPENDIX B. GIS Data Survey Forms	22
APPENDIX C. GIS Implementation Considerations	24
1 System Design	24
1.1 Software choice	24
1.2 Platform configuration	24
1.3 Processing configurations	24
2 Cost of Implementation	25
2.1 Initial costs of software	25
2.2 Initial costs of hardware	25
2.3 Network costs	25
2.4 Data acquisition costs	26
2.5 Applications development costs	26
2.6 Maintenance and upgrade costs	26
3 Time to Implement	26
3.1 The learning curve	26
3.2 Database design and implementation	27
4 System Maintenance and Support	28
4.1 Software upgrades	28
4.2 Data management	28
4.3 Application design	29

1 Introduction

A National Laboratory has the continuing obligation to maintain and monitor the physical environment and infrastructure within its boundaries and in the surrounding areas. Several groups at BNL are charged with meeting this obligation. They include the Safety and Environmental Protection (SEP) Division, the Plant Engineering (PE) Division, and the Office of Environmental Restoration (OER). Although each group has specific needs for the acquisition, analysis, and display of data, there is a significant amount of overlap. It is increasingly clear that the use of standard Geographic Information Systems (GIS) would greatly enhance the analysis and display of the locational and environmental data associated with this endeavor.

A joint effort to develop and maintain a commonly accessible data base and GIS will benefit each group separately as well as the Laboratory as a whole. Potential benefits include time and cost savings and data quality control. This site-specific need has been underscored by DOE Order 4320.1B (1/7/91), which covers Site Development Planning for U.S. Department of Energy sites. Section 7h states, "Sites shall include in the planning process the acquisition and use of standard Geographic Information Systems (GIS), which promote coordinated development, use, sharing and dissemination of surveying, mapping, and related spatial data. Cooperation with State and local governments and the private sector is encouraged to the fullest practical extent."

A GIS covering the entire Laboratory and surrounding areas (a site-wide GIS) would facilitate data sharing within the Laboratory and with other organizations that currently have a GIS in place (for example, the U.S. Geological Survey (USGS), the Suffolk County Department of Health, and State and local governments). Many consulting firms under contract with the Laboratory use GIS in their work and could deliver their results directly in GIS formats. Communications between BNL and outside agencies would be enhanced by expanded analytical capabilities, rapid access to site environmental data, and the ability to generate high quality maps, graphs and reports. An integrated site-wide GIS would also facilitate Brookhaven's participation in DOE's inter-Laboratory Information Exchange group on Use of GIS in Environmental Restoration/Hazardous Waste Programs at DOE Facilities.

At the request of the BNL Office of Environmental Restoration, a GIS Task Group was formed to address the issues involved in developing a common GIS for the three operational groups (OER, SEP, and PE). The Task Group was asked to assess the needs of the groups involved for acquiring, managing, analyzing, and presenting spatially referenced data, and to recommend one or more software and hardware configurations for use as a BNL site-wide GIS. These two tasks have been undertaken in parallel.

This report addresses Task 1, the needs assessment. In order to determine the GIS configuration best suited to the common and particular needs of the

operational groups, we have undertaken an assessment of the organizational functions, data needs, current status, and required deliverables of each group. At the same time, in the process of evaluating software for Task 2, several Geographic Information Systems have been demonstrated on site. This has proven useful for the assessment, since the demonstrations have helped people understand what a GIS is and how it can be used as a tool in their work.

Task 2, the GIS software evaluation and site configuration recommendations, will be discussed in a subsequent report.

2 Approach

A needs assessment requires more than a listing of the self-perceived needs of the departments involved. It requires an understanding of current and anticipated future deliverables and knowledge of hardware systems and software already in place.

The first component of the assessment was a series of interviews which were conducted with personnel in OER, SEP, and PE. Each group was interviewed at least twice. In addition, overall Laboratory concerns were discussed with Laboratory administrators and division heads.

A questionnaire was developed for use in the first round of interviews. The questions focused on:

- Software and hardware currently in use
- Critical functions
- Outputs/deliverables
- Probable/potential users of a GIS
- Data requirements
- Commitments, requirements, drivers

More detailed responses to a subset of questions were obtained from the most likely immediate users of the GIS. The results of these interviews are being compiled into a database, which will identify and prioritize the GIS applications, thematic data layers, and data sources that will be required. To date these detailed interviews have been held formally only with people in OER.

Copies of both questionnaires can be found in Appendix A. The draft database forms are in Appendix B.

A list of applications, data sources, and computing environments for an integrated GIS was distilled from the results of the interviews. These requirements are discussed in Sec. 3 and are summarized in Tables 1-3.

The second component of the assessment comprised a series of software demonstrations. Several candidate GISs were identified, based on the results of the interviews and on our knowledge of and experience with GIS software. Vendors were asked to demonstrate their software using graphical BNL site data from the PE AutoCAD system and an abbreviated set of SEP's historical monitoring well data. These demonstrations helped us to evaluate the various software packages, since each vendor built the demo around the same set of graphical and attribute data. At the same time they provided an opportunity to educate the participants about Geographic Information Systems in general.

The third step in the assessment process involved meetings with environmental consultants and hardware vendors. Several consulting firms provide turn-key GIS systems, GIS support, and/or applications development. They discussed how local government agencies and other Federal facilities are using GIS in support of environmental monitoring and restoration. Hardware vendors provided information on the requirements and capabilities of GIS-related hardware.

Finally, available options were considered and a set of conceptual models for an integrated site-wide GIS at BNL was developed. Section 4 outlines options and presents three basic models for consideration. One of these models will be recommended in the final version of this report.

3 Assessment Results

This section summarizes pertinent information gathered from each department during the interview process.

3.1 Office of Environmental Restoration

3.1.1 Mission. The mission of the Office of Environmental Restoration is to oversee the cleanup of hazardous waste areas at the Laboratory in accordance with CERCLA/SARA requirements.

