#### Illinois Wesleyan University

#### From the SelectedWorks of Meg Miner

September 20, 2013

### Shining a Light on the Digital Dark Ages

Meg Miner, Illinois Wesleyan University



This work is licensed under a Creative Commons CC\_BY-NC-SA International License.





# Shining a Light on the Digital Dark Ages

Meg Miner
University Archivist and Assistant Professor

Faculty Colloquium September 20, 2013



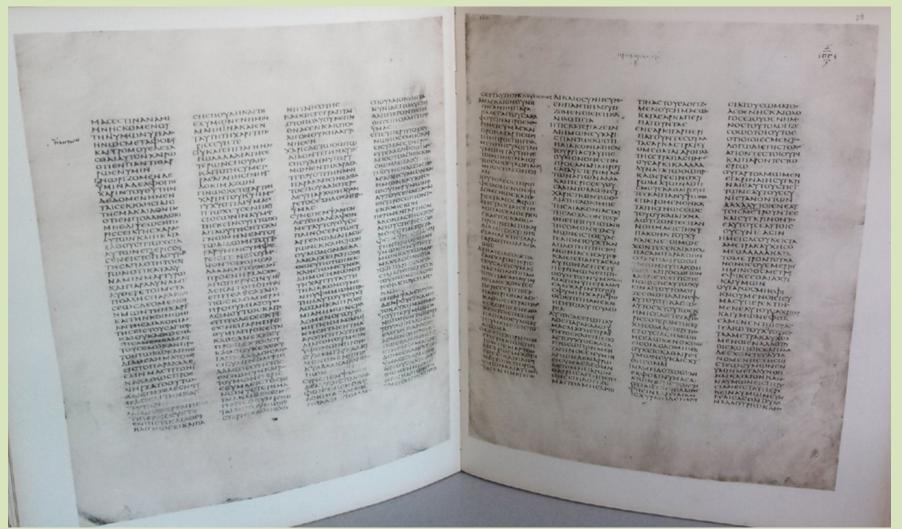
### Why "Dark Ages"?

- "Digital Dark Ages" is a term coined in 1997
- Need to care for humanities' electronic heritage
- Recalls the efforts of monks in the Middle Ages
  - practice of copying texts and dispersing them saved knowledge that predated the Dark Ages

Connotes loss of knowledge or lack of intellectual growth due to a loss of understanding of what has come before.



# Codex Sinaiticus 4<sup>th</sup> century mss. Bible, 20<sup>th</sup> century reprint





### Why now?

- IWU's Digital Dark Ages date from late-1990s
  - Our historical record is fragmented: servers, hard drives, external drives and institutional memory.
     Oh my!
- 2012 request for long-time administrator.
- 2011 request for 1920s student life.

### Concerns for IWU Heritage

- Our Web site is our default archive, but
  - we are not consistent in how we post.
  - our ability to understand change is in doubt.
  - our file types are not stable for the long term.
- Which version of a record is the official IWU word on a subject?
  - If there is a discrepancy, how will it be resolved?
  - Where does the original live? Who cares for it?
     The archives does if found but some slip by.

### Concerns for Heritage (cont'd)

- Content Management Systems
  - Unlimited ability to create and distribute
  - You are our monks...but not all on the same page
- Social Media & Third Party platforms
  - Issues about ownership, ongoing access, file types

### Moving IWU Out of the Dark Ages

Using the Web for promotion of our work is important, necessary and useful but we make assumptions about the safety of our records that are misleading.

We must start making decisions about what we value for our cultural heritage and secure that content in ways that will ensure its survival.



# Shining a Light on the Digital Dark Ages

True Confessions of a Book- and Paper-Loving Librarian

Meg Miner

University Archivist and Assistant Professor

Faculty Colloquium September 20, 2013

### Assumptions in the Digital Age

Storage is cheap and everywhere!

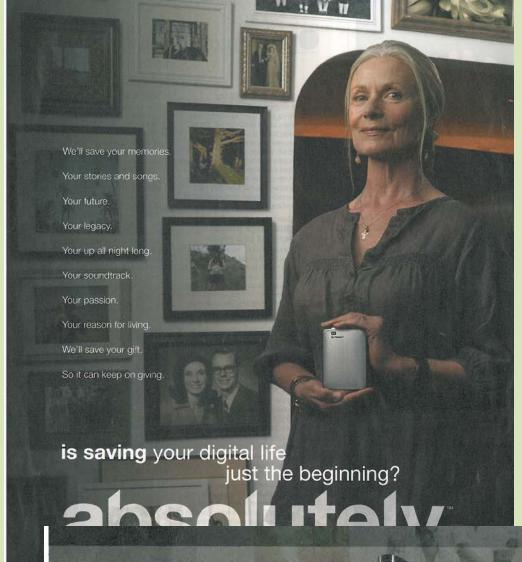
Is the content secure/accessible? How? By whom?

If it is digital it will always be available!

Will it be useable? If software changes or files decay how will we know what is gone before it is too late?

### Storage

Storage options are cheap and everywhere!



Storage ≠ Preservation

Save Digital Storage

Western Digital may be mertio Share

Personal Cloud Storage Stream

HD Routers **Enjoy** Media

Media Players



- Storage options are cheap and everywhere!
  - But human intervention is not (and we are so busy that it is easy to put off)
    - Files need naming, tagging, organizing
       AKA metadata
  - Do your backed up versions all match?
    - Do you spend time synching devices? How much?

Don't forget: all drives fail eventually



"Do you have a back-up plan?" by John K http://www.flickr.com/photos/johnkay/5200871042/

Media needs to be replaced every five years (aka migration)

"Bit Rot" by The Joy of the Mundane



http://www.flickr.com/photos/mundane\_joy/2316921309/

- Storage options are cheap and everywhere!
  - But human intervention is not (and we are so busy that it is easy to put off)
  - Do your backed up versions all match?
  - Remember: you get what you pay for with physical storage and in the Cloud!

