

Developing Digital Repositories for Long-Term Preservation and Access to International Resources:

Technology and Funding Issues

Mark Lawrence Kornbluh

Michigan State University

TICFIA Project Directors Meeting
The University of Southern California
March 15, 2003

My Perspective

- Historian
- H-Net: Humanities and Social Sciences OnLine
- MATRIX: The Center for Humane Arts, Letters, and Social Sciences Online
- Humanities technology

The Challenge of Working in an Age of Revolutionary Technological Change

- Speed and breadth of change
- Shift in political and economic contexts
- Opaqueness of the future

Historians Look to the Past to Understand the Present and the Future

The Gutenberg Paradigm:

- The printed book is:
 - Reproducible
 - Immutable
 - Basis of modern educational and cultural systems

Understanding Technology in an Economic Context

- The wood pulp industry, the economics of publishing, and the democratization of education
- Vast international inequalities characterized the twentieth-century print world

The Digital Revolution Represents a Fundamental Paradigm Shift

- Changing economics of publishing
- Proliferation of publishing outlets
- Democratization of access
- Speed of communication
- Mutability of information
- Fundamental changes to the meaning of intellectual property

The Internet and Access to International Resources

- Increased ability to locate and organize resources
 - Online catalogues, finding aids, portals
- Increased access to fundamental library and archival resources: text, image, databases
- Increased access to multimedia: sound, video, 3D images
- Building complex objects

Phase One of Digital Web Projects:

- Portals, links, identifying and organizing content
- The online card catalogue and the page-view webpage ([LOC example](#))
- HTML web projects: collections, exhibits, educational materials, interpretations ([NARA example](#))

The Project Model

- Project-by-project development
 - Responding to individualized or team vision
 - Page views of individual objects
 - Idiosyncratic approaches to building complex objects
- Three routes
 - HTML hand coding
 - Use of commercial software (Cold Fusion)
 - Custom scripting (Perl, CGI, Java)

The Project Model

- Long run limitations and problems
 - Projects are stand-alone, not interoperable
 - Too expensive to develop and build
 - Unable to maintain and preserve

Time to Move Beyond the Project

Approach: Designing Digital Repositories

- To meet the preservation and access responsibilities of libraries, museums, and archives
- To meet the complex pedagogical and publishing needs of cultural heritage and educational institutions
- To take full advantage of networked digital media

Building Digital Repositories

- Need to plan for long-term preservation and access from the start
- Must design interoperable, distributive systems
- Design for multimedia from the outset
- Use and develop international best practices
- Design for re-use, change, growth
- Minimize labor costs

A Framework: Open Archive Information System Reference Model (OAIS)

- Common vocabulary, analytical framework to envision and plan repositories
- Allows OAIS compliant partners to exchange metadata and digital objects in predetermined fashion
- Allows for integrated searching across multiple repositories
- ISO Archiving Standard

Metadata Encoding and Transmission Standard (METS)

