

Digitization 101

The place for staying up-to-date on issues, topics, lessons learned and events surrounding the creation, management, marketing and preservation of digital assets.

About Me



Jill Hurst-Wahl

Jill Hurst-Wahl, MLS, is a digitization consultant and owner of Hurst Associates, Ltd. She also teaches digitization for Syracuse University's School of Information Studies. You can read her bio [here](#) and additional info on her digitization work. To contact Jill, use these [methods](#).

[View my complete profile](#)



This work/blog is licensed under a Creative Commons Attribution-NonCommercial-ShareAlike 3.0 License.

Want to receive updates from Digitization 101?

[Subscribe to Digitization 101](#)



Thursday, November 13, 2008

"Old Fulton NY Post Cards" is more than it seems

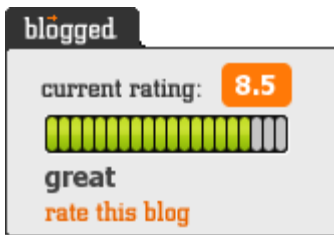
This blog post is long overdue...

During the summer, I ventured to Fulton, NY to visit the public library and to talk to Tom Tryniski, who runs the site "[Old Fulton NY Post Cards](#)." First of all, the site isn't really about postcards. Yes, it started as a small postcard digitization program, but has grown to be so much more. When I saw him, Tryniski has digitized over 6.12 million old New York State historical newspaper pages and would soon be placing another 1.3 million pages online. That's more than 2.8 terabytes of content from 1837 to 2007. This has translated into the site being a popular destination for people who are doing research (more than 1.4 million hits in 30 days).

When I first heard about Tom Tryniski earlier in the summer, I was amazed that I hadn't heard of the project sooner. When I met him, I understood why. This isn't a library doing a project or some other organization. This project is a labor of love by a former IT executive, who has the skills, technology and time to create a growing archive. At first, Tryniski digitized postcards that he had received and placed them online. When they got positive feedback, he digitized some other materials. Soon he stumbled upon content that was rich in information -- newspapers -- and began digitizing newspapers from microfilm. The amount of content he has digitized is amazing.

Okay...so what about his "process"? Is he following recommended practices? As we talked, he told me about his equipment, etc., and - yes - he is following recommended practices. In talking about the equipment he uses, Tryniski says on the site, "Old Newspapers found on this site has have been Scanned by production grade Wicks and Wilson Microfilm scanners which in the authors opinion are the Finest available." Yes, he backs up and backs up and backs up. (This is a former IT guy, remember.)

Does he copyright clear the materials? Tryniski has a good understanding of copyright and the public domain. When things are not in the public domain, he seeks permission. In some cases,



Digitization Resource List (created April 2008)

Digitization 101: The Workshop Resource List

Jill's Speaking & Travel Schedule

All Feeds Calendar [XML](#)

12/10/2008 (Wed)
11:00 AM
AIIP Webinar
Why You Need to Care about Virtual Worlds

1/23/2009 (Fri)
10:30 AM
SMPS Southern Regional Conference
E-Networking & the Benefits of Being Hyper-Connected (presentation)

2/10/2009 (Tue)
9:00 AM
SEFLIN
Digitization 101: Planning & Management of Digitization Projects

2/11/2009 (Wed)
9:00 AM
SEFLIN
Digitization 101: Planning & Management of Digitization Projects

2/26/2009 (Thu)
9:30 AM
SCRLC
Practical Digitizing Series: Planning & Management of Digitization Projects (workshop)

RSSCalendar.com
Powered by RSSCalendar

he has worked directly with content sources on digitizing materials and they've given him permission. When taking about copyright, Tryniski often points people to Mary Minow's article "[Library Digitization Projects and Copyright](#)." (BTW I mentioned Peter Hirtle's [chart](#) to him, which I find very valuable.)

