30% Accessible - A Survey of The UK Wide Web

Dave Beckett

Computing Laboratory, University of Kent at Canterbury
Canterbury, Kent, CT2 7NF, England

D.J.Beckett@ukc.ac.uk

Abstract

This paper describes a comprehensive survey of the UK domain and of UK Web site Home Pages. The survey determined the features of the typical WWW page and analysed the HTTP and HTML in general terms, for the use of standards and for accessibility. Finally, the features found were used to calculate a figure for the overall accessibility of the UK Web pages.

Introduction

Surveys of the *content* of the Web have been done before using Web crawler technology in [Woodruff1996] with Inktomi(1) and [Bray1996] with Open Text(2) systems respectively. However, the emphasis in these papers has been on *comprehensive* surveys of large numbers of documents to determine, amongst other things, current HTML tag use, connectivity and to find novel ways to the visualise complex web connectivity.

There has to date been little published work that surveys how *accessible* the web is -- how the use of HTTP and HTML affects the usability of the web to *all* -- such as those using graphical browsers with images turned off, text only browsers, browsers with small screens or browsers that do not support the latest cute feature. This survey attempts to discover the accessibility of the UK Web and took place on the 1st December 1996.

The first step taken was to determine the scope and structure of the UK Internet domains and the Web based upon them and from that, to summarise the web pages found. Having retrieved the pages, the HTTP headers and the content of each web page was analysed, and the typical web page described. An analysis of the *accessibility* issues arising from each of the elements surveyed was then carried out as these elements were found.

The UK Wide Web (UKWW)

The UKWW is a mature and fast growing part of the global Internet. It has been evolving over several decades from earlier networks using other protocols (such as UUCP, Coloured Book). The UK is slightly unusual in having two country codes UK (*United Kingdom*) and GB (*Great Britain*) and hence two top level domains `.uk` and `.gb`; but `.uk` is the main one used.

The structure of the `.uk` domain is hierarchical by category of the organisation, as described in [UKNIC1996] at the UK Network Information Centre or UKNIC(3). The authority for the sub-domains is currently delegated to bodies in three communities -- UKERNA(4) for `ac.uk` (Further and Higher Education) and `gov.uk`; CCTA(5) for other governmental domains (`nhs.uk`, `mod.uk`, ...), and Nominet UK(3) for the remaining domains. UKERNA and Nominet are run as not-for-profit limited companies. Nominet organisation members currently include most UK Internet Service Providers (ISPs) and UKERNA.

The UK domain structure

The detailed domain structure was found from records in the UK Domain Name System (DNS)[Mockapetris1987] using the `host`(6) program which can do a recursive walk of a domain to list the entire contents. The domain survey was limited to *sites* where a *site* is at the organisation level, rather than internal to an organisation.

On December 1st 1996, 39162 domains were found under `.uk`, with the main top-level domains structured as shown in **Table 1**.

Top Level Domain	Sub-domains	Total domains	% of Total
co.uk	5	35406	90.40%
org.uk	2	1791	4.57%
ac.uk	0	871	2.22%
gov.uk	1	413	1.05%
ltd.uk	0	321	0.81%
net.uk	0	115	0.29%
plc.uk	0	80	0.20%
nhs.uk	16	25	0.06%
icnet.uk	0	19	0.04%
mod.uk	0	12	0.03%
police.uk	0	5	0.01%
parliament.uk	0	1	0.01%
sch.uk	73	93	0.23%
Total uk domains		**39162**	**100.00%**

Table 1: The 13 Main Top Level .uk domains(*)

Note: (*) There are some *sites* that have domains at the top level that are not included in this table such as `.jet.uk` (which was placed at the top level before this structure was developed) and top-level sites such as `www.nic.uk`(3), the UK NIC. In addition there are some depreciated *domains* such as `govt.uk` and `orgn.uk`, for which sites are now stored under `gov.uk` and `org.uk` respectively.

Accessibility Discussion

Table 1 shows that most of the domains are in the commercial `co.uk` domain which remains the most rapid growing domain. This large domain and fast registration rate means that the accessibility issues in this area are related to the problem of clashing requests for domains names and getting access to domains representing UK trademarks and company names. A new registrant may find that the domain representing its UK trademarks and/or company name may already have been registered by another entity. This causes two further problems -- confusion by third parties that it actually represents the company and difficulties for people searching domains for the company they know.