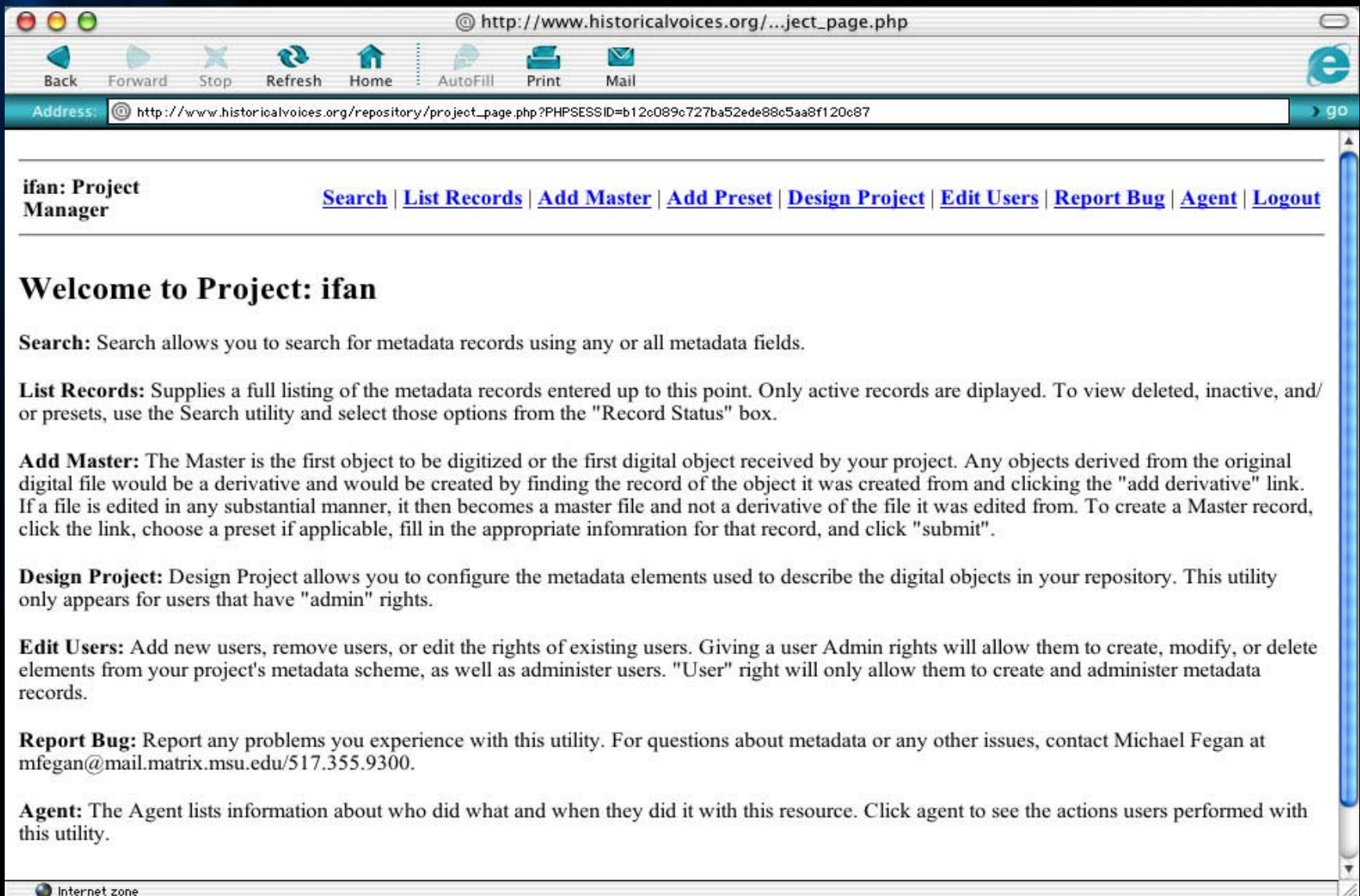
- Standardized, but extremely flexible, xml metadata schema
- Incorporates all the elements of metadata necessary for long-term preservation
- “Bucket” approach to structural, descriptive, and administrative metadata
- Positioned to be a tool for exchange of metadata and digital objects
- Like OAIS, high level agreement to facilitate communication

MATRIX Digital Repository:

Phase I: PHP/MYSQL Implementation of METS

- Interoperability through a single database solution
- Schematized database design allowing for dynamic project creation
- Modular approach to delivery applications
- Facilitates creation of complex digital objects
- Full flexibility searching
- XML ready

MATRIX Modular Project Design



The screenshot shows a web browser window with the address bar containing the URL: http://www.historicalvoices.org/repository/project_page.php?PHPSESSID=b12c089c727ba52ede88c5aa8f120c87. The browser's navigation bar includes buttons for Back, Forward, Stop, Refresh, Home, AutoFill, Print, and Mail. The page content is titled "ifan: Project Manager" and features a navigation menu with links for Search, List Records, Add Master, Add Preset, Design Project, Edit Users, Report Bug, Agent, and Logout. The main content area provides a "Welcome to Project: ifan" message and detailed instructions for various project management utilities.

ifan: Project Manager [Search](#) | [List Records](#) | [Add Master](#) | [Add Preset](#) | [Design Project](#) | [Edit Users](#) | [Report Bug](#) | [Agent](#) | [Logout](#)

Welcome to Project: ifan

Search: Search allows you to search for metadata records using any or all metadata fields.

List Records: Supplies a full listing of the metadata records entered up to this point. Only active records are displayed. To view deleted, inactive, and/or presets, use the Search utility and select those options from the "Record Status" box.

Add Master: The Master is the first object to be digitized or the first digital object received by your project. Any objects derived from the original digital file would be a derivative and would be created by finding the record of the object it was created from and clicking the "add derivative" link. If a file is edited in any substantial manner, it then becomes a master file and not a derivative of the file it was edited from. To create a Master record, click the link, choose a preset if applicable, fill in the appropriate information for that record, and click "submit".

