

The Open Source Evaluation Project at the National Archives

Jo Pettitt
24 June 2003

Stance

- All new IT systems in government departments must be e-GIF compliant
- E-GMS compliant as well
- Assumes XML wherever (in any layer) practical
- Government wants departments to investigate OSS wherever practical

Archivists Proactive Role

- Archivists need to anticipate technology dependencies for future transfers
- Need to influence (if possible) technologies adopted by creators so that...
- Costs controlled and speed of availability supported

Our Approach To Digital Preservation

- Practical approach taken
 - Digital Archive built
 - PRONOM released internally
- Both are e-GIF compliant
- Both use XML
- And both have thrown up issues relating to file format longevity

Archive Search Screen


Digital Archive - Search Form - Microsoft Internet Explorer provided by Public Record Office

File Edit View Favorites Tools Help

Back Forward Stop Search Favorites History

Address http://ds2wck.digitalarchive/templates/_security_check Go Links

Google Search Web Search Site News PageRank PageDiffs Uo Highlight


the national archives

Home **Search** Browse Log Off Change Password Help

Search

Search

The fields marked with * are mandatory:

Search phrase: *

From year:

To year:

Department:





































Number of matches per page:

Look-up by record ID

Please input the ID of a record and press the button:

Record ID:

Done

Start                                     <

PRONOM Search Screen

PRONOM - File and Product Report Generator - Microsoft Internet Explorer provided by Public Record Office

File Edit View Favorites Tools Help

Address http://harris/PRONOM/Product/proProductSearch.asp?tab=tblReport

Google Search Web

Public Record Office
The National Archives

PRONOM

PRONOM Home Vendor Product Version File Format Review Contact Us

+ Include ALL products

+ Include only products whose names begin with ✓ X

+ Include only products whose names contain ✓ X

+ Include ALL vendors

+ Include only vendors whose names begin with ✓ X

+ Include only vendors whose names contain ✓ X

Go Show me all products from all vendors.

Save As XML Save As CSV Print Friendly Listing

Current Report: Listing of ALL products from ALL vendors.
Matching Records: 43

Product	Version	Vendor
Access	2000 (9.0)	Microsoft
Adobe Acrobat	5.0	Adobe
AutoCAD	2004	Autodesk
Corel Draw	11	Corel
Excel	2000 (9.0)	Microsoft
Excel	97 (8.0)	Microsoft

Show me all products from all vendors.

Local intranet

PRONOM XML Schema

```
<?xml version="1.0" encoding="UTF-16" ?>
- <PRONOM-Report xmlns:rs="urn:schemas-microsoft-com:rowset" xmlns:z="#RowsetSchema">
- <report_product_detail>
  - <Product>
    <ProductID>20</ProductID>
    <ProductName>Excel</ProductName>
    <VendorID>1</VendorID>
    <VendorName>Microsoft</VendorName>
    <Version>2000 (9.0)</Version>
    <SupportName>Microsoft</SupportName>
    <SupportID>1</SupportID>
    <ThirdPartySupportID />
    <ThirdPartySupportName />
    <DefaultExtensionID>132</DefaultExtensionID>
    <DefaultExtension>xls</DefaultExtension>
    <FileTypeID>32</FileTypeID>
    <FileTypeDescription>Microsoft Excel 2000 Workbook</FileTypeDescription>
    <Category>Spreadsheet</Category>
    <CategoryID>3</CategoryID>
    <MediaFormat>CD-ROM</MediaFormat>
    <MediaFormatID>1</MediaFormatID>
    <ReleaseDate>10 Jun 1999</ReleaseDate>
    <WithdrawnDate>30 Jun 2006</WithdrawnDate>
    <CopyrightOwner>Microsoft</CopyrightOwner>
    <CopyrightStatus>Copyright Owner</CopyrightStatus>
```

