



Calcolatori, Internet e il Web

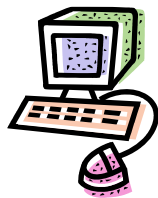
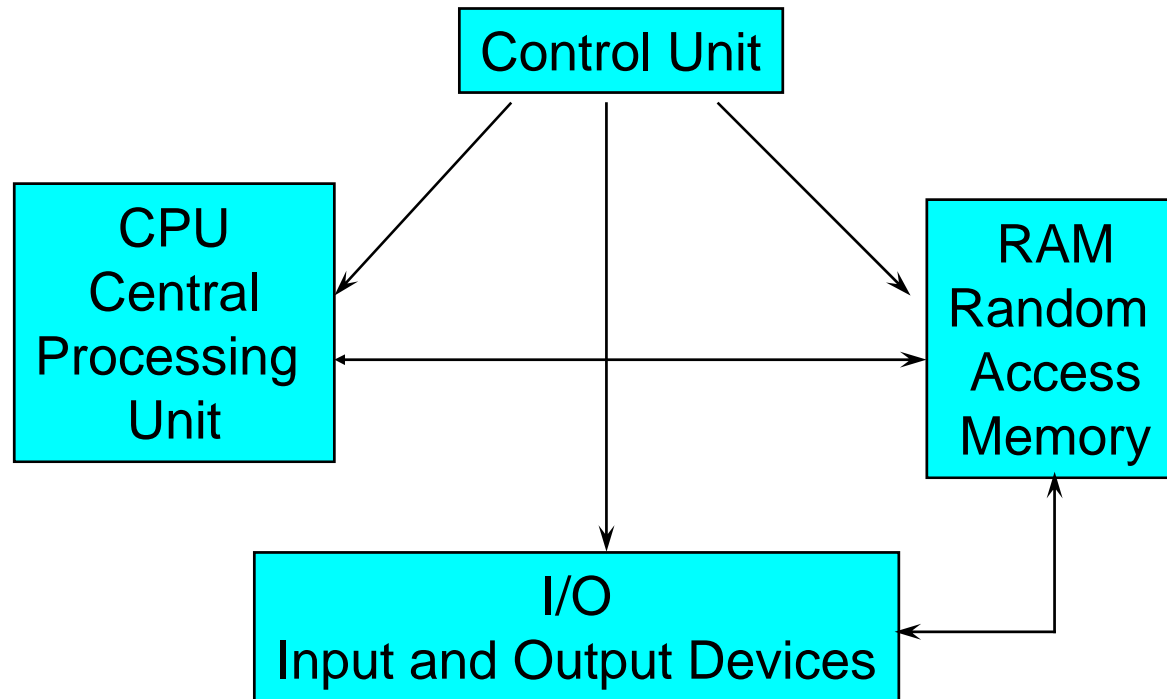
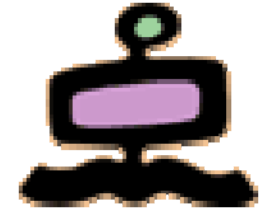


Refresher on Computer Fundamentals and Networking

- History of computers
- Architecture of a computer
- Data representation within a computer
- Computer networks and the Internet
- The World Wide web



Basic components of a computer





Evolution of computer technology



- First Generation
mechanical/electromechanical
- Second Generation
vacuum tubes
- Third Generation
discrete transistors (solid state devices)
SSI, MSI, LSI integrated circuits
- Fourth Generation
VLSI integrated circuits

VLSI = Very Large Scale Integration



Evolution of computer components



- Computer technology
 - CPU on integrated chips
 - From KHz to MHz to GHz
 - Random Access Memories
 - RAM – from KB to GB
 - External memories
 - Tapes, hard disks, floppy disks
 - CDs
 - DVDs
 - Memory sticks
 - SSD
 - from MB to GB to TB to PB to EB



Size of digital information



1000	k	kilo
1000 ²	M	mega
1000 ³	G	giga
1000 ⁴	T	tera
1000 ⁵	P	peta
1000 ⁶	E	exa
1000 ⁷	Z	zetta
1000 ⁸	Y	yotta



Evolution of computer market (1940-2000)



- Military applications in early 40s
- Scientific/research applications in late 40s
- Commercial applications appear in early 50s
- Monopoly of IBM starts with 650, 701, 702
- Monopoly of IBM continues with 7070, 7090 and the 360 series, starting the “mainframe era” (in the 60s and 70s)
- Arrival of the “minicomputers” in the 70s
- Arrival of the PC in the 80s
- Arrival of the Internet in the 90s
- Arrival of the Web in the 90s



A “mainframe” in the 60’





A “mainframe” in the 70’



Photograph: Dominic Hart/NASA Ames



Evolution of computer market (1940-2000)



- Military applications in early 40s
- Scientific/research applications in late 40s
- Commercial applications appear in early 50s
- Monopoly of IBM starts with 650, 701, 702
- Monopoly of IBM continues with 7070, 7090 and the 360 series, starting the “mainframe era” (in the 60s)
- Arrival of the “minicomputers” in the 70s
- Arrival of the PC in the 80s
- Arrival of the Internet in the 90s
- Arrival of the Web in the 90s



Minicomputers





Early PCs





Evolution of technology



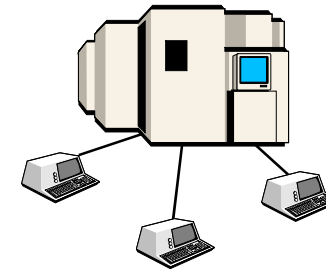
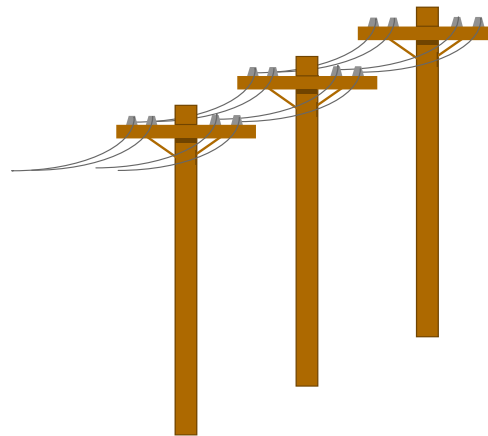
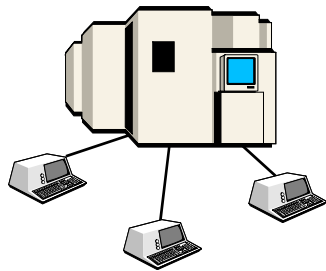
- Computer technology
 - CPU and integrated chips
 - Random Access Memories
 - RAM – from KB to GB
 - External memories
 - Tapes, hard disks, floppy disks
 - Memory sticks
 - CDs
 - DVDs
 - from MB to GB to TB to PB to EB
- Communication technology (networks)
 - (Telephone) line speed
 - Point to point (leased lines)
 - Local Area Networks
 - Inter-networking (TCP/IP)



Early computer communication



From mainframe to mainframe
through telephone lines
(point to point connection)



Telephone lines:
slow
expensive
regulated



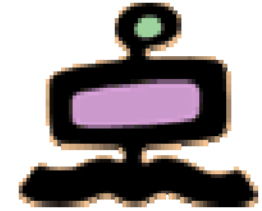
Networking



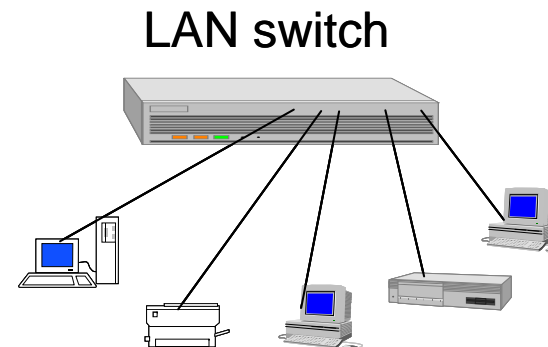
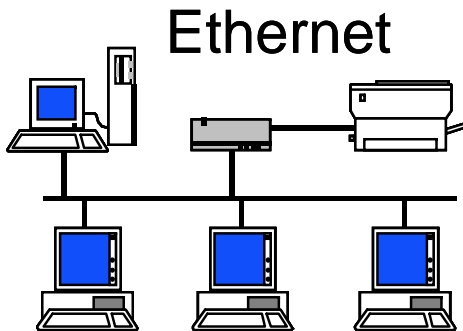
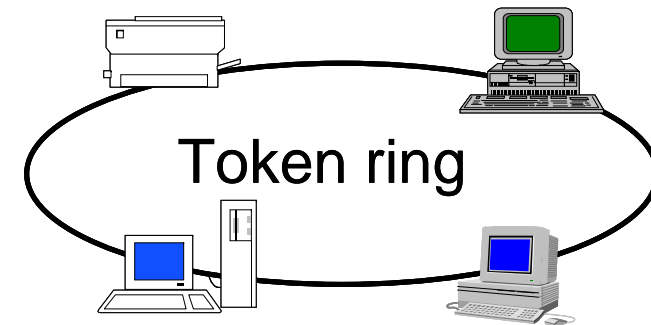
- In the sixties, first studies on “networking”
 - Networking means communication between node A and node B through one or more intermediate nodes
- In the seventies, fragmentation of the market with the arrival of “minicomputers” provided further motivation for research on networking
- At the same time (in the seventies), the arrival of the LANs (Local Area Networks) provided the final impulse for the development of networking



