

WWW Basics

406.424 Internet Applications

Jonghun Park

jonghun@snu.ac.kr

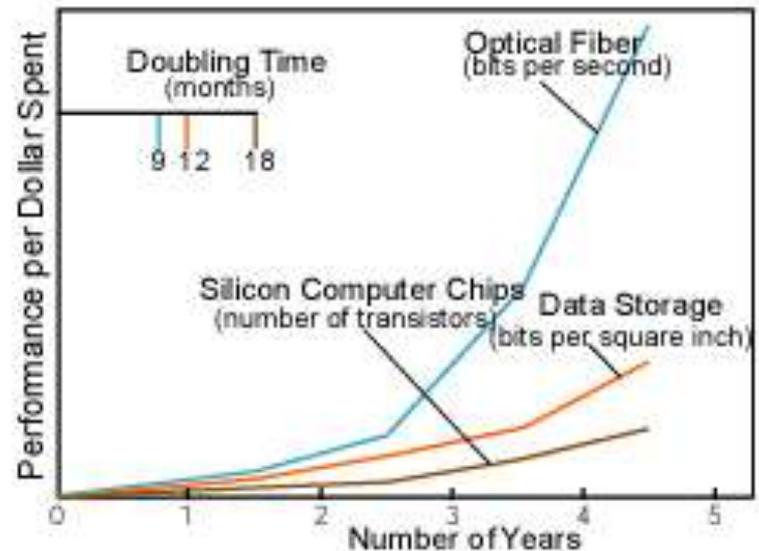
**Dept. of Industrial Eng.
Seoul National University**

9/1/2010



Network Exponentials

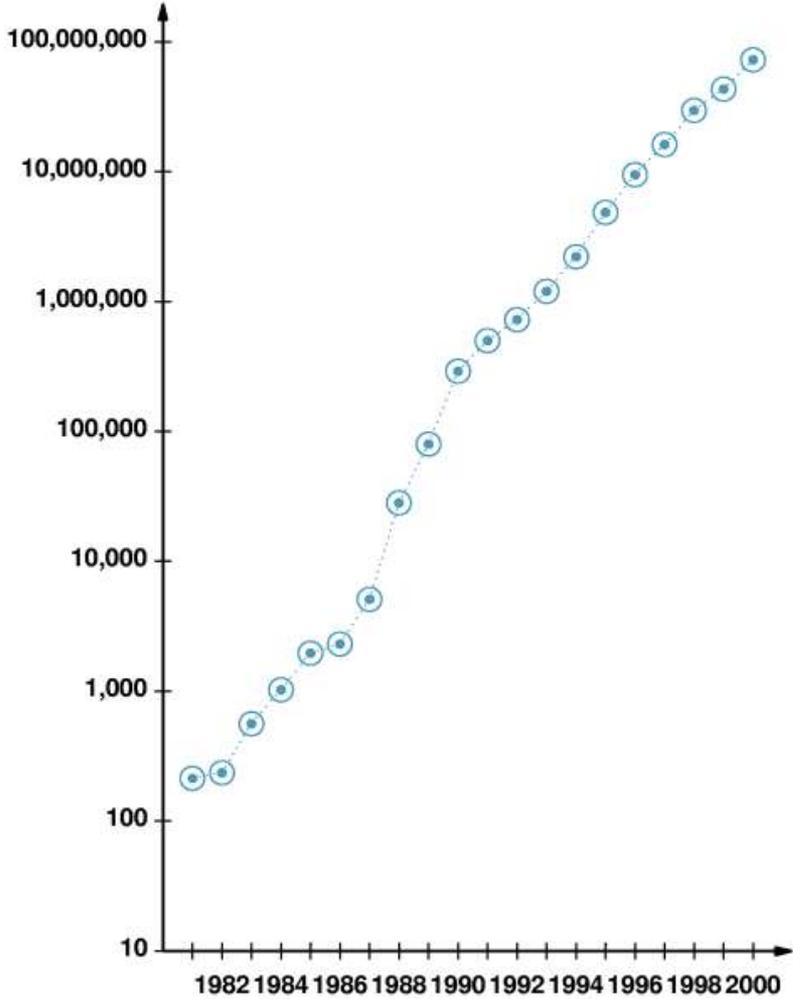
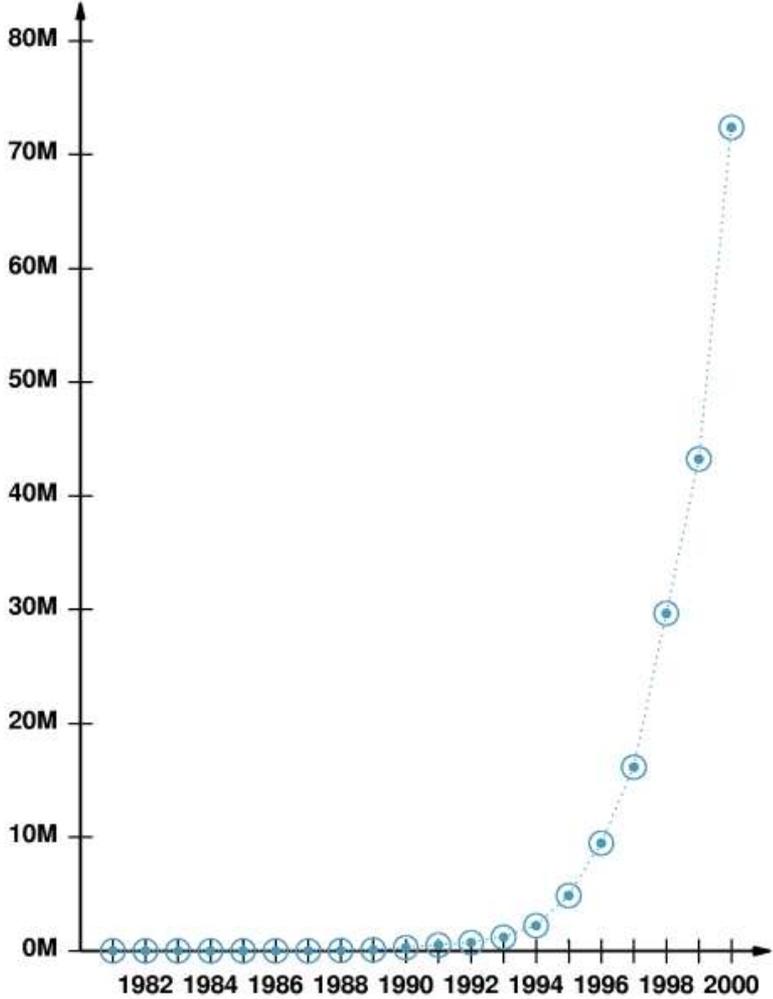
- network vs. computer performance
 - Computer speed doubles every 18 months
 - Network speed doubles every 9 months
 - Difference = order of magnitude per 5 years
- 1986 to 2000
 - computers: x 500
 - networks: x 340,000
- 2001 to 2010
 - computers: x 60
 - networks: x 4000



Source: Scientific American (Jan-2001)

