41 Toolbars 45 Mutual Funds 64 Copernic Summarizer

ONLINE®

The Leading Magazine for Information Professionals JANUARY/FEBRUARY 2004

Harvesting USDA's Agricultural Information

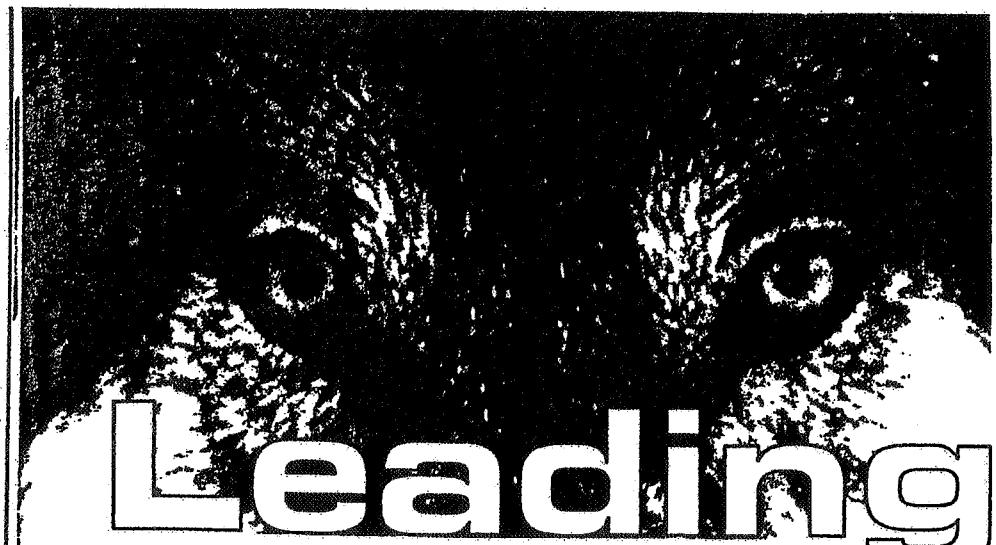
Volume 28 · Number 1

Display until February 29

0 74470 85534 7 01>

ONLINE

Information Today, Inc.



Leading the Pack:

Librarians Create an Original Model for a Customer-Friendly Publications System

by Nancy Allmang and Jo Ann Remshard

Foresight and imagination are nothing new to libraries and information professionals. For the past decade, we have been coping with, and capitalizing on, evolving technologies. Developing new electronic systems presents challenges to define and carry forward the essence of a process while leaving cumbersome traditional nonessentials behind. The most useful electronic products turn out to be much more than automated versions of old processes.

Our organization took a complex collection of electronic and manual publishing processes that no one wanted to use and developed a cohesive, unified system model that includes a digital library and a library online catalog.

Once fully implemented, the new system will enable authors to submit manuscripts for review and to track the status of their submissions. It will allow supervisors to generate reports. It will empower researchers inside and outside of the agency to locate and retrieve full-text electronic versions of agency publications from several access points. By means of crosswalk scripting, it will convert bibliographic information about published documents in the database to Machine-Readable Cataloging (MARC) records in the library catalog.

NATIONAL INSTITUTE OF STANDARDS AND TECHNOLOGY

The Information Services Division (ISD) of Technology Services at the National Institute of Standards and Technology (NIST) [1] supports the NIST research activities through a comprehensive program of knowledge management. Under the ISD umbrella are a publishing department (the Electronic Information and Publications Group), a scientific and technical research library (the Research Library and Information Group), and a Museum and History Program. The Electronic Information and Publications Group (EIPG) provide both Washington and Boulder NIST campuses with a wide range of electronic information and publication services. Some are secretariat services for the NIST manuscript review and approval processes, as well as editorial, print, electronic layout, and design services. The Research Library maintains a collection of about 300,000 volumes and 1,150 journal subscriptions. The collection also includes approximately 6,000 agency publications.