Backups on cheap and/or fragile media

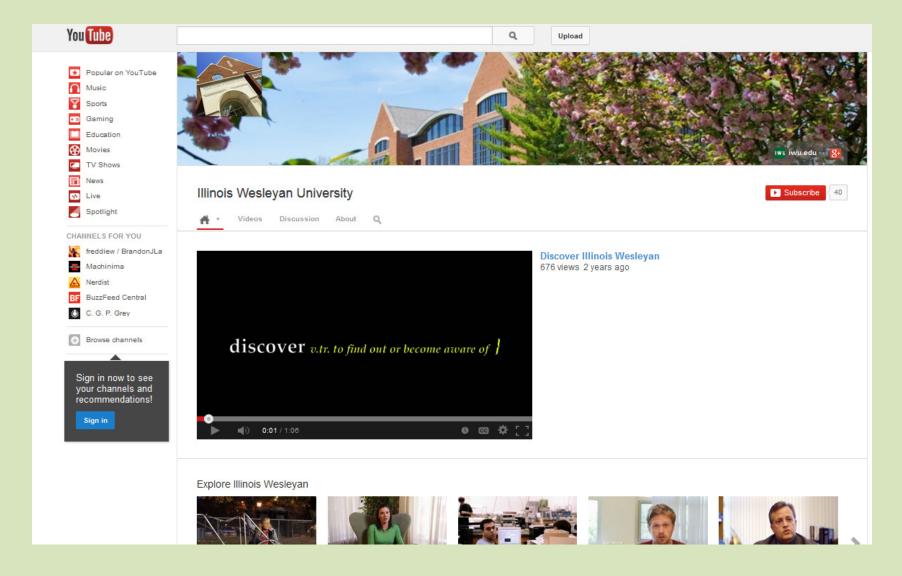
"Data deteriora Univ. of VA Libra Preservation Se http://www.flic photos/uvalibp ion/803308599



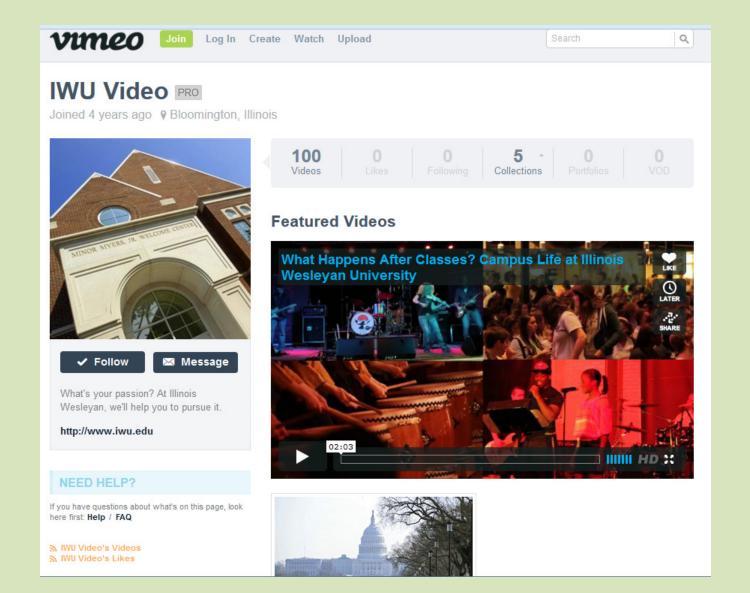
## Cloud providers may not = stable, consistent access

- What security measures do they promise?
- Do they back up anywhere? How will you know?
- Do they offer an exit strategy if they go out of business?
- If the person who set up the account dies, who has access?

### Served ≠ preserved



### Compressed files



### Third party contracts





# Sharing & convenience? Yes! Continued access? Maybe.

### Google

#### Drive

Welcome to Google Drive, the new home for Google Docs





Access everywhere



Store files safely



Collaborate with Google Docs

Learn more

### Homepages only on IWU servers

Where is Student Life?



#### Blogs at Illinois Wesleyan

#### **Student Teaching**

Just another Blogs at Illinois Wesleyan site

#### ITS @ IWU

Comments on Illinois Wesleyan University's Information Technology: planning and operations

## Some is here... sometimes.

#### IWU SAAC

Student-Athlete Advisory Committee

#### **Hispanic Studies**

¿Qué pasa? Otro blog de Illinois Wesleyan

#### ACTION RESEARCH CENTER at Illinois Wesleyan University

Use the Navigation Menu to Learn More About ARC's Connection to the Community!

#### NSLS at IWU

Welcome to the National Society of Leadership and Success



The John Updike Society



#### GREENetwork

Tips and updates from Illinois
Wesleyan University's
GREENetwork

#### IWU Anthropology Club

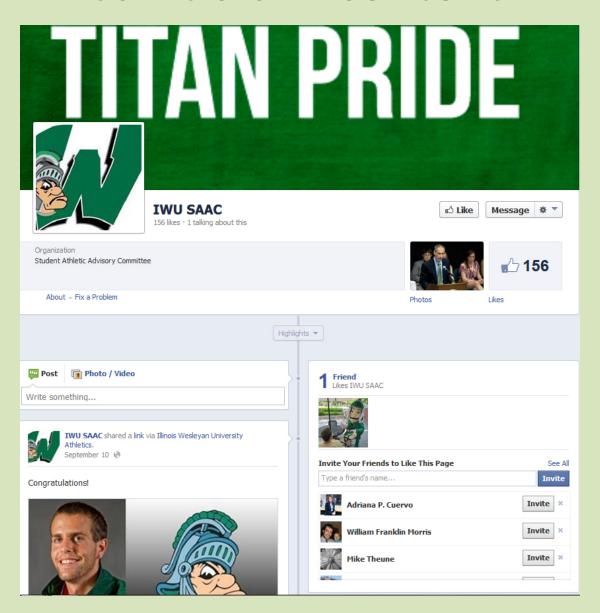
Information for students interested in anthropology

#### Around the House

IWU English Department News

### Can't crawl content

Student Life & Culture in the Cloud = History at Risk



- "All that's on the internet, it'll always be there."
  - Quality of what's there? Who is responsible?



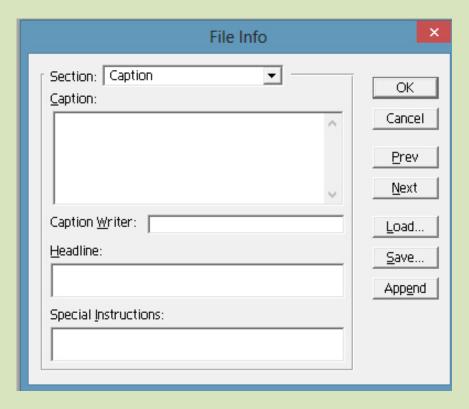
Is it usable?

