

Standards, Tools and Web 2.0

Web Programming

Uta Priss
ZELL, Ostfalia University

2013

Outline

Guidelines and Tests

Logfile analysis

W3C Standards

Tools

Web 2.0

Nielsen's 10 Usability Heuristics (1995)

- ▶ Visibility of system status
 - ▶ Match between system and the real world
 - ▶ User control and freedom
 - ▶ Consistency and standards
 - ▶ Error prevention
 - ▶ Recognition rather than recall
 - ▶ Flexibility and efficiency of use
 - ▶ Aesthetic and minimalist design
 - ▶ Help users recognise, diagnose, and recover from errors
 - ▶ Help and documentation
- usability, interaction design, user experience

Information Architecture for the Web (around 2000)

- ▶ Information architecture attempts to balance the technical requirements with user requirements and information system requirements.
- ▶ Systems analysis techniques can be applied to information design.
- ▶ A variety of common practices (such as use of metaphors, common elements of websites) can be identified and gathered in informal guidelines.

Information Architecture: Design Guidelines

- ▶ Make it easy to find.
- ▶ Provide clear ordering options.
- ▶ Provide rich functionality.
- ▶ Make related items available.
- ▶ Keep forms simple.
- ▶ Ensure security of information.
- ▶ Metaphors: “home page”, “chat room”, “guest book” .
- ▶ Consistent and recognisable terminology: “site search”, “about”, “contact us”, “home” .
- ▶ Icons and metadata

Navigation

- ▶ How many links per webpage?
- ▶ Is the navigational hierarchy deep or shallow?
- ▶ Are there many paths to a single page or is there only one path?
- ▶ Navigational elements on a single page: “home”, “back”, “up”
- ▶ Use of a sitemap

Testing and Quality Assurance

- ▶ software testing (unit testing, link checker, html validation, html tidy, ...)
- ▶ SEO (search engine optimisation), logfile and web statistics analysis
- ▶ usability testing
- ▶ accessibility testing
- ▶ security testing

Apache access log

```
www.upriss.org.uk:80 220.181.108.161 - -  
[22/Sep/2013:06:43:12 +0000]  
"GET / HTTP/1.1" 200 469 "-"  
"Mozilla/5.0 (compatible; Baiduspider/2.0; +http://www.baic
```

Also of interest: Apache error log


Which pages are most visited

Top 30 of 680 Total URLs					
#	Hits		KBytes		URL
1	8798	15.08%	349968	16.48%	/fca/examples/liveinwater.png
2	8363	14.34%	430503	20.28%	/fca/examples/liveinwater.at2.png
3	1991	3.41%	639	0.03%	/emadr.gif
4	1389	2.38%	5833	0.27%	/python/PythonCourse.html
5	921	1.58%	3183	0.15%	/
6	721	1.24%	2153	0.10%	/python/session1.html
7	641	1.10%	94464	4.45%	/perlcgi/l2.pdf
8	536	0.92%	176	0.01%	/fca/fca.ess
9	512	0.88%	2228	0.10%	/fca/fca.html
10	348	0.60%	17540	0.83%	/fca/fca.jpg
11	345	0.59%	85756	4.04%	/maths/mlec1.pdf
12	319	0.55%	1165	0.05%	/python/session2.html
13	279	0.48%	2434	0.11%	/fca/fcasoftware.html
14	249	0.43%	1230	0.06%	/python/session3.html

