

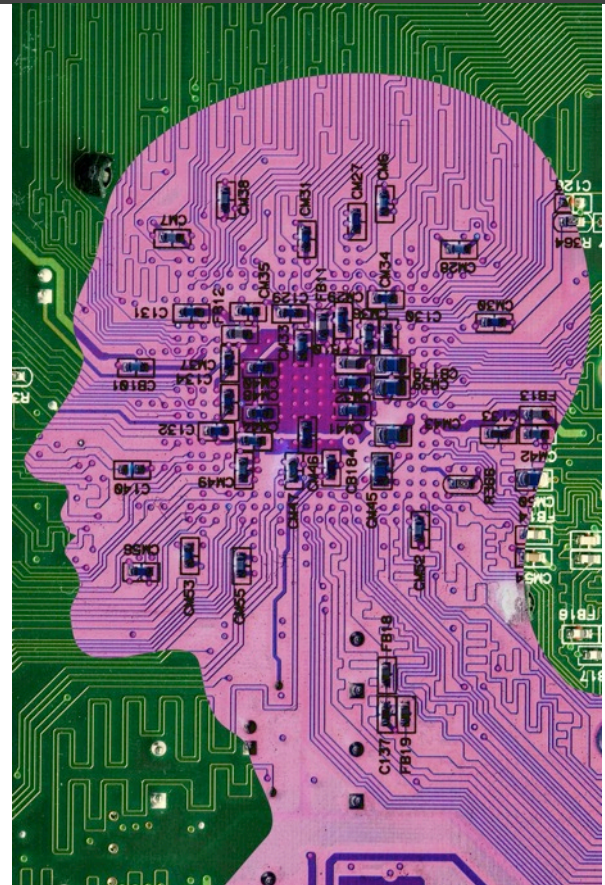


“Regular or digital?”

Forecasting the Future of Digitization

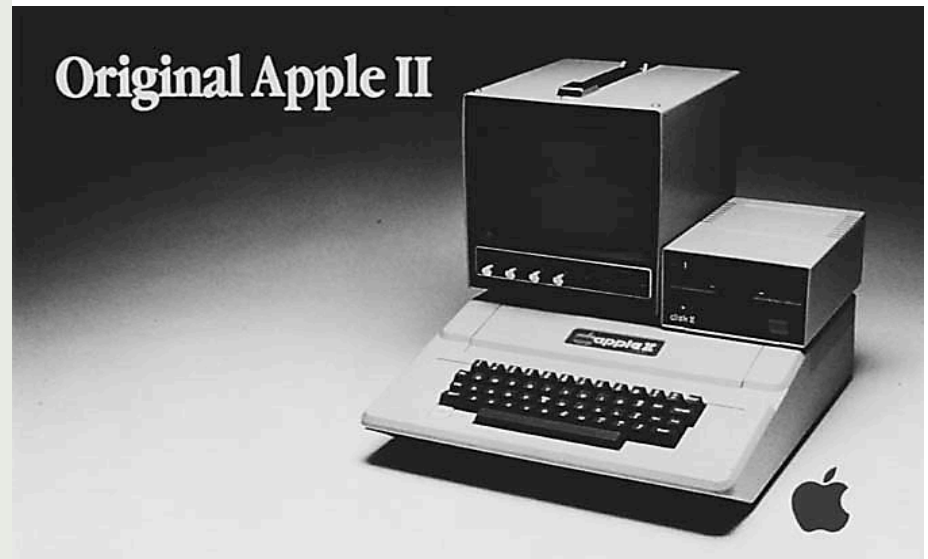
“You may delay, but time will not.” – Benjamin Franklin

- Digesting/Creating more information and documents
- 10x more information will be produced in 2011 vs. 2006.
- Digital technologies keep moving forward, librarians/archivists unsure of what to grab onto.
- File formats discussion will perhaps provide a more steady feeling towards digital content
- Digital content increases access to material not a preservation method.



First Personal Computer

- 1977 – Apple II
- Need not initially obvious
- Typewriter vs. Born Digital
- Low Quality
- Storage / Floppy Disks
- DOS proprietary file types.
- Many records lost.



