

ENSURING LONG-TERM ACCESS TO
GOVERNMENT DOCUMENTS THROUGH
VIRTUALIZATION

Geoffrey Brown
Indiana University
Department of Computer Science



[GPO
ACCESS](#)

[FDLP
Desktop](#)

[University of
Chicago](#)

[University of
Illinois](#)

[Indiana
University](#)

[University of
Iowa](#)

[University of
Michigan](#)

[Michigan
State
University](#)

[University of
Minnesota](#)

[Northwestern
University](#)

[Ohio State
University](#)

[Pennsylvania
State
University](#)

[Purdue
University](#)

[University of
Wisconsin-
Madison](#)

CIC Floppy Disk Project

The CIC Floppy Disk Project is a partnership between the [U.S. Government Printing Office](#) and the [Indiana University, Bloomington Libraries](#) on behalf of the [Committee on Institutional Cooperation \(CIC\)](#), making publications that were distributed to federal depository libraries on floppy disk available via FTP over the Internet. The site is designed to provide a central location through which Federal Data, made available on floppy diskettes, can be identified, located and downloaded. While the project provides the information as it was originally presented on the floppy disks, it cannot provide software which might be needed to fully operate the disk. Records provided within the FDP inform the user about the specifications and may refer users to "read me" files. Consultation with local computer administrators may be required. With the ever-increasing number of products being released by the Government Printing Office, **FDP** enables libraries to either fill in gaps to their collections, or provide an immediate access point for patrons. While the collection represents over 200 entries, our collection is not complete. We would welcome contributions from your library's holdings or suggested titles.

A search will retrieve a title list from which individual titles can be "clicked" and down-loaded (click on the file name next to "link.")

- 1. When clicking on the file link, you will be prompted to open or save the file.**
- 2. Select SAVE and an appropriate location (e.g. floppy disk or hard drive). All files are compressed.**
- 3. Once the file is downloaded, double click on it to "explode" the compressed files and directories (the files will be extracted within the directory where the compressed file resides).**

Every effort has been made to maintain the original file/directory structure of the diskettes. Please let us know if you have any [problems or suggestions](#).

[Click here to display titles with full bibliographic detail](#)

Last updated: Sat Apr 22

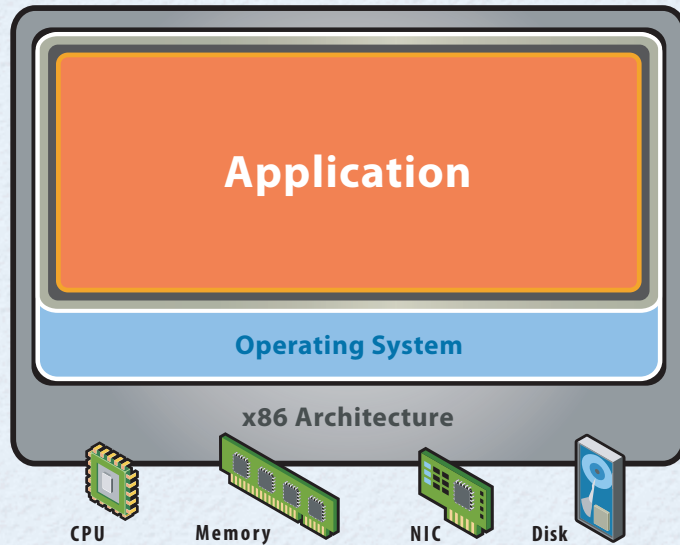
Questions and comments: lbgpd@indiana.edu

Copyright 2003, The Trustees of [Indiana University](#)

