Taming a Wild Workflow and Domesticating our Digital Data

Session Number C202

Nancy Allmang, Reference Librarian/Leader, Redesign Team, *nancy.allmang@nist.gov*
Paula Deutsch, Information Specialist/Member, Redesign Team, *paula.deutsch@nist.gov*

The tempo of change in the publishing field is almost dizzying, and those of us responsible for organizing knowledge created by our institutions and making that knowledge accessible are beginning to see that we have a lot of work to do.

A seven-member team in the Information Services Division (ISD) of the National Institute of Standards and Technology (NIST) Technology Services (TS) Operating Unit (OU) was tasked with studying the organization's complicated publications system from manuscript submission by an author to retrieval of the published document by an end-user. The team made recommendations for the development of an efficient, streamlined system model.

The team discovered that the publications system was a mass of print and digital processes that NIST authors hated but needed to use to get their work published. NIST policy requires editorial review and approval of all "official writing." This requirement includes review and approval both within the originating operating unit and by a NIST Editorial Review Board. Policy and review procedure are outlined in the NIST Administrative Manual. Within the organization's 45 divisions, different manuscript approval workflows evolved over time. Electronic parts of the system were unfriendly and difficult to use. There was much duplication of effort within and across operating units and no full-text retrieval from a central source.

The team conceived a system model, the NIST Integrated Knowledge EditorialNet (NIKE), to streamline the manuscript submissions workflow and connect it with a library catalog and a digital library. A scripted crosswalk will automatically change information entered by authors into MARC records for the library catalog.

A formal document, the Requirements Analysis and Development Plan delineated the project into seven "modules" and spelled out requirements for each. The plan was approved by management and ISD staff are now working closely with TS IT staff on the details of NIKE.

The NIKE project consists of four separate elements: (1) Web interface for internal customers (NIST staff); (2) database that holds information about publications (including bibliographic data and process information for tracking); (3) server for storing full text documents; (4) public access to metadata and publications through the library's online catalog.

The Web interface will be used to submit manuscripts for approval, edit manuscripts, track manuscripts through the approval process, and generate status reports and bibliographic lists. It will also be used within the Information Services Division to perform maintenance functions on the database. Each type of user will have a tailored portal to the NIKE system. The team is keenly aware that this is where it will need to sell the application. If the internal customers find the system difficult to use, they will be less likely to supply timely and accurate data and the whole application could fail.

The information collected through the customer interface will be stored in the NIKE database. The database will hold bibliographic metadata and NIST organizational information, and will track NIST-authored work from manuscript to publication. Its current iteration includes over 50 tables. These can be summarized as user tables, NIST organization tables, tables for tracking where a publication is in the production process, authority tables, and publication tables.

Two types of documents will be stored on the document server, manuscripts and published documents. Manuscripts that have not yet been approved and published will only be accessible to those involved in their creation or their editorial review. Once items have been published, a final version will be placed on the document server which will be accessible to the public.

The final piece of the puzzle is public access to NIST publications which will be provided through the online catalog. Bibliographic metadata in the NIKE database will be converted into the MARC format through an automated crosswalk. These records will include a URL to the full text document. Procedures will be put in place to ensure the integrity of these bibliographic records.

This massive project will be rolled out in two phases. A majority of the functionality will be available in the first phase. The second phase will provide an opportunity to review the effectiveness of phase one as well as provide additional capabilities.

Taming a Wild Workflow and Domesticating our Digital Data

Nancy Allmang, *nancy.allmang@nist.gov*
Paula Deutsch, *paula.deutsch@nist.gov*

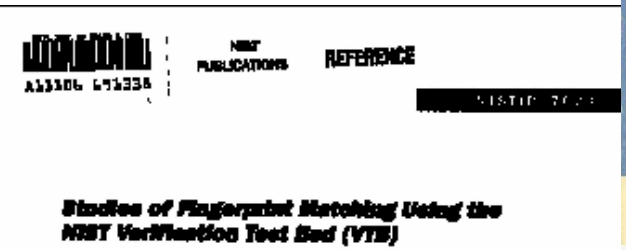
Information Services Division
Technology Services