LAN - Local Area Networks



Private networks
Up to several kilometers
Speed up to 100 Mb/sec





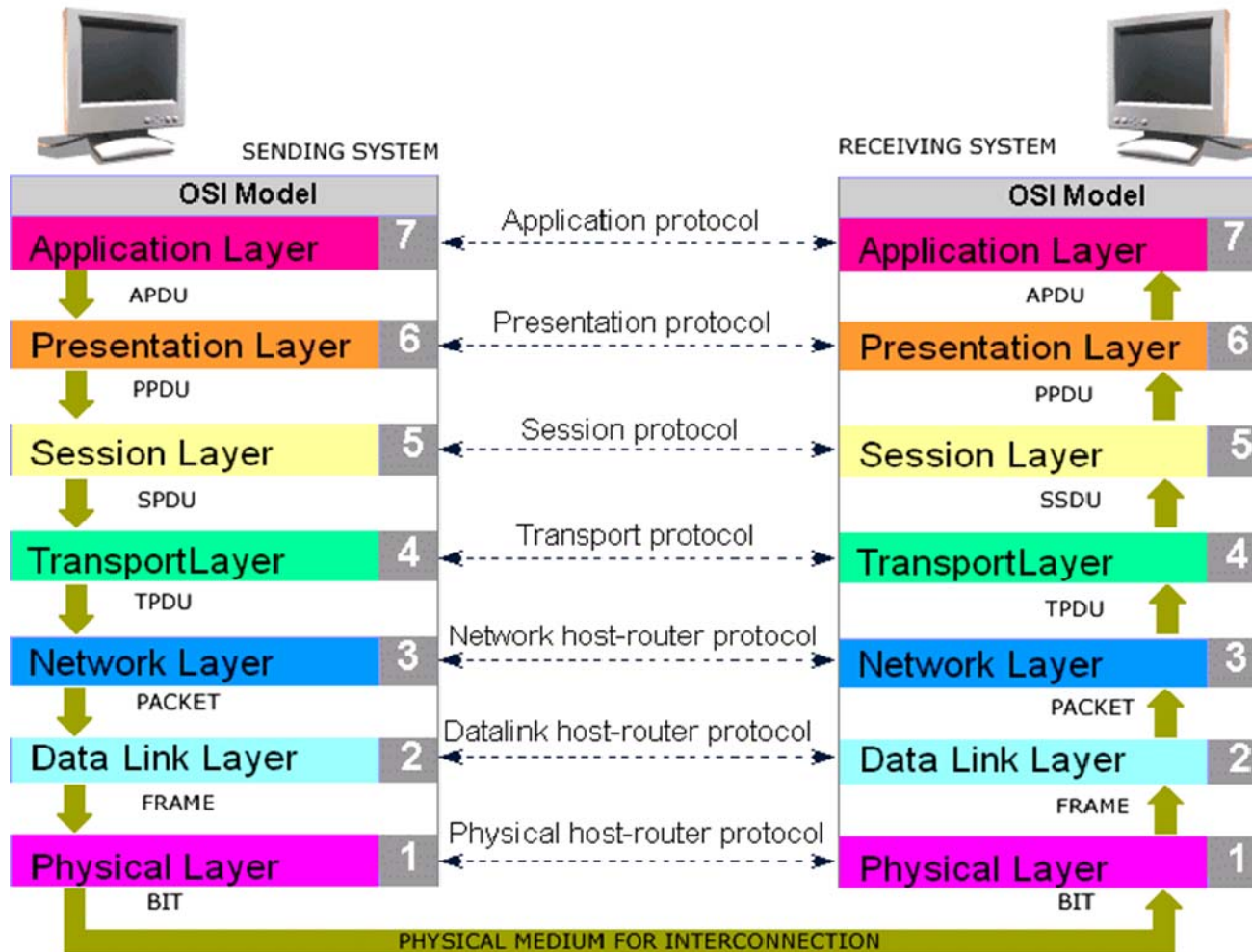
Research on networking



- Starting in the late sixties, many research projects on networking, both from universities and industry
 - Arpanet, Cyclades, SNA, DECnet
- In the late seventies ISO (International Standard Organization), under pressure of a group of computer manufacturer, started the work for the proposal of a “new” communication standard, called OSI: Open System Interconnection
- The OSI model, though no longer in use today, has established a number of networking concepts and is still used as a “reference model”
- The main concept introduced by OSI is the “communication layer”



The OSI model

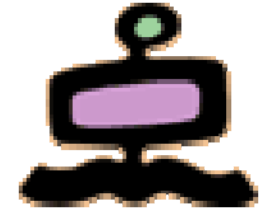


Protocol:
formats and rules
for exchanging
messages between
“partners”

Packet switching:
messages are
broken down into
“packets”, and each
packet gets to
destination
independently from
the others.



Mnemonics for OSI layers



All	Application
People	Presentation
Seem	Session
To	Transport
Need	Network
Data	Data Link
Processing	Physical

Please	Physical
Do	Data Link
Not	Network
Throw	Transport
Sausage	Session
Pizza	Presentation
Away	Application



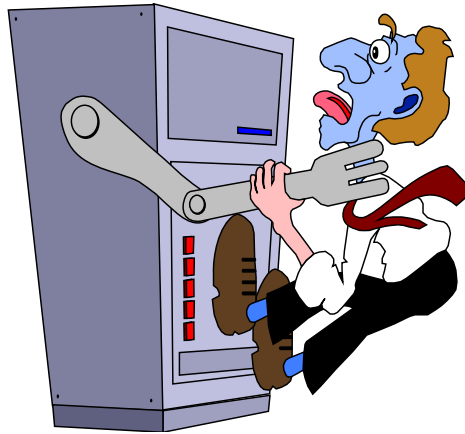
OSI and Internet



- The OSI effort provided a sound and durable foundation for networking, but never became a “market leader”
 - Slow development
 - Initial opposition from IBM
 - “Designed by a Committee”
 - Expensive development
 - Heavy and slow in operation
- In the same period the Internet was defining a number of “light weight” protocols
- Most of the market preferred them to OSI



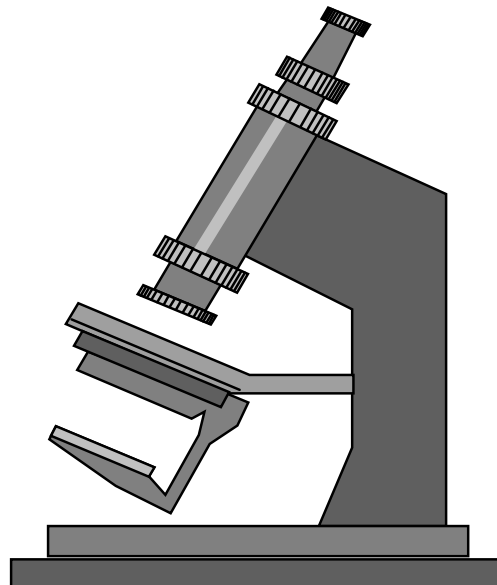
Internet evolution



Reserach
Network
NSF
Internet
In the 80'



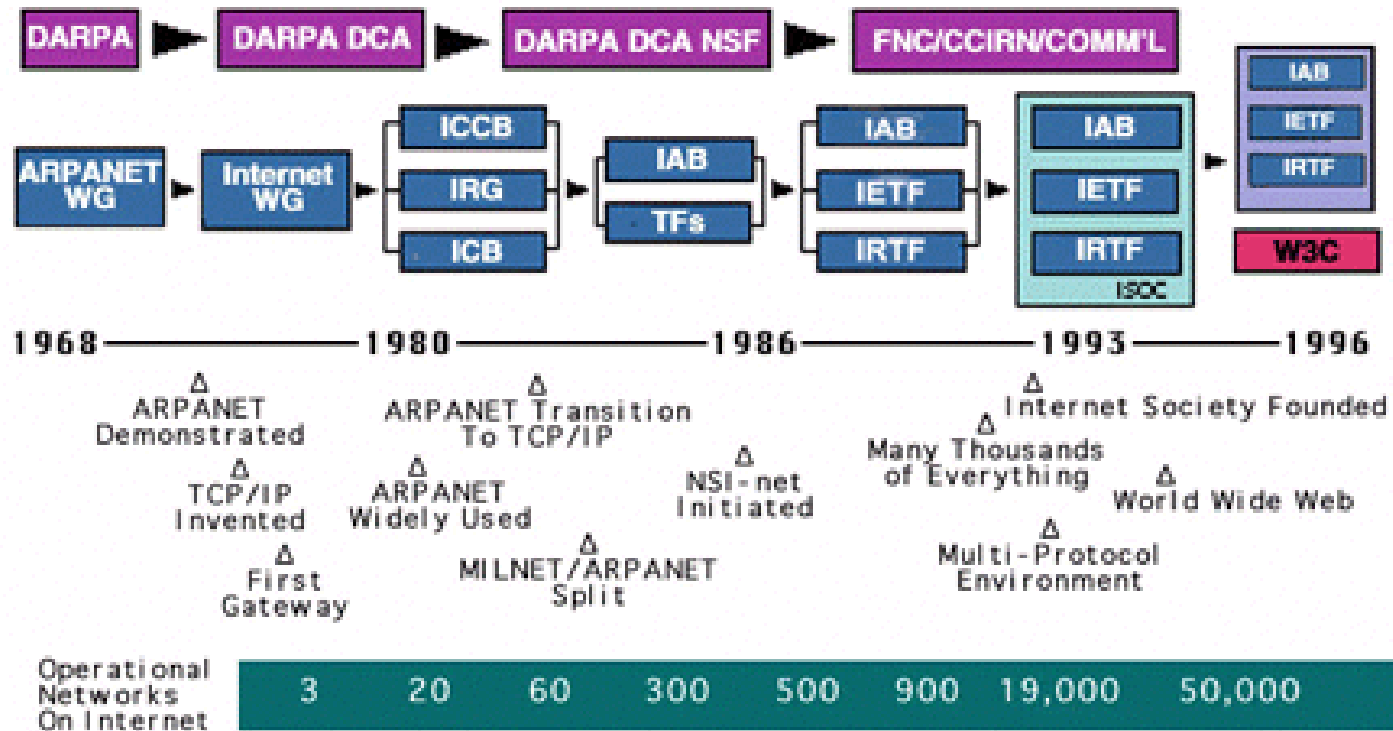
Experimental
Network
DARPA
Arpanet
In the 60' and 70'



Communication
Infrastructure
Private and public
sectors
The Web
In the 90'



Internet timeline

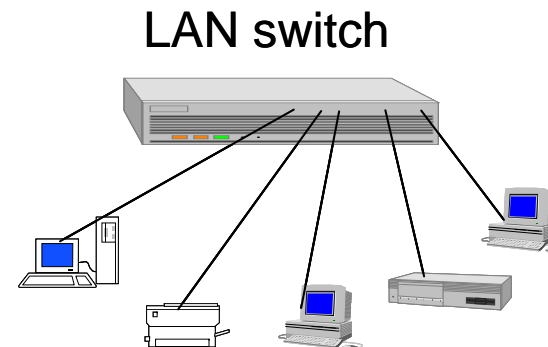
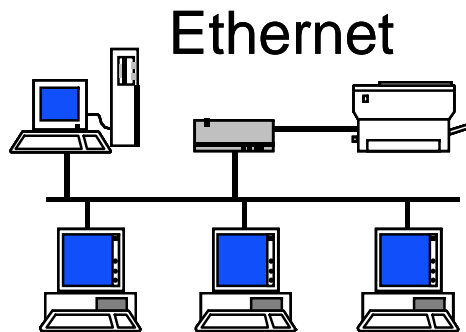
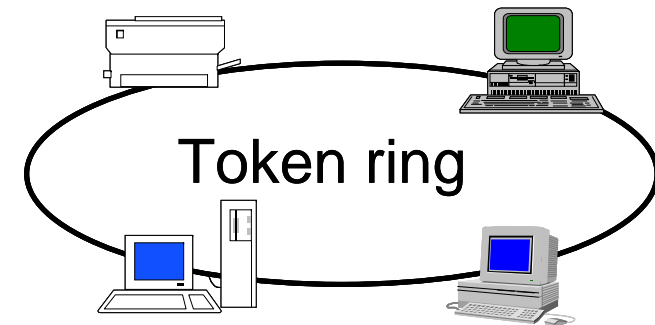