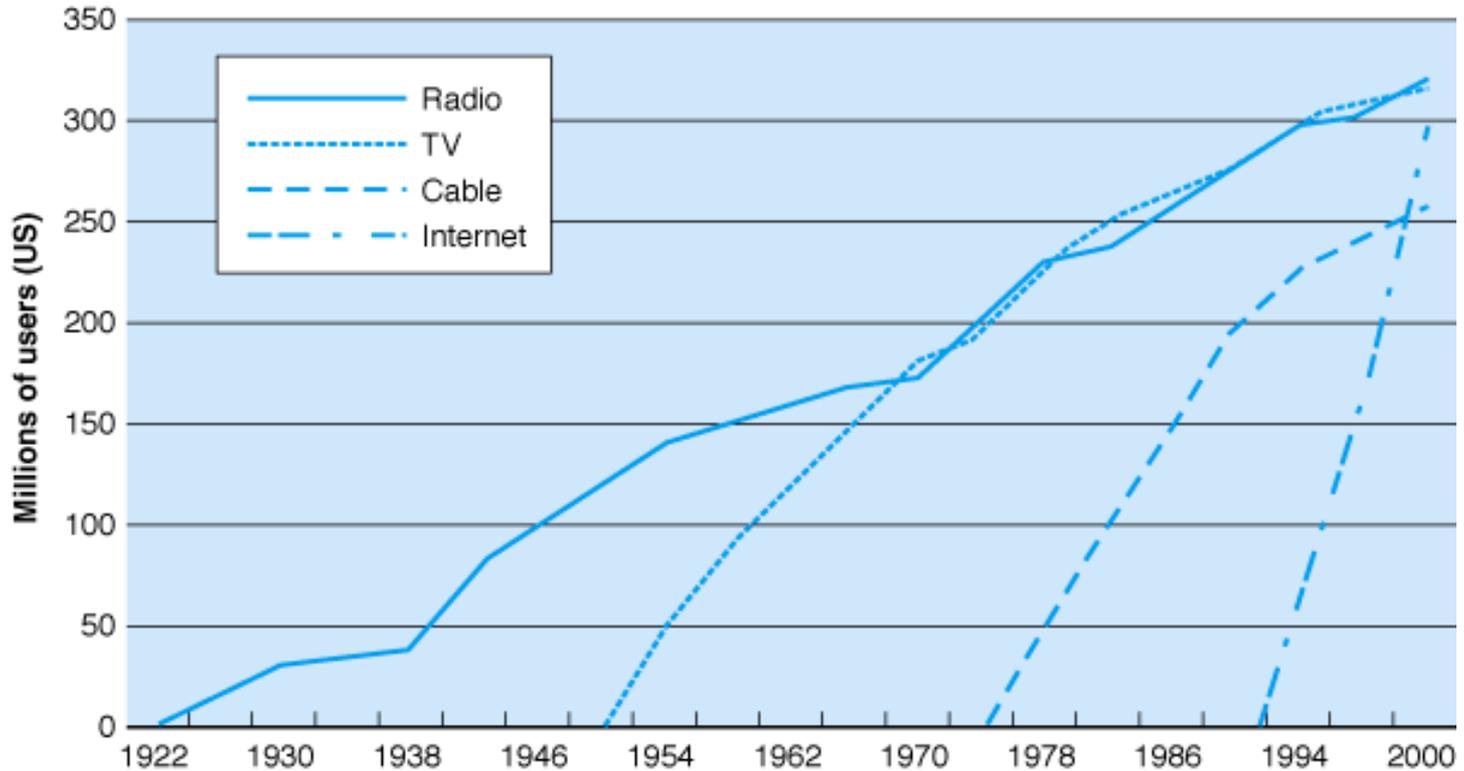


of computers attached to the Internet



internet is the fastest growing medium

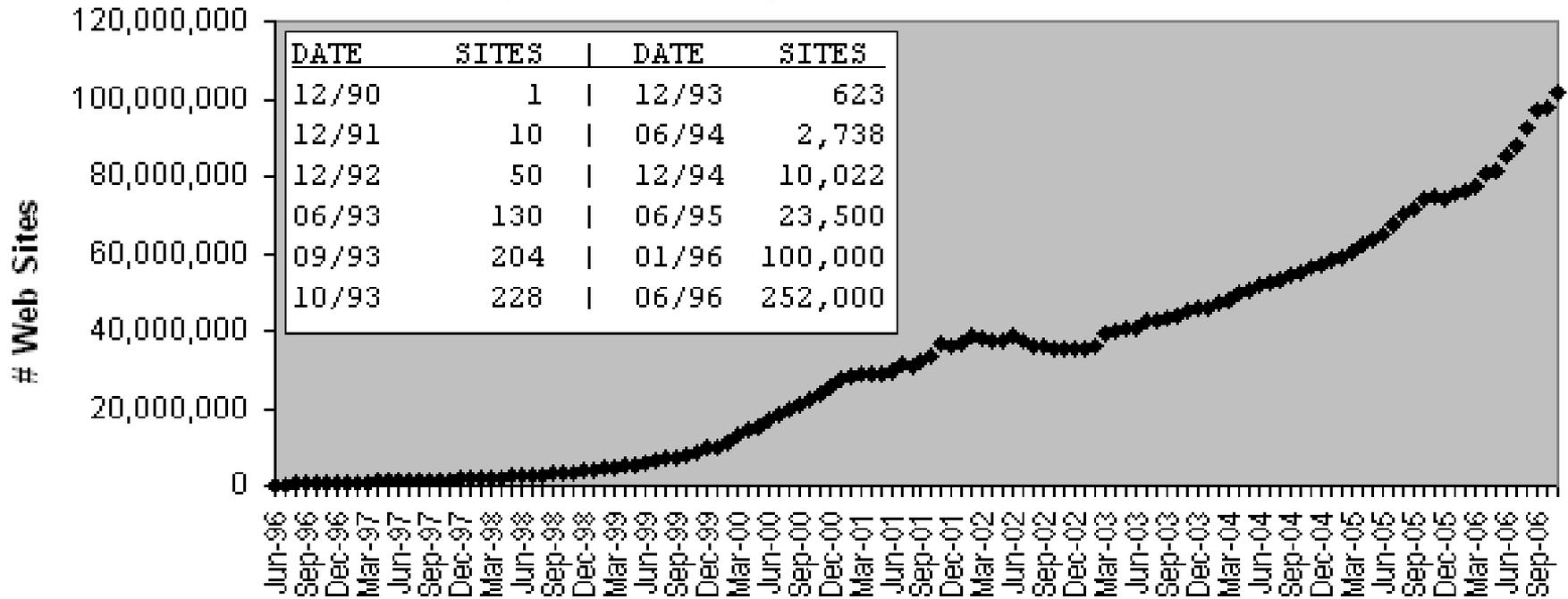
- length of time it took for each ad media to reach 300 million users



internet timeline

- www.zakon.org/robert/internet/timeline/

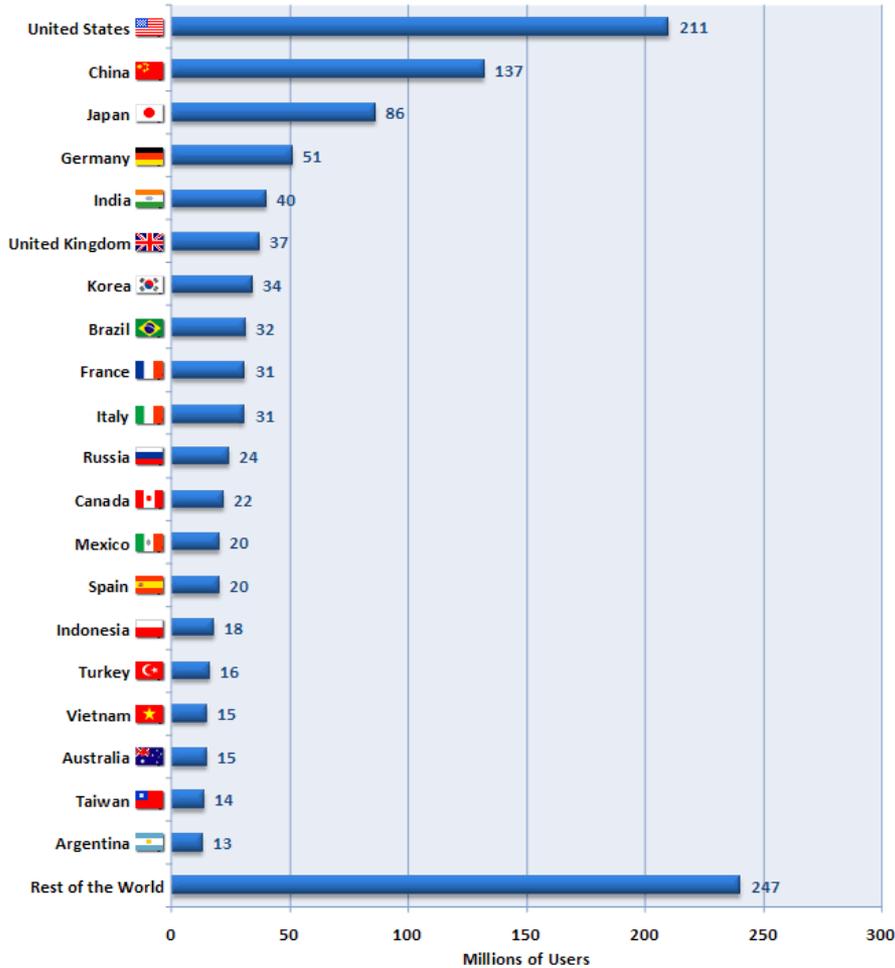
Hobbes' Internet Timeline Copyright ©2006 Robert H Zakon
<http://www.zakon.org/robert/internet/timeline/>



