

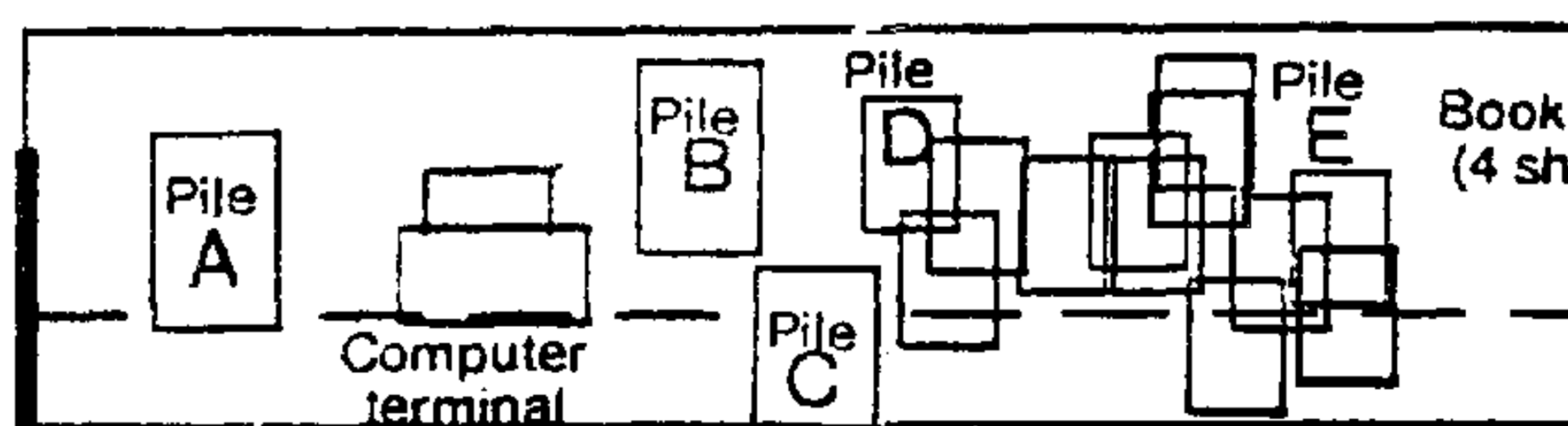
CSCI 1377

Tools for Thought

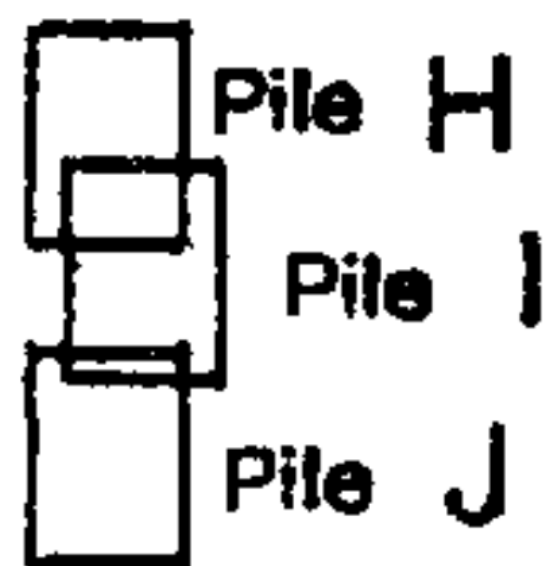
Hypertext II

The World Wide Web

“Trying to fix HTML is like trying to graft arms and legs onto [a] hamburger.” — Ted Nelson (1999)



Bulletin board



Table



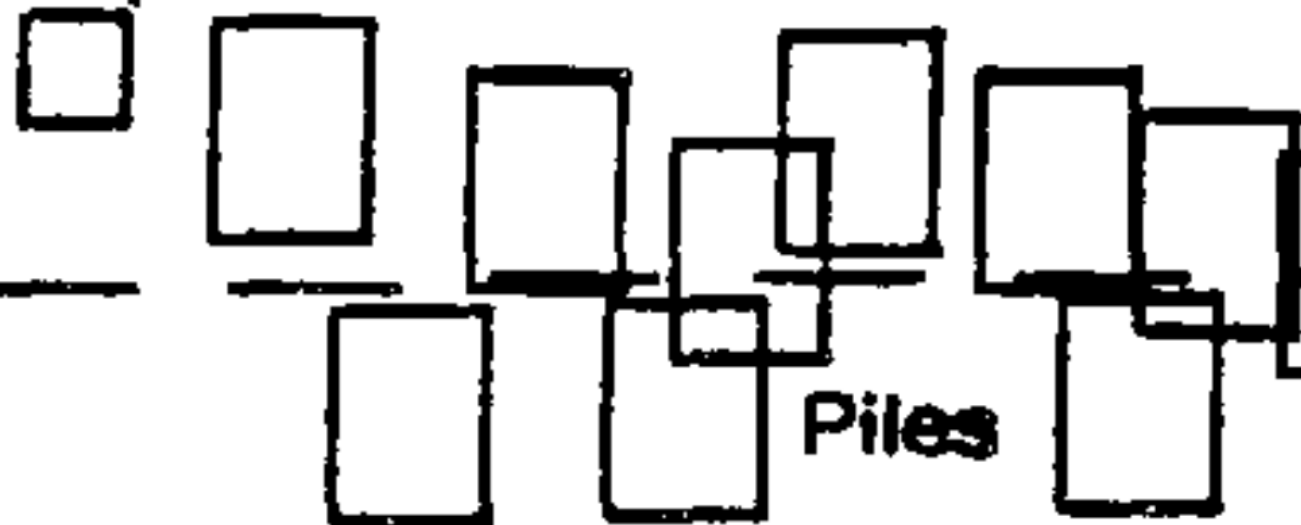
Blackboard

Boxes



Desk

Telephone



Bookshelves
(3 shelves)

Bulletin board



Design goals for Licklider's library of the future

- Handle both documents and facts.
- Permit several different categories of input, ranging from authority-approved formal contributions (e.g., papers accepted by recognized journals) to informal notes and comments.
- Converse or negotiate with the user while he formulates his requests and while responding to them.
- Adjust itself to the level of sophistication of the individual user, providing terse, streamlined modes for experienced users working in their fields of expertness, and functioning as a teaching machine to guide and improve the efforts of neophytes.
- Facilitate its own further development by providing tool-building languages and techniques to users and preserving the tools they devise and by recording measures of its own performance and adapting in such a way as to maximize the measures.