Most frequent visitors to my site

Top 30 of 11154 Total Sites									
#	Hits		Files		KBytes		Visits		Hostname
1	3238	5.55%	971	2.31%	41618	1.96%	343	4.09%	crawl-66-249-73-41.googlebot.com
2	1016	1.74%	686	1.63%	3765	0.18%	165	1.97%	208-115-111-72.reverse.wowrack.com
3	956	1.64%	641	1.53%	3406	0.16%	162	1.93%	208.115.113.88
4	482	0.83%	445	1.06%	2334	0.11%	18	0.21%	hosted-by.leaseweb.com
5	462	0.79%	456	1.09%	2229	0.10%	3	0.04%	ukspider4.wise-guys.nl
6	449	0.77%	443	1.05%	35742	1.68%	5	0.06%	csxcrawlstorage01.list.psu.edu
7	422	0.72%	60	0.14%	3006	0.14%	0	0.00%	62-84-58-30.customers.almanet.kz
8	320	0.55%	4	0.01%	11	0.00%	1	0.01%	183.60.244.29
9	320	0.55%	4	0.01%	11	0.00%	1	0.01%	183.60.244.46
10	304	0.52%	241	0.57%	12382	0.58%	96	1.15%	spider-100-43-83-143.yandex.com

From which countries

Countries (Top 25) - Full list				
Countries			Pages	Hits
	United States	us	2,291	11,003
	Great Britain	gb	523	2,997
	China	cn	459	977
	India	in	454	1,611
	Germany	de	451	2,484
	Canada	ca	262	1,169
	France	fr	238	1,394
	Netherlands	nl	159	740
	Ukraine	ua	118	265
	Australia	au	90	713
	Italy	it	85	627

From which websites



The screenshot shows a table with the following data:

	Direct address	Bookmark	Link in email
Links from an Internet Search Engine - Full			
- Google	940	3,517	
- Baidu	16	24	
- Unknown search engines	12	20	
- Yahoo!	11	30	
- Microsoft Bing	9	29	
- Ask	4	13	
- WebCrawler	3	4	
- Yandex	2	3	
- DMOZ	1	1	
- Alexa	1	1	
- Digg (Social Bookmark)	0	38	
- AOL	0	1	

Which browsers

Browsers (Top 10)	
	Google Chrome
	Firefox
	Safari
	MS Internet Explorer
	Opera
	Android browser (PDA/Phone browser)
	Mozilla
	Unknown
	iPhone (PDA/Phone browser)
	NetNewsWire (RSS Reader)
	Others

Question:

What kind of information from logfiles is most relevant for your website?

(If you don't have a website, imagine you are maintaining a website for a small, local company.)

World Wide Web Consortium (W3C)

- ▶ Founded by Tim Berners-Lee
- ▶ International consortium
- ▶ Members: businesses, nonprofit organisations, universities, governmental entities, and individuals
- ▶ Criticised as being dominated by larger organisations
- ▶ Developers do not always comply with the standards (e.g. browsers incompatibilities)

W3C Standards

- ▶ Web Design and Applications (HTML, XHTML, CGI, CSS, SVG, WCAG)
- ▶ Web Architecture (protocols, URI, HTTP)
- ▶ Semantic Web (linked data, OWL, SPARQL, SKOS)
- ▶ XML Technology (XML, XPath, XQuery, XSLT)
- ▶ Web of Services (SOAP, WDSL)
- ▶ Web of Devices (mobile, voiceXML)
- ▶ Browsers and Authoring Tools (ATAG)

Validating whether websites comply with W3C standards:

Markup validation: <http://validator.w3.org>

Mobile content: <http://validator.w3.org/mobile/>

Broken links: <http://validator.w3.org/checklink>

Internationalisation: <http://validator.w3.org/i18n-checker/>

CSS and others: <http://www.w3.org/QA/Tools/>

Not W3C:

Accessibility and usability evaluation

Guidelines for content

Security testing

Browser-based testing tools:

Example: Firefox (user contributed tools)

- ▶ Development: Firebug, Web Developer
- ▶ change headers: TamperData, Modify Headers, RefControl
- ▶ JavaScript Debugger, JavaScript Deobfuscator
- ▶ Security: SQL Inject Me, HackBar, XSS Me, SQL Injection

Webmaster tools:

- ▶ SEO tools: keyword analysis tools, search position checkers, metatag generators, spam block list checkers
- ▶ Web utilities: DNS/WHOIS/IP address lookup
- ▶ Tools for finding and generating domain names
- ▶ Other tools: CSS menu generators, colour pickers, anti-malware scanning, sitemap generators

(Free tools available on-line. Hosting packages often come with such tools.)