To solve these problems for the commercial domains, Nominet has recently developed a process (that is intended to be automated) as described in [Carey1996] which can be used to generate domain names from the legal, registered UK company names. These new domains are stored under `ltd.uk` for Limited liability companies and `plc.uk` for Public Limited Companies (PLCs). This process has two distinct advantages: it provides unique domain names which will not clash and these are based on the legal, registered company names so the existing UK companies can be guaranteed to get the names that they have.

At the survey date there were only 391 `ltd.uk` and `plc.uk` domains. If `ltd.uk` and `plc.uk` domains are included, then commercial domains account for 35807 domains or 91.4% of all the `.uk` domains.

In conclusion: the new process should give accessibility for commercial organisations which do not want a `.co.uk` domain where the existing problems still apply.

WWW Sites

Domains may be registered, but it is not necessarily the case that they are being actively *used* i.e. there may not be a WWW site for the domain. To determine this, for each domain, a name resolve was attempted for the ''standard'' WWW site *www*`.foo.co.uk` for domain `foo.co.uk`. If the name existed, then an HTTP request was attempted for the URL `http://www.foo.co.uk/` -- the *Home Page* of the site. The LibWWW-Perl(7) library was used to perform the HTTP protocol requests. The WWW pages used were fresh (i.e. < 7 days old at survey date).

WWW pages may be hosted by ISPs, but domains are also registered by specialised name registration companies. Hence each WWW page is not *necessarily* unique -- multiple domains can point to the same WWW page. To determine this situation, the HTML bodies of the WWW pages were checksummed using MD5[Rivest1992], and the unique pages found. Note that the HTTP headers are excluded from the checksum because they have fields that include date information.

The survey found 39162 domains, of which 18754 domains (47.89%) had no WWW sites and 20408 domains (52.11%) were represented by 13312 unique WWW home pages. 31068 domains (79.33%) were unique, but one WWW page represented 1457 domains, and several others represented hundreds of domains. It was likely many more domains are also registered and not currently used; but they could not be identified as duplicates if no WWW page could be retrieved.

For each unique WWW site, the retrieved information from the HTTP `GET` requests consisted of two parts -- the HTTP response headers and an HTML body. The response headers were analysed by counting fields, while the HTML pages were subjected to more extensive analysis including validation, detailed checks on the used tags, attributes and content of the page.

HTTP Headers

For each HTTP `GET` request done, there were usually 5 or 6 headers present in the response (for 78.08% of responses) and there were 40 different header types seen, as summarised in **Table 2**. The top 5 WWW Servers seen (with and without version numbers), are summarised in **Table 3**.

Order	HTTP Header	Frequency	Percent	Order	HTTP Header	Frequency	Percent
1	Content-Type	13301	99.92%	14	Content-Base	128	0.96%
2	Server	13291	99.84%	15	Pragma	88	0.66%
3	Date	13117	98.54%	16	Content-Transfer-Encoding	70	0.53%
4	Last-Modified	11101	83.39%	17	Keywords	10	0.08%
5	Content-Length	10732	80.62%	18	PICS-Label	8	0.06%
6	MIME-Version	1402	10.53%	19	Accept	7	0.05%
7	Accept-Ranges	778	5.84%	20	Security-Scheme	4	0.03%
8	Allow-Ranges	556	4.18%	21	Connection	3	0.02%
9	Set-Cookie	542	4.07%	22	Refresh	3	0.02%
10	Expires	434	3.26%	23	URI	3	0.02%
11	Allow	289	2.17%	24	Version	2	0.02%
12	Message-ID	167	1.25%	25	Description	2	0.02%
13	Link	140	1.05%	26..40	(15 more)	15	0.15%
					Total Responses	**13312**	**100%**

Table 2: HTTP Response Headers

Order	WWW Servers With Version	Count	Percent	Order	WWW Servers Without Version	Count	Percent
1	Apache/1.1.1	3002	22.55%	1	Apache	7289	54.76%
2	Apache/1.0.5	1551	11.65%	2	Microsoft-IIS	998	7.50%
3	Apache/0.8.14	1344	10.10%	3	NCSA	935	7.02%
4	CERN/3.0	633	4.76%	4	CERN	659	4.95%
5	Microsoft-IIS/1.0	502	3.77%	5	Netscape-Commerce(*)	553	4.15%
	Total (200 servers)	**13312**	**100%**		**Total (90 servers)**	**13312**	**100%**

Table 3: Top 5 WWW Servers - With / Without Version Numbers

Note: (*) The total of all Netscape servers rather than just the Commerce one is 1319 which would be #2 with 9.91% of total.