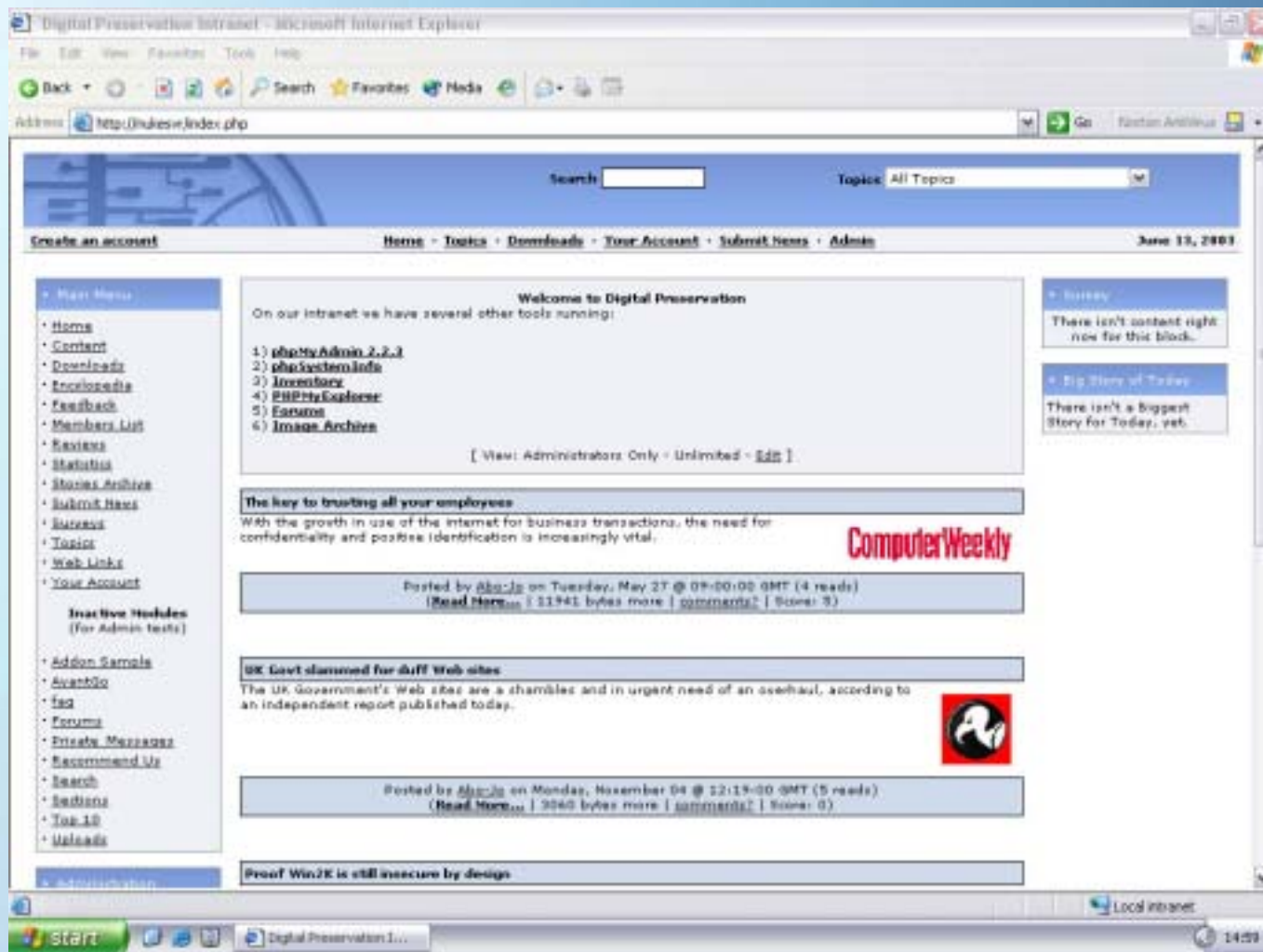
Trials And Pilots

- The need to resolve
 - File format longevity
 - And other issues discovered
- Have led to:
 - Trials and pilots programme
 - Work on open source products have been included

Open Source trials

- PHP Nuke
- Post Nuke
- MySQL
- Dspace
- E-Prints
- Image Magick
- Open Office
- SuSE Linux 8.2
- PHP MySQL asset tracking application

PHP Nuke & Post Nuke



MySQL Database

nuke.nuke_users running on localhost - phpMyAdmin 2.2.3 - Microsoft Internet Explorer

File Edit View Favorites Tools Help

Back Forward Stop Search Favorites Media

Address http://localhost/phpMyAdmin/ Go Norton Antivirus

Database nuke - table nuke_users running on localhost

[Browse] [Select] [Insert] [Empty] [Drop]