Two types of tools for creating content:

CMS: Content Management Systems
(Joomla!, Drupal, Typo3)

WAF: Web application frameworks
(Django, Ruby on Rails, Zend Framework for PHP)

Features which CMS and WAF have in common:

- ▶ separation of content and design (content stored in database)
- ▶ user administration (access rights)
- ▶ plugins and components

Examples of plugins for CMS:

search facility, guestbook, wiki, e-commerce solutions

CMS vs WAF

CMS	WAF
for static websites (which might change often)	for dynamic websites, web applications, web services
collaborative development project management	program-driven development environment

Agile development methodology (used for WAF)

MVC (Model-View-Controller)

- ▶ Model: data, functions, rules
- ▶ View: templates for the representation
- ▶ Controller: connecting model and view

DRY: Don't repeat yourself

ORM (Object-relational mapping)

Classes in the programming language correspond to tables in the database.

Persistence (where does a programming language store its data?)

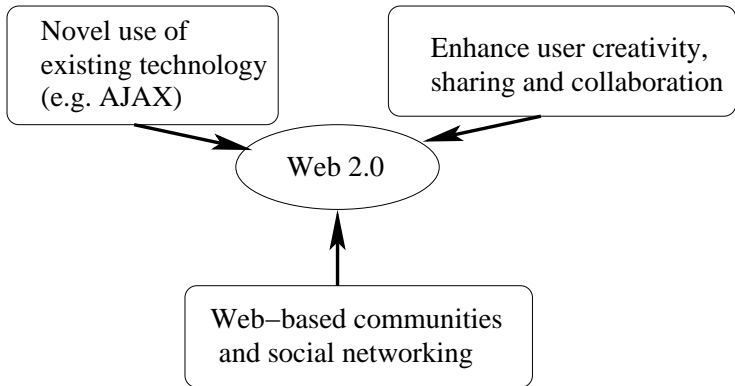
“Convention over Configuration”

History: Web 2.0

Web 2.0 was a buzzword from 2004,
nevertheless it is interesting to look at what it meant.

It is followed by Web 3.0 = the Semantic Web.

Web 2.0



Examples: eBay, Wikipedia, del.icio.us, Skype, YouTube.

Overview of some of the technologies

- ▶ Rich Internet applications
(websites that mimic desktop applications)
- ▶ Web services
(API access to remote web servers)
- ▶ Cloud-based services
- ▶ And ...

Social software and social networking sites

- ▶ Older: Usenet, Listserv, Talk, Email.
- ▶ Blog, forum, wiki.
- ▶ Networking: friends sites and recommender systems.
- ▶ User generated content and self-description sites.
- ▶ FOAF and XFN: protocols for friends networks.
- ▶ But: privacy concerns.

Examples: MySpace, Facebook, ...

Folksonomies

Collaborative/social + classification/indexing/tagging

- ▶ Shared bookmarking.
- ▶ Bookmarks/tags are annotated with keywords.
- ▶ Visualisations: tag clouds.
- ▶ APML: protocol for import/export of tag clouds.

XHTML social RSS tagging
blogs Web 2.0 bookmarking XML
web services Javascript AJAX wiki

Examples: Flickr, del.icio.us

Web feeds: RSS and Atom

- ▶ RSS 2.0: Really Simple Syndication.
- ▶ RSS 1.0: RDF Site Summary.
- ▶ Automated updates for frequently changing content: blogs, news, podcasts.
- ▶ Aggregation of content from different sites (Mashups).
- ▶ Used in feed readers, browsers or desktop widgets.



Freemium business model

Venture capitalist Fred Wilson's Blog on 23 March 2006:

Give your service away for free, possibly ad supported but maybe not, acquire a lot of customers very efficiently through word of mouth, referral networks, organic search marketing, etc., then offer premium priced value added services or an enhanced version of your service to your customer base.

Used for Shareware and Open Source software.