Accessibility Discussion

The HTTP response headers do not much impact on the accessibility of WWW sites since HTTP was designed to be ''*Future Compatible*'' i.e. new headers do not affect old implementations -- they can safely ignore them. There are two exceptions to this from the commonly seen headers:

- `Set-Cookie` header for ‘‘Cookies’’ but very few sites *require* them.
- `Refresh` header which tells the *browser* to re-GET the page after a certain time period, and may also add a new URL to make the browser fetch a different page. This is used on a few sites to generate *Splash Pages* -- pages that appear for a short while with logos and then disappear, to be replaced with the main content of the site. This is not very accessible if no actual anchor link is given on the page.

HTML Validation

The current version of HTML is *HTML 3.2*[Raggett1997] and has been specified by the W3C(8) to update HTML 2.0 (but remain compatible with it) by adding commonly deployed features such as tables, applets and text flow around images. This recommendation describes that ‘‘HTML documents are SGML documents’’ and ‘‘HTML 3.2 is a SGML application conforming to International Standard ISO 8879 -- Standard Generalized MarkUp Language’’. This means that HTML has an SGML Document Type Definition (DTD) and that it can be validated against it with an SGML parser.

In fact, there are several HTML DTDs for different versions of HTML. For a particular HTML document to be considered a correct SGML document, the DTD should be present in a `<DOCTYPE>` declaration at the start of the document. In the 13312 WWW pages, 202 pages had illegal `<DOCTYPE>` syntax, mostly missing terminating > or text before the `<DOCTYPE>`. In total, 3490 DTDs were seen, of which 90 were unique but the top 7 DTDs corresponded to 2995 -- 85.82% of all the DTD seen. These results are shown in **Table 4**.

Order	HTML DTDs	Frequency	Percent
1	-//IETF//DTD HTML//EN	1094	31.35%
2	-//SQ//DTD HTML 2.0 + all extensions//EN	673	19.28%
3	-//W3C//DTD HTML 3.2//EN	478	13.70%
4	-//W3O/DTD HTML//EN	259	7.42%
5	-//SQ//DTD HTML 2.0 HoTMetaL + extensions//EN	219	6.28%
6	-//IETF//DTD HTML 3.0//EN	183	5.24%
7	-//IETF//DTD HTML 2.0//EN	89	2.55%
	Total of Top 7 DTDs	**2995**	**85.82%**
	Total DTDs seen	**3490**	**100%**

Table 4: Top 7 HTML DTDs Present

If no DTD was found, a default was used -- the latest HTML 3.2 DTD. Reading the DTDs list in **Table 4** more carefully, it can be seen that entries 1,2,4,5 and 7 are HTML 2 (with some extensions) for a total of 2334 or 66.88% of all DTDs, and entries 3 and 6 imply HTML 3 or HTML 3.2 for 661 or 18.94% of all DTDs.

The validation of the document was then carried out using the NSGMLS(9) SGML parser. The results are presented in **Table 5**.

DTD Source	Validation						Totals	
	Succeeded		Failed		Not Possible			
Known DTD in document	210	1.58%	3187	23.94%	-	-	3397	25.52%
Unknown DTD in document	-	-	-	-	93	0.70%	93	0.70%
HTML 3.2 DTD	655	4.92%	9167	68.86%	-	-	9822	73.78%
Totals	865	6.50%	12354	92.80%	93	0.70%	13312	100%

Table 5: DTDs and HTML Validation

The above two tables show that the use of valid HTML is virtually non-existent, and the use of the `<DOCTYPE>` tag to indicate a DTD is inconsistent -- it does not imply that the DTD inside is useful (common) or is an indication of a valid document. The evidence is that most authors do not validate documents or use tools that enforce validation. However the ‘SQ’ DTDs in **Table 4** refer to SoftQuad(10) products which **do** enforce DTD use.

Accessibility Discussion

What impact does validation have on accessibility? That is difficult to say. A minority of pages have DTDs but it seems most authors do not used them as intended -- the HTML they are using is beyond the DTDs, perhaps due to *Feature-Creep* in the browsers. Since so many pages do not validate against a DTD, further analysis of the use of the HTML was necessary to see what effects the individual tags had on accessibility.

