

Lecture

CS571 - Course Introduction

CS571: Web Technologies

- Instructor: Prof. Marco Papa
- office: PHE 514
- email: papa@usc.edu
- office hours: Thursday 4:30-5:15 PM (by appointment only)
- quick way to ask a “personal” question: Private message to Instructors on Piazza

General Rule

NO D-CLEARANCE

Unless you are a “superstar” undergrad ☺

Course Objectives

- This course focuses on the phenomenon known as the World Wide Web
- Core technologies are:
 - HyperText Markup Language (HTML) and Cascading Style Sheets (CSS) for laying out (formatting) pages that contain text, graphics, audio and video components
 - HyperText Transfer Protocol (HTTP), the communication standard whereby Web browsers and Web servers communicate
 - Web servers, their configuration and performance properties
 - Extensible Markup Languages (XML), a mechanism for defining new tag sets and interchanging data
 - Server-Side programming using PHP
 - Client-side programming using JavaScript
 - Ajax Development Style
- Other Technologies of Interest
 - Responsive Website Design (Bootstrap, etc.)
 - Web services (REST)
 - Web security
 - Web technologies for mobile phones
 - Cloud computing

General Information

- Lectures:
 - Section 1: Tuesday - Thursday (N/A)
 - Section 2: Tuesday - Thursday 5:30PM - 7:20PM
 - Section 3: Tuesday - Thursday 7:30PM - 9:20PM
- Last 30 minutes of session: Lynda.com video [will be used in exams]
- TA & Producers: office hours on course website
- The course website is located at
<http://cs-server.usc.edu:45678/>
- Assignments - yes
- Two **Exams**, one in week 7 and one towards the end. Exam "rules" on Course Info class website
- Mobile Application final assignment
- Attendance - up to you

Software and Storage

- **Student Disk space on cs-server.usc.edu:**
 - Upgraded from 150MB to 0.5GB
- **CGI software: PHP**
- **Software for projects:**
 - Apache 2: version 2.2.22
 - PHP: version 5.4.5
- **Website in the cloud**
 - Amazon's Elastic Compute Cloud (AWS)
 - Google Cloud Platform App Engine

Reading Materials

- No required textbook
- Class Slides are available online on the class website
Click on the  or  icon
- Videos, Example and Links online on the class website
- Recommended reading
 - the W3C site contains the formal specifications for (almost) all of the technologies; www.w3c.org
- O'Reilly & Associates publishes many relevant books:
 - *HTML & XHTML: The Definitive Guide*, 5th Edition by Chuck Musciano and Bill Kennedy, O'Reilly and Associates;
 - *JavaScript: The Definitive Guide*, 5th Edition by David Flanagan, O'Reilly and Associates;
 - *HTTP: The Definitive Guide*, Gourley, et. al. O'Reilly & Associates
- For PHP:
 - *Learning PHP, MySQL, JavaScript, CSS & HTML5* by Robin Nixon, O'Reilly & Associates
- For Mobile SDKs:
 - Online material at developer.apple.com, developer.android.com, dev.windows.com
 - Several books on iOS, Android and Windows 10 SDKs

Other Issues

- Class Sign up list
 - On the class "Home" page click on the green "**Sign Up**" button at the right of your Section details; fill in the form; remember to record your Class ID: it will be used to look up your scores.
 - Use the Class ID to modify your Sign Up data. Available in the "Grades" page.
- Class news group
 - We use Piazza. Activate your membership by joining at:
piazza.com/usc/fall2016/csci571/home
- Academic Integrity Policy
 - Do NOT submit the same program; you can discuss the project with fellow students, but do not develop code with other students; we have tools to check for similar code; do not change your exams after the fact. We scan all the exams. See "Academic Integrity Policy".
- Downloading course software
 - all software can be downloaded from the class website; and all installation instructions correspond to the listed version of the software.

Take a Look at Piazza

general advice
retrieved from
last semester

The screenshot shows a web browser window with the following details:

- Title Bar:** File Edit View History Bookmarks Tools Help
- Tab:** CSCI 571 (20 unread)
- Address Bar:** https://piazza.com/class/hxnr3xbzvl56m
- Piazza Header:** piazza CSCI 571 Q & A Course Page Manage Class
- User Profile:** Ellis Horowitz
- Navigation:** Unread Updated Unresolved Following
- Search Bar:** Search or add a post...
- Class at a Glance:** Updated 1 minute ago. Reload

 - 20 unread posts** (red exclamation mark icon)
 - no unanswered questions** (green checkmark icon)
 - no unresolved followups** (green checkmark icon)
 - 21 total posts**
 - 172 total contributions**
 - 0 instructors' responses**
 - 0 students' responses**
 - n/a avg. response time**

- Student Enrollment:** 13 enrolled ..out of 350 (estimated)
- Share Your Class:** Professors appreciate Piazza best when they see how it is being used. Allow colleagues to view your class through a demo link - a restricted, read only version of your class where all students' names are anonymized and all student information hidden.
- Demo Link:** https://piazza.com/demo_login?nid=hxnr3xbzvl56m&auth=ff4abc4
- Note:** Opening this link in the same browser will log you out as horowitz@usc.edu
- Homework List:** WEEK 7/13 - 7/19
 - Private Introduce Piazza to your stud... 7/15/14
 - Private Get familiar with Piazza 7/15/14
 - Private Tips & Tricks for a success... 7/15/14
 - Welcome to Piazza! 7/15/14

Web Technology Videos ("Discussions")

(see Additional Resources | Videos)

The screenshot shows a web browser window for 'CSCI 571 - Additional Res' at 'cs-server.usc.edu:45678/resources.html'. The page header includes the USC Viterbi School of Engineering logo and 'CSCI571 - Web Technologies Fall 2015'. Below the header, there are navigation links for Home, Lectures, Homeworks, Grades, Software, More, Resources, Textbooks, and Videos. The 'Videos' link is currently selected. The main content area is titled 'CSCI 571 Lynda Web Videos'. It contains a paragraph about Lynda.com and a list of video categories:

- CSS
 - CSS Web Site Design
 - CSS Fundamentals
 - CSS Page Layout
 - CSS: Core Concepts
 - CSS Styling Forms
 - CSS for Developers
 - CSS Positioning Best Practices
- HTML 5
 - HTML5: Document Editing in Depth
 - HTML5: Messaging and Communications in Depth
 - HTML5: File API in Depth
 - HTML5: Geolocation in Depth
 - HTML5: Managing Browser History
 - HTML5: Drag and Drop in Depth
 - HTML5: Graphics and Animation with Canvas
 - HTML5: Video and Audio in Depth
 - HTML5: Local Storage and Offline Applications
 - HTML5: Structure, Syntax, and Semantics
 - HTML5 First Look
 - HTML5: Web Forms in Depth

The screenshot shows a web browser window for 'lynda.com | IT Services' at 'itserVICES.usc.edu/lynda/'. The page header includes the USC University of Southern California logo and 'USC Information Technology Services'. Below the header, there are navigation links for Services, Information Security, Alerts, and About. A search bar is present. The main content area is titled 'lynda.com' and contains the following information:

- lynda.com is an online training provider offering more than 2,380 video-based courses on a broad range of computing and technology topics.
- Features**: Topics covered include Google Apps, Adobe products, Microsoft Office, web design and development, audio/video production, computer programming, and mobile devices, among others.
- Eligibility**: All USC faculty, staff and students.
- Getting Started**: USC faculty, staff and students can access Lynda.com directly by going to [lynda.com](#) and entering your USC NetID and password.
- Please note that your lynda.com session will time out after 15 minutes of no activity and you will need to log back in.

Characterizing the Web

- How many web sites are there?
- How many web pages are there?
- Invisible Web – what is missing
 - Searchable Databases – Most of the invisible web is made up of the contents of thousands of specialized searchable databases that you can search via the Web.
 - Excluded Pages – There are some types of pages that search engine companies exclude by policy.
- For more information on the invisible web see,

<http://www.robertlackie.com/invisible/index.html>

<http://www.lib.berkeley.edu/TeachingLib/Guides/Internet/InvisibleWeb.html>

Characterizing Web Content

There are very few studies that examine the types of content on the web, however . . .

(From IEEE Spectrum, Jan. 2004, pp. 75) :

- Claim: 30% of the web is porn
- Claim: 30% of the web is duplicate information
- 50,000,000 pages are either new or changed each day
- 65% of the web pages are in English

(From Personal Computer World, Optenet, Sep. 2008) :

- Claim: 35% of the web is porn, 11% is e-commerce
- <http://www.optenet.com/en-us/new.asp?id=162>

(From Forbes, Sept. 2011) :

- Claim: 4% is porn, 13% are porn Web Searches
- <http://www.forbes.com/sites/julieruvolo/2011/09/07/how-much-of-the-internet-is-actually-for-porn/>

(From BBC, July 2013) :

- Claim: is porn 4% or 37%?
- <http://www.bbc.com/news/technology-23030090/>

Sample Web Sites (Modest Size)

- Running a web site can get complicated; here is one example.
- The facts:
 - `www.fogdog.com`, online sale of sporting goods
 - Revenues: \$5 million per year
 - 2.2 million page views per month
 - average of 20,000 unique visitors per day
- The solution (in-house):
 - commodity hardware
 - Linux server running Apache 2.0 web servers
 - Using MySQL data base
 - They moved to Ebay!
 - `http://stores.ebay.com/fogdog/welcome`
 - Citrix Netscaler OS, Apache-Coyote/1.1 web server

Sample Web Sites (Medium size)

- Here is a popular, alternate strategy for maintaining a web site
- The facts:
 - www.autobytel.com, new and used cars
 - Market Cap: \$144M (July 2016)
 - Yearly Revenues: \$143M (July 2016)
 - 500,000 purchased vehicles in 2013
 - Mobile version launched in 2012
 - Stock symbol: ABTL (Nasdaq)
- Original Microsoft solution:
 - Microsoft Windows Server
 - Microsoft IIS 7.5 web server
 - Microsoft SQL server database
 - Akamai CDN
- Today:
 - Linux OS, MI/7.5 web server
 - Akamai International CDN (Netherlands)

Sample Web Sites (large size)

- The facts:
 - www.etrade.com, online investing services and resources
 - Market Cap: \$6.22B (July 2016)
 - Yearly Revenues: \$1.90B (July 2016)
 - 60 million page views per month
 - average of 53,000 unique visitors per day
 - 4.9 million accounts (Jan. 2015)
 - 25,000 new retail accounts opened (Oct 2015)
 - 1,952,000 customer transactions per month
 - Stock Symbol: ETFC (Nasdaq)
- The solution:
 - IBM 90 xSeries running Linux/Citrix Netscaler, Apache and Tomcat web servers
 - Hardware facility for load balancing and redundancy
 - Oracle database system
 - Proprietary programming systems

Web Server Farms

- Until recently all serious web sites were maintained using web server farms;
 - A group of computers acting as servers and housed in a single location;
 - Internet Service Providers (ISP's) provide web hosting services using a web server farm
- Hardware and software is used to load balance requests across the machines
- Other issues addressed by web server farms include:
 - **Redundancy** eliminates single point of failure; backup and failover strategy is required
 - **Security**, secure areas are placed behind firewalls which monitor web traffic, network address translation, port translation, SSL

Popular Web Hosting Services

- ***For individuals and small business:***

- **1&1**

http://www.1and1.com/linux-web-hosting?_lf=Static&linkOrigin=&LinkId=ct.tab.hosting&stage=hosting