	Field	Type	Attributes	Null	Default	Extra	Action
<input type="checkbox"/>	uid	int(11)		No		auto_increment	Change Drop Primary Index Unique Fulltest
<input type="checkbox"/>	name	varchar(50)		No			Change Drop Primary Index Unique Fulltest
<input type="checkbox"/>	uname	varchar(25)		No			Change Drop Primary Index Unique Fulltest
<input type="checkbox"/>	email	varchar(255)		No			Change Drop Primary Index Unique Fulltest
<input type="checkbox"/>	femail	varchar(255)		No			Change Drop Primary Index Unique Fulltest
<input type="checkbox"/>	url	varchar(255)		No			Change Drop Primary Index Unique Fulltest
<input type="checkbox"/>	user_avatar	varchar(30)		Yes	NULL		Change Drop Primary Index Unique Fulltest
<input type="checkbox"/>	user_regdate	varchar(20)		No			Change Drop Primary Index Unique Fulltest
<input type="checkbox"/>	user_icq	varchar(15)		Yes	NULL		Change Drop Primary Index Unique Fulltest
<input type="checkbox"/>	user_occ	varchar(100)		Yes	NULL		Change Drop Primary Index Unique Fulltest
<input type="checkbox"/>	user_from	varchar(100)		Yes	NULL		Change Drop Primary Index Unique Fulltest
<input type="checkbox"/>	user_intrest	varchar(150)		Yes	NULL		Change Drop Primary Index Unique Fulltest
<input type="checkbox"/>	user_sig	varchar(255)		Yes	NULL		Change Drop Primary Index Unique Fulltest
<input type="checkbox"/>	user_viewemail	tinyint(2)		Yes	NULL		Change Drop Primary Index Unique Fulltest
<input type="checkbox"/>	user_theme	int(3)		Yes	NULL		Change Drop Primary Index Unique Fulltest
<input type="checkbox"/>	user_aim	varchar(18)		Yes	NULL		Change Drop Primary Index Unique Fulltest
<input type="checkbox"/>	user_yim	varchar(25)		Yes	NULL		Change Drop Primary Index Unique Fulltest
<input type="checkbox"/>	user_msnm	varchar(25)		Yes	NULL		Change Drop Primary Index Unique Fulltest
<input type="checkbox"/>	pass	varchar(40)		No			Change Drop Primary Index Unique Fulltest
<input type="checkbox"/>	storynum	tinyint(4)		No	10		Change Drop Primary Index Unique Fulltest

Local intranet 15:00

DSpace



The screenshot shows a Microsoft Internet Explorer browser window displaying the DSpace website. The browser's address bar shows the URL <https://hdl.handle/1725.1/1793>. The website header features the DSpace logo and the MIT Libraries logo. The main content area is titled "Technology and Law Program" and "Collection home page". It includes a search bar with the text "Technology and Law Program" and a "Go" button. Below the search bar, there are buttons for "Titles", "Authors", and "By Date". A "Submit to This Collection" button is also present. The right sidebar lists "Recent Submissions" with links to various documents. The bottom of the browser window shows the Windows taskbar with several open applications.

DSpace at MIT: Technology and Law Program - Microsoft Internet Explorer provided by Public Record Office

File Edit View Favorites Tools Help

Back Forward Stop Search Favorites History

Address <https://hdl.handle/1725.1/1793>

Google dspace Search Web Search Site News PageRank Page Off Up Highlight dspace

DSpace™ About DSpace MIT Libraries

Search DSpace: Go
[Advanced Search](#)

[Home](#)

Browse

- [Communities & Collections](#)
- [Titles](#)
- [Authors](#)
- [By Date](#)

Sign on to:

- [Receive email updates](#)
- [My DSpace](#) authorized users
- [Edit Profile](#)
- [Help](#)

[DSpace at MIT >](#)
[Center for Technology, Policy, and Industrial Development \(CTPID\) >](#)

Technology and Law Program

Collection home page

In: Technology and Law Program

Search for Go

or browse [Titles](#) [Authors](#) [By Date](#)

[Submit to This Collection](#)

Subscribe to this collection to receive daily e-mail notification of new additions [Subscribe](#)

The Technology and Law Program offers a cluster of graduate-level subjects associated with the Technology and Policy Program (TPP) at MIT, as well as research opportunities at the interface of law and technology. The program offers a two-semester sequence in Environmental Law and Policy and a one-semester course in Technology, Law, and the Working Environment. Students from MIT, the Harvard School of Public Health, and Harvard Law School participate in these interactive, Socratic teaching offerings. Law, Technology, and Public Policy, which is a core subject in the Technology and Policy Program, is also offered. The subject, Sustainability, Trade, and Environment, offered through both the School of Management and TPP, involves the

