

## The Digital Library: Serving a Business Purpose in an R&D Laboratory

*Robert S. Gresehover, DeAnna T. Jones, and Paul S. Tarantino*

**T**he R. E. Gibson Library and Information Center of APL is presently rebuilding its capabilities to serve the business purpose of a research and development laboratory. This effort is focusing on the technological, information resource, and collaborative work components of modern digital libraries. The rebuilt library will not be a physical place, but rather a dynamic information environment where Library staff can rapidly, and with relative ease, tailor information products to the specific needs of its customers. (Keywords: Collaboration, Digital library, Digital resources, Information technologies.)

### INTRODUCTION

The concept of the digital library has received significant national attention in recent years owing, in large part, to an interagency federal program called the Digital Libraries Initiative.<sup>1</sup> This initiative—originally supported by the National Science Foundation, the Defense Advanced Research Projects Agency, and NASA—funded the research and development of tested activities in a number of digital library areas. It is now in its second phase and has broadened its scope of objectives as well as its list of government and corporate sponsors.<sup>2</sup> Meanwhile, individual public and private libraries have begun to evolve into digital libraries serving the unique requirements of their particular communities,<sup>3,4</sup> and specialized digital libraries are being created to meet highly specific information needs.<sup>5</sup>

This article describes efforts within APL to create a digital library that integrates three components: (1) applied information technology, (2) digital business and

technical information resources, and (3) collaboration between APL information professionals and their customers (the APL user community). We believe that these components define a digital library. Each is described in more detail later in the article.

A traditional science and technology library was established at APL in 1945, only 3 years after the founding of the Laboratory. Since its inception, the APL Library has been part of a system of libraries corresponding to the various divisions of The Johns Hopkins University. Each library develops collections and services that align with the particular needs of its primary clientele while serving as an active resource for the whole University.

A new library building was constructed in 1963 on the Laboratory's Howard County Campus. It was subsequently dedicated as the R. E. Gibson Library in 1969 in honor of APL's retiring Director, Ralph Gibson. Throughout its long history, the Library has regularly

rebuild its service capabilities to meet the changing information needs of the Laboratory, to manage an ever-expanding quantity of information, and to apply information technologies to maintain processing efficiencies and service delivery effectiveness.

The Library was an early adopter of new information technologies to support APL's diverse technical program areas and new business development efforts. It manages computer systems that provide Laboratory-wide access to a large number of commercial, academic, and government resources, as well as information resources created within APL. In addition it offers an information environment for individuals and groups to access and share information while ensuring compliance with the licensing agreements and copyright protections of the publishers.

We believe that the successful operation of a digital library depends on regular collaboration between our staff of information professionals (and support personnel) and our customers. Library staff participate in customer planning teams, become directly involved in technical projects, support Laboratory-wide training efforts, and attend technical and military conferences and seminars along with members of the APL technical staff. The increasing availability of digital information and information technologies permits our Library staff to focus less on issues associated with managing collections and more on managing collaborative efforts with our customers. The digital library is not a physical place; although there remains a physical building, the Library is really a dynamic information environment, where our information specialists can quickly and easily tailor information products to specific needs.

## PLANNING THE DIGITAL LIBRARY

Plans for the digital library began in earnest in 1990 with the Library hosting a series of customer focus groups comprising staff from APL's technical departments. Participants were led through discussions to identify their collective vision of a future library based on their experiences with libraries in other organizations, their present information needs, and their predictions about the future impact of technology. They were also asked to identify specific products and services that should be available through the Library, and then to rank these in terms of their relative importance. The over 100 responses were grouped into five categories which clearly showed what our customers were looking for, i.e., the ability to

1. Access APL-generated and externally published information online
2. Acquire or borrow materials from a broad number of sources outside of APL

3. Access a broad subject scope of periodical literature to include business and world strategic information
4. Order information materials online
5. Perform end-user searches of online bibliographic and full-text databases

The compilation of customer needs derived from these focus groups served as the basis for our strategic direction into greater use of applied information technologies.

The process of planning for our digital library is iterative and continuing. We recently assembled a Laboratory-wide team to examine issues associated with the creation and management of APL-generated documents. A number of document-processing models were developed and discussed by the team to achieve a balance between local control and Laboratory-wide identification of internal documents. Central to the team's input was the need for applied information technologies to support APL-wide search and retrieval, version control, and document creation work flow.