Detailed HTML analysis

For each of the 13320 WWW pages retrieved, the HTML was parsed using routines in the LibWWW-Perl(7) library. During the parsing several aspects were measured:

- Tag use.
- Use of fonts and colours.
- Details on the HTML document size / length.
- Mentions of popular products in phrases such as ‘‘**X** *required*’’ or ‘‘*best viewed with* **X**’’
- Use of `<SCRIPT>` languages and `<APPLET>`s.
- Character entities, (&...), presence of non-ISO 8879-1986 (Latin 1) characters.

HTML Tags

Counts were made of the use of HTML tags in the WWW pages and these are shown in **Figure 1**, for the top 10 average number of tags per document, and **Figure 2**, for the top 10 tags used over all documents.

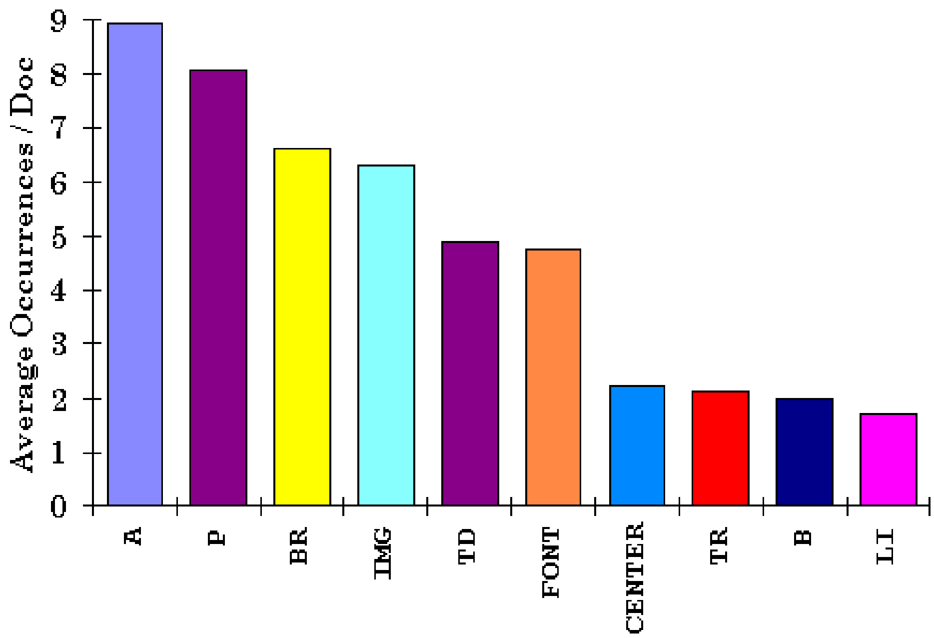

Figure 1: Top 10 of Average Tag Occurrences Per Document

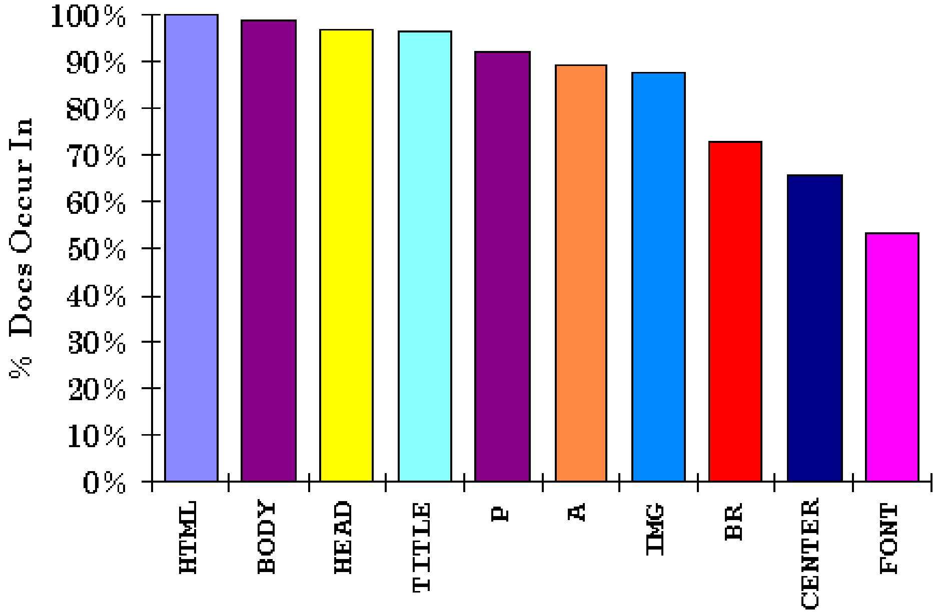

Figure 2: Top 10 Tags Present Per Document(*)

Note: (*) The value of 100% for HTML is an artifact of the HTML parser.