Design goals for Licklider's library of the future

- Provide the flexibility, legibility, and convenience of the printed page at input and output and, at the same time, the dynamic quality and immediate responsiveness of the oscilloscope screen and light pen.
- Present flexible, wide-band interfaces to other systems, such as research systems in laboratories, information-acquisition systems in government, and application systems in business and industry.
- Essentially eliminate publication lag.
- Evidence neither the ponderousness now associated with overcentralization nor the confusing diversity and provinciality now associated with highly distributed systems.

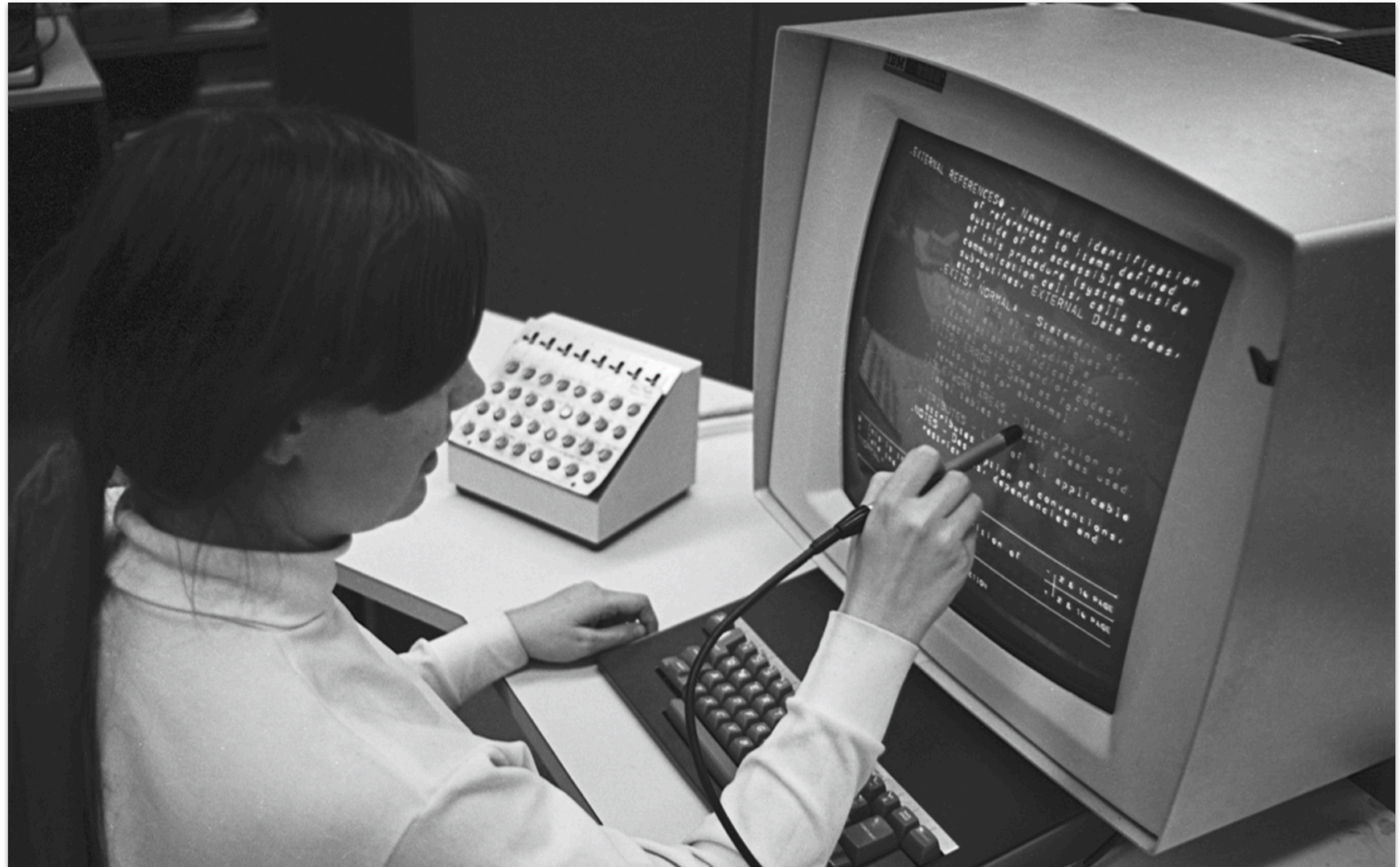
“Hypertext is the combination of natural-language text with the computer’s capacities for interactive, branching or dynamic display, *when explicitly used as a medium.*”

Ted Nelson, “Getting It Out of Our System” (1967)

