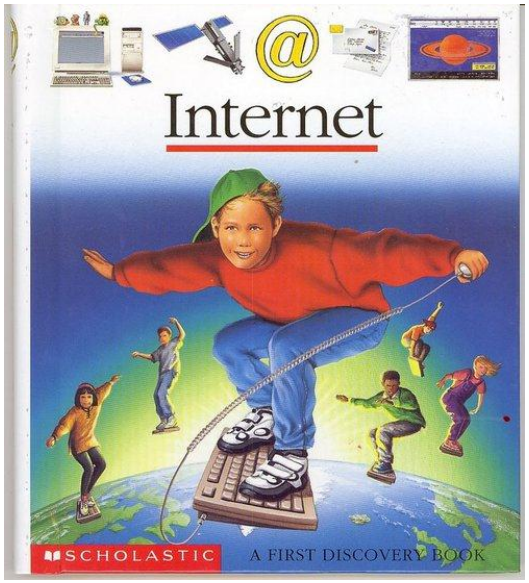
THE WORLD WIDE WEB

The World Wide Web: Past, Present and Future **Tim Berners-Lee** [1996]

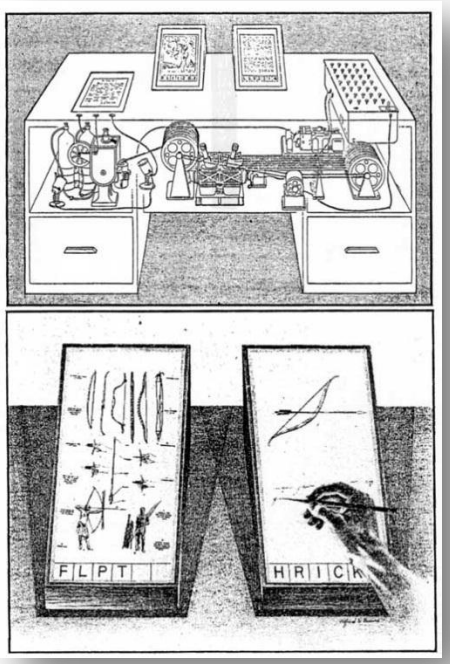
The world-wide web **T.J. Berners-Lee, R. Cailliau and J.-F. Groff** [1992]

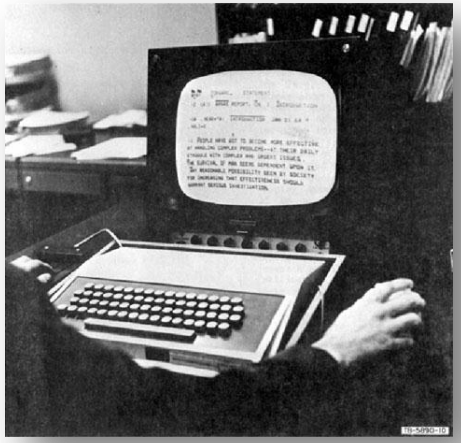
What is the world wide web?

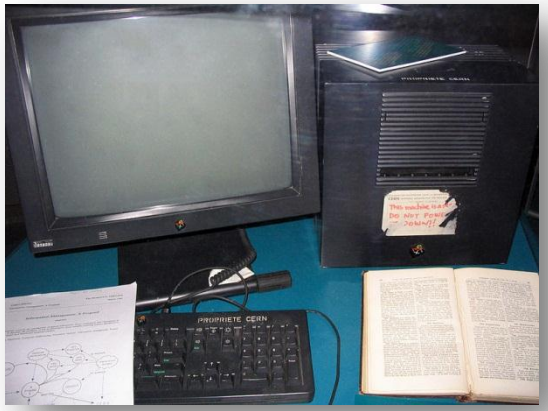


Before the web

- Silos of information
  - Incompatible protocols, networks and data formats
- At a time when more and more information is being recorded electronically
- As We May Think – Vannevar Bush [1945]
  - “The summation of human experience is being expanded at a prodigious rate, and the means we use for threading through the consequent maze to the momentarily important item is the same as was used in the days of square-rigged ships.”
  - “This is the essential feature of the memex. The process of tying two items together is the important thing.”