Accessibility Discussion

The high count for `<TR>` and `<TD>` tags imply a large use of tables for formatting and `<FONT>` and `<CENTER>` tags indicate that the *look* of the document is very important to the authors. The `<B>` is a hint that physical emphasis tags are in extensive use. The total of use for all physical emphasis tags (TT, I, B, U, STRIKE, BIG, SMALL, SUB, SUP, BLINK, CENTER) is 67770, or 7.92%; and for logical emphasis or structural tags (CODE, EM, STRONG, DFN, SAMP, KBD, VAR, CITE, DIV) is 16140, or

1.89%. Physical emphasis wins by a factor of 4 but accessibility should not be affected -- tags that are not understood in this area can usually be ignored safely, as long as the new tags are used carefully.

Fonts / Faces

The FACE attribute when used with <FONT> or other tags allows the use of specific fonts in WWW pages. [Note that this is not in the current HTML 3.2 Recommendation.] **Table 6** lists the top 10 faces in use -- there were 144 different fonts in total seen in 1524 uses, but the *Arial* font alone accounted for 48.82% of the total. Microsoft's Internet Explorer 3.0(11) browser first introduced the FACE attribute and consequently the free TrueType fonts Microsoft provides(12) -- which include Arial -- have the most use. Lists of faces are also allowed in the FACE attribute to give alternative suitable fonts and this was seen in 27.71% of the uses of fonts.

Order	Face	Count	Percent	Order	Face	Count	Percent
1	Arial	744	48.82%	6	Times	21	1.38%
2	Helvetica	251	16.47%	7	Arial Narrow	19	1.25%
3	Times New Roman	101	6.63%	8	Courier New	18	1.18%
4	Arial Black	27	1.77%	9	Verdana	17	1.12%
5	Comic Sans MS	26	1.71%	10	MS Sans Serif	14	0.92%
					Total	1524	100%

Table 6: Top 10 Fonts / Faces Seen

Accessibility Discussion

The <FONT> tag can be abused -- for example, using it instead of structured markup -- but when used properly it can enhance the design without affecting the *content*. In general, for browsers that either do not understand it, or do not have the particular font mentioned, use of fonts does not imply lack of accessibility.

Colours

Table 7 shows the top 10 colours in use for HTML text and background -- 4175 different colours were seen in 25053 uses, but most authors are still thinking in Monochrome - the top two colours were White and Black.

Order	Colours	Count	Percent	Order	Colours	Count	Percent
1	#FFFFFF (White)	4625	18.46%	5	#808080 (Gray)	473	1.89%
2	#000000 (Black)	3389	13.53%	6	FFFFFF (White)(*)	472	1.88%
3	#FF0000 (Red)	1884	7.52%	7	#FFFF00 (Yellow)	471	1.88%
4	#0000FF (Blue)	1790	7.14%	8	#000080 (Mid Blue)	345	1.38%
9	#C0C0C0 (Light Grey)	264	1.05%	10	000000 (Black)(*)	238	0.95%
					Total	**25053**	**100%**

Table 7: Top 10 Colours Seen

Note: (*) The white and black colours are duplicated with bad hex format syntax but this does not affect the ordering.

Accessibility Discussion

Bad choices of colours can make a page or links invisible or disappear if clicked and many of these pages would probably appear in one colour on a monochrome display. A little thought can alleviate this and prevent the damage to accessibility.

The Typical WWW Page

The typical WWW page has the features given in **Table 8**.

