Internet and World Wide Web

S.Chellammal

13.03.2014

Outline

- An overview of Internet
- World Wide Web
  - URL,
  - Web server, Web browser, HTTP Protocol
  - Finding information on the Web
- Evolution of Web
- Today's Web
- Web in Future

Which came first?

- Internet or World Wide Web?

The Birth of ARPA

- In 1958, US forms the Advanced Research Projects Agency (ARPA)
- ARPA directly reports to the US Department of Defense (DoD)

ARPANET

- In 1969, the ARPANET commissioned by DoD for research into networking

- Only 4 nodes comprise the ARPANET

The first ARPANET

- Nodes are connected by AT&T 50kbps lines.
- Node 1: University of California Los Angeles
- Node 2: Stanford Research Institute
- Node 3: University of California Santa Barbara
- Node 4: University of Utah, Host

Hosts

- IBM 360/75

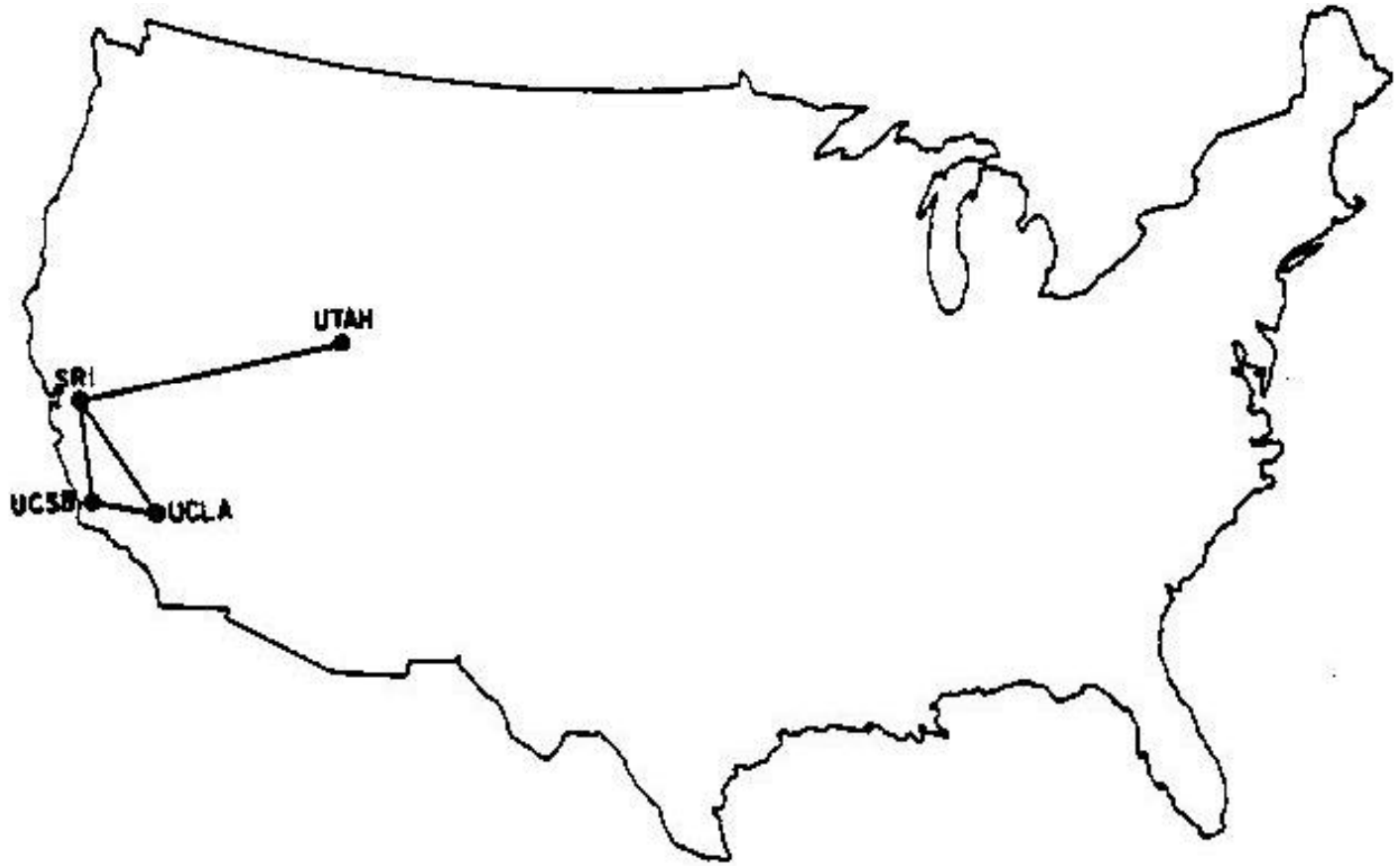
DEC PDP-10

Interface Message Processor (IMP)

4 IMPs were connected,
ARPANET was born.

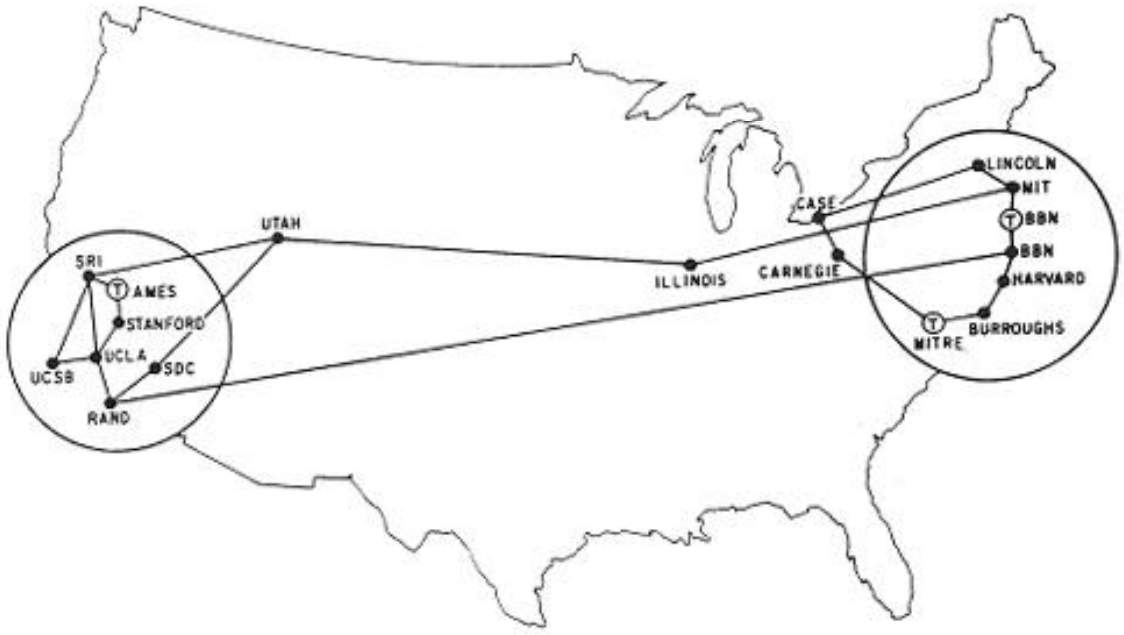
Interface Message Processor (IMP)