LAN - Local Area Networks

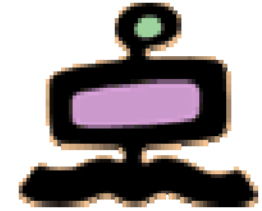


Private networks
Up to several kilometers
Speed up to 100 Mb/sec

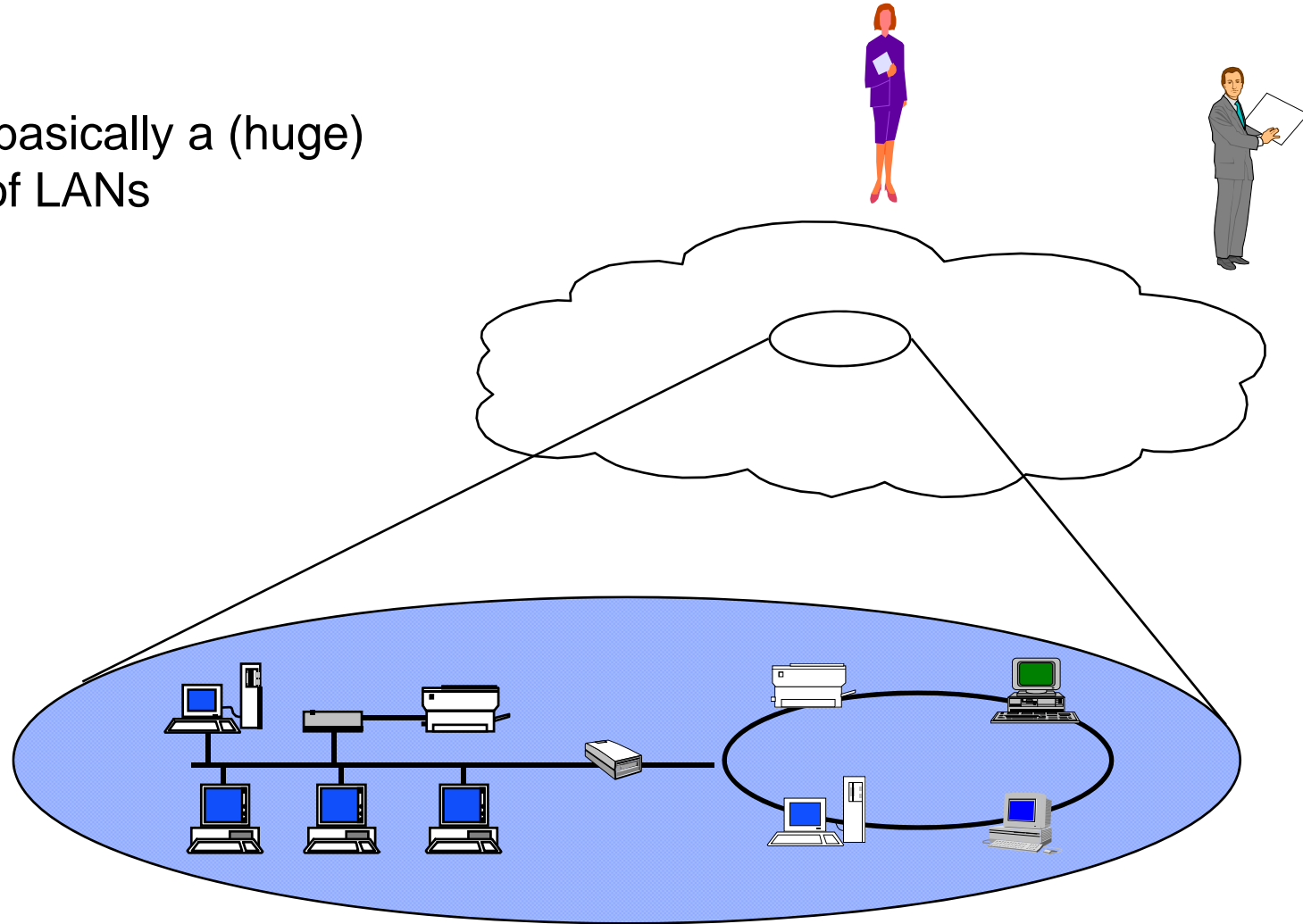




Inter-networking



Internet is basically a (huge) collection of LANs

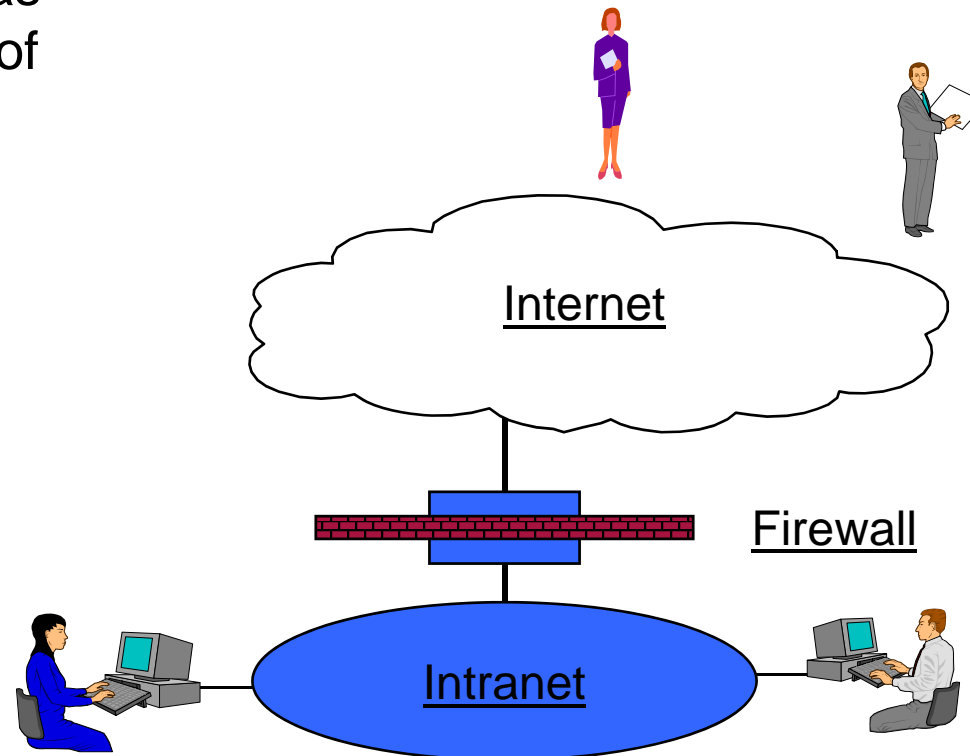




Internet and Intranets

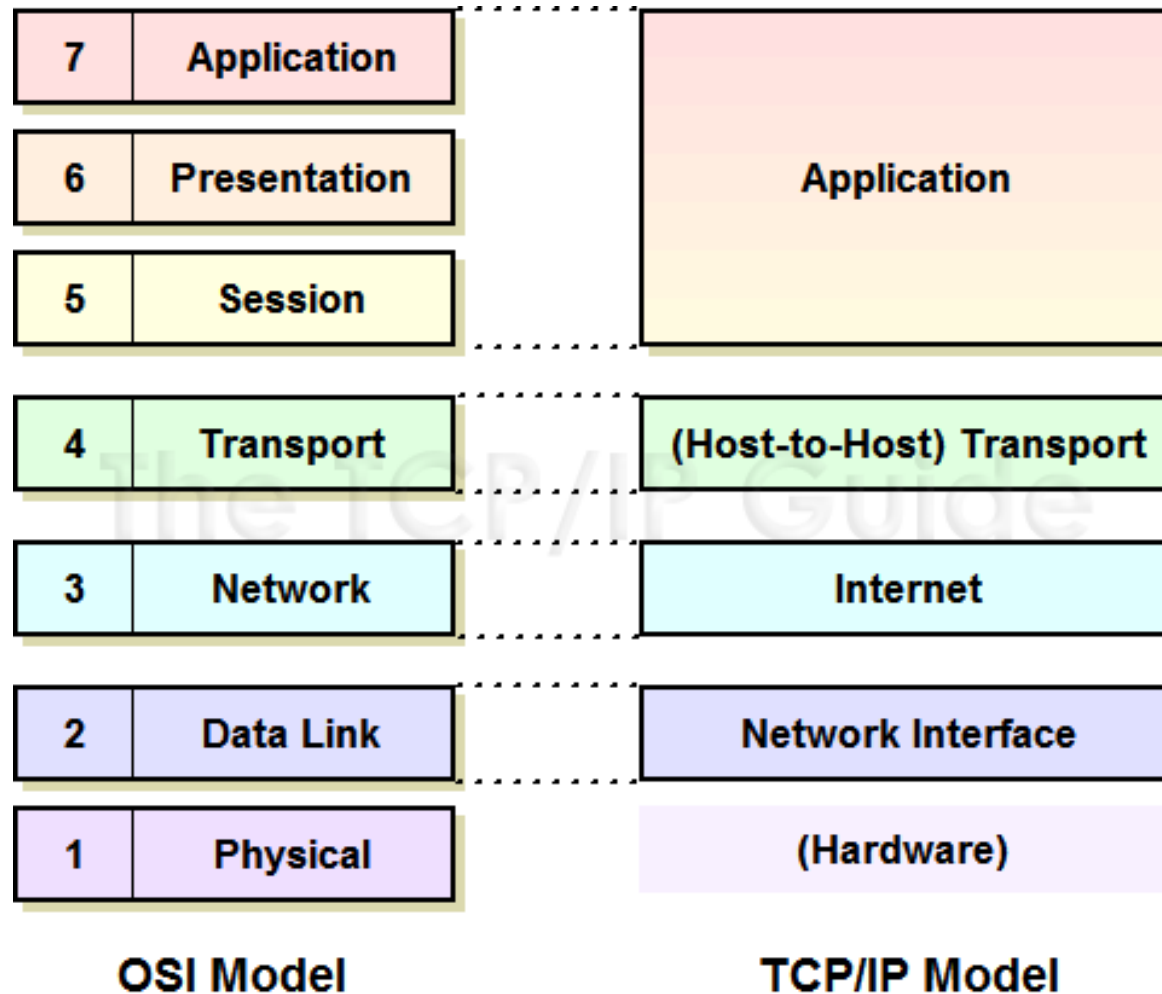


The growth of Internet was also due to the adoption of the Internet protocols by private companies



