

The Ohio State University's Institutional Repository How Policy Informs Practice

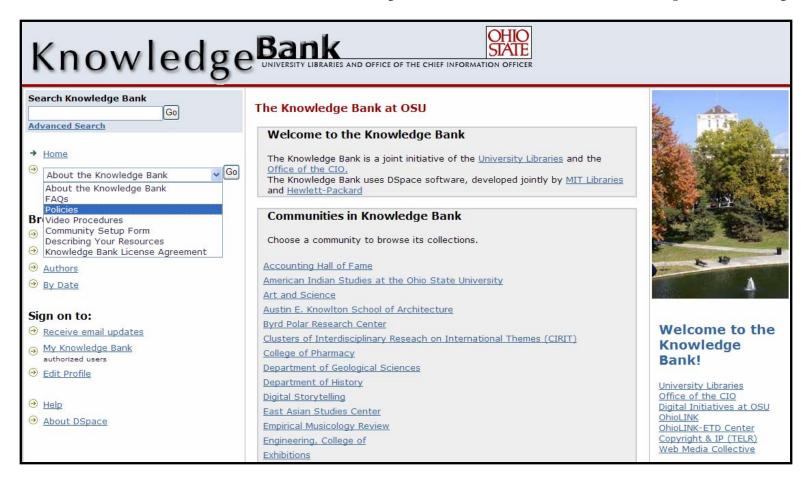
Maureen P. Walsh & Tschera Harkness Connell, OSU Librarie

OhioDIG 3rd Quarter Meeting

September 14, 2007



The Ohio State University's Institutional Repository





Mission

The Ohio State University's Institutional Repository: How Policy Informs Practice

"The mission of the institutional repository of The Ohio State University Knowledge Bank, is to collect, preserve, and share the intellectual output of the faculty, staff, and students that is in a digital format."

Source: http://library.osu.edu/sites/kbinfo/policies.html retrieved 09-10-2007

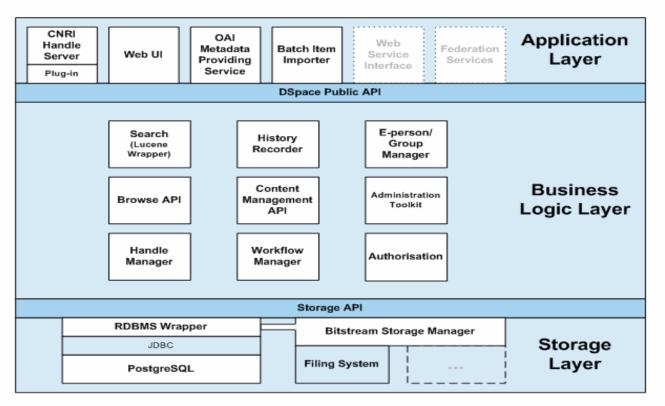


Software

The Ohio State University's Institutional Repository: How Policy Informs Practice

DSpace Captures, Indexes, Preserves, and Distributes

Architecture Diagram



Source: http://libraries.mit.edu/dspace-mit/technology/architecture-diagram.html retrieved 09-10-2007

September 14, 2007

OhioDIG 3rd Quarter Meeting



Knowledge Bank Communities





Policies



Knowledge Bank Policies

KB Community & Collection Policies

KB File Format Support



KB Community & Collection Policies

A Knowledge Bank Community has the right to:

- decide policy regarding content to be submitted
- decide who may submit content
- limit access to content
- customize interfaces to community content



KB Community & Collection Policies

The Knowledge Bank agrees to:

- retain, maintain, and distribute content according to community decisions
- preserve content using accepted preservation techniques
- comply with University policy on issues which affect the KB



KB Community & Collection Policies

The Knowledge Bank retains the right to:

- redistribute or amend metadata for items in the KB
- remove from view items or collections according to KB Withdrawal Policy
- migrate items if format is in danger of obsolescence
- establish quotas and determine service fee structures



Knowledge Bank File Format Support

- We will preserve all bitstreams as-is and allow them to be retrieved
- We will recognize as many file formats as possible
- We will support as many known file formats as possible



Levels of Digital Preservation

- **Bit preservation**: ensures that a file remains exactly the same over time
- Functional preservation: the file does change over time so that the material continues to be immediately usable in the same way it was originally