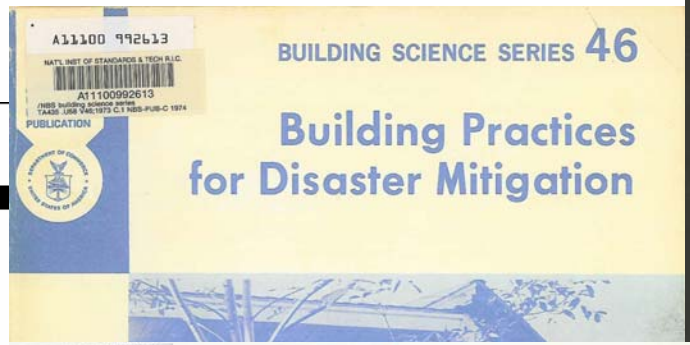
Project Overview

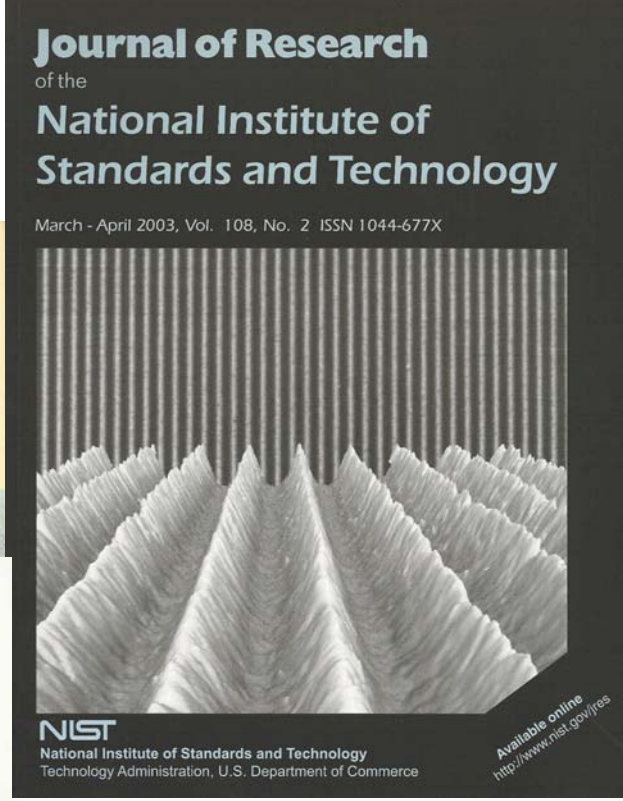
A group of librarians and publishing specialists:

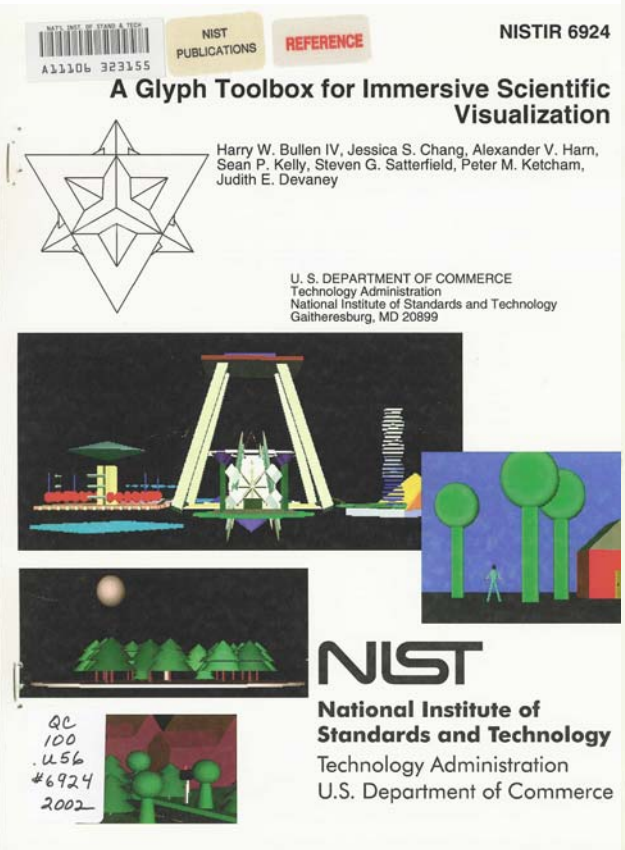
- Began with a mass of paper and digital publishing processes that had evolved over 101 years
- Developed a streamlined workflow system linked to an online library catalog and a digital library