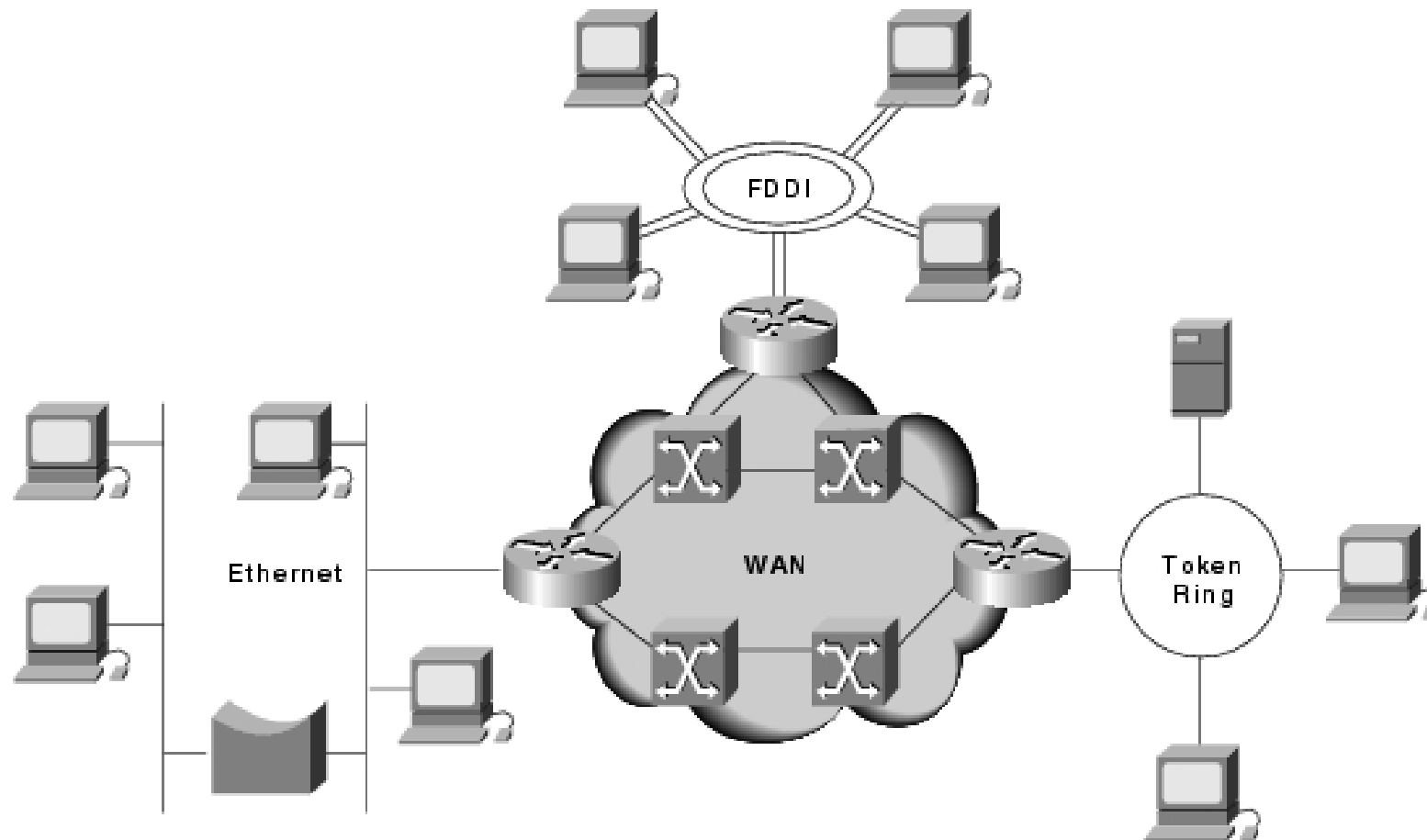


OSI and TCP/IP



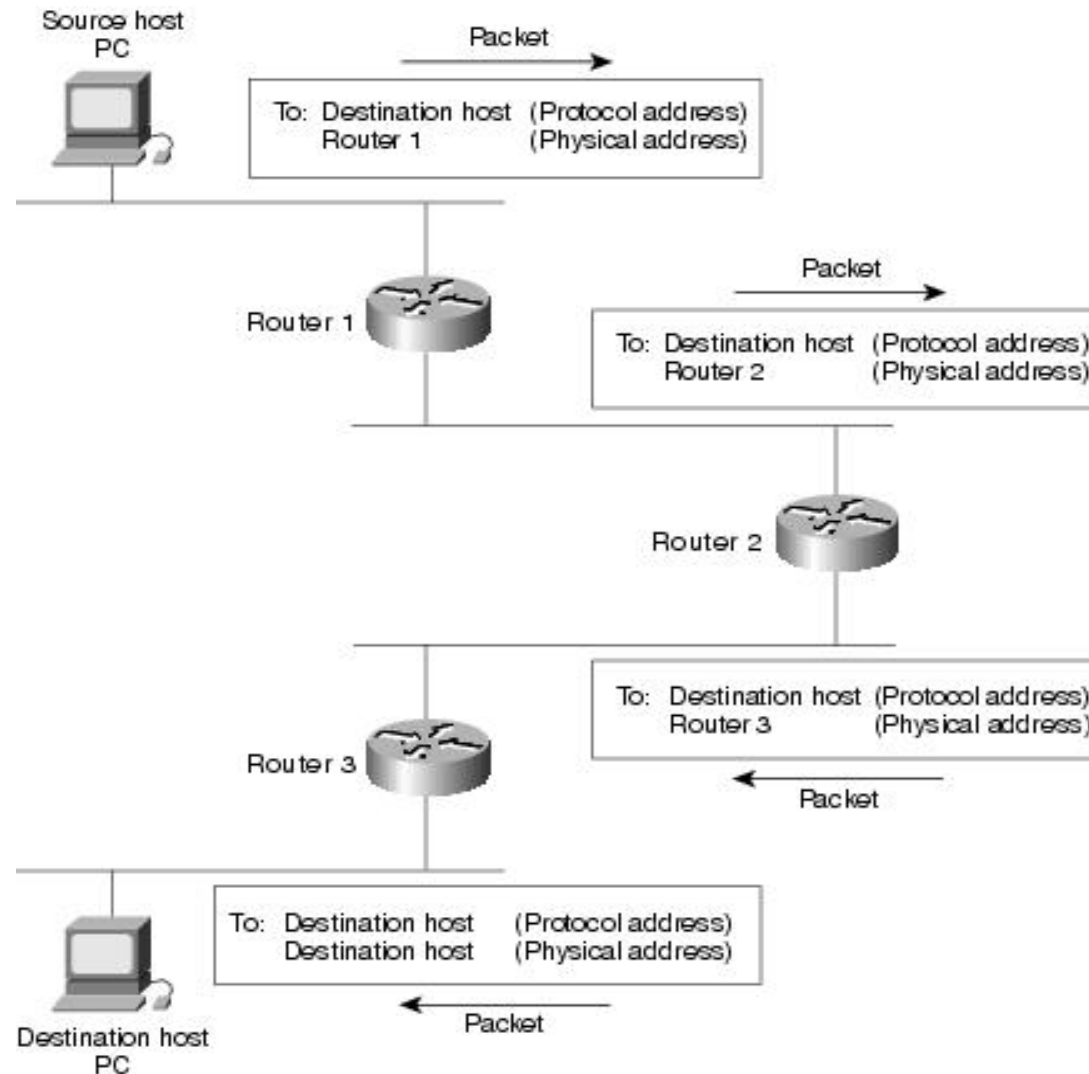


Internetworking





Routing in Internet

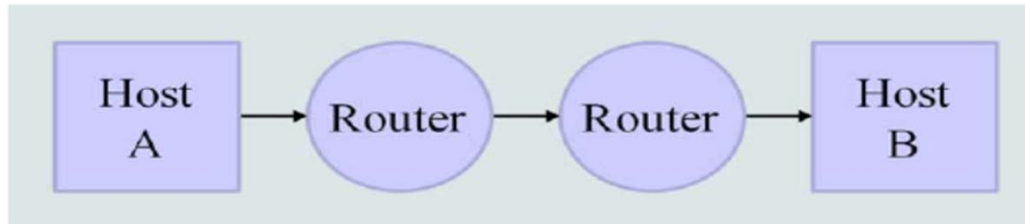




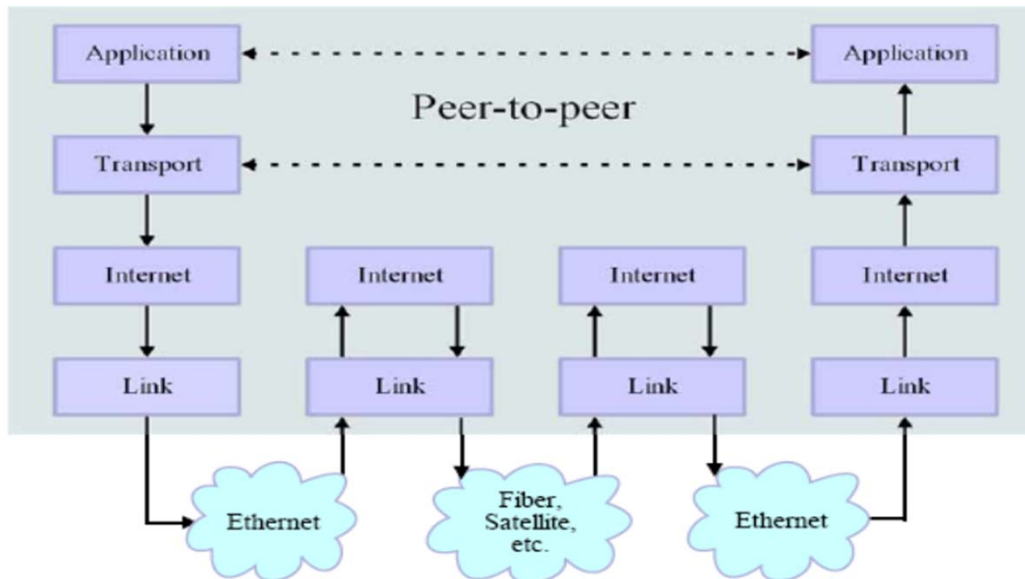
Internet protocols



Network Connections



Stack Connections



Application protocols

TCP/UDP

IP

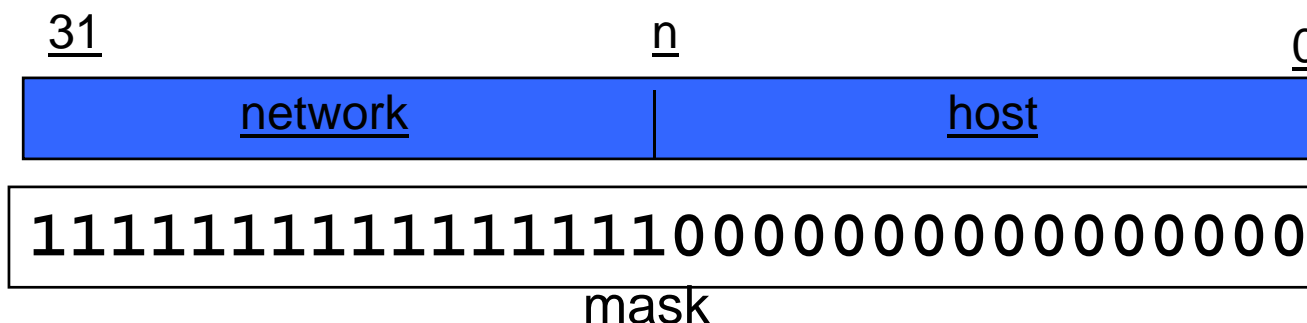
ETHERNET



IPv4 addressing



- Each node in the Internet is identified by (one or more) IP address, and each IPv4 address has 32 bits (4 bytes)
- An IP address is (was) made of two parts: the network address and the node address within the network
- The boundary between the parts is variable, and is identified by the “network mask”
- The 1s in the mask identify the net portion and the 0s the host portion





IPv4 addresses



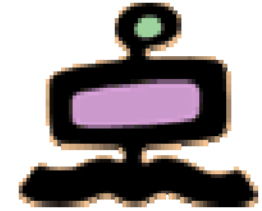
- An IP address is usually indicated with four numbers (from 0 to 255) corresponding to the 4 bytes of the address

IP address: 131.114.1.30
mask: 255.255.255.0
network address 131.114.1
host address 30

- Three classes of network addresses (255.0.0.0, 255.255.0.0, 255.255.255.0)
- No more IPv4 addresses available today
 - Network Address Translation (NAT) commonly used
- IPv6 (128 bits) slowly replacing IPv4



Evolution of computer market



- Military applications in early 40s
- Scientific/research applications in late 40s
- Commercial applications appear in early 50s
- Monopoly of IBM starts with 650, 701, 702
- Monopoly of IBM continues with 7070, 7090 and the 360 series, starting the “mainframe era” (in the 60s)
- Arrival of the “minicomputers” in the 70s
- Arrival of the PC in the 80s
- Arrival of the Internet in the 90s
- Arrival of the Web in the 90s



The World Wide Web



- Combination of computer technology and communication technology
- It all started with the “hyperlink”
- Then came the “browser” (Mosaic)
- Then came the first wave
- Then came the “dot come, dot gone”
- Then came the second wave
- Finally came the “information explosion”
 - An estimate of 500 to 1000 million hosts
 - An estimate of 30 to 50 billion pages on line
- And now we are in Web 2.0 (with Web 3.0 already happening)