Geographical Position



The growth of ARPANET

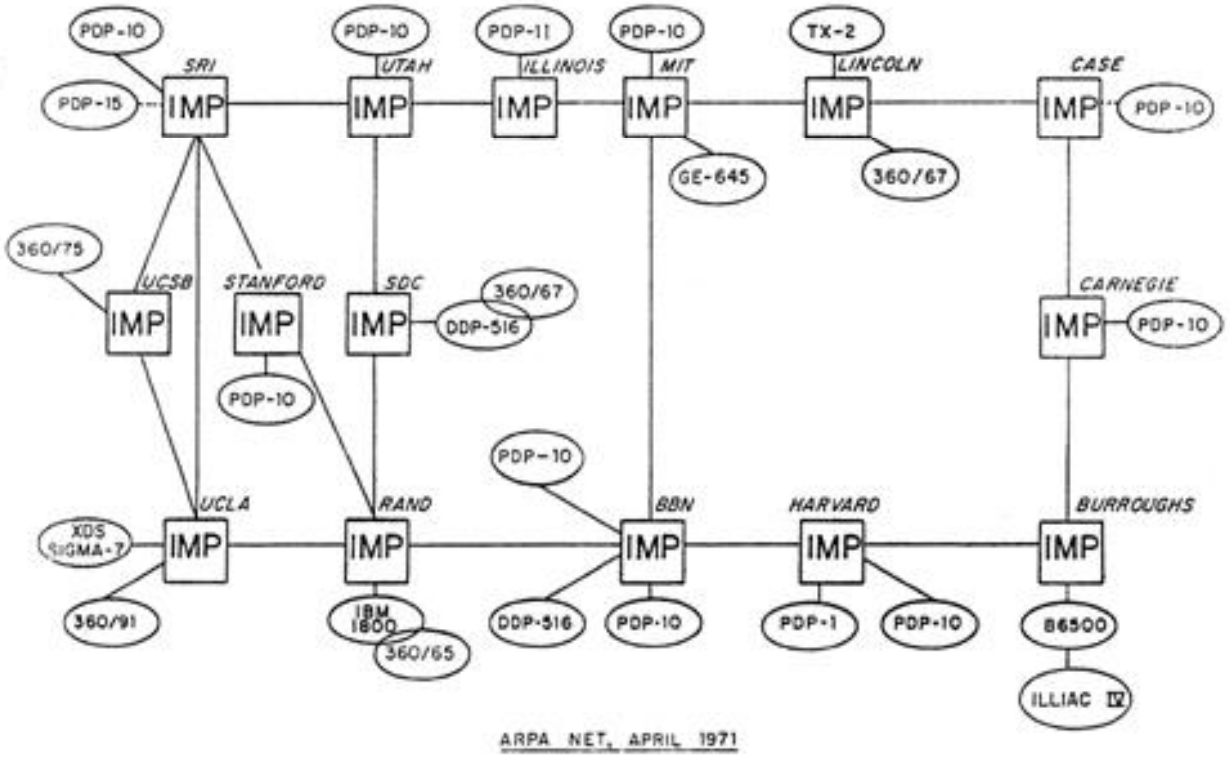
- 1971, 15 nodes (23 hosts): UCLA, SRI, UCSB, Univ of Utah, MIT, Harvard, Stanford, UIUC, CMU, NASA, etc…



MAP 4 September 1971

The growth of ARPANET

- 1971, 15 nodes (23 hosts): UCLA, SRI, UCSB, Univ of Utah, MIT, Harvard, Stanford, UIUC, CMU, NASA, etc…



Late 1970s, Early 1980s

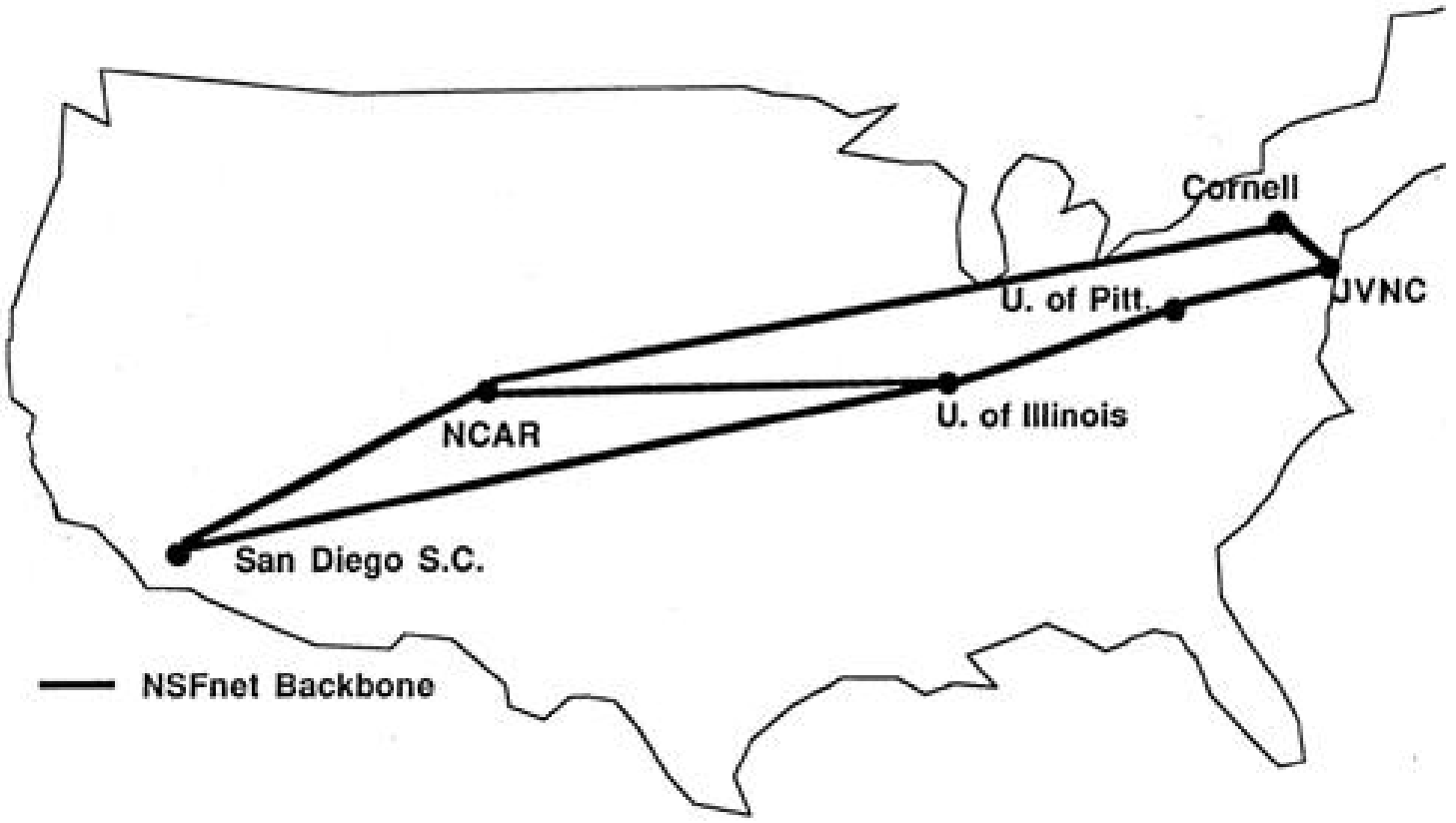
- Many networks were built
  - In 1981, BITNET, the “Because It’s Time NETwork” started as cooperative network.
  - CSNET (Computer Science NETwork) seeds grant support by National Science Foundation (NSF) and provides connection between universities.
  - EUnet (European UNIX Network)
  - JUNET (Japan UNIX Network)
  - JANET (Joint Academic Network) in UK

NSFNET

- In 1986, NSFNET was created (backbone speed of 56Kbps) (National Science Foundation)

- Connected 5 supercomputing centers.
  - JVNC@Princeton
  - PSC@Pittsburgh
  - SDSC@UCSD
  - NCSA@UIUC
  - Theory Center@Cornell

