



Digital libraries: A bit of history

some historical perspectives
and precursors



This work is licensed under a

[Creative Commons Attribution-NonCommercial-ShareAlike 3.0 United States License](https://creativecommons.org/licenses/by-nc-sa/3.0/)



DL history? What?

is it a bit presumptuous to talk about DLib *history*?

- DLs are around for real only for about a decade or two
- But many relevant ideas were around for a long time
- And libraries and technological change were connected forever



ToC

- Some connections between libraries and technology → shown graphically
- Some ideas influencing DLib concepts → Vannevar Bush
J.C.R. Licklider
- A few major related developments → Information retrieval
OPACs; Networks; Web
- DLib research & practice → a few generalizations

Libraries & technology

- Already mentioned:
 - technological imperative: libraries always depended, among others, on technology
 - technology keeps transforming libraries
- Contemporary information & communication technology is doing the same thing
 - in a **BIG** way
- And because of it many **new** players that were not in the library business are entering into the arena of digital libraries
- Technology: Tails that wags the dog

Then: Oldest surviving “book”



**18th dynasty, ancient
Egypt, circa 1550 B.C.**

Title: *Word of Khakheperreseneb.*

Content: a literary discourse concerning
personal and social chaos

Technology: wooden board,
layer of gypsum; hole for a
cord to hang on a peg

Now: Digitized or borne digital

New Jersey Digital Highway

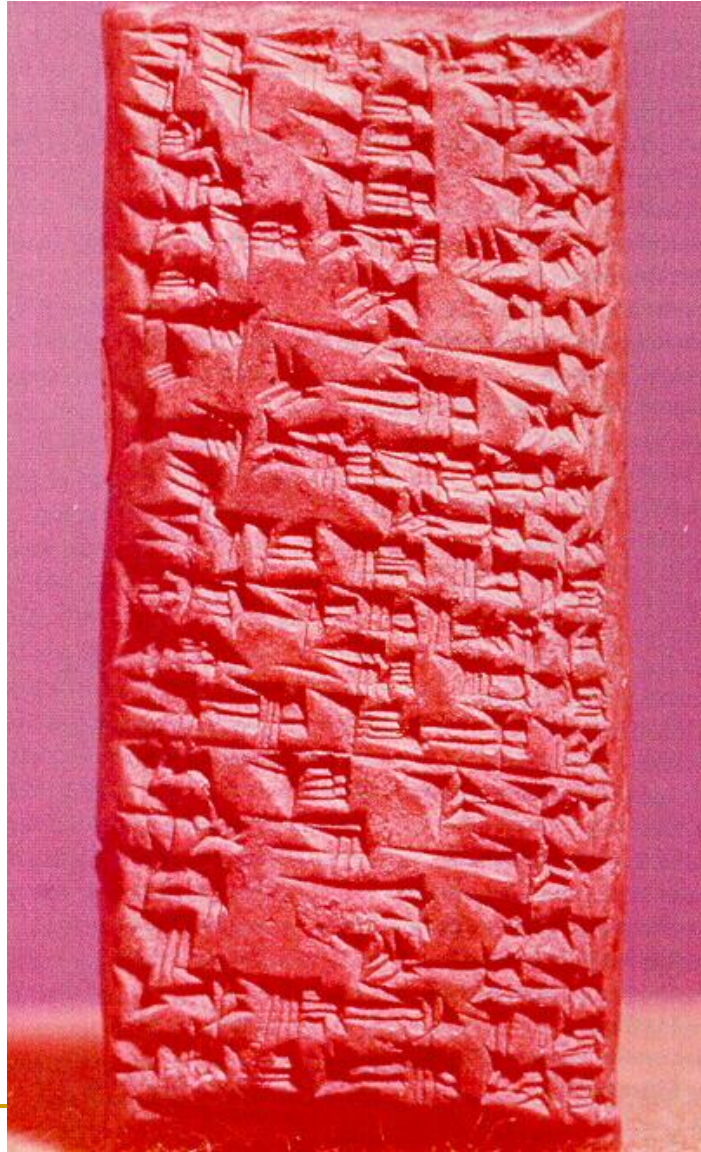
“...designed for the in-depth study of New Jersey history.”
 Developed at [Scholarly Communication Center](#),
 Rutgers University Libraries