Compact Discs

- Introduced to market in 1982
- No cooperation with computer companies
- Burned scans much lower quality than pressed.
- Easily worn away, scratched, and damaged.



PDF (Portable Document Format)

- Early 1990s - Text and Graphics for printing presses.
- Adobe Acrobat > Adobe Reader > Basic Image Viewer – Open Access
- PDF/A - ISO 19005-1



Optical Character Recognition

- Ultimate Metadata
- Tagging
- Searchable Content
- ReCaptcha



The image shows a screenshot of the reCAPTCHA website homepage. A large, curved orange arrow points from the reCAPTCHA logo on the left towards the main content area on the right. The reCAPTCHA logo is a red 'C' with a white 'reCAPTCHA' text inside. The main content area features a navigation menu with links: 'WHAT IS reCAPTCHA', 'GET reCAPTCHA', 'PROTECT YOUR EMAIL', 'MY ACCOUNT', and 'RESOURCES: DOCS & PLUGINS'. Below the navigation menu, there is a section titled 'reCAPTCHA IS A FREE ANTI-BOT SERVICE THAT HELPS DIGITIZE BOOKS.' This section includes a snippet of text from a digitized book: 'steamboat train, from New York City. This **morning** ran off the track near New-London. Four cars plunged...'. A red box highlights the word 'morning' in the text. Below the text snippet is a reCAPTCHA challenge box showing the words 'morning' and 'upon' in a stylized font. A red arrow points from the word 'morning' in the text snippet to the word 'morning' in the challenge box. Below the challenge box is a red button that says 'LEARN HOW reCAPTCHA WORKS'. Further down is another red button that says 'USE reCAPTCHA ON YOUR SITE'. Below this button are three bullet points: 'STRONG SECURITY', 'ACCESSIBLE TO BLIND USERS', and '30+ MILLION SERVED DAILY'. At the bottom of the page, there is a 'NEW' announcement: 'Our more secure audio CAPTCHA helps transcribe historical media!'. The footer contains links for 'Blog | About Us | Contact | FAQs' and the copyright notice '© 2009 Carnegie Mellon University.'

reCAPTCHA™

- WHAT IS reCAPTCHA
- GET reCAPTCHA
- PROTECT YOUR EMAIL
- MY ACCOUNT
- RESOURCES: DOCS & PLUGINS

reCAPTCHA IS A FREE ANTI-BOT SERVICE THAT HELPS DIGITIZE BOOKS.

steamboat train, from New York City. This **morning** ran off the track near New-London. Four cars plunged...

morning upon

Type the two words:

reCAPTCHA™ stop spam. read books.

→ LEARN HOW reCAPTCHA WORKS

USE reCAPTCHA ON YOUR SITE

- STRONG SECURITY
- ACCESSIBLE TO BLIND USERS
- 30+ MILLION SERVED DAILY

NEW Our more secure audio CAPTCHA helps transcribe historical media!

Blog | About Us | Contact | FAQs
© 2009 Carnegie Mellon University.

gwap news about contact

We make games with a purpose.

When you play a game at Gwap, you aren't just having fun. You're helping the world become a better place. By playing our games, you're training computers to solve problems for humans all over the world.

1 You play the game

Let's take the ESP Game for example. You and a partner see the same image and are asked to type in a tag for it. When you agree on a tag, you move on and are awarded points. After just a minute of play, you've agreed on six or seven tags.

2 Computers get smarter

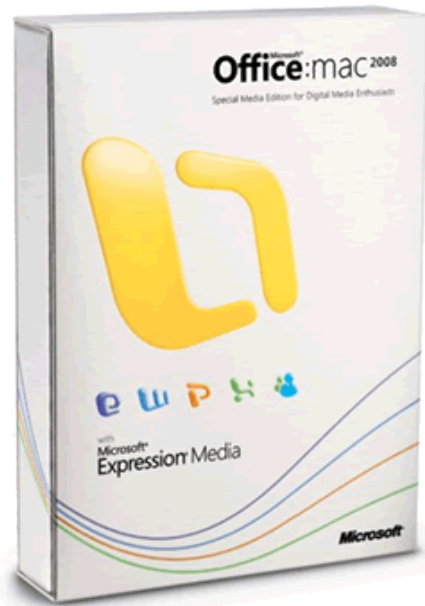
We record those six or seven tags and associate them with the images.