IWU Greek System page 1997



http://web.archive.org/web/19971007185105/http://www.iwu.edu/organizations/greeksystem/index.html

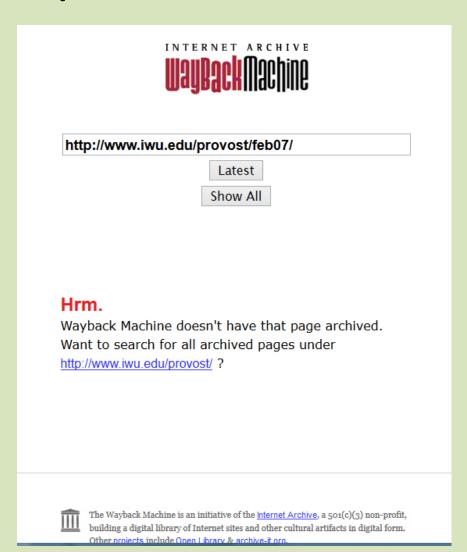
Is it useful?
Not without metadata!





NOT always going to be there

IWU Provost 2007



http://web.archive.org/web/\*/http://www.iwu.edu/provost/feb07/\*

#### Illinois Wesleyan University Disclaimer of Content

Information in the Illinois Wesleyan University Web is provided by many different people. While the authors of these pages try to keep them accurate and up-to-date, the University cannot guarantee that they always will be. If you see something in an IWU document that should be corrected or updated, send mail to the address on the document (or to webmaster@iwu.edu if the document is unsigned.) Be sure to give the full URL of the document in your letter.

Unless otherwise noted, the web information does not represent official statements or views of Illinois Wesleyan University.

"...cannot guarantee that they always will be...."

http://www.iwu.edu/disclaimer.html Accessed September 14, 2013

- "All that's on the internet, it'll always be there."
- Digital files deteriorate.

### They call it Bit Rot

"The stakes are high. Mistakes 30 years ago mean that much of the early digital age is already a closed book (or no book at all) to historians."

The Economist April 28, 2012



### Digital data Bit rot

The world is losing its ability to reconstruct history. Better regulation could fix that

Like 939

Apr 28th 2012 | From the print edition

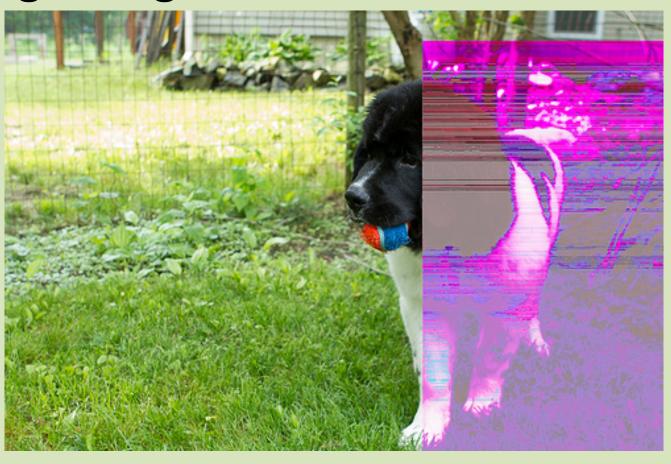


PICTURE yourself as a historian in 2035, trying to make sense of this year's American election campaign. Many of the websites and blogs now abuzz with news and comment will have long since perished. Data stored electronically decays. Many floppy disks from the early digital age are already unreadable. If you are lucky, copies of campaign material, and of e-mails and other materials (including declassified official documents), will be available in public libraries.

But will you be able to read them? Already, NASA has lost data from some of its earliest missions to the moon because the machines used to read the tapes were scrapped and

### Regarding Permanence

Files do
deteriorate &
with no sign
until you try
to open them



"And this is why you backup your files..." By npurmort http://www.flickr.com/photos/npurmort/7158077037/

- "All that's on the internet, it'll always be there."
- Digital files deteriorate.
- Software changes make the ability to read file content uncertain over time.

### Software and "Permanent"

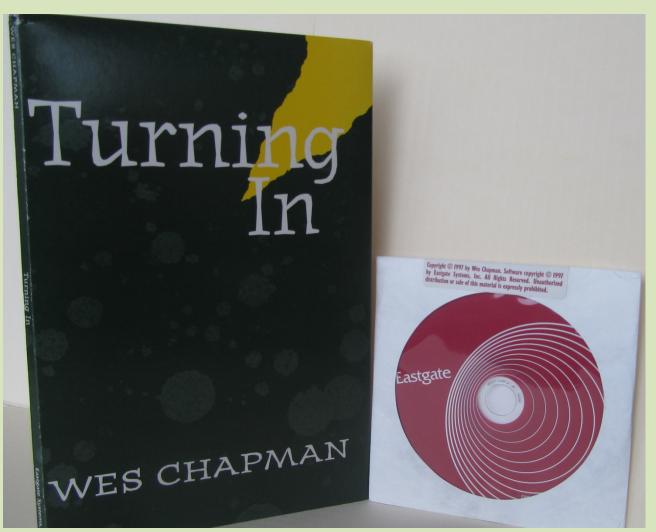


#### Dear Enid:

Thankó verù mucè foò á lovelù luncheoî a imagină yoō no÷ iî Indiá ană wondeò iæ y samă Indianó witè whoî wå talkeă yearó a werå quitå remarkable¬ buô thå political atomiã bombó foò Indiá witè eager¬ burni noô tī kilì anù insects® (Several haä t silk so that no insect would get caught



# Software and "Permanent" (1997)



# Software and "Permanent" (2001)



# Perceptions of Permanence

- "All that's on the internet, it'll always be there."
- Digital files deteriorate.
- Software changes make the ability to read file content uncertain over time.
- Some online content should not be permanent.
  - "Official" files are superseded and old content not removed from the Web site.