3.1.2 Computing environment. OER currently operates with a variety of PC-based software, including AutoCAD¹ and R:Base.² They are not currently linked to the Laboratory Ethernet network. However, OER is a relatively new department at BNL and considerable development work is now going on. More powerful UNIX-based database management systems (DBMS) are being evaluated to manage their environmental data. Two Sun workstations, including one file server, are on order and an Ethernet network (using PC-NFS³ software) is being developed.

1 © AutoDesk

2 © MicroRIM

3 © Sun Microsystems, Inc.

3.1.3 Needs. OER has been the moving force behind the current effort to establish a site-wide GIS. They are committed to the use of a GIS, and recognize their own need for environmental data generated by SEP, SEP historical data, and physical plant data from Plant Engineering. At the same time, OER is willing to share their data with other Laboratory entities and sees a site-wide GIS as a powerful integration tool.

In addition to data exchange and integration, OER requires a GIS that can easily interface with other software that they will be using extensively, including modeling programs and statistical packages. Because OER is just now starting to develop an information management system, they are anxious to have a GIS (site-wide or otherwise) in place as soon as possible. They want to begin analyzing existing conditions and collecting new data in the GIS environment, rather than having to transfer the information at a later date.

3.2 Safety and Environmental Protection

3.2.1 Mission. The Safety and Environmental Protection Division is charged with monitoring and protecting Laboratory personnel and the site's natural environment. This charge extends to off-site populations and environments that might be impacted by Laboratory activities. The three groups which have been involved in this site-wide GIS needs assessment are Environmental Monitoring, Computer Support, and Emergency Response, the latter two both part of the Technical Support Group.

3.2.2 Computing environment. SEP's monitoring database currently resides in R:Base in a PC environment. The PCs and two Sun Sparc II workstations are connected through Ethernet using VINES⁴ network software. The division's data management requirements have long since outstripped the capabilities of their current DBMS and they are now in the early stages of developing a new system using Sybase,⁵ which is supported by CCD. They anticipate that it will be several months before the new database is in place. One aspect of upgrading SEP's monitoring database will be to ensure that their data (e.g., locations and attributes of groundwater monitoring wells) agree with those of OER and Plant Engineering.

3.2.3 Needs. SEP suggested a number of potential applications for a GIS. These range from visual display of database queries (e.g., plotting water table levels and annual average sampling data from a set of wells) to geographic analyses (where are the high concentrations? what has been the progress of a groundwater plume over time?) to planning applications (where do we put our next wells?). Some of these applications would require a GIS to be able to communicate with SAS,⁶ the statistical package which will be used by the

4 © Banyan Systems, Inc.

5 © Sybase, Inc.

6 © SAS Institute, Inc.

division. SEP would also like to be able to generate simple maps rapidly (< 1 day) rather than having to wait for processing by Plant Engineering.⁷

The most pressing need expressed by SEP was to upgrade their database. SEP has expressed interest in GIS in the past but has been financially constrained from implementing a GIS independently. They would find a site-wide GIS useful in terms of direct applications as well as data sharing and integration, but stress the importance of an easy-to-use system which will minimize the time and effort required for use.

3.3 Plant Engineering

3.3.1 Mission. The Plant Engineering Division maintains the infrastructure of the Laboratory, including utilities, buildings, and grounds.

3.3.2 Computing environment. All site maps are maintained in AutoCAD. Computerized data are stored in dBASE⁸ databases and Excel spreadsheets; some data are not maintained electronically. PE utilizes both PCs (DOS and Windows environments) and Macintoshes. The Division has access to Ethernet using PC-NFS.

3.3.3 Needs. At present PE produces hardcopy maps on demand, not only for OER and SEP but also for other divisions, such as Graphic Arts. However, they find this process cumbersome; the PE AutoCAD specialist estimated that he spends up to 50 percent of his time producing maps for others. Plant Engineering has two basic needs that they would hope to address with a GIS: (1) to facilitate the transfer of site maps in electronic form; and (2) to expand their computerized database to include storage tanks and the locations of hazardous materials. PE sees its principal GIS application as Facilities Management.

Plant Engineering has indicated a willingness to support a site-wide GIS effort and will purchase a UNIX workstation if necessary. However, they plan to continue to use AutoCAD for their mapping needs and prefer to maintain their PC environment.

3.4 Summary

Among the three divisions considered in this report, SEP and OER are likely to be the principal beneficiaries of an integrated site-wide GIS. Sharing of data between these two groups will be facilitated and each will be able to produce the maps and reports it needs without depending directly on another group. Because Plant Engineering plans to continue using AutoCAD to create

⁷ Graphic Arts, which was not interviewed for this needs assessment, has expressed interest in a site-wide GIS for the same reason.

⁸ ©Ashton-Tate Corporation

the site maps, the main immediate benefit to PE will be having more time to do its own work and less demand on its time to produce work for others. However, many of the benefits anticipated for SEP and OER will require the participation of Plant Engineering.

From the interview process two kinds of needs became apparent: those specific to an individual or division, and those which were common to all three groups. The common needs serve as a list of basic requirements for a site-wide GIS at Brookhaven. The list addresses two fundamental concerns. First, few people in these divisions have any experience with any kind of GIS. Second, it is necessary to keep costs to a minimum while obtaining a system of adequate power and functionality, not just for the present but for anticipated future expansion. Initial capital and personnel costs can be kept low and development of productive GIS applications can be hastened by choosing site-wide GIS which has a short learning curve and utilizes hardware and software already in place.