Our very own Andy created a prototype hypertext word processor in 1967



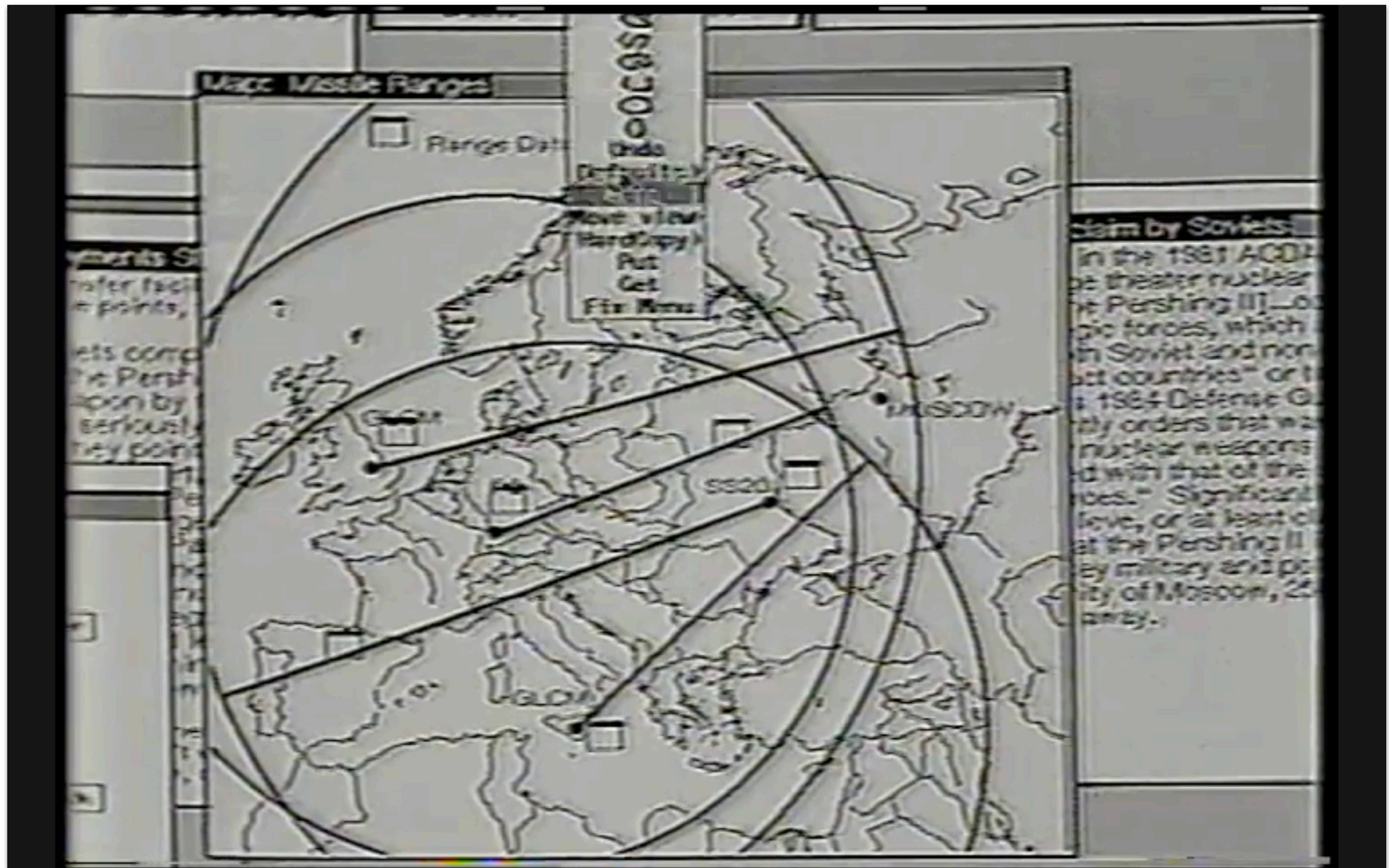
Hypertext
Editing
System
(HES)



Check out 1951V!

<https://cs1951v-2023.vercel.app>

Xerox PARC NoteCards



<https://youtu.be/YvB2xU6xjfY?si=YlWES3wB-e50TbGn&t=549>

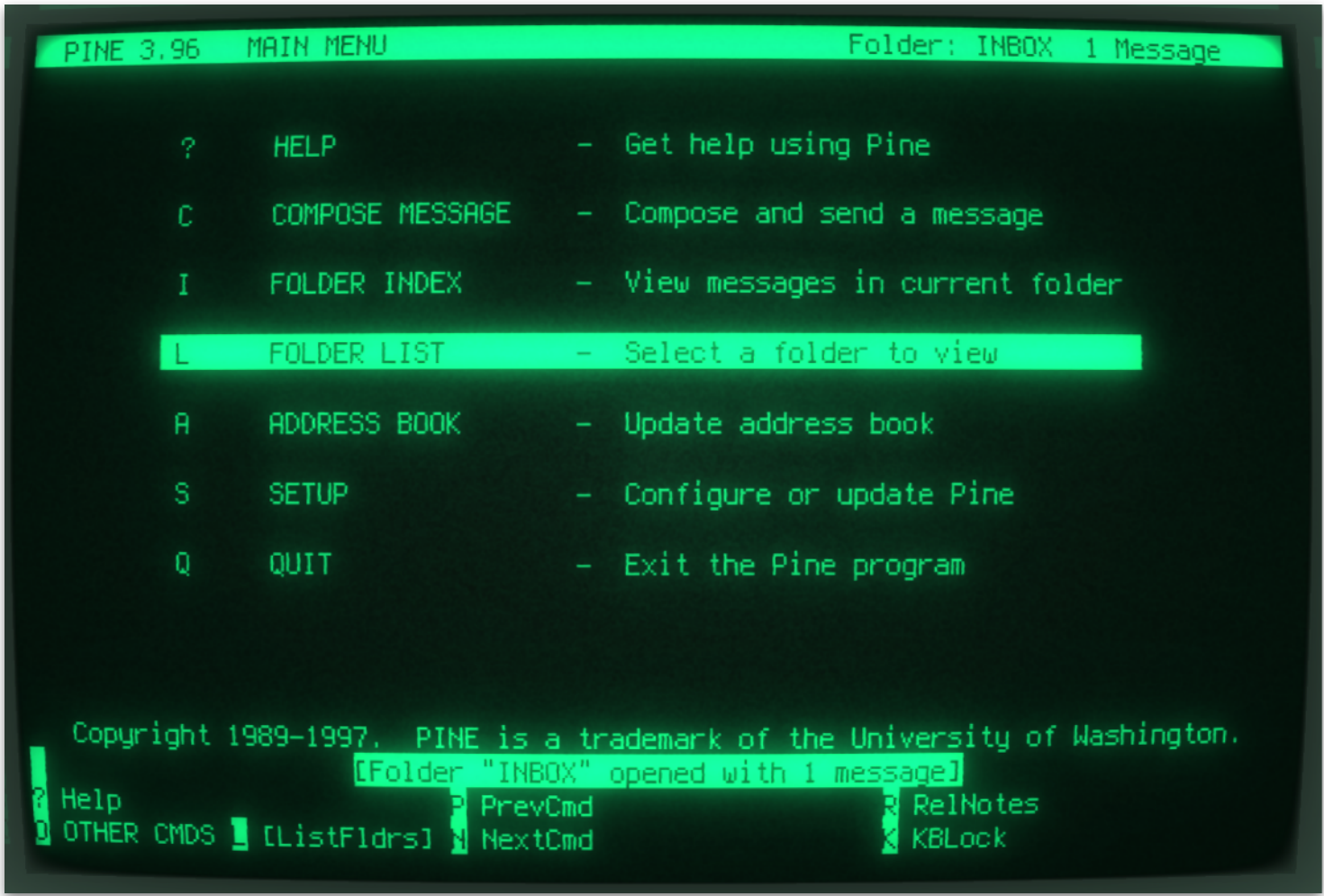
Users needed more features out of hypertext

- Search and query in a hypermedia network
 - “Access to information [...] requires query-based access to complement navigation”
- Composites—augmenting the basic node and link model
 - “a way of representing and dealing with groups of nodes and links as unique entities separate from their components.”
- Virtual structures for dealing with changing information
 - “As the user’s knowledge of the information space evolves, previous organizational commitments (titles and filing categories) become obsolete”
- Versioning
 - “manipulate a history of changes [...] explore several alternative configurations”

Meanwhile, on the Internet...