internet users

- www.internetworldstats.com/top20.htm

20 Top Countries in Internet Usage



TOP 20 COUNTRIES WITH HIGHEST NUMBER OF INTERNET USERS						
#	Country or Region	Internet Users, Latest Data	Population (2007 Est.)	Internet Penetration	Source and Date of Latest Data	% Users of World
1	United States	211,108,086	301,967,681	69.9 %	Nielsen//NR Dec/06	18.9 %
2	China	137,000,000	1,317,431,495	10.4 %	CNNIC Dec/06	12.3 %
3	Japan	86,300,000	128,646,345	67.1 %	eTForcasts Dec/05	7.7 %
4	Germany	50,471,212	82,509,367	61.2 %	Nielsen//NR Dec/06	3.6 %
5	India	40,000,000	1,129,667,528	3.5 %	IWS Nov/06	3.6 %
6	United Kingdom	37,600,000	60,363,602	62.3 %	ITU Sept/06	3.4 %
7	Korea (South)	34,120,000	51,300,989	66.5 %	eTForcasts Dec/05	3.1 %
8	Brazil	32,130,000	186,771,161	17.2 %	ITU Mar/07	2.9 %
9	France	30,837,592	61,350,009	50.3 %	Nielsen//NR Jan/07	2.8 %
10	Italy	30,763,848	59,546,696	51.7 %	Nielsen//NR Jan/07	2.8 %
11	Russia	23,700,000	143,406,042	16.5 %	eTForcasts Dec/05	2.1 %
12	Canada	22,000,000	32,440,970	67.8 %	ITU Mar/07	2.1 %
13	Mexico	20,200,000	106,457,446	19.0 %	AMIPCI Oct/06	1.8 %
14	Spain	19,765,032	45,003,663	43.9 %	Nielsen//NR Jan/07	1.8 %
15	Indonesia	18,000,000	224,481,720	8.0 %	eTForcasts Dec/05	1.6 %
16	Turkey	16,000,000	75,863,600	21.1 %	ITU Sept/06	1.4 %
17	Vietnam	14,913,652	85,031,436	17.5 %	VNNIC Jan/07	1.3 %
18	Australia	14,729,191	20,984,595	70.2 %	Nielsen//NR Dec/06	1.3 %
19	Taiwan	14,500,000	23,001,442	63.0 %	FIND Dec/06	1.3 %
20	Argentina	13,000,000	38,237,770	34.0 %	Indec Sept/06	1.2 %
TOP 20 Countries		867,138,708	4,174,463,557	20.8 %	IWS - Mar. 10/07	77.8 %
Rest of the World		247,135,718	2,400,202,860	10.3 %	IWS - Mar. 10/07	22.2 %
Total World - Users		1,114,274,426	6,574,666,417	16.9 %	IWS - Mar. 10/07	100.0 %



of internet users in korea

- isis.nida.or.kr



KR 도메인 등록건수

787,627개

(2007년 03월 26일 현재)



IP주소

IPv4 보유수 : 52,452,608개

IPv6 보유수 : 5,185개

(2007년 2월 현재)



초고속인터넷가입자

14,102,888명

(2007년 1월 현재)



인터넷 이용자수 및 이용률

34,120천명 (74.8%)

(2006년 12월 현재)



오양 비디오.



how is internet used? (korea, multiple responses)



comparative analysis based on sex and age (korea)

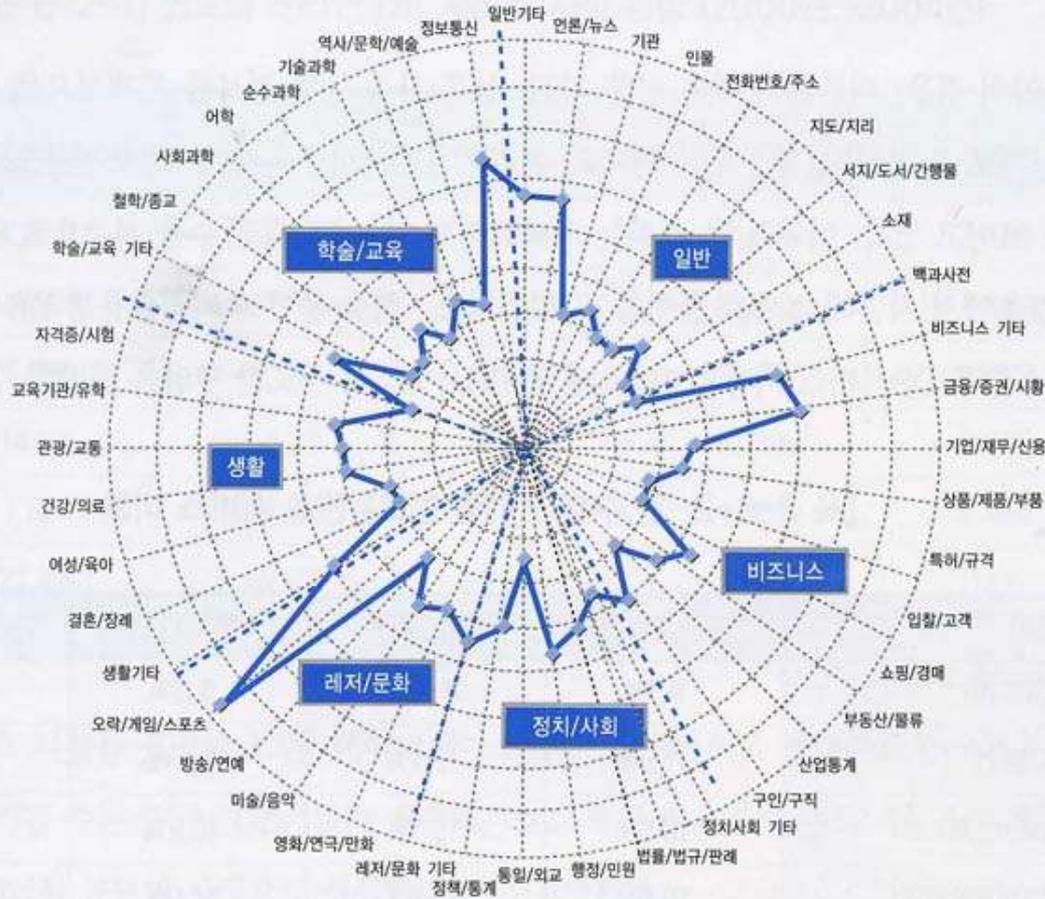
		정보검색	이메일	블로그 미니홈피	동호회 커뮤니티	쇼핑·예약	음악듣기	게임
성	남성	90.1	89.5	86.7	79.1	67.9	58.2	59.8
	여성	88.8	87.5	84.2	76.4	75.1	64.8	50.7
연령	12-19세	85.5	79.9	87.8	75.1	56.8	66.2	67.4
	20대	89.1	90.9	<u>96.1</u>	<u>88.8</u>	73.9	67.1	61.7
	30대	<u>91.4</u>	<u>91.7</u>	85.7	77.3	77.4	57.6	49.5
	40대	90.8	89.1	69.1	66.0	71.7	54.7	45.3

		카페/커뮤니티 이용	퍼나르기 경험	본인 블로그/ 미니홈피 운영	가끔 또는 자주 댓글 작성
성	남성	79.1	64.1	49.9	52.2
	여성	76.4	59.6	<u>55.1</u>	38.4
연령	12-19세	75.1	<u>63.9</u>	<u>54.7</u>	<u>44.3</u>
	20대	<u>88.8</u>	<u>74.4</u>	<u>76.3</u>	<u>49.3</u>
	30대	<u>77.3</u>	58.5	46.3	44.0
	40대	66.0	48.6	27.0	44.2



content distribution (korea)