The ten basic requirements for a site-wide GIS at BNL are as follows:

1. UNIX-based. A GIS operating under a UNIX operating system will facilitate movement between hardware platforms, provide enough computing power to manage large amounts of data, and have the ability to generate high resolution (1280 x 1024) graphics. This choice also anticipates future development, since most software development is now occurring in a UNIX environment.
2. Supported on current hardware. A site-wide GIS will be feasible only if the software can run on and be accessed by hardware already in place in the participating groups. Thus the GIS must be accessible by both PCs and workstations. Additional hardware purchases initially should be minimal, with funding going mainly to the software purchase and database and applications development.
3. AutoCAD compatibility. All of the site data currently in digital form are in AutoCAD. Experience at other sites has also indicated that CAD is the preferred method for digitizing new data for GIS use. A GIS that can easily access AutoCAD drawings will shorten start-up time for the input of graphical data.
4. Seamless access to a variety of standard (SQL) database management systems. The GIS must be able to access attribute data stored in powerful DBMSs that are adequate for the needs identified by the participating groups. The system should have the capability of utilizing data from outside sources, and also to use the database software currently in place. The GIS should be able to access more than one database format simultaneously without requiring translation.

5. Access to other software. In addition to accessing attribute databases, the GIS will require links to statistical and modeling software which provide functions not available within the GIS. Direct access to such software is rare. However, the GIS should support transparent access (without having to exit from the GIS) through programmed applications.
6. Multiple data exchange formats available. Data will be shared internally and exchanged with outside entities such as DOE, regulatory agencies, data vendors and consultants. Since virtually all GISs utilize a proprietary format, a site-wide GIS must support the import and export of data in a large number of these formats. Three which are definitely required are:
 - (a) the Spatial Data Transfer Standard (SDTS), which was approved as a FIPS standard in July, 1992, and formally goes into effect on February 15, 1993;
 - (b) DXF, the AutoCAD format which has long served as an ad hoc transfer standard among GISs; and
 - (c) ARC/EXPORT, an ASCII transfer format used by ARC/INFO,⁹ currently the most widely used commercial GIS on Long Island.
7. Multiple coordinate systems supported. Integration of data from many sources requires compatibility between file formats and demands the ability to match data that were originally registered to different spatial coordinate systems or transformed to different map projections. A BNL GIS must support non-standard systems, such as the BNL system, as well as standard coordinate systems such as geographic (latitude/longitude) and Universal Transverse Mercator (UTM).
8. Both vector and raster capability. Spatial data in a GIS can be stored in two basic ways. A vector GIS stores spatial information in the form of points, lines, and polygons; a raster GIS stores spatial information in a regular matrix of cells. The principal reason to favor a vector GIS is that BNL's current spatial database, in PE's AutoCAD system, is a vector system. Vector formats are also preferred for some applications. For example, they permit inclusion of directional information (e.g., direction of flow in a sewer line) and the generation and mapping of contour lines from irregularly spaced sampling points. Raster capability is also needed, however, principally for utilization of digital of aerial photos and certain spatial modeling techniques. A raster format is also used for 3-D displays of data surfaces.
9. User-friendly. The majority of potential GIS users at BNL have had no experience with a GIS. For a system to be cost-effective, it is imperative that the learning curve be as short as possible within its level of complexity. This will entail a menu-driven user interface, a set of built-in common applications, screen-form data entry, and extensive on-line help.

⁹ © Environmental Systems Research Institute

10. Training and support available. Because previous experience working with GIS is scarce at BNL, the software vendor must provide the needed training and ongoing support to the Laboratory. On-site training and unlimited telephone support, and contact with other users are highly desirable.

Discussions with software and hardware vendors have indicated that there are no major technical impediments to implementing an integrated GIS for use by OER, SEP and PE. Serious consideration of software will be limited to a few major commercial GIS products. Several database management systems and many hardware platforms are available and may be considered separately.

4 Conceptual Models

4.1 Implementation Considerations

The needs assessment identified several key issues which must be addressed in order to construct GIS implementation models. These issues include system design (software options and hardware configurations); implementation costs, both initial and ongoing; the time required to have the system operational; and system maintenance and support options.

The software will be chosen on the basis of the ten general requirements listed in Sec. 3.4, on specific departmental needs, and on cost. The cost of the software will depend on the vendor chosen and on the number of users on site. For a full-function GIS the general range is \$15,000 for a single user on a workstation to \$100,000 for a server-based system accessible by many users across a network.

Hardware needs will depend on the hardware currently in place and the number and type of users in the department. Existing PCs may suffice for occasional or light users, but heavy users may need to be upgraded to a workstation. Current hardware and projected use will also affect the network requirements for each group and the overall implementation plan. It will be crucial to determine if the existing communications network is adequate or requires expansion.

The time required to implement the GIS will depend on the issues of software complexity and database design. Consultants may be used to implement the system more rapidly. Although full-function GISs are inherently complex, learning time will be greatly reduced for those individuals who are computer literate and who participate in formal training.

Once the GIS has been implemented, system maintenance, on-site user support, and applications development will be required. These tasks may be accomplished by a specifically appointed group or may be apportioned among the participating groups.

These implementation considerations are discussed in greater detail in Appendix C.

On the basis of these considerations three models can be outlined. The differences among them lie mainly in the areas of data management, applications development and analysis, and on-site GIS support.

4.2 Model Options

4.2.1 Centralized model. In a centralized GIS system, all responsibility is located in a single entity. This core GIS support group includes expertise both in the technical side of GIS and in environmental and facilities management applications.

- *Data management:* The GIS Support Group (GSG) obtains AutoCAD files from PE and converts or imports them into GIS spatial data layers. Attribute data that are collected by the departments are passed on to the GSG for inclusion in the BNL site data system. Data QA/QC may occur within the responsible departments, but the database maintained by the GSG is the one with which people are expected to work. Differential security may be applied so that users in different departments have access to different parts of the data; only the GIS support group has full write and delete access.
- *Applications and analysis:* Development of applications and GIS-based analyses are also the responsibility of the core GIS support group. Department personnel indicate their needs to the group, which generates maps, plots, and reports or creates applications to allow others to produce such output within their departments. These "GIS consumers" do not in fact need to know how to use the GIS.
- *Local support:* User support for the GIS is centered in the GIS support group. Since the GSG constitutes the only real user, the group will be the point of contact with the vendor. The GSG also provides assistance to people in determining potential new applications and works with CCD on technical and hardware problems.