Levels of Format Preservation

- Supported formats will be functionally preserved using either format migration or emulation techniques
- Known formats are those that we can't promise to preserve, such as proprietary or binary formats, but for which migration tools will likely emerge
- Unsupported formats are those that we don't know enough about to do any sort of functional preservation



Format Preservation

- For all three levels (supported, known, unsupported)
 DSpace does bit-level preservation
- The Knowledge Bank attempts to keep the percentage of supported format material as high as possible
- The Knowledge Bank continually monitors formats and techniques to ensure we can accommodate needs as they arise



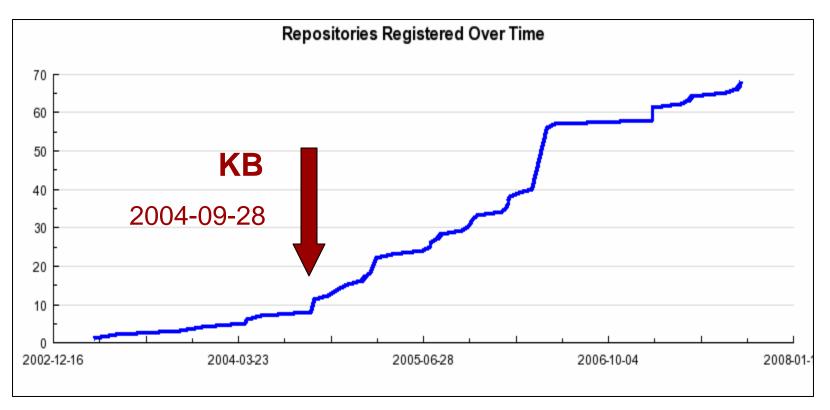
Where we are



38 communities 27,319 records



DSpace Repositories in US



Source: ROAR (Registry of Open Access Repositories) http://roar.eprints.org retrieved 09-10-2007

September 14, 2007

OhioDIG 3rd Quarter Meeting



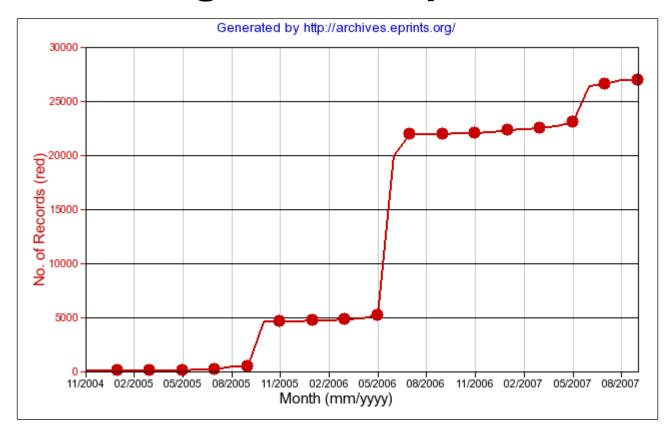
KB Repository Ranking by Deposits

- 2nd of 68 DSpace Repositories in the U.S.
- 7th of 253 DSpace Repositories
- 8th of 219 U.S. Repositories on any System Software
- 38th of 925 Repositories on any System Software
 - University of Michigan: Deep Blue (39,659 records)
 - Ohio State University: Knowledge Bank (26,949 records)
 - DSpace at MIT (25,670 records)
 - Georgia Tech's IR: SMARTech (18,375 records)
 - eCommons@Cornell (6990 records)

Source: ROAR (Registry of Open Access Repositories) http://roar.eprints.org retrieved 09-10-2007



Knowledge Bank Deposits



Source: ROAR (Registry of Open Access Repositories) http://roar.eprints.org retrieved 09-10-2007