F L P T
H R I C k




PROPRIETE CERN

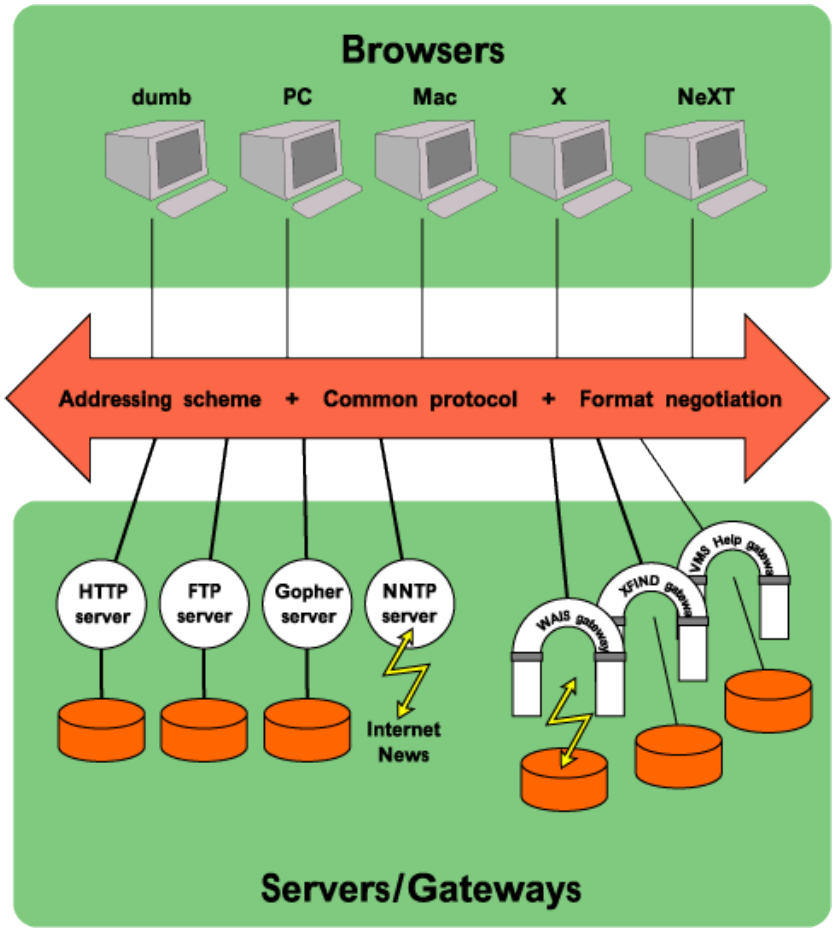
Goal

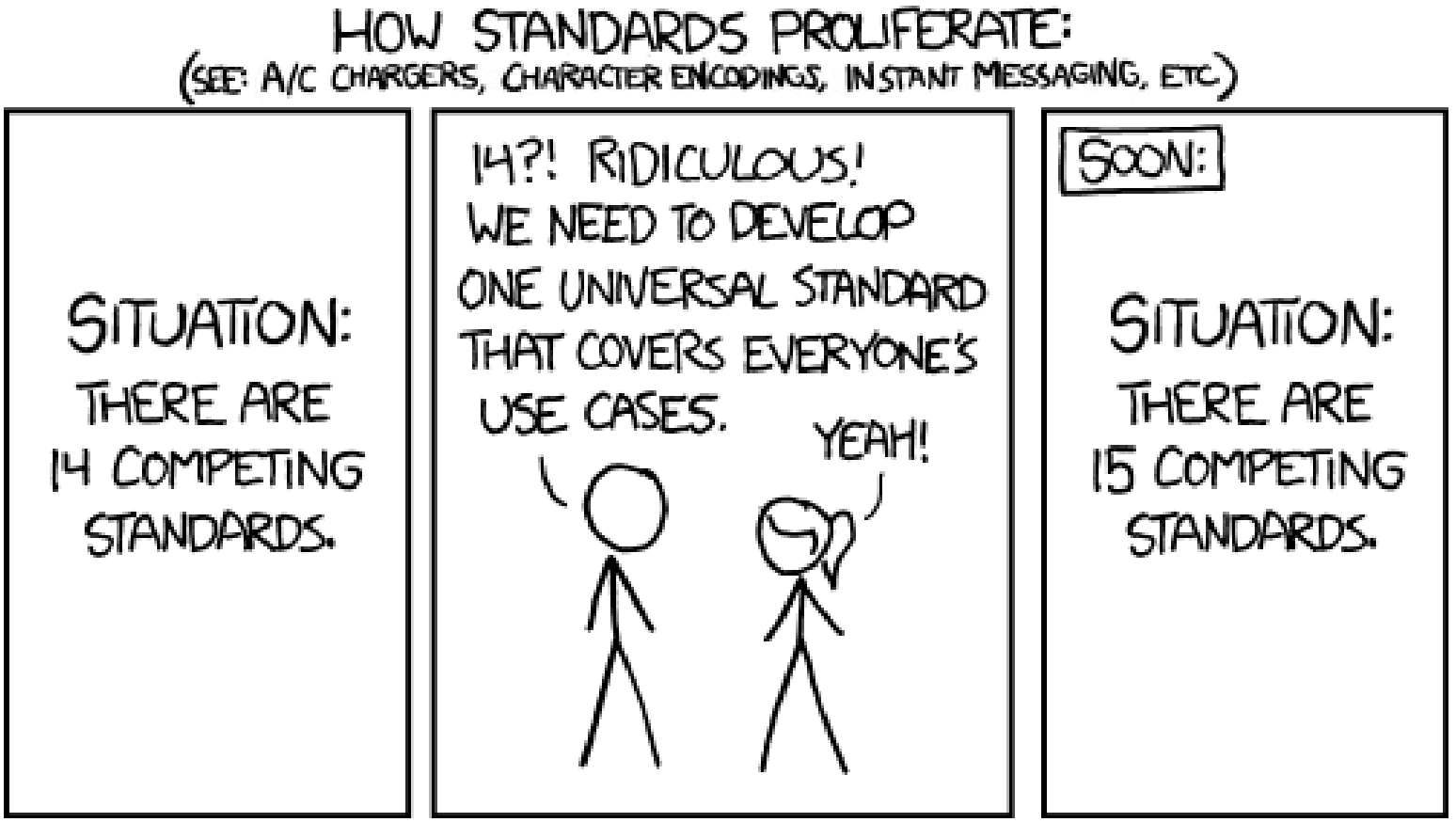
- Shared information space through which people and machines could communicate.
  - Globally shared information space
  - Machine analysis of work patterns, interactions and thoughts