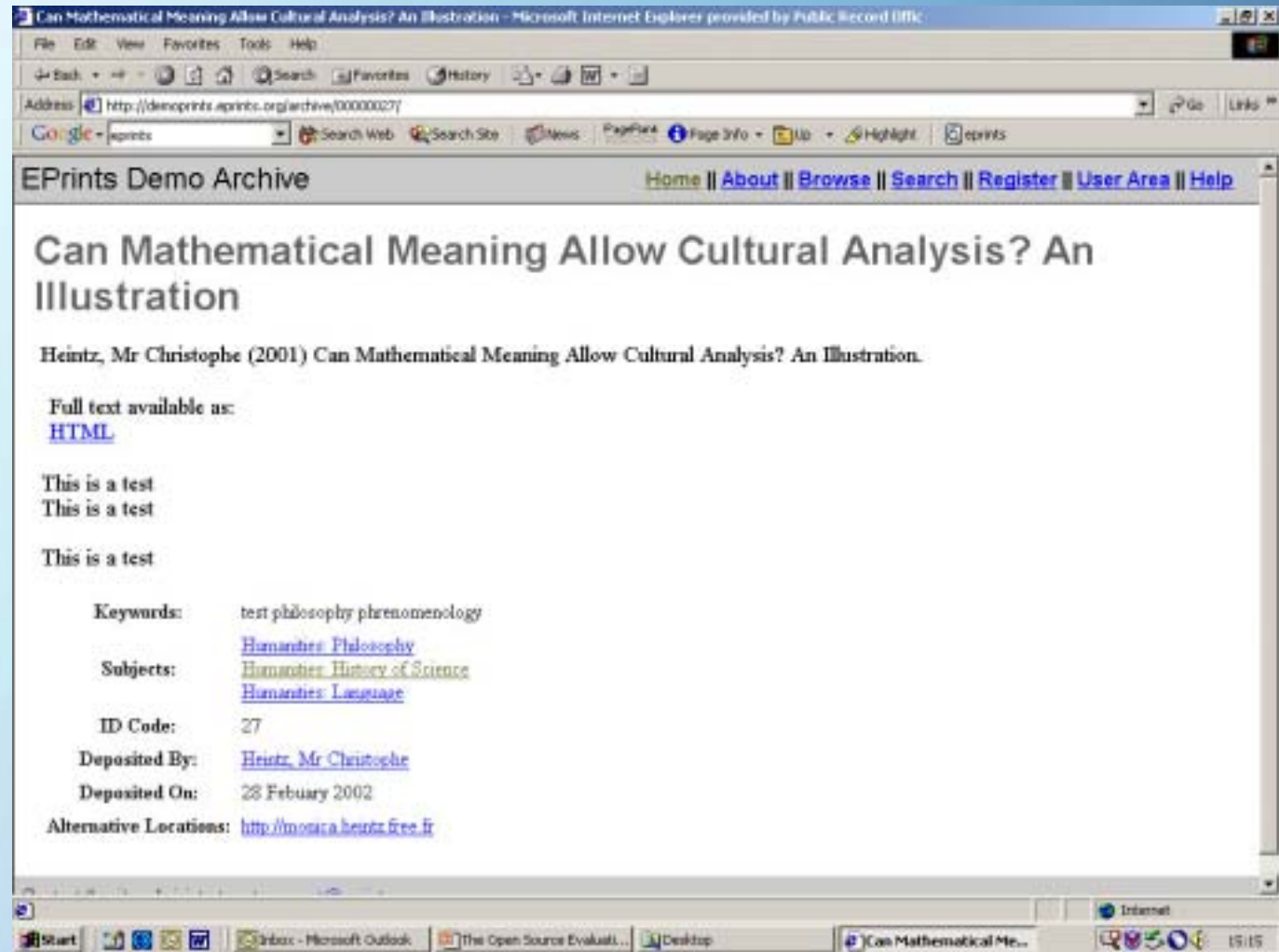
Recent Submissions

- [An Innovation-Based Strategy for a Sustainable Environment](#)
- [Justice in the Global Work Life: The Right to Know, to Participate, and to Benefit in Sustainable Industrial Transformations](#)
- [Pathways to Sustainability: Evolution or Revolution?](#)
- [Implementing A Precautionary Approach in Decisions Affecting Health, Safety, and the Environment: Risk, Technology Alternatives, and Tradeoff Analysis](#)
- [Social and Policy Implications of Low-Level Exposures to Chemicals](#)