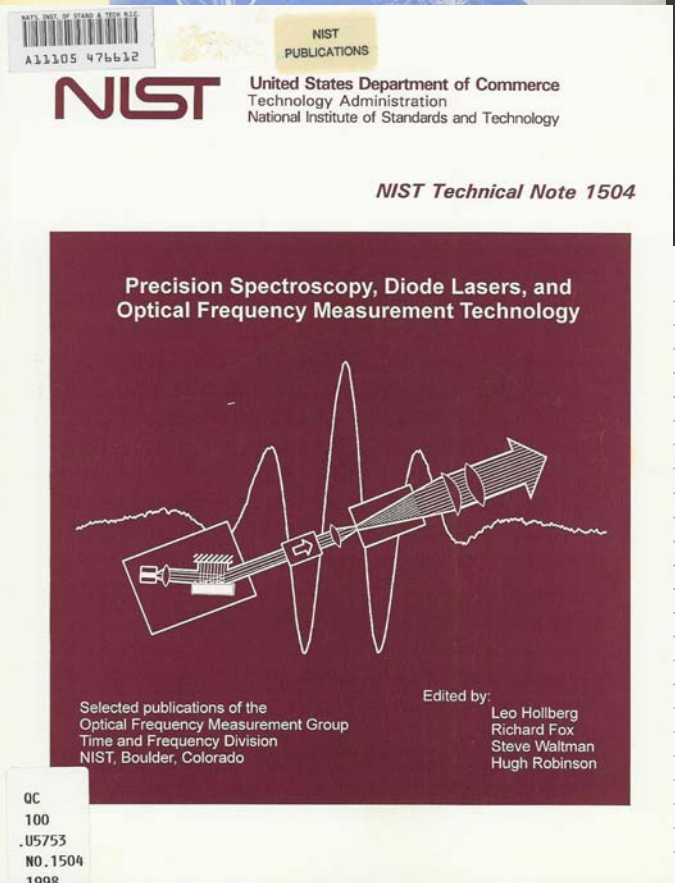
NIST Series Publications











Other NIST Publications

- Journal articles
- Proceedings papers
- Books
- Book chapters
- Video and audio tapes
- CDs
- Databases

Current System

- Not user-friendly
- Frequent crashes
- Difficult to put information in
- Difficult to get information out
- No full text
- Confusing array of “homegrown” databases

Assembled Team

- Knowledge Management Librarian
- Technical Services Librarian
- ILL Librarian
- Web Specialist
- Editorial Review Board Secretary
- Sirsi Administrator
- Reference Librarian

Gathered Data

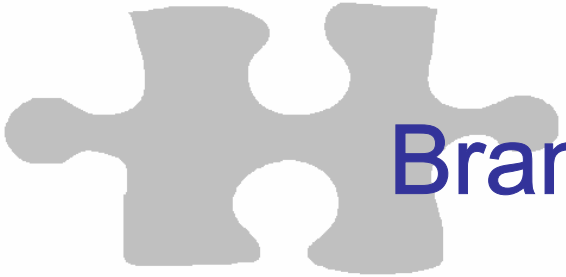
1. Studied other research institutions' solutions
2. Evaluated commercial and open-source digital library products and manuscript submission systems
3. Investigated 7 NIST homegrown databases
4. Compiled user requirements

Envisioned Future Application

- Streamlined, efficient submissions process
- Online tracking, editing, and approving
- Flexible reporting
- Efficient one-stop searching
- Full-text linking
- Customer-customized screens

Made Recommendations

- Adapt best NIST homegrown system
- Add features and make it serviceable for ALL customers