Design Project: Design Project allows you to configure the metadata elements used to describe the digital objects in your repository. This utility only appears for users that have "admin" rights.

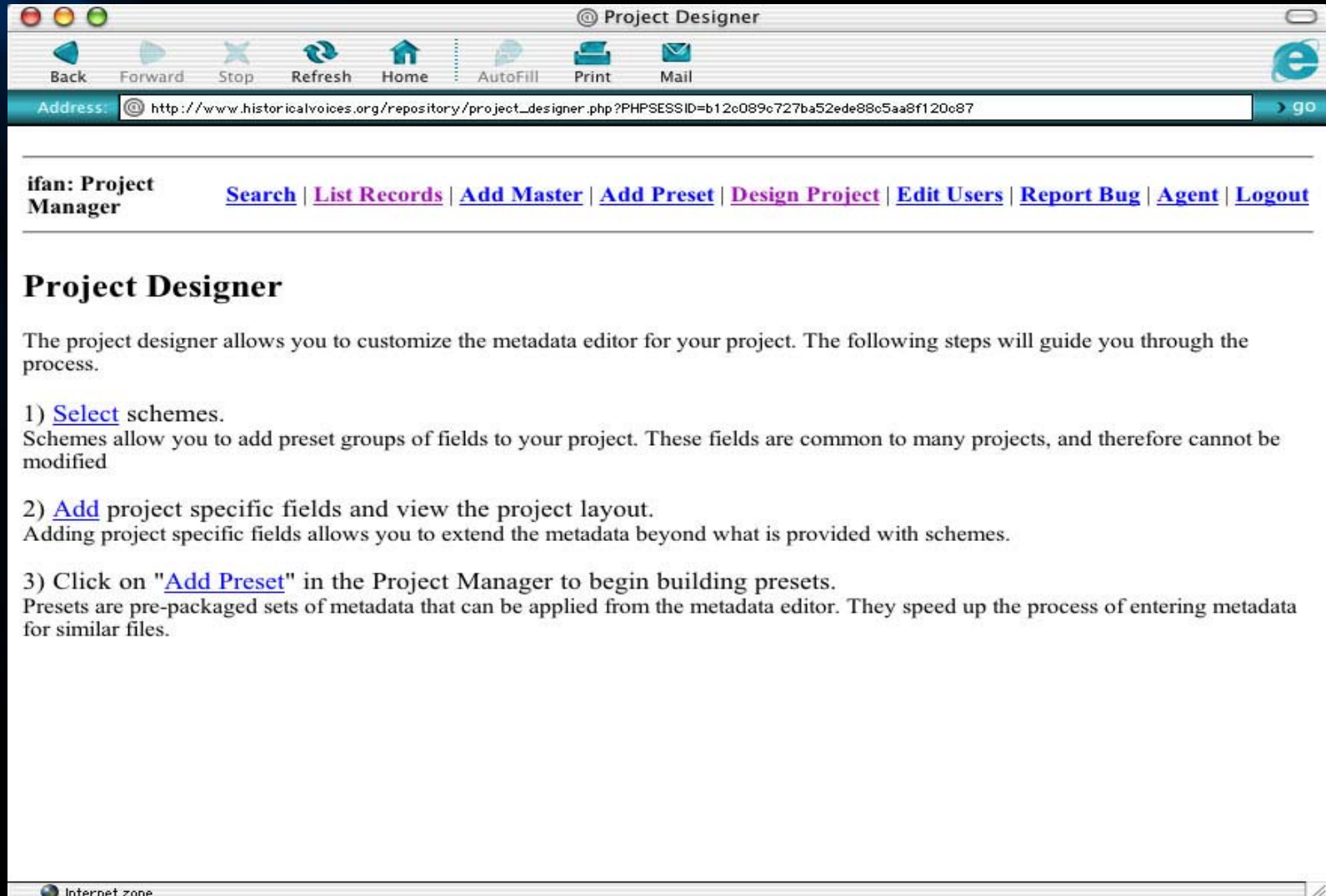
Edit Users: Add new users, remove users, or edit the rights of existing users. Giving a user Admin rights will allow them to create, modify, or delete elements from your project's metadata scheme, as well as administer users. "User" right will only allow them to create and administer metadata records.

Report Bug: Report any problems you experience with this utility. For questions about metadata or any other issues, contact Michael Fegan at mfegan@mail.matrix.msu.edu/517.355.9300.