NSFNET - Backbone



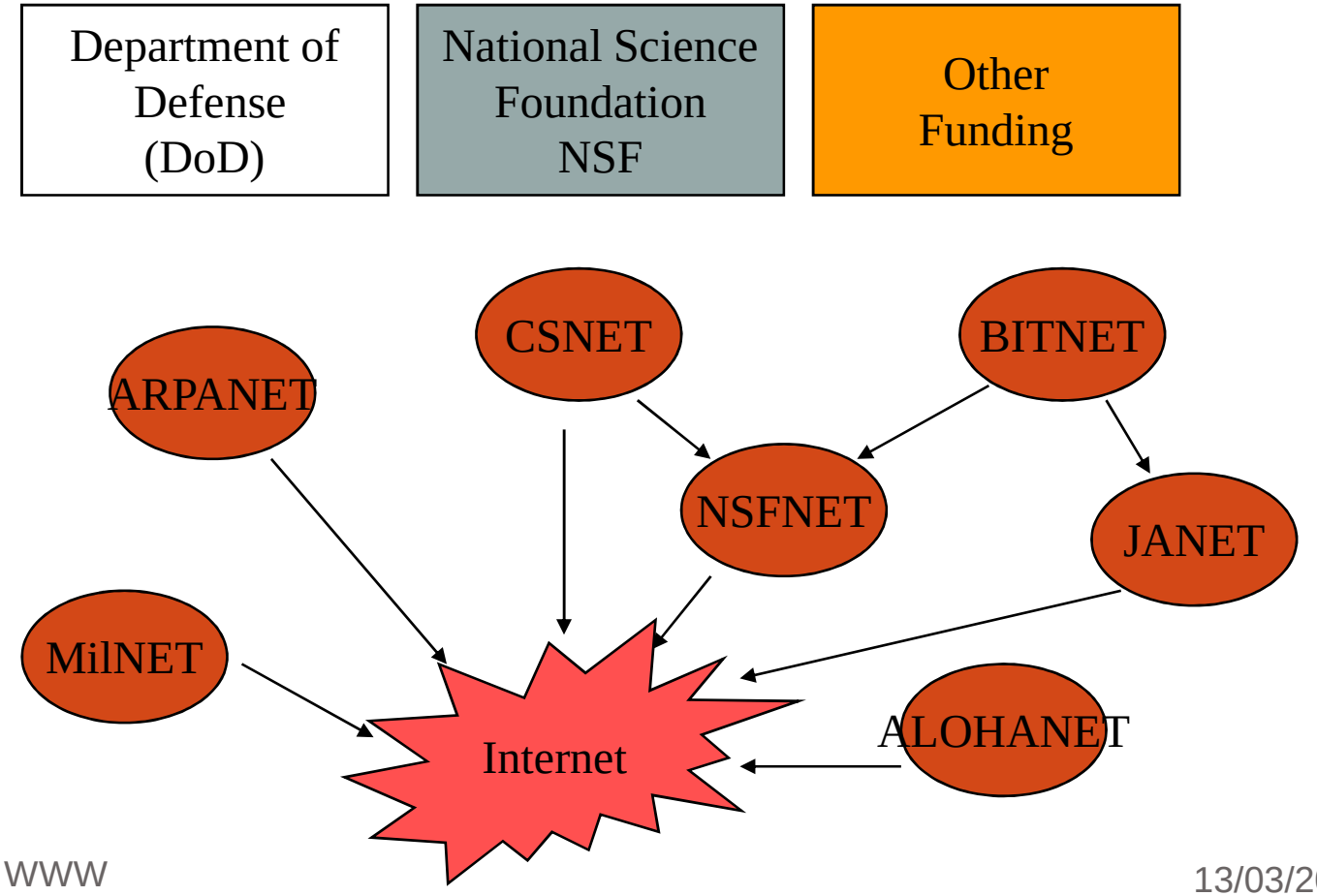
Expansion of hosts

- Number of hosts breaks 10,000 in 1987
- NSFNET backbone upgraded to T1 (1.5M) 1988

- Number of hosts breaks 100,000 in 1989
- NSFNET upgraded to T3 (44.736Mbps) 1991

- Number of hosts breaks 1,000,000 in 1992

The emergence of the Internet

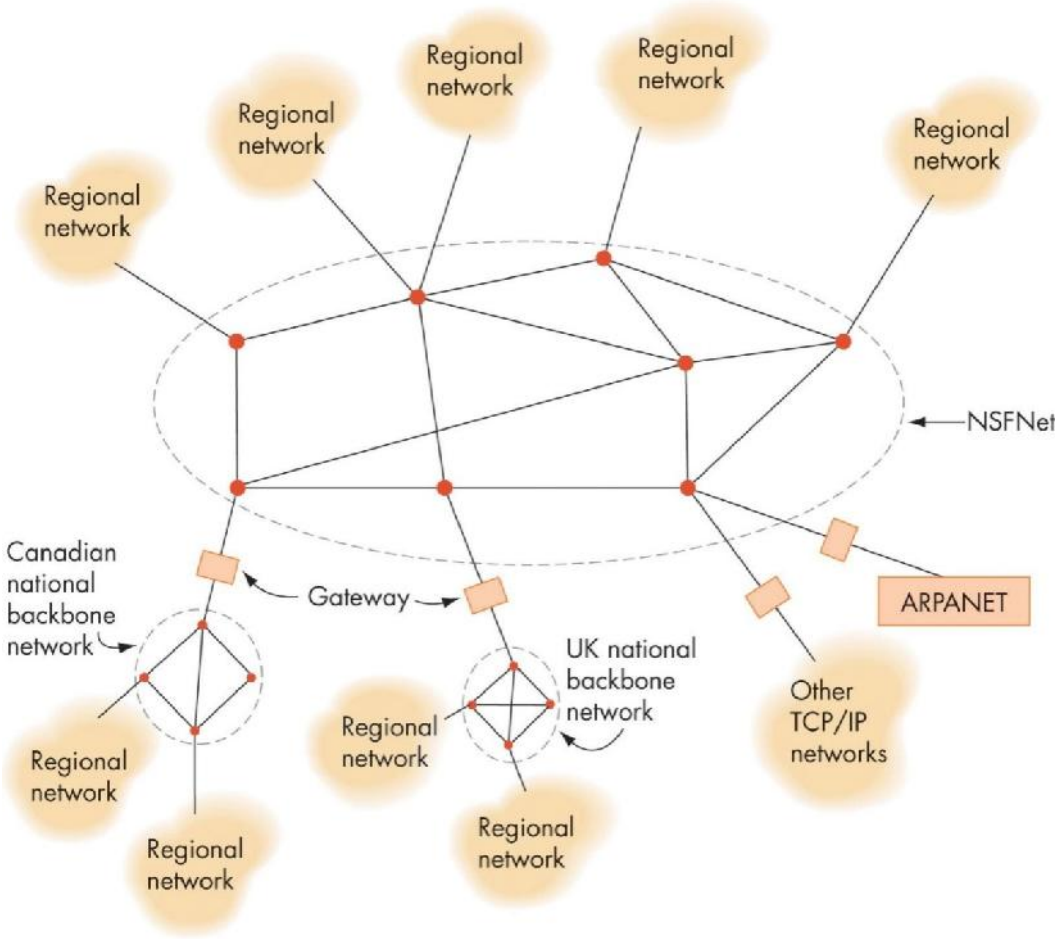


Internet

- Commercial Internet Service Provider

- Internet service providers start offering Internet access once provided by the ARPANET and NSFNet

Internet

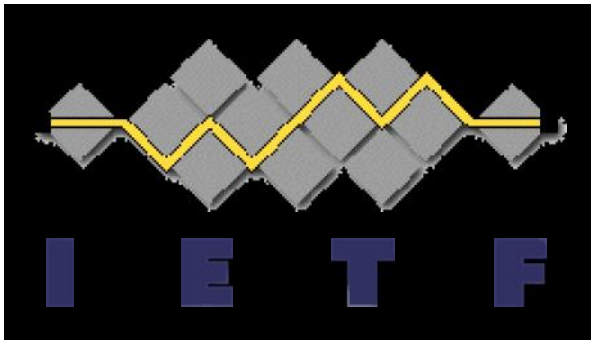

Regional network
Regional network
Regional network
Regional network
Regional network
Regional network
NSFNet
Canadian national backbone network
Gateway
ARPANET
UK national backbone network
Other TCP/IP networks
Regional network
Regional network
Regional network

Who governs the Internet?

- NOBODY!!
- Internet Society (ISOC)
  - Professional membership society
- World Wide Web Consortium
  - develops technologies (specifications, guidelines, software, and tools)
- Internet Network Information Center (InterNIC)
  - Domain registration

Who governs the Internet?

- Internet Corporation for Assigned Names and Numbers (ICANN)
  - responsibility for the IP address space allocation, protocol parameter assignment, domain name system management, and root server system management
- Internet Engineering Task Force (IETF)
  - a large open international community of network designers, operators, vendors, and researchers



Who manages IP address?