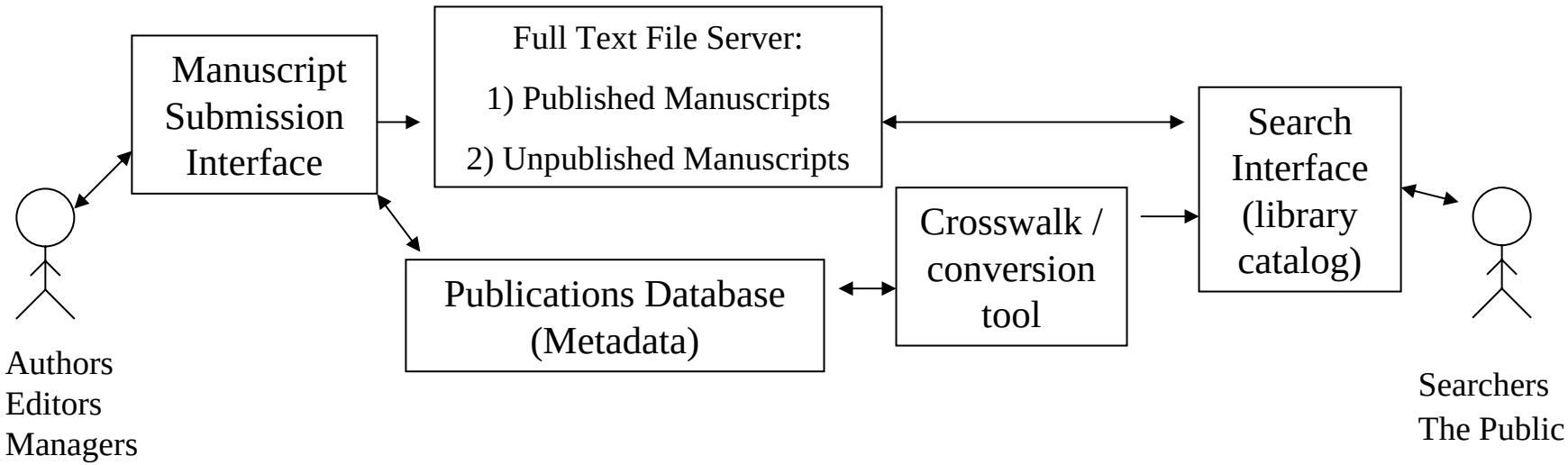
Branded

isd
INFORMATION
SERVICES
DIVISION

“NIKE”

NIST Integrated Knowledge EditorialNet

NIKE Publications Model Overview



Chose Sirsi Catalog as Search Interface

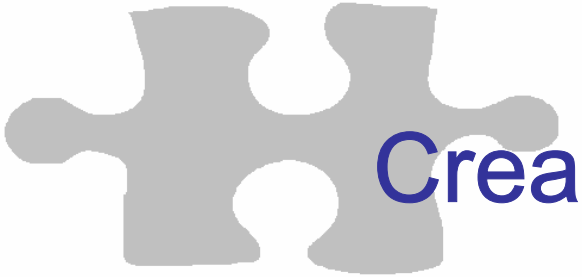
- Supports sophisticated searching capabilities
- Supports sophisticated authority control
- Allows for easy linking to full text
- Provides logical access point

Developed Plan in “Modules”

- User Interface
- Database
- File Servers
- Crosswalk from Oracle → Sirsi
- Sirsi Content Clean-up
- Conversion of Legacy Data
- ERI Submission and Reporting Fixes

Prepared Legacy Records

- “Cleaned up” and enhanced online catalog records
- “Cleaned up” and enhanced existing database records
- Will migrate database records to new Oracle database



Created Digital Full Text

- Located and captured any NIST pubs already on NIST webspace
- Scanned print NIST pubs to searchable PDFs
- Developed directory structure
- Stored on server