- **GoDaddy.com**

<http://www.godaddy.com/products/secure-hosting.aspx?ci=72738>

- **Yahoo**

<http://www.iwebhostingplans.com/yahoo/yahoowebhosting.asp>

- ***For companies willing to pay MUCH higher costs:***

- **Rackspace**

http://www.rackspace.com/index.php?CMP=Google_hosting

- **Network Solutions**

<http://www.networksolutions.com/web-hosting/index.jsp>

- ***Reviews and price comparisons:***

- <http://www.hosting-review.com>

- See next slide

- <http://www.pcmag.com/category2/0,2806,2269,00.asp>

Web Hosting Services

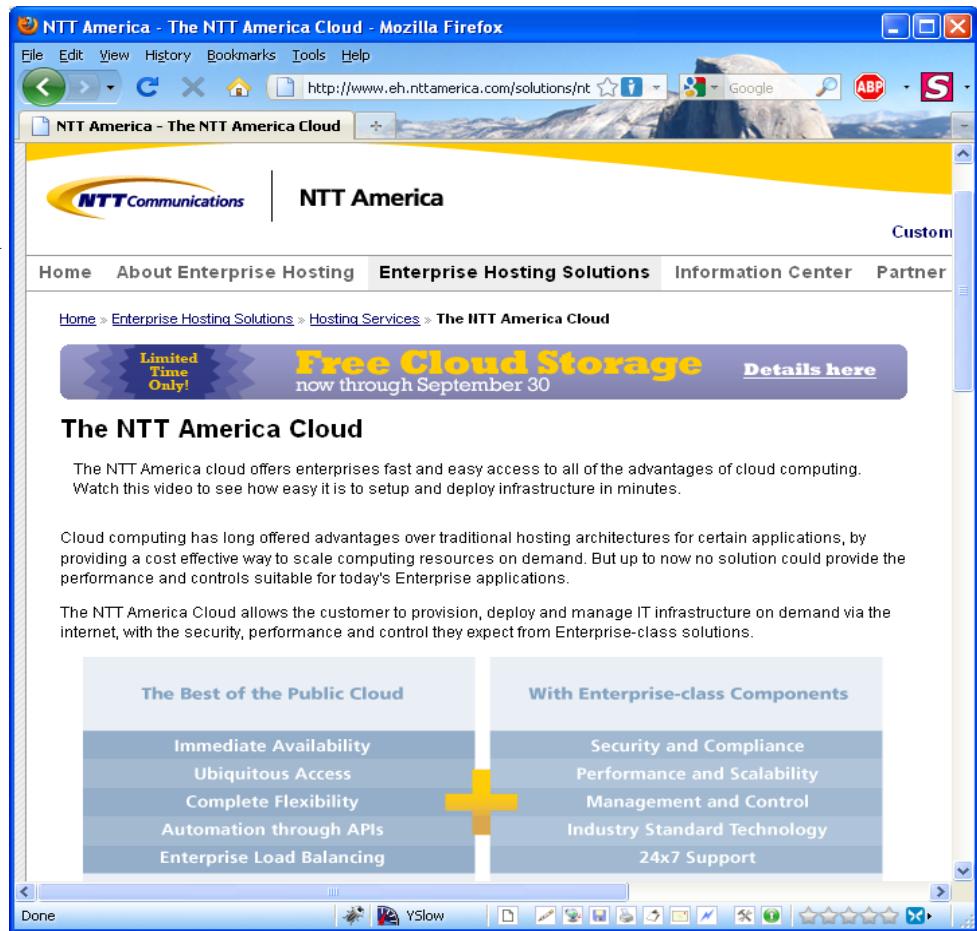
TOP 10 WEB HOSTING PROVIDERS - Updated July 2016										
RANK	SHARED WEB HOST	PRICE	SALES INDEX	TREND	UPTIME SPEED	CUSTOMER REVIEWS	EDITOR'S REVIEW	SCORE INDEX	VISIT	
1	 HostPapa MORE ▾	\$3.95				 113+  20-	 Read	94%	Visit Site	
2	 eHost.com MORE ▾	\$2.75				 40+  0-	 Read	92%	Visit Site	
3	 iPage MORE ▾	\$3.25				 133+  52-	 Read	89%	Visit Site	
4	 HOSTGATOR MORE ▾	\$3.22				 45+  24-	 Read	88%	Visit Site	
5	 iX Webhosting MORE ▾	\$1.95				 73+  39-	 Read	87%	Visit Site	
6	 GoDaddy.com MORE ▾	\$6.29				 60+  30-	 Read	86%	Visit Site	
7	 TMDHosting MORE ▾	\$1.99				 35+  4-	 Read	85%	Visit Site	
8	 bluehost MORE ▾	\$4.95		-		 53+  28-	 Read	85%	Visit Site	
9	 Abaco Smart Business Market locally MORE ▾	\$2.49		-		 40+  21-	 Read	84%	Visit Site	
10	 A2 HOSTING MORE ▾	\$3.92				 58+  4-	 Read	83%	Visit Site	

Cloud Computing

- **Cloud computing** is Internet-based computing, whereby shared resources, software, and information are provided to computers and other devices **on demand**, like the electricity grid.
- Users no longer have need for expertise in, or control over, the technology infrastructure "in the cloud" that supports them.
- It typically includes web-based tools or applications that users can access and use through a web browser as if it were a program installed locally on their own computer.¹
- Typical cloud computing providers deliver common business applications online that are accessed from another Web service or software like a Web browser, while the software and data are stored on servers.
- The major cloud service providers include Microsoft, Salesforce, Skytap, HP, IBM, Amazon, Google and Apple (iCloud).

An Example – NTT America

- For example, NTT America Cloud offers two types of service plans — by the hour with no commitments or minimum fees, or a monthly pre-paid plan. Under both plans the customer is only charged for the CPU and RAM usage when the server is actually running.
- See the link below for details
- Data Center locations are in:
 - San Jose, CA
 - Santa Clara, CA
 - Sterling, VA
 - Ashburn, VA
 - New York, NY
 - Frankfurt, Germany
 - London, England
 - Madrid, Spain
 - Paris, France
 - Singapore



http://www.eh.nttamerica.com/solutions/ntt_america_cloud.cfm

An Example – Amazon’s Elastic Compute Cloud

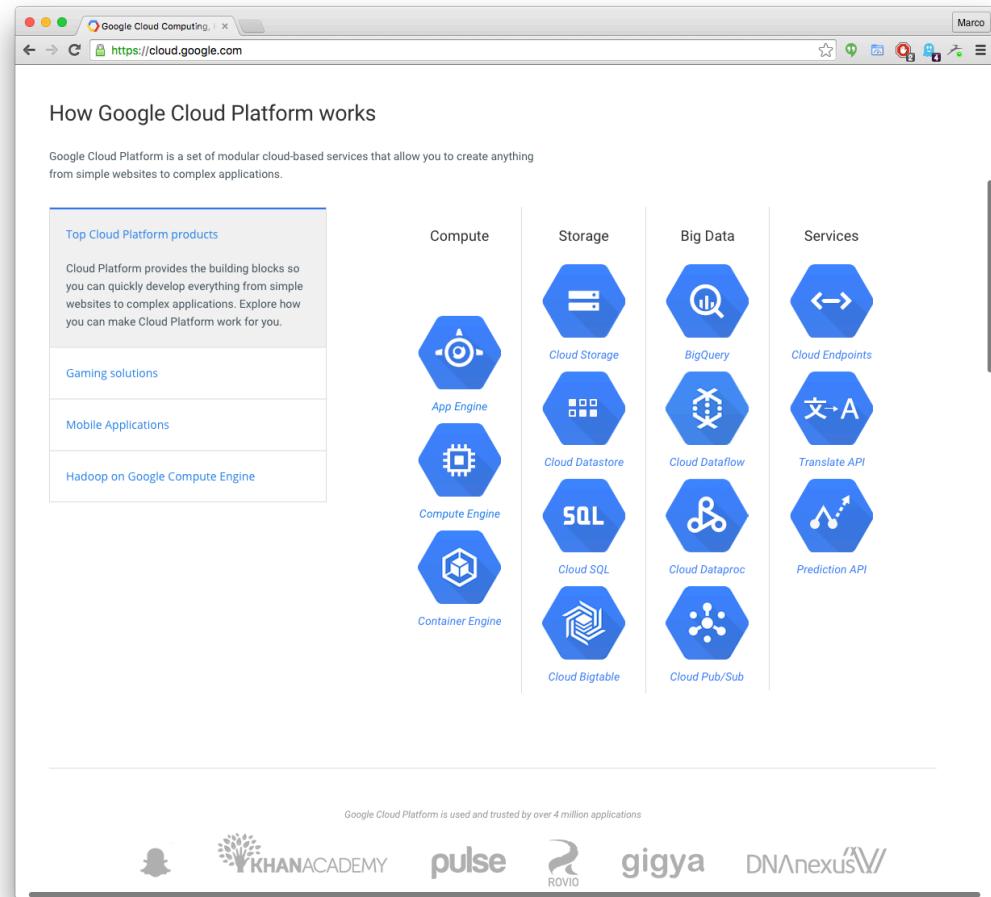
- A web service providing resizable compute capacity
- The “elastic” nature means the service instantly scales to meet demand with no up-front investment
- Users create an Amazon Machine Image (AMI), a virtual computer running your selected operating system (Linux, Windows, etc)
- Users use Amazon’s Simple Storage Service (S3) for large-scale, persistent storage
- You only pay for running AMI
- All accounts are limited to 5 Elastic IPv4 addresses per region
- See: aws.amazon.com/ec2



Amazon currently runs in 8 regions: US East, US West (Oregon), US West (Northern CA), Ireland, Asia Pacific (Singapore), Asia Pacific (Tokyo), Asia Pacific (Sydney), South America (Sao Paulo)

An Example - Google Cloud Platform

- A web service providing basic Compute, Storage, Big Data and Services.
- Additional services for massively scalable Gaming solutions, Mobile Applications backend, and Apache Hadoop.
- App Engine – A platform for building scalable web applications and mobile backends. App Engine scales applications automatically in response to the amount of traffic it receives.
- Compute Engine - Offers predefined virtual machine configurations: Debian, CentOS, CoreOS, SUSE, Ubuntu, Red Hat, FreeBSD, or Windows 2008/2012.



Google uses software-defined networking technology to route packets across the globe and enable fast edge-caching so that data is where it needs to be to serve users.

A Familiar Sample Web Site - USC

The screenshot shows a web browser window with the title bar "Web Server Statistics for [my organisation]" and the URL "www.usc.edu/stats/prev/uscweb-monthly-server-stats.html". The page content includes a header "Monthly Server Report", a note about monthly server statistics, program start time, analysis period, and a "General Summary" section with various statistics.

Note: This Monthly SERVER statistics report contains mostly server statistics for last month (various summaries, reports, and only the most-visited top-level directories). To get a full alphabetical listing of all USCweb directories with more than 10 hits, check out the [Monthly Directory Report](#).

Program started at Sun-03-Jan-2016 04:18.
Analysed requests from Tue-01-Dec-2015 00:00 to Thu-31-Dec-2015 23:59 (31.00 days).

General Summary

(Go To: Top | General Summary | Daily Summary | Hourly Summary | Daily Report | Domain Report | File Type Report | Browser Summary | Browser Report | Status Code Report | File Size Report | Directory Report | User Report | User Failure Report | Referring Site Report | Referrer Report | Failed Referrer Report | Request Report | Failure Report)

This report contains overall statistics.

Figures in parentheses refer to the 7-day period ending 31-Dec-2015 23:59.