The screenshot shows the New Jersey Digital Highway website. At the top right is the Rutgers logo. The main header features the NJ DH logo with the tagline "Where history, culture, and learning merge" and a row of historical images. Below this is a navigation bar with links for [Home], Librarians & Curators, Educators, and Students. The main content area is divided into several sections:

- About NJDH:** Includes links for About Us, Site Policies, Links for Everyone, Discover New Jersey, Finding Books, Cultural Heritage Orgs, Immigration Sources, Preserve Your Family History, Research Family History, Oral History, and Digital Exhibitions.
- Search:** A search bar with a "Search NJDH" button and a link to "Advanced Search". Below it, it says "You are searching all content."
- Spotlight:** A section highlighting featured content. It includes a post titled "Bayonne Public Library partners with Rutgers University to digitize historic maps" with a "View" link, and another post titled "Botto House Silk Walk Project 2013" with a "View" link.
- Right Sidebar:** Features the "NEW JERSEY HISTORY E-JOURNAL" and a "News & Updates" section. It includes a post about the Bayonne Public Library partnership and another about the New Jersey Governor's conference on November 16, 2013.

At the bottom left of the website interface, there are two promotional boxes: "E-ZStart" for digital projects and "From Page2Pixel" for a blog about the Digital Dark Age.

Then: A catalog card



600 B.C

**Niniveh,
Babylonia**

.

Technology:
Clay –
glazed, baked

Now: OPACs (Online Public Access Catalogs)



At Rutgers University Library & many other libraries:

Technology:
digital MARC records,
computers,
online access

Then: most famous library



**Alexandria
library**

~330- 1 B.C.

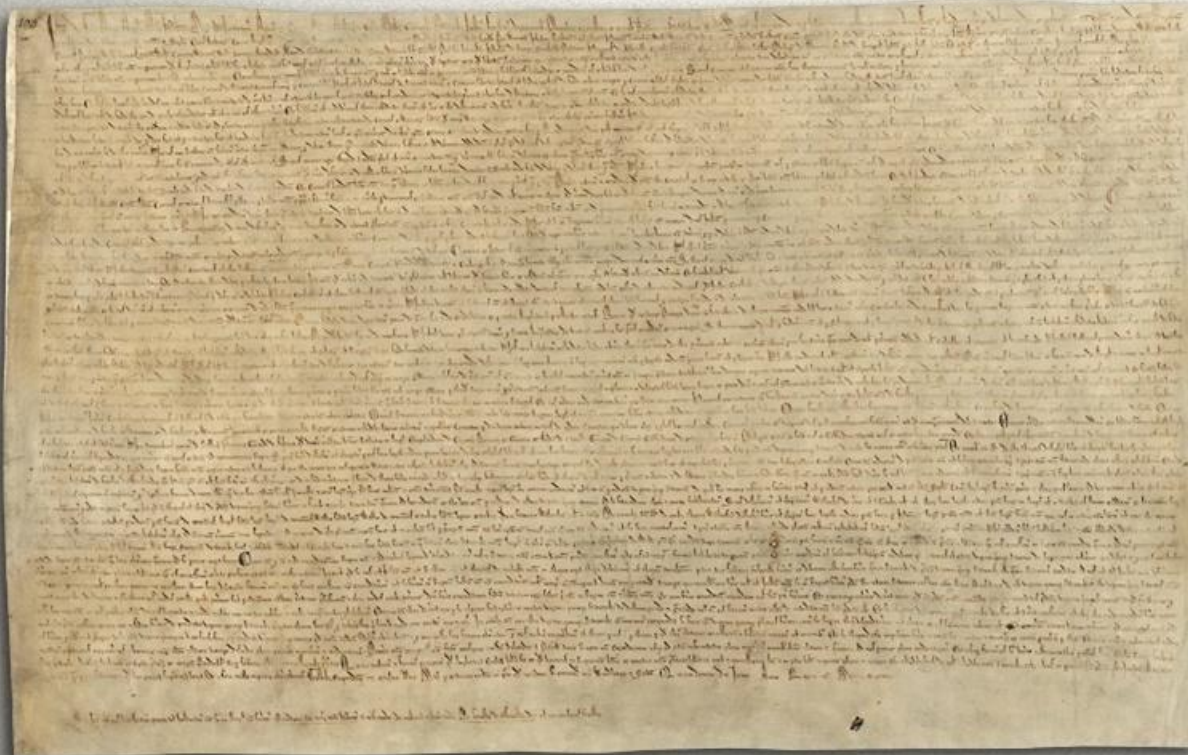
(depiction)