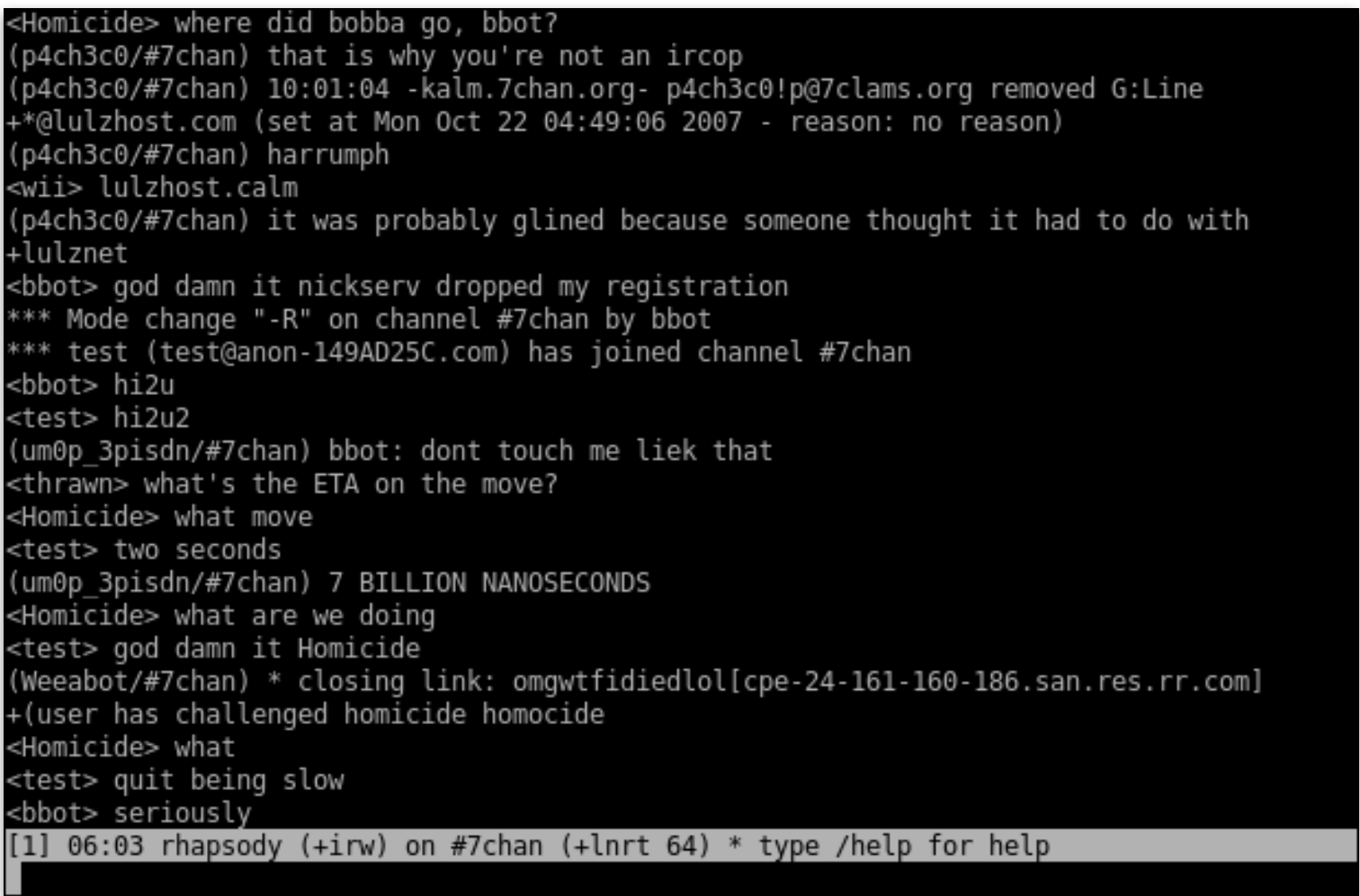
The Internet existed before the Web

Majenko, CC0, via Wikimedia Commons



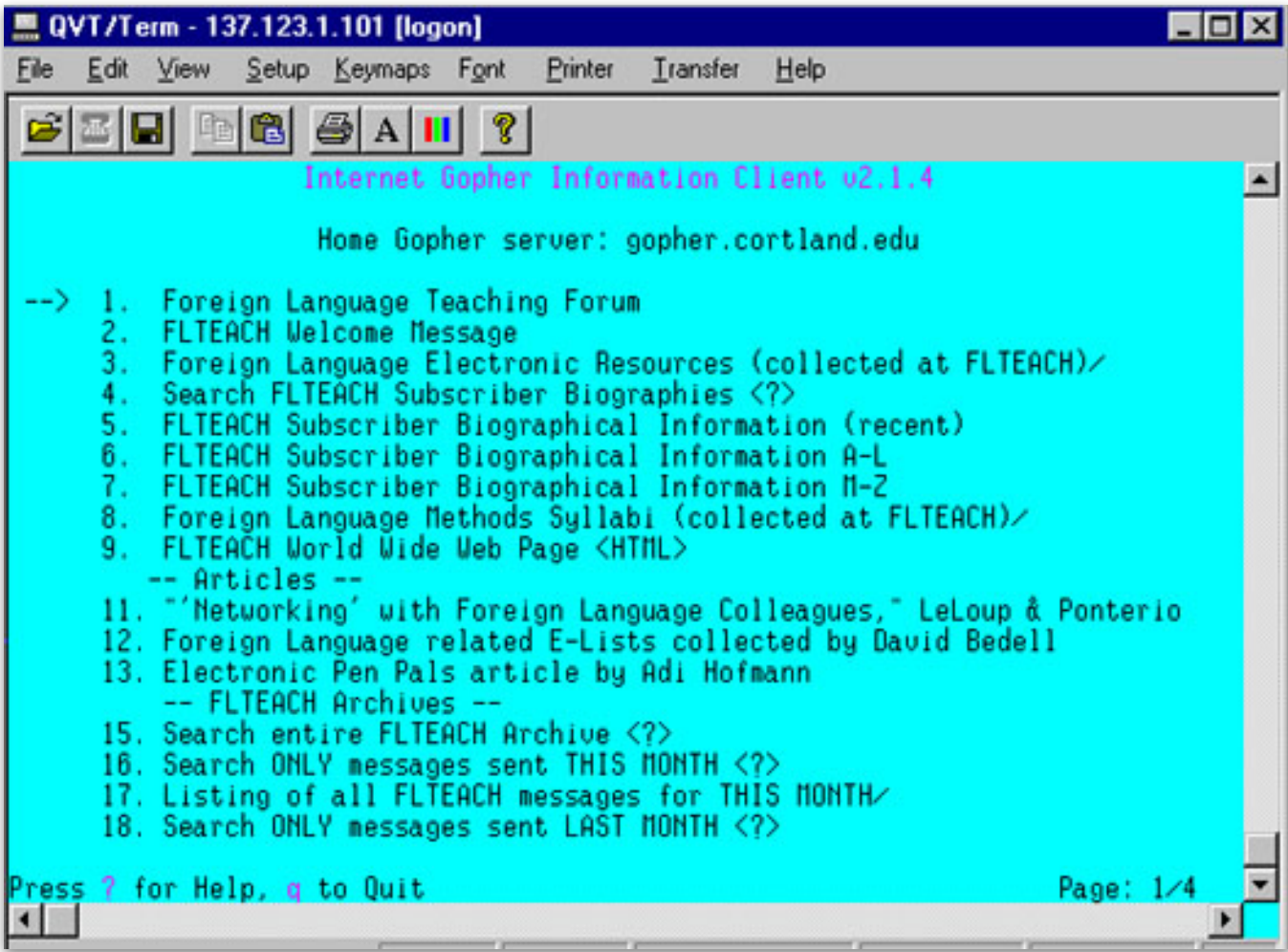
Email over SMTP (1983)

Rhapsody Scarlet, BSD, via Wikimedia Commons



IRC (1988)

via SUNY Cortland



Gopher (1991)

The World Wide Web started at CERN

- CERN needed to organize a lot of technical knowledge
 - Turnover: “information is constantly being lost”
 - Churn: “CERN is constantly changing as new ideas are produced, as new technology becomes available, and in order to get around unforeseen technical problem”
 - Questions: “Where is this module used? Who wrote this code? Where do they work? Which systems depend on this device? What documents refer to this one?”
- Existing hierarchical / tree-like organizations were insufficient
 - e.g. `HELP` → `COMPILER` → `SOURCE_FORMAT` → `PRAGMAS` → `DEFAULTS`
- Keywords were also insufficient: “two people never chose the same keywords”

WWW influenced by CERN's requirements

- **“Remote access across networks:** CERN is distributed, and access from remote machines is essential.
- **Bells and Whistles:** Storage of ASCII text, and display on 24x80 screens, is in the short term sufficient, and essential.
- **Non-Centralisation:** A new system must allow existing systems to be linked together without requiring any central control
- **Private links:** One must be able to add one's own private links to and from public information. One must also be able to annotate links, as well as nodes, privately.”

WWW provides hypertext infrastructure

Hypertext Transfer Protocol (HTTP)

```
GET /csci-1377-f26/ HTTP/2
Host: cel.cs.brown.edu
User-Agent: Mozilla/5.0 [...]
Connection: keep-alive
```

```
HTTP/2 200
server: GitHub.com
content-type: text/html;
charset=utf-8
content-length: 14043
last-modified: Wed, 30 Sep
2026 20:00:25 GMT
```

Hypertext Markup Language (HTML)

```
<!DOCTYPE html>
<html>
  <head>
    <title>
      Tools for Thought
    </title>
  </head>
  <body>
    <h1>Course Notes</h1>
    <a href="/lectures">
      Lectures
    </a>
  </body>
</html>
```