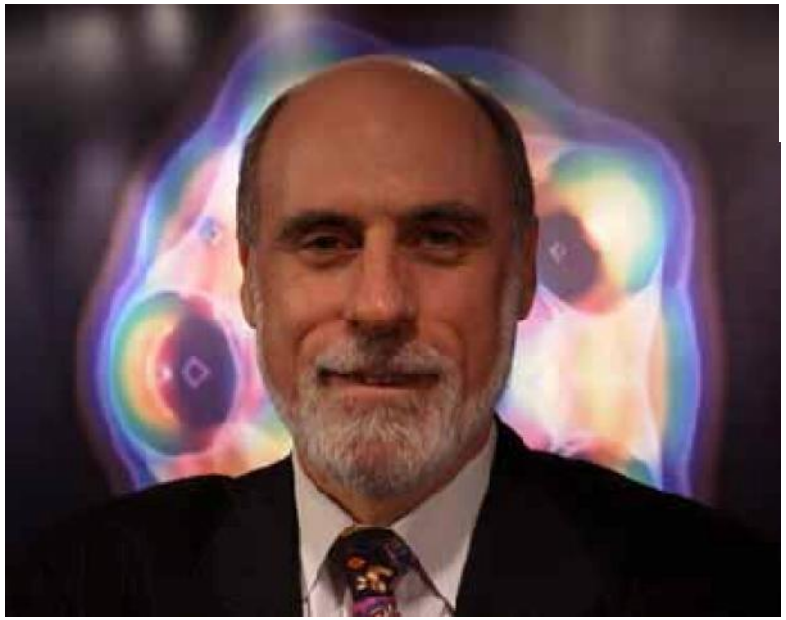
- Internet Assigned Numbers Authority (IANA)

- Asia Pacific Network Information Centre

- American Registry for Internet Number

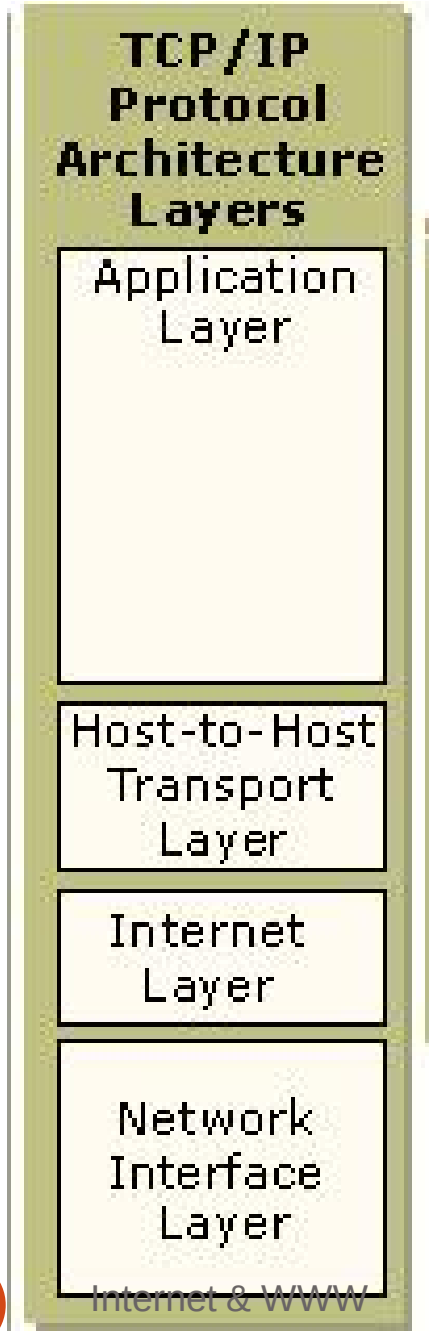
The father of the Internet

- Vint Cerf defines the network protocol and breaks the independent self-contained networks, forms TCP/IP which becomes the standard



Communication Protocols

- A protocol
  - A mutually agreed upon set of rules, conventions, and agreements for the efficient and orderly exchange of information
- TCP/IP
  - The Internet protocol hierarchy
  - Governs the operation of the Internet
  - Five layers



Application layer

Refers to standard network services like http, ftp, telnet as well as communication methods used by various application programs

Also ***defines compatible representation*** of all data

Transport layer

Manages the transfer of data by using connection oriented (TCP) and connectionless (UDP) transport protocols

Manages the connections between networked applications

Internet layer

Manages ***addressing of packets and delivery of packets between networks***

Fragments packets so that they can be dealt with by lower level layer (Network Interface layer Network)

Network Interface layer

Delivers data via physical link (Ethernet is the most common link level protocol)

Provides ***error detection and packet framing***

Internet services

- **Electronic Mail (E-mail)**
- **FTP (File Transfer Protocol)**
- **Telnet**
- **Voice / Video Communication – Skype**
- **Social networks**
- **World Wide Web**
- **Search engines**

The World Wide Web

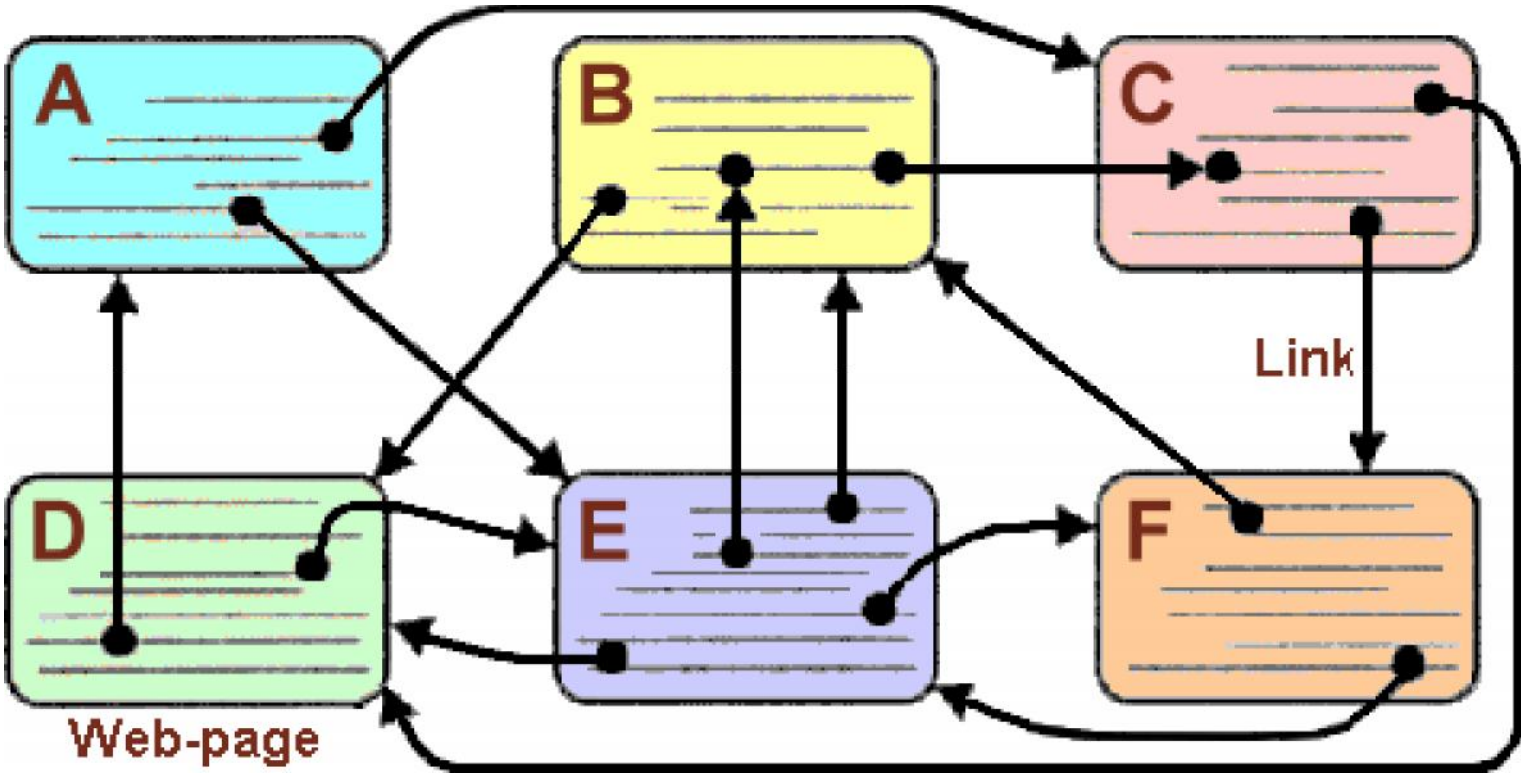
- 1989 – WWW concept by Tim Berners-Lee at European Laboratory for Particle Physics in 1991