Successful requests: 86,565,090 (12,915,339)
Average successful requests per day: 2,792,484 (1,845,048)
Successful requests for pages: 7,195,934 (1,065,768)
Average successful requests for pages per day: 232,132 (152,252)
Failed requests: 1,468,279 (229,324)
Redirected requests: 5,883,350 (1,165,360)
Distinct files requested: 1,416,770 (216,839)
Distinct hosts served: 1,233,102 (274,132)
Corrupt logfile lines: 467,780
Unwanted logfile entries: 2,008,157,064
Data transferred: 2.85 terabytes (527.34 gigabytes)
Average data transferred per day: 94.04 gigabytes (75.33 gigabytes)

86 million page requests and approx 2.7 million requests per day for the 7-day period ending in Dec, 2015

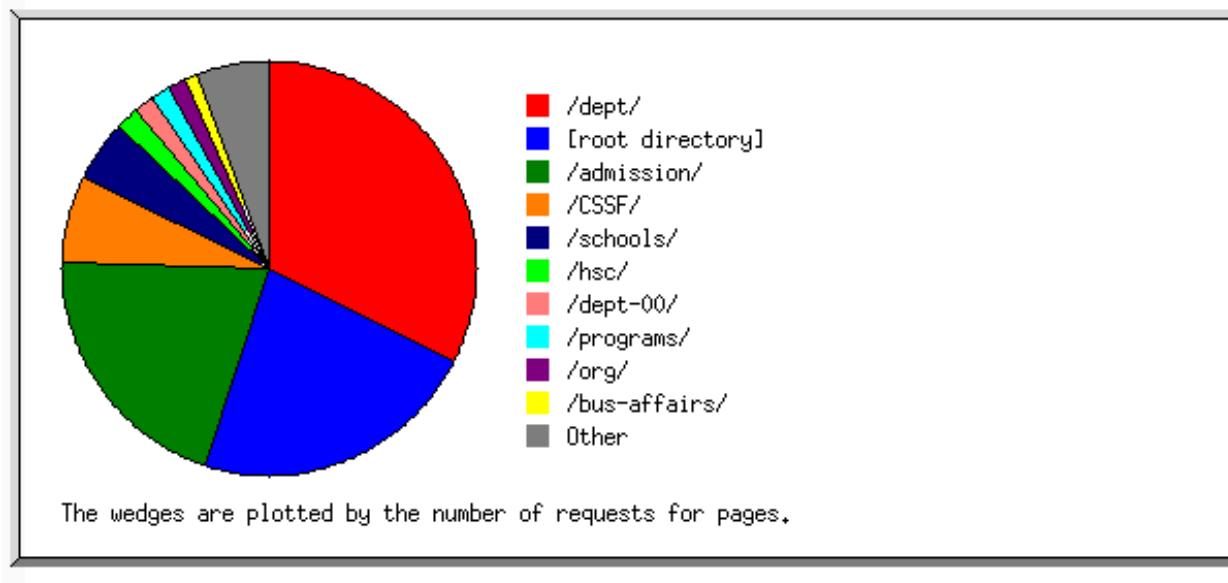
<http://www.usc.edu/stats/prev/uscweb-monthly-server-stats.html>

www.usc.edu server - Directory Report

Directory Report

(Go To: [Top](#) | [General Summary](#) | [Daily Summary](#) | [Hourly Summary](#) | [Daily Report](#) | [Domain Report](#) | [File Type Report](#) | [Browser Summary](#) | [Browser Report](#) | [Status Code Report](#) | [File Size Report](#) | [Directory Report](#) | [User Report](#) | [User Failure Report](#) | [Referring Site Report](#) | [Referrer Report](#) | [Failed Referrer Report](#) | [Request Report](#) | [Failure Report](#))

This report lists the directories from which files were requested. (The figures for each directory include all of its subdirectories.)



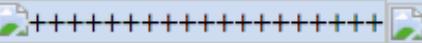
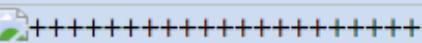
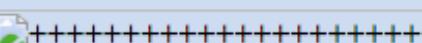
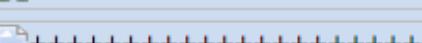
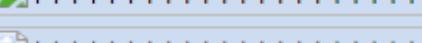
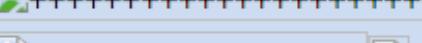
www.usc.edu server – Daily Summary

Daily Summary

([Go To: Top](#) | [General Summary](#) | [Daily Summary](#) | [Hourly Summary](#) | [Daily Report](#) | [Domain Report](#) | [File Type Report](#) | [Browser Summary](#) | [Browser Report](#) | [Status Code Report](#) | [File Size Report](#) | [Directory Report](#) | [User Report](#) | [User Failure Report](#) | [Referring Site Report](#) | [Referrer Report](#) | [Failed Referrer Report](#) | [Request Report](#) | [Failure Report](#))

This report lists the total activity for each day of the week, summed over all the weeks in the report.

Each unit () represents 40,000 requests for pages or part thereof.

day	pages	% pages	
Sun	732006	10.17%	 
Mon	867403	12.05%	 
Tue	1364925	18.97%	 
wed	1345239	18.69%	 
Thu	1213021	16.86%	 
Fri	962051	13.37%	 
Sat	711289	9.88%	 

Busiest day is Thursday

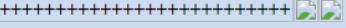
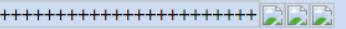
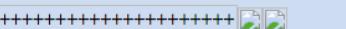
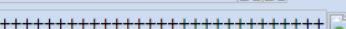
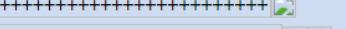
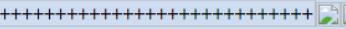
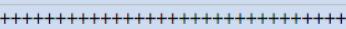
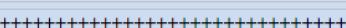
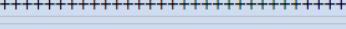
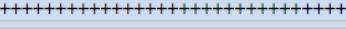
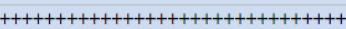
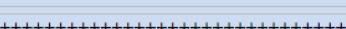
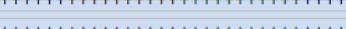
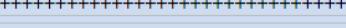
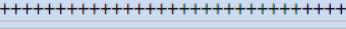
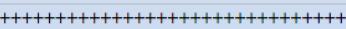
www.usc.edu server - Hourly Summary

Hourly Summary

(Go To: [Top](#) | [General Summary](#) | [Daily Summary](#) | [Hourly Summary](#) | [Daily Report](#) | [Domain Report](#) | [File Type Report](#) | [Browser Summary](#) | [Browser Report](#) | [Status Code Report](#) | [File Size Report](#) | [Directory Report](#) | [User Report](#) | [User Failure Report](#) | [Referring Site Report](#) | [Referrer Report](#) | [Failed Referrer Report](#) | [Request Report](#) | [Failure Report](#))

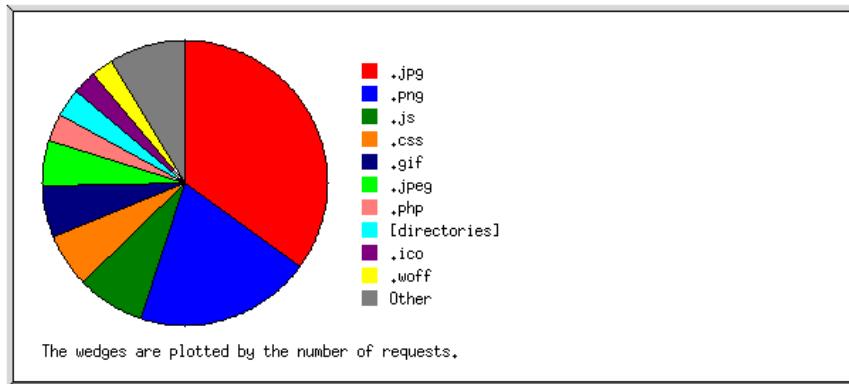
This report lists the total activity for each hour of the day, summed over all the days in the report.

Each unit (+) represents 10,000 requests for pages or part thereof.

hour	pages	% pages	
0	250696	3.48%	
1	224118	3.11%	
2	202032	2.81%	
3	286419	3.98%	
4	325238	4.52%	
5	230540	3.20%	
6	248610	3.45%	
7	272650	3.79%	
8	312770	4.35%	
9	354727	4.93%	
10	371446	5.16%	
11	366253	5.09%	
12	338097	4.70%	
13	340064	4.73%	
14	332571	4.62%	
15	349152	4.85%	
16	314820	4.37%	
17	300479	4.18%	
18	304097	4.23%	
19	303599	4.22%	
20	300673	4.18%	
21	298898	4.15%	
22	287526	4.00%	
23	280459	3.90%	

Heaviest usage
occurs around
10:00am and 3-4:00pm

www.usc.edu server - File Type Report



Listing extensions with at least 100 requests, sorted by the number of requests.

reqs	%reqs	extension
30281917	34.98%	.jpg
17335800	20.03%	.png
6587858	7.61%	.js
5336069	6.16%	.css
5092427	5.88%	.gif
4380596	5.06%	.jpeg
2791088	3.22%	.php
2743043	3.17%	[directories]
2336201	2.70%	.ico
2284802	2.64%	.woff
1813001	2.09%	.html
1457093	1.68%	.pdf
939910	1.09%	.html
864324	1.00%	.xml
838308	0.97%	.svg
386244	0.45%	.15
243588	0.28%	.eot
219621	0.25%	.ttf
165153	0.19%	.htm
144583	0.17%	.xls
52259	0.06%	.txt
42058	0.05%	JPG
38769	0.04%	[no extension]

gif/jpg/png account
For 62% of files
requested

USC Has Many Web Servers Running

The screenshot shows a web browser window for Netcraft. The URL is searchdns.netcraft.com/?host=usc.edu&x=12&y=10. The page title is 'Search Web by Domain'. On the left, there's a sidebar with various Netcraft services like Phishing & Security, Internet Data Mining, and Internet Exploration. The main content area shows a search bar with 'site contains usc.edu' and a 'lookup!' button. Below it, a table lists 83 results for usc.edu, including entries like www.usc.edu, www-scf.usc.edu, and structure.usc.edu.

Site	Site Report	First seen	Netblock	OS
1. www.usc.edu		august 1995	university of southern california	unknown
2. www-scf.usc.edu		august 1995	university of southern california	unknown
3. www-bcf.usc.edu		april 1996	university of southern california	unknown
4. anrg.usc.edu		october 2009	university of southern california	linux - ubuntu
5. you.usc.edu		february 2015	university of southern california	unknown
6. libguides.usc.edu		january 2009	softlayer technologies inc.	linux
7. structure.usc.edu		september 2007	university of southern california	macosx

- Netcraft lists **83** separate web servers with usc.edu in their name, e.g.
 - www.usc.edu
 - mat.usc.edu
 - www.cs.usc.edu
 - dornsife.usc.edu
 - web-applusc.edu
 - www-scf.usc.edu
- However, some may not be connected to USC, e.g.
 - www.usc.edu.au

Web Browsers Use Standard Layout Engines

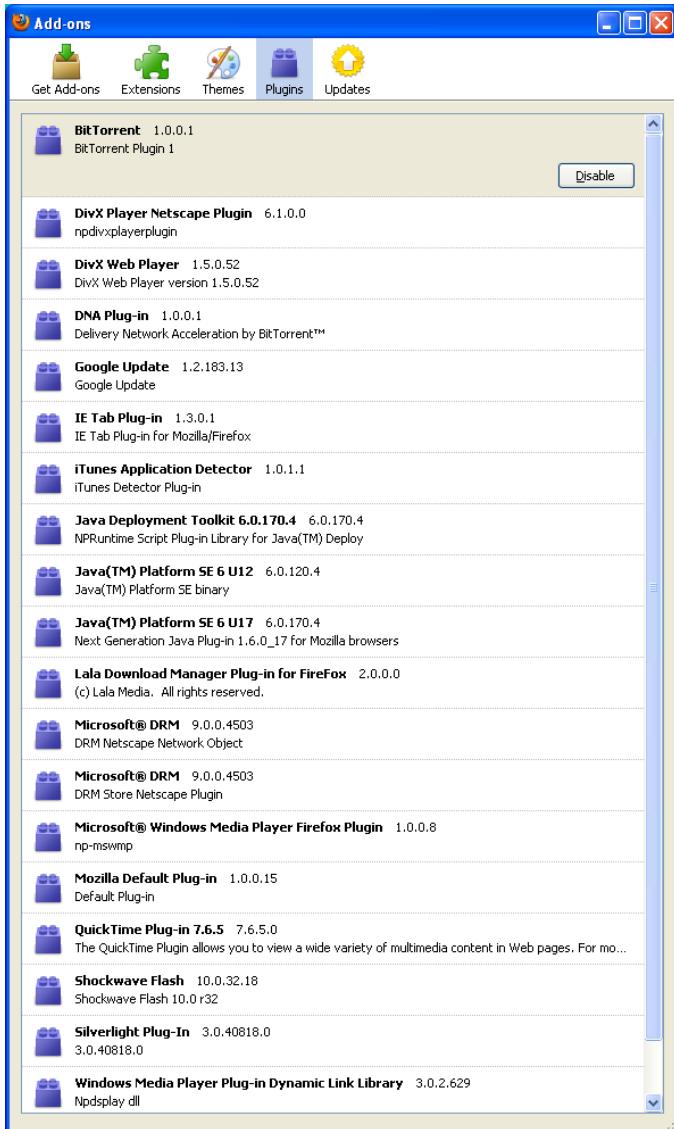
- **WebKit** is a software component used to render web pages; it is open source.
 - It is used by Google's Chrome and Apple's Safari web browsers
 - WebKit is also the name of the Mac OS X system framework version of the engine that's used by Safari, Dashboard, Mail, and many other OS X applications;
- **Gecko** is a layout engine developed by Mozilla Corporation, known as the layout engine of the Firefox web browser.
 - It is used to display web pages and, in some cases, an application's user interface.
 - It offers a rich programming API that makes it suitable for a wide variety of roles in Internet-enabled applications, such as web browsers
 - Its development originated with Netscape Communications Corporation
- Some web kits and the browsers that use them
 - **Gecko-based:** FireFox (Mozilla), Flock, Netscape
 - **Trident-shells:** Internet Explorer (Microsoft)
 - **EdgeHTML:** Edge (Microsoft), fork of Trident 7
 - **WebKit-based:** Chrome and Android (Google), Midori, Safari and Mobile Safari (Apple), Symbian^3 (Nokia) and many others
 - **Presto-based:** Opera, Nintendo DS, Opera Mini, Opera Mobile
 - **Java-based:** HotJava, Lobo