Uniform Resource Locator (URL)

```
http://
cel.cs.brown.edu/
csci-1377-f26/
index.html
#lectures
?visited=true
```


Hypertext design dimensions of the Web

- **Nodes:** web pages
 - Pages have sub-nodes based on HTML structure, e.g. by ID or tag name
- **Links:** anchors via `<a>` tags
 - URLs link to pages or page fragments
 - Some webapps provide custom fragments, e.g., GitHub line ranges
 - Links are **unidirectional**, so the linked page does not know about being linked
 - `Referer` (sp.) header added in 1996 to address this
- **Content:** arbitrary applications without OS access
 - Cascading Style Sheets (1996) provides rich 2D UI layout and rendering
 - JavaScript (1995) provides interaction and computation
 - WebAssembly (2015) provides high-performance execution

Hypertext design dimensions of the Web

- **Navigation:** depends on browser and webpage
 - Browser: back/forward buttons, local browsing history, bookmarks
 - Webpage: can provide breadcrumbs
- **Finding:** depends on browser, web page, search engine
 - Browser: Ctrl+F to find text on the current page, bookmarks for browsing old pages
 - Webpage: can provide table of contents, sitemap
 - Search engine: use keywords (Google) or natural language (LLMs) to find pages
- **Viewing:** depends on OS, browser, webpage
 - OS: window management for viewing pages concurrently
 - Browser: tabs for multiple pages, tab groups more recently
 - Webpage: can provide app-specific support for viewing previews, multiple docs

Hypertext, search engines, and profit motives: is your tool for thought adversarially robust?

- Lewandowski et al. 2021: ~80% of results for popular Google queries show signs of non-trivial search engine optimization
- Bevendorff et al. 2024:
 - “higher-ranked pages are on average more optimized, more monetized with affiliate marketing, and they show signs of lower text quality”
 - “search engines measurably target SEO and affiliate spam with their ranker updates. Google’s updates in particular are having a noticeable, yet mostly short-lived, effect.”
- Let’s hope our LLMs fare better!