【그림 2-6-2-2】 온라인 정보콘텐츠 서비스 주제 분야별 분포

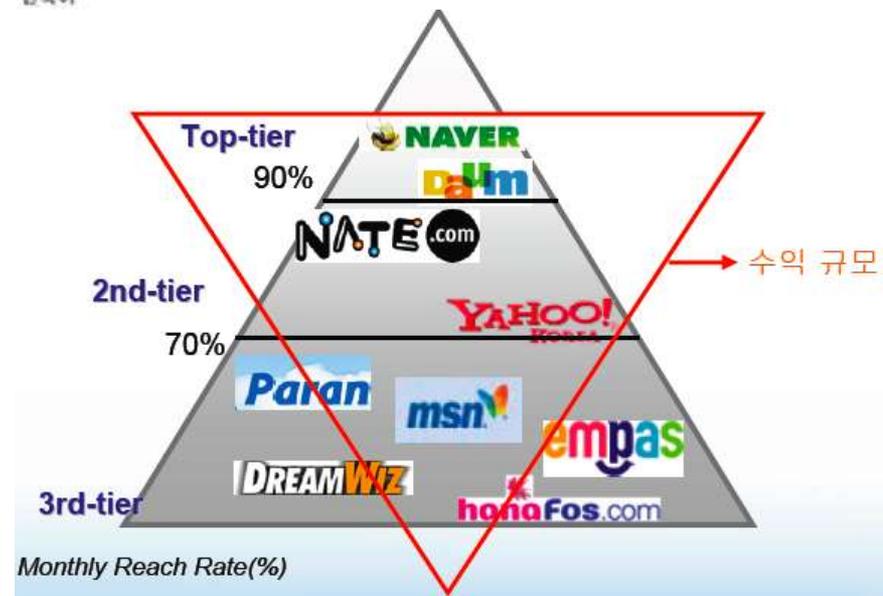


* 출처 : 한국데이터베이스진흥센터, 2004



top sites in korea

주요 포털 사이트 & 구글 한국어 월간 방문자수 변화 추이



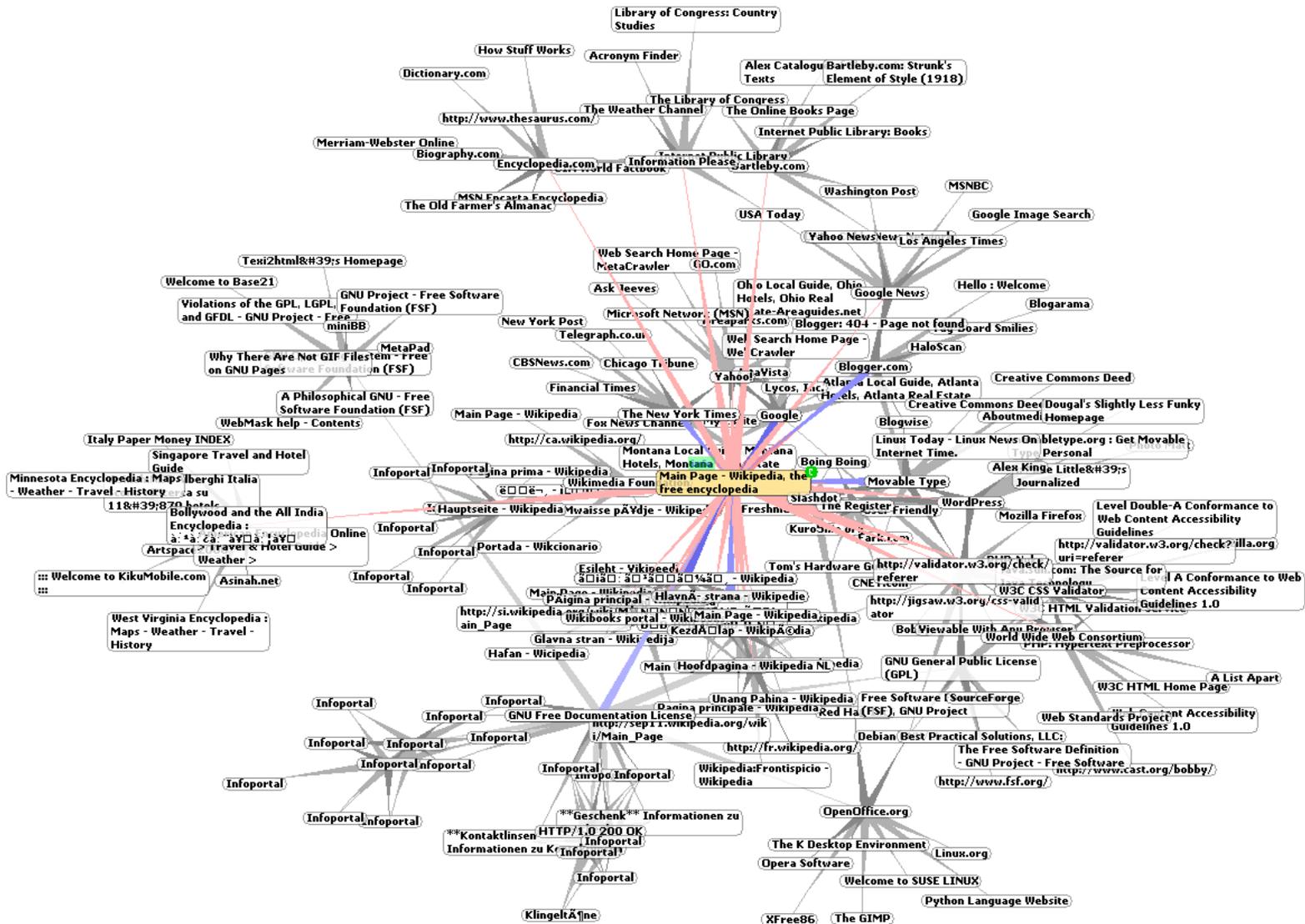
origin and growth of www

- networked application that **links** users and services distributed across computers around the world
- one of the internet services and allows for the **exchanging of documents** (video, text, music, images) over the internet using **HTTP** (hyper text transfer protocol)
 - web pages are written in **HTML** (hyper text markup language)
 - pages on the web are connected by **hyperlinks**
- first proposed by Tim Berners-Lee in 1989
 - linking information present on various machines at CERN, the European Lab. for Particle Physics near Geneva
 - invented HTML, the first server and the first web browser (Lynx)
- technologies predating WWW: FTP, Gopher, Archie, WAIS
- mosaic browser: by M. Andreessen and E. Bina, 1992



graphic representation of the www

(<http://en.wikipedia.org/wiki/Image:WorldWideWebAroundWikipedia.png>)



internet space (opte.org)

	<p>Date: Jan 15 2005 Graph Engine: LGL 1024x1024 (png, 1,329 K) 4000x4000 (png, 1,505 K) 400x400 (png, 229 K)</p> <p>LGL Data: LGL File: Yes LGL 2D Coords: yes LGL Edge Colors: yes</p>	<p>Here is a little updated version of the database at 30%. There is an obvious problem with the color code that must be corrected, however, the image itself is pretty neat. I'll be correcting the color tomorrow for a more updated view without the black on black routes and the missing .mil, .gov.</p> <p>net, ca, us com, org mil, gov, edu jp, cn, tw, au de, uk, it, pl, fr br, kr, nl unknown</p>
	<p>Date: Jan 11 2005 Graph Engine: LGL 1024x1024 (png, 1,309 K) 4000x4000 (png, 1,495 K) 400x400 (png, 225 K)</p> <p>LGL Data: LGL File: Yes LGL 2D Coords: yes LGL Edge Colors: yes</p>	<p>This image is only 10% of the new style database. As you can see, there is a lot more detail in this image with just 10% compared to some of the older images. When the database is populated with all of the routes, we will have the most complex image of the Internet ever seen. I also updated the color system a little, I am not sure if I am happy with it yet. Color mappings are:</p> <p>net, ca, us com, org mil, gov, edu jp, cn, tw, au de, uk, it, pl, fr br, kr, nl unknown</p>
 <p>The Opte Project</p>	<p>Date: Nov 23 2003 Graph Engine: LGL 4000x4000 (png, 11,108 K) 400x400 (png, 288 K) 4096x4096 (png, 2,810 K) 700x700 (png, 867 K)</p> <p>LGL Data: LGL File: Yes LGL 2D Coords: yes LGL Edge Colors: yes</p>	<p>This graph is by far our most complex. It is using over 5 million edges and has an estimated 50 million hop count. We will be producing more maps like this on a dialy basis. We still have yet to fix the color system, but all in due time.</p> <p>Graph Colors:</p> <p>Asia Pacific - Red Europe/Middle East/Central Asia/Africa - Green North America - Blue Latin American and Caribbean - Yellow RFC1918 IP Addresses - Cyan Unknown - White</p>



open directory project (dmoz.org)



open directory project

In partnership with
AOL search

[about dmoz](#) | [suggest URL](#) | [help](#) | [link](#) | [editor login](#)

Search [advanced](#)

Arts

[Movies](#), [Television](#), [Music](#)...