Agent: The Agent lists information about who did what and when they did it with this resource. Click agent to see the actions users performed with this utility.

Internet zone

MATRIX Modular Project Design



The screenshot shows a web browser window with the title "Project Designer". The address bar contains the URL: http://www.historicalvoices.org/repository/project_designer.php?PHPSESSID=b12c089c727ba52ede88c5aa8f120c87. The browser's navigation bar includes buttons for Back, Forward, Stop, Refresh, Home, AutoFill, Print, and Mail. Below the address bar, there is a navigation menu with the following links: [Search](#), [List Records](#), [Add Master](#), [Add Preset](#), [Design Project](#), [Edit Users](#), [Report Bug](#), [Agent](#), and [Logout](#). The main content area features a heading "Project Designer" followed by a paragraph: "The project designer allows you to customize the metadata editor for your project. The following steps will guide you through the process." Below this, there are three numbered steps: 1) [Select](#) schemes. Schemes allow you to add preset groups of fields to your project. These fields are common to many projects, and therefore cannot be modified. 2) [Add](#) project specific fields and view the project layout. Adding project specific fields allows you to extend the metadata beyond what is provided with schemes. 3) Click on "[Add Preset](#)" in the Project Manager to begin building presets. Presets are pre-packaged sets of metadata that can be applied from the metadata editor. They speed up the process of entering metadata for similar files. At the bottom left of the browser window, there is a small icon and the text "Internet zone".

ifan: Project Manager

[Search](#) | [List Records](#) | [Add Master](#) | [Add Preset](#) | [Design Project](#) | [Edit Users](#) | [Report Bug](#) | [Agent](#) | [Logout](#)

Project Designer

The project designer allows you to customize the metadata editor for your project. The following steps will guide you through the process.

- 1) [Select](#) schemes.
Schemes allow you to add preset groups of fields to your project. These fields are common to many projects, and therefore cannot be modified
- 2) [Add](#) project specific fields and view the project layout.
Adding project specific fields allows you to extend the metadata beyond what is provided with schemes.
- 3) Click on "[Add Preset](#)" in the Project Manager to begin building presets.
Presets are pre-packaged sets of metadata that can be applied from the metadata editor. They speed up the process of entering metadata for similar files.

Internet zone

MATRIX Modular Project Design

ifan: Project Manager

[Search](#) | [List Records](#) | [Add Master](#) | [Add Preset](#) | [Design Project](#) | [Edit Users](#) | [Report Bug](#) | [Agent](#) | [Logout](#)

Scheme Selection

Schemes are metadata sets that can be added to your project. These fields are common to many projects, and therefore cannot be modified. Although the fields for a scheme can't be modified, you can add fields to the core set of metadata that only will appear for your projects. If you think the metadata schemes need to be altered in some way, please contact MATRIX.

These are the schemes that are currently being used:

These are the schemes that you can add to the project:

Internet zone

MATRIX Modular Project Design

Field Editor

The field editor allows you to create or modify the information about a particular field that will describe an object for your project. Fill in the appropriate information for the metadata field and press the submit button. Click [here](#) for definitions of each of the forum fields in the field editor.

Group: or if other:

Name:

Required

Sequence:

Description:

Caption:

Type:

Note:

Internet zone

MATRIX Modular Project Design

METADATA EDITOR: ifan-a0a0a7-a

Back Forward Stop Refresh Home AutoFill Print Mail

Address: http://www.historicalvoices.org/repository/metadata_editor_form.php?pbd=ifan-a0a0a7-a&PHPSESSID=b12c089c727ba52ede88c5aa8f120c87

METADATA EDITOR: ifan-a0a0a7-a

To create a record, fill in the appropriate information for that record and the option to update the record by clicking edit for that particular record on the right.

data filled in at a later time

Internet Explorer Script Alert

! Improper format. Please enter a name in the format: lastname, firstname

OK

descriptive

Title:

Author/Interviewee:

Interviewer:

Keywords:

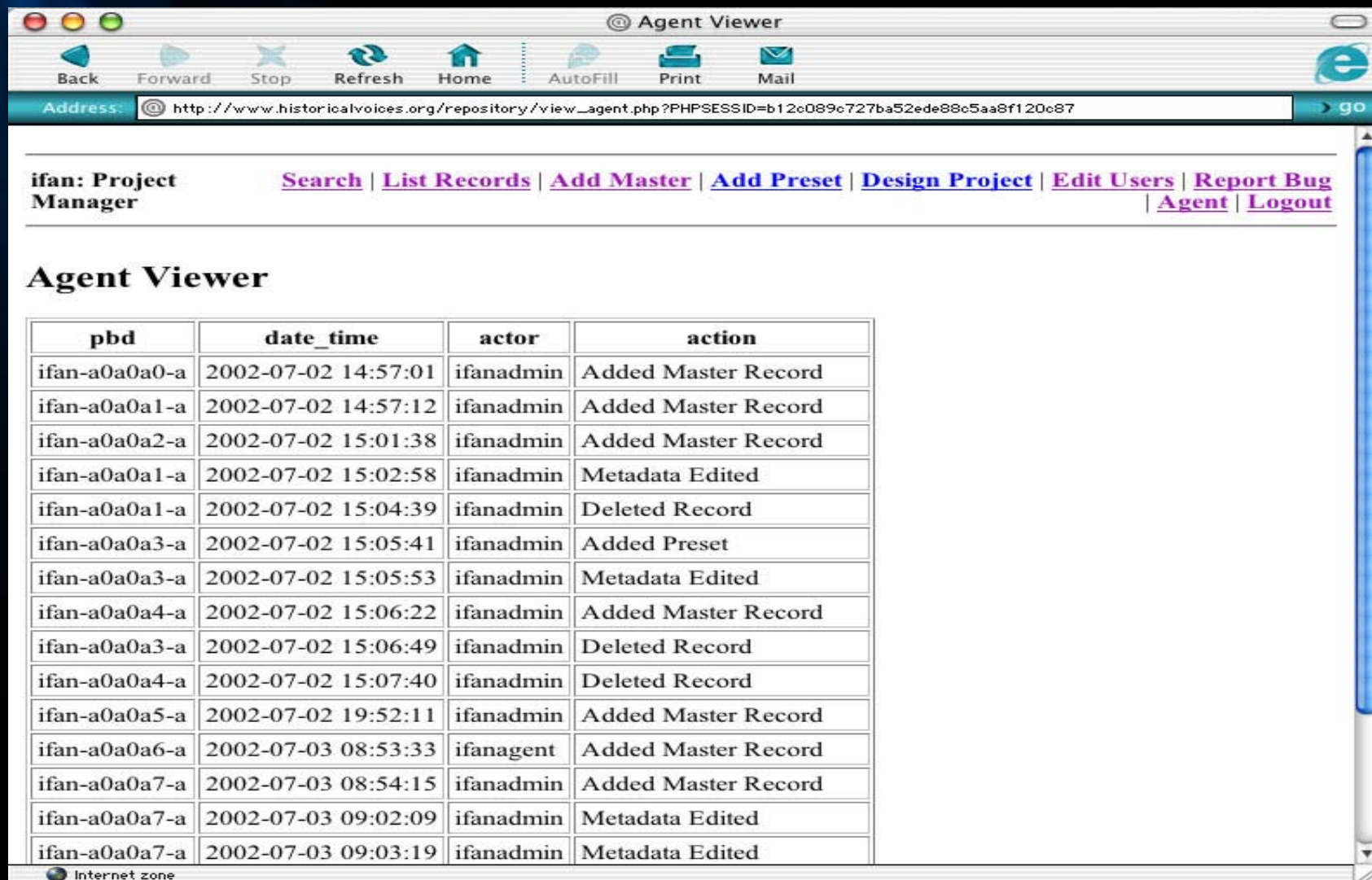
Description:

Publication Date:

Language:
french
Wolof

Internet zone

MATRIX Modular Project Design



ifan: Project Manager

[Search](#) | [List Records](#) | [Add Master](#) | [Add Preset](#) | [Design Project](#) | [Edit Users](#) | [Report Bug](#) | [Agent](#) | [Logout](#)