# Perceptions of Permanence

# Version control

#### Illinois Wesleyan: Alcohol Policy and Adjudication Guidelines

IWU Philosophical Statement. PHILOSOPHY. A review of alcohol issues at IWU begins with guiding principals upon which **policies** and programs are built. www.iwu.edu/judicial/Alcohol\_Policy.html

#### Illinois Wesleyan: Student Handbook Policy and Regulations

Campus Environment. Illinois Wesleyan encourages the entire community to engage in positive dialogue involving different life experiences. Intercultural ... https://www.iwu.edu/judicial/.../StudentHandbookPolicies.html

#### Illinois Wesleyan: Judicial Affairs - Whistleblower Policy

Whistleblower Policy. I. PURPOSE AND APPLICABILITY. The purpose of this policy is to set forth Illinois Wesleyan University's policy on Board member, faculty, ... https://www.iwu.edu/judicial/policy/WhistleBlowerPolicy.html

#### Illinois Wesleyan: ITS Usage Policy

IT USAGE **POLICY**. Just as knowledge and discussion are at the heart of teaching and learning, so too are information and communications technology such as ... https://www.iwu.edu/its/about/itsusage.html

#### Illinois Wesleyan: Student Handbook Alcohol and Substance Policy

Alcohol and Substance Policy. Illinois Wesleyan University endorses the Drug Free Schools and Campuses Act and the required provisions of that act may be ... https://www.iwu.edu/judicial/.../StudentHandbookAlcoholPolicy.html

#### Illinois Wesleyan: Sexual Harassment and Sexual Violence Policy

Sexual Harassment and Sexual Violence **Policy**. Section 1: Statement of Purpose . Illinois Wesleyan University ("University" or "IWU") is committed to creating, ... www.iwu.edu/judicial/handbook/SexualHarassment.html

#### Illinois Wesleyan: Judicial Affairs Federal Policies

Federal Policies. Federal Drug-Free Schools and Communities Amendment Act. The federal government has enacted two phases of legislation aimed at ... www.iwu.edu/judicial/MandatedPolicies.html

#### Social Media Policy This policy provides guidelines for ...

File Format: PDF/Adobe Acrobat

Social Media **Policy**. This **policy** provides guidelines for communicators using social media on behalf of. Illinois Wesleyan University academic and ... www.iwu.edu/writingguide/social-media-**policy**.pdf

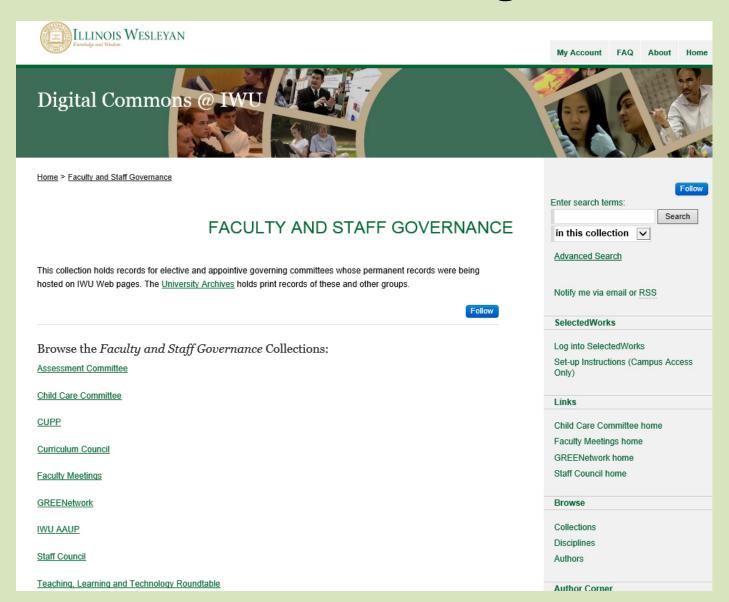
# Two steps towards preservation

Making back up copies of digital objects vs.

Making back up copies of digital objects and storing them in different places

- The minimum recommendation
- Geographic distance = more protection
  - IWU Web site and networked servers are backed up to another city.
  - Specialized software like Digital Commons

# Searchable, stable, migrate-able



My Account

Enter search terms:

Advanced Search

in this repository -

FAQ

About

Search





#### Browse Journals and Peer-Reviewed Series

#### Constructing the Past (History Department)

A publication of the Nu Gamma chapter of Phi Alpha Theta and the Department of History at Illinois Wesleyan University

Visit Journal



#### Pseudonym Required (English)

This website is home to the creative writing of student authors at IWU. At Pseudonym Required, our goal is to help students become better writers by providing feedback to the authors and offering them a chance for their pieces to be published. Our main focus is helping authors improve their writing, but our submission and revision process is also designed to give us an opportunity to showcase the creative talents of our fellow IWU students. All of the pieces published this year as well as our archives of past publications are located on this site. Take a look at some of our student writers' work!

Visit Journal

#### Res Publica - Journal of Undergraduate Research (Political Science Department)

Res Publica - Journal of Undergraduate Research is a publication of Pi Sigma Alpha, the Political Science Honorary Society. It is funded by the Illinois Wesleyan University Student Senate.

Visit Journal



RES PUBLICA

Notify me via email or RSS SelectedWorks Log into SelectedWorks Set-up Instructions (Campus Access Only) **Browse** Collections Disciplines Authors **Author Corner** Author FAQ Submit Research Copyright Information



# Full digital preservation

Deciding which digital objects need to be saved, describing them,

backing up in geographically distant places, and monitoring the digital files for deterioration.

4 Ds and 1 M
Decision making + Description +
Duplication + Distance + Monitoring =
Digital Preservation

### Official definition: LoC

"Digital preservation is the active management of digital content over time to ensure ongoing

ASK A LIBRARIAN

DIGITAL COLLECTIONS

**Digital Preservation** 

LIBRARY CATALOGS

Fostering a national outreach and education network to encourage

the preservation of digital content.

See how you can learn more a

Search Loc.gov

□ Print Subscribe Share/Save QGive Feedback

stewardship.

GO

LIBRARY OF CONGRESS

The Library of Congress > Digital Preservation

DIGITAL PRESERVATION

> Web Archiving at the Library of

Congress

access."