Capabilities of a Browser

- Web browsers fetch and display documents from other WWW sites; their capabilities include:
 - A mouse-driven graphical user interface
 - Display of
 - Hypertext documents conforming to latest HTML standard
 - Text with fonts, styles, and varying point sizes
 - Foreign-language character sets conforming to ISO-8859
 - Forms composed of edit boxes, check boxes, radio boxes, lists, text areas, etc.
 - Graphics in different formats (GIF, JPEG, MPEG, PNG, XBM) including monochrome, color

GIF = graphic interchange format, MPEG = Motion Picture Experts Group, JPEG = Joint Photographic Experts Group, PNG = Portable Network Graphics, XBM = x bitmap

Capabilities of a Browser



- Knowledge of Internet services such as FTP, TELNET
- Ability to invoke helper applications and plug-ins, e.g.
 - *Adobe Acrobat* - used to view pdf files
 - *Windows Media Player* to play digital sound files
 - *Adobe Flash Player*, used to display video e.g. YouTube, etc.)
- Ability to communicate over a secure channel, using SSL
- Ability to maintain and exchange digital certificates
- Ability to run Java applets, JavaScript, and Active X components

The Browser Wars - Desktop Statistics

Browser Statistics and Trends



Statistics are important information.

From the statistics below (collected from W3Schools' log-files since 2002), you can read the long term trends of browser usage.

We see that Google Chrome, Firefox, and Internet Explorer are the most used browsers today.

Browser Statistics

Click on the browser names to see detailed browser information:

2016	Chrome	IE	Firefox	Safari	Opera
May	71.4 %	5.7 %	16.9 %	3.6 %	1.2 %
April	70.4 %	5.8 %	17.5 %	3.7 %	1.3 %
March	69.9 %	6.1 %	17.8 %	3.6 %	1.3 %
February	69.0 %	6.2 %	18.6 %	3.7 %	1.3 %
January	68.4 %	6.2 %	18.8 %	3.7 %	1.4 %

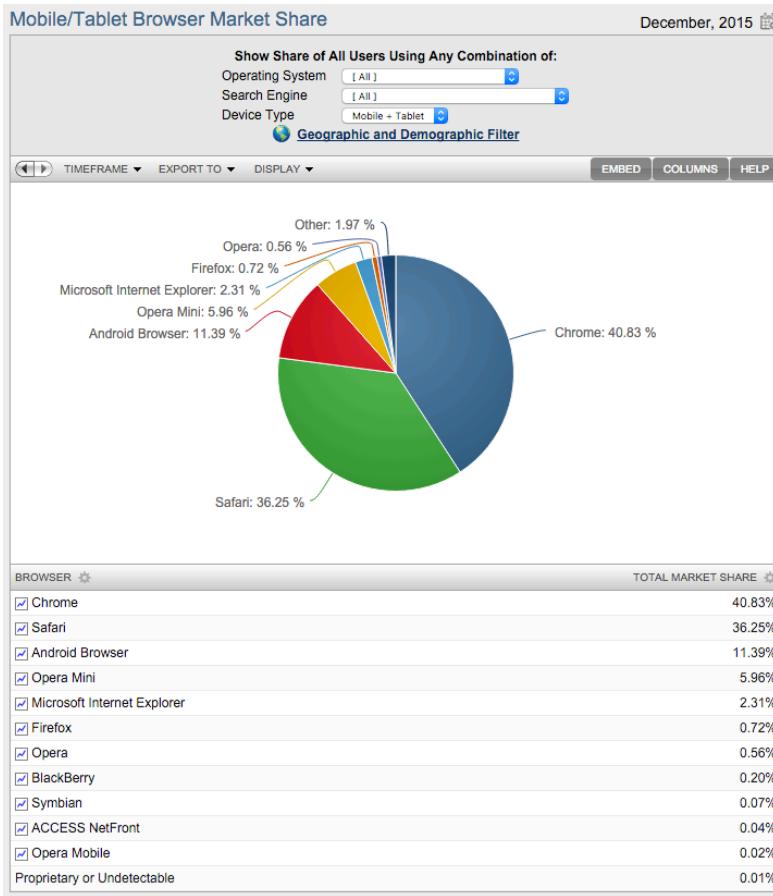
2015	Chrome	IE	Firefox	Safari	Opera
December	68.0 %	6.3 %	19.1 %	3.7 %	1.5 %
November	67.4 %	6.8 %	19.2 %	3.9 %	1.5 %
October	66.5 %	6.9 %	20.0 %	3.8 %	1.4 %
September	65.9 %	7.2 %	20.6 %	3.6 %	1.4 %
August	64.0 %	6.6 %	21.2 %	4.5 %	2.2 %
July	63.3 %	6.5 %	21.6 %	4.9 %	2.5 %
June	64.8 %	7.1 %	21.3 %	3.8 %	1.8 %
May	64.9 %	7.1 %	21.5 %	3.8 %	1.6 %
April	63.9 %	8.0 %	21.6 %	3.8 %	1.5 %
March	63.7 %	7.7 %	22.1 %	3.9 %	1.5 %
February	62.5 %	8.0 %	22.9 %	3.9 %	1.5 %
January	61.9 %	7.8 %	23.4 %	3.8 %	1.6 %

- See http://www.w3schools.com/browsers/browsers_stats.asp
- See also <http://www.upsdell.com/BrowserNews/stat.htm>

Conclusion of the above study:

- Chrome is the clear winner
- Firefox comes second
- Internet Explorer is losing ground
- Safari and Opera having small percentages

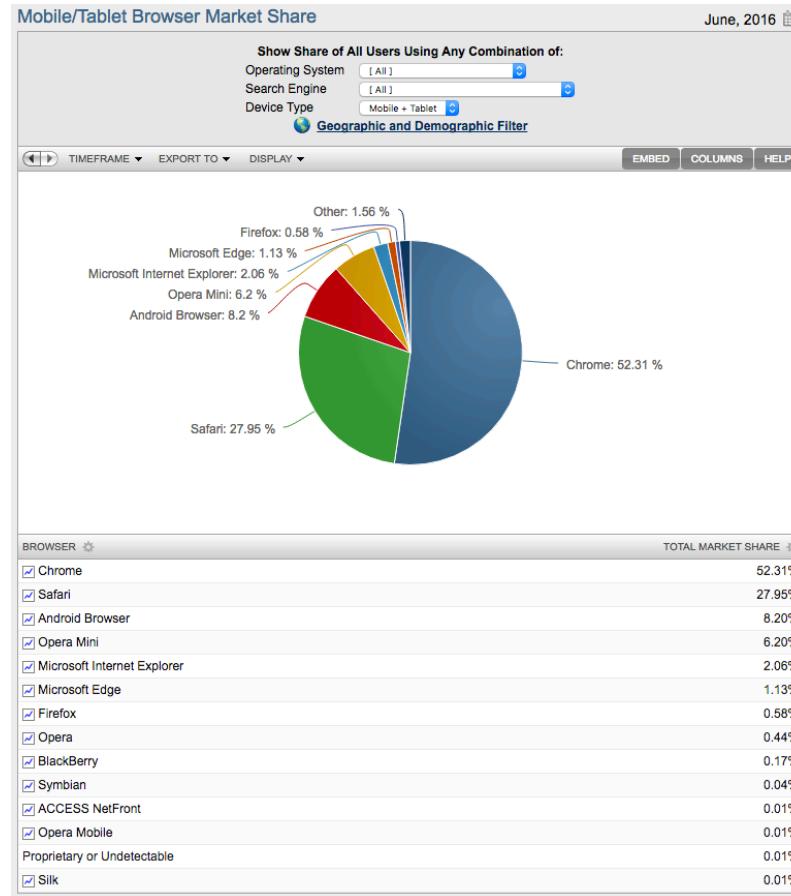
Mobile/Tablet Browser Market Share Statistics



December, 2015

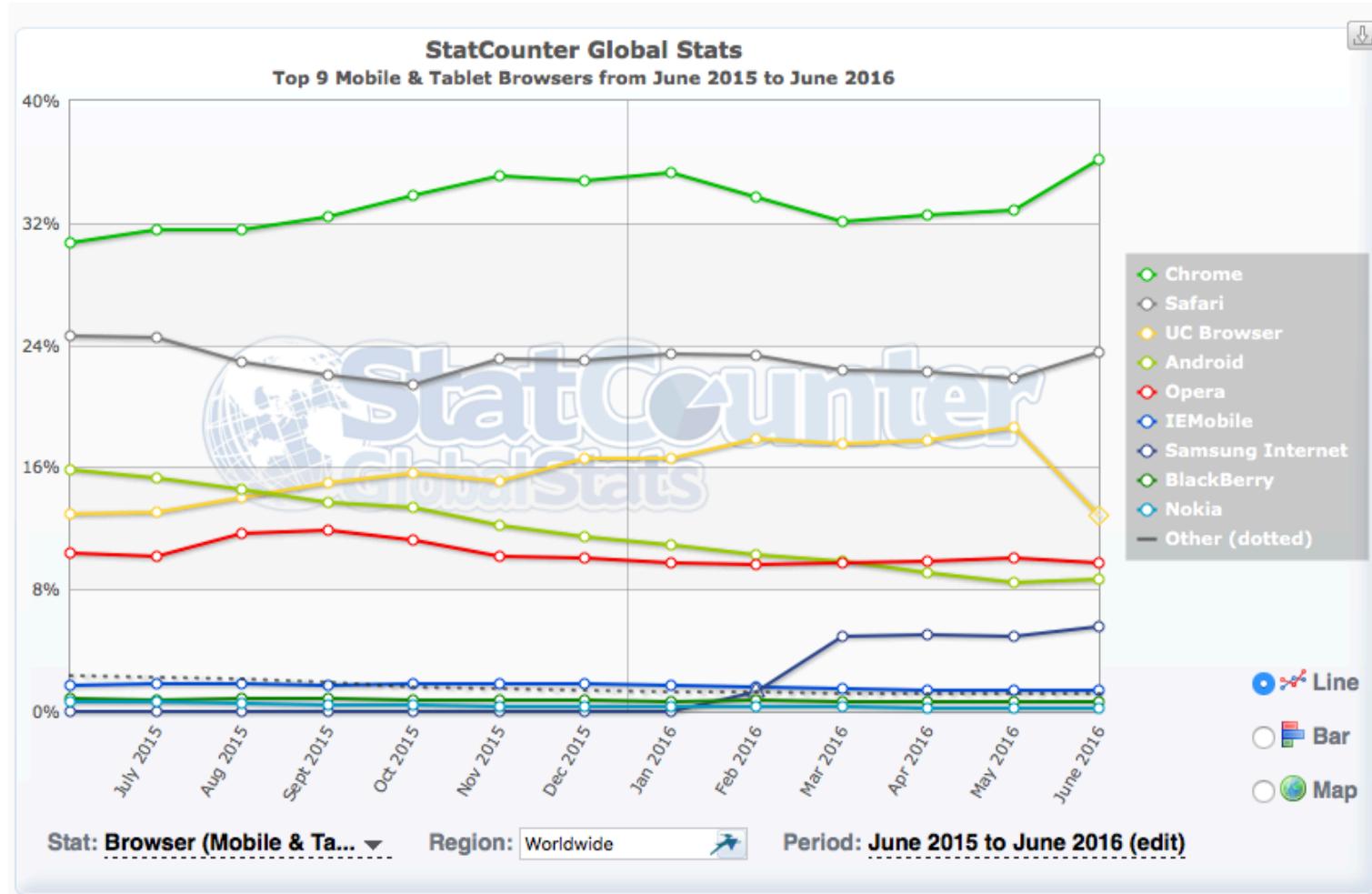
<http://www.netmarketshare.com/>

Chrome leads with about 50% market share over Safari 28%, which surpassed Android browser Webkit (used by iOS, Chrome, Android, Blackberry & others) has over 90% market share



June, 2016

The Browser Wars Comparison (cont'd)



StatCounter Global Stats, June 2015 – June, 2015, See

<http://gs.statcounter.com/>

Chrome has the lead with about 36% followed by Safari at 24%

Browsers are the Gateway to the Web/Internet

Despite Netscape's failure, there is now a new business model for browsers; Google will pay Mozilla \$300 million/year for 3 years to keep Google its default search engine. (Dec. 25, 2011)

Similarly, Google is rumored to have been paying Apple \$1 billion in 2011 to keep Google the default search engine for Safari on iOS devices and OS X.