ISSUES WE ARE TRYING TO SOLVE

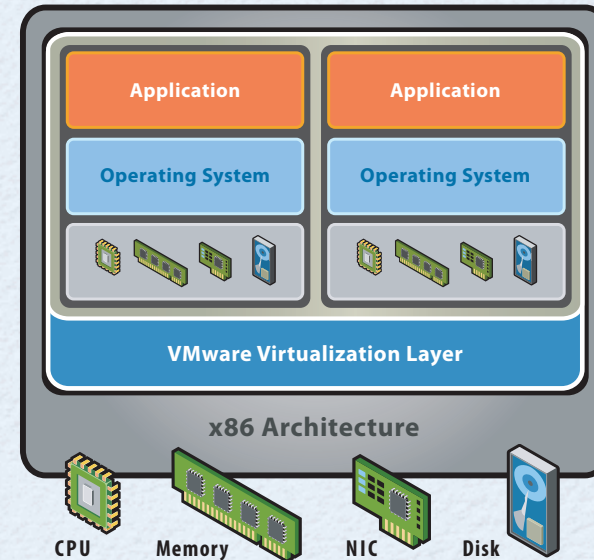
- Documents in FDP (for example) require obsolete applications and operating systems
- Installing documents to access them requires specialized expertise
- These problems generalize to SUDOC documents on CD-ROM

VIRTUALIZATION



Before Virtualization:

- Single OS image per machine
- Software and hardware tightly coupled
- Running multiple applications on same machine often creates conflict
- Underutilized resources
- Inflexible and costly infrastructure



After Virtualization:

- Hardware-independence of operating system and applications
- Virtual machines can be provisioned to any system
- Can manage OS and application as a single unit by encapsulating them into virtual machines

Mozilla Firefox

File Edit View Go Bookmarks Tools Help

file:///D:/bytitle/bytitle.html

Getting Started Latest Headlines

- [National ambulatory medical care survey.](#)
- [National biennial RCRA hazardous waste report.](#)
- [The National health and nutrition examination surveys: a selective bibliography, 1980-93/96.](#)
- [Natural gas annual.](#)
- [NED/SIPS: a stand inventory processor and simulator program for forests of the Northeastern U States.](#)
- [NEWILD Users Manual.](#)
- [NIH/DRG/ISB.](#)
- [Non-residential buildings energy consumption survey: NBECS.](#)
- [Oak Ridge uranium market model: ORUMM.](#)
- [Oil Market Simulation Model\(OMS\).](#)
- [PCAEO 90.](#)
- [PC-AEO Forecasting Model \(Annual Energy Outlook\).](#)
- [Performance profiles of major energy producers \(selected tables\).](#)
- [A Pilot Standard National Course Classification System for Secondary Education.](#)
- [President Clinton's economic plan : and additional budgetary information.](#)
- [Public library data/National center for education statistics.](#)
- [Residential energy consumption survey.](#)
- [The role of leadership in sustaining school reform : voices from the field.](#)
- [Staff data handbook : elementary, secondary, and early childhood education.](#)
- [State energy data system census: SEDS.](#)
- [State energy price & expenditure data system : SEPEDS.](#)
- [State library agencies data FY 1996.](#)
- [Status of open recommendations.](#)
- [STF 3A, updated software.](#)
- [Student data handbook : elementary/secondary and early childhood education.](#)
- [Sugar and sweetener situation and outlook yearbook.](#)
- [Suspended-sediment budget for the Kankakee River Basin, 1993-95.](#)
- [Taxes & You.](#)
- [Toxic chemical release inventory reporting year.](#)
- [TransVU.](#)

Done

- My Computer
- My Documents
- Internet Explorer
- Network Neighborhood
- Recycle Bin
- Shared
- Inbox

PC-AEO forecasting model :annual energy outlook - Mozilla Firefox

File Edit View Go Bookmarks Tools Help

file:///D:/bytitle/Library_10.html

Getting Started Latest Headlines

Title: PC-AEO forecasting model :annual energy outlook

Link: [Library 10.zip](#)

SuDoc classification number: E 3.1/5:989/floppy 1-3

Published: Washington, D.C. : Dept. of Energy, Energy Information Administration; Springfield, VA : Available from National Technical Information Service, Office of Data Base Services, 1989

Summary: Contains forecasts of national energy trends by fuel, region, and sector, by representing the most significant characteristics of U.S. energy markets. Annual forecasts to the year 2000 derived from the model are published in EIA's Annual energy outlook.