The editors



- Text processing applications started already in the early days of the computers (sixties)
- A “text processor” (or editor) has two main functions:
 - processing the text (delete, replace, insert, etc.)
 - specifying the format (bold, center, new line, etc.)
- The first editors were using a “mark up” language (i.e. commands intermixed with the text) to provide formatting instructions (only limited interactivity available through typewriter-like terminals)
- The “second generation” editors were using the WYSIWYG paradigm: What You See Is What You Get (much better interactivity available with display and mouse)



The hyperlink



- The idea of the “hyperlink” was (experimentally) proposed in the sixties, as a feature of a “smart editor”
 - selecting a portion of the text, it was possible to open a second document, in addition to the one being edited (very awkward to use on a typewriter-like terminal)
- With the arrival of display screens and the mouse (eighties) the hyperlink came back in “3D documents”
 - clicking on a portion of the text it was possible to open a second document, which was maintained as a second (virtual) screen behind the first one
- With the arrival of the (fast) internet, it became the “web hyperlink”
 - clicking on a portion of the text it was possible to open a second document, coming from a different computer



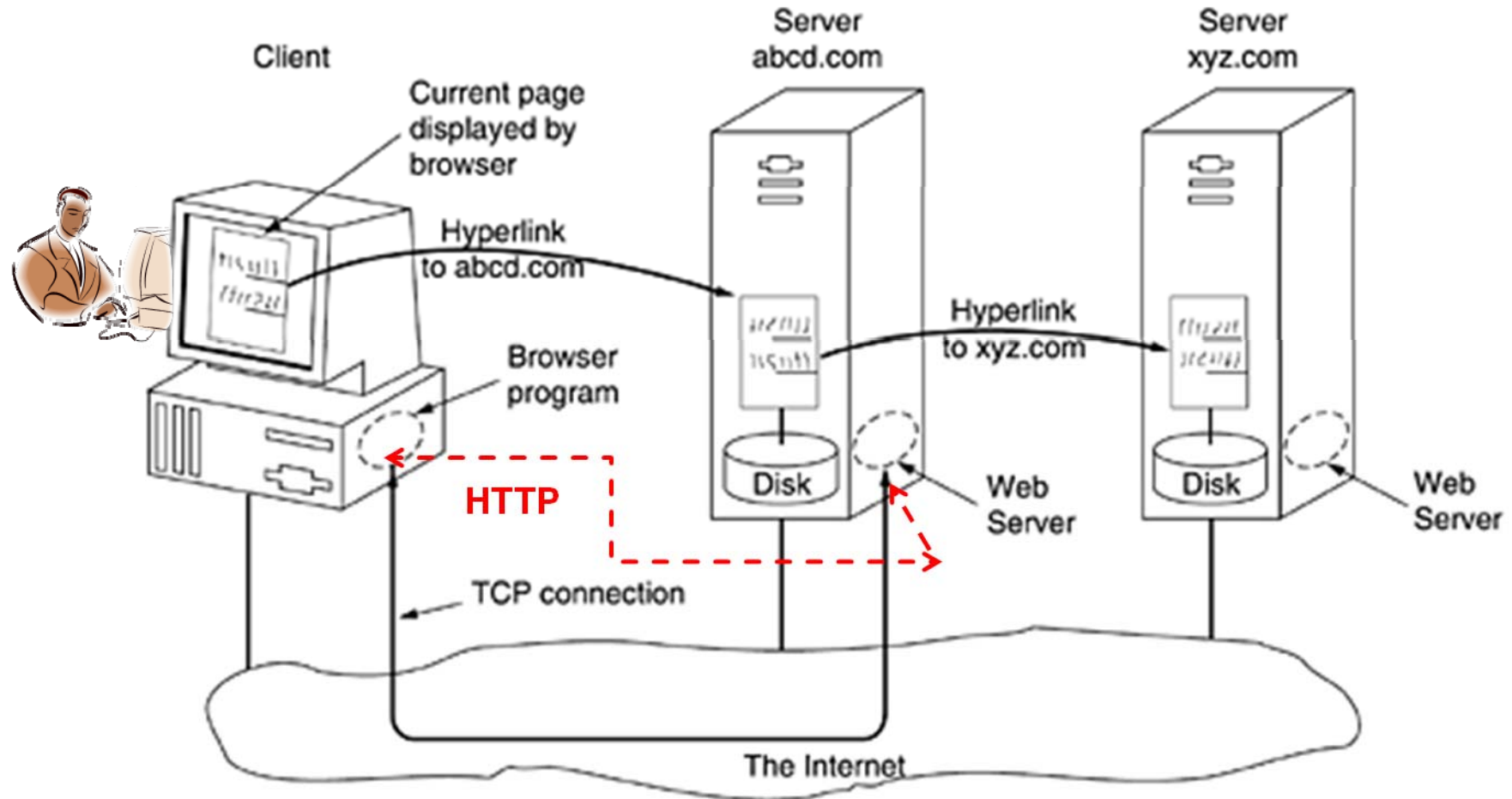
The browser



- With the arrival of the (web) hyperlink, the problem was then how to properly display a (web) page that had been generated on a different computer, possibly with a different (wysiwyg) editor
- The solution was the definition of HTML (Hyper Text Markup Language), i.e. a standard mark up language, and the implementation of smart editors (the browser) capable of correctly displaying pages formatted with HTML, regardless of where they were coming from



The Web architecture





The World Wide Web



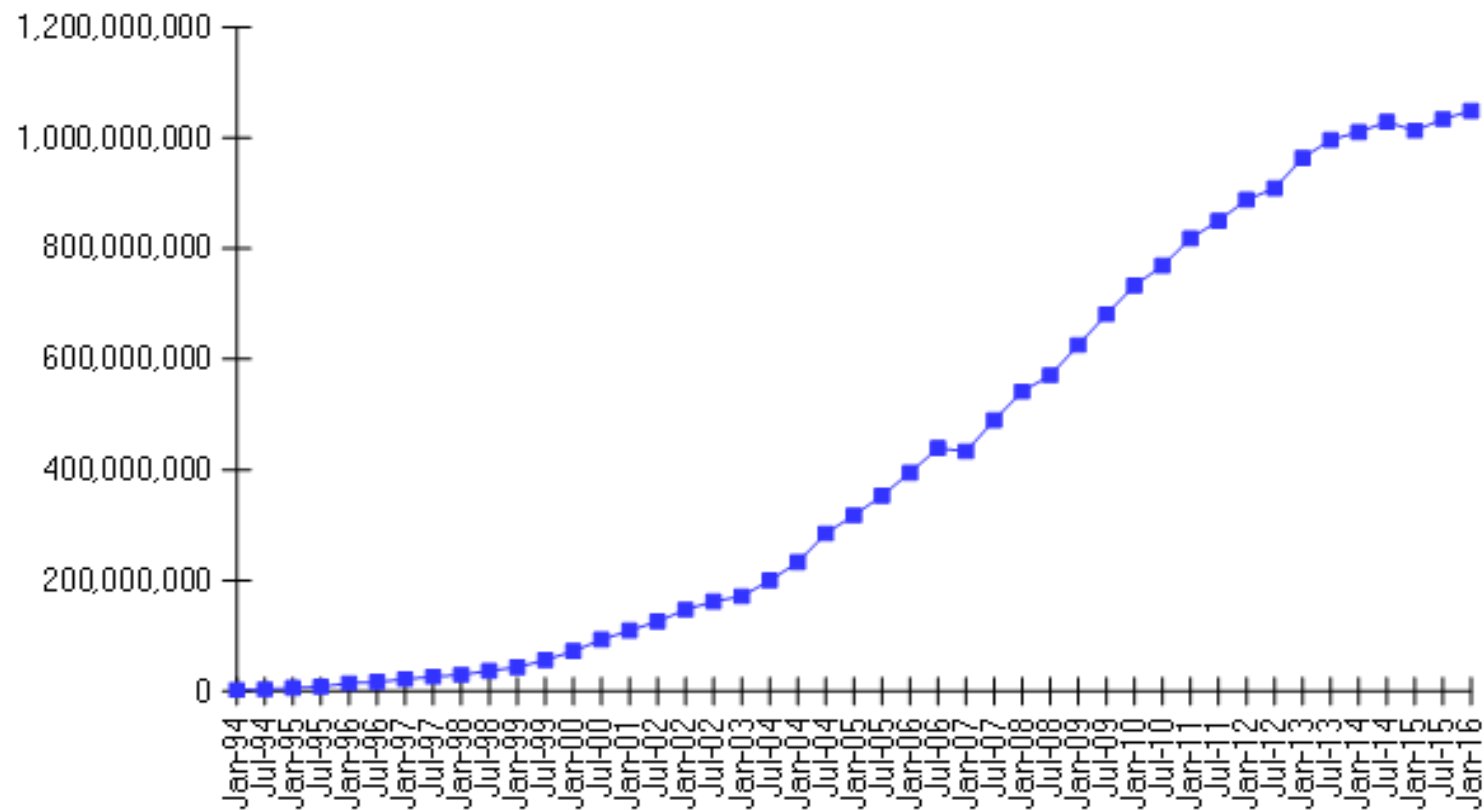
- Combination of computer technology and communication technology
- It all started with the “hyperlink”
- Then came the “browser” (Mosaic)
- Then came the first wave
- Then came the “dot come, dot gone”
- Then came the second wave
- Finally came the “information explosion”
 - An estimate of 500 to 1000 million hosts
 - An estimate of 30 to 50 billion pages on line
- And now we are in Web 2.0 (with Web 3.0 already happening)



Number of hosts



Internet Domain Survey Host Count



Source: Internet Systems Consortium (www.isc.org)



