Appendix A

Internet mirroring robot (Pavuk) configuration

```
% Don't show any messages on the screen
Quiet: false
% Run in background mode
BgMode: false
% Checks if it is running as background job
% if so do not output any messages
CheckIfRunnigAtBackground: false
% Show retrieving progress while running in
% the terminal
ShowProgress: false
% Output debugging information
Debug: true
% Debugging information will contain details
% on URLs that could not be mirrored
DebugLevel: limits
% Number of directory levels to omit in local tree.
BaseLevel: 1
% Assign unique local filename to unique URL
UniqueDocName: true
% replace special characters from URL into local filename
TrChrToChr: "\"\\ ?&*$;:,@=+[](){}<>'`|!%`"" "_"
% Local directory where all retrieved documents are stored
WorkingDir: pavuk_data
```

```
% System information will be printed in this file
LogFile: ../pavuk.log
% Summary of system information will be printed in this file
SLogFile: ../pavuk.slog

StoreDocInfoFiles: true
SeparateInfoDir: ../pavuk_info
DumpFD: -1
DumpUrlFD: -1
SinglePage: false
MaxLevel: 3
MaxDocs: 0
SleepBetween: 1
RandomizeSleepPeriod: false
MaxRetry: 2
MaxRegets: 2
MaxRedirections: 5
CommTimeout: 0.5
RegetRollbackAmount: 0
UseCache: false
UseRobots: true
AllowHTTP: true
AllowSSL: true
AllowFTP: false
AllowFTPS: false
AllowGopher: false
AllowCGI: true
AllowGZEncoding: true
AllowRelocation: false
FTPhtml: false
FTPListCMD: false
AllowFTPRecursion: true
StoreDirIndexFile: true
ForceReget: true
DisallowedSuffixes: .css,.gif,.jpeg,.jpg,.jpe,.png,.mpeg,.mpg,.mp3,.smi,.smil,.qt,
    .mov,.pdf,.eps,.ps,.doc,.rtf,.ppt,.xls,.exe,.avi,.ra,.ram,.rm,.wav,.au
AllowedMIMETypes: text/html,text/html;*
HTTPAuthorizationScheme: Basic
AuthReuseDigestNonce: false
AuthReuseProxyDigestNonce: false
EmailAddress: webmaster@eaccess.rince.ie
```

```
SendFromHeader: false
DefaultMode: normal
GopherViaHTTPProxy: false
FTPViaHTTPProxy: false
FTPDirtyProxy: false
MaxSize: 0
MinSize: 0
HTTPProxyAuth: Basic
ShowDownloadTime: true
RemoveOldDocuments: false
ActiveFTPData: false
NLSMessageCatalogDir: /usr/local/share/locale
Language: en_US
Reschedule: 0
SiteLevel: 0
DontLeaveSite: false
DontLeaveDir: false
LeaveSiteEnterDirectory: false
PreserveTime: true
PreservePermisions: false
PreserveAbsoluteSymlinks: false
LeaveLevel: 0
CookieSend: true
CookieRecv: true
CookieUpdate: false
CookieCheckDomain: true
CookiesMax: 0
DisableHTMLTag: IMG,SRC;INPUT,SRC;BODY,BACKGROUND
IndexName: _auto_index.html
CheckSize: true
BgMode: false
MaxRate: 0.000
MinRate: 0.000
ReadBufferSize: 32
FileSizeQuota: 0
TransferQuota: 225
FSQuota: 0
EnableJS: false
AllLinksToLocal: false
SelectedLinksToLocal: false
AllLinksToRemote: false
```

```
AutoReferer: false
UrlSchedulingStrategy: level
RemoveAdvertisement: false
SendIfRange: true
SchedulingCommand: at -f %f %t %d.%m.%Y
UniqueLogName: false
SSLVersion: ssl23
UniqueSSLID: false
UseHTTP11: true
DeleteAfterTransfer: false
RetrieveSymlinks: false
LimitInlineObjects: true
FixWuFTPDBrokenLISTcmd: false
PostUpdate: false
HackAddIndex: false
SkipURLREMatchPattern: ^http(s?)://([^/]+)/~
```

Appendix B

Case-variations of Formal Public Identifiers

- HTML 4.01 Strict

    - `"-//W3C//DTD HTML 4.01//EN"` (typical usage)

    - `"-//w3c//dtd html 4.01//en"`

- HTML 4.01 Transitional

    - `"-//W3C//DTD HTML 4.01 Transitional//EN"` (typical usage)

    - `"-//w3c//dtd html 4.01 transitional//en"`

    - `"-//W3C//DTD HTML 4.01 transitional//EN"`

- HTML 4.01 Frameset

    - `"-//W3C//DTD HTML 4.01 Frameset//EN"` (typical usage)

    - `"-//w3c//dtd html 4.01 frameset//en"`

    - `"-//W3C//DTD HTML 4.01 frameset//EN"`

- HTML 4.0 Strict

  - `"-//W3C//DTD HTML 4.0//EN"` (typical usage)

  - `"-//w3c//dtd html 4.0//en"`

- HTML 4.0 Transitional

  - `"-//W3C//DTD HTML 4.0 Transitional//EN"` (typical usage)

  - `"-//w3c//dtd html 4.0 transitional//en"`

  - `"-//W3C//DTD HTML 4.0 transitional//EN"`

- HTML 4.0 Frameset

  - `"-//W3C//DTD HTML 4.0 Frameset//EN"` (typical usage)

  - `"-//w3c//dtd html 4.0 frameset//en"`

  - `"-//W3C//DTD HTML 4.0 frameset//EN"`

- HTML 3.2

  - `"-//W3C//DTD HTML 3.2 Final//EN"` (typical usage)

  - `"-//W3C//DTD HTML 3.2 Draft//EN"` (typical usage)

  - `"-//W3C//DTD HTML 3.2//EN"` (typical usage)

  - `"-//W3C//DTD HTML 3.2 final//EN"`

  - `"-//w3c//dtd html 3.2 final//en"`

  - `"-//W3C//DTD HTML 3.2 draft//EN"`

  - `"-//w3c//dtd html 3.2 draft//en"`

  - `"-//w3c//dtd html 3.2//en"`

- HTML (2.0)

  - "-//IETF//DTD HTML//EN" (typical usage)

  - "-//IETF//DTD HTML 2.0//EN" (typical usage)

  - "-//IETF//DTD HTML Level 2//EN" (typical usage)

  - "-//IETF//DTD HTML 2.0 Level 2//EN" (typical usage)

  - "-//IETF//DTD HTML Level 1//EN" (typical usage)

  - "-//IETF//DTD HTML 2.0 Level 1//EN" (typical usage)

  - "-//IETF//DTD HTML Strict//EN" (typical usage)

  - "-//IETF//DTD HTML 2.0 Strict//EN" (typical usage)

  - "-//IETF//DTD HTML Strict Level 2//EN" (typical usage)

  - "-//IETF//DTD HTML 2.0 Strict Level 2//EN" (typical usage)

  - "-//IETF//DTD HTML Strict Level 1//EN" (typical usage)