Description: 3 computer disks; 5 1/4 in. + user's manual

Version: 89C

Operating requirements: System requirements: IBM or compatible personal computer with 286 or 386 processor, 1M RAM; at least 640K conventional and 768K expanded memory; DOS 3.1 or higher; Lotus 1-2-3 version 2.01 or higher.

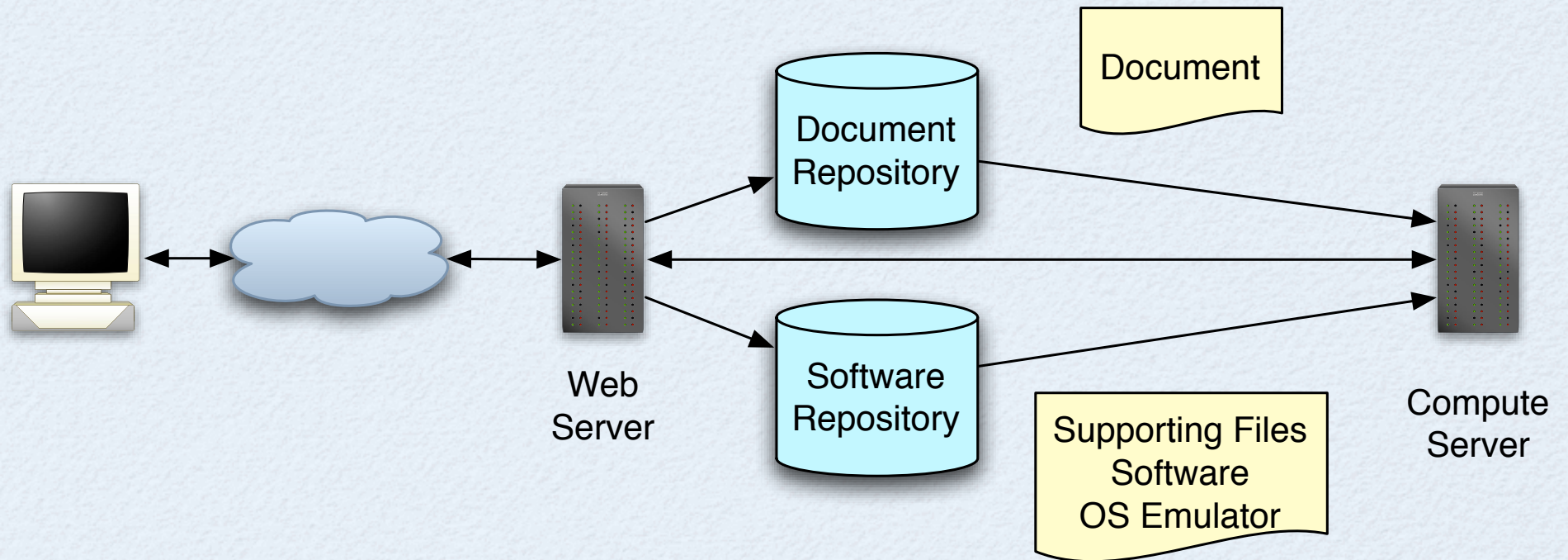
Notes: Spreadsheets. Title from disk label. Shipping list no.: 91-017-E. "September 1989"--User's manual. "DOE/ELA-M040(89)"--User's manual.

Subject headings: Power resources--United States--Forecasting--Databases. Energy consumption--United States--Forecasting--Databases.