Technology:
papyrus
scrolls

Now: digitized rare documents in collections in libraries

LIBRARY
HS1LIBBTREASURES IN FULL
MAGNA CARTAFrom: British Library – [Turning the pages](#)

Magna Carta - The Great Charter



Then: Creating & preserving books



Scriptorium
- middle ages

Technology:
pen, paper, ink;
copying

Followed by: printing books



Johannes
Gutenberg
1397-1468



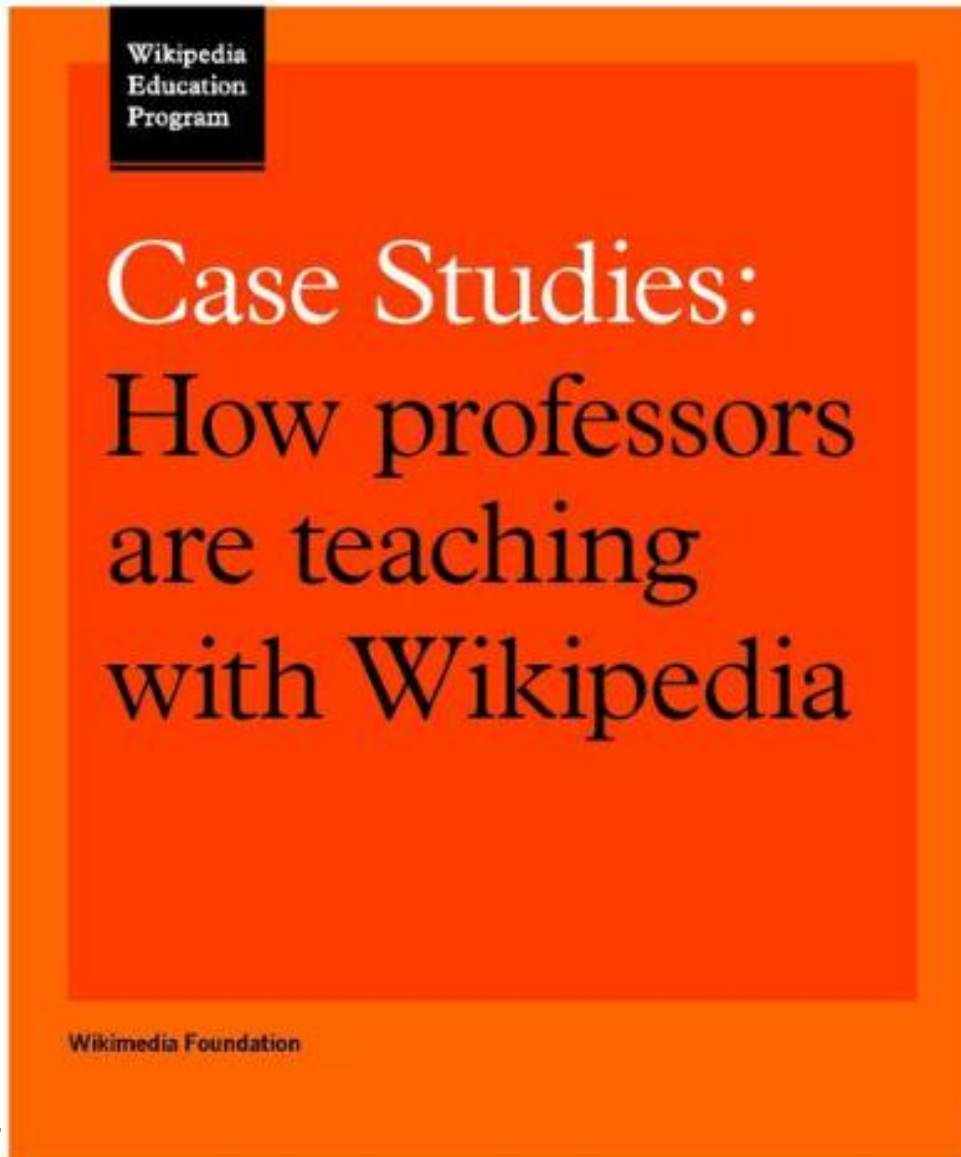
Technology: combined four skeins: ink, paper, movable type & printing press
Created a revolution – role in Renaissance & libraries
Basic idea still operational