  - "-//IETF//DTD HTML 2.0 Strict Level 1//EN" (typical usage)

  - "-//ietf//dtd html//en"

  - "-//ietf//dtd html 2.0//en"

  - "-//ietf//dtd html level 2//en"

  - "-//ietf//dtd html 2.0 level 2//en"

  - "-//ietf//dtd html level 1//en"

  - "-//ietf//dtd html 2.0 level 1//en"

- "-//ietf//dtd html strict//en"

- "-//ietf//dtd html 2.0 strict//en"

- "-//ietf//dtd html strict level 2//en"

- "-//ietf//dtd html 2.0 strict level 2//en"

- "-//ietf//dtd html strict level 1//en"

- "-//ietf//dtd html 2.0 strict level 1//en"

- XHTML

  - "-//W3C//DTD XHTML 1.0 Strict//EN" (typical usage)

  - "-//W3C//DTD XHTML 1.0 Transitional//EN" (typical usage)

  - "-//W3C//DTD XHTML 1.0 Frameset//EN" (typical usage)

  - "-//W3C//DTD XHTML 1.1//EN" (typical usage)

Appendix C

Results of surveys conducted on the web samples of Irish, UK, French and German web sites (May 2003)

C.1   Web content mirroring process results

Table C.1: **Results of the web content sampling process**

	IE	UK	FR	DE	Overall
Data amount retrieved	29 MB	552 MB	167 MB	377 MB	1.1 GB
Web pages considered for mirroring	3,319	67,598	22,319	58,278	151,514
Web pages correctly retrieved (size>0 and no retrieval errors)	2,817 (84.9%)	59,248 (87.6%)	19,727 (88.4%)	52,825 (90.6%)	134,617(88.8%)
Web pages in web sites qualified for conformance investigations	2,288 (68.9%)	45,857 (67.8%)	16,312 (73.1%)	38,062 (65.3%)	102,519(67.7%)
Web sites qualified for conformance investigations	123 (45.2%)	2,380 (41.7%)	747 (48.3%)	1,636 (38.5%)	4,886 (41.5%)

Table C.2: **Details regarding the web sites that didn't qualify for investigations**

	IE	UK	FR	DE	Overall
Web sites not qualified for investigations	149	3,322	798	2,615	6,884
No data retrieved	9 (6%)	270 (8.1%)	106 (13.3%)	147 (5.6%)	532 (7.7%)
Only the main page retrieved (size>0)	91 (61.1%)	1,669 (50.2%)	328 (41.1%)	1,159 (44.3%)	3,247 (47.2%)
Less than 3 pages (2) and at least 100 KB retrieved	1 (0.7%)	4 (0.1%)	4 (0.5%)	1 (0%)	10 (0.1%)
Less than 100 KB and at least 3 pages retrieved	46 (30.9%)	1,211 (36.5%)	301 (37.7%)	1,177 (45%)	2,735 (39.7%)
Less than 3 pages (2) and less that 100 KB (but more than 0) retrieved	2 (1.3%)	168 (5.1%)	59 (7.4%)	130 (5.0%)	359 (5.2%)

C.2 WCAG 1.0 conformance investigation results

Table C.3: **Details regarding Bobby executions**

	IE		UK		FR		DE		Overall	
Web sites considered by Bobby	123		2,380		747		1,636		4,886	
Bobby abnormally terminated	3	(2.4%)	137	(5.8%)	63	(8.4%)	101	(6.2%)	304	(6.2%)
Bobby "locked up"	6	(4.9%)	92	(3.9%)	45	(6%)	90	(5.5%)	233	(4.8%)
Bobby completed successfully	114	(92.7%)	2,151	(90.4%)	639	(85.5%)	1,444	(88.3%)	4,348	(89%)

Table C.4: **Bobby Priority 1 diagnostics triggered during the investigation**

ID		IE		UK		FR		DE		Overall	
g9	Provide alternative text for all images	106	(93%)	1,989	(92.5%)	625	(97.8%)	1,336	(92.5%)	4,056	(93.3%)
g39	Give each frame a title	25	(21.9%)	579	(26.9%)	263	(41.2%)	735	(50.9%)	1,602	(36.8%)
g38	Each FRAME must reference an HTML file	24	(21.1%)	548	(25.5%)	251	(39.3%)	726	(50.3%)	1,549	(35.6%)
g240	Provide alternative text for all image map hot-spots	25	(21.9%)	429	(19.9%)	241	(37.7%)	265	(18.4%)	960	(22.1%)
g10	Provide alternative text for all image-type buttons in forms	9	(7.9%)	183	(8.5%)	82	(12.8%)	126	(8.7%)	400	(9.2%)
g21	Provide alternative text for each APPLET	10	(8.8%)	158	(7.3%)	45	(7%)	110	(7.6%)	323	(7.4%)
g20	Provide alternative content for each OBJECT	0	(0%)	5	(0.2%)	3	(0.5%)	3	(0.2%)	11	(0.3%)

Figure C.1: **Bobby Priority 1 diagnostics**
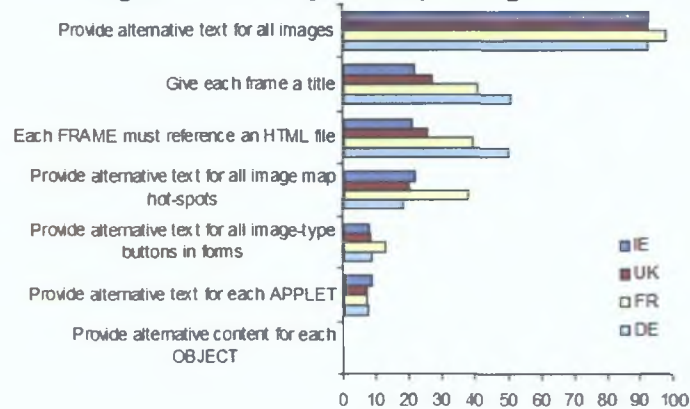
Table C.5: **Density of Bobby Priority 1 diagnostics**

	IE		UK		FR		DE		Overall	
Web sites with Bobby completed successfully	114		2,151		639		1,444		4,348	
Web sites that triggered 0 distinct diagnostics	7	(6.1%)	119	(5.5%)	9	(1.4%)	66	(4.6%)	201	(4.6%)
Web sites that triggered 1 distinct diagnostic	55	(48.2%)	1,021	(47.5%)	199	(31.1%)	461	(31.9%)	1,736	(39.9%)
Web sites that triggered 2 distinct diagnostic	25	(21.9%)	415	(19.3%)	155	(24.3%)	200	(13.9%)	795	(18.3%)
Web sites that triggered 3 distinct diagnostics	16	(14%)	377	(17.5%)	135	(21.1%)	463	(32.1%)	991	(22.8%)
Web sites that triggered 4 distinct diagnostics	9	(7.9%)	186	(8.6%)	110	(17.2%)	220	(15.2%)	525	(12.1%)
Web sites that triggered 5 distinct diagnostics	2	(1.8%)	33	(1.5%)	30	(4.7%)	33	(2.3%)	98	(2.3%)
Web sites that triggered 6 distinct diagnostics	0	(0%)	0	(0%)	1	(0.2%)	1	(0.1%)	2	(0%)