GPO item number: 429-J-3

Done

MODEL ARCHITECTURE



ACTIONS IN RESPONSE TO PATRON REQUEST

- A pre-configured emulator is allocated
- Emulator is customized
 - Document file system mounted
 - Document specific installation executed
 - Shared file directories created for patron use
- Link to emulator and web accessible file system provided through patron browser
- Emulator executes remotely under patron control

SOFTWARE REQUIRED FOR FDP

- Windows 98 (most disks were for msdos, win 3.1)
- DBase III -- (dbfviewer2000)
- WordPerfect
- Lotus 1-2-3 (smartsuite, smartsuite viewer)
- Microsoft Word (we use free msword viewer)
- Various Archive Tools
- Browser (we use Firefox)
- Generic Postscript Printer Driver
- Software we didn't install -- pascal, fortran, sas, arcinfo

SUDOC VIRTUALIZATION PROJECT

- Approximately 2500 SUDOC CD-ROMs in IU Library
- We've built an online database of roughly 1800 CD-ROMs (about 1500 items)
- Analyzed software requirements for about 1000



SUDOC Virtualization Project

This site provides access to an online database containing government documents originally published on CDROM. There are two interfaces, a [Database](#) interface that enables browsing the entire collection and a [Search](#) interface that allows searching for specific documents using [CQL](#) queries.

Through either interface it is possible to access the original CDROM contents exported as [WebDAV](#). These WebDAV directories can be directly "mounted" on some operating systems (OS X, Linux). In Windows they can be accessed through the network neighborhood or mounted using Novell Netdrive. WebDAV access does require a password. Please contact Geoffrey Brown for access.

On the Indiana University Bloomington campus, the CDROMs can be imported using windows file sharing. Interested users may contact Geoffrey Brown for details.

Search Site

Go

Home Page

Contacts

Database

Search

Help

RSS

Search

Query:

Submit

8 Records:

1. [The coastal use issues course](#)
A 1.142: C 63/CD
30000057033734
2. [Coastal Zone 01](#)
Cleveland Ohio July 15-19 2001.
C 55.54: C 58/CD
30000078791138
3. [C-CAP Coastal Change Analysis Program](#)
changes in land cover in the Yakutat Bay Alaska Region 1986-1993.
C 55.54: C 63/2
30000050360068
4. [Coastal GeoTools](#)
proceedings of the ... biennial Coastal GeoTools Conference.
C 55.54: C 63/5/CD 2001
30000053688416
5. [Alabama coastal hazarde assessment](#)

Search Site [Home Page](#)[Contacts](#)[Database](#)[Search](#)[Help](#)[RSS](#)

Database

<< < > >> 47

Page: 47 of 112 Records: 1672

<input type="checkbox"/>	<u>Title</u>	<u>Call Number</u>	<u>Barcode Number</u>
<input type="checkbox"/>			
X	Sorted By: Sort ascending		
<input type="radio"/>	World ocean atlas 1998 figures h[computer file] / cprepared by the Ocean Climate Laboratory National Oceanographic Data Center.	C 55.297: OC 2/3/DISC.4 FIGURES	30000065731923
<input type="radio"/>	Biological Atlas of the Arctic Seas 2000 h[computer file] : bplankton of the Barents and Kara seas / cMurmansk Marine Biological Institute ; Ocean Climate Laboratory National Oceanographic Data Center/NOAA.	C 55.297: OC 2/V.2/CD	30000078802729
<input type="radio"/>	Hydrochemical atlas of the Sea of Okhotsk 2001 h[electronic resource] / cLaboratory of the Marine Ecology Russian Federal Research Institute of Fisheries and Oceanography ; World Data Center for Oceanography Ocean Climate Laboratory National Oceanograp	C 55.297: OC 2/V.3/CD	30000078802737
<input type="radio"/>	ODINAFRICA h[electronic resource] / cprepared by the Ocean Climate Laboratory National Oceanographic Data Center.	C 55.297: OD 3/CD	30000085793465
<input checked="" type="radio"/>	World ocean atlas 1998 figures h[computer file] / cprepared by the Ocean Climate Laboratory National Oceanographic Data Center.	C 55.297: T 24/2/DISC.1-3	30000067278824
<input type="radio"/>	International Ocean Atlas and Information Series Atlas of	C 55.297: T	