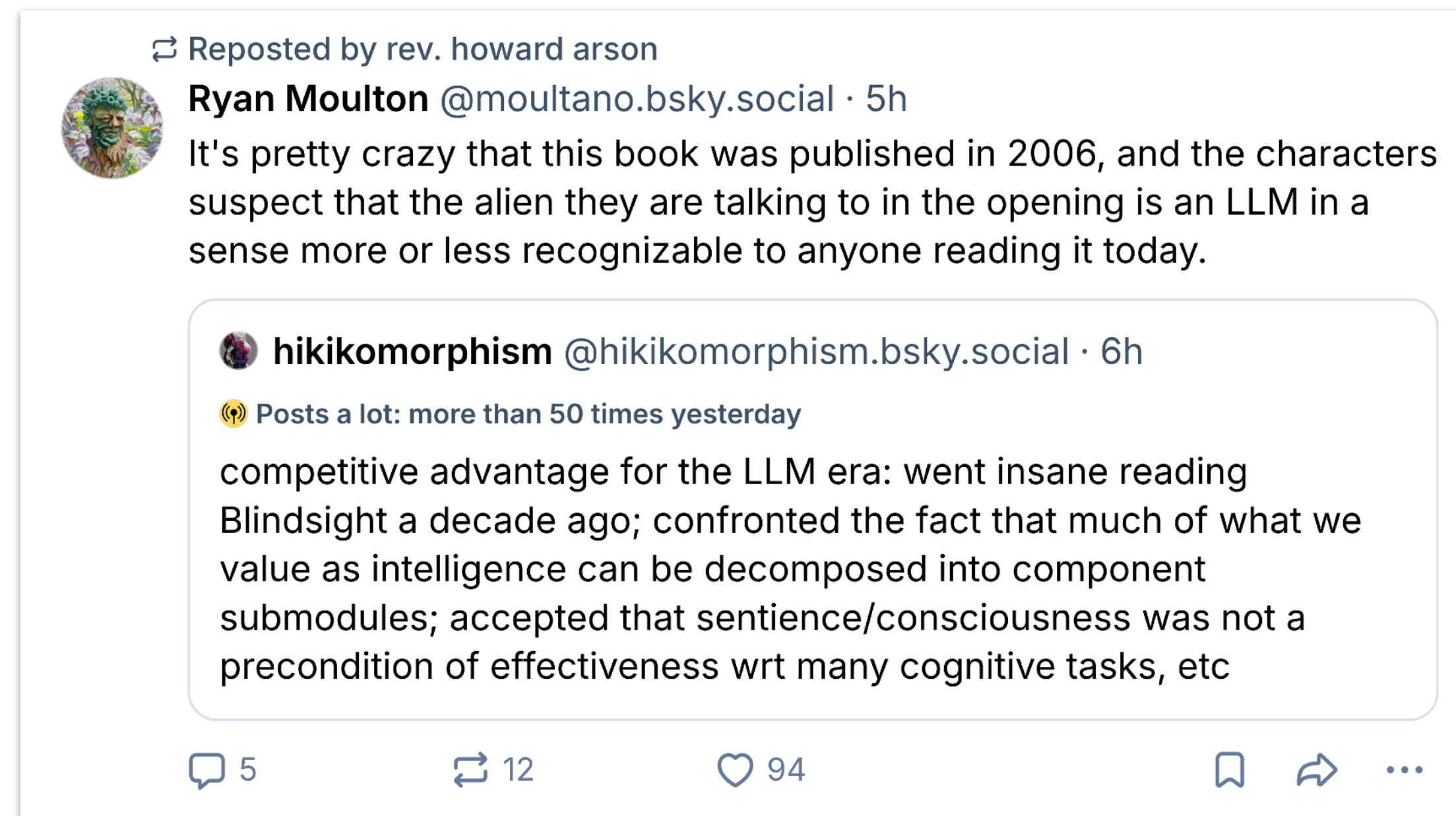
Hypertext design dimensions of the Web

- **Authoring:** requires a web site and a server
 - WYSIWYG web builders like Squarespace have a long history, e.g. Adobe Dreamweaver (1997) and NaviPress (1994)
- **Collaboration:** entirely application-driven
 - Operational Transforms and Conflict-Free Replicated Datatypes provide the basis for modern collaborative editing
- **Annotation:** entirely application-driven
 - E.g., Google Docs / Notion / Overleaf have bespoke support
 - Web Annotation Protocol (2017) could change this! Supported by Hypothes.is
- **Versioning:** only latest version available, can disappear any time
 - “Link rot”: ~half of links in Supreme Court opinions are dead! (Zittrain et al. 2014)