September 14, 2007

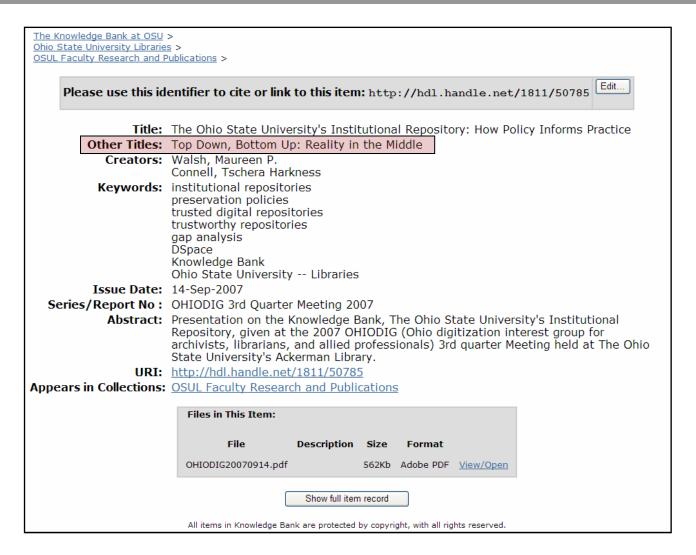
OhioDIG 3rd Quarter Meeting



How are we doing?



The Ohio State University's Institutional Repository: How Policy Informs Practice





Trustworthy Repository?



Developing Standards & Research

- Reference Model for an Open Archival Information System (OAIS)
- Trusted Digital Repositories: Attributes and Responsibilities
- Trustworthy Repositories Audit & Certification:
 Criteria and Checklist
- Digital Archive Policies and Trusted Digital Repositories



Where are the gaps?



Trustworthy Repositories Checklist

- Organizational Infrastructure
- Digital Object Management
- Technologies, Technical Infrastructure, and Security

Please note: Selected concepts from the Checklist are used on the following slides to illustrate our process of gap analysis.



Organizational Infrastructure

Governance & Organizational Viability

- Has mission statement
- Has formal succession plan

Contracts, Licenses & Liabilities

- Maintains agreements
- Tracks & manages intellectual property rights



Organizational Infrastructure

Procedural Accountability & Policy Framework

- Has publicly accessible definitions and policies in place to dictate how preservation requirements will be met
- Has mechanisms for review of policies & procedures
- Has documented history of changes



Digital Object Management

Ingest: Acquisition of Content

- Identifies properties Repository will preserve
- Specifies information that needs to be associated with digital material



Digital Object Management

Ingest: Creation of the Archival Package

- Has written definition for each class of information
- Uses persistent naming convention
- Acquires preservation metadata for associated Content Information

The Ohio State University's Institutional Repository: How Policy Informs Practice

Preservation Metadata

Element	Qualifier	Value			Language		
date	accessioned	2007-09-11T04:31:50Z		<u>^</u>		Remove	
date	available	2007-09-11T04:31:50Z				Remove	
description		Submitted by Maureen Walsh (walsh.260@osu.edu) on 2007-09-11T04:31:49Z No. of bitstreams: 1		en	Remove		
description provenance		OHIODIG20070914.pdf: 575696 bytes, checksum: 04c2b2f1a2a75d161403d167f5f1fbb8 (MD5) Previous issue date: 2007-09-14			en	Remove	
format	extent	575696 bytes				Remove	
format	mimetype	application/pdf				Remove	
Primary Bitstream	Name	Source	Description	Format	User For	mat Description	n
	OHIODIG20070914.pdf.txt	Written by MediaFilter org	Extracted text	(Text)			View Remove
	license.txt	Written by org.dspace.cor		(License)			View Remove
0	OHIODIG20070914.pdf	OHIODIG20070914.pdf		(Adobe PDF	=)		View Remove



Digital Object Management

Information Management

- Articulates minimum metadata requirements
- Captures or creates minimum descriptive metadata and ensures that it is associated with archived object

Access Management

- Documents & communicates to communities what access & delivery options are available
- Demonstrates that all access requests result in a response of acceptance or rejection



Digital Object Management

Preservation Planning

- Has documented preservation strategies
- Has mechanisms for monitoring when Representation Information approaches obsolescence

Archival Storage & Preservation/Maintenance of AIPs

- Employs documented preservation strategies
- Preserves Content Information of archival objects
- Repository monitors integrity of archival objects



"Employs documented preservation strategies"

Formats?