- 1990 – first browser/editor program

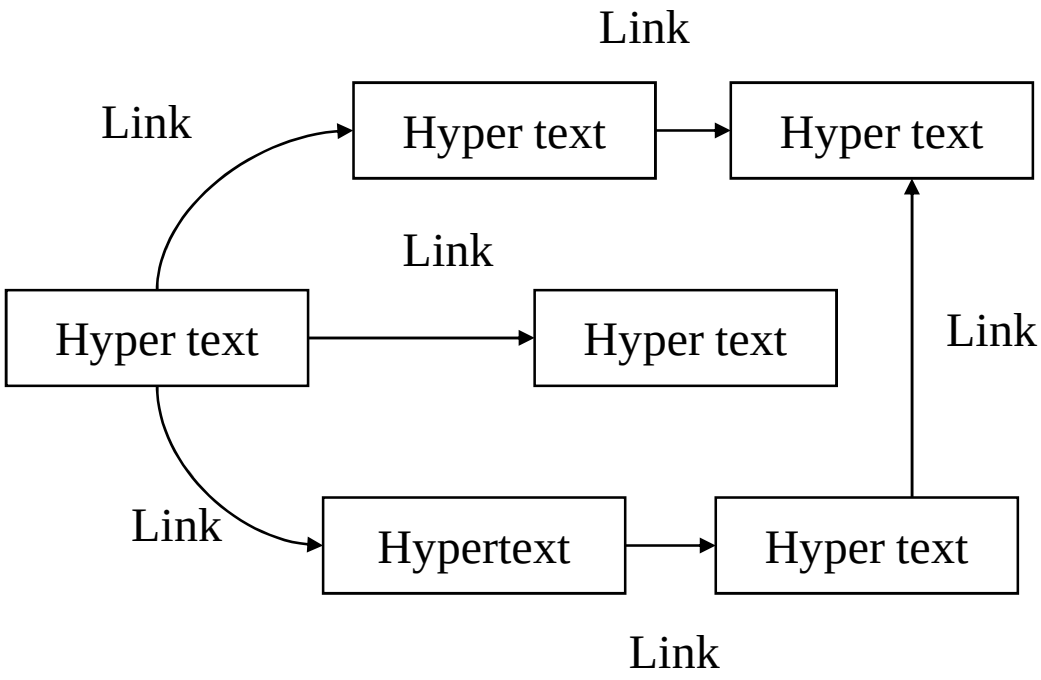
What is WWW?

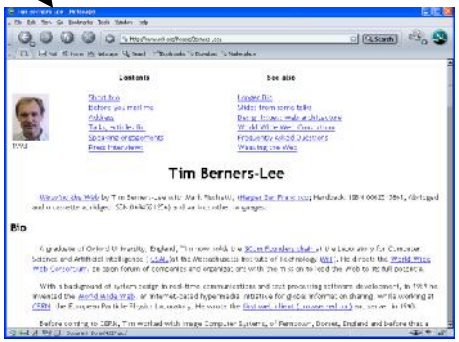
- The World Wide Web (abbreviated as WWW/ web) is a system of interlinked hypertext documents accessed via the Internet. With a browser, one can view web pages that may contain text, images, videos, and other multimedia and navigate between them via hyperlinks.

World Wide Web

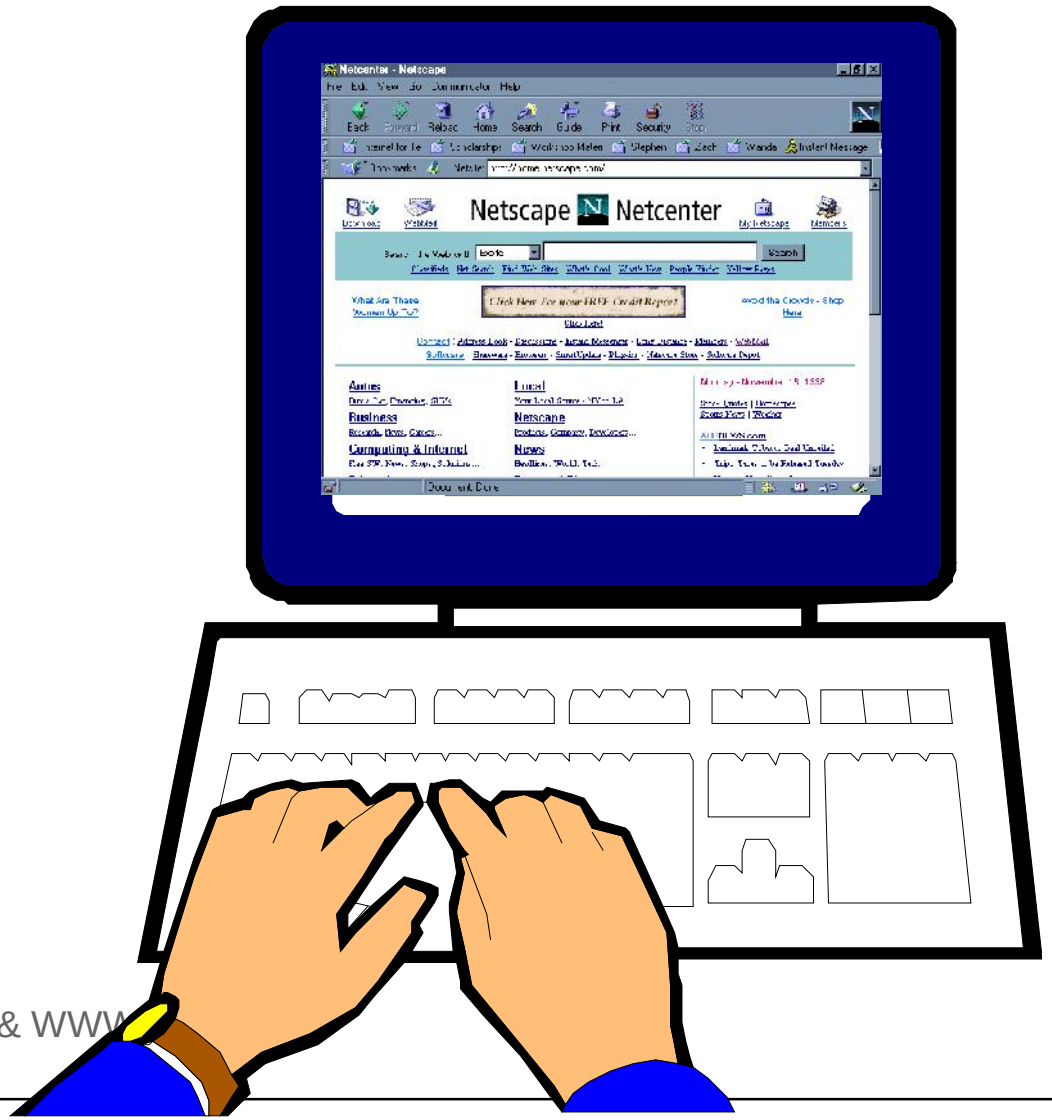


World Wide Web



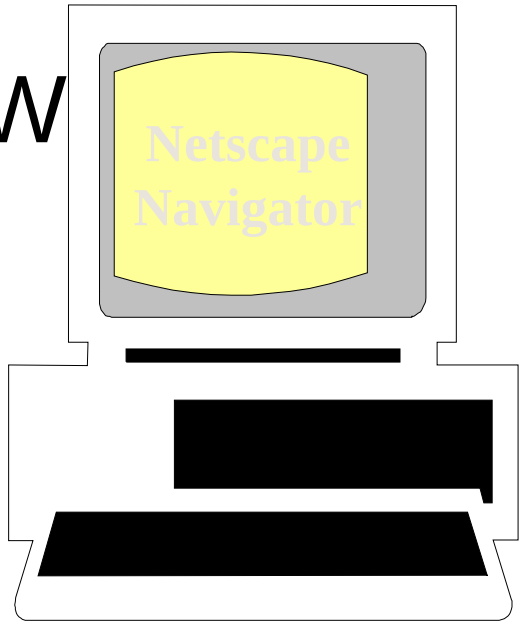


- Resources are linked together forming the web
- no distinction between the resources
- no distinction between the links