Search this site GO Home About This is the home of the National Digital Stewardship Alliance, Digital Preservation Outreach and Education and the National Partners Digital Information Infrastructure and Preservation Program. Collections Tools Showcase **Recent Blog Posts** National Digital Stewardship Alliance Multimedia & Publications An initiative of the National Digital Information Infrastructure and > Dear Husband: I'm So Sorry for Your Meetings & Events Preservation Program, NDSA members work together to make a Data Loss Education & Training sustained contribution to digital stewardship. Welcome, NDSR Inaugural Class! Contact Us > Snow Byte, the Video! http://www.digitalp Read about the NDSA » Read The Signal Blog » **Preserving Your Digital Memories** > Digital Formats Sustainability Just as libraries are preserving the nation's digital memory, Library of Congress Digital > Federal Agencies Digitization individuals need to take steps to preserve their own digital **Preservation Newsletter** Guidelines Initiative memories. Here are some tips on what you can do. Preservation and Conservation at the Library of Congress View the tips » > Section 108 Study Group Standards at the Library of Digital Preservation Outreach and Education Congress Newsletter provides updates on digital

reservation.gov/ab out

### Official definition: ALCTS

"Combines **policies**, **strategies** and **actions** to ensure access to content that is born digital or converted to digital form **regardless of the challenges** of file corruption, media failure and technological change.

"The goal of digital preservation is the **most accurate** rendering possible of authenticated content over time."

Preservation and Reformatting Section, Association for Library Collections & Technical Services
 Association for Library Collections & Technical Services, a division of the American Library Association; http://www.ala.org/alcts/resources/preserv/2009def

### Then & now: the BIG difference

DP is widespread, cultural heritage concern.

- Time elapsing between technological changes and scale of potential problems ahead.
  - Consider span of time between technological changes: handwriting to print to audio/visual media to digital media.
  - Consider the relative difficulty in creating earlier types of media and the proliferating paths to creating and storing that are available today.
- Attention to our needs now = survival.

### A time travel survivor



#### Reasons to act

#### Hardware and media become obsolete









Chamber of Horrors <a href="http://dpworkshop.org/dpm-eng/oldmedia/chamber.html">http://dpworkshop.org/dpm-eng/oldmedia/chamber.html</a>

## Reasons to act, pt. 2

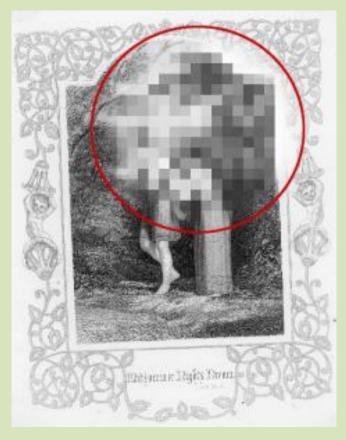
Loss of original file fidelity



"Migrant Mother" by Dorothea Lange. Courtesy of the Library of Congress. http://tinyurl.com/lmh4g6q

# Reasons to act, pt. 3

Intentional editing





# Reasons to act, pt. 4

Quantity being saved and loss of context that gives digital objects their meaning

Channel the NSA ...just a little!

Name	Date modified	Type	Size
	12/1/2000 10100 1111	71 LO IIIIGGE	2,110110
■ _DSC0089	12/7/2006 10:01 PM	JPEG image	1,475 KB
■ _DSC0090	12/7/2006 10:01 PM	JPEG image	1,455 KB
■ _DSC0091	12/7/2006 10:01 PM	JPEG image	1,787 KB
■ _DSC0093	12/7/2006 10:01 PM	JPEG image	1,436 KB
■ _DSC0094	12/7/2006 10:02 PM	JPEG image	1,431 KB
■ _DSC0095	12/7/2006 10:03 PM	JPEG image	1,475 KB
■ _DSC0096	12/7/2006 10:03 PM	JPEG image	1,464 KB
■ _DSC0099	12/7/2006 10:04 PM	JPEG image	1,464 KB
■ _DSC0100	12/7/2006 10:04 PM	JPEG image	1,452 KB
■ _DSC0102	12/7/2006 10:05 PM	JPEG image	1,444 KB
■ _DSC0103	12/7/2006 10:05 PM	JPEG image	1,468 KB
■ _DSC0106	12/7/2006 10:05 PM	JPEG image	1,461 KB
■ _DSC0108	12/7/2006 10:06 PM	JPEG image	1,423 KB
■ _DSC0109	12/7/2006 10:07 PM	JPEG image	1,452 KB
■ _DSC0112	12/7/2006 10:07 PM	JPEG image	1,456 KB
■ _DSC0113	12/7/2006 10:08 PM	JPEG image	1,440 KB
■ _DSC0116	12/7/2006 10:08 PM	JPEG image	1,444 KB
■ _DSC0117	12/7/2006 10:09 PM	JPEG image	1,456 KB
■ _DSC0119	12/7/2006 10:10 PM	JPEG image	1,465 KB
■ _DSC0121	12/7/2006 10:10 PM	JPEG image	1,462 KB
■ _DSC0123	12/7/2006 10:11 PM	JPEG image	1,461 KB



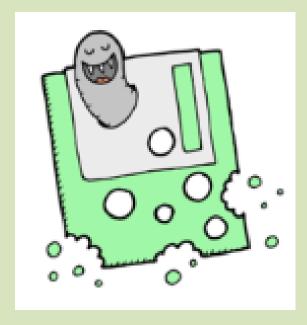
# Taking a pause for a pulse & our students

### "Active management" job opportunities

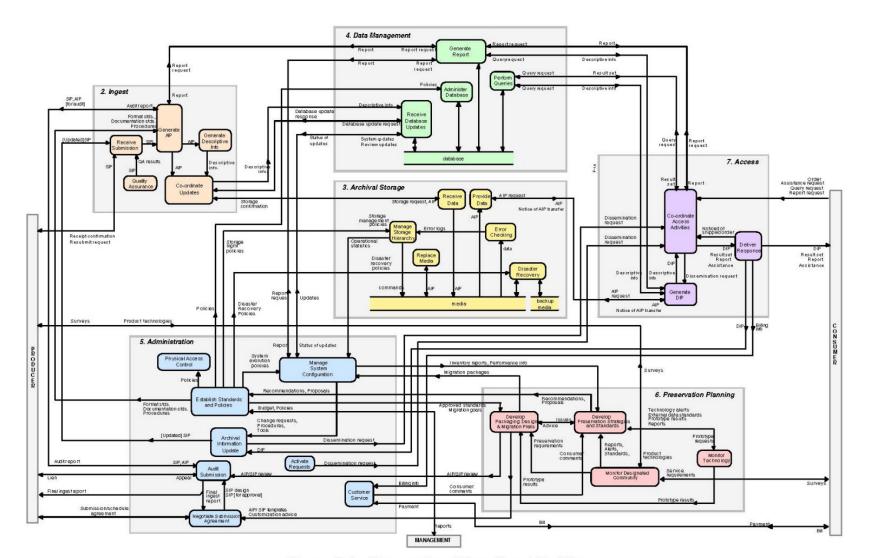
- Archivists, Librarians & IT: organize data, data mining, search/discovery processes
- Computer scientists & Mathematicians: digital forensics work on old media and obsolete systems, encrypting or breaking encryption
- Lawyers: intellectual property, contracts
- Sociologists, anthropologists, psychologists: hoarding as a phenomenon, coping with loss
- Archivists: decision making/selections, long-term planning for patron-base needs