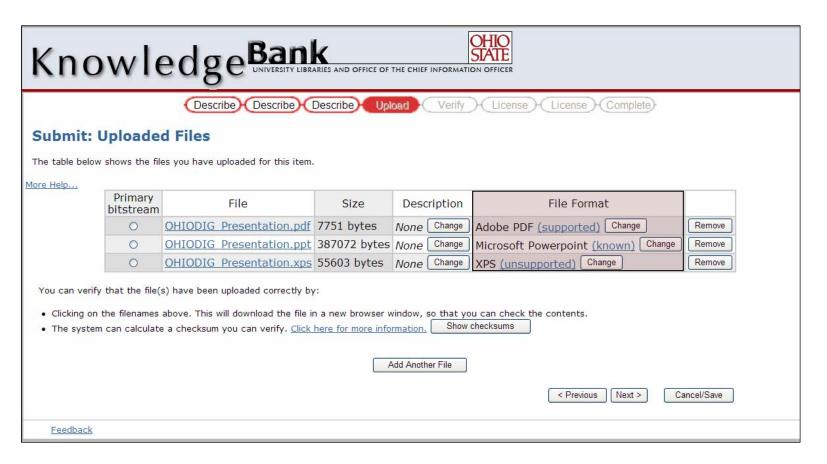


KB Bitstream Format Registry

MIME Type	Description	Extension	Level	MIME Type	Description	Extension	Level
application/marc	MARC	marc, mrc	supported	audio/x-aiff	AIFF	aiff, aif, aifc	supported
application/mathematica application/msword application/octet-stream application/ogg	Mathematica Microsoft Word Unknown Ogg Media Type	ma doc (anything not liste ogg, OggS	known known d)unsupported known	audi o/basic audi o/x-mpeg audi o/x-pn-real audio audi o/x-wav	audio/basic MPEG Audio RealAudio WAV	au, snd mpa, abs, mpeg ra, ram way	known known known known
application/pdf application/postscript	Adobe PDF Postscript	pdf ps, eps, ai	supported supported	image/gif image/jpeg	GIF JPEG	gif jpeg, jpg	supported supported
application/sgml application/vnd.ms-excel	SGML Microsoft Excel	sgm, sgml xls	known known	image/png image/tiff	PNG TIFF	png tiff, tif	supported supported
N ■ ■ 10.00 mile (1996 of 1990) 1990 (■ 19 (1990) ■ 19 (1990) 1990 (1990) 29 (1990) 1990 (1			known	image/x-ms-bmp	BMP	bmp	known
application/vnd.ms-project	Microsoft Project	mpp, mpx, mpd	known	image/x-photo-cd	Photo CD	pcd	known
application/vnd.visio	Microsoft Visio	vsd	known	text/html	HTML	html, htm	supported
application/wordperfect5.1	WordPerfect	wpd	known	text/plain	Text	txt	supported
application/x-dvi	TeXdvi	dvi	known	text/richtext	Rich Text Format	rtf	supported
application/x-filemaker	FMP3	fm	known	text/xm1	XML	xml	supported
application/x-latex	LateX	latex	known	video/mpeg	MPEG	mpeg, mpg, mpe	known
application/x-photoshop	Photoshop	psd, pdd	known	video/quicktime	Video Quicktime	mov, qt	known
application/x-tex	TeX	tex	known	FIREDOMORROUS TROPINS NO. 10 LODGE			

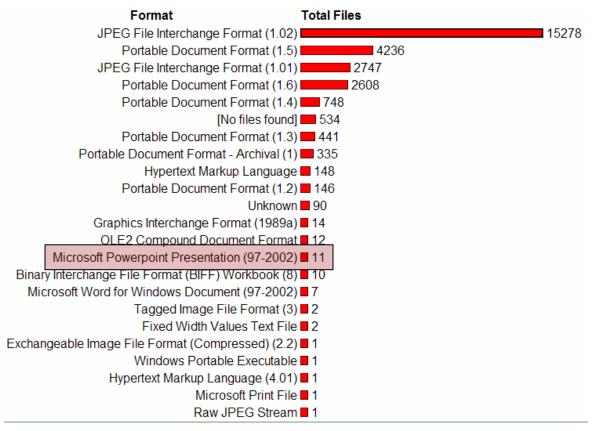


Preservation Support Level Advice at Ingest





File Format Profile for Stored Objects*



^{*}Excludes files not hosted on the same server as the OAI interface and files over 2 megabytes in size. Source: ROAR (Registry of Open Access Repositories) http://roar.eprints.org retrieved 09-10-2007