Obviously, the availability of information technologies, particularly Web-based technologies, has also played an important role in guiding our planning effort. The relative ease of implementing computer servers and creating documents, along with the proliferation of common gateway-interface (CGI) routines that allow users to interact with traditional database systems via browsers, has spurred the explosion of digital government and commercial information. In turn, the increasing quantity of digital information has caused us to identify cost-effective ways to bring it to our customers.

At the same time, the move to increase the availability of digital information has raised concerns about providing access to archival digital publications. Although some publishers are beginning to articulate their plans to offer continued access to archival collections of their digital publications, doubts remain regarding the intent and ability of others to maintain such access. The same issue applies to government information sources, particularly descriptive publications of military platforms where configuration information on older platforms may be significant. Until we feel confident in our ability to access archival digital collections, we will continue to maintain access to printed documentation.

## COMPONENTS OF THE DIGITAL LIBRARY

### Applied Information Technology

The first computer technology used within the Library included computer terminals and modems to access remote bibliographic databases, e.g., Defense Technical Information Center (DTIC) and Dialog. As application software for libraries became available, we acquired a

system to help manage our books and periodical collections and provide users with an online public access catalog (OPAC). OPAC enhanced our customers' ability to locate books within the collection by giving them multiple points of entry (e.g., author, title, subjects, keywords), and made the processing of materials (acquisition, cataloging, and circulating) much more efficient and less labor intensive for Library staff.

With the incorporation of an OPAC came the need for Library staff to have workstations to interact not only with our own OPAC but also with remote databases to find, search, and download the online cataloging records of other institutions. In addition, users wanted to be able to search OPAC from their own workstations. This need focused our effort to bring information and resources to the users' desktops, first through serial connections (telnet sessions) and later using TCP/IP over Ethernet.

The goal of making all Library information available to all APL staff at their desktops is ambitious. It assumes that the information is available in digital form, that it is in a format that can be delivered to each desktop (regardless of workstation platform), and that customers can navigate to and through information resources once they are available. The Library has experimented with information delivery for many years. One early attempt—the online delivery of a popular newsletter, *Inside the Navy*, by Inside Washington Publishers—proved very successful. The publisher was willing to deliver the information to the Library as attachments to e-mail from its site, but it was up to Library staff to format the text so that it could be viewed within APL via browser (Netscape/Internet Explorer). The Library staff developed a Microsoft Word macro to automatically convert the newsletter, formatted for printing, into HTML for access and viewing on the Laboratory's internal web, i.e., the "intranet" (Fig. 1). Through early efforts such as this, Library staff became familiar with the implications of providing online information in usable formats to customers.

While networking technologies were evolving, CD-ROMs became a popular distribution medium of commercial and government publishers. The challenge was to provide the same level of remote access to these CD-ROMs as we were able to provide with OPAC. Achieving remote access was problematic since the database structure and the search and retrieval application

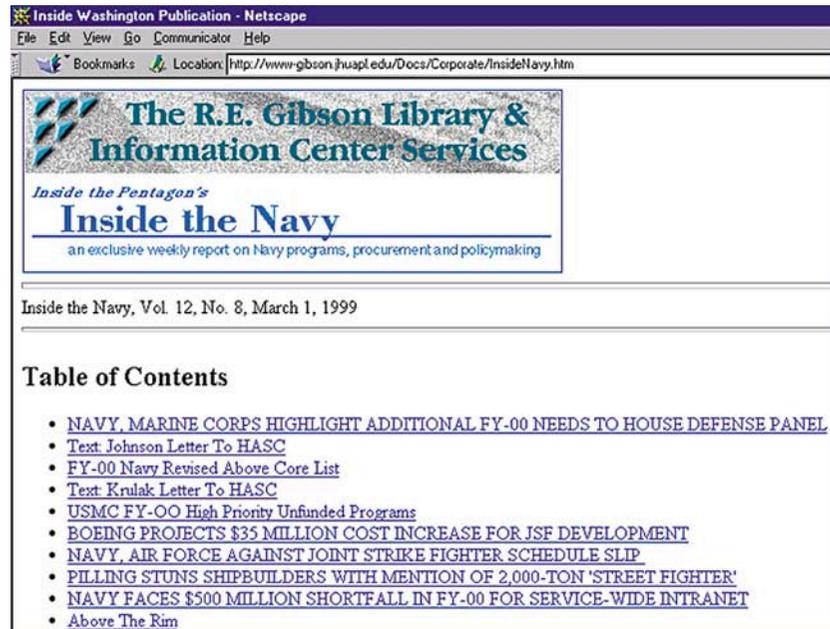


Figure 1. Web newsletter page with table of contents created from an e-mail attachment.