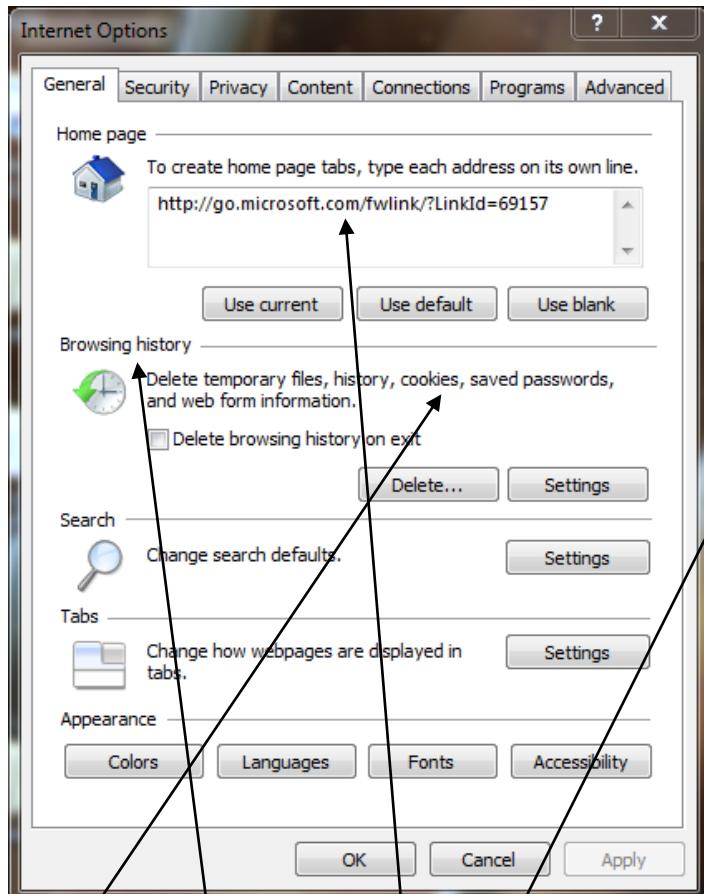
(http://articles.businessinsider.com/2012-03-09/tech/31138467_1_google-maps-ben-schachter-google-searches)

June, 2014: Apple announces DuckDuckgo will be another built-in search engine on Safari (no user tracking)

The screenshot shows a Mozilla Firefox browser window. The title bar reads "Netscape Founder Backs New Browser - NYTimes.com - Mozilla Firefox". The main content area displays an article titled "Netscape Founder Backs New Browser" by MIGUEL HELFT, published on August 13, 2009. The article discusses Marc Andreessen's support for RockMelt, a new browser he is backing. It includes a photo of Andreessen and a smaller photo of him speaking at a podium. On the right side of the browser window, there is a sidebar with social sharing options like "SIGN IN TO RECOMMEND", "E-MAIL", "SEND TO PHONE", "PRINT", "REPRINTS", and "SHARE". Below the sidebar, there is an advertisement for "WHIP IT OCT 9". The bottom of the browser window shows standard navigation and search controls.