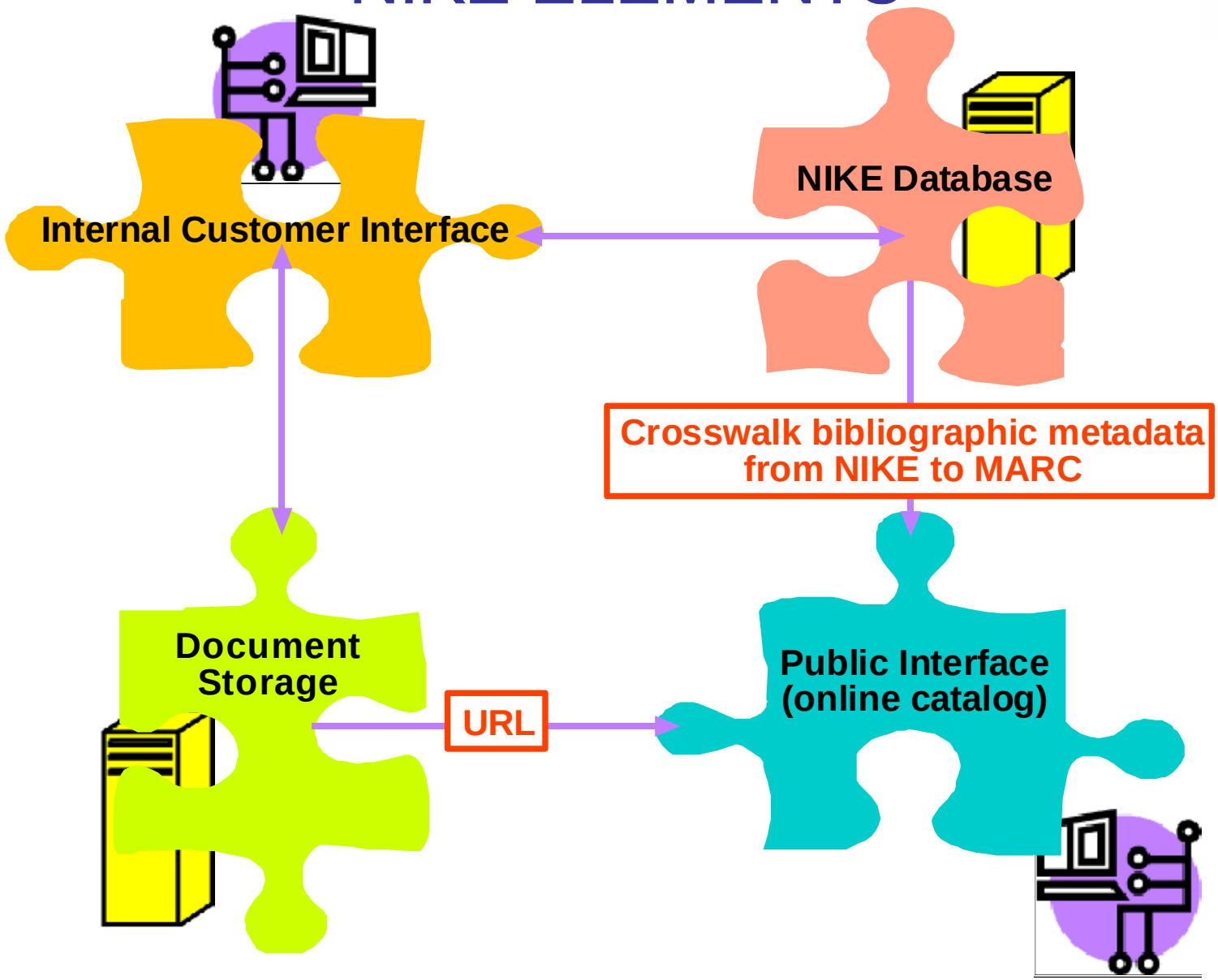
Considered Policy Issues

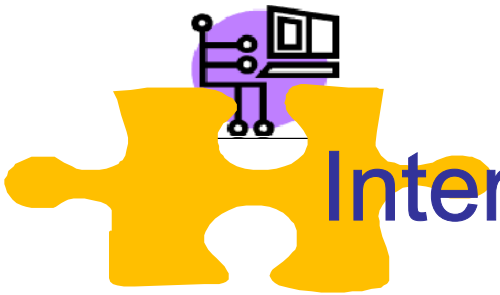
- Copyright
- Full text format
- File size constraints
- Electronic signatures

Planned Post-Launch Activities

- Usability testing
- Marketing
- Evaluation

NIKE ELEMENTS

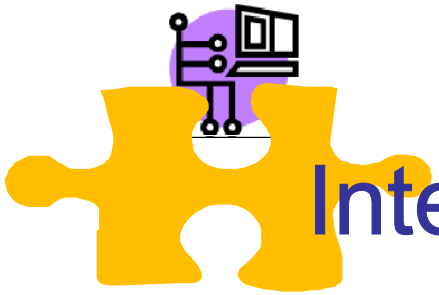




Internal Customer Interface

WEB PORTAL

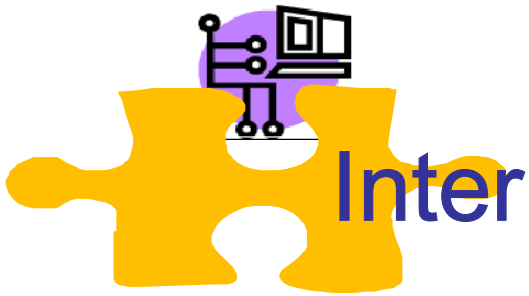
- Web page design and story boarding created by Information Services Division staff
- Functionality behind the pages coded in Perl by TS Information Systems Group



Internal Customer Interface

FIVE TYPES OF CUSTOMERS:

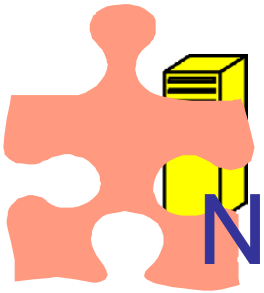
- Authors
- Operating-unit level readers and reviewers
- Managers
- NIST Editorial Review Board members
- Information Services Division staff