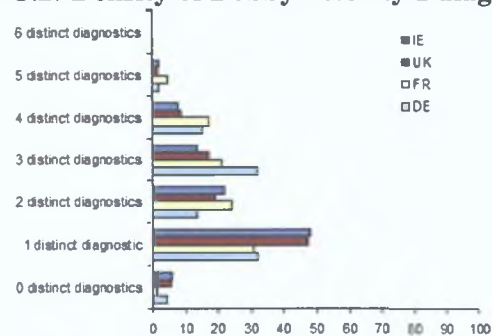Figure C.2: **Density of Bobby Priority 1 diagnostics**

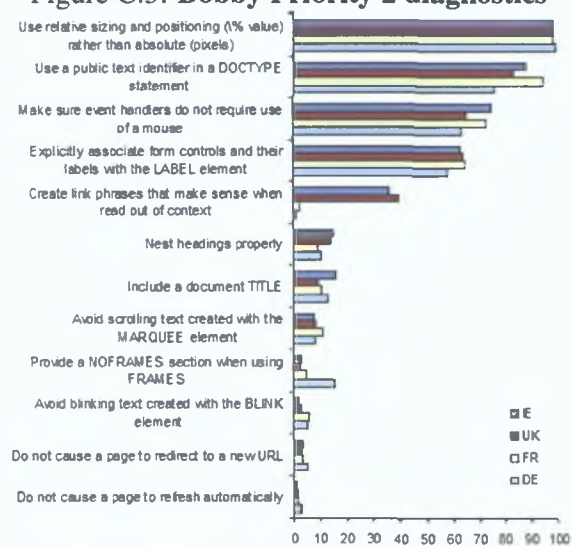Figure C.3: **Bobby Priority 2 diagnostics**
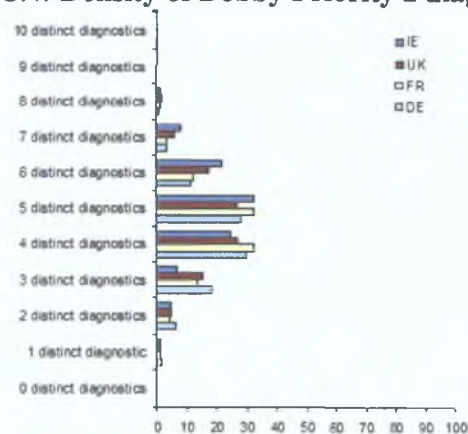

Figure C.4: **Density of Bobby Priority 2 diagnostics**

6

Table C.6: **Bobby Priority 2 diagnostics triggered during the investigation**

ID		IE		UK		FR		DE		Overall	
g104	Use relative sizing and positioning (% value) rather than absolute (pixels)	112	(98.2%)	2,113	(98.2%)	627	(98.1%)	1,429	(99.0%)	4,281	(98.5%)
g271	Use a public text identifier in a DOCTYPE statement	100	(87.7%)	1,775	(82.5%)	602	(94.2%)	1,093	(75.7%)	3,570	(82.1%)
g269	Make sure event handlers do not require use of a mouse	85	(74.6%)	1,404	(65.3%)	464	(72.6%)	915	(63.4%)	2,868	(66.0%)
g41	Explicitly associate form controls and their labels with the LABEL element	71	(62.3%)	1,368	(63.6%)	412	(64.5%)	836	(57.9%)	2,687	(61.8%)
g34	Create link phrases that make sense when read out of context	41	(36.0%)	840	(39.1%)	15	(2.3%)	18	(1.2%)	914	(21.0%)
g2	Nest headings properly	17	(14.9%)	298	(13.9%)	56	(8.8%)	150	(10.4%)	521	(12.0%)
g273	Include a document TITLE	18	(15.8%)	191	(8.9%)	68	(10.6%)	182	(12.6%)	459	(10.6%)
g5	Avoid scrolling text created with the MARQUEE element	9	(7.9%)	178	(8.3%)	69	(10.8%)	119	(8.2%)	375	(8.6%)
g37	Provide a NOFRAMES section when using FRAMES	3	(2.6%)	46	(2.1%)	30	(4.7%)	219	(15.2%)	298	(6.9%)
g41	Avoid blinking text created with the BLINK element	2	(1.8%)	55	(2.6%)	37	(5.8%)	80	(5.5%)	174	(4.0%)
g254	Do not cause a page to redirect to a new URL	4	(3.5%)	62	(2.9%)	19	(3%)	78	(5.4%)	163	(3.7%)
g33	Do not cause a page to refresh automatically	1	(0.9%)	35	(1.6%)	9	(1.4%)	37	(2.6%)	82	(1.9%)

Table C.7: **Density of Bobby Priority 2 diagnostics**

	IE		UK		FR		DE		Overall	
Web sites with Bobby completed successfully	114		2,151		639		1,444		4,348	
Web sites that triggered 0 distinct diagnostics	0	(0%)	4	(0.2%)	0	(0%)	0	(0%)	4	(0.1%)
Web sites that triggered 1 distinct diagnostic	1	(0.9%)	22	(1.0%)	7	(1.1%)	25	(1.7%)	55	(1.3%)
Web sites that triggered 2 distinct diagnostics	5	(4.4%)	100	(4.6%)	26	(4.1%)	92	(6.4%)	223	(5.1%)
Web sites that triggered 3 distinct diagnostics	8	(7%)	329	(15.3%)	87	(13.6%)	265	(18.4%)	689	(15.8%)
Web sites that triggered 4 distinct diagnostics	28	(24.6%)	576	(26.8%)	207	(32.4%)	431	(29.8%)	1,242	(28.6%)
Web sites that triggered 5 distinct diagnostics	37	(32.5%)	577	(26.8%)	207	(32.4%)	406	(28.1%)	1,227	(28.2%)
Web sites that triggered 6 distinct diagnostics	25	(21.9%)	374	(17.4%)	80	(12.5%)	167	(11.6%)	646	(14.9%)
Web sites that triggered 7 distinct diagnostics	9	(7.9%)	125	(5.8%)	19	(3%)	47	(3.3%)	200	(4.6%)
Web sites that triggered 8 distinct diagnostics	1	(0.9%)	34	(1.6%)	6	(0.9%)	9	(0.6%)	50	(1.1%)
Web sites that triggered 9 distinct diagnostics	0	(0%)	9	(0.4%)	0	(0%)	2	(0.1%)	11	(0.3%)
Web sites that triggered 10 distinct diagnostics	0	(0%)	1	(0%)	0	(0%)	0	(0%)	1	(0%)