Games

[Video Games](#), [RPGs](#), [Gambling](#)...

Kids and Teens

[Arts](#), [School Time](#), [Teen Life](#)...

Reference

[Maps](#), [Education](#), [Libraries](#)...

Shopping

[Autos](#), [Clothing](#), [Gifts](#)...

World

[Deutsch](#), [Español](#), [Français](#), [Italiano](#), [Japanese](#), [Nederlands](#), [Polska](#), [Dansk](#), [Svenska](#)...

Business

[Jobs](#), [Real Estate](#), [Investing](#)...

Health

[Fitness](#), [Medicine](#), [Alternative](#)...

News

[Media](#), [Newspapers](#), [Weather](#)...

Regional

[US](#), [Canada](#), [UK](#), [Europe](#)...

Society

[People](#), [Religion](#), [Issues](#)...

Computers

[Internet](#), [Software](#), [Hardware](#)...

Home

[Family](#), [Consumers](#), [Cooking](#)...

Recreation

[Travel](#), [Food](#), [Outdoors](#), [Humor](#)...

Science

[Biology](#), [Psychology](#), [Physics](#)...

Sports

[Baseball](#), [Soccer](#), [Basketball](#)...

[Become an Editor](#) Help build the largest human-edited directory of the web



Copyright © 1998-2007 Netscape

4,830,584 sites - 75,151 editors - over 590,000 categories



web archives (www.archive.org)

Technorati

Member Sign-In

HOT NEWS

URL:

Blogdex Top 10

Get Link Cosmos

Find out who's linking to you.

Cosmos allows you to find what's new in the blogging universe, and find URL for a blog, website, or interesting article on the web, and Cosmos

Right now, we're still building this site, so lots of things are still broken that most of the kinks are worked out before we go live, but in the mea

What Cosmos results mean

There are two ways of ranking the inbound links to a blog - either by sorted in order of Authority, blogs that have the most inbound links are relative authority of the site itself - if lots of people who have authority interesting content. Very few authoritative links to a blog may mean that considering interesting to many bloggers. Perhaps it is a diamond in the

The freshness ranking places the most updated blogs first in the search allows you to find newly updated links quickly, which usually mean in

Some examples:

Hit the (Cosmos) link to see the Link Cosmos for each of these recentl

1. [Extra Extra](#) 2 links (Cosmos)
2. [dille schreeuwt het van de daken!](#) 5 links (Cosmos)
3. [stead.ca: A boy named Jay](#) 1 link (Cosmos)
4. [e version 2.0](#) 19 links (Cosmos)
5. [Juan Gato's Bucket o' Rants](#) 53 links (Cosmos)
6. [chronopolis | NEW YORK](#) 77 links (Cosmos)
7. [8Bit Joystick.com](#) 44 links (Cosmos)
8. [Jeffrey A. Cross](#) 3 links (Cosmos)
9. [Outlander](#) 5 links (Cosmos)
10. [Space Waitress Gate A](#) 7 links (Cosmos)
11. [Waslander.nu](#) 1 link (Cosmos)
12. [Sam Ruby](#) 112 links (Cosmos)
13. [The Road to Surfdom](#) 147 links (Cosmos)
14. [Samizdata.net](#) 140 links (Cosmos)
15. [Solonator's Ink Well](#) 1 link (Cosmos)
16. [tweneblog](#) 21 links (Cosmos)
17. [Banners And Us](#) 9 links (Cosmos)
18. [zem](#) 32 links (Cosmos)
19. [Full Bleed: Confessions of a Zine Girl](#) 133 links (Cosmos)
20. [x.log](#) 4 links (Cosmos)

Last Updated: 11:06AM
11/25/2002 PST

1. [Guardian Unlimited Observer | Special reports | Full text: bin Laden's 'letter to America'](#) (Cosmos)

2. [Guardian Unlimited Observer | Special reports | Osama issues new call to arms](#) (Cosmos)

3. [The Register](#) (Cosmos)

4. [HometownAnnapolis.com, Naval Academy - Academy seizes computers from nearby 100 mids](#) (Cosmos)

5. [NY Daily News - Daily Dish - Make theirs a double - Bush twins turn 21](#) (Cosmos)

6. [Patricia Breen - Treat Seeker Ornament](#) (Cosmos)

7. [THREAT MADE IN LOUISIANA SENATE RACE: LANDRIEU WARNS OPPONENT 'THIS IS YOUR LAST CAMPAIGN'](#) (Cosmos)

8. [A Small Victory: feeding the IDF: a call to action](#) (Cosmos)

9. [DDN | Piqua's library has to flesh out its own Web site](#) (Cosmos)

10. [Playtrot Gallery](#) (Cosmos)

WHAT'S NEW

[CONTACT](#) | [SITE MAP](#) | [PRIVACY](#) | [LEGAL](#) | [WEBMASTER](#)

© 2002 Sify Consulting All rights reserved.

Home - Popular - Favorites - Watchlists - WTF

Signed in as dbrute02 [Edit] - Sign out - Help

Technorati™

WTF?
What's hot, and why!

Top Searches

1. Twitter
2. Steven Furtick
3. Youtube
4. Myspace
5. Paris Hilton
6. Iran
7. Dell
8. Sanjaya
9. Swards
10. Catry Selop

More in popular -

Top Tags

birthday bush comedy dance iran islan me money ps3 tags technorunch tecnologia spa war alliant cat children global-warming human-rights leia mp3 scribd twitter web-20 web2.0 wordpress y_combinator youtube 300 2007

More top tags -

Technorati Tools

Claim your blog

Ping us

Technorati Mini

View constantly updated "conversation" in a mini window on your desktop.

Technorati Tools

Plugins, applications and widgets to help you get the most out of Technorati.

Search for

in blog posts Search

Advanced search

Newest posts from your Favorite blogs

So That's How Jesus Did It...

2 days ago in ___/___/___
So That's How Jesus Did It... [whalesharkwann, 458x336 [pg]] It's no longer a secret how Jesus managed to feed the 5,000 with just five loaves of bread and two fish. Those fish must've been a couple of these whoppers. The giant whale shark was caught off the coast of China by hardcore fishermen who managed to harness the ten-metre, - Show details

Man Attacked by Stingray in NZ!

3 days ago in ___/___/___
- Show details

Giant Squid Headed for Oven?

3 days ago in ___/___/___
- Show details

It's Just Like Being Bit by a Dog?

4 days ago in ___/___/___
- Show details

The Shark is Still Working... Now Spread the W...

5 days ago in ___/___/___
- Show details

View your Favorites -

Popular

Music Videos Movies Games DVDs News Buzz

The most popular music bloggers are linking to right now.

 Como Ama Una Mujer Jennifer Lopez
12 new links to this album

 We Were Dead Before the Ship Even Sank Modest Mouse
10 new links to this album

 J.Lo Jennifer Lopez
10 new links to this album

 Rebirth Jennifer Lopez
10 new links to this album

 Elliott Yamin Elliott Yamin
7 new links to this album

 Snakes & Arrows Rush
7 new links to this album

Your Blogs

- [부리부리막사 블로그](#)

No blogs link here yet



Advertise on Technorati

Featured Bloggers

Members with recently claimed blogs



Your face here: Add your blog!



The World Wide Web Consortium (W3C) develops interoperable technologies (specifications, guidelines, software, and tools) to lead the Web to its full potential. W3C is a forum for information, commerce, communication, and collective understanding. On this page, you'll find [W3C news](#), links to [W3C technologies](#) and way to [get involved](#). New visitors can find help in [Finding Your Way at W3C](#). We encourage organizations to learn more [about W3C](#) and [about W3C Membership](#).