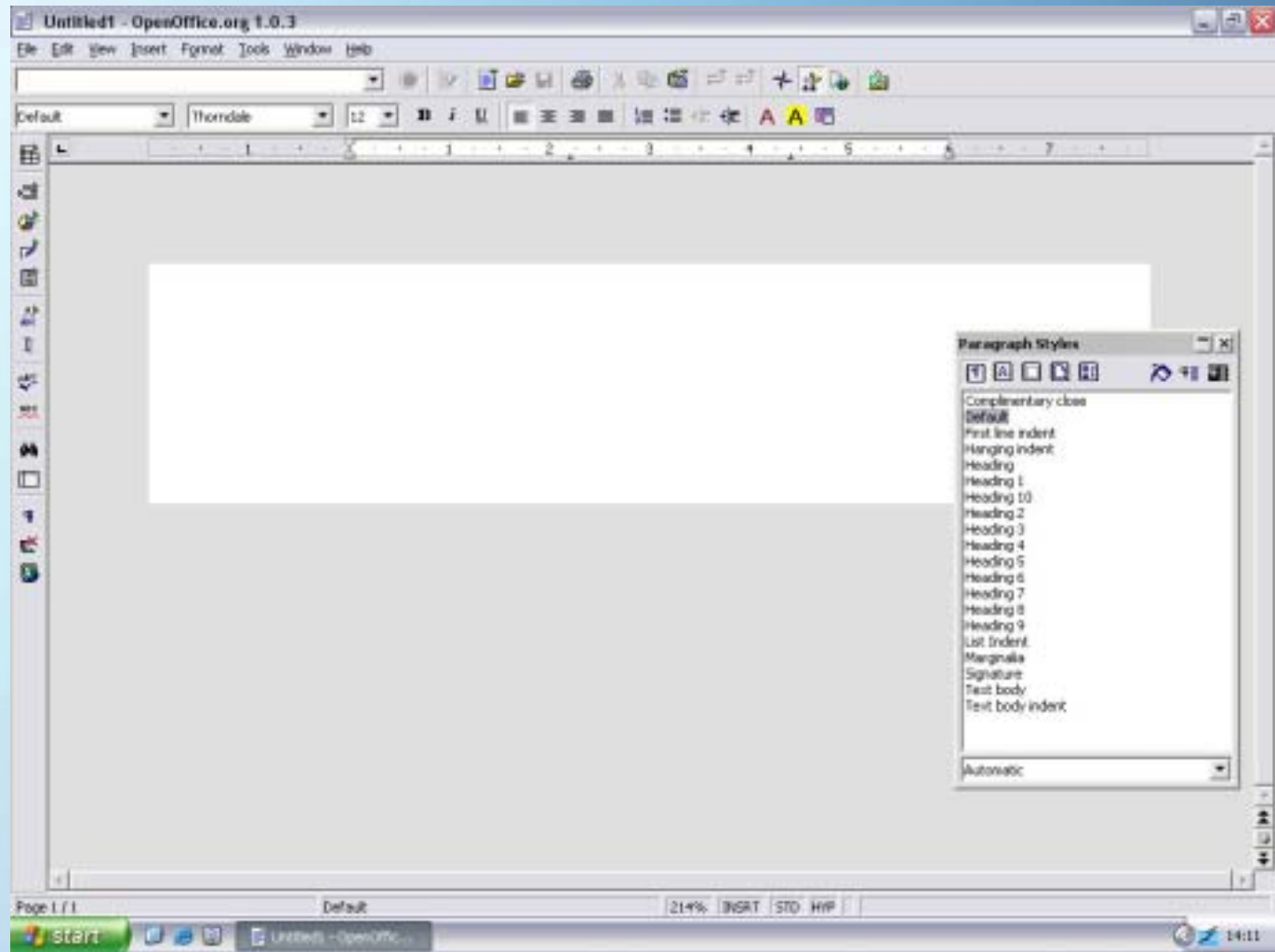
Done Internet

Start Inboxes - Microsoft Outlook The Open Source Evaluat... Desktop DSpace at MIT Techn... 15:23