- Design criteria
  - Allow random associations between any objects
  - Independent systems (scalable)
  - No constraints on language/OS/user mental model of data
  - Easy to use (browse and add new information)
  - Supports all platforms and forward compatible

Architecture

- Flexible
  - Minimal constraint
    - Major factor in adoption
  - Independent specifications
    - Allow parts of design to be replaced
- Addressing scheme
  - URI
- Common protocol
  - HTTP
- Format negotiation
  - Best suited content




HOW STANDARDS PROLIFERATE:
(SEE: A/C CHARGERS, CHARACTER ENCODINGS, INSTANT MESSAGING, ETC)
SITUATION:
THERE ARE
14 COMPETING
STANDARDS.
14?! RIDICULOUS!
WE NEED TO DEVELOP
ONE UNIVERSAL STANDARD
THAT COVERS EVERYONE'S
USE CASES.
YEAH!
SOON:
SITUATION:
THERE ARE
15 COMPETING
STANDARDS.

URI

- Points to any resource of any kind in the universe of information.
  - Prefix indexes into specified “information space” (e.g. HTTP/NNTP)
- Dangling reference
  - Allowing references to be made without consultation with the destination.
  - Provides scalability at cost of bad link/link rot

- Question: URI are generally treated as opaque strings: client software is not allowed to look inside them and draw conclusions about the object referenced.


MARIO
069550
×12
WORLD
1-4
TIME
235
404 PAGE NOT FOUND
THANK YOU MARIO!
BUT OUR PRINCESS IS IN
ANOTHER CASTLE!

HTTP + Format negotiation

- Protocol for traversing hypertext links
  - “Simple stateless search and retrieve protocol”
    - Cookies
  - “… using one TCP/IP connection per search or retrieve operation”
    - HTTP keep-alive (default in HTTP 1.1)
- Format negotiation (HTTP Header)
  - Client
    - Specify preferences in terms of language and data format.
  - Server
    - Select specific object best fit client needs
  - Independence between HTTP and HTML
  - Example
    - Return GIF instead of PNG (IE4)
    - Mobile?

HTML

- “Because HTML is a high-level markup, it allows the same logical structured text to be represented optimally whatever the capabilities of the client platform. For example … a dumb terminal browser may use capital letters when an X-windows browser uses a different font.”
  - Early HTML (no CSS): use special tags and browser chooses formatting
  - CSS introduced to separate content from format information

Current situation (1996)

- Leverage HTTP to transport arbitrary data formats
- Format negotiation
  - Not widely deployed
  - Leads to browser make/version check by server
  - Difficult to introduce new clients
- Formation of WWW Consortium @ MIT to create and maintain standards (1994)
- Platform for Internet Content Selection
  - Questions
    - Wide spread adoption?
    - Require technical savvy parents to setup filters and extra cost?
    - No standard rating (label) policy?

Future (now?)