Feature	Mean Value	Median Value	Median Frequency	Median Percent	Min Value	Max Value
Size (bytes)	2802	10213	330	2.48%	0	464341
Length (lines)	80	107	430	3.23%	1	57882
% of `<IMG>` with ALT text	39.72	0	4778	35.89%	0	100
Number of Internal (same document) links	0.47	0	11760	88.34%	0	116
Number of Local (same site) links	13.07	0	1019	7.65%	0	4105
Number of Remote (outgoing) links	2.61	0	4211	31.63%	0	243
Number of Java Applets	1.26	0	13096	98.38%	0	7

Table 8: Typical Page Features

Overall Accessibility

Table 9 summarises the core HTTP and HTML features that may affect accessibility. A feature that is a *mention* of a product means that the text of the page contained a phrase like ‘‘*requires* **X**’’ or ‘‘**X** *recommended*’’ or a link to a WWW site for the product. For each feature, the final column of the table contains the *probable* effect on accessibility. The categories are: ‘‘None’’ for no affect, ‘‘Less’’ meaning the WWW page is less accessible, and ‘‘Benign’’ which means there is an effect but it is probably not important.

Feature	Frequency	Percent	Accessibility Effect
HTML validation failed(*)	12447	93.50%	None
Some missing `<IMG>` ALT Text	8817	66.23%	Less
Has `<META>` tag	8732	65.56%	None
All `<IMG>` tags have ALT Text	2821	21.19%	None
Mentions Netscape Navigator(13)	1653	12.42%	Less
Uses `<META HTTP-Equiv...>`	1049	7.88%	Benign(+)
Mentions Microsoft Internet Explorer(11)	898	6.75%	Less
Uses JavaScript	889	6.68%	Less
HTML validation succeeded(*)	865	6.50%	None
Has multiple newline types	746	5.60%	Benign
Bad character entities syntax	675	5.07%	Benign
Uses `<FRAME>`s with `<NOFRAMES>`	456	3.42%	None
Non ISO 8879-1986 (Latin-1) characters	274	2.06%	Benign
Uses `<FRAME>`s without `<NOFRAMES>`	267	2.01%	Less
Uses HTTP `Refresh` header	242	1.82%	Less
Uses Java	216	1.62%	Less
Mentions Shockwave(14)	80	0.60%	Less
Uses Visual Basic Script	18	0.14%	Less

Table 9: Features Used in Each WWW Page

Notes:
(*) Duplicated here for use in comparison with other features.
(+) Benign effect here but the `Refresh` header case is covered below.

The table shows that most of the features that make the pages less accessible appear under 13% of the time, **except** for the missing `<IMG> ALT` text which happens in 66.23% of documents. Taking this feature as a requirement for accessibility, and considering the other features, there are several analyses that can be done.

Being pessimistic, mentions of products may actually be requirements implying they are necessary to view the pages, and the use of the scripts and languages are required. In that case, the pages that are accessible are those that have `<IMG> ALT` text present, have none of the ‘‘Less’’ features, and can have any of the ‘‘Benign’’ ones.

> The total of pages that match this feature set is 3567, or 26.78% of all pages.

Alternatively, being more optimistic, the mentions of products are benign simply implying the products would be useful to access the pages, but are not required. The scripts and languages still remain necessary.

A total of 3990 or 29.95% of all pages match this category.

Finally if we are *very* optimistic and assume that the scripts and languages are not required to access the page, we get:

A total of 4195 or 31.49% of all pages match this category.

These analyses give the total figure for accessibility at around 30% of the total number of unique WWW pages, but are likely to be too low for several reasons:

- Surveys show that most people use ``*bleeding-edge*'' browsers that **do** support many of the features above and hence are not affected by the accessibility issues.
- Browsers improve and start implementing more of the newer features, especially those in the emerging standards.

The figures are also likely to be too high because of reasons including the following:

- The use of client-side and server-side image maps is growing and is often seen without alternatives. The typical use is a home page, with a single image map for a large image and no other links (maybe using `<FRAME>`s too).
- Different pages are sent or formatted differently depending on the browser being used. This caused one manufacturer to use the `User-Agent` name of another browser in order that correct WWW pages would be delivered by servers.

Conclusions

A survey of the state of the UK Web was presented and discussion of the accessibility issues for the Web made. Around 30% of unique WWW pages were found to be accessible to all.

Over time, as ongoing W3C HTML standardisation and the implementation of new features continues, accessibility should improve. Unfortunately, there is still the problem that as new features are added to browsers, they may not be compatible with existing software and will consequently reduce accessibility in the push to gain a commercial edge. Standardisation efforts like HTML 3.2[Raggett1997] are very important to standardise the changing state of HTML so browsers can provide access to all wherever possible, and yet allow the Web to move forward.