8

C.3 HTML technical standards conformance investigation results

Table C.8: **Details regarding OpenSP executions**

	IE		UK		FR		DE		Overall	
Web sites processed by OpenSP	123		2,380		747		1,636		4,886	
Web sites where all the web pages had usable document type information	2	(1.6%)	35	(1.5%)	1	(0.1%)	52	(3.2%)	90	(1.8%)
Web sites considered for HTML defects analysis	36	(29.3%)	663	(27.9%)	140	(18.7%)	603	(36.9%)	1,442	(29.5%)

Table C.9: **Document Type Declaration in web pages considered by OpenSP**

	IE		UK		FR		DE		Overall	
Web pages processed by OpenSP	2,288		45,857		16,312		38,062		102,519	
No document type information	1,732	(75.7%)	34,564	(75.4%)	13,837	(84.8%)	24,258	(63.7%)	74,391	(72.6%)
Incorrect document type information	42	(1.8%)	965	(2.1%)	427	(2.6%)	2,692	(7.1%)	4,126	(4%)
Usable document type information	465	(20.3%)	9,562	(20.9%)	1,788	(11%)	10,681	(28.1%)	22,496	(21.9%)
Considered for HTML defects analysis (in web sites with at least 3 web pages and 100 KB of data with usable document type information)	454	(19.8%)	9,291	(20.3%)	1,672	(10.3%)	10,431	(27.4%)	21,848	(21.3%)

9

Table C.10: **The 10 most common OpenSP diagnostics triggered in the investigation**

ID		IE		UK		FR		DE		Overall	
108	Undefined attribute	33	(91.7%)	600	(90.5%)	116	(82.9%)	549	(91%)	1,298	(90%)
64	Element not allowed by document type	29	(80.6%)	509	(76.8%)	101	(72.1%)	486	(80.6%)	1,125	(78%)
127	Required attribute not specified	19	(52.8%)	496	(74.8%)	102	(72.9%)	463	(76.8%)	1,080	(74.9%)
79	End tag for not opened element	27	(75%)	487	(73.5%)	84	(60%)	416	(69%)	1,014	(70.3%)
68	Missing required end tag (element DTD)	20	(55.6%)	351	(52.9%)	65	(46.4%)	327	(54.2%)	763	(52.9%)
25	General entity not defined and no default entity	17	(47.2%)	326	(49.2%)	57	(40.7%)	269	(44.6%)	669	(46.4%)
139	Non SGML character number	20	(55.6%)	270	(40.7%)	66	(47.1%)	268	(44.4%)	624	(43.3%)
73	End tag for element not finished	11	(30.6%)	254	(38.3%)	52	(37.1%)	276	(45.8%)	593	(41.1%)
131	Attribute value not allowed	10	(27.8%)	261	(39.4%)	59	(42.1%)	217	(36%)	547	(37.9%)
82	Attribute value must be literal	16	(44.4%)	251	(37.9%)	54	(38.6%)	212	(35.2%)	533	(37%)

Figure C.5: **The 10 most common OpenSP diagnostics**
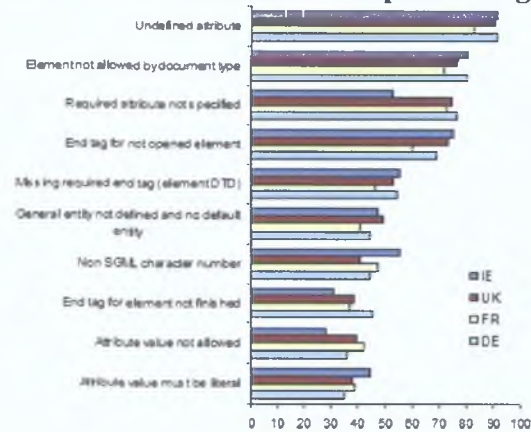
10

Table C.11: **Density of OpenSP diagnostics**

	IE		UK		FR		DE		Overall	
Web sites considered for HTML defects analysis	36		663		140		603		1,442	
Web sites that triggered 0 distinct diagnostics	1	(2.8%)	29	(4.4%)	16	(11.4%)	29	(4.8%)	75	(5.2%)
Web sites that triggered 1 distinct diagnostic	1	(2.8%)	8	(1.2%)	2	(1.4%)	9	(1.5%)	20	(1.4%)
Web sites that triggered 2 distinct diagnostics	2	(5.6%)	20	(3%)	3	(2.1%)	6	(1%)	31	(2.1%)
Web sites that triggered 3 distinct diagnostics	1	(2.8%)	19	(2.9%)	3	(2.1%)	19	(3.2%)	42	(2.9%)
Web sites that triggered 4 distinct diagnostics	2	(5.6%)	37	(5.6%)	11	(7.9%)	21	(3.5%)	71	(4.9%)
Web sites that triggered 5 distinct diagnostics	1	(2.8%)	46	(6.9%)	6	(4.3%)	29	(4.8%)	82	(5.7%)
Web sites that triggered 6 distinct diagnostics	1	(2.8%)	40	(6%)	10	(7.1%)	44	(7.3%)	95	(6.6%)
Web sites that triggered 7 distinct diagnostics	3	(8.3%)	58	(8.7%)	13	(9.3%)	56	(9.3%)	130	(9%)
Web sites that triggered 8 distinct diagnostics	3	(8.3%)	51	(7.7%)	9	(6.4%)	63	(10.4%)	126	(8.7%)
Web sites that triggered 9 distinct diagnostics	1	(2.8%)	56	(8.4%)	7	(5%)	52	(8.6%)	116	(8%)
Web sites that triggered 10 distinct diagnostics	5	(13.9%)	54	(8.1%)	11	(7.9%)	49	(8.1%)	119	(8.3%)
Web sites that triggered 11 distinct diagnostics	5	(13.9%)	60	(9%)	10	(7.1%)	48	(8%)	123	(8.5%)
Web sites that triggered 12 distinct diagnostics	4	(11.1%)	52	(7.8%)	11	(7.9%)	39	(6.5%)	106	(7.4%)
Web sites that triggered 13 distinct diagnostics	2	(5.6%)	37	(5.6%)	10	(7.1%)	43	(7.1%)	92	(6.4%)
Web sites that triggered 14 distinct diagnostics	2	(5.6%)	38	(5.7%)	3	(2.1%)	20	(3.3%)	63	(4.4%)
Web sites that triggered 15 distinct diagnostics	0	(0%)	19	(2.9%)	3	(2.1%)	22	(3.6%)	44	(3.1%)
Web sites that triggered 16 distinct diagnostics	1	(2.8%)	20	(3%)	6	(4.3%)	22	(3.6%)	49	(3.4%)

11

Table C.11: **Density of OpenSP diagnostics**

	IE		UK		FR		DE		Overall	
Web sites that triggered 17 distinct diagnostics	0	(0%)	8	(1.2%)	3	(2.1%)	7	(1.2%)	18	(1.2%)
Web sites that triggered 18 distinct diagnostics	0	(0%)	7	(1.1%)	2	(1.4%)	10	(1.7%)	19	(1.3%)
Web sites that triggered 19 distinct diagnostics	0	(0%)	2	(0.3%)	0	(0%)	8	(1.3%)	10	(0.7%)
Web sites that triggered 20 distinct diagnostics	0	(0%)	1	(0.2%)	1	(0.7%)	5	(0.8%)	7	(0.5%)
Web sites that triggered 21 distinct diagnostics	0	(0%)	0	(0%)	0	(0%)	1	(0.2%)	1	(0.1%)
Web sites that triggered 22 distinct diagnostics	1	(2.8%)	1	(0.2%)	0	(0%)	0	(0%)	2	(0.1%)
Web sites that triggered 23 distinct diagnostics	0	(0%)	0	(0%)	0	(0%)	1	(0.2%)	1	(0.1%)