software contained on each CD-ROM were specifically tailored to each information product, and the application software was platform specific. We were able to ameliorate some of the latter problem using WinFrame server software<sup>6</sup> and Independent Computing Architecture (ICA) protocols that allow applications to run on a server while passing screen display information to the workstation. The last piece of enabling technology was the World Wide Web with its suite of publishing, searching, and object-linking capabilities. When a CGI became available in 1995 for our OPAC, we made this interface available to our customers as well.

The Library maintains four Windows NT-based servers that support OPAC, a document management system, a networked CD-ROM system, corporate and government information sites, a conferencing system, and a customer information/help desk system. (The amount of mass storage required for these systems and associated files is approximately 150 GB.) Except for OPAC, all of these systems reside behind the Laboratory's firewall, i.e., they are only accessible to APL staff through the intranet.

## Information Resources

We have replaced numerous hard-copy subscriptions with online information services and posted ever-increasing amounts (and categories) of information for APL clients online via the intranet. By making this information available electronically, and by continually educating our customers in the use of these resources, we reduce the number of simple information retrievals that require hands-on intermediation by Library staff.

Internet-savvy clients can get much of the information they require immediately and without assistance. Figure 2 shows the gradual decline in customer requests fulfilled by Library staff and the sharp increase in the number of customer accesses (hits) to our Web resources. The digital information resources managed by the Library fall into three categories: (1) open literature published commercially or by technical and scientific societies, (2) government (military) information, some of which is classified or has limited distribution, and (3) corporate, proprietary APL information, including internal memoranda and reports.

Regardless of the kind of resource (open, government, or internal), our customers expect immediate online access to all information. They are no longer content to search a database of citations and then wait for the identified primary information to be delivered some time later. The hyperbolized claims of some commercial Internet providers have created expectations of immediacy that cannot be met in reality. Much of the scientific and technical information needed to support the business needs of an R&D laboratory remains available only in printed form. However, that reality is changing rapidly (as is evident, again, in Fig. 2). The Library continues to acquire digital information resources as they become available and adopt information technologies that will help APL access and use these new formats.

### Open Literature

To date, the Library provides APL staff with desktop access to over 1000 digital journals, ranging in subject matter from biomedicine to space science. We offer customers a complete listing of the available digital publications as well as hypertext links to publications via our Information Navigation System<sup>7</sup> (Fig. 3), which

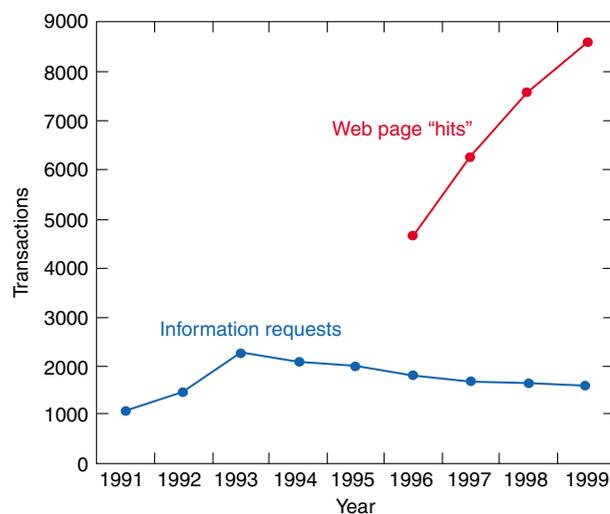


Figure 2. In-person information requests versus Web access.