Why does he do this? Tryniski recognizes that there is a lot of valuable content in newspapers and other analogue sources. He also knows that many small libraries and historical societies don't have the capacity to do this work (or even pay for it). Therefore, he set himself on a mission that bring a huge smile to his face when he talks about it. And he is open to working with any organization to further this cause. His only stipulation is that whatever he digitizes, he is able to add to his collection and give people free access to it.

What concerns does he have about this work? Actually, he and I have the same concern about various efforts to compile historic information online. I worry about projects, [like this one](#) by Joyce Tice, that are done by volunteers. I worry about the content surviving long-long-term. For example, what happens when the volunteers are old and unable to maintain the site? Have they made provisions for someone to take over? Personally, I think it would be great if there were someone (or an organization) that was willing to "inherit" these sites in order to ensure their longevity. (BTW if you search for my name on [that site](#) -- once you figure out how to search the site -- you'll see that I have a vested interest in having this site not disappear.) I did not ask Tryniski what specific long-term plans he had made, but given his passion, I'm sure he's already got that documented.

And what do I think of the "Old Fulton NY Post Cards" site? Wow...what a lot of content! What is missing for me is information on the project's background and standards, as well as a list of content sources. I wish the site was less animated, which some may find inhibiting. It would be great to have resources available to help students, teachers and researchers understand and use the materials. The things I want are things that you will see in digitization projects that have been done by institutions, but this is one person and a labor of love. I wonder if he would be open to having interns, who could add supporting materials to the site?

Now that I have told you about it, go check [it](#) out!

Technorati tag: [Digitization](#)

Posted by Jill Hurst-Wahl at 8:32 AM 0 comments [Links to this post](#)

Labels: [Digitization](#)

Wednesday, November 12, 2008

[JISC study on Digital Preservation](#)

Labels

Abraham Lincoln (2)
ALA (4)
Analysis Software (1)
Archives (2)
Blogging (6)
Book Digitization (48)
Book Recommendation (5)
Brewster Kahle (3)
CIL2007 (13)
CIL2008 (21)
Collaboration (1)
Communications (1)
Conference (3)
Content Management System (4)
Copyright (56)
Crowdsourcing (1)
Definition (5)
Digital Asset Management (6)
Digital Audio (6)
Digital Curation (3)
Digital Imaging (6)
Digital Library (12)
Digital Preservation (90)
Digital Repositories (1)
Digitization (245)
Document Processing (1)
Ebooks (2)
Evaluation (1)
Federated Search (9)
Google (26)
Greenstone (9)
Harrisburg (8)
IL2007 (21)
Image Manipulation (1)
Innovation (1)
Institutional Repository (13)
Internet (1)
Internet Archive (2)
JHW (25)
JPEG (1)
JPEG2000 (13)
Knowledge management (2)
Libraries (12)
Library of Congress (2)
LOCKSS (2)
Marketing (16)
Metadata (26)
Microfilm (2)
Microforms (1)
Microsoft (2)
New York State (15)
Nylink (3)
NYS (13)
OCLC (1)
OCR (3)
Open Content Alliance (3)
Optical Storage (2)

Policies

In case you haven't seen this announcement in email:

The JISC are pleased to announce the publication of a study on Digital Preservation Policies which can be downloaded in PDF format from <http://www.jisc.ac.uk/publications/publications/jiscpolicyfinalreport.aspx>

A major business driver in all universities and colleges over the past decade has been harnessing digital content and electronic services and the undoubted benefits in terms of flexibility and increased productivity they can bring. The priority in recent years has been on developing e-strategies and infrastructure to underpin electronic access and services and to deliver those benefits. However any long-term access and future benefit may be heavily dependent on digital preservation strategies being in place and underpinned by relevant policy and procedures. This should now be an increasing area of focus in our institutions.

This JISC funded study completed by Charles Beagrie Ltd aims to provide an outline model for digital preservation policies and to analyse the role that digital preservation can play in supporting and delivering key strategies for Higher and Further Education Institutions. Although focussing on the UK Higher and Further Education sectors, the study draws widely on policy and implementations from other sectors and countries and will be of interest to those wishing to develop policy and justify investment in digital preservation within a wide range of institutions.