Internet users 2016



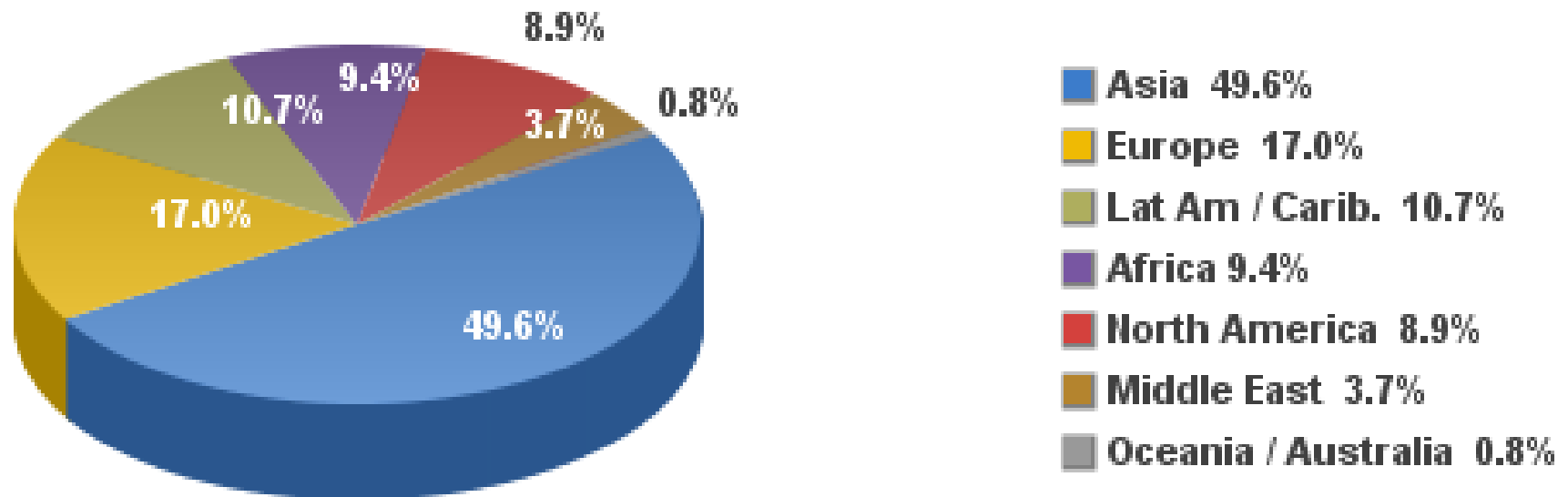
WORLD INTERNET USAGE AND POPULATION STATISTICS JUNE 30, 2016 - Update						
World Regions	Population (2016 Est.)	Population % of World	Internet Users 30 June 2016	Penetration (% Population)	Growth 2000-2016	Users % of Table
<u>Africa</u>	1,185,529,578	16.2 %	339,283,342	28.6 %	7,415.6%	9.4 %
<u>Asia</u>	4,052,652,889	55.2 %	1,792,163,654	44.2 %	1,467.9%	49.6 %
<u>Europe</u>	832,073,224	11.3 %	614,979,903	73.9 %	485.2%	17.0 %
<u>Latin America / Caribbean</u>	626,054,392	8.5 %	384,751,302	61.5 %	2,029.4%	10.7 %
<u>Middle East</u>	246,700,900	3.4 %	132,589,765	53.7 %	3,936.5%	3.7 %
<u>North America</u>	359,492,293	4.9 %	320,067,193	89.0 %	196.1%	8.9 %
<u>Oceania / Australia</u>	37,590,704	0.5 %	27,540,654	73.3 %	261.4%	0.8 %
<u>WORLD TOTAL</u>	7,340,093,980	100.0 %	3,611,375,813	49.2 %	900.4%	100.0 %



Distribution by world regions



Internet Users in the World by Regions June 2016



Source: Internet World Stats - www.internetworldstats.com/stats.htm

Basis: 3,611,375,813 Internet users on June 30, 2016

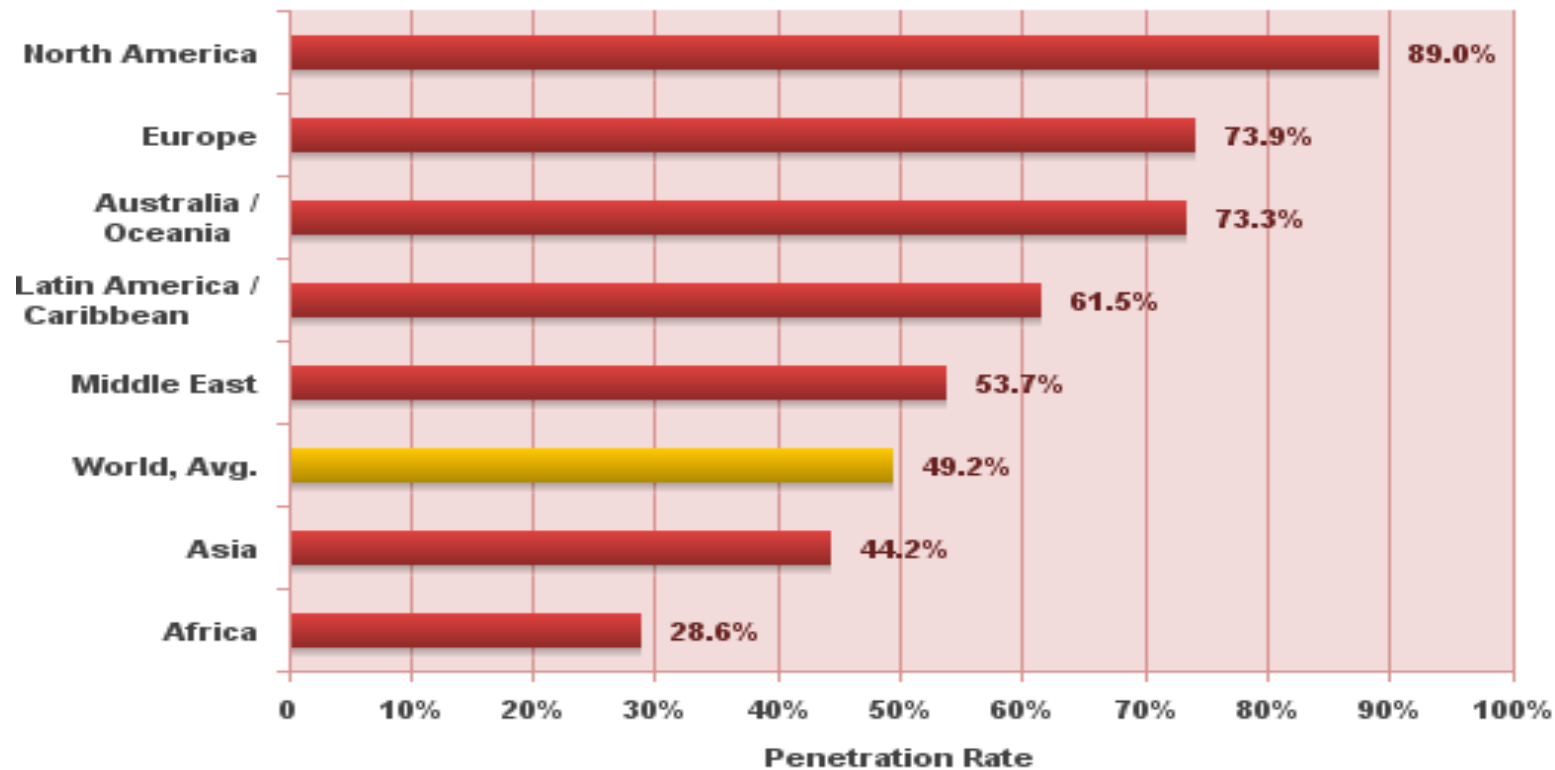
Copyright © 2016, Miniwatts Marketing Group



Internet world penetration



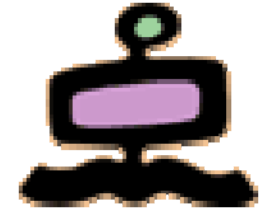
Internet World Penetration Rates by Geographic Regions - June 2016



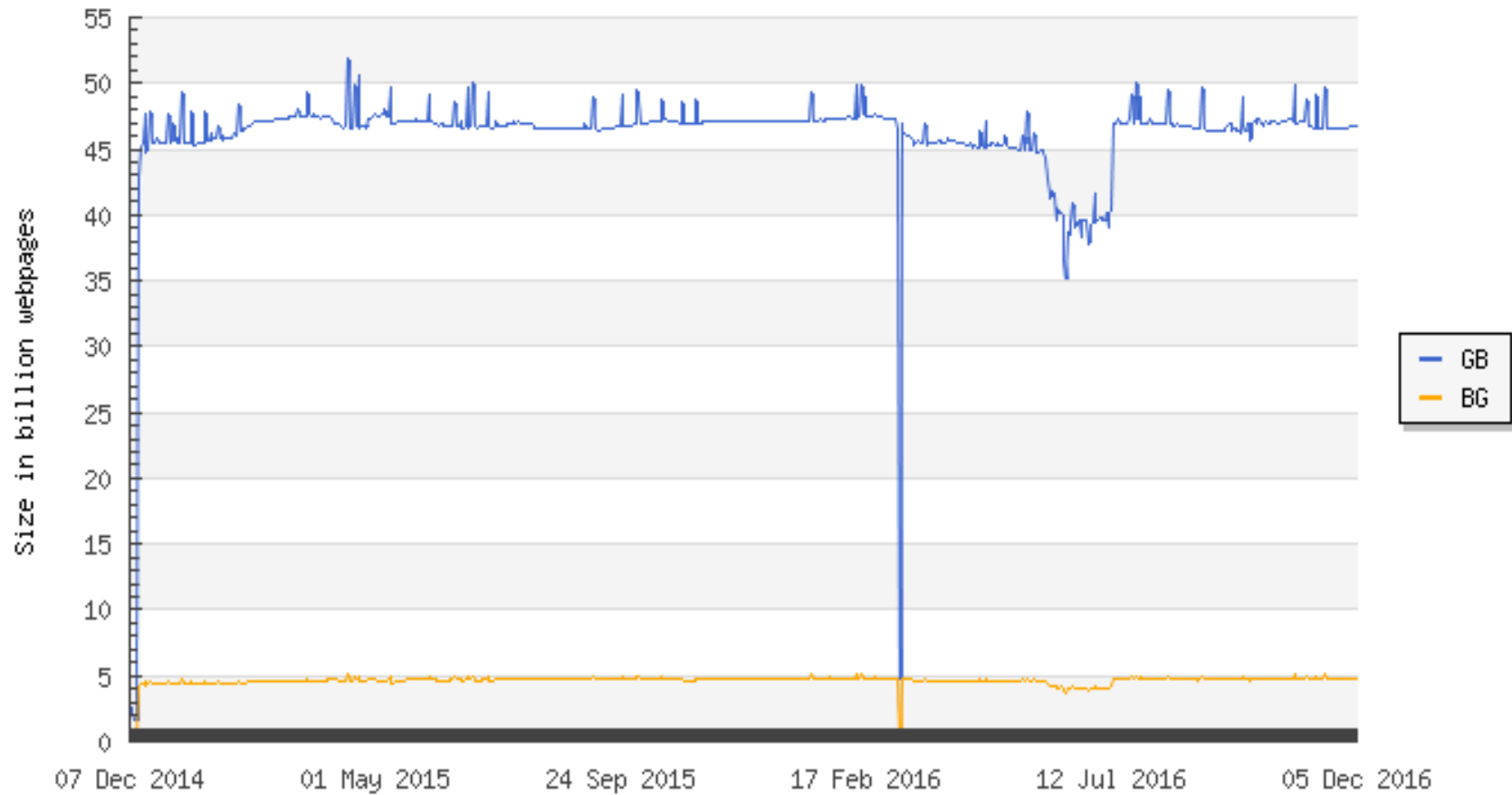
Source: Internet World Stats - www.internetworldstats.com/stats.htm
Penetration Rates are based on a world population of 7,340,093,980 and 3,611,375,813 estimated Internet users on June 30, 2016.
Copyright © 2016, Miniwatts Marketing Group



The size of the indexed web



The size of the indexed World Wide Web
(Number of webpages)