4.2.2 Distributed model. In a fully distributed system, each operational group holds ultimate responsibility for meeting its own GIS needs and for cooperating effectively with the other groups. While each operational group determines its own level of effort with regard to implementing a GIS, individual GIS needs must be considered with respect to implementation of the overall system. There is no core support group within the Laboratory.

- *Data management:* Plant Engineering creates the BNL spatial data layers and converts or imports them into GIS format. PE is responsible for informing the other two departments about any new drawings which are put up on the system. Attribute databases are maintained separately by each department. Responsible individuals in each department ensure that their databases conform to GIS requirements (e.g., consistent data formats, naming conventions, etc.). They inform the other groups when structural changes in the database, including software upgrades, are anticipated, so that applications which access the data can be revised if necessary. Database security and access restrictions for both spatial and attribute data are overseen by the departmental data managers.

- *Applications and analysis:* All data analysis and applications development are done within the individual operating groups. Individual users may operate independently, or a department may decide to designate one or two people to become GIS specialists.
- *Local support:* Since no core GIS support group exists within this model, individual users will probably need to form their own formal or informal user support group. Such a group would enable users to share problems, tips, and applications. Arrangements must be made with the GIS software vendor concerning which users have access to their user support facilities.

4.2.3 Composite model. Composite models fall into the large middle ground between a completely centralized system and a completely distributed one. The attraction of a centralized model is the high degree of coordination among participating departments which is built into the system. It also focuses the effort of implementing a site-wide GIS in a small group of people whose only job is GIS. A distributed system allows each department to maintain control over its own data, including the critical element of security. It also leaves each group to find its own way through the process of learning to use the GIS and imposes the additional responsibility of maintaining the data within the context of the larger system. A composite model comprises elements of the other two, including distributed responsibility coupled with a core group of GIS expertise.

- *Data management:* In the composite data management model for a BNL GIS, each department creates and maintains a particular subset of the site data, while a GIS Coordination Group (GCG) made up of representatives from each operational group is formed to deal with data-sharing issues. This allows each group to use the software it prefers and permits physical as well as electronic access control while facilitating critical coordination efforts.
- *Applications and analysis:* The composite model retains the concept of a GIS support group but includes as well distribution of GIS analyses and applications development. The menu-driven GIS software with built-in elementary applications (Sec. 3.4) enables casual users to query the system and to generate simple maps and reports for their own use. The support group initially serves as applications developers for all users. High priority for initial development is the creation of data layers and associated applications which all divisions have in common (e.g., screen and print maps of the site boundary, wells, concentrations of specific compounds). With time each department will begin to develop specific applications in-house. The GSG then functions as a consultant to the user community.
- *Local support:* On-site support to GIS users is centered in the GIS Coordination Group and the GIS Support Group. These two groups may be one and the same, and each may be an independent entity or may be made up of user representatives from the participating groups. The GSG may be headed by an experienced GIS expert brought to the Lab to help implement and manage the site-wide GIS.

Because BNL does not have a large experienced user community, the effectiveness of a support group made up solely of local users will depend on the ease of learning and user friendliness of the software, on the quality and quantity of training provided by the vendor, on the availability of locally based vendor support, and on the amount of other demands placed on the time of the group's members.

5 Issues to be Resolved

The conceptual models for a site-wide GIS for the Laboratory take into account the mutual needs and requirements of the three participating groups and suggest different ways in which responsibility for data management, applications development, spatial analysis, and user support may be configured. Many questions have been identified that will affect the selection of the most appropriate GIS model for BNL. Some of these have been alluded to directly or indirectly in previous discussions in this report. Others will arise during the course of implementation of the GIS. Among these issues are the following:

Hardware issues:

- Will the GIS be required to run in a PC or Macintosh environment?
- Will the PCs access the GIS as dumb terminals (X-terminal emulation software)?
- Will each department be required to purchase additional hardware to access the GIS (UNIX workstations, plotters, additional disk space, etc.)?
- Will the current BNL network be sufficient to handle the GIS information flow?

Software issues:

- The need to interact with AutoCAD and other software has been well documented in the departmental assessments. How will this interaction be established?

Data management issues:

- Will each department be responsible for maintaining its data in a "GIS-ready" format or will there be a separate support group to do this?
- How will the development of GIS data layers be prioritized?

Application, analysis, and support issues:

- Will each department be responsible for developing its own applications, data analysis, and supporting its portion of the GIS system, or will there be a separate support group available to all?
- How will multi-departmental applications be handled?

- How will development of all GIS applications and analyses be prioritized?

General:

- How much flexibility must be built into the system to accommodate expansion, changing needs, and changing requirements?
- How will the cost of the system be shared among the participating departments?

To attempt to answer these questions without the participation of the three operational groups is beyond the scope of the tasks assigned to the GIS Task Group. However, the resolution of these issues will have a direct effect on the conceptual model chosen, the final specification for selection and purchase of a GIS, and the implementation plan.