Figure C.6: **Density of OpenSP diagnostics**

Appendix D

Results of surveys conducted on the web samples of Irish, UK, French and German web sites (December 2003)

D.1   Web content mirroring process results

Table D.1: **Results of the web content sampling process**

	IE	UK	FR	DE	Overall
Data amount retrieved	26 MB	527 MB	162 MB	378 MB	1.1 GB
Web pages considered for mirroring	3,043	62,491	20,406	57,598	143,538
Web pages correctly re-trieved (size>0 and no re-trieval errors)	2,541 (83.5%)	54,135 (86.6%)	17,956 (88%)	51,473 (89.4%)	126,105(87.9%)
Web pages in web sites qualified for conformance investigations	2,064 (67.8%)	42,189 (67.5%)	14,874 (72.9%)	37,974 (65.9%)	97,101 (67.6%)
Web sites qualified for conformance investiga-tions	111 (40.8%)	2,269 (39.8%)	722 (46.7%)	1,649 (38.8%)	4,751 (40.4%)

Table D.2: **Details regarding the web sites that didn't qualify for investiga-tions**

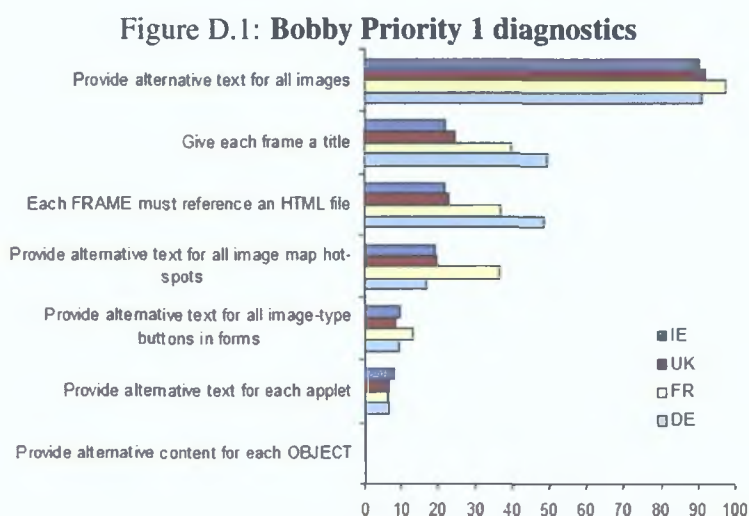|	IE	UK	FR	DE	Overall
Web sites not qualified for investigations	161	3,434	823	2,602	7,020
No data retrieved	15 (9.3%)	417 (12.1%)	124 (15.1%)	178 (6.8%)	734 (10.5%)
Only the main page re-trieved (size>0)	102 (63.4%)	1,770 (51.5%)	358 (43.5%)	1,209 (46.5%)	3,439 (49%)
Less than 3 pages (2) and at least 100 KB retrieved	1 (0.6%)	3 (0.1%)	3 (0.4%)	4 (0.2%)	11 (0.2%)
Less than 100 KB and at least 3 pages retrieved	40 (24.8%)	1,084 (31.6%)	285 (34.6%)	1,082 (41.6%)	2,491 (35.5%)
Less than 3 pages (2) and less that 100 KB (but more than 0) retrieved	3 (1.9%)	158 (4.6%)	53 (6.4%)	128 (4.9%)	342 (4.9%)

D.2   WCAG 1.0 conformance investigation results

Table D.3: **Details regarding Bobby executions**

	IE		UK		FR		DE		Overall	
Web sites considered by Bobby	111		2,269		722		1,649		4,751	
Bobby abnormally terminated	3	(2.7%)	133	(5.9%)	51	(7.1%)	102	(6.2%)	289	(6.1%)
Bobby "locked up"	4	(3.6%)	76	(3.3%)	28	(3.9%)	75	(4.5%)	183	(3.9%)
Bobby completed successfully	104	(93.7%)	2,059	(90.7%)	643	(89.1%)	1,471	(89.2%)	4,277	(90%)

Table D.4: **Bobby Priority 1 diagnostics triggered during the investigation**

ID		IE		UK		FR		DE		Overall	
g9	Provide alternative text for all images	94	(90.4%)	1,896	(92.1%)	627	(97.5%)	1,343	(91.3%)	3,960	(92.6%)
g39	Give each frame a title	23	(22.1%)	511	(24.8%)	257	(40%)	730	(49.6%)	1,521	(35.6%)
g38	Each FRAME must reference an HTML file	23	(22.1%)	477	(23.2%)	238	(37%)	715	(48.6%)	1,453	(34%)
g240	Provide alternative text for all image map hot-spots	20	(19.2%)	401	(19.5%)	235	(36.5%)	243	(16.5%)	899	(21%)
g10	Provide alternative text for all image-type buttons in forms	10	(9.6%)	170	(8.3%)	85	(13.2%)	135	(9.2%)	400	(9.4%)
g21	Provide alternative text for each APPLET	8	(7.7%)	134	(6.5%)	38	(5.9%)	92	(6.3%)	272	(6.4%)
g20	Provide alternative content for each OBJECT	0	(0%)	3	(0.1%)	3	(0.5%)	3	(0.2%)	9	(0.2%)

Figure D.1: **Bobby Priority 1 diagnostics**

Table D.5: **Density of Bobby Priority 1 diagnostics**

	IE		UK		FR		DE		Overall	
Web sites with Bobby completed successfully	104		2,059		643		1,471		4,277	
Web sites that triggered 0 distinct diagnostics	7	(6.7%)	123	(6%)	10	(1.6%)	69	(4.7%)	209	(4.9%)
Web sites that triggered 1 distinct diagnostic	50	(48.1%)	1,008	(49%)	204	(31.7%)	496	(33.7%)	1,758	(41.1%)
Web sites that triggered 2 distinct diagnostic	25	(24%)	400	(19.4%)	166	(25.8%)	207	(14.1%)	798	(18.7%)
Web sites that triggered 3 distinct diagnostics	12	(11.5%)	352	(17.1%)	131	(20.4%)	474	(32.2%)	969	(22.7%)
Web sites that triggered 4 distinct diagnostics	8	(7.7%)	152	(7.4%)	107	(16.6%)	197	(13.4%)	464	(10.8%)
Web sites that triggered 5 distinct diagnostics	2	(1.9%)	24	(1.2%)	24	(3.7%)	27	(1.8%)	77	(1.8%)
Web sites that triggered 6 distinct diagnostics	0	(0%)	0	(0%)	1	(0.2%)	1	(0.1%)	2	(0%)