Two tools have been created in this study:

- 1) a model/framework for digital preservation policy and implementation clauses based on examination of existing digital preservation policies;
- 2) a series of mappings of digital preservation to other key institutional strategies in UK universities and colleges including Research, Teaching and Learning, Information, Libraries, and Records Management.

Our aim has been to help institutions and their staff develop appropriate digital preservation policies and clauses set in the context of broader institutional strategies.

Technorati tag: [Digital Preservation](#)

Posted by Jill Hurst-Wahl at 11:30 AM 0 comments [Links to this post](#) 

Labels: [Digital Preservation](#)

Pennsylvania (8)
Planning (7)
Preservation (5)
Print on Demand (1)
Privacy (3)
Project Management (1)
Publicity (1)
RAC (13)
Regents (13)
Report (2)
Scrapbooking (1)
Second Life (22)
SLA2007 (9)
SLA2008 (6)
Social Networking (16)
Staffing (1)
Standards (4)
state library (6)
Subscribe (1)
Surveys (2)
Syracuse University (2)
Thesaurus (1)
UNYUNC (2)
Web 2.0 (4)
Web archiving (1)

Blog Archive

▼ 2008 (274)

▼ November (9)

"Old Fulton NY Post Cards" is more than it seems

JISC study on Digital Preservation Policies

For New Yorkers: Notes from NYLA legislative & lib...

Blog post: Quality of Book Digitization

MOS: le magazine du stockage et de la gestion d'in...

Video: Brewster Kahle: A digital library, free to ...

The Vietnam Center and Archive (and John McCain)

U.S. Presidential Election Jill's November schedule

► October (24)

► September (23)

► August (20)

► July (22)

► June (26)

► May (33)

► April (40)

► March (25)

► February (23)

► January (29)

► 2007 (353)

Monday, November 10, 2008

For New Yorkers: Notes from NYLA legislative & library leaders, and Commissioner Mills

On Friday, I was able to attend two events at the New York Library Association annual conference where the future of library funding in New York State was discussed. The speakers were:



- Assemblywoman Amy Paulin, District 88 and Chair, Assembly Committee on Libraries and Education Technology
- ~~Assemblyman Marcus Molinaro, District 103~~ (corrected 11/12/2008)
- Assemblyman Robert Reilly, District 109
- Mark Leinung, Governor's Assistant Secretary for Education
- Richard Strauss, President, New York State Association of Library Boards
- Kathy Miller, Chair, NYLA Legislative Committee
- Regent James C. Dawson
- Jeffry Cannell, Deputy Commissioner for Cultural Education and Interim State Librarian
- Richard P. Mills, President of The University of the State of New York ([USNY](#)) and Commissioner of Education (Commissioner Mills has announced that he will [step down](#) from his position in June 2009.)

Introduced was [Bernard A. Margolis](#) who will assume the position of New York State Librarian in January 2009. Among the 100+ people in the audience were staff from the State Library, representatives of various library consortia, and members of the Regents Advisory Council on Libraries (Gerald Nichols, Sara Kelly Johns and myself).

Notes below are grouped by speaker and are not a transcript. Some are just sound-bites.

Reilly:

- 20% of the State's yearly income comes from taxes paid by Wall Street (business and personal taxes, etc.)
- One-third of children in NYS do not receive a high school diploma.
- Just getting more money for libraries doesn't solve the (fiscal) crisis that is occurring.

- ▶ 2006 (398)
- ▶ 2005 (528)
- ▶ 2004 (70)



- Need consolidation of services.
- Need to straighten out structures.
- We tend to cut services, but don't fix bureaucracies.
- We need to work our way out (of this mess), not cut our way out.