<h3>W3C Supporters</h3>	<h3>News</h3>	<h3>Search</h3>
<p>Help W3C by making a donation through the W3C Supporters Program.</p>	<h2>EGovernment and the Web: Workshop Report</h2>	
<h3>Mobile Web Initiative</h3>		<p>Search W3C</p> <input type="text"/> <input type="button" value="Go"/>
<p>The goal of the Mobile Web Initiative is to make Web access from a mobile device as simple, easy, and convenient as Web access from a desktop device. Become an MWI sponsor.</p>	<p>2007-03-20: The European W3C Symposium on eGovernment Report has been published. Participants discussed specific government and citizens' needs related to eGovernment services, to identify aspects that put Web interoperability at risk and find how governments can deliver better and more efficient services through computer technologies. Held 1-2 February (press release), in Gijón, Asturias, Spain, the symposium was organized by the W3C Spanish Office and Fundación CTIC, and supported by the Principality of Asturias Government. (Permalink)</p>	<p>Search W3C Mailing Lists</p>
<h3>Employment</h3>	<h3>► Last Call: EARL 1.0 Schema</h3>	<h3>Testimonials</h3>
<p>Current job opportunities at W3C are HTML and WebAPI specialist, SVG and Compound Documents specialist, and Web Accessibility Engineer. Current W3C Fellows Program openings are Business and Technology Communications Specialist, Web / Graphic Designer, and Software Engineer.</p>	<p>2007-03-23: The WAI ERT Working Group released a Last Call Working Draft of EARL 1.0 Schema. Comments are welcome through 20 April. The Evaluation and Report Language (EARL) is a format to exchange, combine, and analyze results from different evaluation tools and checkers. EARL is introduced in the Evaluation and Report Language (EARL) Overview. Visit the Web Accessibility Initiative (WAI) home page. (Permalink)</p>	<h3>Centre Virtuel de la Connaissance sur l'Europe</h3>
<h3>W3C A to Z</h3>	<h3>► HTTP Vocabulary in RDF:</h3>	
<ul style="list-style-type: none"> • Accessibility • Amaya • Annotea • CC/PP • Compound Document Formats • CSS • CSS Validator 		<p>The CVCE is a public undertaking and developer of the European Navigator knowledge base on the history of Europe. The Centre operates in the field of multimedia, multisource and multilingual digital libraries where it supports and implements several W3C standards. (Member testimonials)</p>
		<h3>Members</h3>
		<ul style="list-style-type: none"> • Member Home Page • Member Submissions • Current Members • Meetings • Fellows (New Openings)



W3C

- develops HTML, CSS, and most Web technology
- founded in 1994, directed by Tim Berners-Lee
- has 380 companies and organizations as members
- located at MIT (US), Inria (France), Keiko (Japan)
- W3C players
 - Members (\$50,000 per year), Team, Advisory board, Technical Architecture Group, Working Groups, ...
- documents
 - Working Drafts, Candidate Recommendations, Proposed Recommendations, Recommendations
 - Working Group Notes, Member Submissions, Staff Comments, Team Submissions
- principles
 - consensus among members
 - limited intellectual property rights
 - free Web access to technical reports (unlike ISO)



IETF Search:



The Internet Engineering Task Force

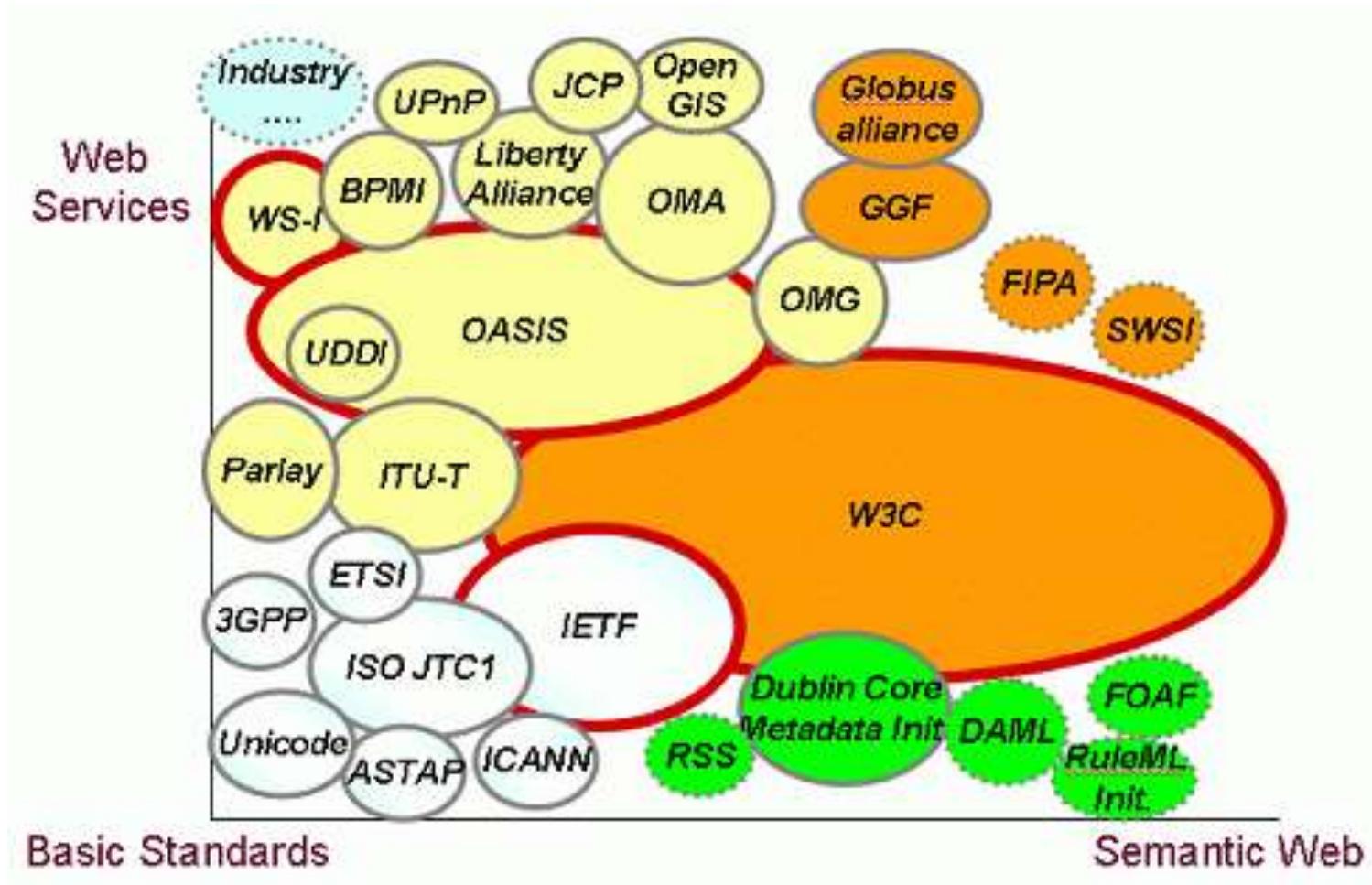
- [Overview of the IETF](#)
- [The Internet Standards Process](#)
- [IETF Working Groups](#)
- [WG Chairs Web Page](#)
- [Internet-Drafts](#)
- [RFC Pages](#)
- [Educational Materials](#)
- [Mailing Lists](#)
- [IESG Activities/Actions](#)
- [Meetings](#)
- ★ [68th IETF - Prague, Czech Republic](#)
(March 18-23, 2007)
- [Proceedings](#)
- [IETF Liaison Activities](#)
- [IETF IPR Disclosure Page](#)
- [The NomCom](#)
- [IETF Secretariat](#)

Related Web Pages: • [IASA](#) • [IAB](#) • [RFC Editor](#) • [IANA](#) • [IRTF](#) • [ISOC](#)

The IETF is an organized activity of the 



other organizations for standardization



web browsers

- an application program that **request** documents from computers connected to the internet (servers) around the world and then **displays** the information in the browser window
- displays the information **according to HTML instructions**
- lynx, is the first web browser
 - a **text based** browser invented by Tim Berners-Lee
- mosaic is the first **graphical** web browser which allowed you to view multimedia files on the Web
 - invented by Marc Anderson, a student at UIUC in 1992
- netscape in 1994 by Marc Anderson
- internet explorer in 1995 by Microsoft after congress passed bill to open web for commerce



the internet movie



components of the web

- uniform resource identifiers (URI)
 - universal **naming mechanism** for identifying resources on the Web
 - consists of protocol, server name, and resource name
- hypertext markup language (HTML)
 - hypertext: **nonlinear** writing or linking related documents for navigation
 - standard representation for hypertext documents in ASCII format
 - derived from SGML (Standard Generalized Markup Language)
- hypertext transfer protocol (HTTP)
 - for communication between web clients and servers
 - defines the **syntax of the messages** and how the fields in each line of the message should be interpreted
 - request-response protocol
 - stateless protocol