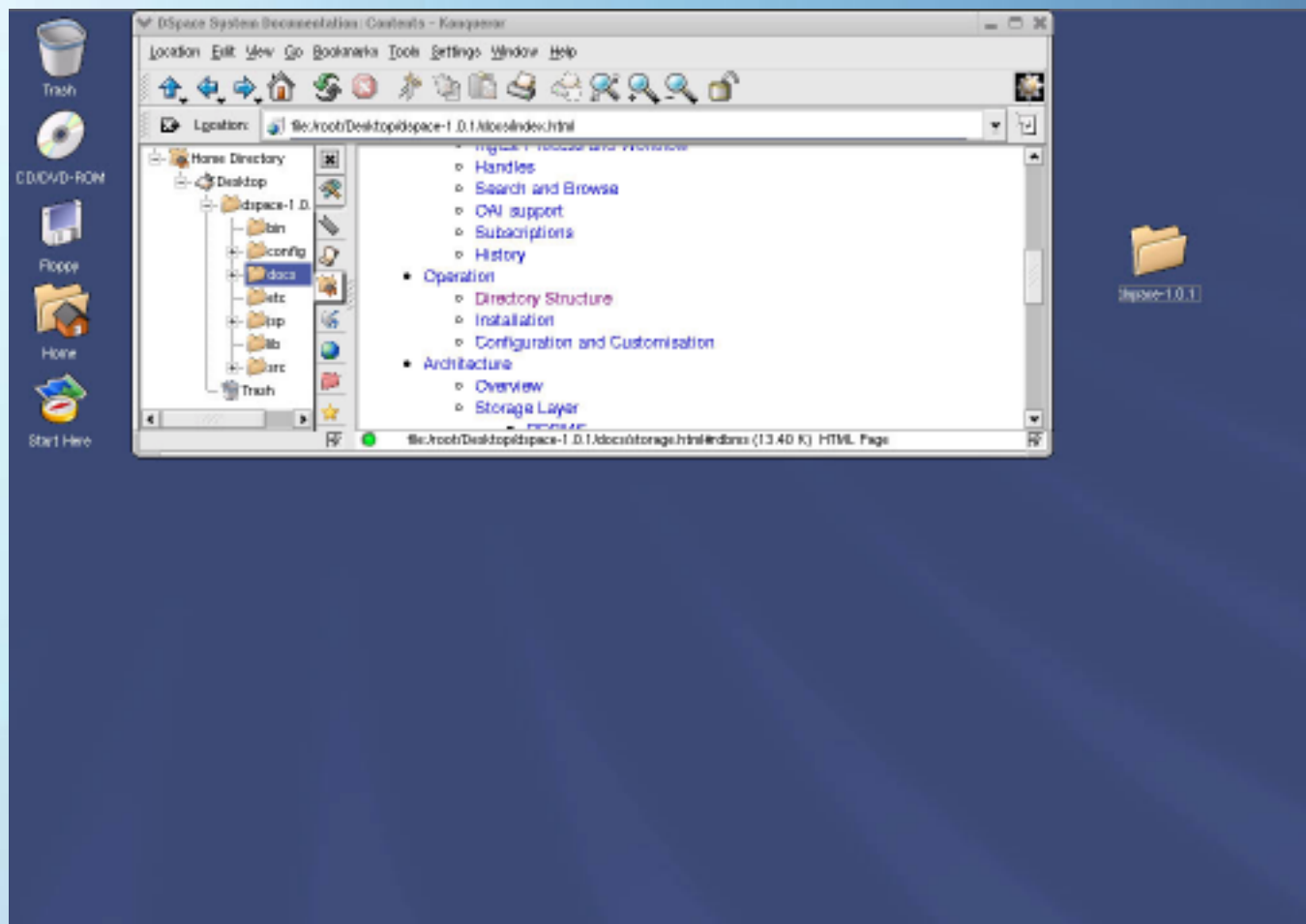
EPrints



Open Office



RedHat 9.0 Linux



SuSE 8.2 Linux

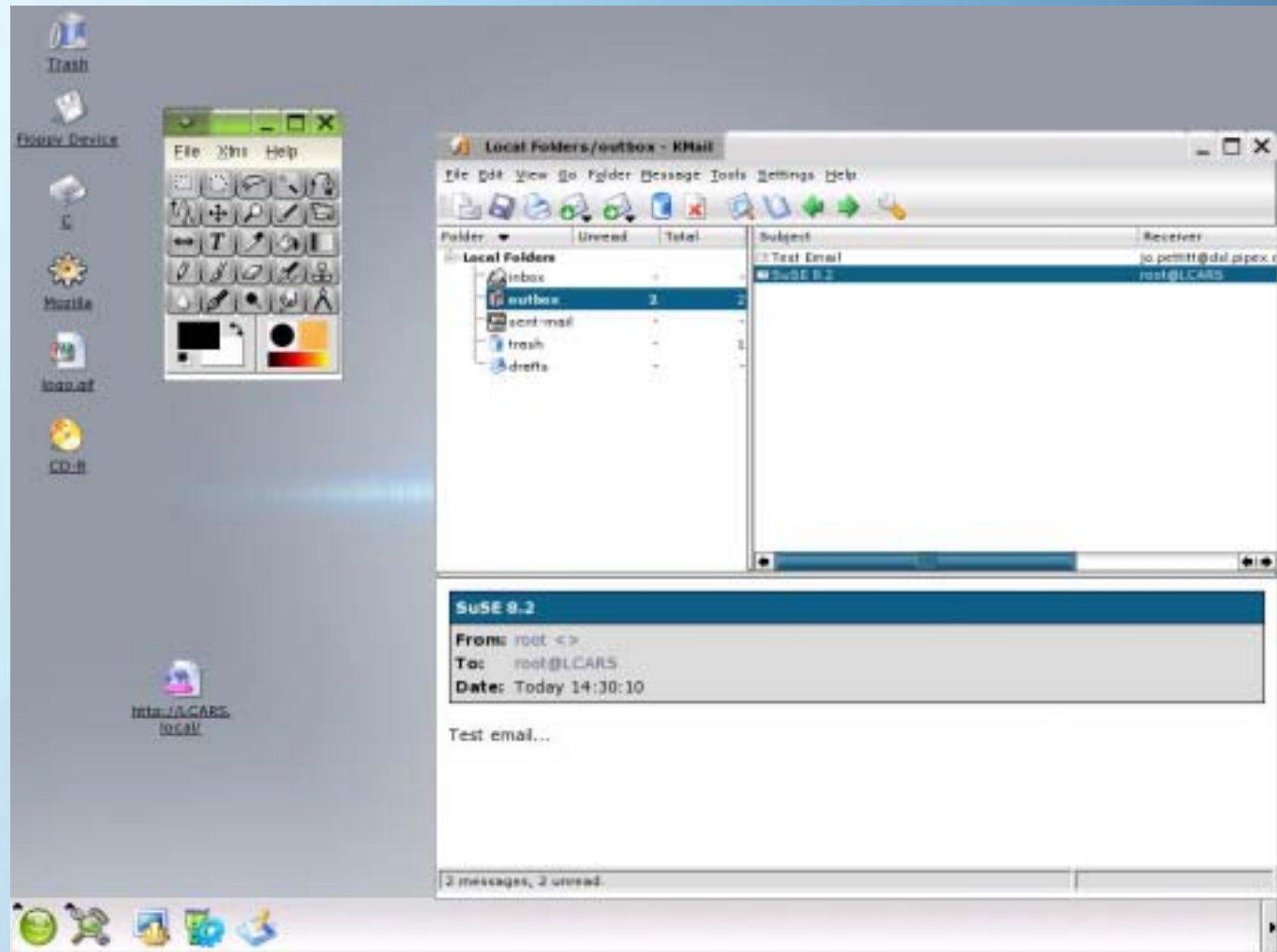
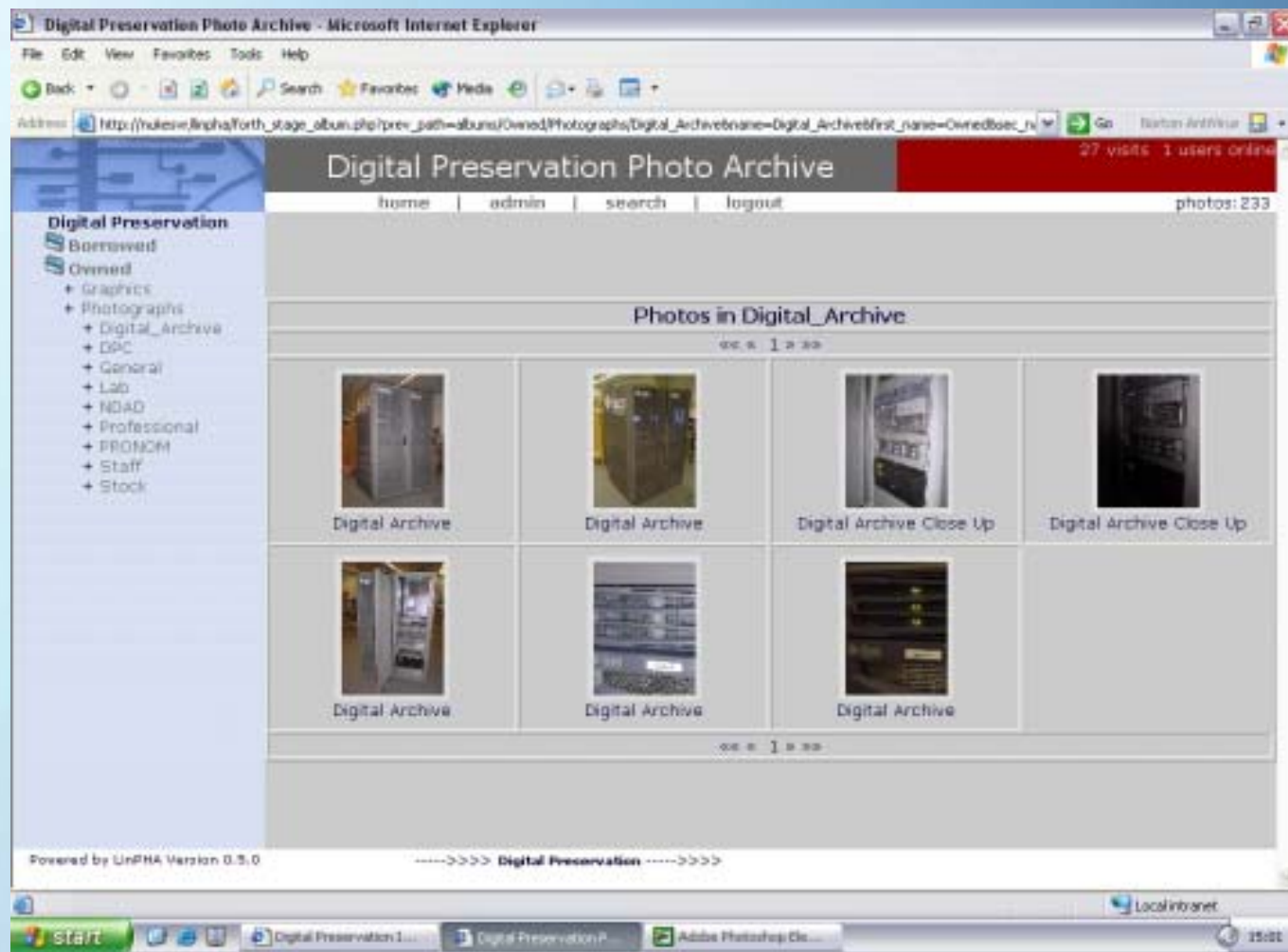
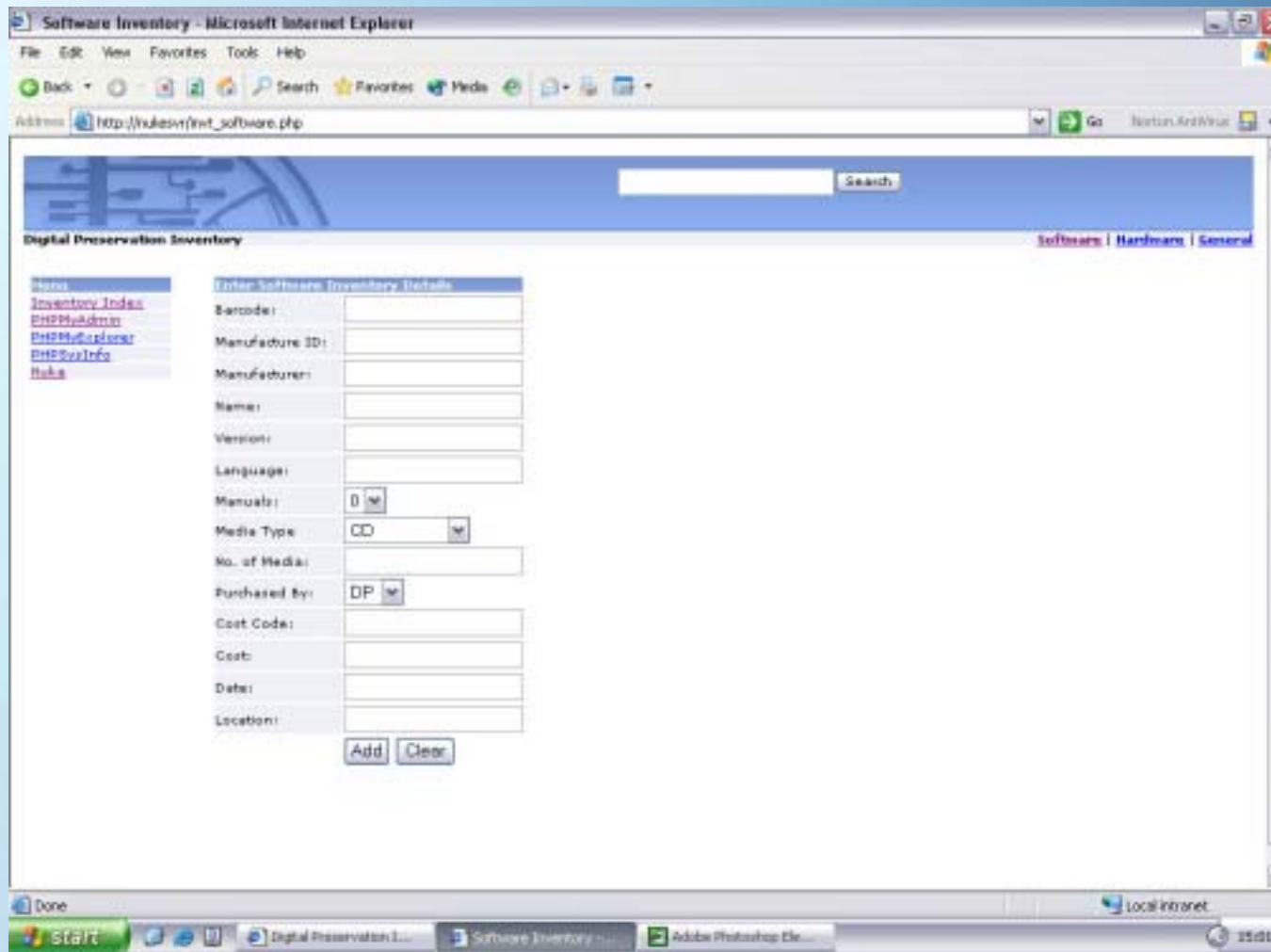


Image Magick



PHP MySQL Asset Tracking



The screenshot shows a web browser window titled "Software Inventory - Microsoft Internet Explorer". The address bar displays "http://nukesvn/inet_software.php". The page content includes a search bar at the top right. Below the search bar, there are three tabs: "Software", "Hardware", and "General". On the left side, there is a sidebar with a "Home" link and a list of links: "Inventory Index", "PHPMyAdmin", "PHPMyAdmin", "PHPMyAdmin", "PHPMyAdmin", and "Links". The main content area is titled "Enter Software Inventory Details" and contains a form with the following fields: Barcode, Manufacture ID, Manufacturer, Name, Version, Language, Manual (with a dropdown menu), Media Type (with a dropdown menu), No. of Media, Purchased By (with a dropdown menu), Cost Code, Cost, Date, and Location. At the bottom of the form are "Add" and "Clear" buttons. The status bar at the bottom of the browser window shows "Done", "Local intranet", and a taskbar with icons for "Start", "Digital Preservation I...", "Software Inventory...", and "Adobe Photoshop El...".

Upside

- Cheap/free
- Code is open
- No hidden spyware
- No strict EULAs

Downside

- Some immature releases
- Poor support
- Difficult to set-up
- Poor documentation



Thank you!