ACRONYMS

AOC	Area of Concern
BNL	Brookhaven National Laboratory
CAD	Computer Aided Design
CCD	Computing and Communications Division
CERCLA	Comprehensive Environmental Response, Compensation and Liability Act
CPU	Central Processing Unit
DBMS	Data Base Management System
DOE	U.S. Department of Energy
FIPS	Federal Information Processing Standard
GCG	GIS Coordination Group
GIS	Geographic Information System
GSG	GIS Support Group
HSR	Historic Site Review
LILCO	Long Island Lighting Company
LIRR	Long Island Rail Road
NFS	Network File System
OER	Office of Environmental Restoration
OU	Operable Unit
PC	Personal Computer (here refers to IBMs and IBM clones)
PE	Plant Engineering
PI	Preliminary Investigation
QA/QC	Quality Assurance/Quality Control
RA	Remedial Action
RCRA	Resource Conservation and Recovery Act
SDTS	Spatial Data Transfer Standard
SEP	Safety and Environmental Protection
SARA	Superfund Amendments Reauthorization Act of 1986
SI	Site Identification
SCWA	Suffolk County Water Authority
SGPA	Special Groundwater Protection Area
SQL	Structured Query Language
USGS	United States Geological Survey
UTM	Universal Transverse Mercator

Table 1. Identified GIS Applications at BNL

These GIS applications were identified from the needs analysis conducted with OER, SEP, and PE.

APPLICATION	OER	SEP	PE
Produce maps and figures through a user-friendly interface	X	X	X
Plot and display county-wide dosimeter measurements		X	
Track groundwater plume over time	X	X	
Storage tank location and data		X	X
Location and status of fire hydrants		X	
Tracing the distribution and abundance of flora, fauna	X	X	
Water table maps	X	X	X
Water quality maps using a variety of display techniques	X	X	
Storm water runoff	X	X	X
Supervising field work	X		
Risk analysis to private well owners	X		
Interface with PE CAD system	X	X	
Interface with site environmental databases	X	X	
Interface with spatial statistics package	X	X	
Connection with air dispersion model		X	
Set up and post-process data files for groundwater modeling	X	X	
Integration with other application software	X	X	X
User-friendly interface to enter/transfer existing data	X	X	X
Tracking and managing field data	X		
Geologic log database	X	X	
Hazardous chemical inventory		X	
QA for environmental sampling	X	X	
Sewer line sampling		X	
Fire extinguisher sprinkler system maintenance	X	X	
Aerial photo inventory and retrieval system	X	X	X
PE historic drawing storage and retrieval system	X	X	X
Potential AOC photo inventory and retrieval system	X		
Historic Site Review photo inventory and retrieval system	X		
Coordination with other departments on restoration projects	X	X	X

Table 2. Projected GIS Data Requirements at BNL

This list of data layers was developed based on the applications identified in Table 1. The data layers have been divided into environmental and facilities groups. Each group is further divided into data layers generated and maintained within the Laboratory and those that are maintained by outside (external) agencies. Part A identifies which groups could be responsible for the maintenance of these layers in the BNL site-wide GIS. Part B identifies the current format in which the data exist, either at BNL or off site.

A. Potential Maintenance Responsibility

DATA LAYER	OER	SEP	PE
Environmental - internally maintained			
AOCs	X		
Current and former landfill data	X	X	
Digital Historic Site Review (HSR) photos	X		
Digital potential AOC photos	X		
Field supervision data	X		
Groundwater analysis data	X	X	
Hazardous chemical database		X	
HSR photograph location cover	X		
Operable Unit (OU) boundaries	X		
Potential AOC photograph location coverage	X		
Potential AOCs	X	X	
Project locations (PA/SI, OU tasks, etc.)	X		
RA locations	X		
Radioactive soil areas		X	X
Sample points/monitoring wells data	X	X	
Site fauna	X	X	
Site flora	X	X	
Soil	X	X	
Spray aeration	X	X	
Surface hydrology		X	X
Topographic contours		X	X
Environmental - externally maintained			
208 zones	X	X	
Coastline	X		
Heritage survey data	X		
Off-site surface hydrology	X	X	
Off-site water table elevations	X	X	
Parcel centroids	X		

Table 2A, continued

DATA LAYER	OFR	SEP	PE
Environmental - externally maintained			
Private well contamination data	X	X	
Public water customers	X		
Regional development activities	X	X	
Regional potential contamination sites	X	X	
SGPA zones	X		
Facilities - internally maintained			
BNL grid			X
BNL roads			X
BNL special areas (e.g., biology fields)			X
Building floor plans			X
Building footprints			X
Digital air photos		X	X
Digital historic drawing images	X	X	X
Drainage basins	X	X	X
Fire extinguisher and sprinkler data		X	X
Fire hydrant data		X	X
Historic drawing index	X	X	
Lab property boundary		X	X
Misc. BNL areas (tennis courts, ball fields)			X
Potable and supply wells	X	X	X
Sewer lines		X	X
Storage tanks		X	X
Storm water drainage		X	X
Utilities			X
Facilities - externally maintained			
County-wide dosimeter data		X	
LILCO	X		
LIRR	X		
Public roads	X		
SCWA wells/future sites	X	X	

B. Current Data Availability

Because some of these data layers may be maintained by more than one division (Table 2A), the data may exist in more than one form. Non-graphical formats include spreadsheets, relational databases, and flat (ASCII) files.