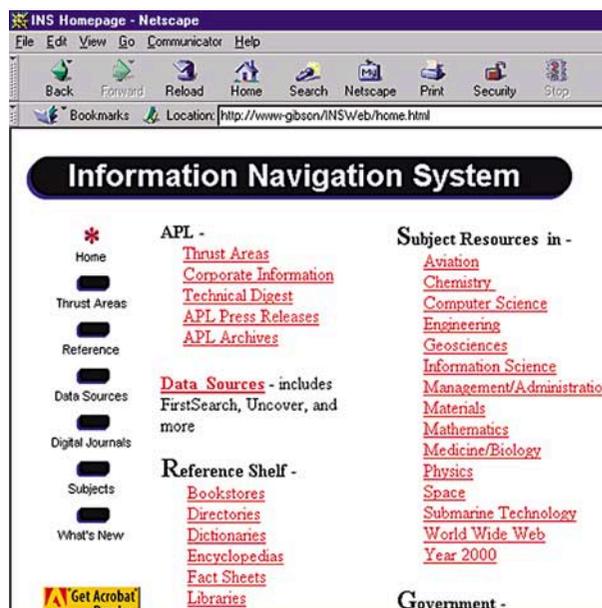


Figure 3. The Library's Information Navigation System (INS).

includes hypertext links to the digital publications via the searchable OPAC. APL staff can read the full text of these digital journals from desktop workstations since the entry process developed by many publishers recognizes the IP address of APL during logon. When use of the IP address is not available from the publisher, staff link to a Library Web server containing an automated script that transparently accomplishes the user identification and password negotiations with the publisher's online system.

The Library must often adjust its information delivery systems to match the evolving technological advancements of publishers. For example, the Jane's series, one of the most popular information resources used by our customers, has applied a number of such improvements to its digital product offering. The challenges of early versions of the digital Jane's, as with other valuable resources originally published in print, were how to deliver the information with an easily searchable interface and how to include all of the information and image files to many different workstation platforms and operating systems. Jane's initially delivered the information on CD-ROMs, often changing the search interface and expanding the number of disks as the content expanded. These changes required the Library to change its CD-ROM network configuration and its instructions for users. The publisher then experimented with delivering the databases as raw information in HTML, requiring the use of third-party search engines to find and use the information. Currently, Jane's is available with a Web-based interface, allowing Library customers to search and retrieve text and images using a standard Web browser.

In addition to primary publications such as digital periodicals, newsletters, handbooks, and other reference works, the Library provides APL staff with direct access to several digital publications of secondary information like online indexes and abstracts. The most notable of these end-user online resources is First Search, which allows online access to the union catalog of the Online Computer Library Center, Inc. (OCLC). In addition, a number of general and specialized journal article indexes are made available, e.g., UnCover, which enables the means to search the table of contents of thousands of journals, and Dialog Web, which gives users access to the world's largest online collection of science and business information.

### **Government Information**

The Library also manages government information, specifically military documents. Among the many government-created digital files made available to APL staff are *Commerce Business Daily*, Defense/other Federal Acquisition Regulations (FARs) and FAR supplements, DoD Hazardous Materials Information System, DoD R&D Program Element Descriptive Summaries, DoD/Navy/other directives and instructions, Federal Supply Catalog System data, DoD Lessons Learned, DTIC files, Federal Register/Code of Federal Regulations, DoD/federal/NASA specifications and standards, Government-Industry Data Exchange Program reports, NAVSEA technical manual index, and Navy Tactical Information Compendium.

Customer use of our online government information has steadily increased since we began making Web-based resources available. Although we suspect that the amount of such digital information will expand, the Library retains much of it in paper form. We hold more than 5000 government-created printed documents, including many special DoD publication series such as Joint Munitions Effectiveness Manuals (JMEMs), Naval Warfare Publications (NWP), Navy Operational Requirements and Master Plans, technical manuals covering NAVAIR/NAVSEA/AF/Army equipment, Data Item Descriptions (DIDs), and publications of the Tactical Weapons Guidance and Control Information Analysis Center (GACIAC) and DoD's Reliability Analysis Center (RAC). Military, NASA, and other sponsor/government handbooks, directories, dictionaries, budgets, security classification guides, and indexes are also available. National and international rules, regulations, and guides for telecommunication and spectrum management are provided as well. Approximately 100 titles are maintained covering DoD/federal news, policy, research solicitations, program/project management, and technology.