Netcenter - Netscape
Netscape Netcenter

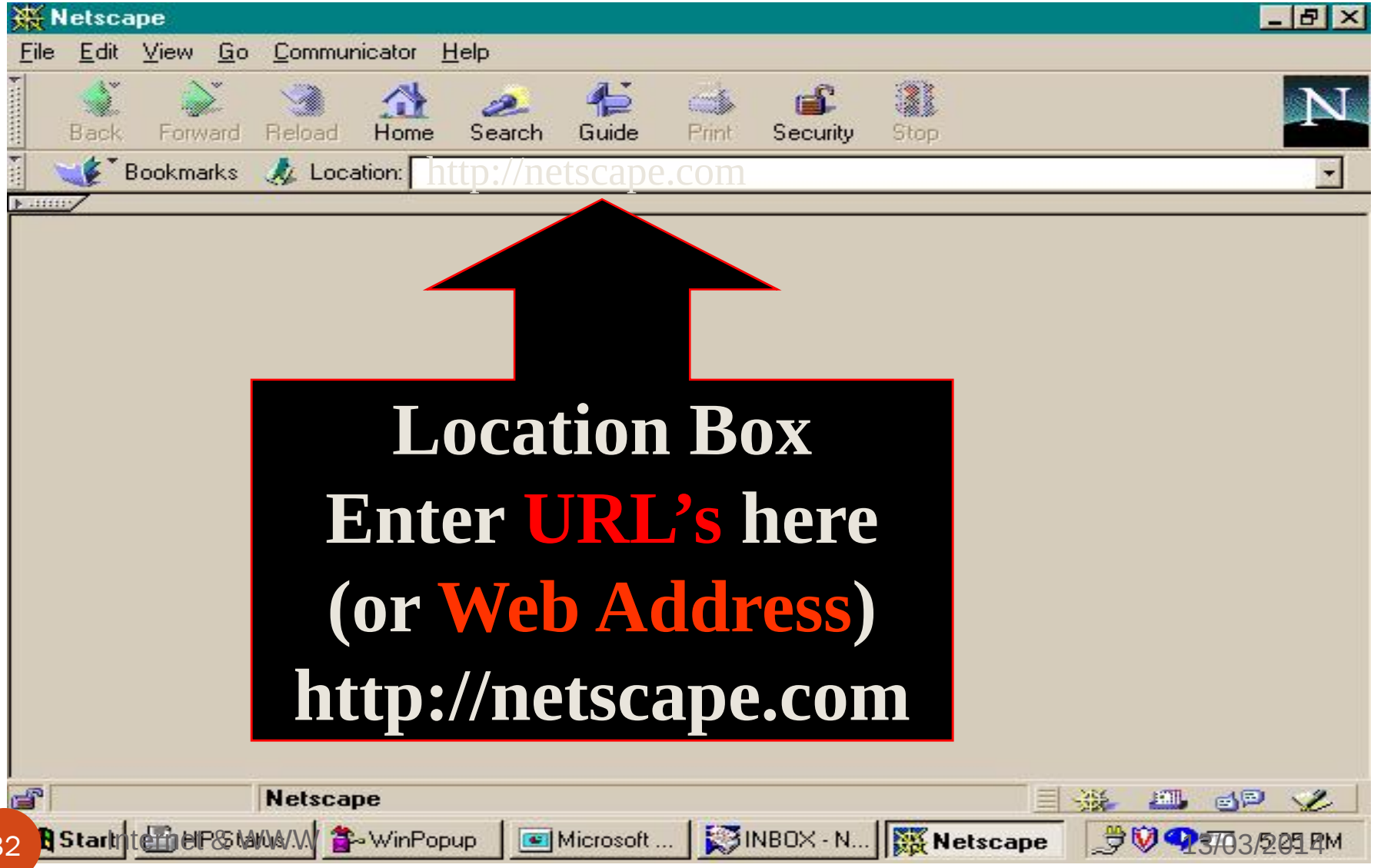
Web Browser

- Software programs used to access the World Wide W



  - Microsoft Internet Explorer
  - Netscape Navigator
  - Mozila firefox; Google chrome

Entering a URL



Universal Resource Locator

http://www.msu.edu/~urquhar5/tour/active.html

http://
identifies type
of transfer

www.msu.edu
Domain Name -
name of remote computer

/~urquhar5/tour/active.html
File Location on Remote Computer

HTTP Request

METHOD URL VERSION<crlf>

HEADER: Value<crlf>

. . .

HEADER: Value<crlf>

Empty line <crlf>

BODY OF THE REQUEST

GET http://en.kioskea.net/ HTTP/1.0

Accept: text/html

HTTP Response

VERSION-HTTP CODE EXPLANATION<crlf>

HEADER: Value<crlf> . . .

. . .

HEADER: Value<crlf>

Empty line <crlf> BODY OF THE RESPONSE

HTTP/1.0 200 OK

Content-Type: text/HTML

Content-Length: 1245

<html><body>Welcome to WWW</body></html>

HTTP Response Format

status-line (HTTP-version response-code response-phrase)
headers (0 or more)
<blank line>
body

- Second type of HTTP message: *response*
  - Web servers construct and send response messages
- Typical HTTP response:
  - HTTP/1.0 301 Moved Permanently

    Location: http://www.wisc.edu/cs/index.html

HTTP Response Codes

1xx – Informational – request received, processing

2xx – Success – action received, understood, accepted

3xx – Redirection – further action necessary

4xx – Client Error – bad syntax or cannot be fulfilled

5xx – Server Error – server failed

Web Pages

Web Pages - documents on the WWW

Written and constructed using HTML

HTML - Hyper Text Markup Language

- a special language or code used to design and publish documents on the Web

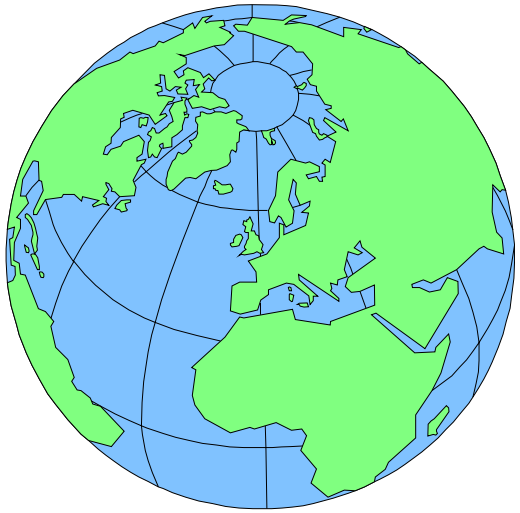
```
<html>
<head>
<title> welcome page</title>
</head>
<body> HAI GREETINGS
</body>
</html>
```

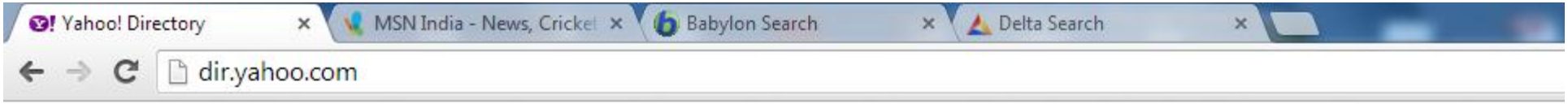
Finding information on Web

using

Directories

Search Engines





New User? Register | Sign In | Help

Get new Yahoo Mail apps

YAHOO!
DIRECTORY SUBMIT

Search

Yahoo! Directory

Arts & Humanities Photography, History, Literature...	**News & Media** Newspapers, Radio, Weather, Blogs...
Business & Economy B2B, Finance, Shopping, Jobs...	**Recreation & Sports** Sports, Travel, Autos, Outdoors...
Computer & Internet Hardware, Software, Web, Games...	**Reference** Phone Numbers, Dictionaries, Quotes...
Education Colleges, K-12, Distance Learning...	**Regional** Countries, Regions, U.S. States...
Entertainment Movies, TV Shows, Music, Humor...	**Science** Animals, Astronomy, Earth Science...
Government Elections, Military, Law, Taxes...	**Social Science** Languages, Archaeology, Psychology...
Health Disease, Drugs, Fitness, Nutrition...	**Society & Culture** Sexuality, Religion, Food & Drink...
New Additions 10/5, 10/4, 10/3, 10/2, 10/1...	**Subscribe via RSS** Arts, Music, Sports, TV, more...

Help us improve the Yahoo! Directory - Share your ideas

Search engines

Alta Vista	http://www.altavista.digital.com
Excite	http://www.excite.com
HotBot	http://www.hotbot.com
InfoSeek	http://www.infoseek.com
Lycos	http://www.lycos.com
WebCrawler	http://www.webcrawler.com
Google	http://www.google.com

How does a search engine work?