4

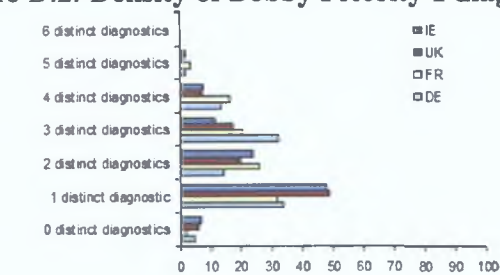## Figure D.2: **Density of Bobby Priority 1 diagnostics**
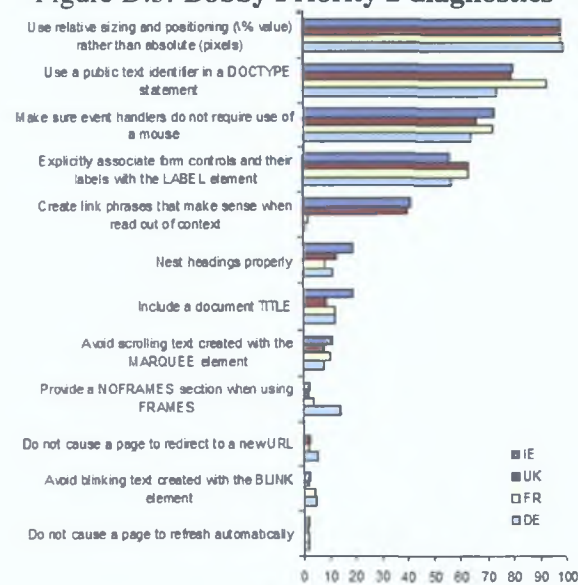

Figure D.3: **Bobby Priority 2 diagnostics**
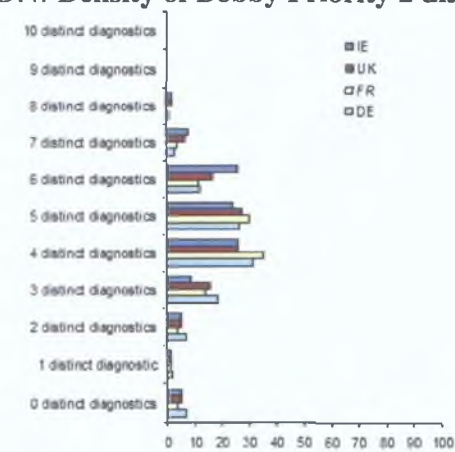

Figure D.4: **Density of Bobby Priority 2 diagnostics**

5

Table D.6: **Bobby Priority 2 diagnostics triggered during the investigation**

ID		IE		UK		FR		DE		Overall	
g104	Use relative sizing and positioning (% value) rather than absolute (pixels)	102	(98.1%)	2,020	(98.1%)	630	(98%)	1,454	(98.8%)	4,206	(98.3%)
g271	Use a public text identifier in a DOCTYPE statement	83	(79.8%)	1,631	(79.2%)	596	(92.7%)	1,085	(73.8%)	3,395	(79.4%)
g269	Make sure event handlers do not require use of a mouse	76	(73.1%)	1,356	(65.9%)	464	(72.2%)	948	(64.4%)	2,844	(66.5%)
g41	Explicitly associate form controls and their labels with the LABEL element	58	(55.8%)	1,303	(63.3%)	407	(63.3%)	836	(56.8%)	2,604	(60.9%)
g34	Create link phrases that make sense when read out of context	43	(41.3%)	822	(39.9%)	15	(2.3%)	19	(1.3%)	899	(21%)
g2	Nest headings properly	20	(19.2%)	263	(12.8%)	55	(8.6%)	167	(11.4%)	505	(11.8%)
g273	Include a document TITLE	20	(19.2%)	185	(9%)	77	(12%)	177	(12%)	459	(10.7%)
g5	Avoid scrolling text created with the MARQUEE element	12	(11.5%)	158	(7.7%)	65	(10.1%)	114	(7.7%)	349	(8.2%)
g37	Provide a NOFRAMES section when using FRAMES	3	(2.9%)	48	(2.3%)	25	(3.9%)	207	(14.1%)	283	(6.6%)
g41	Avoid blinking text created with the BLINK element	3	(2.9%)	43	(2.1%)	29	(4.5%)	73	(5%)	148	(3.5%)
g254	Do not cause a page to redirect to a new URL	0	(0%)	53	(2.6%)	16	(2.5%)	79	(5.4%)	148	(3.5%)
g33	Do not cause a page to refresh automatically	2	(1.9%)	29	(1.4%)	13	(2%)	32	(2.2%)	76	(1.8%)

Table D.7: **Density of Bobby Priority 2 diagnostics**

	IE		UK		FR		DE		Overall	
Web sites for which Bobby completed successfully	104		2,059		643		1,471		4,277	
Web sites that triggered 0 distinct diagnostics	0	(0%)	4	(0.2%)	1	(0.2%)	0	(0%)	5	(0.1%)
Web sites that triggered 1 distinct diagnostic	1	(1%)	24	(1.2%)	8	(1.2%)	25	(1.7%)	58	(1.4%)
Web sites that triggered 2 distinct diagnostics	5	(4.8%)	99	(4.8%)	23	(3.6%)	99	(6.7%)	226	(5.3%)
Web sites that triggered 3 distinct diagnostics	9	(8.7%)	328	(15.9%)	89	(13.8%)	273	(18.6%)	699	(16.3%)
Web sites that triggered 4 distinct diagnostics	27	(26%)	535	(26%)	226	(35.1%)	459	(31.2%)	1,247	(29.2%)
Web sites that triggered 5 distinct diagnostics	25	(24%)	562	(27.3%)	194	(30.2%)	387	(26.3%)	1,168	(27.3%)
Web sites that triggered 6 distinct diagnostics	27	(26%)	339	(16.5%)	74	(11.5%)	176	(12%)	616	(14.4%)
Web sites that triggered 7 distinct diagnostics	8	(7.7%)	134	(6.5%)	25	(3.9%)	41	(2.8%)	208	(4.9%)
Web sites that triggered 8 distinct diagnostics	2	(1.9%)	28	(1.4%)	3	(0.5%)	10	(0.7%)	43	(1%)
Web sites that triggered 9 distinct diagnostics	0	(0%)	5	(0.2%)	0	(0%)	1	(0.1%)	6	(0.1%)
Web sites that triggered 10 distinct diagnostics	0	(0%)	1	(0%)	0	(0%)	0	(0%)	1	(0%)

D.3 HTML technical standards conformance investigation results

Table D.8: **Details regarding OpenSP executions**

	IE		UK		FR		DE		Overall	
Web sites processed by OpenSP	111		2,269		722		1,649		4,751	
Web sites where all the web pages had usable document type information	2	(1.8%)	50	(2.2%)	2	(0.3%)	52	(3.2%)	106	(2.2%)
Web sites considered for HTML defects analysis	39	(35.1%)	720	(31.7%)	142	(19.7%)	647	(39.2%)	1,548	(32.6%)