- Infrastructure improvement
  - Server scalability
  - Handle dynamic load
  - Fault tolerance
  - Today – distributed systems, cloud computing (provisioning VMs on demand)
- Human communication
  - Group tool at all scales (usage model as a medium to exchange ideas between coworkers, not pron and cat pictures?)
  - Question: decentralized and cooperative usage model?
    - "… run an existing shell interface to the database from your server script"

Future (now?)

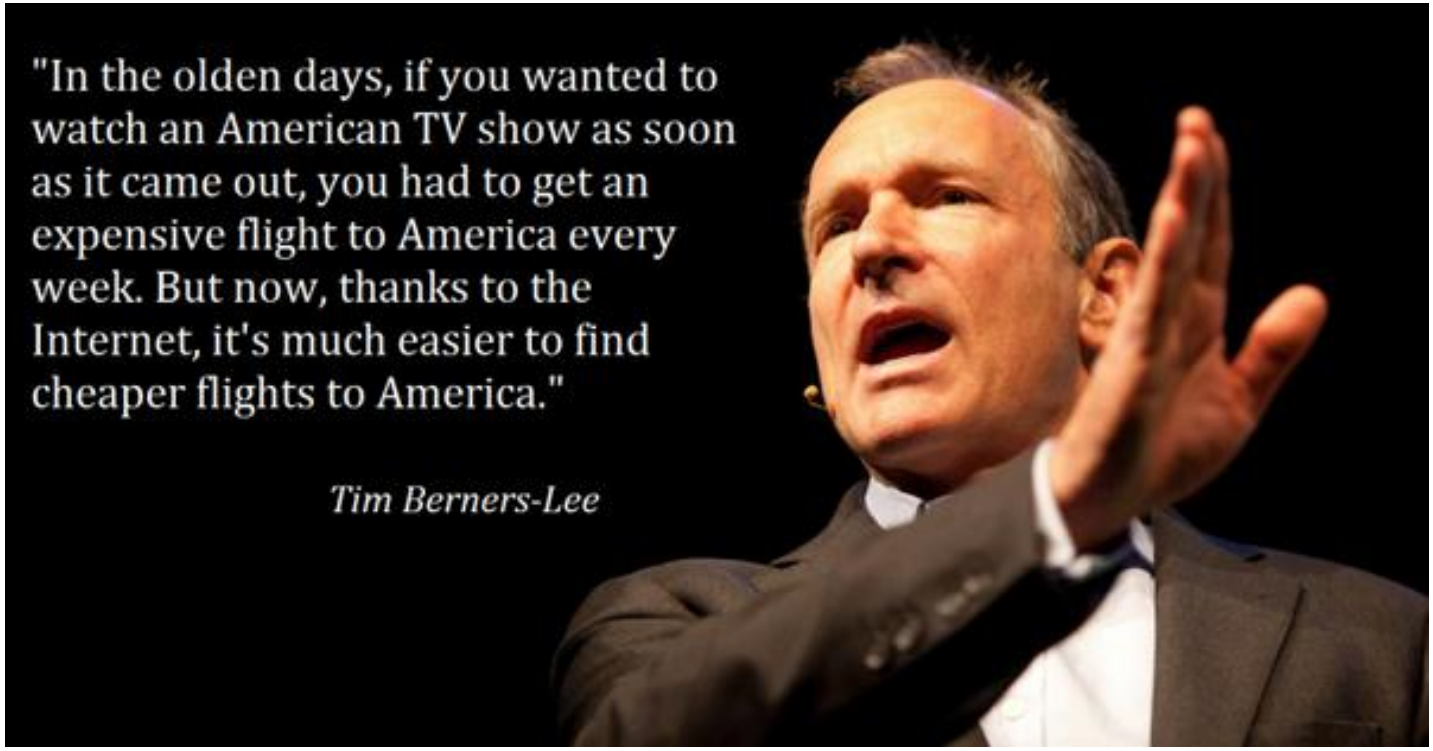
- Machine interactions
  - Machine-readable form (Electronic Document Interchange)
  - Agents (computers) to facilitate business on behalf of users
    - Computerized trading?
    - IOT?

Ethical and social concerns (relevant today)

- Engineers deigning protocol should be aware of associated social/ethical issues
- Copyright and freedom of information
  - Pirate Bay
  - WikiLeaks

Conclusion

- “End to end” functionality
- Platform independent, consistent, reliable and fair
  - Law and technology need to work together
  - Net neutrality?


"In the olden days, if you wanted to watch an American TV show as soon as it came out, you had to get an expensive flight to America every week. But now, thanks to the Internet, it's much easier to find cheaper flights to America."
Tim Berners-Lee