& then by born digital books

From:

[Wikimedia
Outreach](#)

(when there, click
on the image to
download)



Go to page



A remarkable thing (point repeated & underscored)

- Basically, libraries deal with collecting, organizing, preserving & providing access to human knowledge records
- This role is constant across civilizations, history, time, geography
- And it is constant regardless of enormous technological changes
- Digital libraries are not an exception
 - as to role they have strong & long historical roots

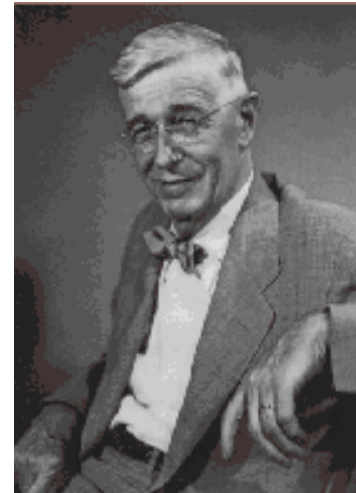
Vannevar Bush

Problems addressed

- Head of US scientific effort during WWII
- In 1945 published a highly influential article:

As We May Think (1945):

- Defined problem as “... ***the massive task of making more accessible of a bewildering store of knowledge.***”
- Problem still with us & growing – “information explosion”
 - joined now also with “communication explosion”