# In the beginning

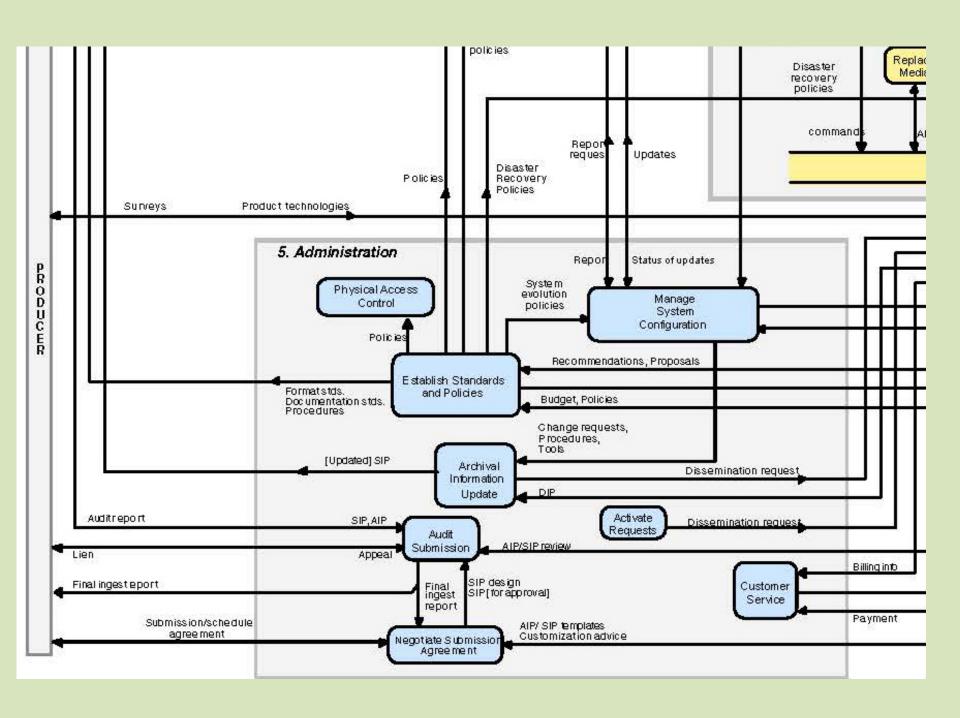
- Analog media is okay alone for long periods
  - Changes are usually visible to the eye or discernible by the nose!
- My education in digital preservation (2007)
  - grant mandates, workshops
  - Bit Rot seemed of most concern;
     digital file degradation is not obvious until too late



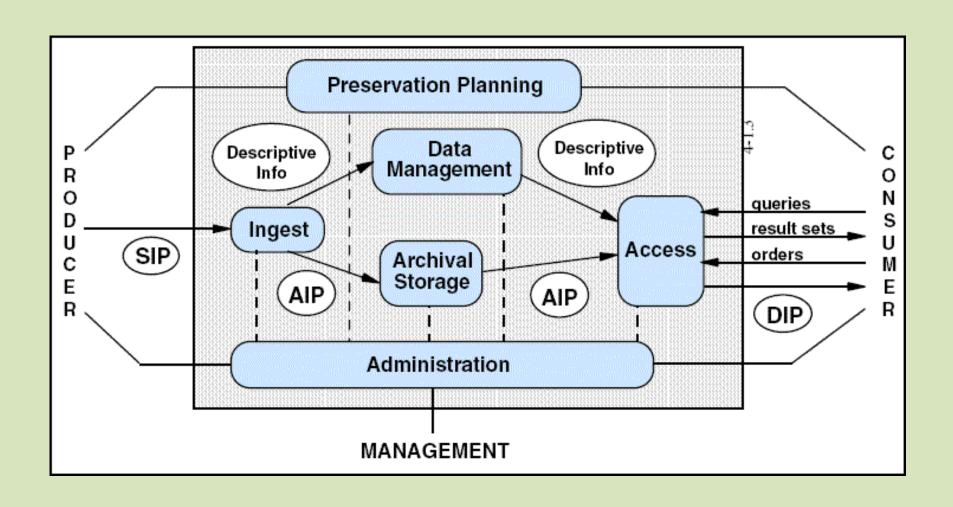
# Digital preservation solved



**Figure F-1: Composite of Functional Entities** 



# Digital Preservation at a Glance



### The IMLS Grant

- No consortium (CARLI) or Big 10 help available.
- Do what librarians do best: collaborate!
- National Leadership Grant from federal agency awarded Fall 2011.



"Advises the President and Congress on plans, policies, and activities to sustain and increase public access to information and ideas."

