NIST researchers use a variety of vehicles for disseminating information about their research. These include agency publications—such as handbooks, special publications, technical notes, and interagency reports—along with the NIST *Journal of Research*. NIST researchers also publish articles in scholarly commercial and professional society journals. Approximately 2,000 manuscripts targeted to agency or external publications are submitted for review and approval each year.

ORIGINAL SYSTEM

The current publications workflow system consists of a number of components:

- An Access database, "NIST Pubs," containing information on documents that have been authored or co-authored by NIST employees
- A Web-based input system called the Editorial Review Intranet (ERI) designed to facilitate entry and maintenance of data into the current system through a Web interface
- An R:Base publications database used by researchers on the NIST Boulder campus
- Editorial Review Board (ERB) processes in Washington (WERB) and in Boulder (BERB)
- A Publications Calculator
- A Web-based service called Technipubs which points to bibliographic information about approved manuscripts

A huge snag in the workflow system has been the Web-based ERI input system. Recently, customers in many of our operating units or laboratories reported that it locks frequently, making their submission of manuscripts a tedious and painful process. An Access database of publications slowed down because it is strained to capacity.

As a result of these difficulties, several laboratories came up with solutions for keeping track of their own publications. Ironically, many of these homegrown solutions made use of publications data imported from the database our division maintains. In addition to

this variety of separate databases, researchers in many labs have been posting the full text of their publications on division or lab Web sites. With no "pointers," interested searchers can't find them. While the labs took these actions to meet an immediate need, the result is a distributed system with no central control and no easy way for customers to find agency and NIST-authored publications.

The Research Library uses Sirsi to catalog a collection that includes NIST publications. Both NIST researchers and external customers use the online catalog with the Technipubs interface at our Web site [http://nvl.nist.gov] to hunt down citations to NIST-authored publications. They then find ways to get paper copies: NIST researchers call or come to the library. Users from off-campus frequently order print copies at a nominal fee from the Government Printing Office or National Technical Information Service. Some outside customers place interlibrary loan requests for documents through their local libraries.

Given the inefficiency of the manuscript input system, the overtaxing of the database, the duplication of effort by "splinter" laboratory databases, and ISD's desire to provide electronic full text and an efficient means of searching for all customers, the time was clearly ripe for change.

A CALL FOR CHANGE

A seven-member team, assembled by the chief of ISD, examined the current system, gathered customer requirements, and explored a variety of options to implement a new publications system for NIST. Members came to the team with a variety of important skills. The group included a knowledge management librarian, the Washington Editorial Review Board secretary and editor, ISD's Sirsi system administrator, a Web expert/writer/editor, a librarian/cataloger, a systems/ILL/document delivery librarian, and a reference librarian as team leader. The ISD chief charged us to transform the current array of processes into a unified Publications Knowledge Management System. She envisioned a whole greater than the sum of its parts.

The new system was to accomplish the following:

- Take a completed, signed manuscript from submission to publication with tracking and reporting mechanisms.
- Eliminate ongoing problems with the current submissions system (ERI).
- Eliminate frequent crashes and slowness of the current database.
- Eliminate duplication of effort (a number of individual labs and divisions had begun running "splinter" databases of their own).
- Offer searchers an efficient way to locate NIST publications.
- Provide the full text of publications.

GATHERING REQUIREMENTS

We knew that our customers were key to laying the groundwork for the new system. To learn how customers would use the system and what information they would want in the new database, team members met with them over a period of weeks. Our customers were authors, managers, administrative support staff, internal and external reviewers, NIST researchers, outside researchers, businesspeople, and members of the public at large. All had different requirements for the system.

While some members of the team gathered customer requirements, others looked at open source software options, commercially available products, and systems developed by other institutions. We found no one system that met all of our requirements: an all-electronic manuscript submission system with editing, tracking, and reporting for authors, reviewers, and managers; a database with bibliographic and status information about agency publications; a means to house the full text of published NIST documents and journal articles; and a way to automatically catalog and connect the full text after publication with the library's online catalog.