Non-Supported Files in the KB – MS PPT

Record Identifier	URL	Date
oai:kb.osu.edu:1811/150	/dspace/bitstream/1811/150/2/Tull_wilsworld.ppt	2006-05-14 16:53:16
oai:kb.osu.edu:1811/183	/dspace/bitstream/1811/183/1/Potterylecture.ppt	2006-05-14 16:53:17
oai:kb.osu.edu:1811/202	/bitstream/1811/202/1/Learning+Group+report.ppt	2006-05-14 16:53:17
oai:kb.osu.edu:1811/270	/dspace/bitstream/1811/270/5/ALAO_workshop.ppt	2006-05-14 16:53:2
	/dspace/bitstream/1811/270/2/LO_activity.pdf	2006-05-14 16:53:22
	Portable Document Format (1.4)	
	/dspace/bitstream/1811/270/3/LO_resources.pdf	2006-05-14 16:53:22
	Portable Document Format (1.4)	
	/dspace/bitstream/1811/270/4/LO_EvalTool.pdf	2006-05-14 16:53:22
	Portable Document Format (1.4)	
oai:kb.osu.edu:1811/682	/dspace/bitstream/1811/682/1/ALAO-TSIG20050304.ppt	2006-05-14 16:53:42
oai:kb.osu.edu:1811/683	/dspace/bitstream/1811/683/1/ALAO+Presentation.ppt	2006-05-14 16:53:42
oai:kb.osu.edu:1811/6024	ace/bitstream/1811/6024/4/LewisOSU-20060316.ppt	2006-05-14 16:57:4
	/dspace/bitstream/1811/6024/5/Lewis-kbdirect.html	2006-05-14 16:57:43
	Hypertext Markup Language	
oai:kb.osu.edu:1811/24	/dspace/bitstream/1811/24/1/lzzo_UD.ppt	2006-05-16 17:00:05
oai:kb.osu.edu:1811/24793	tream/1811/24793/1/CIL07_black_kilzer_final.ppt	2007-04-17 17:59:51
oai:kb.osu.edu:1811/2481	71/Creating+Online+Tutorials+%28w+Chinese%29.ppt	2007-04-25 06:07:1
	ream/1811/24817/2/Creating Online Tutorials.pdf	2007-04-25 06:07:17
	Portable Document Format (1.5)	
oai:kb.osu.edu:1811/2481	8/1/Developing+Online+Course+%28w+Chinese%29.pp	t 2007-04-25 06:09:3
	tream/1811/24818/2/Developing+Online+Course.pdf	2007-04-25 06:09:37
	Portable Document Format (1.5)	

Source: ROAR (Registry of Open Access Repositories) http://roar.eprints.org retrieved 09-10-2007

September 14, 2007

OhioDIG 3rd Quarter Meeting



"Information Management"

Quality Metadata?



Community Metadata Application Profile

ELEMENT NAME	DEFINITION	EXAMPLE
Title	Title of book	Samuel Beckett: humanistic perspectives
Creator(s)	Author(s) of book [Use form on catalog record]	
		Beja, Morris
Creator(s)	Author(s) of book	Gontarski, S. E.
Creator(s)	Author(s) of book	Astier, Pierre A. G.
Contributor(s)	Other contributors e.g., illustrator	
Date of Issue	Publication year	1983
Publisher	Publisher	Ohio State University Press
Identifiers.Other	ISBN	0814203345
Identifiers.Other	OCLC Number [Precede number by: OCLC#]	OCLC #8589285
Identifiers.Other	LCCN [precede number by LCCN]	LCCN 82-12468
Description	Extent of print item.	x, 217 p., [2] folded leaves of plates : ill. ; 24 cm.
Language	Language(s) of the intellectual content of the dig	English
Туре	Work type	book
Abstract	Abstract, if available.	
Description	Description. For example, TOC.	
Series	Series title	Victorian Critical Interventions
Subjects. Keywords	Keywords to enhance subject headings, if obvious	
Subjects.LCSH	LCSH subject heading	Beckett, Samuel, 1906-1989 Criticism and interpretation



The Knowledge Bank at OSU

Browse by Author

Showing authors 1-21 of 23103.

Showing authors 20071-20091 of 23103.