Table D.9: **Document Type Declaration in web pages considered by OpenSP**

	IE		UK		FR		DE		Overall	
Web pages processed by OpenSP	2,064		42,189		14,874		37,974		97,101	
No document type information	1,435	(69.5%)	30,528	(72.4%)	12,578	(84.6%)	23,422	(61.7%)	67,963	(70%)
Incorrect document type information	78	(3.8%)	873	(2.1%)	457	(3.1%)	2,411	(6.3%)	3,819	(3.9%)
Usable document type information	509	(24.7%)	10,192	(24.2%)	1,758	(11.8%)	11,726	(30.9%)	24,185	(24.9%)
Considered for HTML defects analysis (in web sites with at least 3 web pages and 100 KB of data with usable document type information)	500	(24.2%)	9,916	(23.5%)	1,646	(11.1%)	11,474	(30.2%)	23,536	(24.2%)

8

Table D.10: **The 10 most common OpenSP diagnostics triggered in the investigation**

ID		IE		UK		FR		DE		Overall	
108	Undefined attribute	36	(92.3%)	631	(87.6%)	117	(82.4%)	588	(90.9%)	1,372	(88.6%)
127	Required attribute not specified	28	(71.8%)	560	(77.8%)	104	(73.2%)	515	(79.6%)	1,207	(78%)
64	Element not allowed by document type	31	(79.5%)	520	(72.2%)	99	(69.7%)	511	(79%)	1,161	(75%)
79	End tag for not opened element	28	(71.8%)	506	(70.3%)	86	(60.6%)	446	(68.9%)	1,066	(68.9%)
68	Missing required end tag (element DTD)	22	(56.4%)	342	(47.5%)	53	(37.3%)	344	(53.2%)	761	(49.2%)
25	General entity not defined and no default entity	20	(51.3%)	349	(48.5%)	63	(44.4%)	287	(44.4%)	719	(46.4%)
139	Non SGML character number	24	(61.5%)	285	(39.6%)	64	(45.1%)	313	(48.4%)	686	(44.3%)
73	End tag for element not finished	16	(41%)	270	(37.5%)	52	(36.6%)	294	(45.4%)	632	(40.8%)
131	Attribute value not allowed	15	(38.5%)	282	(39.2%)	59	(41.5%)	230	(35.5%)	586	(37.9%)
82	Attribute value must be literal	18	(46.2%)	261	(36.3%)	51	(35.9%)	222	(34.3%)	552	(35.7%)

Figure D.5: **The 10 most common OpenSP diagnostics**
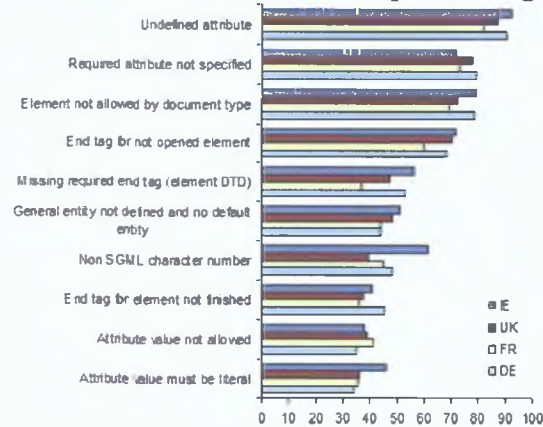
9

<div style="text-align:center">Table D.11: **Density of OpenSP diagnostics**</div>

	IE		UK		FR		DE		Overall	
Web sites considered for further analysis	39		720		142		647		1,548	
Web sites that triggered 0 distinct diagnostics	1	(2.6%)	37	(5.1%)	17	(12%)	26	(4%)	81	(5.2%)
Web sites that triggered 1 distinct diagnostic	1	(2.6%)	8	(1.1%)	3	(2.1%)	6	(0.9%)	18	(1.2%)
Web sites that triggered 2 distinct diagnostics	1	(2.6%)	17	(2.4%)	1	(0.7%)	10	(1.5%)	29	(1.9%)
Web sites that triggered 3 distinct diagnostics	1	(2.6%)	29	(4%)	4	(2.8%)	30	(4.6%)	64	(4.1%)
Web sites that triggered 4 distinct diagnostics	0	(0%)	32	(4.4%)	12	(8.5%)	28	(4.3%)	72	(4.7%)
Web sites that triggered 5 distinct diagnostics	3	(7.7%)	48	(6.7%)	6	(4.2%)	29	(4.5%)	86	(5.6%)
Web sites that triggered 6 distinct diagnostics	1	(2.6%)	55	(7.6%)	13	(9.2%)	47	(7.3%)	116	(7.5%)
Web sites that triggered 7 distinct diagnostics	3	(7.7%)	67	(9.3%)	8	(5.6%)	53	(8.2%)	131	(8.5%)
Web sites that triggered 8 distinct diagnostics	1	(2.6%)	59	(8.2%)	16	(11.3%)	60	(9.3%)	136	(8.8%)
Web sites that triggered 9 distinct diagnostics	2	(5.1%)	60	(8.3%)	7	(4.9%)	59	(9.1%)	128	(8.3%)
Web sites that triggered 10 distinct diagnostics	6	(15.4%)	63	(8.8%)	10	(7%)	63	(9.7%)	142	(9.2%)
Web sites that triggered 11 distinct diagnostics	6	(15.4%)	55	(7.6%)	10	(7%)	53	(8.2%)	124	(8%)
Web sites that triggered 12 distinct diagnostics	3	(7.7%)	54	(7.5%)	7	(4.9%)	33	(5.1%)	97	(6.3%)
Web sites that triggered 13 distinct diagnostics	1	(2.6%)	44	(6.1%)	11	(7.7%)	46	(7.1%)	102	(6.6%)
Web sites that triggered 14 distinct diagnostics	2	(5.1%)	32	(4.4%)	1	(0.7%)	22	(3.4%)	57	(3.7%)
Web sites that triggered 15 distinct diagnostics	4	(10.3%)	24	(3.3%)	3	(2.1%)	20	(3.1%)	51	(3.3%)
Web sites that triggered 16 distinct diagnostics	0	(0%)	17	(2.4%)	5	(3.5%)	26	(4%)	48	(3.1%)

Table D.11: **Density of OpenSP diagnostics**

	IE		UK		FR		DE		Overall	
Web sites that triggered 17 distinct diagnostics	1	(2.6%)	8	(1.1%)	1	(0.7%)	11	(1.7%)	21	(1.4%)
Web sites that triggered 18 distinct diagnostics	1	(2.6%)	4	(0.6%)	3	(2.1%)	6	(0.9%)	14	(0.9%)
Web sites that triggered 19 distinct diagnostics	1	(2.6%)	5	(0.7%)	3	(2.1%)	8	(1.2%)	17	(1.1%)
Web sites that triggered 20 distinct diagnostics	0	(0%)	2	(0.3%)	1	(0.7%)	5	(0.8%)	8	(0.5%)
Web sites that triggered 21 distinct diagnostics	0	(0%)	0	(0%)	0	(0%)	5	(0.8%)	5	(0.3%)
Web sites that triggered 22 distinct diagnostics	0	(0%)	0	(0%)	0	(0%)	1	(0.2%)	1	(0.1%)