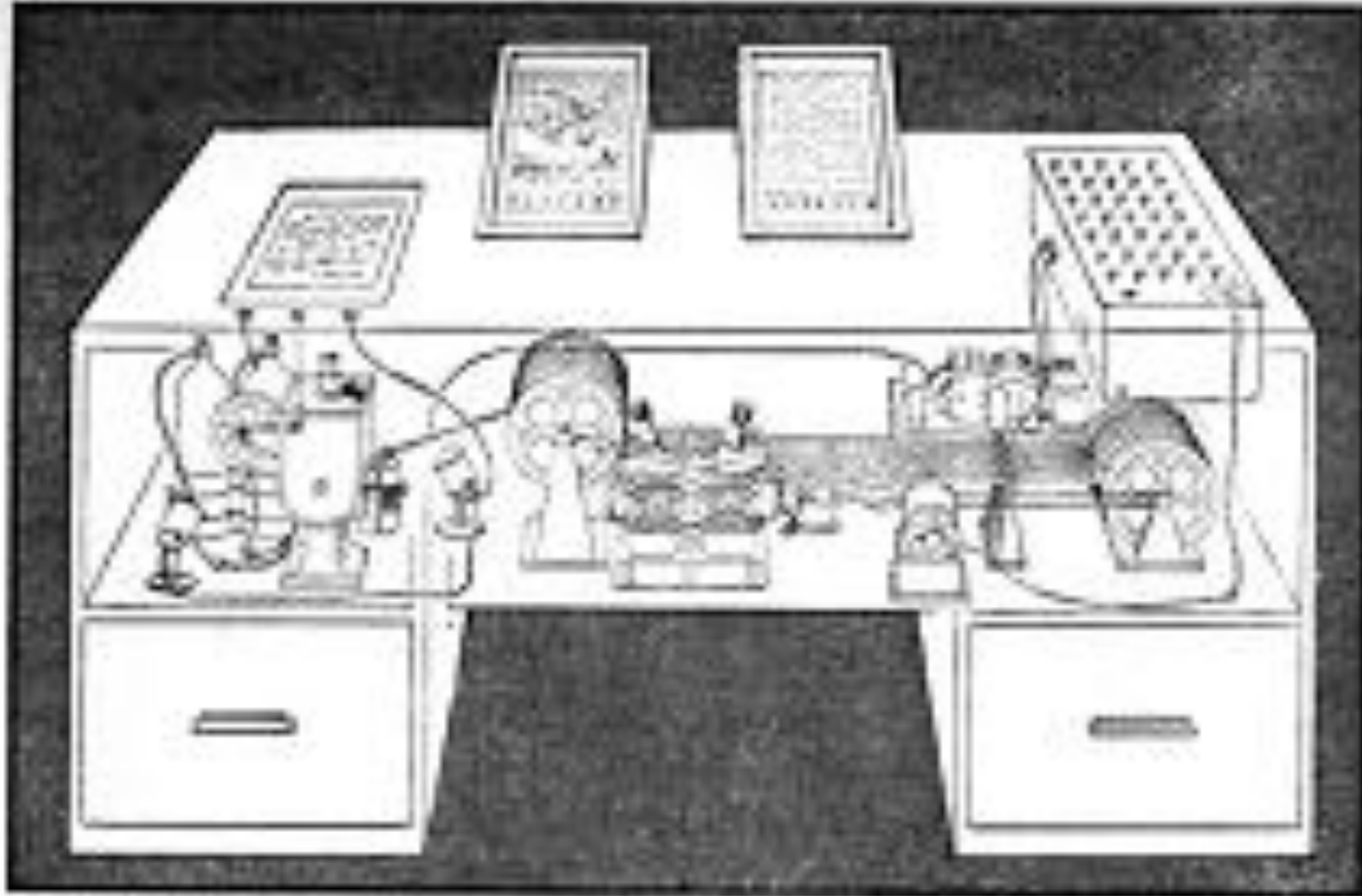


Vannevar Bush
1890-1974

... solution

- Bush suggested a machine: *“Memex ... association of ideas ... duplicate mental processes artificially.”*
 - Memex-like machine has yet to be built
- But goal constant: Technological fix to problem
- This influenced many efforts
 - funding from US government
 - including decades later DLib projects
- Bush was not the only one with such idea
 - but was most influential

Bush's memex



Science fiction, of course.

Note use of microfiche

As visualized by Life Magazine in 1945.

J.C.R. Licklider

An early visionary & leader

- Worked on computer networks
- Influential works:
 - The Computer as a Communication Device* (1968)
 - Libraries of the future* (1965) MIT Press
- Suggested that future libraries
 - will be highly innovative
 - different in structure, processing & access
 - all through applications of technology
- Affected thinking on the suggested vision
 - to a large degree turned into reality



J.C.R. Licklider
1915-1990

Information Retrieval (IR) – definition & objective

Mooers coined term “information retrieval” & defined:

“ IR: ... intellectual aspects of description of information,
... search, ... & systems, machines...”

