TORPEDO:

Electronic Dissemination of Physics Journals and Technical Reports on Campus Networks

Laurie E. Stackpole, and

Roderick D. Atkinson

U.S. Naval Research Laboratory

Robert A. Kelly

The American Physical Society

The Naval Research Laboratory (NRL) Library and the American Physical Society (APS) are engaged in a two-year cooperative arrangement to distribute two APS journals, *Physical Review Letters* and *Physical Review E* in electronic format to NRL researchers. Additionally, the Library is experimenting with network dissemination of its own unclassified unlimited NRL technical reports along with the APS journals. The project itself is called TORPEDO (The Optical Retrieval Project: Electronic Documents Online) and consists of scanning the paper copies of the journals and the reports and making them available in page-image format to selected campus networks.

TORPEDO Access

Under the terms of the NRL-APS agreement NRL employees, NRL on-site contractors and ONR Headquarters staff may search and view the APS journals in TORPEDO. Access to the TORPEDO system is only available from the campus networks of NRL/ONR and the APS.

TORPEDO access is provided through the Library's Web Home Page (URL - http://infonext.nrl.navy.mil/). The Library's Home Page serves as a jumping point from which end-users can learn about the TORPEDO project, access the Home Pages of the associated participants (APS and Los Alamos National Laboratory), download TORPEDO's freely distributable client software and user guides (beginning in May 1995), and launch TORPEDO. The computer workstation requirements for accessing TORPEDO are identical to those for running NCSA Mosaic.

To serve electronic journals and technical reports to end-users TORPEDO uses commercial imaging software from Excalibur Technologies, called EFS (Electronic Filing Systems). EFS is predominantly client/server based and comes with freely distributable client software for MS Windows and Macintosh

workstations. UNIX access to EFS is currently supported through an X-Window interface, and a true UNIX software client is scheduled for release in the next major EFS upgrade.

Scanning

The APS journals and NRL technical reports are converted at NRL from paper format to scanned CCITT Group IV images, OCR'ed, associated with bibliographic information, indexed by the EFS database, and stored on 5 1/4" multifunction (read-write) optical disks. This whole process can be performed in one 24hr period. Thus, end-users can have the current journals electronically available at their desktops one day after they are received by the Library.

Searching

Journals and reports are made end-user searchable using OCR'ed full-text data combined with fuzzy logic, bibliographic data elements, or directly through a tree-structure hierarchy. In terms of full-text searching, EFS retrieves documents by recognizing data patterns at a binary level. As a result, data itself automatically directs the creation of indexes that are highly fault tolerant and thereby offer the ability to accurately retrieve information based on an approximation of query terms or phrases. Because the EFS retrieval software seeks patterns rather than exact words or phrases, users can accurately search "dirty" ASCII (raw OCR-processed text) eliminating the need for ASCII clean-up or re-keying.

Major Participants in TORPEDO

The Naval Research Laboratory Library

NRL is the Navy's corporate research and development laboratory, created in 1923 by Congress for the Department of the Navy on the advice of Thomas Edison. The Ruth H. Hooker Research Library and Technical Information Center addresses the information needs of the NRL research community, consisting of about 3,500 Federal staff and about 1,500 contractors at the Washington, DC facility. NRL occupies a 130-acre campus of 152 buildings located on the Potomac river in Southwest Washington, DC. Research facilities are also located in Orlando, FL; Bay St. Louis, MS; and Monterey, CA. The research efforts of the Laboratory are concentrated in 17 broad areas covering physics, electronics, chemistry, and space sciences. In addition, the Library also services NRL's parent organization, the Office of Naval Research (ONR), in nearby Arlington, VA.

The NRL Library has been in the forefront in moving to a totally digital library. The Library began actively scanning its technical reports collection in 1988 and has been scanning close to 10,000 pages a day since early 1993. In addition, the Library has supported a campus-wide information system, called the InfoNet,

since 1992. The InfoNet provides NRL/ONR researchers desktop access to commercial and non-commercial online services on the Internet, the Library online catalog, NRL resources, CD-ROM databases, and electronic books. The Library has been cited for its work as an emerging virtual library in publications such as *Internet World, Government Computer News*, and *Science*.

The American Physical Society

The American Physical Society (APS) is an organization of more than 43,000 physicists worldwide. Since its formation in 1899, it has been dedicated to the advancement and diffusion of the knowledge of physics. The APS publishes some of the world's leading physics research journals: the *Physical Review* series, *Physical Review Letters*, and *Reviews of Modern Physics* . It organizes scientific meetings where new results are reported and discussed. In addition to these primary functions, the Society has many other programs in areas such as education, international affairs, public affairs, and public information.

Since its founding the primary purpose of the APS has been to disseminate the knowledge of physics. Recently the APS is became quite active in projects to disseminate its journals electronically. In addition to working with NRL, the APS is active in several electronic journal dissemination projects, including an archive of the *Physical Review* with Los Alamos National Laboratory and in disseminating its flagship publication, *Physical Review Letters*, through OCLC. As part of its commitment to the electronic publication the APS is moving all of its journal publications to SGML.

Note: A detailed handout coving the TORPEDO project will be made available at the project briefing session.