DATA LAYER	Digital			Non-digital	
	GIS	CAD	Non-Graphic	Hard copy	Not Determ
Environmental - internally maintained					
AOCs				Y	
Current and former landfill data		P	P	Y	
Digital Historic Site Review (HSR) photos				Y	
Digital potential AOC photos				Y	
Field supervision data					Y
Groundwater analysis data		P	P	P	
Hazardous chemical database					Y
HSR photograph location cover				Y	
Operable Unit (OU) boundaries				Y	
Potential AOC photograph location coverage				Y	
Potential AOCs				Y	
Project locations (PA/SI, OU tasks, etc.)				P	
RA locations				Y	
Radioactive soil areas		Y			
Sample points/monitoring wells data		P	P	P	
Site fauna					Y
Site flora					Y
Soil					Y
Spray aeration		P	P	P	
Surface hydrology		Y			
Topographic contours		Y			
Environmental - externally maintained					
208 zones	Y				
Coastline	Y				
Heritage survey data				Y	
Off-site surface hydrology	Y				
Off-site water table elevations			Y		
Parcel centroids			Y		
Private well contamination data			P	P	
Public water customers			Y		
Regional development activities					Y

Table 2B, continued

DATA LAYER	Digital			Non-digital	
	GIS	CAD	Non-Graphic	Hard copy	Not Determ
Environmental - externally maintained					
Regional potential contamination sites	Y		P	P	
SGPA zones	Y				
Facilities - internally maintained					
BNL grid		Y			
BNL roads		Y			
BNL special areas (e.g., biology fields)		Y			
Building floor plans		Y			
Building footprints		Y			
Digital air photos		Y			
Digital historic drawing images				Y	
Drainage basins		Y			
Fire extinguisher and sprinkler data					Y
Fire hydrant data					Y
Historic drawing index				Y	
Lab property boundary		Y			
Misc. BNL areas (tennis courts, ball fields)		Y			
Potable and supply wells		Y	Y		
Sewer lines		Y			
Storage tanks					Y
Storm water drainage		Y			
Utilities		Y			
Facilities - externally maintained					
County-wide dosimeter data					Y
LILCO	P				
LIRR	P				
Public roads	P				
SCWA wells/future sites	P				

KEY: Y = Yes
P = Partial
(blank) = No

Table 3. Computing Environment at BNL

The following list of hardware and software comprises the environment within which a site-wide integrated GIS will have to operate. The list does not include either hardware or software which are unlikely to affect the implementation or use of a GIS. Operating systems are not included; all the workstations run UNIX, and both DOS and Windows exist on the PCs.

HARDWARE	OER	SEP	PE
Workstations	planned	X	X
PCs	X	X	X
Macintoshes		X	X

SOFTWARE	OER	SEP	PE
AutoCAD	X	X	X
Surfer	X	X	
Grapher	X	X	
Excel		X	X
Lotus 1-2-3		X	X
Quatro Pro	X	X	
WordPerfect	X	X	X
WordPerfect Office		X	
DrawPerfect	X	X	
DataEase			X
dBASE		X	X
Oracle			
R:Base	X	X	
Sybase	planned	X	
FoxPro		X	
Progress		X	
SAS	X	X	

NETWORKING	OER	SEP	PE
Modems		X	X
Gandalf lines			
Ethernet connections	planned	X	X
Software:			
PC-NFS	planned		X
TCP-IP	planned	X	
VINES		X	

APPENDIX A

QUESTIONNAIRES

1 NEEDS ANALYSIS FOR A SITE-WIDE GIS REQUIRED INFORMATION

GOAL: The establishment of a site wide system for the management and shared use of spatial data in a network environment using tools appropriate to the user's needs and requirements.

Current status of:

- Data bases
- Software used
- Size

Problems/limitations:

- Graphics capabilities
- Hardware platforms

Critical functions:

- Questions that must be answered
- Outputs
 - Reports
 - Maps
 - Model runs

Who will be using the GIS/DB?

What types of data will be required? (Spatial and non-spatial components)

- Air
- Groundwater, including wells
- Surface water
- Terrain
- Land cover
- Infrastructure
 - Buildings
 - Utilities
 - Roads

Scale/resolution of spatial data:

Commitments, requirements, drivers:

- Environmental

- CERCLA, RCRA
 - Improve accuracy and availability of site infrastructure data
 - Improve efficiency of operations and support organizations
 - DOE-HQ Order 4320.1B (1/7/91) Site Development Planning Policy
 - Section 7-h
 - DOE-RL Operations Directive and Program Guidance

2 USER NEEDS QUESTIONNAIRE

What are your daily responsibilities?

What products are you responsible for producing?

What information do you need to produce these products?

What GIS applications would you like to see developed?

How do you envision your role in utilizing the GIS system (casual user, active developer, etc.)?

APPENDIX B

GIS DATA SURVEY FORMS

GIS DATA SURVEY			
PAGE 1 of 3			
	First	Last	
FILLED OUT BY:			ORGANIZATION:
PHONE EXT.:			DATE:
DATA THEME:			
REASON FOR AUTOMATION:			

SURVEY Edit Table: DATA_THEME Page 1 <17, 1>

GIS DATA SURVEY	
PAGE 2 of 3	
MAINTAINED BY:	
MAINTENANCE FREQUENCY:	
DATA SOURCE:	
GRAPHICAL DATA?	(Yes [Y] or No [N])
NON-GRAPHICAL DATA?	(Yes [Y] or No [N])
DATE ORIGINATED:	(Enter date in MM/DD/YY format)
AUTOMATED?	(Yes [Y] or No [N])
REGISTRATION?	(Yes [Y] or No [N])

SURVEY Edit Table: DATA_SOURCE Page 2 <20,5>

GIS DATA SURVEY	
PAGE 3 of 3	
SCALE:	
FEATURES TO CAPTURE:	

SURVEY

Edit

Table: DATA_SOURCE

Page 3 <20, 5>

APPENDIX C

GIS IMPLEMENTATION CONSIDERATIONS

1 System Design

1.1 *Software choice.* A site-wide Geographic Information System may be based on a single software product or may involve integration of more than one package. The single-product approach is conducive to ensuring that all divisions are working with the same set of information, which is the principal goal of this project, and simplifies applications development and system support. However, use of more than one GIS might facilitate division-specific needs. A critical element in this choice is the degree of compatibility between the systems. The need to import and export files among divisions would place an additional administrative burden on system overseers in order to maintain quality control. The principal multiple-product option to be considered is a system for Plant Engineering that interfaces directly with AutoCAD files plus a full-function system for OER and SEP.