Leinung:

- A strong public library is a part of a strong democratic society.
- Our most precious commodity is information.
- When Wall Street sneezes, NYS get a cold. NYS is getting pneumonia.
- 160,000 people may lose their jobs in the state. 6.5% unemployment rate.
- Deficit is \$47 billion in four years and is the largest in the State's history.
- There is a 35% gap in the general fund used by NYS.
- As the state looks at what to cut, nothing is sacred.

Strauss:

- A library is an economic engine in the community.
- Library trustees must unit to sustain and improve library funding across NYS. Need consistent and stable funding.
- State needs to provide trustee training. Should be mandatory for all trustees in NYS. They are trying to make this a law.

Miller:

- Across NYS, library systems are providing Internet access and continued learning through the library for unaffiliated people.
- Libraries are part of the education infrastructure.
- Libraries are the poster child for efficiency.

Paulin:

- Librarians are forceful, knowledgeable and polite. Librarians need to be less polite as they ask for funding.
- Libraries are a place for people to improve their quality of life.
- Need to think outside the box in regards to advocacy.
- Need to get users involved in our advocacy efforts and create a bigger movement.
- Need to tell legislators what is really happening in our libraries.
- Need to compete against the other interest groups.

Cannell:

- In talking about the budget, he said "everyone needs to get their allowance."

Within the budget discussion, member items are also being discussed.



- This crisis is a change to learn how to be more nimble.
- Need less regulation.

Mills:

- Important to get to fundamentals.
- "I cannot live without books" - Thomas Jefferson. Mills then modified it to say, "I cannot live as a free person without books."
- "A library is the most democratic of all institutions."
- "The reach of libraries is vastly greater" now.
- "A library is where people go to learn the language."
- Libraries...
 - Lost 2% in the last budget
 - Lost 3.35% in operating funds (early this year)
 - Lost an additional 7% in July in operating funds (clarified: 11/12/2008)
 - Have lost \$874,000 this year due to state funding cuts
- Governor wants to cut another \$2 billion from the budget this month.
- Cultural Education's income (day-to-day sustaining budget) is primarily from various fees (e.g., real estate transactions). It also receives some federal money. The fees need to increase in order to sustain the Department (from \$15/transaction to \$22.50). Cultural Education includes the State Library, State Archives and State Museum. (clarified: 11/12/2008)
- Advocacy needs to be local and regional.
- Advocacy needs to be fact-based.
- We need to get legislators into our libraries, so they can see what is happening there.
- All cultural institutions in the State are under the University of the State of New York ([USNY](#)).
- "Everyone needs resources. What value do we add?"
- 1.5 million children have been in free literacy programs.
- The average book in a school library in NYS is 21-25 years old. The amount spent on library resources for pupils needs to increase from \$.625 to \$10,00 per student.
- A family without books cannot educate its children without a library.
- The Regents are working to raise education standards, graduation standards, and graduation rates.

Audience Member:

- Libraries are an economic force in the community. Libraries hire local people and services.
- Librarians have master degrees. Why not use libraries and librarians more in the State as resources to get things done? Libraries are not seen as partners. They are not integrated into other State activities.

Audience Member:

- Libraries are not free, but they provide free access.
- Should be mandatory to have librarians in every school.
- Libraries need "just a little bit of money."

Audience Member:

- Should use library school students to help with advocacy efforts.

Audience Member:

- "We're not far from closing the doors" on our libraries.

Audience Member:

- Cutting libraries is not going to save the budget.
- Think about the future.
- Libraries are a life-line.

I believe that I noted all of the figures mentioned correctly. If I did not, I hope people will leave comments to correct me. (And let me know of other mistakes too! This is important stuff and worth getting correct.)

Thanks to Sara Kelly Johns for allowing me to use her photos.

Corrections and clarifications made on 11/12/2008.