- Performing Keyword Searches

Building index table

Build an index table for every single word on each page- indexation process

Key	Value
$word_1$	URL_1
$word_2$	URL_1
.....	.....
.....	.....
$word_n$	URL_1

Searching

Word	Value
Word1	URL_1 , w_1 ; URL_5 w_2 ; URL_9 w_3
SLR	www.cheapcamera.com(0.6), www.buyithere.com(0.45); ….(www.goodphoto.com(0.3);
Single	www.digcamhelp; www.ehow.com
Lens	www.digcamhelp; www.ehow.com
Reflex	www.digcamhelp; www.ehow.com

Web – an overview

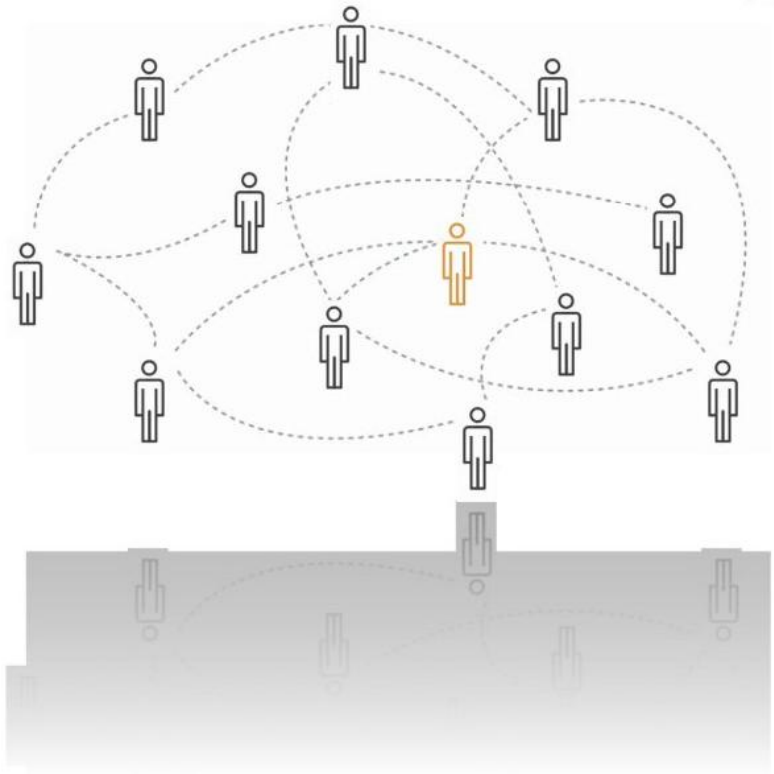
Social Media

Anyone can share anything
with anyone across the globe…

Web 2.0 Provides the Infrastructure for Innovation

Find and connect with people across boundaries

Rapidly distribute ideas, experiences and knowledge



Become aware of what others are doing

Tap into the knowledge of your informal network

Web 2.0 Concepts

- Blogging
- Forums
- Wikis
- Social Networks
- twitter
- Flickr
- You tube

Blogging

- Individuals broadcast ideas to like-minded people
- Business to broadcast latest information to stakeholders
- Citizen journalism
- Receive comments from readers
- Photos, videos, podcasts
- Micro-blogging (twitter)

Forums

- Online message board around one topic
- Discussions with posts and replies
- Threads are collections of posts and replies
- Moderators to clean up spam
- Software communities use forums as part of support platform

Wiki

- Collectively share and edit a body of knowledge
- Ongoing process of creation and collaboration
- Knowledge Management
- Wikipedia
- Enterprise wikis

Social Networking

- Online communities
- Share information
- Connect people with same interest
- Personal, Business, Political
- Facebook, Hyves, Friendster, Orkut, MySpace, LinkedIn, Plaxo
- Flickr, You-Tube, Slideshare, Iens

E-Commerce

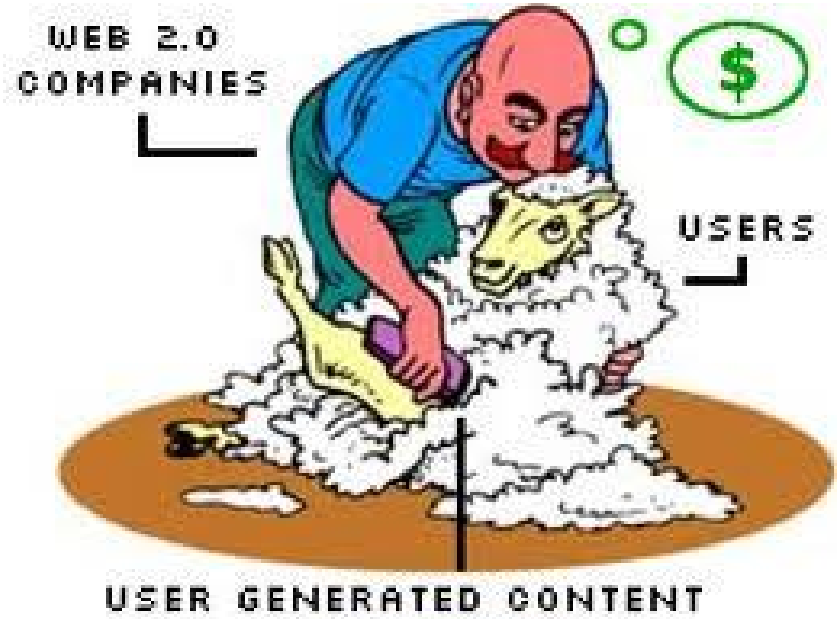
- Connect seller and buyer
- Adds value by providing service
- Ebay, Amazon, Marktplaats
- Kayak, hotels.com, Funda, Monsterboard
- Business model to leverage Web 2.0 technology

Mashup

- Aggregates data from more than one source
- Often using Open API to build services from data sources
- e.g. Real estate data on Google map

Web 2.0

- Web 2.0 – Writable Web



▪-Social networking

- Side effect

Web 3.0

Semantic Web

Web with semantics

Manual integration

- consult a large number of web sites, all different in style, purpose, possibly language…
- We have to mentally integrate all this information to achieve our goals.
- sometimes, this is a long and tedious process!

What is existing?

- the real data is hidden somewhere in databases, XML files, spread sheets etc.
- Some sites such as Expedia and Tripadvisor combine data from other sources with the approval of data owners

What we would like to have?

- We would like to have applications that can

combine all the data in the different Web sites (or underlying databases) in a useful way.

Requirements

- the applications should have mechanisms to access the data
- mechanisms to link data to its related data like Web pages today

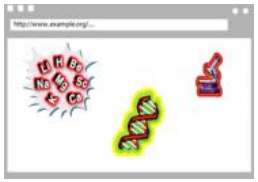
Or put in another way:

- We would like to extend the current Web with a "Web of data"

What to do?

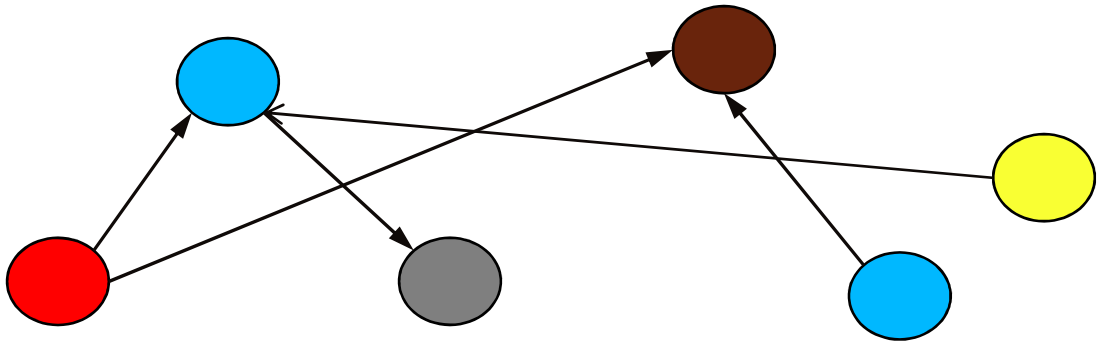
Data should have explicit formal meaning