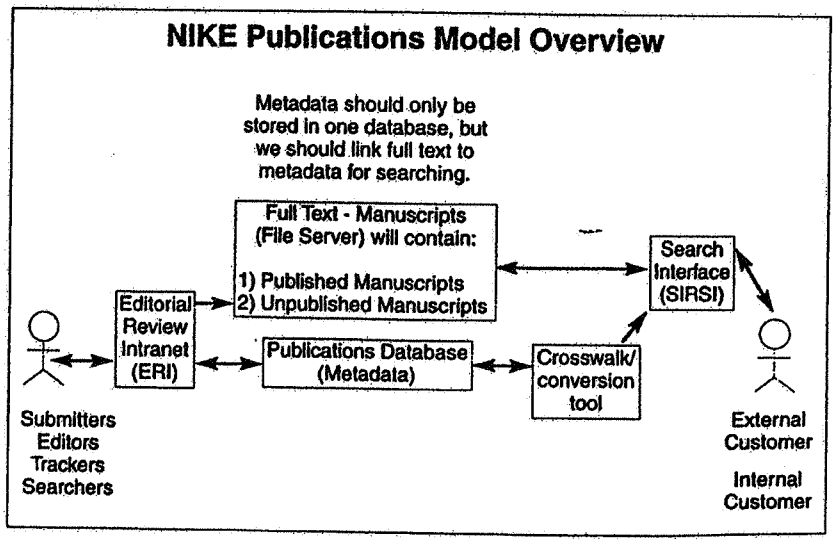
NIKE IS BORN

ISD set about listing requirements and outlining a plan to redesign the entire NIST Publications System. We called it the NIST Integrated Knowledge EditorialNet (NIKE—Greek goddess of victory).

During the gathering of customer requirements, ISD became aware of publications databases developed by several of the labs. One in particular stood out, that of NIST's Electronics and Electrical Engineering Lab (EEEL). EEEL's system, created by scientists for scientists, worked so well that other labs asked to adapt it. The EEEL system uses an Oracle database, a Web interface, and Perl scripting. EEEL was willing to let ISD use its database as a model for a new NIST-wide database that would satisfy requirements for *all* the NIST labs. And the creator of the EEEL database was willing to work with us.

WHAT NEEDED TO HAPPEN

We wanted to spell out our technical requirements clearly for the budgeting of resources and as a basis for an information technology plan and timeline. We divided the project into two phases, the first of which called for repairs to the current system with the addition of a simple reporting mechanism. We wanted customers to be able to generate reports for year-end summaries.



points. (Currently these exist as "brief" records.) We also need to add 950 bibliographic records to the online catalog.

Many records of approved manuscripts in the publications database show an "unpublished" status when in fact they already have been published. ISD intends to manually check the accuracy of the information and bring the records up-to-date.

Conversion of legacy data: A critically important part of the NIKE project is the migration of the data from the Access database and the EEEL Oracle database into the new NIKE Oracle database. In this complicated migration, the merging of duplicate records from both sources into one complete record in which all unique data was retained is of utmost importance. It is essential to avoid duplicate records.

NIKE IT PLAN

The IT support for this project comes from another group within our Technology Services Operating Unit—the Information Services Group (ISG), which had previously been part of the Information Services Division. It was also involved in the early planning for NIKE. The requirements document ISD prepared provided the group with detailed descriptions of the modules explained above. ISG then identified and analyzed development methods and established a comprehensive plan for putting NIKE into action. ISG split the project into three major deliverables: NIKE-A (October 2003), NIKE-B (December 2003), and NIKE-C (April 2004). NIKE-B includes all of the work on the new system except for online revisions and peer review. ISG will finish these up separately as NIKE-C.