Technorati tags: [Regents](#), [New York State](#), [NYS](#), [RAC](#), [nNYLA2008](#)

Posted by Jill Hurst-Wahl at 10:18 AM 1 comments [Links to this post](#) 
Labels: [New York State](#), [NYS](#), [RAC](#), [Regents](#)

Thursday, November 06, 2008

Blog post: Quality of Book Digitization

[Tim O'Reilly](#) asked to use comments by Juliet Sutherland of [Distributed Proofreaders](#) as a [guest blog post](#). Sutherland, whose organization does all the quality checking and correction for [Project Gutenberg](#), is concerned about the quality being produced by mass digitization programs, i.e., missing pages. Her concerns are with projects such as those run by Google and the Internet

Archive, while she says:

Our experience with various (book page) image archives suggests that those archives associated with libraries are usually reasonably good.

Her [comments](#) are well thought out and articulated. Worth looking at.

Technorati tags: [Digitization](#),

[Book](#)

Posted by Jill Hurst-Wahl at 12:09 PM 0 comments [Links to this post](#) 

Labels: [Book](#) [Digitization](#)

Wednesday, November 05, 2008

[MOS: le magazine du stockage et de la gestion d'informations](#)

Francis Pelletier from *MOS* magazine mailed to me the August and September 2008 issues, which contained information on new and interesting products. [MOS: le magazine du stockage et de la gestion d'informations](#) (the magazine of storage and management of information) is written in French, so let me try to give you a bit of information on the articles of interest, as well as links for more details.

The August issue had an article on [Pixelion](#) (Aix-en-Provence, France), a company produces software for image processing. (Check the web site for more information.)

There was also an article on "des microformes de haute densite pour un archivage perenne" (**microforms of high density for a perennial filing**). The photos in the article were intriguing. Since my French is not that good, I checked online for information on the company mentioned and found [this text](#):

La société ARnanO est l'une des dernières start up créée par le laboratoire français [LETI \(Laboratoire d'Electronique et de Technologie de l'Information\)](#), avec la vocation de développer les produits avancés permettant la gravure de graphismes ou de textes infiniment petits et inaltérables.

A rough translation is:

The company is one ARnanO recent start-ups created by the French laboratory [LETI \(Laboratory of Electronics and Information Technology\)](#), with the aim to develop advanced products to burn graphics or text and infinitely small and unalterable.

It looks like the micro-images are stored in very high density on a "wafer" that is 200 mm in diameter.

Also in the August issue was an article on [NanoArk](#) (Rochester, NY), which is using waferfiche(TM) for long-term storage. The NanoArk web site says:

In the Waferfiche(TM) technology, data is stored in such a way that it is visible to the human eye with or without magnification. The data from print, digital or any other media is converted to images as a first step. These images, with the help of photolithography tools and fabrication techniques are then imprinted and etched on silicon wafers. The use of silicon makes the information temporarily resistant to high temperatures (up to 400°C) and water, ensuring longevity, which is very useful in preserving documents. In this technique, since the stored data is not processed or digitized before storing, the data is stored for long periods without any loss of data over time. The added advantage of the new technique is retrieval of the data can be as simple and straightforward as magnifying the image on the silicon wafers thereby eliminating the need for a computer. This feature enables archival of data in a technology free environment. Also depending upon the semiconductor fabrication technique used (smallest feature size in the order of 200 nm, 100nm, 90 nm or below), nano-scale images can be imprinted thereby making it possible to store large amount of data on a single silicon wafer.

I suspect -- but could be wrong -- that ARnanO and NanoArk are working on similar solutions for long-term storage of information. We know that microforms are the best solution for long-term preservation of information. These wafers seem to allow for much more information to be stored and on a medium that is less susceptible to environmental changes.

Finally, in the September issue is an article on **book scanners** produced by [Metis Systems](#). The company has two book scanner: DRS 5070 and DRS A1 Plus. From what I can tell, the operator must turn the pages. They also make a large format scanner (DRS 2AO).

For those of you in Europe, who perhaps have read these articles OR are familiar with the companies mentioned, please leave a comment if you can provide more information. Thank you!