Browser Options Menus for IE and Firefox

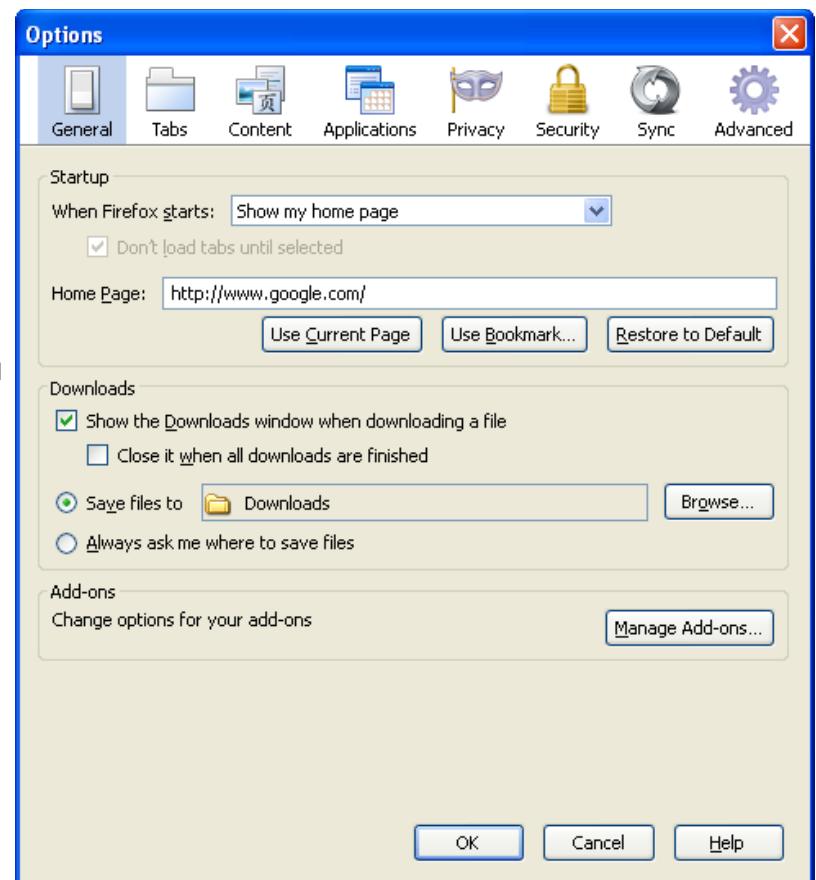
Click on Tools -> Internet Options



Internet Explorer

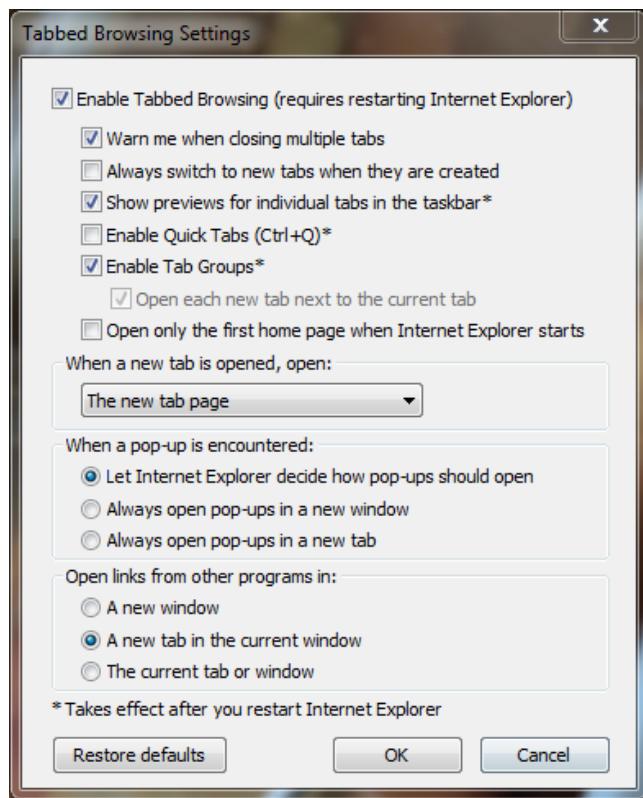
Cookies, History, default opening page

Click on Tools -> Options

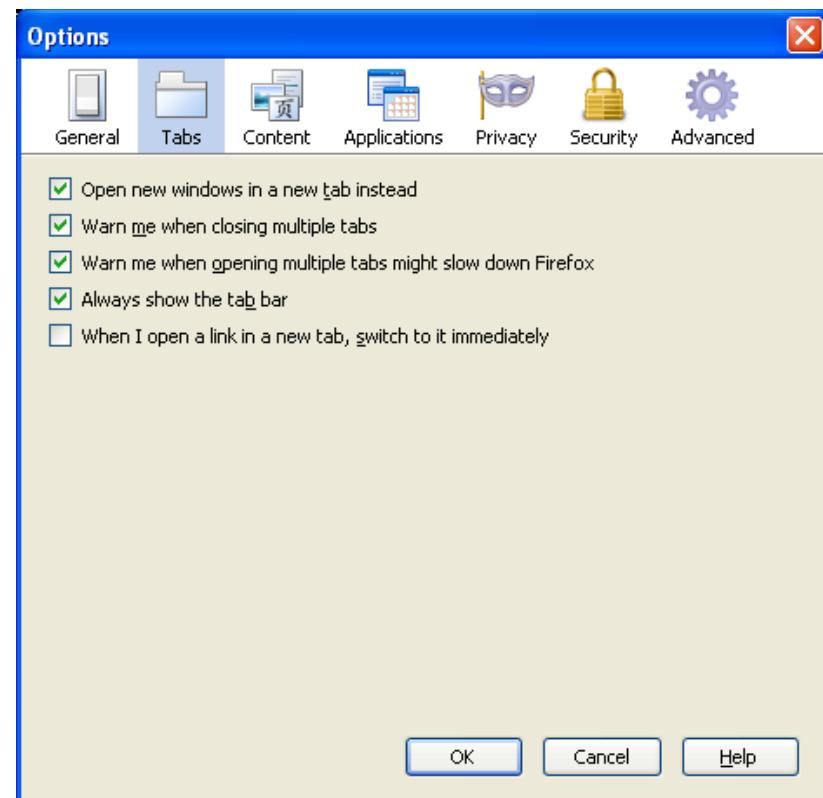


Firefox – Tools | Options

IE and Firefox Tab Control Options

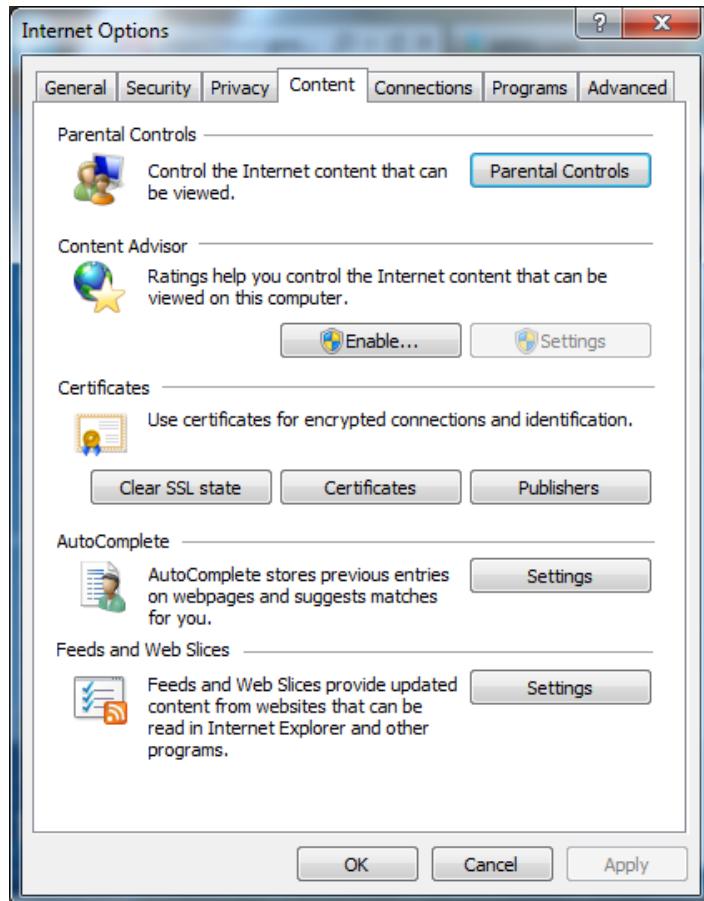


Internet Explorer

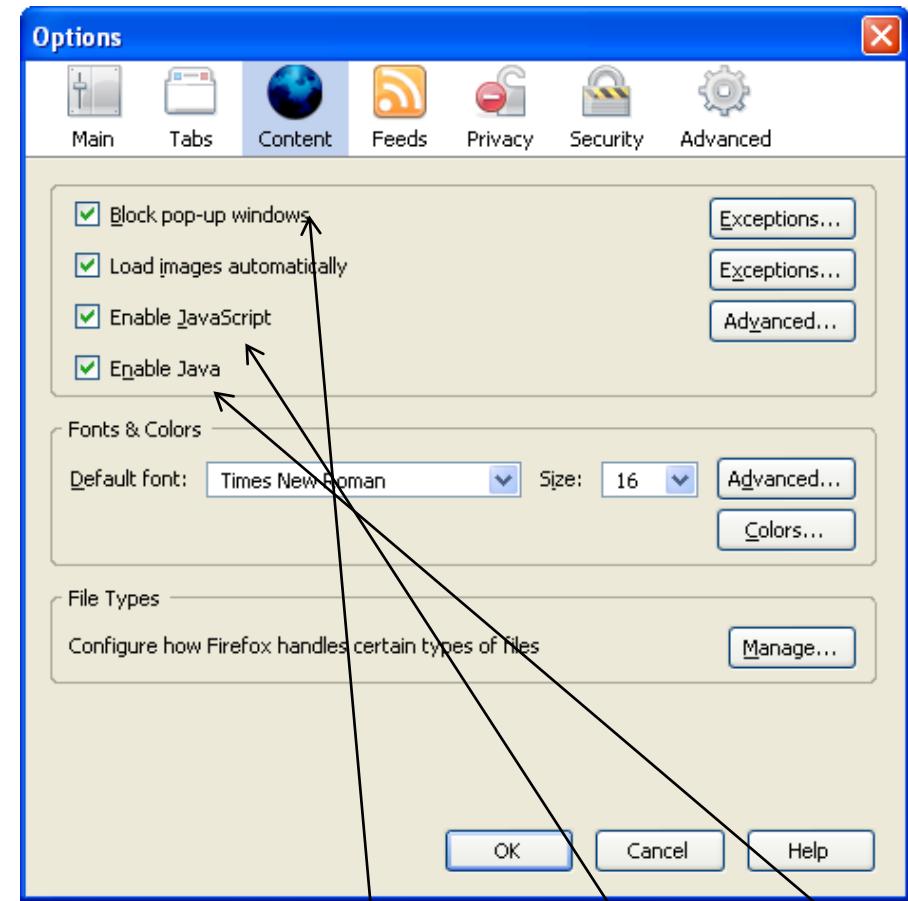


Firefox

IE and Firefox – Content Options

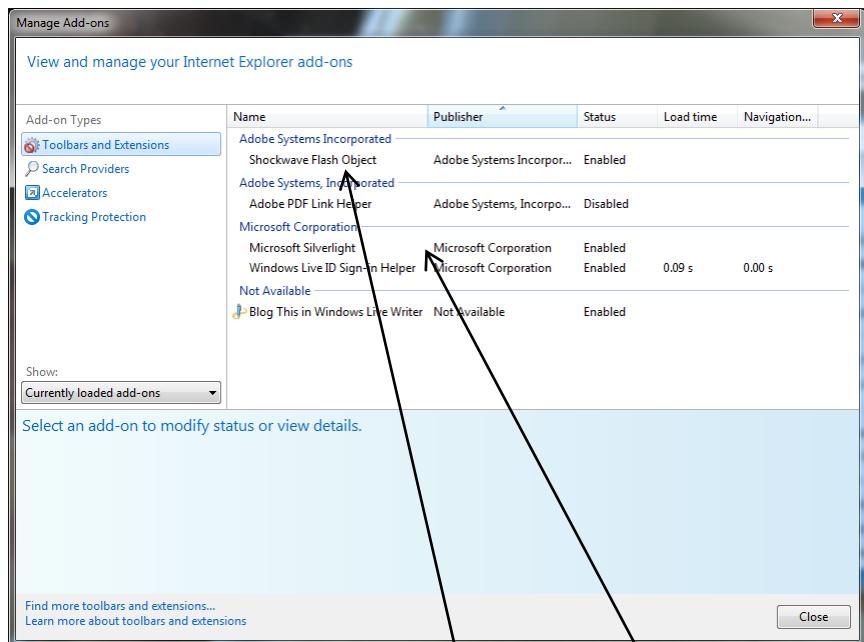


Internet Explorer

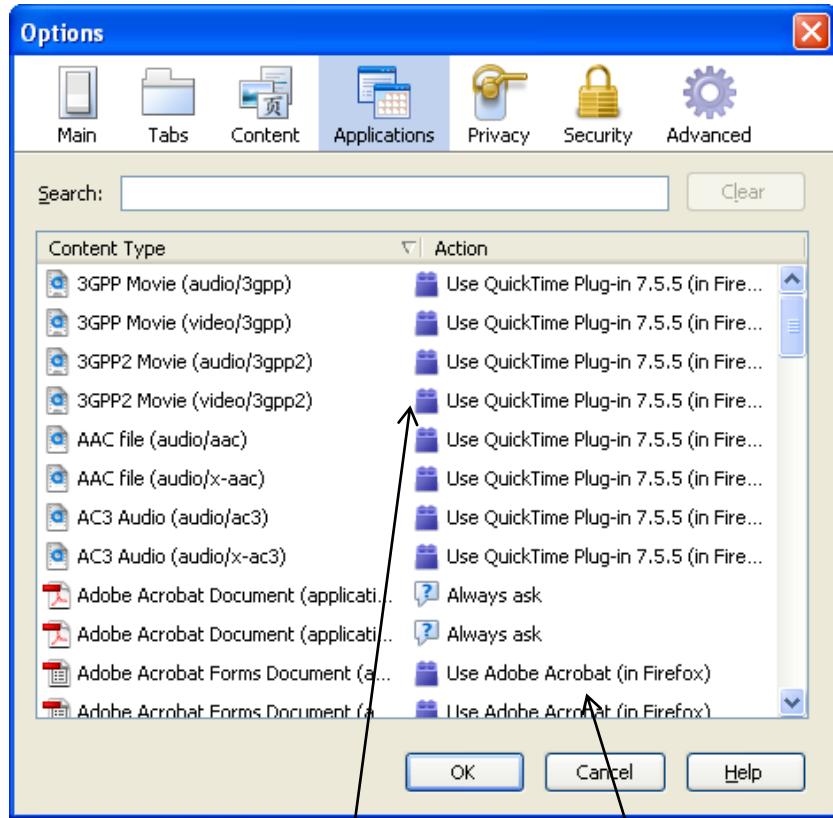


Firefox (PopUps, JavaScript, Java)

IE and Firefox Applications Options

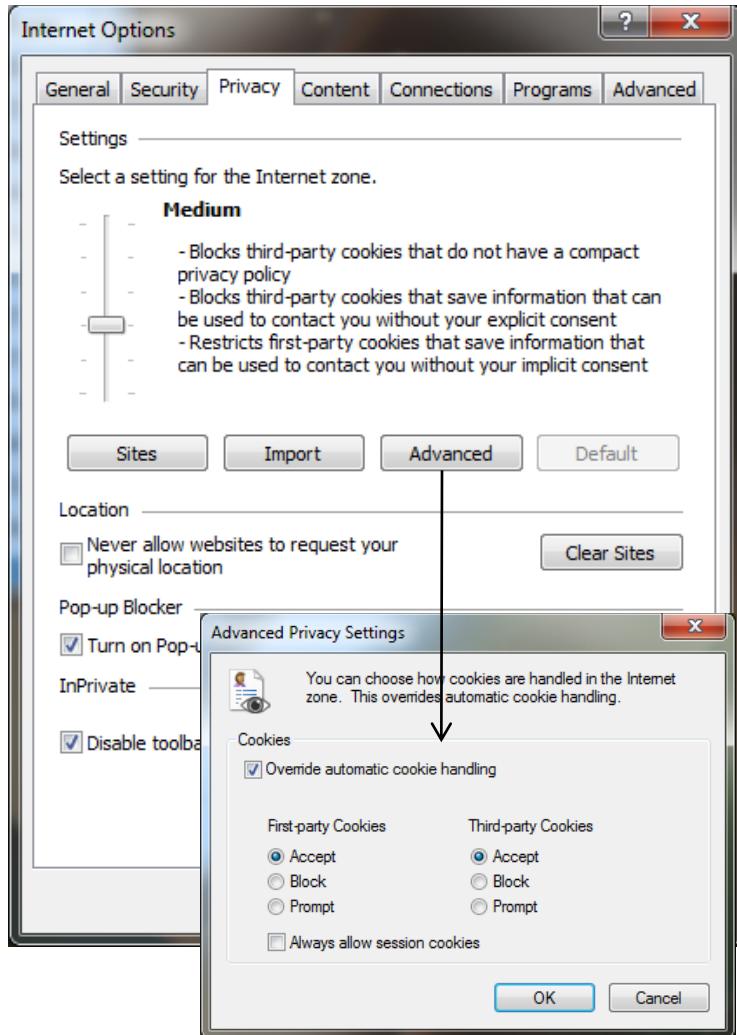


Internet Explorer (Adobe, Silverlight)
(Microsoft has now dropped Silverlight)

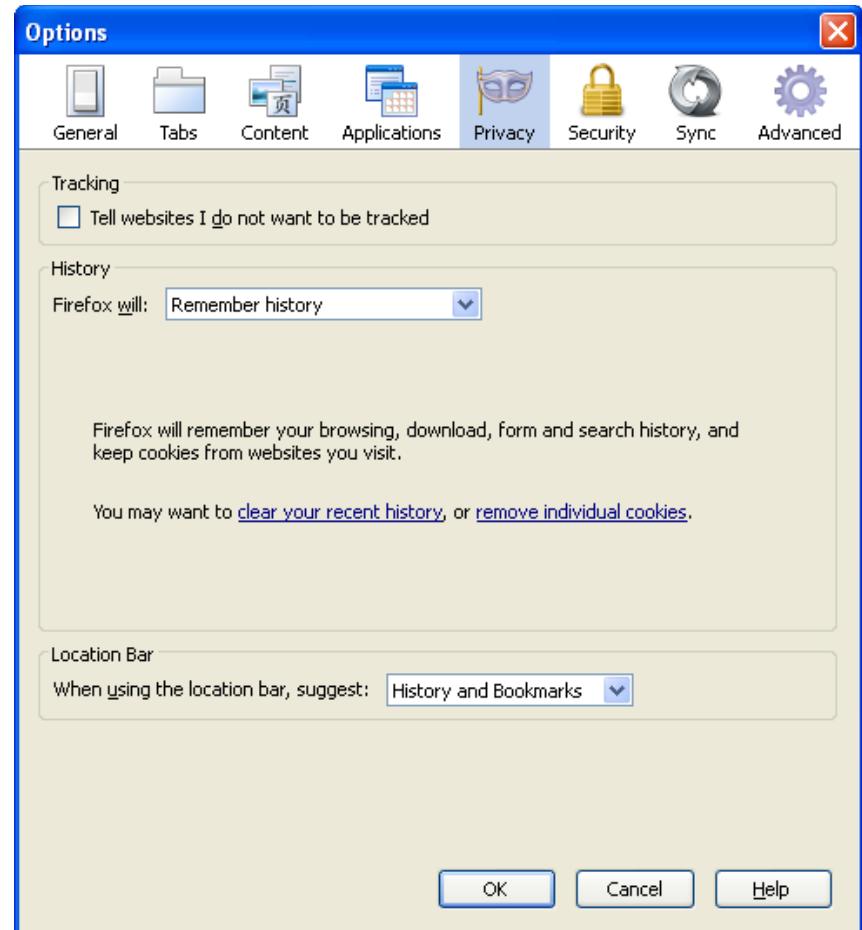


Firefox (Quicktime, Acrobat)