USABILITY AND MARKETING

ISD will test customer and staff interfaces to be sure users can easily submit and search for information as well as generate reports from the system. We plan to conduct usability testing with three user groups: 1) ISD staff members; 2) NIST researchers/authors, support staff, Washington Editorial Review Board members, and others who use the current system; and 3) NIST customers who have never used the NIST publication system.

In order to get NIST-wide buy-in for the new system, ISD launched a marketing campaign well in advance of the system's debut. Marketing activities included articles in Information Services Division and Technical Services newsletters, updates on the Research Library's Virtual Library Web site, presentations to division chiefs and others, articles through publishing avenues, creation and distribution of brochures, posters, a bookmark, a guest book, one-to-one marketing, and new employee orientation and e-mails.

Please indicate in which formats you would like to be able to generate reports.

Format	Very Important	Important	Not Important
HTML	x	x	x
PDF	x	x	
ASCII (plain text)	x	x	x
MS Word	x	x	x
Other (Please Specify)	DB report: SQL, Oracle Pert, Access		

Is it important to have electronic copies of your publications available online, when copyright permits?

Yes	No
x	

Would you like to have the option of not posting your approved manuscript or published document online?

Yes	No
	x

The following table needs your input on the fields you would like to see in the pulications database. Please indicate which fields are important to you and how you expect to use them. In addition, indicate whether or not these fields should be viewable by the general public.

Field	Important for Searching (Y/N)	Important for Reporting (Y/N)	Important just to maintain (Y/N)	Not Important/ Not used (Y/N)	Multiple Entries Needed? (Y/N)	Viewable to the general public? (Y/N)	Your Comments
Author	Y	Y	Y		Y	Y	Access to general public after publication.
Author Affiliation	Y	Y	Y		Y	Y	

ACCELERATING ACCESS TO SCIENTIFIC PAPERS

In recent years, a sea of change has resulted in the creation of numerous digital libraries and transformed the world of scholarly communications. Scientists now read one another's work online shortly after publication, speeding scientific progress.

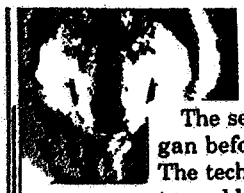
We are excited that NIKE will simplify a complex publications process for our customers and will provide the full text of our publications for all. We feel a distinctive element of the new system will be the automatic translation of bibliographic information from the publications database to the online catalog.

Thanks to much cooperative work with our information technology counterparts, the NIKE project is rapidly moving ahead. The Information Services Division expects the first part of the project to be in place by the time you are reading this. Check our Web site next year for the official unveiling of NIKE.

Disclaimer: The identification of any commercial product or trade name does not imply endorsement or recommendation by the National Institute of Standards and Technology.

Nancy Allmang [nancy.allmang@nist.gov] is reference librarian and laboratory liaison and head of the redesign team, National Institute of Standards and Technology Research Library Information Group. Jo Ann Remshard [joann.remshard@nist.gov] is knowledge management librarian and a member of the redesign team, National Institute of Standards and Technology Electronic Information and Publications Group. Sissy Riley, library technician, National Institute of Standards and Technology Research Library and Information Group, prepared the graphics.

[1] NIST is a nonregulatory federal agency within the U.S. Department of Commerce. Its mission is "to develop and promote measurement, standards, and technology to enhance productivity, facilitate trade, and improve the quality of life." Its work is focused on advancing the nation's technology infrastructure and supporting industry. NIST operates in two locations: Gaithersburg, Maryland, and Boulder, Colorado. It employs about 3,000 scientists, engineers, technicians, and support and administrative personnel. About 1,600 guest researchers complement the staff.



The second phase of the project began before phase one was complete. The technical requirements for phase two addressed six modules: the User Interface, Database, File Servers, Crosswalk scripting to convert data to MARC format, Database Content Cleanup, and Conversion of Legacy Data. To write the technical requirements, we divided the team into groups of two, each assigned to one or two modules. For each module, we wrote objectives, requirements, performance standards, and recommended solutions. We identified applicable documents/paper forms currently in use as well as deliverables, timeframe, and resources.