A'Honro M.E		
A'Hearn, M.F.		
Aaron, Daniel, 1912-		
Abbas, M. M.		
Abbate, M. J.		
ABBATE, S.		
Abbate, S.		
ABBATE, SERGIO		
Abbate, Sergio		
Abbott, Maxine L.		
Abbott, Ralph E.		
Abbott, Richard H.		
ABBOUTI, TEMSAMANI M.		
Abbouti, Temsamani M.		
ABDEL, SHAFY HUSSEIN I.		
Abdel-Halim, F.M.		
ABDEL-HALIM, H.M.		
Abdel-Malek, Albert		
Abdullah, Mohammad		
ABDULNUR, SUHEIL F.		
Abe, Hisashi		

Sztainbuch, I.		
Szudy, J.		
Szymanski, H.		
T-S LIN		
<u>T., Oka</u>		
<u>T.SHIMIZU</u>		
Taatjes, C.A.		
Tabisz, G. C.		
TABISZ, G.C.		
Tabisz, G.C.		
TABOR, WILLIAM J.		
Tack, L. M		
Tack, L. M.		
TACK, L.M.		
Tack, Lesley M.		
Tack, Leslie M.		
Tackett, Brandon S.		
Tadao, Shimizu,		
Taday, P. F.		
Taft, Clarence E.		
TAFT, ROBERT W. JR.		

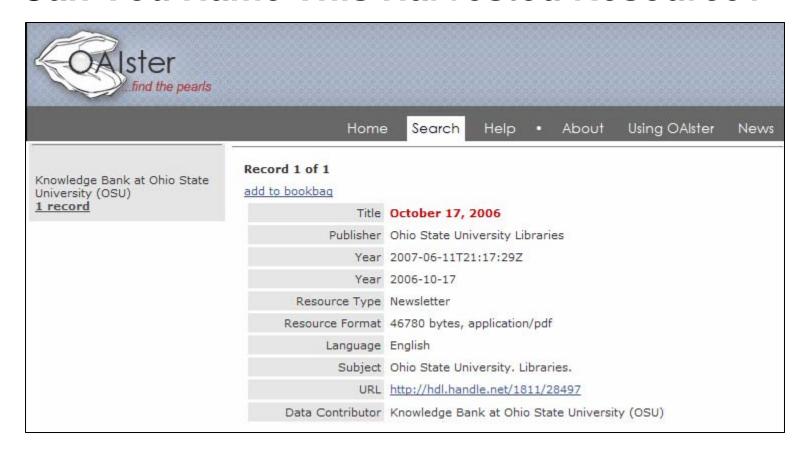


"Information Management"

Shareable Metadata?



Can You Name This Harvested Resource?





The Knowledge Bank at OSU > Ohio State University Libraries > NewsNotes > QuickNotes > Please use this identifier to cite or link to this item: http://hdl.handle.net/1811/28497 Title: October 17, 2006 **Keywords:** Ohio State University. Libraries. **Issue Date:** 17-Oct-2006 **Publisher:** Ohio State University Libraries **URI:** http://hdl.handle.net/1811/28497 **Appears in Collections: QuickNotes** Files in This Item: File **Description Size** Format qn061017.pdf 45Kb Adobe PDF View/Open Show full item record



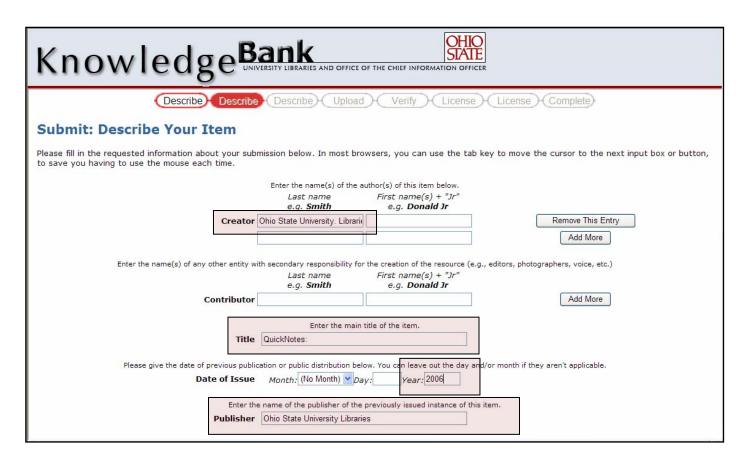
Collection Metadata Application Profile

KB QuickNotes CORE ELEMENT SET

ELEMENT NAME	DEFINITION	EXAMPLE
Title	Title of work	QuickNotes: October 11, 2006
Creator(s)		Ohio State University. Libraries
Contributor(s)	Author(s) of featured content	
Date of Issue	Issue date of work	October 11, 2006
Publisher	Publisher	Ohio State University Libraries
Citation		
Identifier		
Language	Principal language(s) of the work	en_US
Туре	Format(s) of the work	Newsletter
Abstract		
Description		
Subject	Topic of work	Ohio State University Libraries Newsletters