DoD and other federal agencies continue to replace paper documents with online and CD-ROM

publications which are freely accessible through government information Web sites (.mil and .gov domains). There is no accepted agency-wide standard for electronic document format, although Adobe's pdf format is growing in popularity. Likewise, no DoD standard exists for large, searchable documents/databases on CD-ROM, so users of multiple CD-ROM databases must become familiar with several database search engines and system installation procedures. There remains a vast bibliography of "legacy" government documents (e.g., documentation for older military weapon systems, some of which may remain in service for decades to come), which will move to electronic format very slowly, if at all.

Commercial providers of information services led the early conversion of government documents to electronic format, often commanding high prices for repackaging public-domain information in more useful form. The market dominance of commercial providers in many of these information "niches" will fade, however, as more and more new documents become accessible in digital form by the originating government agencies and older documents "age out," reducing the "value added" that the commercial document and information processors can provide.

The publication of electronic government documents with restricted distribution (not just classified information) to nongovernment entities with need-to-know status remains a challenge. Simple access restrictions (e.g., limiting Web page access to clients in the .mil and .gov domains), while widespread, lack needed flexibility and lock out many legitimate users. There is as yet insufficient confidence in the security provisions of the Web to permit restricted information to be posted online via open (vice dedicated, limited-access) networks. (This problem is not unique to government, but is of particular concern because of the volume and, in some cases, extreme sensitivity of limited-access government documents.)

### **Internal APL Documentation**

The last category of information resources in our digital library is APL corporate information, i.e., documents created by APL staff as part of their fulfillment of Laboratory business. The Library maintains a collection of more than 700,000 APL-generated documents. Although the collection dates back to the Laboratory's inception in 1942, we maintain materials on both recently completed and ongoing projects. The document management system used to support the creation and use of these resources permits APL staff to store documents (text and images) in their original format or in pdf. To help in the creation of documents, particularly those resulting from the work of technical project teams, the system automates the edit, review,

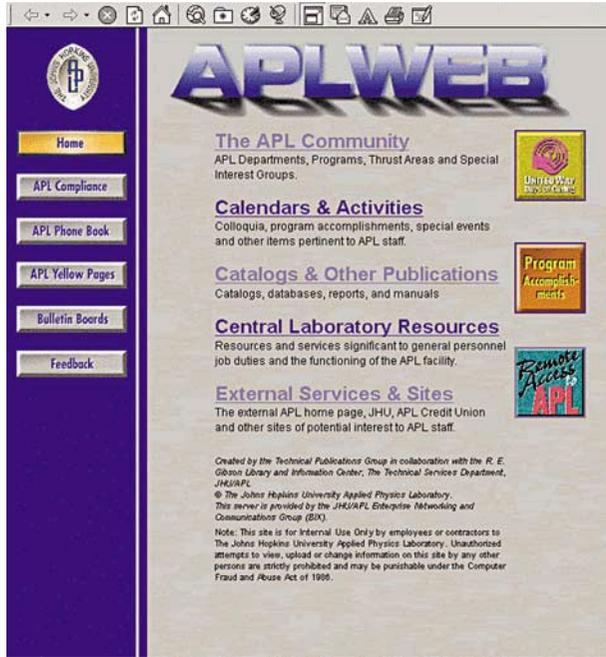


Figure 4. APL's internal home page.

and document approval process through a configurable document work flow process. Through this document management system APL staff can take advantage of document vaulting and version control capabilities.

Accompanying all documents stored in the system are searchable meta-data records that describe content attributes, e.g., author, title, subjects. If users cannot find the needed documents through this method, they may broaden their inquiry by initiating a full-text search of all documents.

The Library does not manage all of the Laboratory's internally developed information resources. Most of APL's administrative and technical departments maintain specific project-oriented digital resources and service directories. To improve access to these resources, the Library manages the Laboratory's intranet home page (Fig. 4), which provides a categorized listing of most departmental and Laboratory-wide information. Our objective in managing an APL-wide document management system and intranet home page is to provide an information resource—and a mechanism for continually enhancing the content of the resource—to help APL staff leverage the intellectual capital of the Laboratory and facilitate collaborative work.