Some other lesser-used hypertext concepts

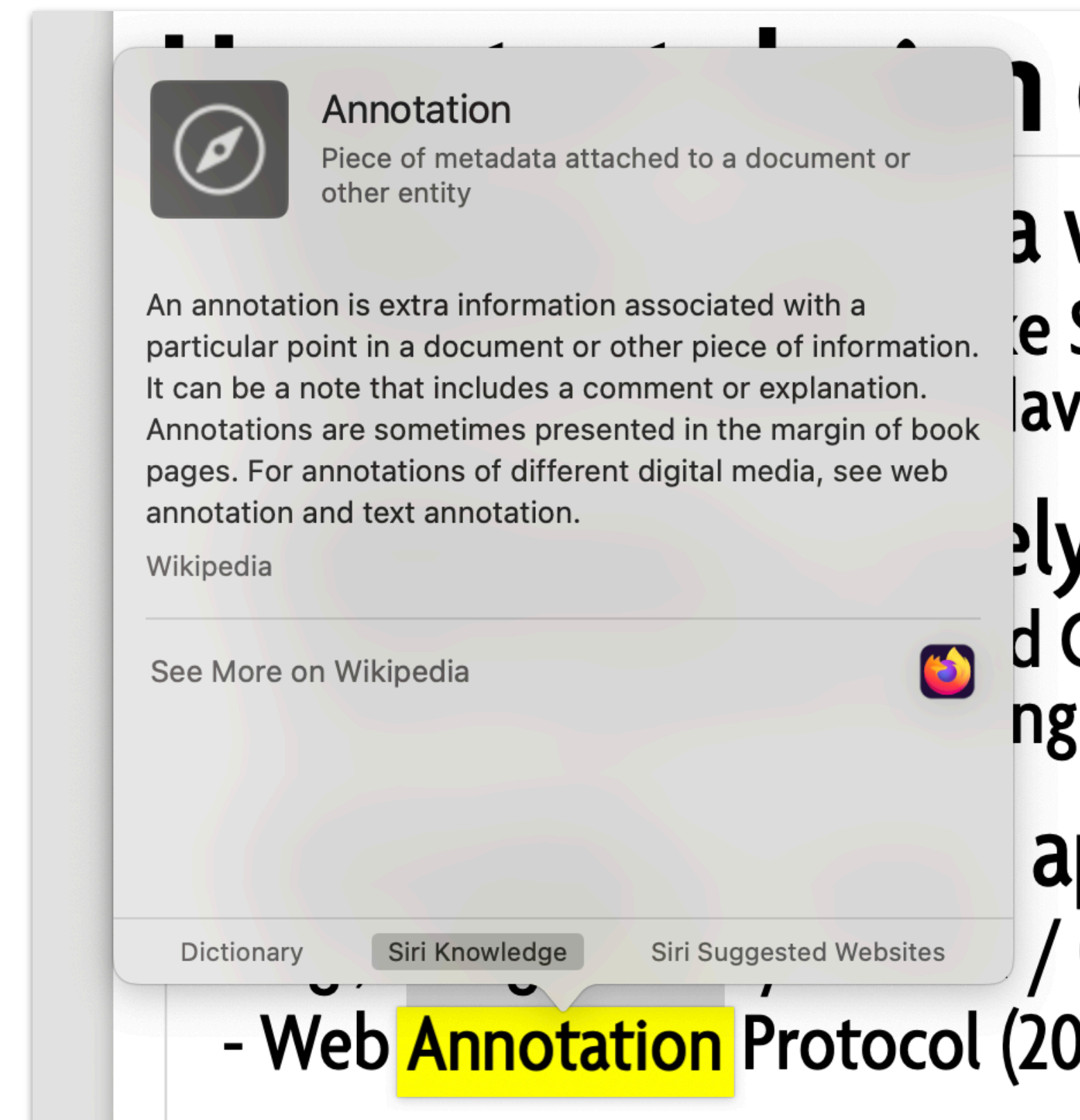
Transclusion

Live inclusion of content from other nodes



Dynamic linking

Reader rather than author determines links



Some other lesser-used hypertext concepts

Bidirectional links

See nodes which link to a given node

The following pages link to **Brown University**

External tools ([link count](#) • [transclusion count](#) • [sorted list](#)) •

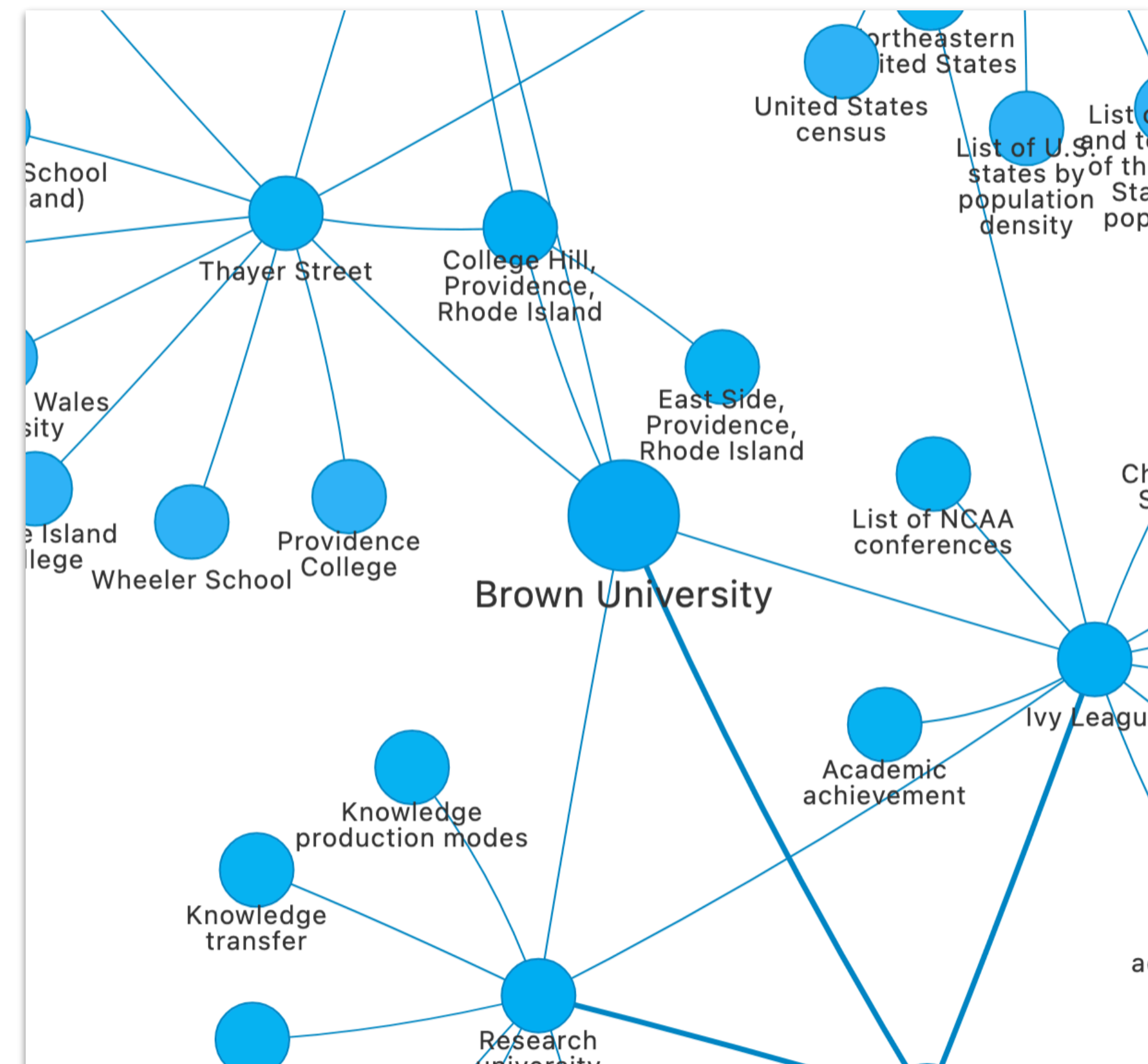
Showing 50 items.