<http://www.worldwidewebsite.com/>



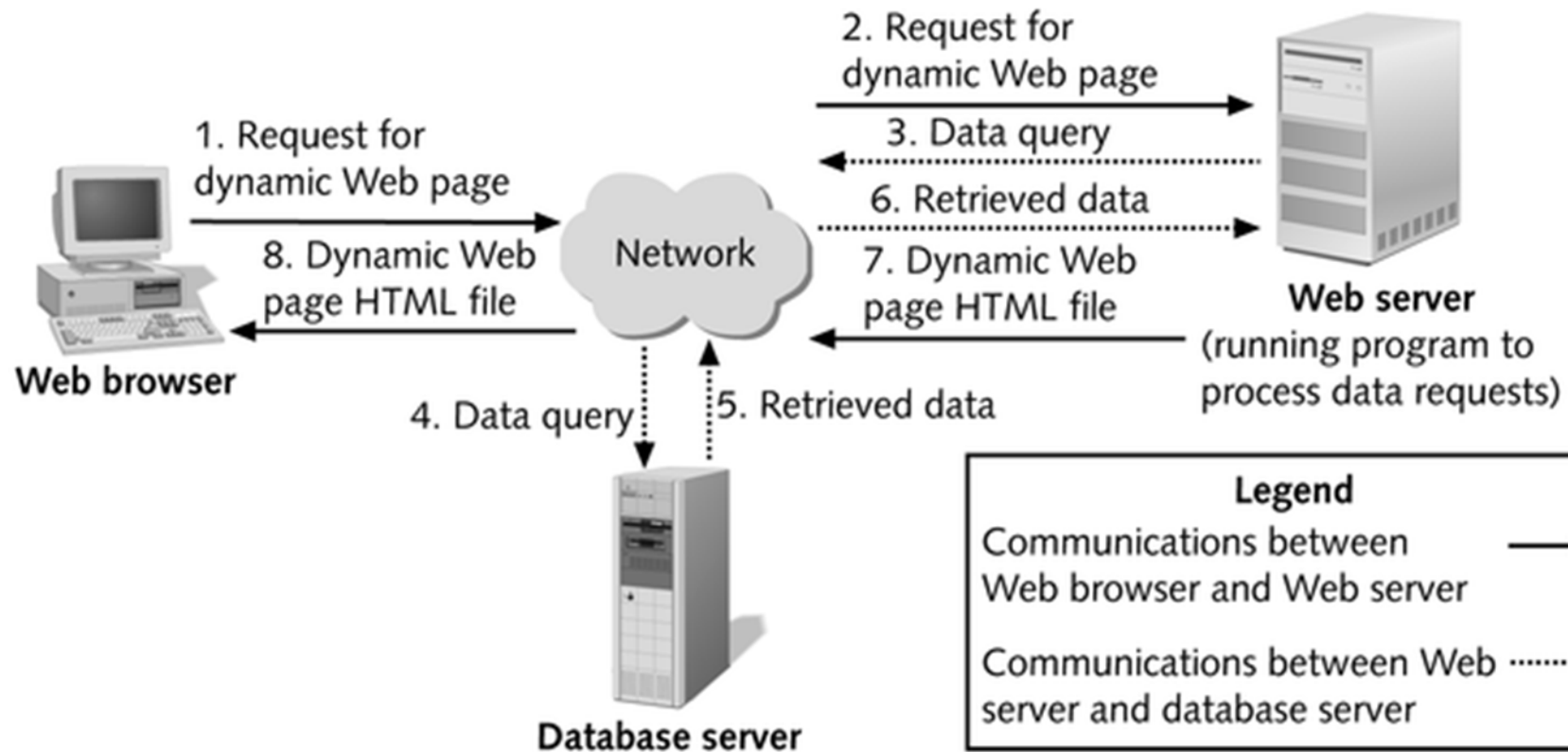
The Depth of the Web



- A URL gives access to a web page.
- That page may have links to other pages (static pages). This is the **surface web**.
- Some pages (dynamic pages) are generated only when some information is provided to the web server.
- These pages cannot be discovered just by crawling. This is the **deep web**.
- The surface web is huge.
- The deep web is “unfathomable”.



Dynamic web pages (data base driven)

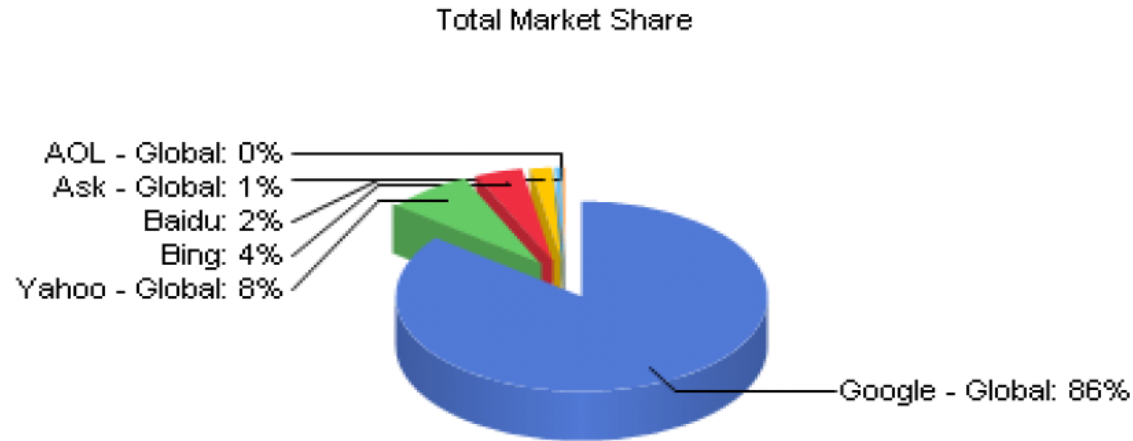




Worldwide queries to search engines September 2012



Week of September 16, 2012



Search Engine	Total Market Share
Google - Global	84.78%
Yahoo - Global	7.51%
Bing	4.01%
Baidu	1.81%
Ask - Global	0.51%
AOL - Global	0.30%
Excite - Global	0.02%
Lycos - Global	0.00%



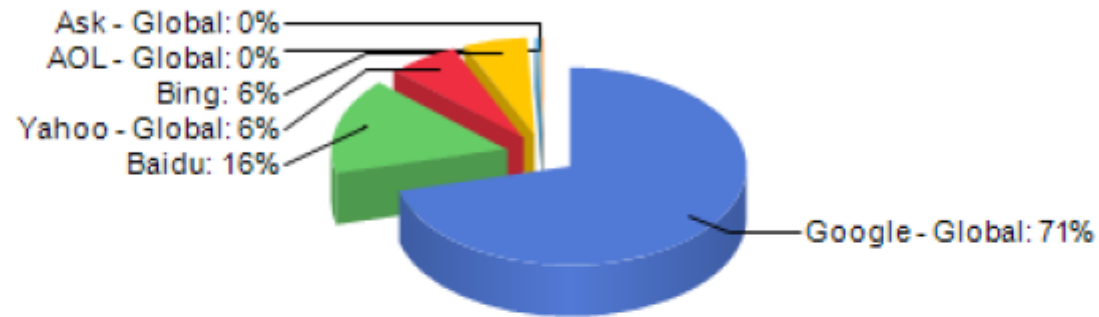
Worldwide queries to search engines September 2013



September, 2013

<http://www.netmarketshare.com/>

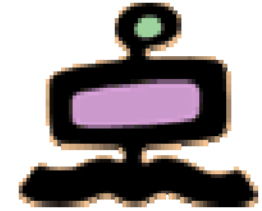
Total Market Share



Search Engine	Total Market Share
Google - Global	70.53%
Baidu	16.04%
Yahoo - Global	6.43%
Bing	5.70%
AOL - Global	0.36%
Ask - Global	0.33%
Excite - Global	0.01%

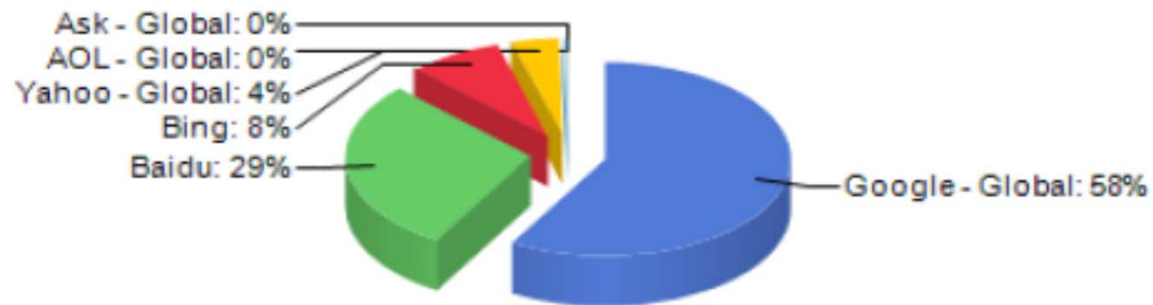


Worldwide queries to search engines October 2014



October, 2014

Total Market Share



Search Engine	Total Market Share
Google - Global	58.01%
Baidu	29.06%
Bing	8.10%
Yahoo - Global	4.01%
AOL - Global	0.21%
Ask - Global	0.10%
Excite - Global	0.00%



Google searches



Year	Annual Number of Google Searches	Average Searches Per Day
2013	2,161,530,000,000	5,922,000,000
2012	1,873,910,000,000	5,134,000,000
2011	1,722,071,000,000	4,717,000,000
2010	1,324,670,000,000	3,627,000,000
2009	953,700,000,000	2,610,000,000
2008	637,200,000,000	1,745,000,000
2007	438,000,000,000	1,200,000,000
2000	22,000,000,000	60,000,000
1998	3,600,000 <i>*Googles official first year</i>	9,800



Google searches



Year	Annual Number of Google Searches	Average Searches Per Day
2016	3,293,250,000,000	9,022,000,000
2015	2,834,650,000,000	7,766,000,000
2014	2,095,100,000,000	5,740,000,000
2013	2,161,530,000,000	5,922,000,000
2012	1,873,910,000,000	5,134,000,000
2011	1,722,071,000,000	4,717,000,000
2010	1,324,670,000,000	3,627,000,000
2009	953,700,000,000	2,610,000,000
2008	637,200,000,000	1,745,000,000
2007	438,000,000,000	1,200,000,000
2000	22,000,000,000	60,000,000
1998	3,600,000 <i>*Googles official first year</i>	9,800



To google



© Original Artist
Reproduction rights obtainable from
www.CartoonStock.com



"Did you Google me before I was born to see if I was going to be a boy or a girl?"



Evolution of the Web



- Web 1.0 (1993-2003/4)
 - Web is a “publishing medium”
 - Users (humans) can only read
- Web 2.0 (2003/4-today)
 - Web is a “social medium”
 - Users (humans) can publish and interact (e.g. Youtube, Wiki, Flickr, Facebook, etc.)
- Web 3.0 (2010/1-today)
 - Users of the Web are “programs” that can interact
 - Users of the Web are “things”, whose programs interact with other things



Evolution of the Web



Web 1.0	Web 2.0	Web 3.0
Mostly Read-Only	Wildly Read-Write	Portable & Personal
Company Focus	Community Focus	Individual Focus
Home Pages	Blogs / Wikis	Lifestreams / Waves
Owning Content	Sharing Content	Consolidating Content
Web Forms	Web Applications	Smart Applications
Directories	Tagging	User Behavior
Page Views	Cost Per Click	User Engagement
Banner Advertising	Interactive Advertising	Behavioral Advertising
Britannica Online	Wikipedia	The Semantic Web
HTML / Portals	XML / RSS	RDF / RDFS / OWL