IE and Firefox Privacy Options

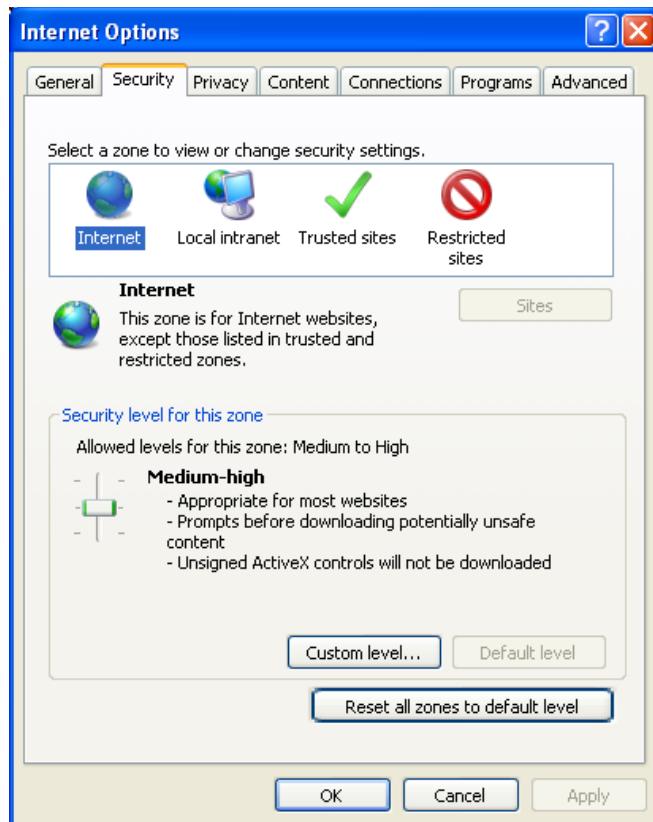


Internet Explorer (Cookies)

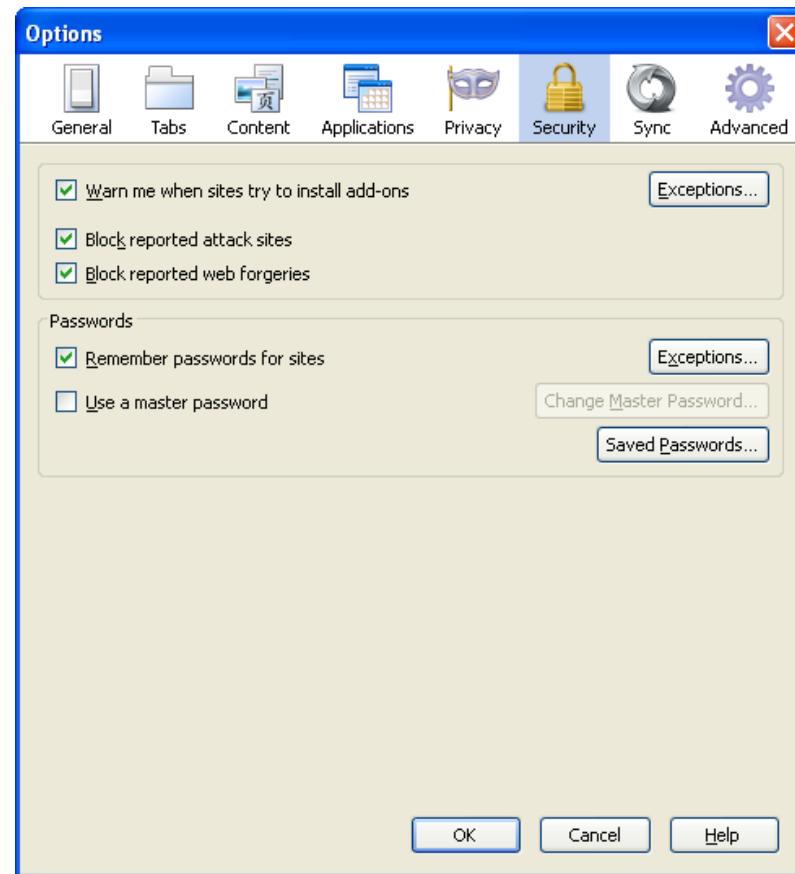


Firefox (History, Cookies)

IE and Firefox Security Options

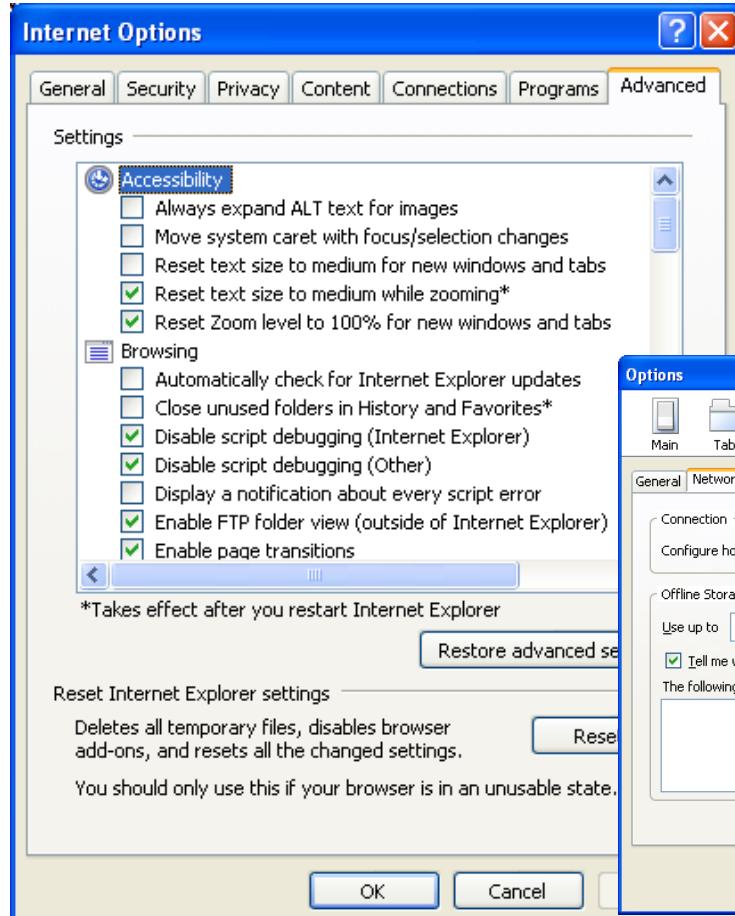


Internet Explorer



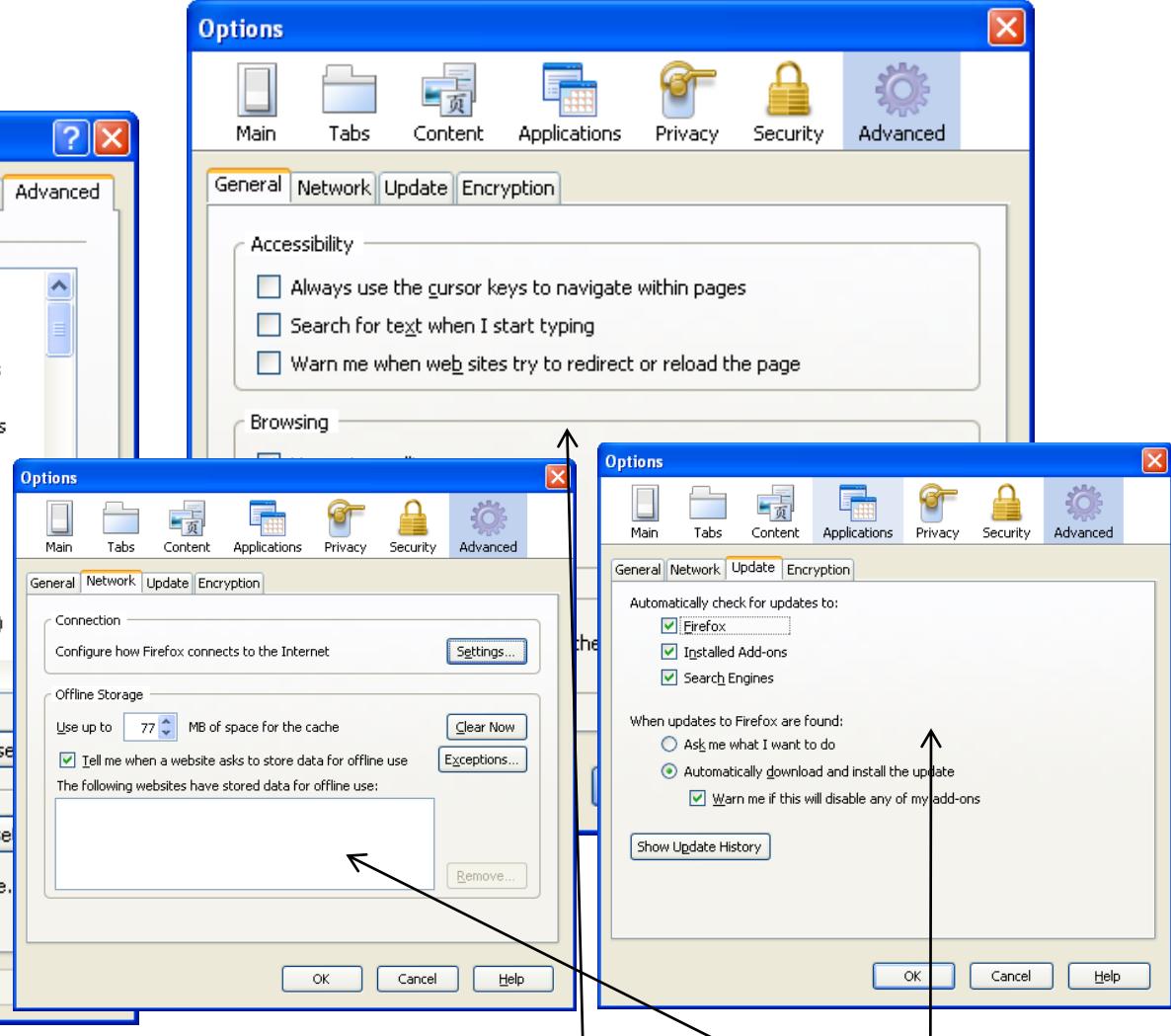
Firefox

IE and Firefox Advanced Options



Internet Explorer

Copyright © Ellis Horowitz & Marco Papa
1999-2016

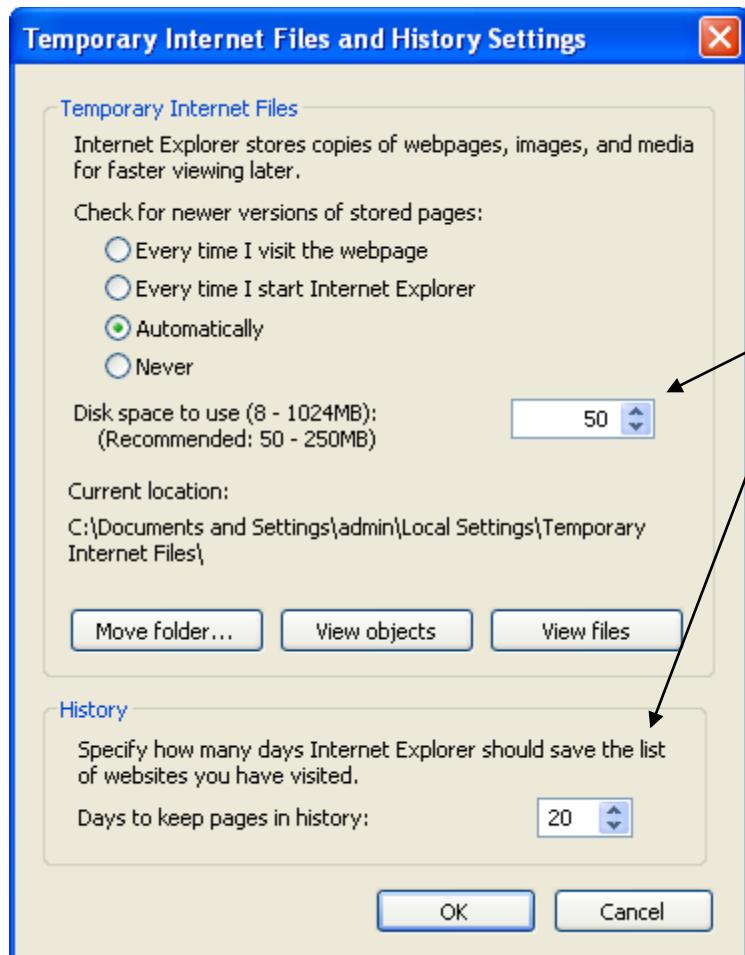


Firefox (General, Network, Update)