PHASE ONE: TECHNICAL REQUIREMENTS

"Fixes" and reporting: The ERI interface was originally developed as a means for customers to enter information that would start the manuscript submission and review process. The interface had technical problems that needed to be addressed. We would accomplish this by editing and fixing the site backend code to allow the site to run more efficiently for all users straightaway. Examples of fixes were widening the window for time-outs, allowing auto-population of previously entered information.

We wanted to give our users the ability to initiate and run reports from the database through the interface as soon as possible. By Fall 2003, we intended to have a simple reporting method ready for our customers.

PHASE TWO: TECHNICAL REQUIREMENTS

User interface: The team determined that a complete redesign of the interface was essential. The new interface will support users through the submission-of-manuscript process. The interface will also be used for editorial review, searching, locating records, and creating reports.

Database: We plan to create a new Oracle database using the EEEL database as a model, generally restructuring and building tables to provide for the needs of all the NIST laboratories.

ISD set high performance standards for multiple customers who might be concurrently querying the database by searching, running reports, adding, updating, and deleting bibliographic information, and performing administrative tasks. To meet our customers' needs, the NIKE team recommended these service-level performance standards to our information technology staff (see chart blow):

To confirm adherence to these performance standards, the team plans to conduct usability testing of the completed Web-based customer interface.

File servers: The requirements included housing for full-text manuscripts as well as final published documents and supplemental files. We also included the size of the initial load of files, the approximate number loaded each year, scalability, integration with existing infrastructure, security, backup, recovery, and technical requirements. Technical requirements included size of the processor, RAM, hard drive, and storage capacity, as well as details on RAID controller, operating system, CD-ROM, and backup.

Performance standards covered the following areas:

- Dedicated file server
- Ability to house multiple file formats
- Scalable to support multiple concurrent users
- Scalable to support 150 GB from the archived collection and 27 GB of new files per year.
- Integrated with existing infrastructure applications
- Have in place security, backup, and recovery

Crosswalk from Oracle to Online Catalog: A software program will automate the process for migrating the captured data in the new Oracle database to the integrated library system (Sirsi-Unicorn). The process will extract relevant data from the NIKE database and convert it into a format that would be imported into the Sirsi system. This automation will greatly reduce the manual editing that might be still required. The crosswalk is intended to work seamlessly between the NIKE and Sirsi systems.

Content cleanup: ISD is fine-tuning existing bibliographic records for system accuracy, including cataloging records of NIST publications in Sirsi and the data in the Access publications database.

Cleaning up the cataloging records entails upgrading 4,460 existing records in the Sirsi online catalog to include author, subject, and series access

Appedix B-Customer Requirements Survey

The Information Services Division is planning major improvements to the NIST Publications Database. Included will be a revision of the online Editorial Review Intranet (ERI), with user editing and more tracking. Please fill in the following survey so we can learn your requirements. Thank you!

Requirements for Publication Management System—Survey

Based on your experiences and usage patterns, whom do you view as the users of the Publications Database? How do they use the database?

	Type of Use			
User	Enter/ Edit Data	Search	Report	Other (Please Specify)
Authors—Internal (NIST)	x	x	x	
Authors—External (non-NIST)	x	x	x	
Administrative Personnel	x	x	x	
Management Personnel	x	x	x	
Reviewers for ERB Submission—Internal (NIST)	x	x	x	
Reviewers for ERB Submission—External (non-NIST)		x	x	
Users—Internal (NIST employees)		x	x	
Users—External (General Public)		x	x	
Other (Please Specify) Journal of Conf Proceeding	x	x	x	

Searches:	4 seconds or less
Reports under 50 pages:	PDF Format in 15 seconds or less
Reports under 50 pages:	ASCII Format in 10 seconds or less
Add/Update/Delete:	3 seconds or less
Retrieve full-text files (if applicable):	5 seconds or less per MB