Customized Submission Template





The Knowledge Bank at OSU > Ohio State University Libraries > NewsNotes > QuickNotes > Please use this identifier to cite or link to this item: http://hdl.handle.net/1811/28497 Title: QuickNotes: October 17, 2006 **Creators:** Ohio State University. Libraries **Keywords:** Ohio State University -- Libraries -- Newsletters Issue Date: 17-Oct-2006 **Publisher:** Ohio State University Libraries **URI:** http://hdl.handle.net/1811/28497 Appears in Collections: QuickNotes Files in This Item: File **Description Size** Format qn061017.pdf 45Kb Adobe PDF View/Open Show full item record

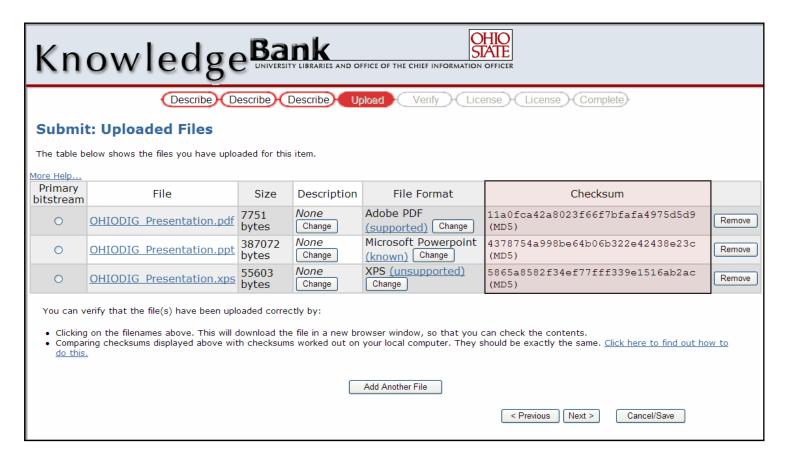


Technologies & Technical Infrastructure

System Infrastructure

- Functions on well supported systems ...
- Ensures adequate hardware & software support
- Has effective mechanisms to detect bit corruption or loss
- Has process for testing effect of critical changes to the system

Checksum Display





Technologies & Technical Infrastructure

Appropriate Technologies

Has hardware and software technologies appropriate to services

Security

- Maintains systematic analysis of data, systems, personnel, physical plant & security needs
- Has suitable written disaster preparedness & recovery plan(s)



What we are working on



Current KB Initiatives

- Automated metadata capture and creation
- Checksum checker
- Format verification tool
- Master file preservation metadata (Dark archives)
- PREMIS in METS
- Shareable metadata
- Name normalization
- Community submission forms and templates
- Metadata application profiles
- Metadata workflows and quality control
- Gap analysis for a Trustworthy Repository



Resources

- Consultative Committee on Space Data Systems, Reference Model for an Open Archival Information System (OAIS). 2002 (ISO Standard 14721) www.ccsds.org/documents/650x0b1.pdf
- Trusted Digital Repositories: Attributes and Responsibilities. Mountain View, CA: RLG, May 2002.
 www.rlg.org/en/pdfs/repositories.pdf
- Trustworthy Repositories Audit & Certification: Criteria and Checklist. Version 1.0. Chicago: CRL; Dublin, OH: OCLC, February 2007. http://www.crl.edu/PDF/trac.pdf
- Smith, MacKenzie and Moore, Reagan W., "Digital Archive Policies and Trusted Digital Repositories", *The International Journal of Digital Curation* 1, no. 2, 2007.



Thank You

Maureen P. Walsh walsh.260@osu.edu

Tschera Harkness Connell connell.17@osu.edu