URL (uniform resource locator)

- the most popular form of a URI that specifies where a resource is located
- provides the **basis for “hyperlink”**
- transparent w.r.t. resource types
- transient vs. permanent URL
- examples
 - <http://imlab.snu.ac.kr/>
 - <http://imlab.snu.ac.kr:8080/>
 - <http://143.248.80.67/>

Name	Used for	Example
http	HTTP	http ://www.cs.vu.nl:80/globe
ftp	FTP	ftp ://ftp.cs.vu.nl/pub/minix/README
file	Local file	file :/edu/book/work/chp/11/11
data	Inline data	data :text/plain;charset=iso-8859-7,%e1%e2%e3
telnet	Remote login	telnet ://flits.cs.vu.nl
tel	Telephone	tel :+31201234567
modem	Modem	modem :+31201234567;type=v32



encoding

- facilitates the **storage** of text in computers and the **transmission** of text through networks
- EUC-KR
 - most korean web portals had used
 - primarily for CJK
 - incompatible with many other foreign languages & browsers
- UTF-8
 - a variable-length character encoding for Unicode
 - able to represent any universal character
 - becoming the preferred encoding for e-mails and web pages
 - necessary for globalization

Delta's Online Corner » SLeek [save this](#)
by knibben to secondlife inventory lists ... [saved by 8 other people](#)

Pegasus News | News and Events for Dallas - Fort Worth - Denton | pegasusnews.com [save this](#)
by cskcz to dallas community entertainment events music local ... [saved by 23 other people](#)

TechCrunch Japanese アーカイブ › ミックス / リップ / マッシュー5つのデータ処理サービスを比較する [save this](#)
Yahoo! Pipesなどの各マッシュアップツールの機能比較
by skh to tools pipes ... [saved by 21 other people](#)

生産的になろう (Aaron Swartz のブログから) [save this](#)
by odwt to column ... [saved by 178 other people](#)

PRO MATRE - FEITO PARA VOCÊ!!! [save this](#)
by lafa.winta to Hospital

No Spin Architecture [save this](#)
by lesl ... [saved by 5 other people](#)

Moteur de recherche en français [save this](#)
by genysab to web2.0 moteur ... [saved by 4 other people](#)

Блог незамутненного оптимизатора :-): Взломанный Xrumer [save this](#)
by georgi3mi to xrumer

Infogistics Ltd. Power Through Knowledge. [save this](#)
by fduozar to job ... [saved by 3 other people](#)

Hot Chick! [save this](#)
It makes you wanna eat it.
by mikalad5 to Funny images

el Periódico de Guatemala » Actualidad » Hospitales guatemaltecos que buscan atender el turismo de salud [save this](#)
by jooejaking to travel



HTML

- markup language
 - contains guidelines for **display** and allows a browser to choose details
 - 2 browsers may display an HTML document differently
- each HTML document is represented as a text file that contains **tags** along with other information
- tag
 - used to specify an immediate **action** or to start a **formatting** operation
 - appears as a tag name bracketed by less-than and greater-than symbols
- can contain references to other documents through **hyperlinks**
- has been extended to express parts of a document in the form of a **script**
- HTML guides
 - <http://sunsite.berkeley.edu/Web/basictags.html>
 - <http://werbach.com/barebones/>



HTML form

- document that contains items a user must supply
- information is **encoded in a URL** for transmission to another document
- example
 - a form contains 3 items named AA, BB, and CC
 - user has supplied values yes, no, and maybe
 - browser creates a string
?AA=yes&BB=no&CC=maybe
which it **appends to the specified URL**
 - <http://citeseer.com/cs?q=Jonghun+Park>



an example HTML document

```
<HTML>
<BODY>
<H1>Hello World</H1>
<P>
<SCRIPT type = "text/javascript">
    document.writeln("<H1>Hello World</H1>"); // Write a line of text
</SCRIPT>
</P>
</BODY>
</HTML>
```

<!-- Start of HTML document-->
<!-- Start of the main body -->
<!-- Basic text to be displayed-->
<!-- Start of new paragraph -->
<!-- Identify scripting language-->
<!-- End of scripting section -->
<!-- End of paragraph section-->
<!-- End of main body -->
<!-- End of HTML section -->



web document types

- HTML, XML, PS, PDF, Script, etc
- type of document is expressed in the form of MIME type

Type	Subtype	Description
Text	Plain	Unformatted text
	HTML	Text including HTML markup commands
	XML	Text including XML markup commands
Image	GIF	Still image in GIF format
	JPEG	Still image in JPEG format
Audio	Basic	Audio, 8-bit PCM sampled at 8000 Hz
	Tone	A specific audible tone
Video	MPEG	Movie in MPEG format
	Pointer	Representation of a pointer device for presentations
Application	Octet-stream	An uninterpreted byte sequence
	Postscript	A printable document in Postscript
	PDF	A printable document in PDF
Multipart	Mixed	Independent parts in the specified order
	Parallel	Parts must be viewed simultaneously



other markup languages

- WML (WAP Markup Language)
- HDML (Hand-held Device Markup Language)
- VoiceXML (Voice Extensible Markup Language)
- XML (Extensible Markup Language)
 - does not assign any meaning to tags
- and a lot more



HTTP

- **Protocol:** a language with a grammar, syntactic structure such as message formats, and semantic rules indicating how the fields of the messages are interpreted
- **HTTP**
 - Proposed by Berners-Lee in March 1990
 - A request-response protocol
 - Stateless
 - Messages are encoded in ASCII
 - Message syntax is based on MIME
 - based on TCP
 - HTTP/1.0 is defined in RFC 1945
 - Designed for extensibility
 - Current version is HTTP/1.1 (RFC 2616, 2617): <ftp://ftp.rfc-editor.org/in-notes/rfc2616.pdf>