Indiana University Department of

ComputerScience

Indiana University School of
informatics
[IU Home](#) | [IUB Home](#) | [Informatics home](#)

Iso [Recent Changes](#) | [Login](#)

-
- [Home Page](#)
 - [Contacts](#)
 - [Database](#)
 - [Search](#)
 - [Help](#)

[RSS](#)

Iso

Title

World ocean atlas 1998 figures
 [computer file] /
 prepared by the Ocean Climate Laboratory National Oceanographic Data Center.

Call number

C 55.297: T 24/2/DISC.1-3

System Requirements

System requirements Microsoft: any DOS Extension that supports CD-ROM player with ISO 9660 software driver
 hard disc drive
 and VGA color monitor.

Installation Notes

No install necessary. Data is accessible through the html file: open_1st.html.

Browse CD Images

[VMware](#) http://www.cs.indiana.edu/sudoc/image_30000067278824/

Options

Clipboard

Send Ctrl-Alt-Del

Refresh

Microphone Tools ?

Search

Bookmarks

TightVNC desktop

entry School of
matics
 e|Informatics home

[Recent Changes](#) | [Login](#)
<http://spitz/cgi-bin/console.cgi?iso=30000067278824>
30000067278824

- Disk 0
- Disk 1
- Disk 2

Index of /

Name	Last Modified	Size
------	---------------	------

[Parent Directory](#)[AGR4ALL.PDF](#)

05-Oct-95 06:35

4.7M

Upload

Browse...

http://www.cs.indiana.edu/sudoc/image_30000067278824/

Location

[News](#) [eBay](#) [bb](#) [nytimes](#) [slashdot](#) [IUB Libraries](#) [Google Reader](#) [sudoc/](#) [B649](#) [START V2 Links](#) [Tigh](#)

Index of /sudoc/image_30000067278824

<u>Name</u>	<u>Last modified</u>	<u>Size</u>	<u>Description</u>
-------------	----------------------	-------------	--------------------

[Parent Directory](#)

-

[30000067278824_1/](#)

10-Jun-1999 02:37

-

[30000067278824_2/](#)

31-Dec-1969 19:00

-




[30000067278824_3/](#)