## Collaboration

Active collaboration with our customers is the third significant component of our digital library. Using knowledge management as our conceptual framework for planning and operating, the Library develops services and products that support programs and projects sponsored by APL customers. Some of these efforts are long term, requiring us to provide updated information on a specific topic over years. Others have a short life span, requiring a rapid response to meet impending deadlines such as responding to new business opportunities. All require that the Library form collaborative relationships with our customers so that we can effectively articulate their needs, identify and locate information that will meet those specific needs, and then create customized products.

In an early effort to understand the challenges associated with supporting electronically augmented collaboration, the Library developed a digital study room (Fig. 5) that emulated an existing physical study room located within the Laboratory. Like its physical counterpart, the digital study room contained pieces of information categorized into topical areas, with each topic further subdivided into summary statements, detailed background information, and insights developed by study room participants. A rudimentary mechanism was developed to allow participants to add comments and identify other information resources. Early efforts such as this lacked the relatively low-cost virtual reality and video conferencing capabilities that are now available via the desktop workstation. Nonetheless, they demonstrated the Library's interest in and ability to support new business and technical teams as a means of sharing knowledge through electronically augmented collaboration, which represents a future direction for our digital library.

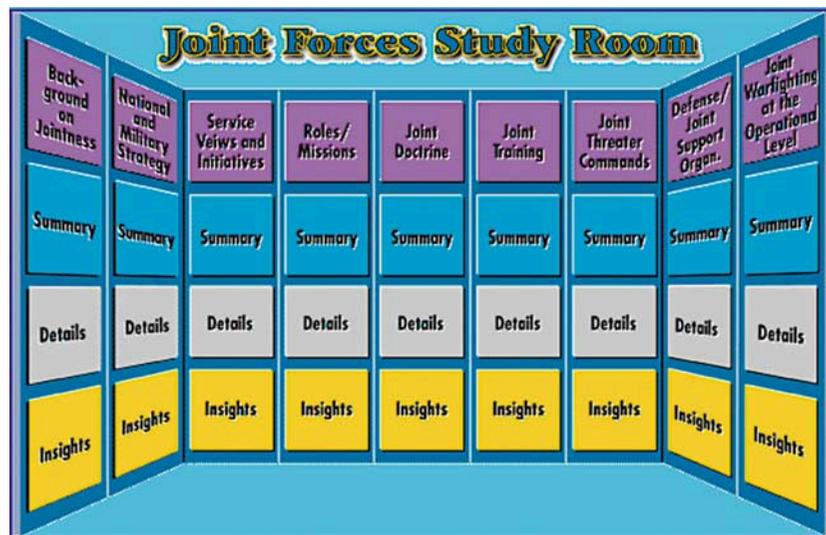


Figure 5. A digital study room developed by the Library in collaboration with customers.

## SUMMARY

Although managing information resources remains an important activity of the Library, we have begun focusing more on distilling information. Our customers are now confronted with vast amounts of information. Even after sifting out the nonrelevant, there often remains far too much to read and comprehend on a particular topic.<sup>8</sup> The challenge is to provide products and services that help staff sort out large quantities of information into manageable and actionable packages. Our efforts in this arena apply to traditional resources as well as newer digital ones. In an R&D organization like APL, we no longer view the size of a collection of books, journals, and reports as an indicator of the Library's service capability, nor its effectiveness. New quality metrics based on the effectiveness of narrowing a body of information to that which is most relevant must be developed as we continue to assess our value to our clients.

The Library regards internally generated information with the same level of importance as externally published literature. Collectively, both represent the Laboratory's knowledge base, one that must be managed and exploited. The Library has implemented systems that provide clients with effective, easy-to-use tools to locate and retrieve specific information, facilitate the management of this information, and once found, help incorporate the information into new documents. Library staff offer the expertise needed to enable customers to integrate automated document management into their document creation process and to link the information created internally with related published information.

We have been particularly fortunate to have the information technologies (hardware, software, and network infrastructure) that permit us to advance the use of digital information and bring information resources and services directly to our customers at their workstations. These tools make it possible to easily

customize products to meet our users' unique needs. We anticipate increasing reliance on information technologies to enhance collaborative efforts within APL as well as between the Laboratory, other divisions of the University, and other institutions and organizations. We continue to explore technologies that will actively support collaborative efforts within APL that are conducted substantially within an electronic environment.