View (previous 50 | [next 50](#)) ([20](#) | [50](#) | [100](#) | [250](#) | [500](#))

- [Alvin Toffler](#) ([links](#) | [edit](#))
- [Alexis Carrel](#) ([links](#) | [edit](#))
- [Aberdeen](#) ([links](#) | [edit](#))
- [Arizona State University](#) ([links](#) | [edit](#))
- [A cappella](#) ([links](#) | [edit](#))
- [Barnard College](#) ([links](#) | [edit](#))
- [Bioterrorism](#) ([links](#) | [edit](#))
- [Zebrafish](#) ([links](#) | [edit](#))
- [California Institute of Technology](#) ([links](#) | [edit](#))
- [Columbia University](#) ([links](#) | [edit](#))

Spatial hypertext

Visualize topology of hypertext graph



Why did the web succeed?

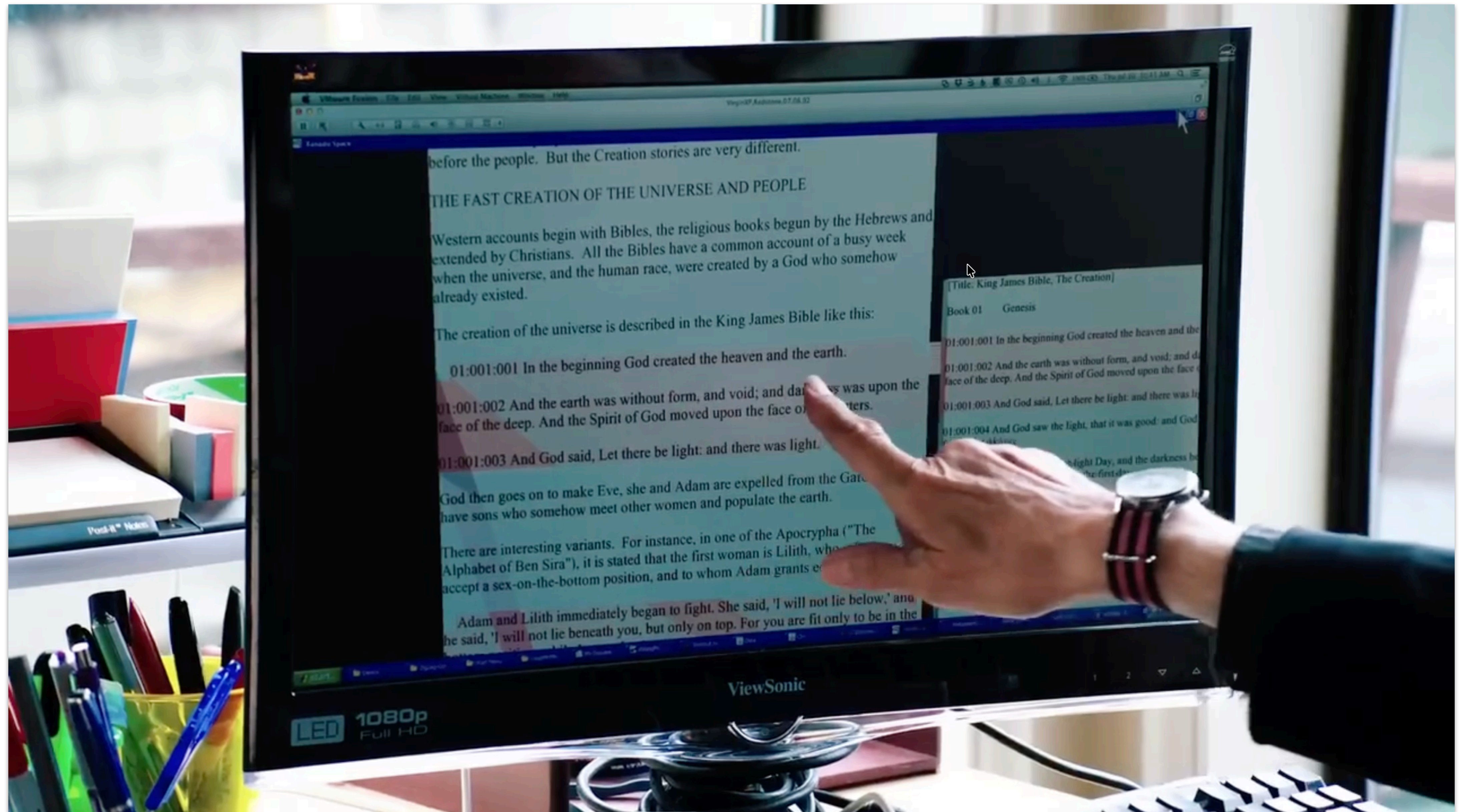
- Highly focused on the concrete problems of an org with deep knowledge work
 - Even though it didn't ultimately find the expected adoption in CERN!
- “Worse is better”: solutions only need to be good enough
 - First mover's advantage + lock-in means middling approaches win over great ones
 - “Universal Document Locator” became “Uniform Resource Locator” by IETF pressure
- Distributed structure allowed anyone to contribute
 - People independently decided to make their own web pages, or browsers!

Where does the Web fall short?

- **Contributing:** many more readers than writers.
 - No mechanisms for annotation (Web Annotation spec, eventually!)
 - Most contributions cordoned off to walled gardens: wikis, forums, social media, ...
- **Stability:** link rot, silent edits, no easy way to archive websites
 - Again, all stability is application-level: Google Docs, Github, etc.
- **Granularity:** documents / “pages” are still the organizing unit

“The Xanadu project did not "fail to invent HTML". HTML is precisely what we were trying to PREVENT-- ever-breaking links, links going outward only, quotes you can't follow to their origins, no version management, no rights management.”

— Ted Nelson (1999)



<https://youtu.be/Bqx6li5dbEY>