31-Dec-1969 19:00

-

Apache/2.0.52 (Red Hat) Server at troy.cs.indiana.edu Port 80

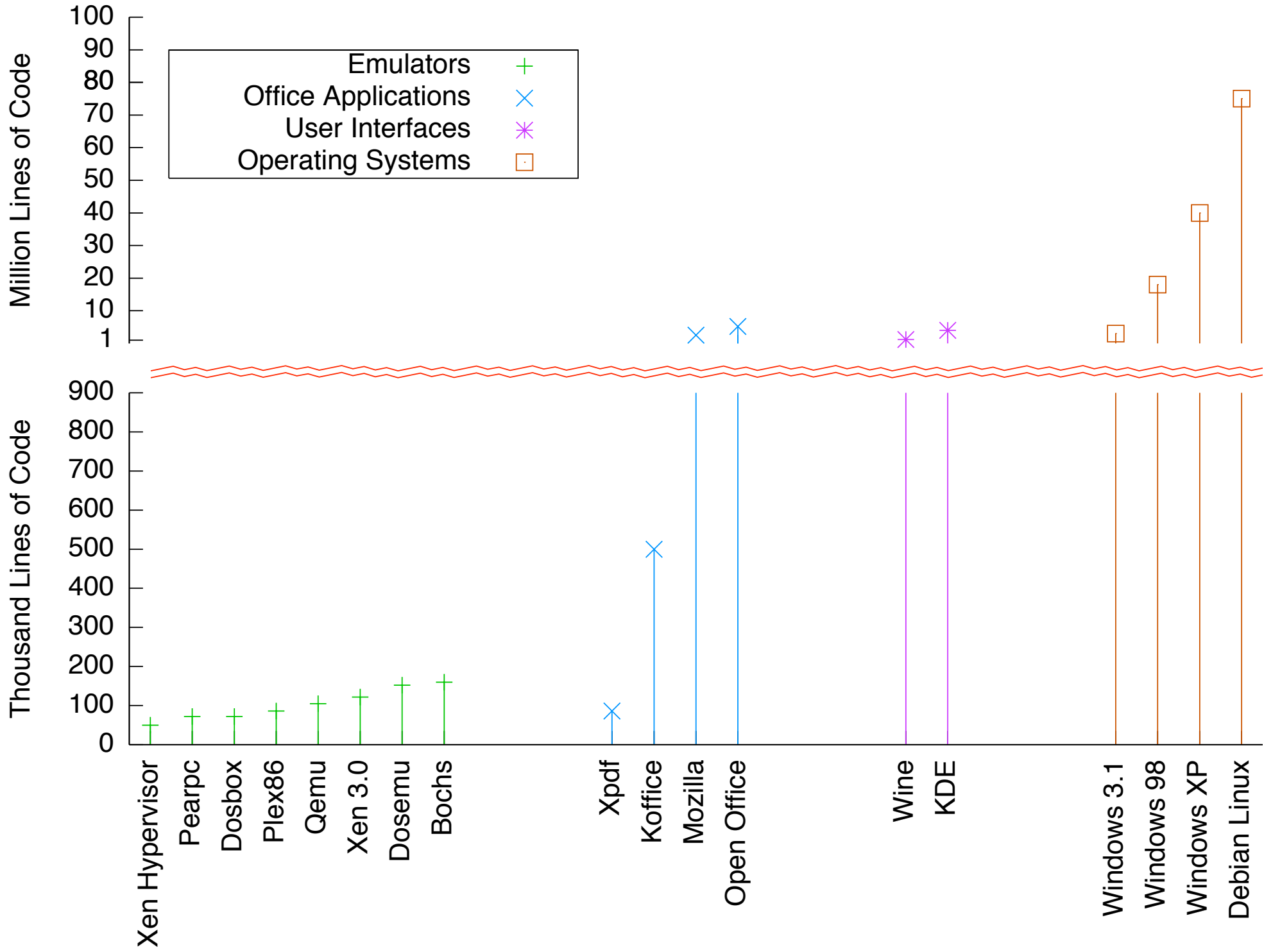
Index of /sudoc/image_30000067278824/30000067278824_1

<u>Name</u>	<u>Last modified</u>	<u>Size</u>	<u>Description</u>
 Parent Directory		-	
 open_1st.htm	01-Jun-1999 09:16	1.5K	
 woaf_cd/	10-Jun-1999 02:37	-	

Apache/2.0.52 (Red Hat) Server at troy.cs.indiana.edu Port 80

PHILOSOPHICAL ISSUES

- Who Preserves the Emulator ?
- Why Not Just Migration ?



WHY NOT JUST MIGRATION

- Loss of information -- e.g. word edits
- Loss of fidelity -- e.g. WordPerfect to Word isn't very good
- Loss of authenticity -- users of migrated document need access to original to verify authenticity
- Not always possible -- closed proprietary formats
- Not always feasible -- costs may be too high
- Emulation may necessary to enable migration

BBC
NEWS
 **OPEN** **BBC News in video and audio**
News services
 Your news when you want it

[News Front Page](#)
[World](#)

[Africa](#)
[Americas](#)
[Asia-Pacific](#)
[Europe](#)
[Middle East](#)
[South Asia](#)
[UK](#)
[England](#)
[Northern Ireland](#)
[Scotland](#)
[Wales](#)
[Business](#)
[Politics](#)
[Health](#)
[Education](#)
[Science/Nature](#)
[Technology](#)
[Entertainment](#)
[Have Your Say](#)
[Magazine](#)
[In Pictures](#)