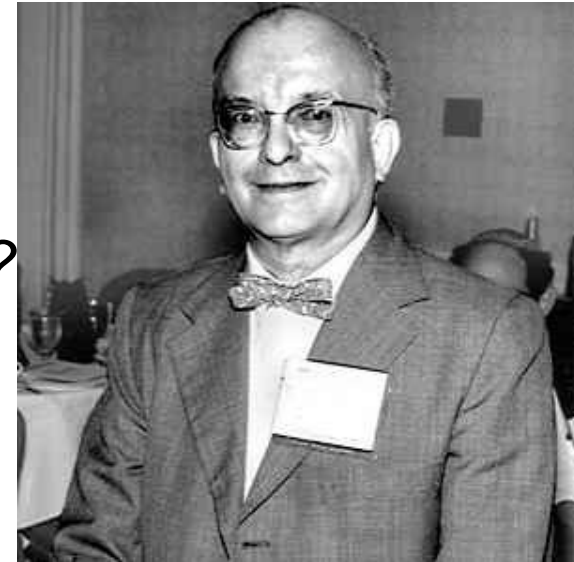
[Calvin Mooers](#), 1951

- Objective: How to provide users with relevant information **effectively**?

For that objective:

1. How to organize information *intellectually*?
2. How to specify the search & interaction *intellectually*?
3. What techniques & systems to use *effectively*?

Basically, same objectives today



Calvin Mooers
1919-1994

Hans Peter Luhn

Early IR developments

- H.P. Luhn, at IBM, developed many IR techniques for computers
 - automatic production of concordances
 - automatic production of KWIC (Keyword in Context) indexes
 - KWOC (out of context) – permuted indexes
 - automatic abstracting, summary
- Basic idea: computers doing representations
 - indexing, concordances ...
 - previously done by humans



[Hans Peter Luhn](#)

1896-1964

Gerard Salton

Pioneered full text IR

- Salton (at Harvard & Cornell) was a leader in the idea of full text information retrieval
 - automatic indexing & searching of full text of articles & other texts
 - done today routinely
- Gerry (as fondly known) developed, among others:
 - IR vector model for indexing
 - SMART system for IR evaluation
 - & brought up many of today's leaders in the field



[Gerard Salton](#)
1927-1995

Contemporary IR

- Spread globally
 - e.g. major IR research communities emerged in China, Korea, Singapore
- Branched outside of information science - “everybody does information retrieval”
 - search engines, data mining, natural language processing, artificial intelligence, computer graphics ...
- Incorporated in
 - DLib indexes & databases, (such as [ACM Digital Library](#))
 - digital libraries as searching mechanisms

Henriette Avram

Machine Readable Cataloging

- MARC developed by Henriette Avram at Library of Congress, 1960s
 - for machines to “understand” different fields in cataloging records
 - what is a “title,” “author,” ...
 - in various forms adapted & used globally
 - even for Web metadata (Dublin Core)
 - originally used for cataloging & by librarians only
 - later adapted (with difficulty) for searching by users – OPACs
- A basic component of digital libraries



Henriette Avram
1919-2006

Fred Kilgour

Library sharing networks

- Fred Kilgour started [OCLC](#) in 1967 in Dublin, Ohio (BTW, thus Dublin Core)
 - (originally **O**hio **C**ollege **L**ibrary **C**enter)
 - libraries sharing cataloging resources
 - started with printed catalog cards from MARC records
 - now a worldwide online member-network of 23,000 libraries, museums & archives from 170 countries & territories
- Resource sharing idea was further developed & reflected in Library Consortia basic to many digital libraries
 - e.g RUL participate in [VALE](#)



Fred Kilgour

1914-2006

Library automation

- Started in early 1960s
 - first as manipulation in cataloging & MARC records
 - expanded to other library functions e.g. collection management, circulation, access ...
 - today nobody calls it “library automation” any more
 - but the activity involves software, hardware & networks in all library processes & aspects
- Early software developed by libraries – most failed
- Now quite a few software companies (many large) compete for library contracts & lease software
- Essential for digital library operations