1.2 *Platform configuration.* PCs, Macintoshes, and UNIX-based workstations are currently in place. No GISs currently available for the Macintosh are adequate for the requirements of a site-wide GIS, so that platform need not be considered.

A decision on platform configuration will be based in part on the software capabilities available on the various platforms (both GIS and database software must be considered) and in part on the cost of acquiring additional hardware. PCs generally lack the speed, memory, graphics capability, and storage capacity of workstations. Speed and memory are critical considerations because digital graphics require an abundance of both. GIS software requires considerable storage space, as do the data layers and the attribute databases. Both initial platform capability and future expandability must be considered.

1.3 *Processing configurations.* A site-wide GIS could exist on several isolated machines. In order for this configuration to be regarded as a "site-wide" system, some centralized GIS entity would be required to ensure that the data and software were identical on each computer. The alternative is to have all GIS users linked through the BNL Ethernet network. Network options include:

- Workstations only (UNIX only). The GIS resides on a file server which each division accesses by way of workstations over the Ethernet network.
- Mixed platforms (UNIX only). The GIS resides on a file server that can be accessed by either workstations or PCs acting as X-terminals.
- Mixed platforms (mixed operating systems). The GIS resides on a file server that can be access by either workstations running UNIX or PCs running DOS or Windows.

2 Cost of Implementation

2.1 *Initial costs of software.* Software costs associated with development of a site-wide GIS will include the purchase of the GIS itself and possibly of other software as well. Commercial GIS software ranges in price from \$500 for a single PC-based package to \$100,000 or more for a workstation-based package licensed for multiple CPUs. In between lie products which are licensed for multiple users (usually about 3) accessing a single CPU. Costs vary widely according to system capabilities. Basic systems may not meet BNL functional requirements and may require the purchase of additional software modules. Low-end/low-cost systems will not be flexible or powerful enough to meet site-wide needs and are likely ultimately to require the purchase of additional copies of the software. Public domain GISs are also available. These may be obtained free over Internet via FTP (File Transfer Protocol). They must then be compiled and set up on the local system. Pre-compiled versions, along with some rudimentary technical support, may be purchased for \$1000-\$1500 from consulting firms and other organizations.

Continued use of existing PC-based database software is not a viable option for Brookhaven, since it is rapidly being outgrown. The Sybase DBMS is currently available from CCD over the Ethernet. However, use of Sybase requires the purchase of applications development software. For example, OER and SEP would have to purchase these modules separately if both groups are to have access to the software. The price of commercial database software varies according to the type of license (numbers of CPUs and multiple users). If the UNIX-based mixed platforms configuration were selected, UNIX or X-terminal emulation software would have to be purchased for PCs.

2.2 *Initial costs of hardware.* Additional workstations would have to be purchased according to the platform configuration selected. A high-end file server might be supplemented with one or more workstations (including 20-inch high-resolution color monitors) in each division where the GIS could be remotely mounted and accessible by PCs and X-terminals. The number and size of workstations would be linked to the allocation of responsibility for system development and maintenance.

In addition to computers, required hardware will include one or more digitizing tablets or tables, scanners, and plotters. This equipment is already in place in Plant Engineering. The most likely additional need will be for plotters and printers as responsibility for map production is decentralized. OER plans to order an electrostatic plotter which they would make available to the other departments.

2.3 *Network costs.* SEP and Plant Engineering are already on the Ethernet system, and OER is in the process of obtaining the necessary communications link and network software. No additional costs beyond the usual and customary would be anticipated unless heavy network traffic becomes a severely limiting constraint.

2.4 Data acquisition costs. Initially data acquisition costs will be relatively low, since the GIS will be built on existing AutoCAD maps and drawings and on existing databases. However, personnel costs will be incurred as data are transferred into the GIS and a new database system. These costs will be reduced if the GIS can work directly with AutoCAD files or if the translation is straightforward, but significant time may be required to make existing CAD layers topologically current. As the GIS is developed, digitizing of non-digital spatial data will be an ongoing cost. Data may also be purchased from government agencies such as the U.S. Geological Survey (USGS) or through commercial sources or consulting firms. It is possible that data acquisition costs could be reduced by sharing data with other users (e.g., county agencies).

Costs associated with development of the attribute databases are a separate issue and would occur regardless of the existence of a GIS. However, the existence of these databases within the context of a networked GIS, with read access to users in each division, would reduce costs by avoiding duplication in database development.

2.5 Application development costs. Applications development costs are difficult to estimate in advance, since they are almost entirely personnel costs. Desired applications must be prioritized according to complexity as well as need, since complex applications may take months to create. The difficulty of the GIS programming language, the level of expertise, and the presence or absence of a menu-driven applications development tool will affect the time required, and hence the cost, especially in the early stages of GIS implementation.

2.6 Maintenance and upgrade costs. Vendors of commercial GISs generally charge an annual maintenance fee after the first year. This fee covers user support services and normally includes the cost of any upgrades during the year. It may not include the cost of purchasing a new release of the software, but would provide for a discount for current licensees. Annual maintenance costs vary with the number of seats and licensed CPUs. In a recent industry survey¹⁰ most vendors who reported their maintenance fee quoted a figure of 10 to 15 percent of the base price of the software.

3 Time to Implement

The time necessary to get a site-wide GIS up and running in a useful way will depend on the tasks required and on the ease with which users become comfortable with the system.

3.1 The learning curve. One aspect of becoming "fluent in GIS" is the familiarity of potential users with the computers and operating systems with which they will be working. Most potential users at BNL are experienced users of PCs and are familiar with DOS and Windows. Fewer are familiar with the

¹⁰ 1991-92 International GIS Sourcebook. 1991. Fort Collins, CO: GIS World, Inc.

UNIX and X-Windows (or Open Windows) environment of workstations. Only one individual within the three divisions has extensive experience with a GIS.