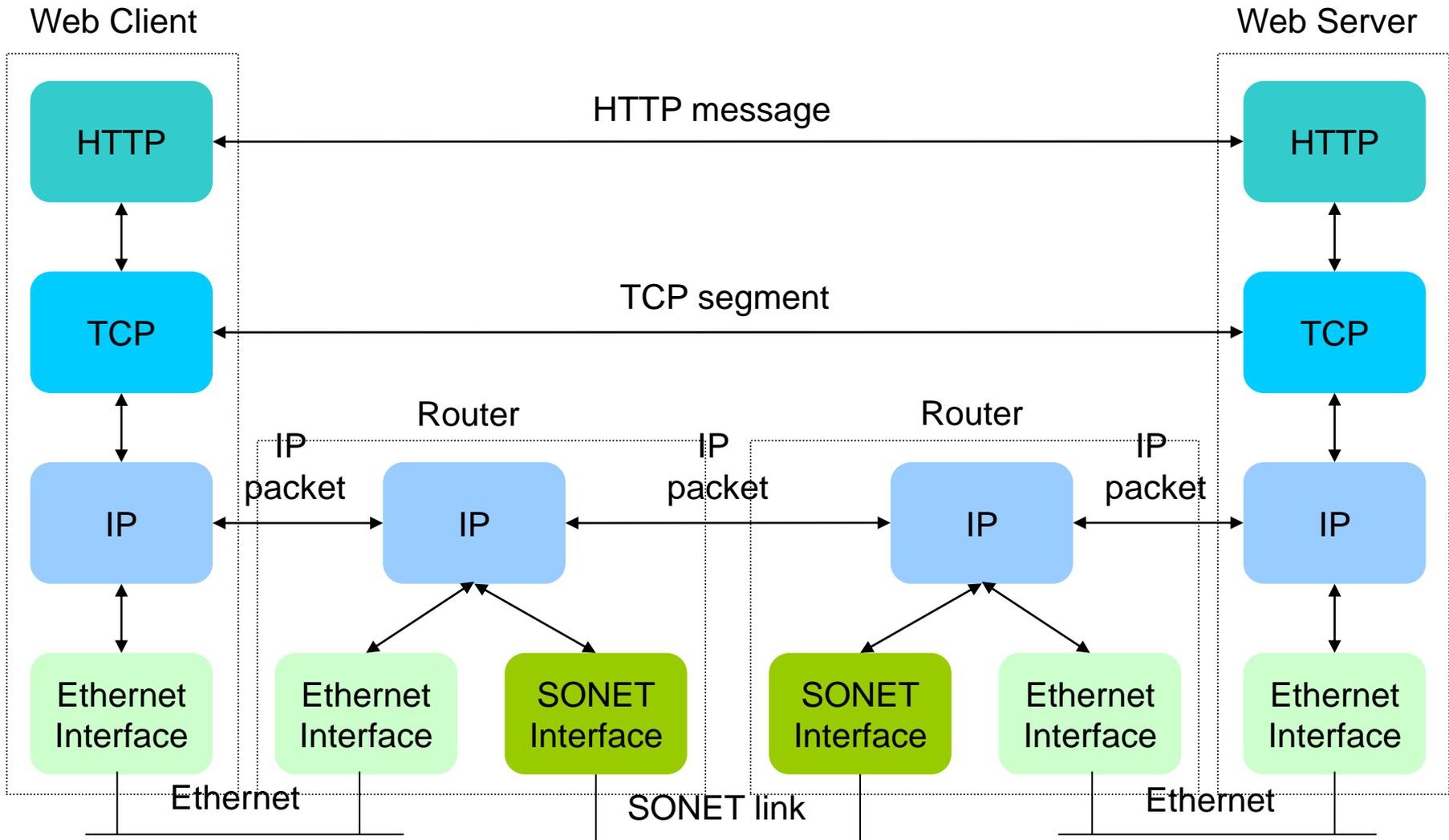


Historical timeline

Date	Document
Mar 1990	CERN labs document proposing Web
Jan 1992	HTTP/0.9 specification
Feb 1992	W3 and WAIS/X.500
Dec 1992	Proposal to add MIME to HTTP
Feb 1993	UDI (Universal Document Identifier) for the Network
Mar 1993	HTTP/1.0 first draft
Jun 1993	HTML (1.0 specification)
Oct 1993	URL specification
Nov 1993	HTTP/1.0 second draft
Mar 1994	URI in WWW
May 1996	HTTP/1.0 Informational, RFC 1945
Jan 1997	HTTP/1.1 Proposed standard, RFC 2068
Jun 1999	HTTP/1.1 Draft standard, RFC 2616
2001	HTTP/1.1 Formal standard



Protocols involved in transferring HTTP messages



key properties of HTTP

- global URI
 - permits resources to reside **anywhere on the Internet** and separates the notion of a **resource** from that of a **response**
 - combination of a **URL** and a **URN**
- request-response exchange
 - an HTTP message is a structured collection of octets in a specific syntax
 - the protocol specifies a set of extensible **request methods** that are used by the client to perform **operations** such as requesting, altering, creating, or deleting a resource
 - a **resource** is an object, service, or collection of entities that can be clearly identified and located anywhere on the network



Key properties of HTTP (cont.)

- **Statelessness**

- The absence of state maintenance across request-response pairs
- A design decision to ensure its **scalability**

- **Resource metadata**

- Metadata is information that relates to a resource but is not part of the resource itself
- e.g., size, type, last modification time

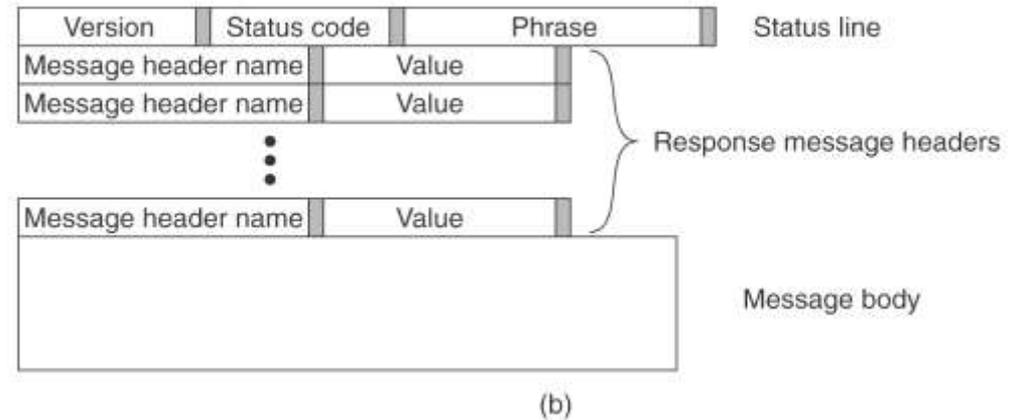
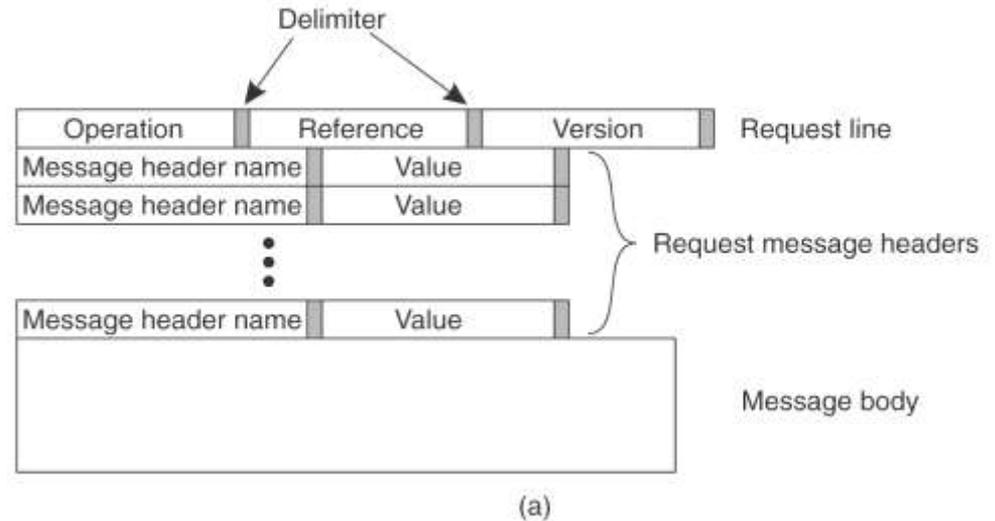
- **Data format**

- The data format is defined as a **MIME type** in HTTP



HTTP messages

- Request message
- Response message
- Request line is mandatory
- Status code is 3-digit



Request and response message

- Request message

- Request line	→	GET /motd HTTP/1.0
General header	→	Date: Wed, 22 Mar 2000 08:09:01 GMT
Request header	→	Pragma: No-cache
Entity header	→	From: jonghun@snu.ac.kr
CRLF	→	User-Agent: Mozilla/4.03
Entity body (Optional)		CRLF

- Response message

- Status line	→	HTTP/1.0 200 OK
General header	→	Date: Wed, 22 Mar 2000 08:09:03 GMT
Response header	→	Server: Netscape-Enterprise/3.5.1
Entity header	→	Content-Length: 23
CRLF		CRLF
Entity body (Optional)	→	Welcome to my world



HTTP/1.0 request methods

- Notifies an HTTP server **what action should be performed** on the resource identified by the URI
 - GET
 - HEAD
 - POST
 - PUT
 - DELETE
 - LINK
 - UNLINK
- Defined in HTTP/1.0
- Not part of HTTP/1.1



HTTP/1.0 response classes

- Every HTTP response message **begins with the Status-Line**
- **Status-Line** has 3 fields: the server's protocol version number, the response code, and a natural language reason phrase
- **Response classes**
 - Informational class: 1xx
 - No actual response codes were allocated until HTTP/1.1
 - Success class: 2xx (e.g., 200 for “OK”)
 - Redirection class: 3xx
 - Client error class: 4xx (e.g., 400 for “bad request”)
 - Server error class: 5xx (e.g., 500 for “internal server error”)



HTTPS

- A protocol that **uses SSL** for transporting the HTTP message
- Everything in the communication between the client and the server is **encrypted**
- HTTPS runs typically on TCP port **443** (not 80)
- URLs start with **https:**
- Typically browsers display an icon (closed lock) if the connection is using SSL
- Several choices are available for a browser
 - SSL 2.0 or SSL 3.0
 - Data Encryption Standard (DES) of NIST or RC2 and RC4 of RSA company



refs for web site design & implementation

- open source web design
 - <http://www.oswd.org/>
- HTML goodies
 - <http://www.htmlgoodies.com/>
- Matt's script archive
 - <http://www.scriptarchive.com/>
- experts exchange
 - <http://www.experts-exchange.com/>
- web developer's virtual library
 - <http://www.wdvl.com/>
- webmonkey
 - <http://webmonkey.wired.com/webmonkey/>