Last Updated: Monday, 2 May, 2005, 17:18 GMT 18:18 UK

 [E-mail this to a friend](#)
 [Printable version](#)

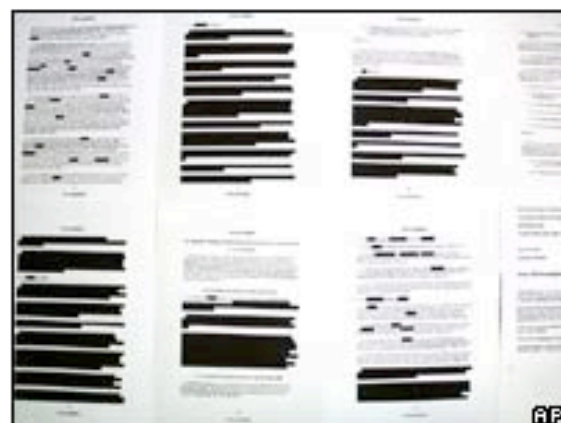
Readers 'declassify' US document

When news started circulating in Italy that a heavily censored Pentagon report into the death of secret agent Nicola Calipari had been decrypted, many thought it must be the work of some top-notch hacker.

In fact, it turned out that the classified document, containing top-secret details - such as the name of the soldier who fired the deadly rounds of ammunition - could be made readable with two simple clicks of your computer mouse.

A few hours after the Pentagon published the report on its website, a few Italian readers found they could make the blacked-out paragraphs reappear by cutting and pasting them from the site into a Word document.

Salvatore Schifani, a 30-year-old IT worker, spotted the document at about 0300 local time (0100 GMT) on Saturday night.



Someone found a simple cut-and-paste job could restore the text

BBC NEWS:VIDEO AND AUDIO
How the censored parts of the report were made public
 **VIDEO**
THE STRUGGLE FOR IRAQ
KEY STORIES

- ▶ [New Iraqi plan to curb violence](#)
- ▶ [Baghdad authorities lift curfew](#)
- ▶ [Baghdad mosque attack kills 10](#)
- ▶ [Iraq war fuels terror - US report](#)
- ▶ [Saddam thrown out for third time](#)

BACKGROUND AND ANALYSIS

Tide of violence

Complete US victory in Iraq may be out of the question, argues John Simpson.

- ▶ [Saddam Hussein on trial](#)
- ▶ [Iraq's federal fracas](#)
- ▶ [War justifications laid bare](#)
- ▶ [War dead figures](#)
- ▶ [Saddam trial one timeline](#)
- ▶ [Saddam trial two timeline](#)

Digital Media

When Words Come Back From The Dead

David M. Ewalt, 12.13.05, 5:00 PM ET

By This Author



David M. Ewalt

- The Hartford CIO's Wireless LCD Television
- Trust

NEW YORK -In the year since pharmaceutical giant **Merck** withdrew its arthritis drug Vioxx from shelves, the company has been hit with 7,000 personal injury lawsuits—one of which already cost the company \$253 million. But it could be a frequently misused feature of **Microsoft Word** that turns out to be the straw that broke Merck's back.

GOALS

- Deliver key SUDOC collections through virtualization
 - Develop web delivery techniques
 - Improve our software analysis tools
 - Develop image customization techniques (e.g. perform software install on the fly)

HOW THIS WORK MIGHT BE USED

- Libraries share pool of software images and licenses
- Libraries share expertise in supporting various document collections
- Libraries collaborate to provide redundancy
- Patrons access from anywhere without needing to obtain or install special software

ACKNOWLEDGMENTS

- Lou Malcomb (IU Head GIMSS)
- Julianne Bobay (IU Head SLIS Library)