Technorati tags: [Digitization](#), [Book](#), [Microforms](#)

Posted by Jill Hurst-Wahl at 9:55 AM 0 comments [Links to this post](#) 

Labels: [Book Digitization](#), [Digitization](#), [Microforms](#)

Tuesday, November 04, 2008

[Video: Brewster Kahle: A digital library, free to the world](#)

This 20-minute video on the TED.com web site became available

in September 2008, but the talk is from December 2007. At 7 minutes into the video, Kahle talks about using a \$100 laptop as an ebook reader. At 9 minutes, he talks about digitizing books and about using robotic book scanners then about creating their own.

He also talks about capturing and storing film, audio, video (e.g., TV), and software at the Internet Archive.

Kahle is always informative and always make a compelling case. To me -- although he is not always in the limelight -- he is one of the people who is moving book digitization through his actions as well as his words.

Below you may only see two "boxes" asking you to select the video player that you want to use. (At least, that is what I see.) Pick a player and the video should begin.

Technorati tags: [Digitization](#), [Book](#)

Posted by Jill Hurst-Wahl at 7:13 PM 0 comments [Links to this post](#) 

Labels: [Book](#) [Digitization](#)

[The Vietnam Center and Archive](#) [\(and John McCain\)](#)

For those interested in the Vietnam War, [this](#) is a wonderful archive and digital collection. Interestingly, it also contains materials that mention John McCain.

Technorati tag: [Digitization](#)

Posted by Jill Hurst-Wahl at 10:50 AM 0 comments [Links to this post](#) 

Labels: [Digitization](#)

U.S. Presidential Election

Today is the first Tuesday in November, which means it is time for the general election in the United States. Every four years, we elect a president. Below is a short video about the presidential election process.

If you are following people on [Twitter](#) or [FriendFeed](#), for example, it is likely that you are seeing many voting-related posts. This election is a big deal...perhaps the most important election since Kennedy-Nixon in 1960 (Kennedy won). This election has captivated people and it is expected to draw many more people to the polls than normal. What is most interesting about this election is that people have become very involved in this election -- there is more talk, more analysis from ordinary people, etc. "We the people" understand that we have a role in deciding how government is run and who our leader is.

I am sure that this election will be analyzed for years to come. Seth Godin has already posted his analysis -- [Marketing lessons from the US election](#) -- which is worth reading.

I suspect that people will get some work done today in the U.S., but don't expect miracles. We're focused on the election and who our next President of the United States (POTUS) will be!

Posted by Jill Hurst-Wahl at 10:29 AM 0 comments [Links to this post](#) 



Sunday, November 02, 2008

Jill's November schedule

It's November and a time when the weather really changes here in New York State. With



that in mind, this Thursday, I will be driving to Saratoga Spring, NY for the [New York Library Association](#) Annual Conference. Thursday night, I will attend the [Syracuse University iSchool](#) alumni reception. Friday,

I'll be at the conference all day and will be part of a panel called "Meet the Bloggers" (4 p.m.) where I and two others will talk about our experience as library bloggers and provide tips to the audience. If you are attending the conference, please feel free to stop me and say "hi". Who knows...maybe we can grab a cup of coffee?!

I do have two other business trips planned for later in the month, so you might see in an airport to/from Texas (Nov. 12 - 14) or in New York City on Nov. 21 for the [Regents Advisory Council on Libraries](#) meeting. If we do cross paths, remember -- say "hi"!

My December and 2009 schedule is on the left side of this blog. Please take a look. For more information on any of the events on my calendar, please contact the sponsoring organization or [me](#). If you would me to speak at your event, please give me a [shout](#).

Technorati tag: [Social Networking Tools](#)

Posted by Jill Hurst-Wahl at 8:12 PM 0 comments [Links to this post](#) 

Labels: [JHW](#), [Social Networking](#)

[Older Posts](#)

Subscribe to: [Posts \(Atom\)](#)