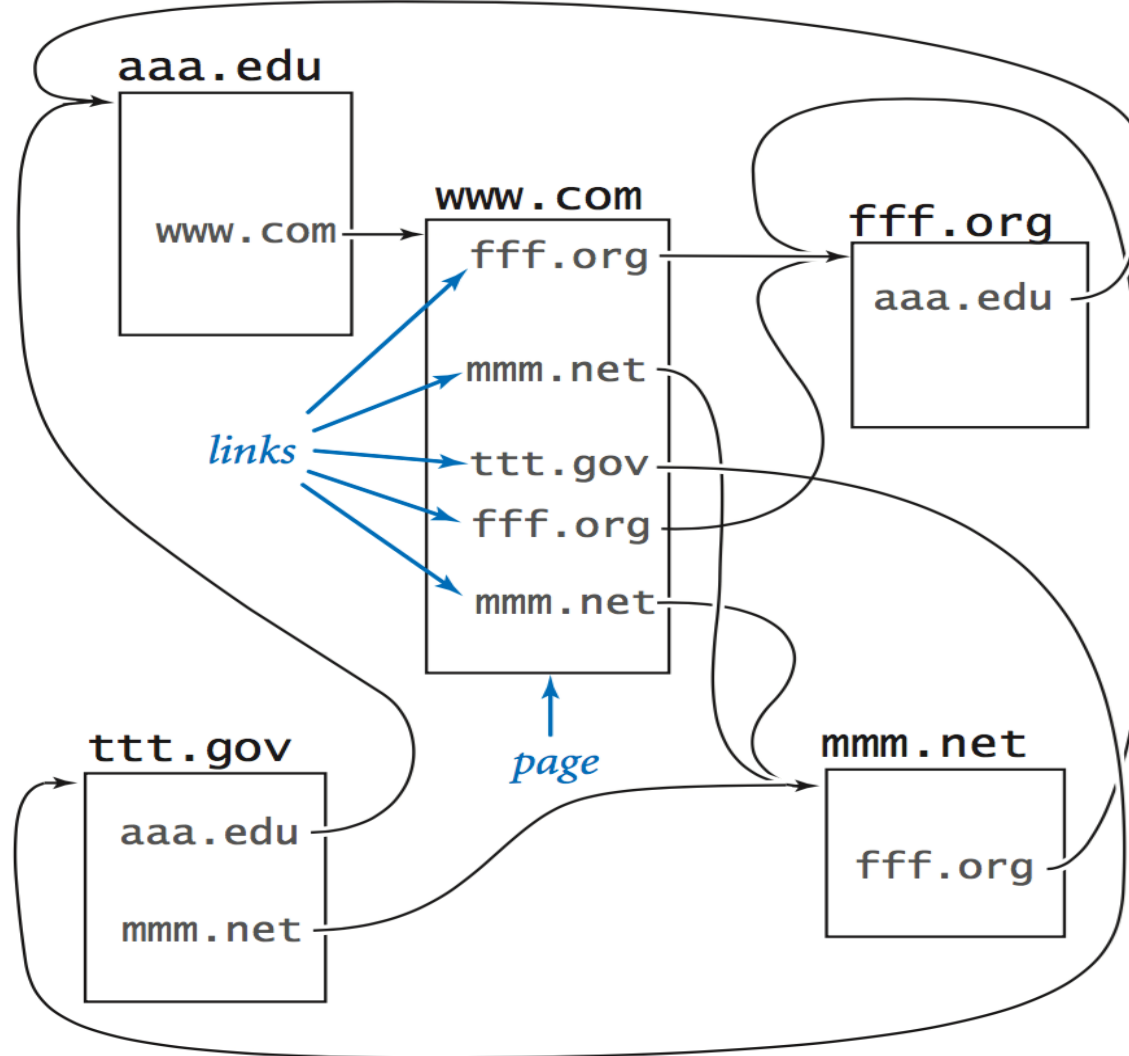
Evolution of information on the Web



- The World Wide Web is called “web” as the information can be seen as a graph of web pages interconnected by hyperlinks
- The hyperlink is “anonymous”, i.e. it does not say what is at the other end. The human, reading the words associated with the hyperlink, can figure out what could be at the other end of the hyperlink (that is very difficult for a program)
- In the Web of Linked (Open) Data the hyperlinks have a “type”, which (in a given vocabulary) provide an indication (to a human or to a program) on what should be expected at the other end of the link



The hyperlinks

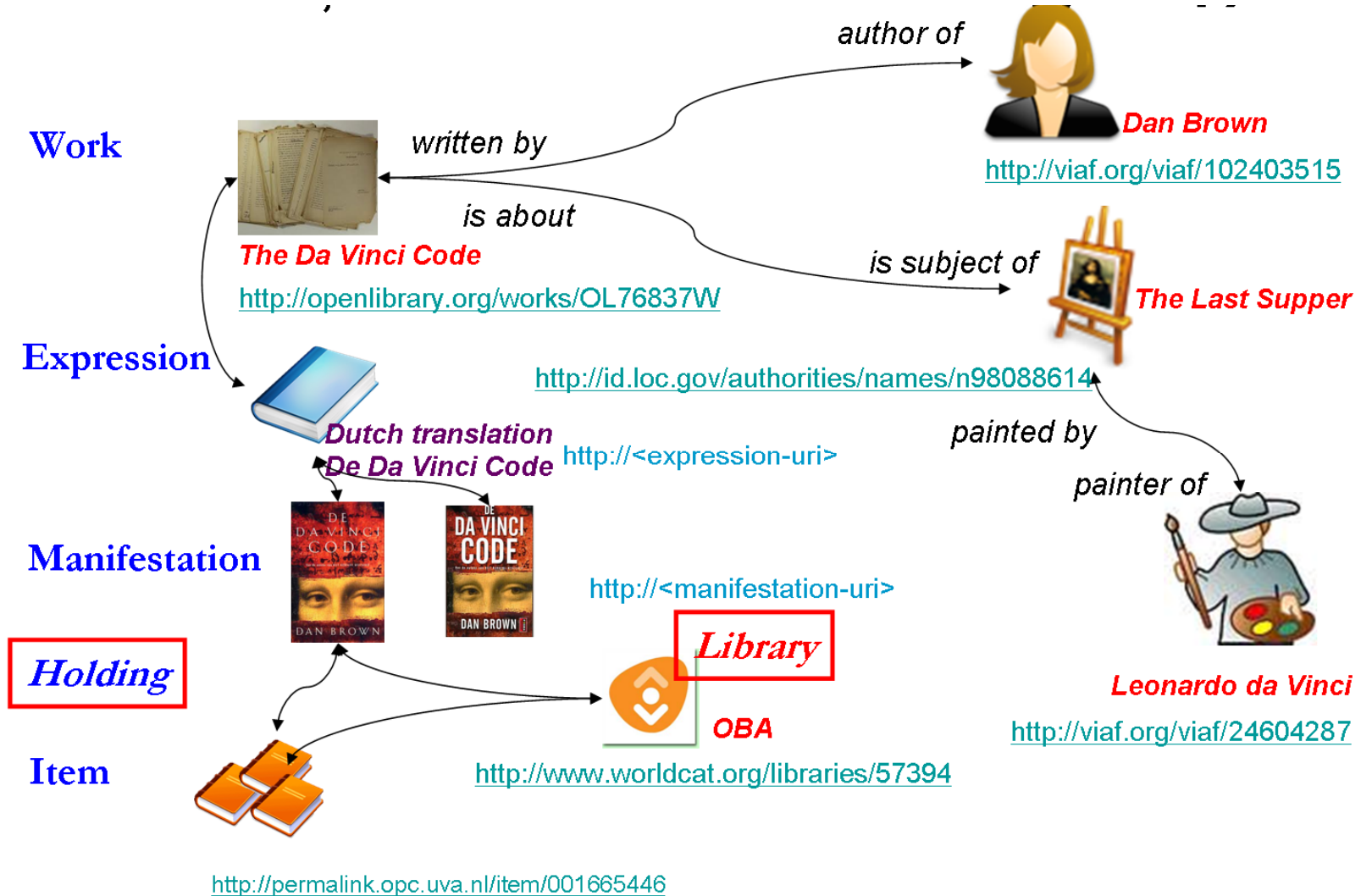




The World Wide Web

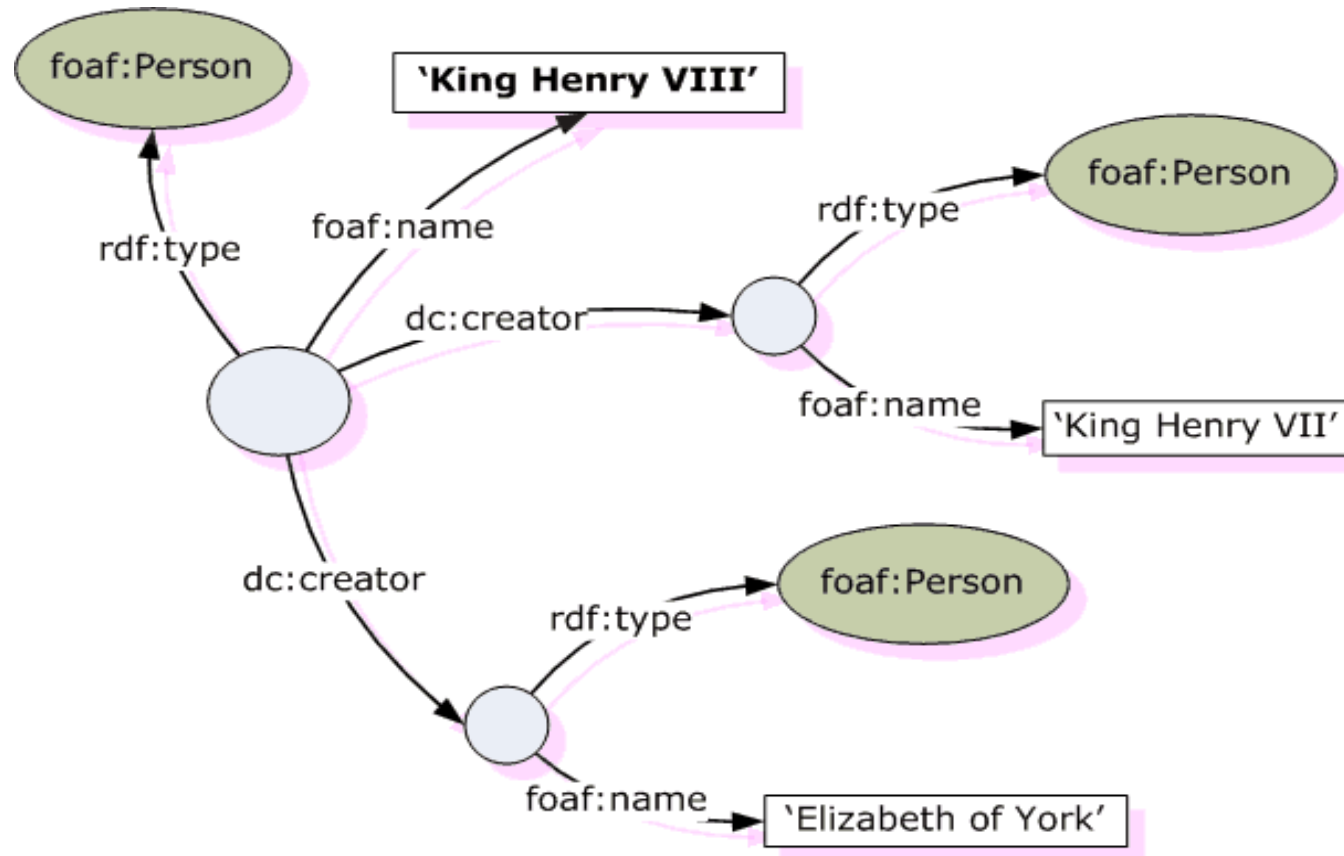


Web of Linked Data





Vocabularies of linked data



```
prefix rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns# >  
prefix dc: <http://purl.org/dc/elements/1.1/>  
prefix foaf: <http://xmlns.com/foaf/0.1/>
```