3 Everyone benefits!

Now a search engine will have a better idea of what's in those images.

Games with a Purpose

- Tagging Photos for search could develop this idea to apply to print/photo/archive materials, method for having fun and tag!
- Creating the most metadata on records makes the digital record more permanent and usable.



DOCs

- Microsoft “Open” file format
- Troubles with Microsoft, broken promises and trouble with new 2007-08 Office
- OpenOffice.org



well, i suppose that
i love you more
than animated .GIFs
too



Image Files

- TIFF (Tagged Image File Format)
- JPEG (Joint Photographers Export Group)
- GIF – should not use for preservation.

Quick File Tips

- No passwords
- Read-Only awareness
- Linking to outside websites use full address (breadcrumbs)
- No compression only zip!
- No internal scripts (JavaScript) or proprietary executable files (QuickTime)
- Metadata, Tagging! Metadata, Tagging! – File/Folder names
- Design
 - Fonts 14 base fonts
 - Helvetica (Bold, Italic, Regular)
 - Courier (Bold, Italic Regular)
 - Times (Bold, Italic, Regular)
 - Symbol
 - Zapf Dingbat
 - Colors
 - CMYK & RGB

Meet the cast:

ABCD
EFGHIJK
LMNOP
QRSTU
VWXYZ

Now see the movie:

Helvetica

A documentary film by Gary Hustwit

A Swiss Date production in association with 	Executive Producers Markus Müller Nicola Müller Michael Götsche Wolfgang Petersen Christoph Waser Andreas Kappeler	Producers Lutz Kunze Michael Götsche Wolfgang Petersen Christoph Waser Andreas Kappeler	Director of Photography Lutz Kunze	Additional Photography Lutz Kunze Christoph Waser Andreas Kappeler	Sound Editor Lutz Kunze	Sound Design Lutz Kunze	Music The Swiss Lovers Andreas Kappeler Christoph Waser Andreas Kappeler	Executive Producers Markus Müller Nicola Müller Michael Götsche Wolfgang Petersen Christoph Waser Andreas Kappeler
---	---	--	---	---	--	--	--	---

Digital Preservation Keys to Success

- LOCKSS – MetaArchive
“Lots of Copies Keep Stuff Safe”
- File Migration Key
- Keep Format/Technology obsolescence in the back of your mind.
- Stay on top of technology trends.



Questions?

Image Sources:

Slide 1 - http://www.danscartoons.com/comp35_demo.gif

Slide 2 - <http://www.cs4fn.org/futurehuman/images/winninghead-sml.jpg>

Slide 3 - <http://www.edwardsamuels.com/illustratedstory/chapter%204/Applell.jpg>

Slide 4 - <http://toknowasiamknown.files.wordpress.com/2007/07/microwaved-dvd.jpg>

Slide 5 - <http://www.waldorp.nl/lourens/literature/pdfLogo.jpg>

Slide 6 - "reCAPTCHA: Stop Spam, Read Books." reCaptcha, <http://recaptcha.net/> (accessed April 26, 2009)

Slide 7 - "Gwap.com About" Games with a Purpose, <http://www.gwap.com/gwap/about/> (accessed April 26, 2009)

Slide 8 - MS Office 2008 Box - http://www.mrgadget.com.au/catalog/images/microsoft_office_2008_mac_special_media_edition.gif

Mac v. PC - <http://rickveneer.com/blog/wp-content/macvspc.png>

Slide 9 - Cartoon - <http://www.toothpastefordinner.com/123006/animated-gif-love.gif>

Flipping Bunny - <http://i8.photobucket.com/albums/a49/toraigan/BunnyHopSmlr.gif>

Slide 10 - <http://futererockstarsofamerica.com/blog/wp-content/uploads/2008/12/helvetica-poster.jpg>

Slide 11 - http://gallery.nofactzone.net/albums/userpics/195810951_9bfd084871_o%7E0.jpg