Agent Viewer

pbd	date_time	actor	action
ifan-a0a0a0-a	2002-07-02 14:57:01	ifanadmin	Added Master Record
ifan-a0a0a01-a	2002-07-02 14:57:12	ifanadmin	Added Master Record
ifan-a0a0a02-a	2002-07-02 15:01:38	ifanadmin	Added Master Record
ifan-a0a0a01-a	2002-07-02 15:02:58	ifanadmin	Metadata Edited
ifan-a0a0a01-a	2002-07-02 15:04:39	ifanadmin	Deleted Record
ifan-a0a0a03-a	2002-07-02 15:05:41	ifanadmin	Added Preset
ifan-a0a0a03-a	2002-07-02 15:05:53	ifanadmin	Metadata Edited
ifan-a0a0a04-a	2002-07-02 15:06:22	ifanadmin	Added Master Record
ifan-a0a0a03-a	2002-07-02 15:06:49	ifanadmin	Deleted Record
ifan-a0a0a04-a	2002-07-02 15:07:40	ifanadmin	Deleted Record
ifan-a0a0a05-a	2002-07-02 19:52:11	ifanadmin	Added Master Record
ifan-a0a0a06-a	2002-07-03 08:53:33	ifanagent	Added Master Record
ifan-a0a0a07-a	2002-07-03 08:54:15	ifanadmin	Added Master Record
ifan-a0a0a07-a	2002-07-03 09:02:09	ifanadmin	Metadata Edited
ifan-a0a0a07-a	2002-07-03 09:03:19	ifanadmin	Metadata Edited

Internet zone

Modularity for Flexible Delivery:

- Alternative Delivery formats
 - Archival: Search, browse, sort
 - Galleries: Content based-groupings
 - Exhibits: Structured presentations
 - Educational classrooms
- Diverse interfaces for diverse audiences
- SMIL and flash presentations of multimedia material

Modularity for Future Growth:

- Developing user-generated content over time to enhance repository
 - Commentary
 - Supporting content
 - Usage information
 - Collateral links

Moving to a Full XML System:

Phase II: MATRIX Digital Repository

- Ingesting and storing native XML
 - The best solution for preservation
 - Designed for interoperability (harvesting)
 - Hierarchical Structure of xml has real cost benefits for metadata entry

- Utilizing Cocoon
 - Pipelines facilitate multi-channel delivery
 - Reusable programming has real cost benefits

Looking Forward to Digital Repositories

- Move from projects to program
- Shift from web pages to web services
- Realizing the digital revolution
 - Access and preservation
 - Depth and richness of multimedia
 - Multiple Voices
 - Interactivity

Funding: Looking Beyond the US Department of Education

- Federal funding
 - NEH, IMLS, NSF
 - Departments of Defense, State, Commerce
- International Funding: EU, UK, Japan
- Charitable Foundations
 - Mellon, Ford, Getty
- Corporate Funding
 - Foundations
 - Partnerships

Major Sources of Federal Funding

■ IMLS

- Museums, libraries, museum-library cooperation

■ NEH

- Preservation and Access, Public Programs, Education

■ NSF

- Digital Library Initiatives (DLI2, IDL, NSDL, DLI3?)
- Research challenges, foreign languages, esoteric content, images, 3d objects, preservation issues, teaching and learning

Farther A Field with Federal Funding

- Department of Defense
 - Language resources
- Department of State
 - Public Affairs (old USIA)
 - USAID (and ALO)
 - Peace Corp
- Department of Commerce
 - Access focus
- Park Service
 - Domestic resources

International Funding

- EU Framework 6
 - Cultural heritage, multilingual
- JISC (UK)
 - NSF counterpart, but more open to cultural heritage
- UNESCO
- Development Funding
 - Cultural heritage has higher visibility in European development activities
- National funding and quasi-independent national foundations
 - Seeking broader visibility for national culture

Charitable Foundations

- Mellon
 - Library resources
 - Portals
 - Licensed content/ economic sustainability/ research
- Ford
 - Libraries, education
- Getty
 - Mostly operating foundation
- Rockefeller
 - Museums

Corporate Foundations and Partnerships

- Telecommunications
 - ATT, Sprint, MCI
 - Baby Bells
- Computer companies
 - Gates, HP, IBM, Dell
- Corporations with large international presence
 - Coke, auto companies
- Corporations with ethnic ties, community ties

Hooks for Funding Digital International Content

- Educational use
- Foreign language content
- Community-based interest group
- Interoperability
- Coordinated partnerships
- Research challenges for IT

Larger Funding Challenges: Meeting Institutional Costs, Development Costs, and Programmatic Costs

- The need for a strong institutional base:
 - Need to on-going staff, computers, etc.
 - On-going preservation costs
 - Interoperability concerns

Final Thoughts: The Big Challenges (Besides Funding):

- IT programming and tool development
- International cooperation and agreement on flexible standards
- Rethinking intellectual property
- Overcoming economic inequalities, especially bandwidth

Partnerships Needed to Build Cultural Heritage Digital Repositories:

- IT skills and participation in international community
- Library science and archival knowledge
- Content-based academic knowledge
- Educational linkages