Tim Berners-Lee

And then came WWW

- Library (& other) online networks existed long before the Web
- Berners-Lee (then at CERN in Switzerland) developed the Web originally for science communication (in 1993)
 - it took off like wildfire – fastest & widest spread technology in history
 - affected development of browsers
 - first Mosaic that morphed into Netscape, and later Internet Explorer, Mozilla Firefox and others
- Basic for any and all digital libraries



Tim Berners-Lee
now Sir Tim Berners-Lee

Digital library: the name

(source: Michael Lesk)

- Until 1993 the normal label was “electronic library”
 - Berners-Lee called it “virtual library”
 - In the UK “elibraries”
- In 1990’s Tom Kalil and Mike Nelson were Clinton-Gore advisers who liked the idea of access to information. Mike Nelson invented the new name, Kalil helped get money for it
- Al Gore was a strong supporter of digital information, which was both good & bad (it politicized the issue)



Tom Kalil



Digital library research (more in lecture on *Research*)

- In the US projects under Digital Library Initiatives (DLI)
 - funded by NSF & other participating agencies
 - DLI 1 ran 1994-1998; DLI 2 1999-2004
 - National Science Digital Libraries (1999-to date)
 - many prototypes & projects in a variety of subjects
 - but by now government research funding for DLlibs from NSF & related agencies mostly ceased
- There were numerous other DLib R&D programs in the US & other countries

Larry Page & Sergie Brin

Unintended byproduct of DLI 1

- Page & Brin (later founders of Google) were graduate students of Hector Garcia-Molina at Stanford who built a system called “Backrub” that rated web pages by the count of pointers to them. This became Google
- In their famous paper explaining basics of the Google approach, Page & Brin in the acknowledgment listed support of the Stanford DLI1 project
- [Origins of Google](#) - an NSF description



Larry Page



Sergei Brin

DL projects in practice

(more in lecture on *Diversity*)

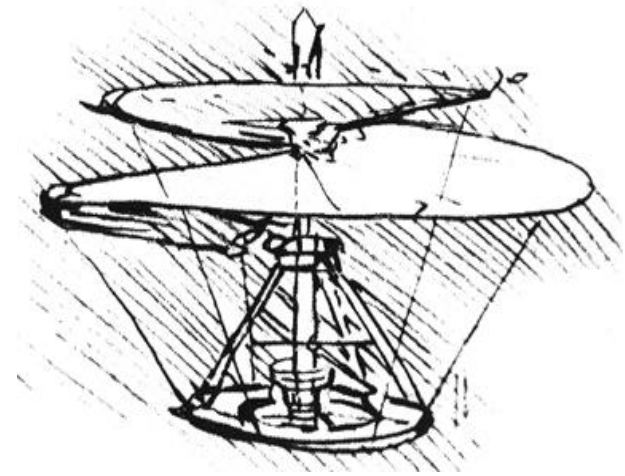
- Heavily oriented toward institutions & their missions
 - in libraries, but also others
 - museums, societies, government, commercial
 - they come in many varieties
- Spread globally, including digitization
- Spending increasing significantly
 - often a trade-off for other resources