As the reader may have noticed, we refer to those who use our services and products as clients or customers, not patrons. That change in terminology is more than semantic. It signifies a realization that information seekers no longer need to use the Library to obtain information. Our customers now have multiple, easy-to-use channels for information searching and delivery. The Library becomes the channel of choice only when it is perceived to be providing value not obtainable from other sources. In a real sense, the Library must compete for our customers' business. Competition drives us to tailor products and services to individual needs. As APL conducts its business in a changing environment, the Library will continue to rebuild its capabilities to maintain its vital role in serving the business purpose of an R&D Laboratory.

## REFERENCES

- <sup>1</sup>Digital Libraries Initiatives, available at <http://dli.grainger.uiuc.edu/national.htm> (accessed 16 Mar 2000).
- <sup>2</sup>Griffin, S. M., "Digital Libraries Initiative—Phase 2: Fiscal Year 1999 Awards," *Bull. Am. Soc. Info. Sci.* **26**(1), 14–21 (Oct–Nov 1999).
- <sup>3</sup>Pack, T., and Pemberton, J., "Library at the Los Alamos National Laboratory," *ONLINE* **23**(2), 35–42 (Mar–Apr 1999).
- <sup>4</sup>Kennedy, M. L., "Building Blocks for Knowledge Management at Digital Equipment Corporation: The Library," *Info. Outlook* **1**(6), 39–42 (Jun 1997).
- <sup>5</sup>D'Alessandro, M. P., D'Alessandro, D. M., Hendrix, M. J. C., and Bakalar, R. S., "The Virtual Naval Hospital: A Digital Library for Naval Primary Care Providers and Patients," *Mil. Med.* **163**(11), 775–780 (Nov 1998).
- <sup>6</sup>Citrix Corporation Web site, available at <http://www.citrix.com/products/winframe.asp> (accessed 16 Mar 2000).
- <sup>7</sup>Skinder, R. F., and Gresehover, R. S., "An Internet Tool for the Technical and Scientific Researcher," *ONLINE* **19**(4), 38–42 (Jul 1995).
- <sup>8</sup>Berghel, H., "Cyberspace 2000: Dealing with Information Overload," *Comm. ACM* **40**(2), 19–24 (Feb 1997).

## THE AUTHORS



ROBERT S. GRESEHOVER received a B.S. degree in biological science from Michigan State University in 1971, an M.S. degree in library science from the University of Maryland in 1975, and an M.S. degree in administration from The Johns Hopkins University in 1982. Before joining APL in 1986, he was Deputy Director of the Johns Hopkins Welch Medical Library. Mr. Gresehover is presently Group Supervisor of the R. E. Gibson Library and Information Center. He is a member of the American Society for Information Science. His e-mail address is [robert.gresehover@jhupl.edu](mailto:robert.gresehover@jhupl.edu).



DEANNA T. JONES received a B.A. in economics from Wilson College, an M.S. in library science from the University of Maryland, and an M.S. in information and telecommunications systems from The Johns Hopkins University. She joined APL in 1980 and is a member of the Senior Professional Staff. She is currently Supervisor of the Database Services Section of the R. E. Gibson Library and Information Center. Ms. Jones has been involved in the selection and installation of two automated integrated library systems and one electronic document management system. Her professional interests include information systems management and large database design and retrieval issues. Her e-mail address is [deanna.jones@jhuapl.edu](mailto:deanna.jones@jhuapl.edu).



PAUL S. TARANTINO received a B.S. in mathematics and history from MIT in 1967 and an M.B.A. from the University of Santa Clara in 1981. He served 23 years in the Navy, retiring as a Commander in 1991. While on active duty, he was a Naval Flight Officer and Mission Commander in P-3 aircraft, served aboard two aircraft carriers, was a financial programmer on the OPNAV staff, and served as a Deputy Program Manager in the Tomahawk Program Office. Mr. Tarantino joined APL in 1991 and was appointed to the Senior Professional Staff in 1993. He is the Assistant Group Supervisor of the R. E. Gibson Library and Information Center and Manager of the APL Government Information Services Project. His e-mail address is [paul.tarantino@jhuapl.edu](mailto:paul.tarantino@jhuapl.edu).