Course Intro

42

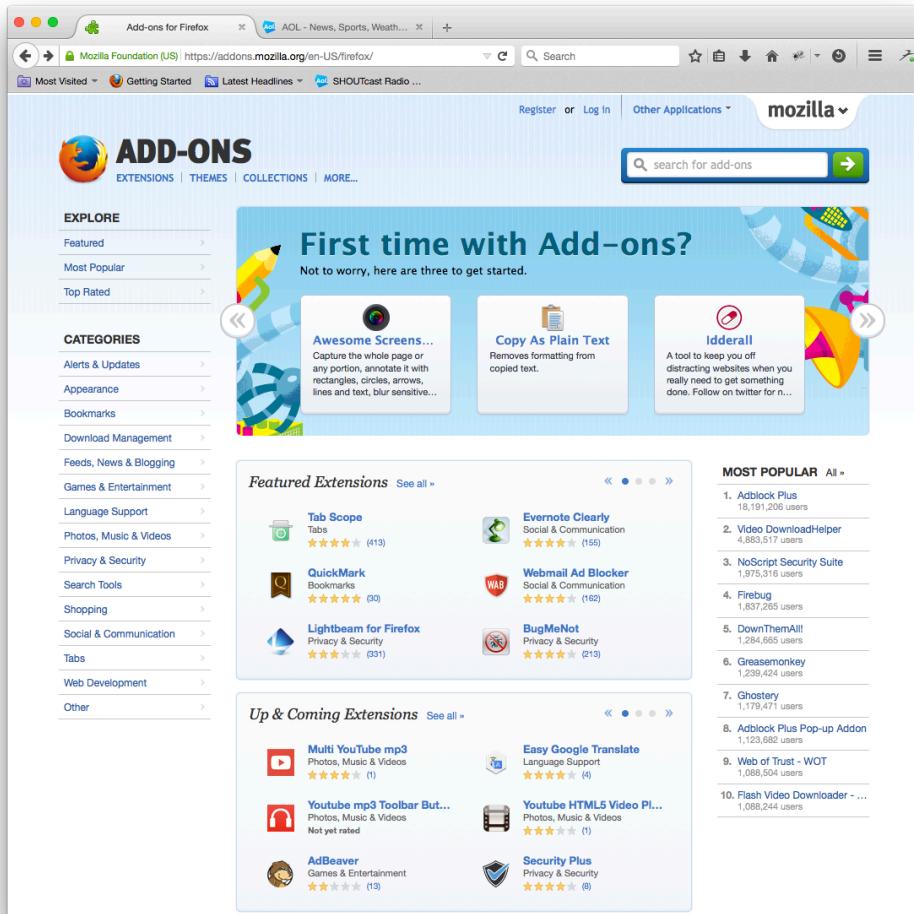
Internet Explorer Browser Caching



- History
 - Links and URLs that have been accessed by the browser over a period of time
- Disk cache
 - Temporary internet files, a folder on the disk that contains cached copies of files
- Memory cache
 - Session-based information that is cached during the session
- Offline content
 - Web content is downloaded when online and viewed offline

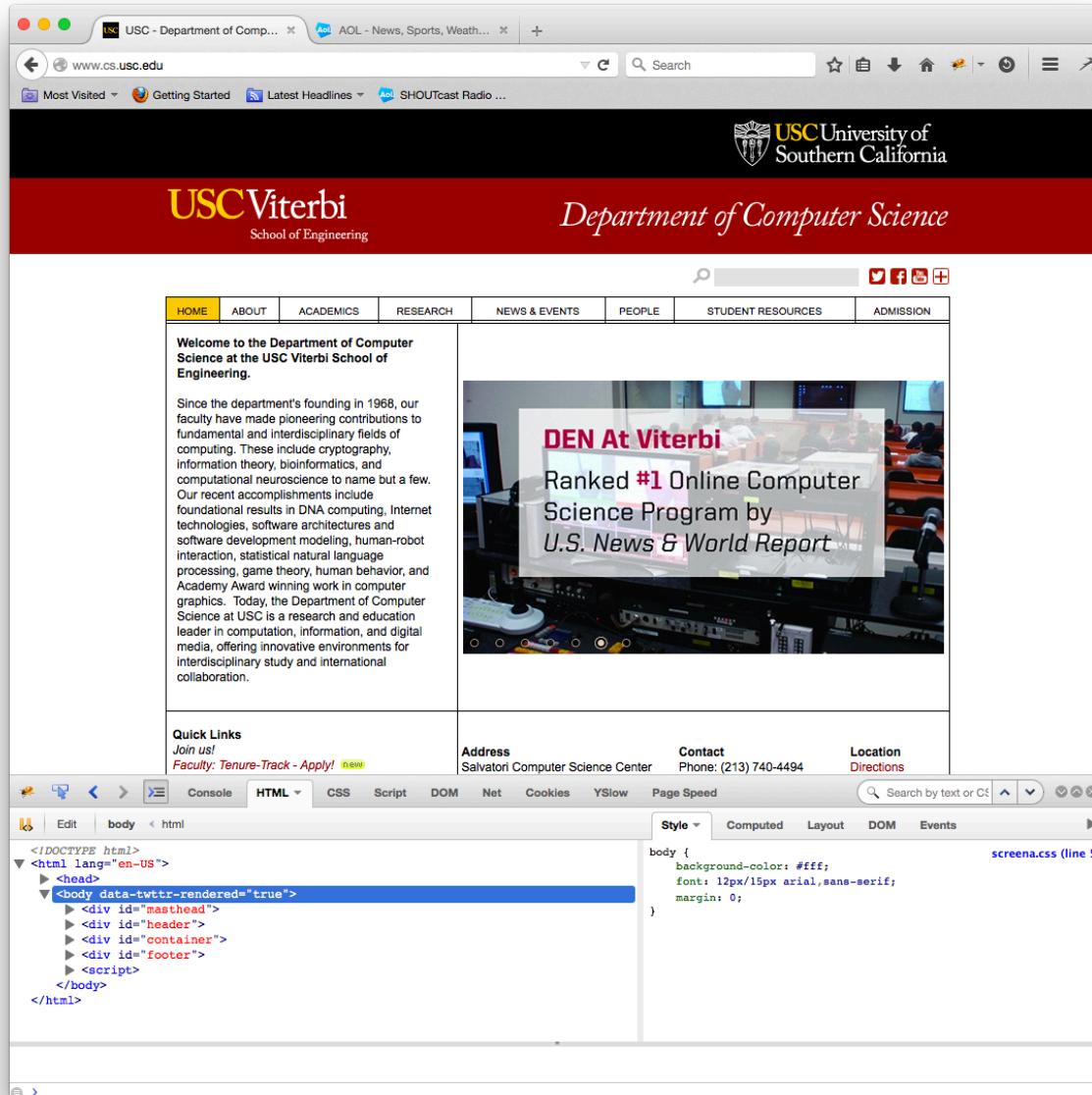
IE caching options screen (Tools | Options | General | Browsing History)

Browsers Have Many Plugins Available

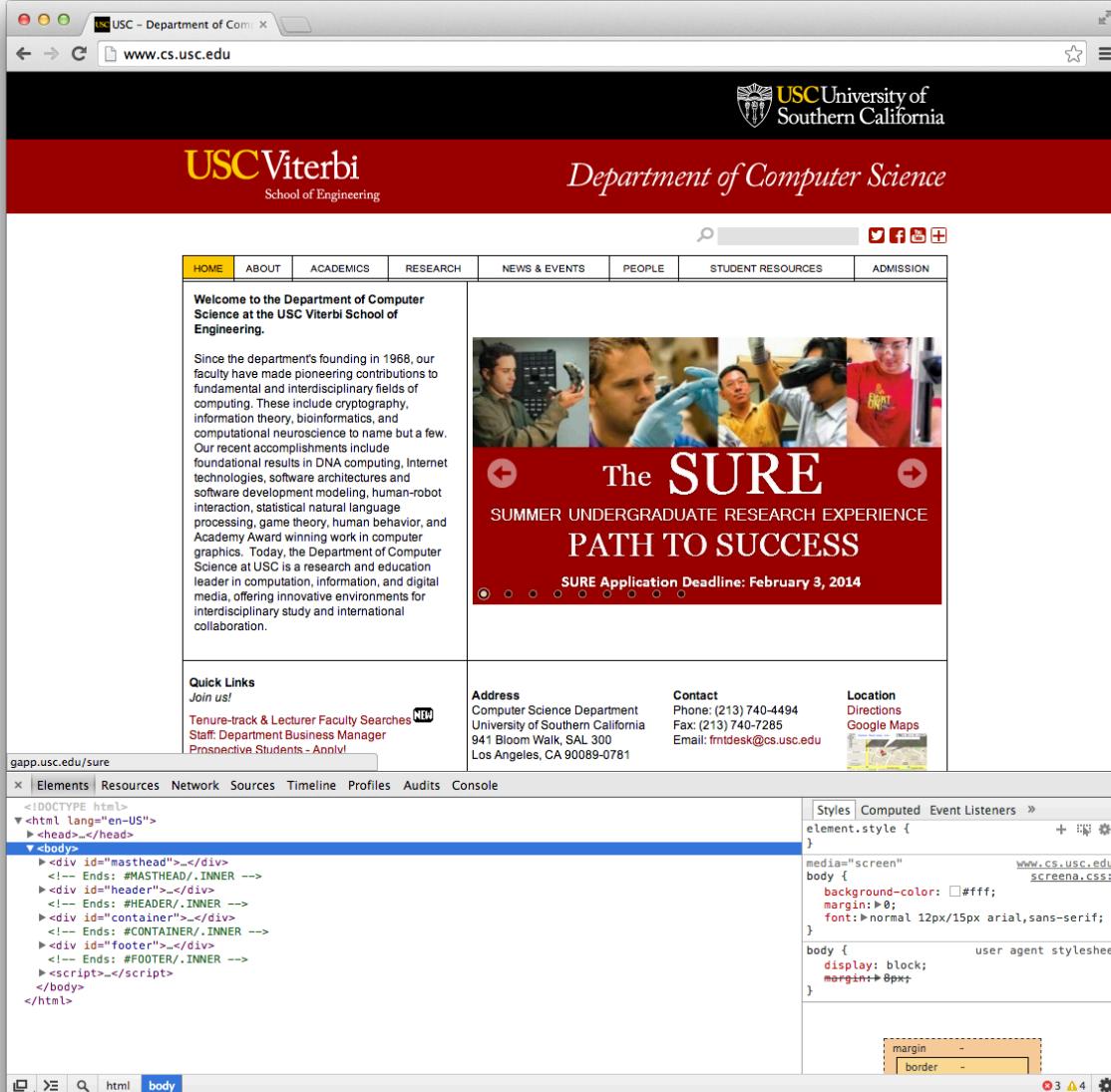


- Three Firefox plug-ins that will be especially useful in this course are:
 - Live HTTP Headers
 - Firebug
 - YSlow
- Available at:
 - addons.mozilla.org
- More about them later on in the semester

Firefox: Tools | Web Developer | Firebug



Chrome: Menu | More Tools | Developer Tools



Evolution of Web Sites