Another important point is the complexity of the software itself. A low-end GIS can be mastered relatively quickly by most users, but its potential applications are limited. With a high-end GIS simple tasks can be carried out by most novice users, but mastery of the system takes time. A full-function GIS is a complicated system that requires both understanding and experience in order to utilize it to its fullest extent. A knowledgeable user must be familiar with the fundamental concepts of Geographic Information Systems and with concepts such as spatial interpolation and spatial autocorrelation.

An intuitive, menu-driven front-end will help novice users accomplish tasks more quickly than if they had to learn commands. Most low-end GISs have a basic, built-in menu system, since they are aimed at non-experts who have a limited number of relatively simple applications in mind. Many major commercial GISs, especially the older ones, require the user to type in the usual command lines, although this is changing as new versions are released. Newer GISs are more likely to be entirely menu-driven. In most GISs Graphical User Interfaces (GUIs) can be developed which are tailored to the users' specific needs. Some GISs offer the user a choice: novice users can use menus, while experienced users may prefer to revert to the command line.

Because virtually none of the potential users of a site-wide GIS at BNL are experienced in GIS, the availability of training is an important consideration. Most low-end GISs, especially those that are PC-based, offer no formal training, although a tutorial usually accompanies the software. Training classes are usually available through the vendor for mid- to high-end software. On-site training is an important option. While more expensive than training classes held at the vendor's headquarters or satellite offices, the costs are usually lower per capita for on-site training when more than two or three people participate.

3.2 Database design and implementation. A GIS remains nothing more than software until spatial and attribute databases have been developed. This can easily be the most time-consuming and costly aspect of creating an operational GIS. The time required to implement a site-wide GIS at BNL will be shortened if a GIS package is selected which can utilize the Laboratory's current digital databases. Nonetheless, transfer of these data into the GIS environment will require careful attention to quality control. Site maps will have to be "cleaned up" (e.g., open polygons must be closed, dangling lines removed) and databases will have to be brought into conformity (e.g., consistent labeling of wells, all locations expressed in the same coordinate system). Time will also be required to add non-digital data to the GIS. This task will include design of new attribute databases and the keying in of data as well as the digitizing of additional spatial information and development of new spatial data layers.

A critical aspect of these tasks will be the coordination of effort among the divisions involved. Design (or redesign) of attribute databases must take

into account the links to the spatial data layers and the potential use of the data by other divisions. Indeed, data management (maintaining spatial metadata and specifying requirements and responsibilities for database changes and upgrades) will be at once the most important and the most difficult aspect of implementing a site-wide GIS. Additional time should be allotted for this coordination effort.

Database design and implementation can be done wholly within the Laboratory. This option could be less costly if only current personnel were involved. However, it would probably take more time, since GIS training will be required first and because the people involved will have other responsibilities. Alternatively, a consulting firm could be used to coordinate the database efforts. A number of consulting companies exist which specialize in developing GIS databases. While costly, these consultants could save time by working simultaneously with all the divisions and would provide experience which BNL lacks.

4 System Maintenance and Support

An operational Geographic Information System is made up of three elements: software, data, and applications. Each of these elements must be continually maintained and supported for a site-wide GIS to be successful.

4.1 *Software upgrades.* The site-wide GIS, including GIS, DBMS, and network software, will be periodically upgraded. Unless the system is completely centralized under the control of a core GIS maintenance group, those responsible for the upgrades will need to apprise the other users in advance of anticipated software changes. It will be necessary to ensure that an upgrade in one part of the system does not compromise its accessibility. System configuration, hard and soft links, and applications must all be considered, since changes in one part of the system may have repercussions throughout the system. Centralization of system maintenance would reduce the risk of system disruption.

4.2 *Data management.* Data ownership and responsibility will potentially be one of the most difficult questions in the development of an integrated Geographic Information System for the Laboratory. Each division currently has responsibility for its own data, and most of the data in the system will continue to be generated by the divisions involved.

One option for dealing with data updates is to centralize the responsibility. Each division would deliver its raw data to a core data management group, which would in turn be responsible for data entry and quality control. Different levels of access would be built into the system, with only the data managers having write and delete privileges. This option would ensure that everyone was using the same information. However, it would be necessary to set up a prioritization scheme, so that the data managers' time was appropriately allocated and critical tasks were completed first. Formal lines of communication with the divisions would have to be instituted and quality

control procedures defined which met the requirements of all users. Under this option, the divisions would effectively relinquish control over the data they generate, although data security and integrity procedures would still be utilized.

Alternatively, some form of the status quo could remain in effect, with each division being responsible for particular data. Data management would take place within the divisions, but levels of access would be defined to allow users in other divisions to read and use the data. Some current methods of tracking these data may have to be changed to allow access by the GIS. Divisional data managers would need to work with all users to define quality control parameters and users would need to be informed whenever an updated version of a database was placed in the system. In some cases updates might simply be done on a stipulated regular basis (weekly, monthly). This option, while perhaps more difficult to implement well, would have the advantage that each group would be responsible for the data that it is most familiar with. The key to success would be inter-divisional cooperation and communication.

A third option might be a compromise, with data management occurring at two levels. The divisions would have primary responsibility for their data, including first-order quality checks. These data would then be made available to the system and also to a central data management group, which would be responsible for integrating and archiving the data. This option would facilitate prioritizing of critical data management tasks while freeing the divisions of long-term maintenance of the data.

4.3 Application design. The third element of a GIS is applications. While many GIS tasks can be accomplished "on the fly," frequently the same task may be required many times. Applications can be developed to automate map generation, routine queries, data building, and standardized reporting.

As with data management, applications development may be carried out by a core GIS entity or by individual divisions. In some cases an application may be equally useful to several divisions, the generation of a simple base map, for example. In other cases the needs of a division, or an individual within a division, may be unique.

END

DATE
FILMED

6 / 29 / 93