# Stretching the Web: Early Experiences with Publishing *Applied Physics Letters Online*

**Timothy Ingoldsby**, Director of New Product Development, American Institute of Physics

**W. Daviess Menefee**, Consulting Product Specialist, OCLC Online Computer Library Center

What are the characteristics of the ideal online journal? Certainly, it must be capable of reproducing all forms of information contained in current print journals. Beyond that, it should support the creation of hyperlinks among the various components of the document, and between the document and other related documents or databases. Eventually, it should support new forms of information transfer, including access to the underlying data used to construct tables and graphs, an ability to manipulate equations contained in documents, and support for information types not possible in a paper-only journal.

Since 1988, the American Institute of Physics (AIP) has been engaged in a process designed to transform its entire 250,000+ annual-page publishing program to provide fully digital "pages," suitable for online as well as print publication. By 1993, AIP had progressed far enough in this project to begin to consider options for the creation of its first online journal. It quickly became clear that the Electronic Journals Online Service developed by OCLC - Online Computer Library Center, featuring the Guidon(tm) user interface, provided the best match of capabilities compared to AIP's criteria.

Accordingly, AIP began negotiations with OCLC, which culminated in the signing of an agreement early in 1994 for the development of *Applied Physics Letters Online*, an online version of AIP's prestigious weekly letters journal. Working closely with OCLC developers, AIP refined its SGML Document Type Definition (DTD), based on the ISO12083 standard. In support of AIP's desire to use this internationally accepted form for SGML information presentation, OCLC developed a translation program to properly process all elements of the AIP DTD including special characters and equations in coded to the ISO standards.

APL Online became available with the first issue of 1995. Already, it has attracted a substantial number of subscribers, which should grow even larger when *Physical Review Letters Online*, a publication of the American Physical Society (APS) - one of AIP's Member Societies - joins *APL Online* on the EJO system in the beginning of July of this year.

At the same time that AIP was negotiating with OCLC, the development of NCSA Mosaic, and its effect on the hypermedia-based World Wide Web, created a great deal of interest by both parties in producing a Web version of EJO. In addition to the obvious similarities between the capabilities of documents based on the Web protocols with the features provided by EJO/Guidon, a Web version of EJO would extend the accessibility of *APL Online* beyond the Microsoft Windows Guidon platform to support Macintosh or X - Windows user communities. When surveys of potential subscribers indicated a substantial population wishing to use Macintosh computers, OCLC and AIP agreed that development of a Web version of EJO would be desirable. The first phase of EJO/Web was completed in time for the product launch in January.

While not yet as full-featured as EJO/Guidon, the Web version represents a dramatic engineering achievement by OCLC, which has transformed the "stateless" Web protocols into a system capable of retaining "session" information. To provide these capabilities, OCLC has stretched the Web protocol to the maximum, and beyond. For example, EJO/Web keeps track of what previous requests have been issued by a user, and can use that knowledge to determine what document should next be presented in response to user requests such as "next document," "next page of results," or "previous page of results."

To unite the stateless and the stateful, a separate gateway session is initiated and maintained at OCLC whenever a user logs on, holding the information required by the OCLC online systems, such as authorization number and profile information. This also allows control of the number of multiple simultaneous users per authorization. A unique, randomly-generated session id is created and tagged to each transaction conducted by the user. As the user interacts with the system, additional information about the session, including current and previous search results, is added to the information stored in the gateway system to allow user navigation within the activity of the session. Once the user logs off or is timed out, the gateway session is terminated.

OCLC's efforts to display the rich content of scholarly scientific research documents has highlighted a variety of shortcomings in the current HyperText Markup Language (HTML) specification, Version 2.0. Beyond the obvious problem of lack of support for special characters (e.g., Greek letters, mathematical symbols, etc.), built-up mathematical formulas (display equations), and overstruck characters, the interaction of HTML files with the capabilities provided by various browsers leave much to be desired. In fact, for physics information, the data displayed on a browser may contain inaccuracies. To alert subscribers to this potential, OCLC was compelled to develop a symbol, a "caution" sign, that is displayed when characters may be inaccurately positioned. Even now, most browsers cannot accommodate super- and subscripting, certainly minimal functionality for the accurate rendition of a printed article. Additionally, users of these browsers have developed high expectations for their local printing capabilities. These expectations are rarely met.

Enhancements of the Web protocols and browser capabilities to meet the needs of scholars and publishers will take time, and will require the combined effort of the community. Projects such as the cooperative efforts of OCLC and its journal publishers, the NSF/ARPA/NASA Digital Libraries projects, and other network publishing initiatives provide the test environment to define and refine the requirements for next-generation Web standards and browser enhancements. It is vital for scholarly societies to assist or even coerce (politely, of course) the developers of the browsers in order to achieve the necessary functionality for scholarly communication. The form that this assistance takes may include minimum specifications for community-wide adoption or even a "recommendation of approval" by a society. The lack of standards in this rapidly changing area will continue to foster the development of popular features at the scholar's expense.

For further information about APL Online, send your comments to aplonline@aip.org.