Common model



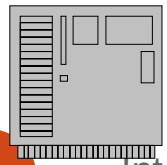
Applications

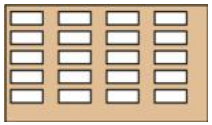
Manipulate
Query
…

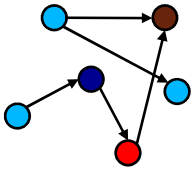


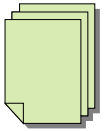
Data represented in abstract format

Map,
Expose,
…





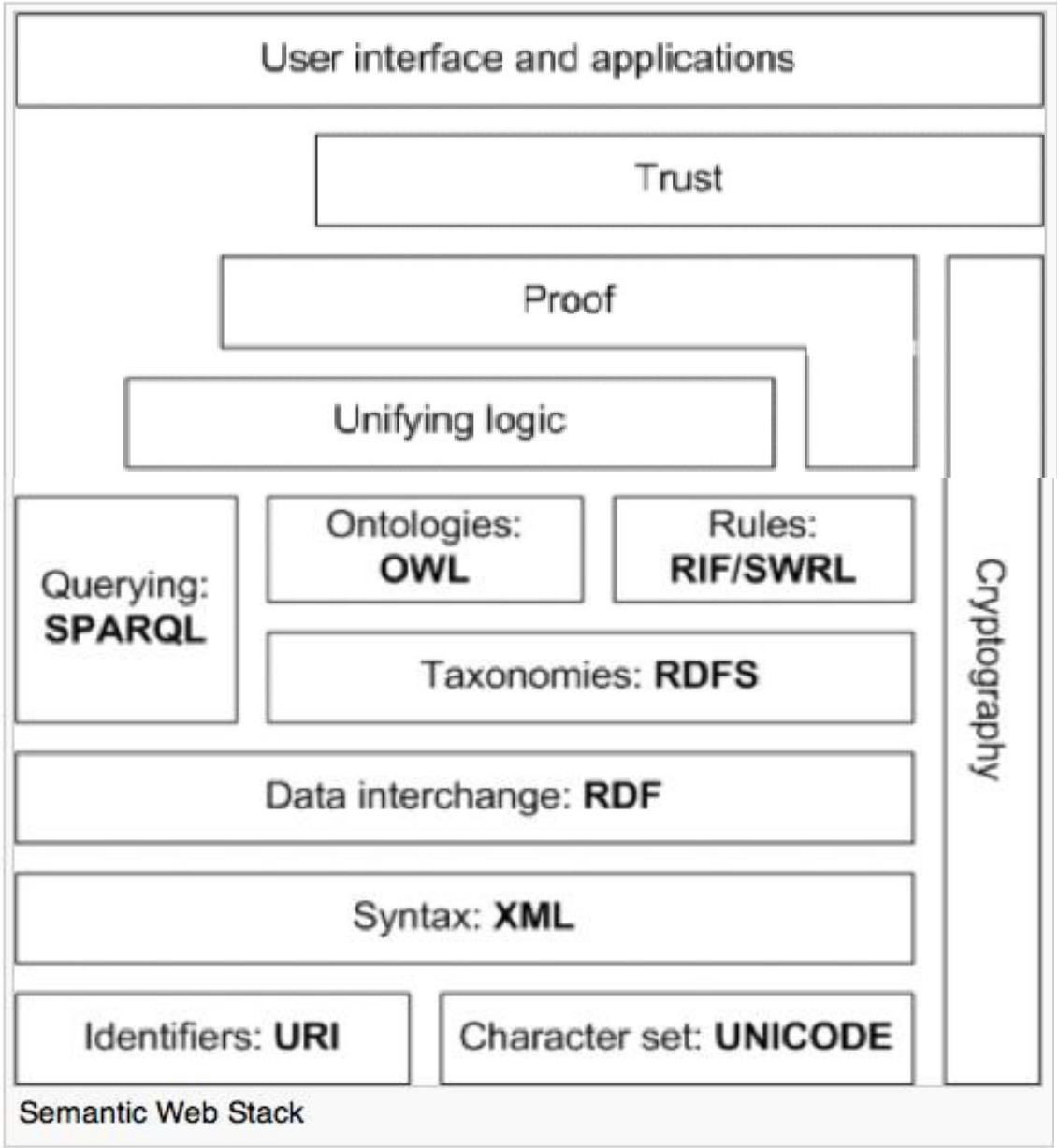




Data in various formats

The Semantic Web Stack

- **Unrealized Semantic Web technologies**
  - **Cryptography**
  - **Trust**
  - **User Interface**

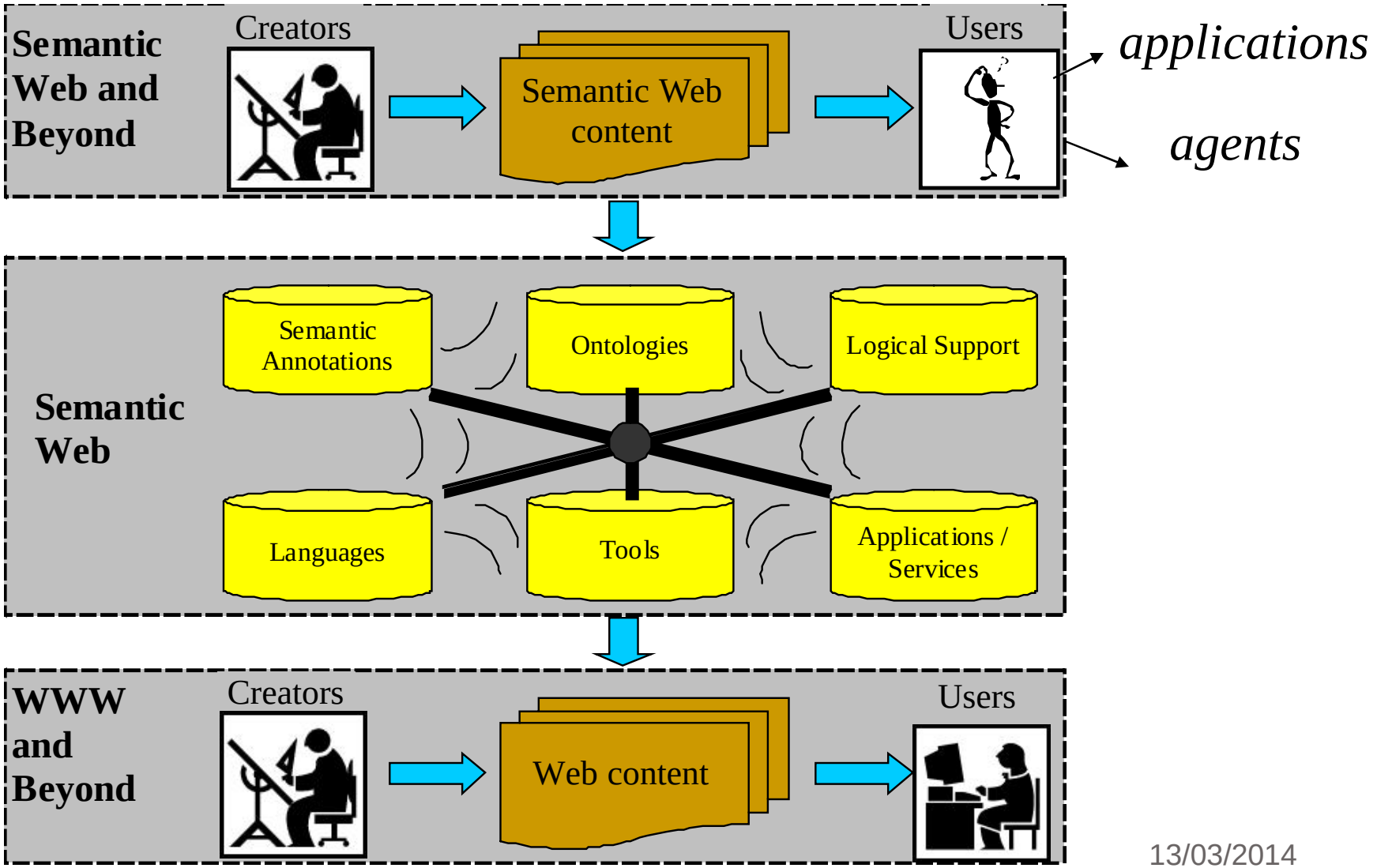


Semantic Web Stack

Tools

- Protégé: an ontology editor
  - create, edit, combine, reason, serialize, visualize ontologies
  - reasoner is pluggable
- Jena: a framework to provide a programmatic environment for RDF, RDFS and OWL
- Reasoners:
  - Fact++: supports OWL DL, open-source
  - Pellet: supports DL, EL, open-source
- RDF Gravity: RDF document visualizer

Towards...


Semantic Web and Beyond
Creators
Semantic Web content
Users
applications
agents
Semantic Web
Semantic Annotations
Ontologies
Logical Support
Languages
Tools
Applications / Services
WWW and Beyond
Creators
Web content
Users