Internal Customer Interface

FUNCTIONALITY

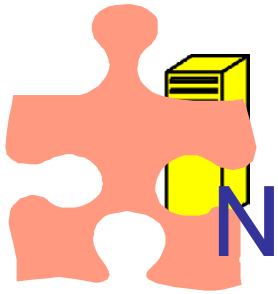
- Submit manuscripts for approval
- Track manuscripts through the approval process
- Post editorial comments or suggestions
- Approve or disapprove manuscripts
- Generate bibliographic and statistical reports
- Maintain the database



NIKE Database

STORES:

- Bibliographic metadata
- NIST organizational information
- Author information
- Authority records
- Process information (for tracking)

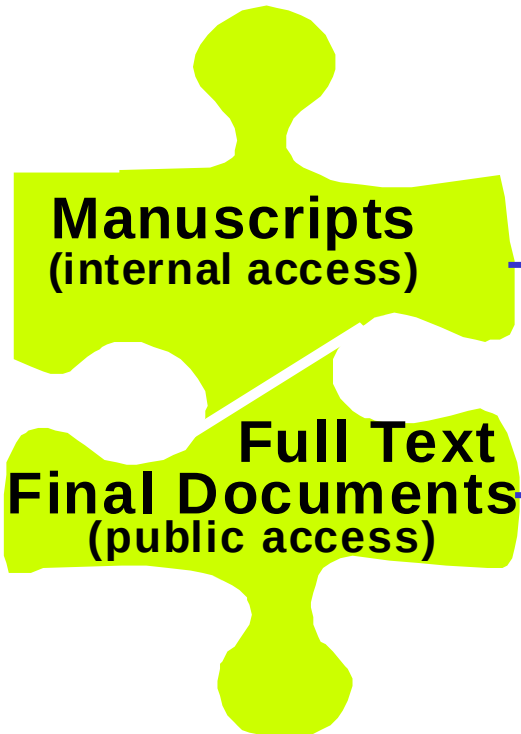


NIKE Database

DATABASE STRUCTURE

- User tables
- NIST organizational structure
- Process tables
  - Approval
  - Publication
- Bibliographic metadata tables

Document Storage



–Originals supplied by authors
–Versions edited by readers and reviewers

–Actual published versions whenever possible
–Supplied by authors or scanned by staff

Public Access: Online Catalog

INCLUDES:

- Bibliographic data stored as MARC record
- URLs (contained in MARC records) that link to full text documents stored on our server

Crosswalk from NIKE to MARC

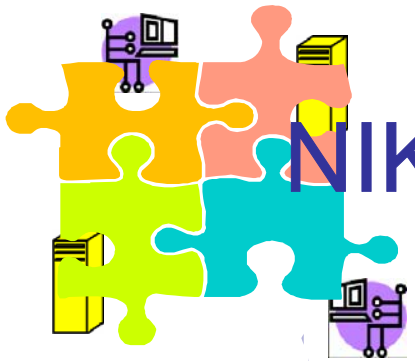
Sample of Crosswalk Map

245 TITLE STATEMENT (PUBLISHED TITLE)		
IND 1	Constant: 0	*(1XX fields not used)*
IND 2	[count of number of non-filing characters]	see Marc21 Initial Articles.doc
$a	db: Title	followed by subfield b--ends in space and colon; followed by subfield c--ends in space /
$b	db: subtitle (portion following ":")	ends in space /
$c	db: authors names: first last, first last.	ends in period

Crosswalk from NIKE to MARC

DATA INTEGRITY

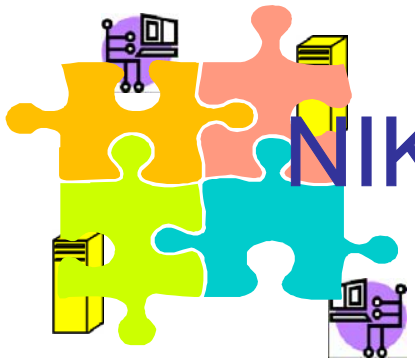
- Journal titles matched against authority list maintained in the online catalog
- Citations verified whenever possible before the crosswalk occurs



NIKE Rollout Plan (CUSTOMER'S PERSPECTIVE)

Production Stage One

- Electronic submission and upload of manuscripts
- Customized input and reporting screens
- Online tracking for NIST Editorial Review Board
- NIKE-interface access to bibliographic records and full text
- Public access to bibliographic records and full text for NIST series publications



NIKE Rollout Plan (CUSTOMER'S PERSPECTIVE)

Production Stage Two

- Online editing of manuscripts
- Online tracking of peer review (operating-unit level review)
- Electronic approval (signatures)
- Public access to bibliographic records and full text for all NIST-authored publications