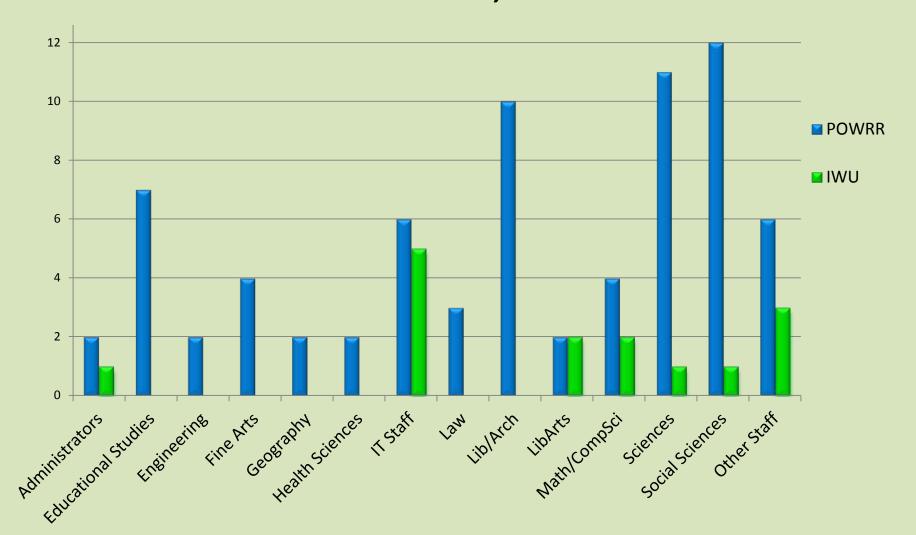
### Objective

Investigate, evaluate and recommend scalable, sustainable digital preservation solutions for libraries with smaller amounts of data and/or fewer resources.

Reports! This is for a government agency, after all. (Final due May 2014)

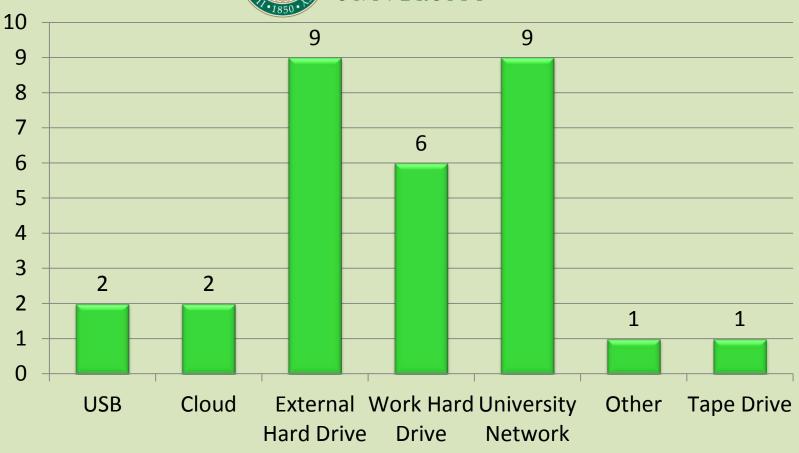
- Case Studies
  - Interviews with faculty, staff, administrators.
  - Explore personal and institutional data management practices.
  - Compare results across institution types.

# Interviewees POWRR = 76 ; IWU = 15

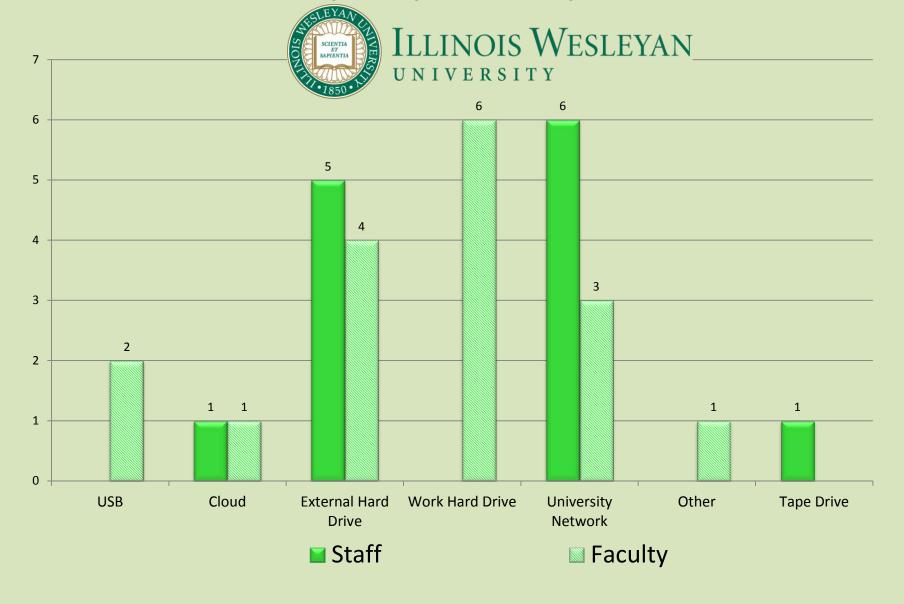


# Where do people keep their data?

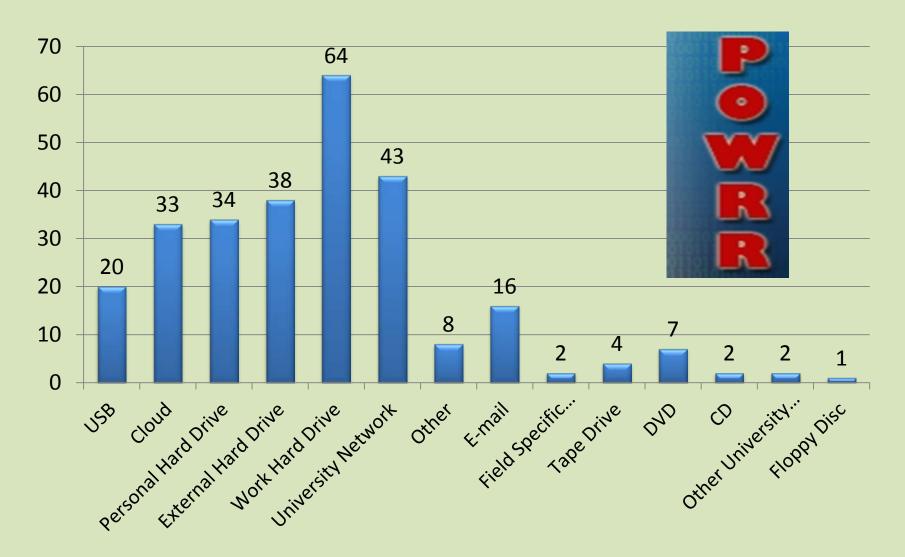




# Where do people keep their data?



# Where do people keep their data?



Reports! This is for a government agency, after all. (Final due May 2014)

- Case Studies
- Analysis of digital preservation products and services
  - Ease of use, availability of support, costs
  - Two parts: Tool Grid and in-depth testing

Evaluating tools and services available today http://digitalpowrr.niu.edu/tool-grid/