The full results from the survey will be made available on-line via a link from my home page(0). As a bonus, try to find the extra data that the on-line version of this paper has if you *access* it without graphics.

Thanks

Thanks to Duncan Langford for comments and encouragement.
This paper was made possible by Perl(15) and GNU Emacs(16) in addition to the software already mentioned.

References

[Woodruff1996]
Allison Woodruff, Paul M. Aoki, Eric Brewer, Paul Gauthier and Lawrence A. Rowe, *An Investigation of Documents from the World Wide Web*, Computer Science Division, University of California at Berkeley, Berkeley, CA 94720-1776, USA, in Proceedings 5th International World Wide Web Conference, Paris, France, May 1996, pp 963--979, <URL:http://www5conf.inria.fr/fich_html/papers/P7/Overview.html>.

[Bray1996]
Tim Bray, *Measuring the Web*, Open Text Corporation, Canada, in Proceedings 4th International World Wide Web Conference, Paris, France, May 1996, pp 994--1005, <URL:http://www5conf.inria.fr/fich_html/papers/P9/Overview.html>.

[UKNIC1996]
UK Network Information Centre (UK NIC), *Domain Names within the UK*, December 1996, <URL:http://www.nic.uk/new/domains.html>.

[Mockapetris1987]
P. Mockapetris, RFC 1034 *Domain Names Concepts and Facilities* and RFC 1035 *Domain Names Implementation and Specification*, USC/Information Sciences Institute, November 1987, <URL:ftp://ftp.uu.net/inet/rfc/rfc1034.Z> and <URL:ftp://ftp.uu.net/inet/rfc/rfc1035.Z>.

[Carey1996]
J. Carey, *Rules for the ltd.uk and plc.uk domains*, Nominet UK, 25th September 1996, <URL:http://www.nic.uk/rules/rup1.htm>.

[Rivest1992]
R. Rivest, *The MD5 Message-Digest Algorithm* RFC 1321, MIT Laboratory for Computer Science and RSA Data Security, Inc., April 1992, <URL:ftp://ftp.uu.net/inet/rfc/rfc1321.Z>.

[Raggett1997]
Dave Raggett, World Wide Web Consortium (W3C), *HTML 3.2 Reference Specification*, Recommendation, 14th January 1997, <URL:http://www.w3.org/pub/WWW/TR/REC-html32.html>.

URLS

(0) *Dave Beckett's Home Page* at <URL:http://www.hensa.ac.uk/parallel/www/djb1.html>.

(1) *Inktomi Corporation, Inc.* at <URL:http://www.inktomi.com/>.

(2) *Open Text Corporation* at <URL:http://www.opentext.com/>.

(3) *Nominet UK / UK NIC* at <URL:http://www.nominet.org.uk/> and <URL:http://www.nic.uk/>.

(4) *UKERNA* -- United Kingdom Education and Research Networking Association, a not-for-profit-company, at <URL:http://www.ukerna.uk/>.

(5) *CCTA: Government Information Service* at <URL:http://www.ccta.gov.uk/>.

(6) *Host* by Eric Wassenaar, Nikhef-H, based on BSD Bind code at <URL:ftp://ftp.nikhef.nl/pub/network/host.tar.Z>.

(7) *Libwww-perl* by Gisle Aas and others at <URL:http://www.sn.no/libwww-perl/>.

(8) *World Wide Web Consortium* (W3C) at <URL:http://www.w3.org/>.

(9) *NSGMLS* SGML parser and validator (part of SP SGML parser suite) by Jim Clark at <URL:http://www.jclark.com/sp/nsgmls.htm>.

(10) *SoftQuad Corporation, Inc.* (SQ) at <URL:http://www.sq.com/>.

(11) *Microsoft Internet Explorer* (MSIE) WWW Browser at <URL:http://www.microsoft.com/ie/>.

(12) *Microsoft Free TrueType fonts for use on the Web* at <URL:http://www.microsoft.com/truetype/css/iexplor/free.htm>.

(13) *Netscape Navigator WWW Browser* (aka Mozilla) at <URL:http://home.netscape.com/>.

(14) *Macromedia Shockwave* at <URL:http://www.macromedia.com/>.

(15) *Perl* by Larry Wall et al, at <URL:http://www.perl.org/>.

(16) *The GNU Project* at <URL:http://www.gnu.ai.mit.edu/>.