In conclusion

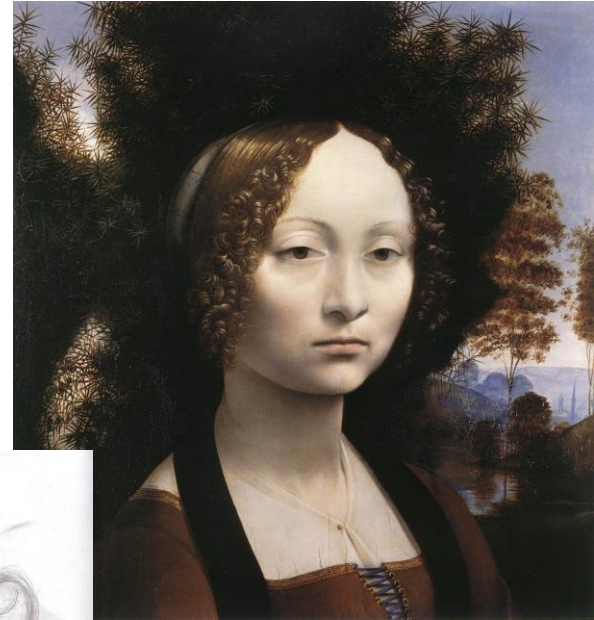
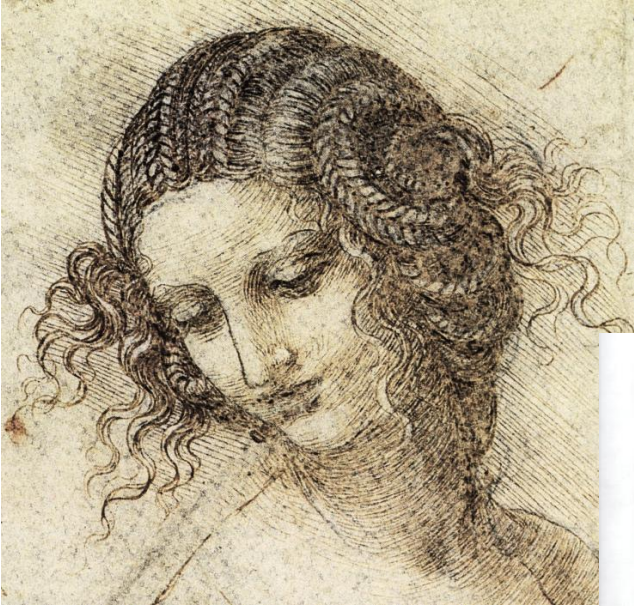
- DLib history is too short to really talk history
- But DLIBs follow many ideas from the past
 - and the trajectory of technological developments in libraries over ages
- In a short period DLIBs are transforming libraries in significant ways
 - and their users & use
 - and requirements for professional competencies - **you**
 - and LIS education
- The future is here

Images

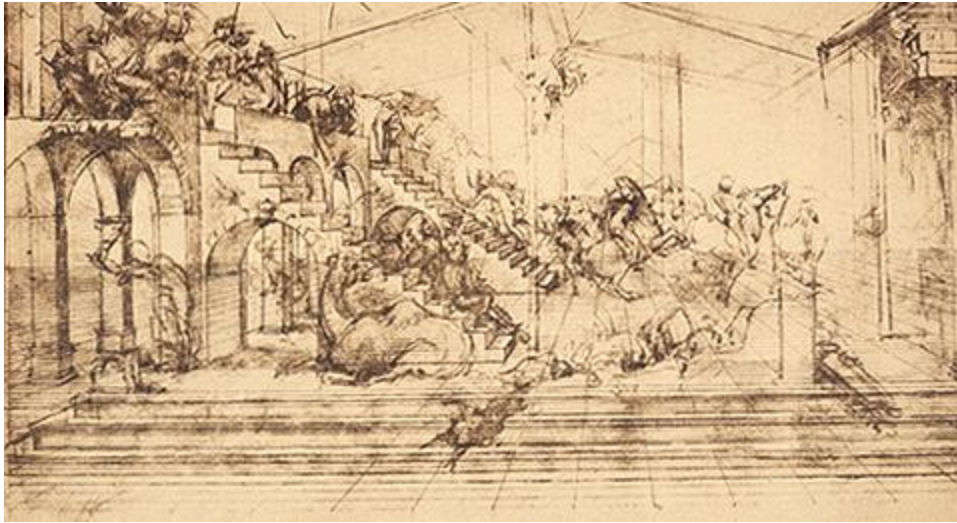
taken from [Museum of Science, Boston](#) and [Webmuseum, Paris](#)



images ...



images ...



and of course...



P.S. a nice site

Poem by Emily Dickenson:

In a library

Who will write a poem:

In a digital library?