	• • ,					٢_									<b>-</b>		51147							
Ingest					Processing					Acc	Access Storage						Maintenance				Other			
Digital POWRR Tool Evaluation Grid	Сору	Fixity Check	Virus Scan	File Dedupe	Auto Unique ID	Auto Metadata Creation	Auto Metadata Harvest	Manual Metadata	Rights Management	Package Metadata	Auto SIP Creation	Public Interface	Auto DIP Creation	Auto AIP Creation	Reliable, Long-Term Bit Preservation	Redundancy	Geographically Despersed Data Storage Model	Exit Strategy	Migration	Monitoring	Auto Recovery	Open Source	Clear Documentation	Cost
ACE (Audit Control Environment)		x																		x		x		Free
AFF Open Source Computer Forensics Software							x															x		Free
Amazon Cloud	X																	X					x	Varies
Archive-It	X	х	X	X	x	x	Х	х	х	х	х	х	х	х		х	x	X		x		x	х	Varies
Archivematica		x	x		x	X	x	x	x	x	х	x	х	х	x					X		x	x	Free
BagIt Library		x			X	x				Х				Х				X						
Baglt Transfer Utilities																								Free
BWF MetaEdit		x						x	X													x		free
Carbonite	х	x										х			x	x		х		X	X			\$149 – \$599
Chronopolis	X	X		X	X	X	X	X	X	X	X			X	X	X	x	X	X	X	X		X	\$1,500 - \$2,200

Testing five tools/services in-depth Curator's Workbench (UNC)









- File management costs depend on total file size (in TBs)
  - Remember: two copies of everything = minimum
  - Six copies of everything = optimum

Reports! This is for a government agency, after all. (Final due May 2014)

- Case Studies
- Analysis of digital preservation products and services
- National and local dissemination (throughout 2014)
  - Conferences and workshops
  - Web-based presentations
  - Meetings like this...thank you!

# Unanticipated outcomes

#### POWRR and collaboration

- UK-based Project SPRUCE (Sustainable PReservation Using Community Engagement): creating a community-owned tool registry that will feed into the POWRR Grid.
- MetaArchive: Piloting technology to support smallscale models.
- NDSA (Library of Congress): Co-planning regional events for NDSA membership drives and POWRR workshops.



### Observations

Local decisions are the most important first steps.

- We will not have an enduring historical record if we rely on http://www.iwu.edu or social media sites.
  - Lack of control/ownership, quality of file types and inconsistent metadata.
  - We must separate promotion from permanence.
     As of now, they must be two distinct things from a digital preservation perspective.

### Observations

- Durability of digital objects is a factor of preserving the context with the file.
  - Metadata: significance of a digital object cannot be understood without it.
- We should not try to keep everything.
  - Some records do not need to remain digital.
  - Cultural/personal habits are against us here!
  - Records creators must be involved but your archivist can help!

### **Observations**

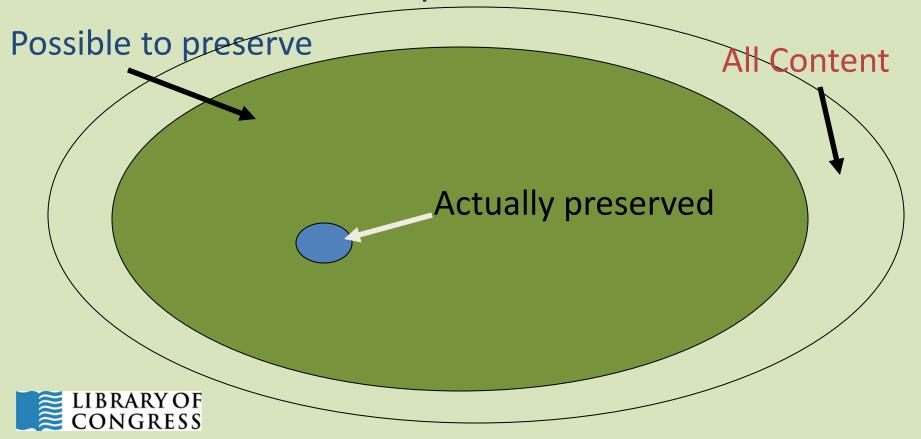
Once we know what is important to keep and for how long, make decisions about

- what types of storage locations the digital objects need, who has access, who is responsible for maintaining.
- what type of long-term preservation system will work.



# Decision making visualized

Good preservation decisions are based on a deep *understanding* of the possible content to be preserved



# Recommendations for office and at home

- If you use specialized software, make sure to check old files in new versions that come out and save them as new.
  - Save a backup copy in a "standardized" format or one that you can migrate.
  - "Normalizing" to a known medium may still be the best hope for survival.
  - Your archivist is ready to help with decisions and format recommendations!



# The best laid plans



# For long-term preservation

Deciding which digital objects need to be saved, describing them,

backing up in geographically distant places, and monitoring the digital files for deterioration.

4 Ds and 1 M
Decision making + Description +
Duplication + Distance + Monitoring =
Digital Preservation



### Institutional Recommendations

#### What we can do in the short term:

- Create a record? Copy your archivist!
- Start a new group? Tell your archivist.
- Want to preserve student activities? Me, too!
   Let's talk.
- Not sure how to start? Consult your archivist.

### Institutional Recommendations

#### Harder questions

- What responsibilities should we take for content that is not on IWU servers?
- What about faculty data?
  - Consider investment in PAT cycle.
  - Consider external pressures like NSF & NEH grant rules for Data Management Plans.
- Solo system implementation not recommended.
   Who will we partner with?
  - How can responsibilities (staffing & systems) be shared—both on campus and with others?

## Final Quiz

### Digital Preservation is...

- A. More than scanning and cloud storage.
- B. Decision making that begins at the moment of creation.
- C. A series of actions that will help secure personal and cultural heritage.
- D. A shared responsibility for our institutional legacy.
- E. All of the above.

# Why should you care?



Geek & Poke, posted December 12, 2007. Available at http://geek-and-poke.com/2007/12/how-to-save-dig.html

#### Thank You!

mminer@iwu.edu x1538

Visit our project Web site http://digitalpowrr.niu.edu



And see the Library of Congress's suggestions about saving your own digital heritage...future archivists will thank you!

http://digitalpreservation.gov/personalarchiving