Figure D.6: **Density of OpenSP diagnostics**

11
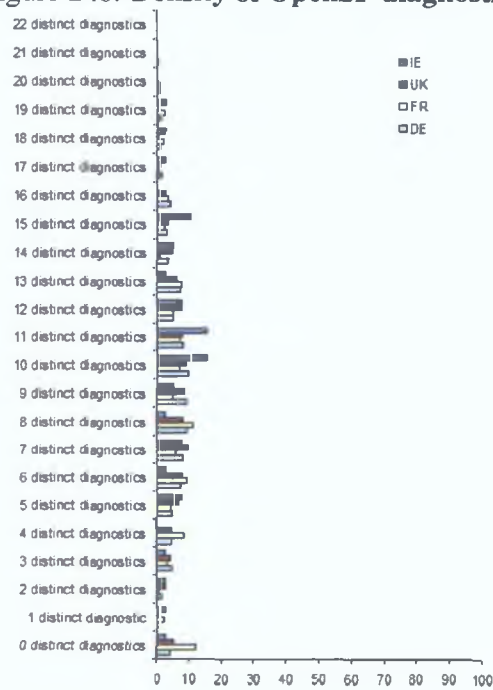
Appendix E

Results of surveys conducted on the web samples of Irish, UK, French and German web sites (May 2004)

E.1   Web content mirroring process results

Table E.1: **Results of the web content sampling process**

	IE	UK	FR	DE	Overall
Data amount retrieved	26 MB	519 MB	157 MB	380 MB	1.1 GB
Web pages considered for mirroring	2,901	60,440	19,148	56,641	139,130
Web pages correctly retrieved (size>0 and no retrieval errors)	2,400 (82.7%)	52,260 (86.5%)	16,388 (85.6%)	50,461 (89.1%)	121,509(87.3%)
Web pages in web sites qualified for conformance investigations	1,966 (67.8%)	40,769 (67.5%)	13,461 (70.3%)	37,245 (65.8%)	93,441 (67.2%)
Web sites qualified for conformance investigations	110 (40.4%)	2,243 (39.3%)	698 (45.2%)	1,646 (38.7%)	4,697 (39.9%)

Table E.2: **Details regarding the web sites that didn't qualify for investigations**

	IE	UK	FR	DE	Overall
Web sites not qualified for investigations	162	3,459	847	2,604	7,072
No data retrieved	21 (13%)	534 (15.4%)	178 (21%)	205 (7.9%)	938 (13.3%)
Only the main page retrieved (size>0)	103 (63.6%)	1,749 (50.6%)	361 (42.6%)	1,226 (47.1%)	3,439 (48.6%)
Less than 3 pages (2) and at least 100 KB retrieved	0 (0%)	5 (0.1%)	1 (0.1%)	4 (0.2%)	10 (0.1%)
Less than 100 KB and at least 3 pages retrieved	34 (21%)	1,004 (29%)	258 (30.5%)	1,054 (40.5%)	2,350 (33.2%)
Less than 3 pages (2) and less that 100 KB (but more than 0) retrieved	4 (2.5%)	167 (4.8%)	49 (5.8%)	115 (4.4%)	335 (4.7%)

E.2 WCAG 1.0 conformance investigation results

Table E.3: **Details regarding Bobby executions**

	IE		UK		FR		DE		Overall	
Web sites considered by Bobby	110		2,243		698		1,646		4,697	
Bobby abnormally terminated	3	(2.7%)	123	(5.5%)	51	(7.3%)	102	(6.2%)	279	(5.9%)
Bobby "locked up"	6	(5.5%)	83	(3.7%)	25	(3.6%)	66	(4%)	180	(3.8%)
Bobby completed successfully	101	(91.8%)	2,036	(90.8%)	622	(89.1%)	1,477	(89.7%)	4,236	(90.2%)

Table E.4: **Bobby Priority 1 diagnostics triggered during the investigation**

ID		IE		UK		FR		DE		Overall	
g9	Provide alternative text for all images	92	(91.1%)	1,840	(90.4%)	603	(96.9%)	1,339	(90.7%)	3,874	(91.5%)
g39	Give each frame a title	22	(21.8%)	491	(24.1%)	235	(37.8%)	721	(48.8%)	1,469	(34.7%)
g38	Each FRAME must reference an HTML file	23	(22.8%)	452	(22.2%)	217	(34.9%)	706	(47.8%)	1,398	(33%)
g240	Provide alternative text for all image map hot-spots	22	(21.8%)	372	(18.3%)	227	(36.5%)	233	(15.8%)	854	(20.2%)
g10	Provide alternative text for all image-type buttons in forms	12	(11.9%)	171	(8.4%)	79	(12.7%)	141	(9.5%)	403	(9.5%)
g21	Provide alternative text for each APPLET	7	(6.9%)	116	(5.7%)	34	(5.5%)	72	(4.9%)	229	(5.4%)
g20	Provide alternative content for each OBJECT	0	(0%)	2	(0.1%)	3	(0.5%)	4	(0.3%)	9	(0.2%)

Figure E.1: **Bobby Priority 1 diagnostics**
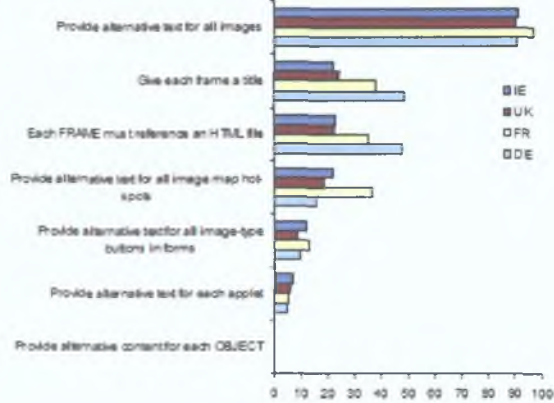
Table E.5: **Density of Bobby Priority 1 diagnostics**

	IE		UK		FR		DE		Overall	
Web sites with Bobby completed successfully	101		2,036		622		1,477		4,236	
Web sites that triggered 0 distinct diagnostics	7	(6.9%)	157	(7.7%)	12	(1.9%)	69	(4.7%)	245	(5.8%)
Web sites that triggered 1 distinct diagnostic	48	(47.5%)	991	(48.7%)	211	(33.9%)	513	(34.7%)	1,763	(41.6%)
Web sites that triggered 2 distinct diagnostic	21	(20.8%)	388	(19.1%)	156	(25.1%)	222	(15%)	787	(18.6%)
Web sites that triggered 3 distinct diagnostics	15	(14.9%)	344	(16.9%)	119	(19.1%)	465	(31.5%)	943	(22.3%)
Web sites that triggered 4 distinct diagnostics	7	(6.9%)	135	(6.6%)	102	(16.4%)	177	(12%)	421	(9.9%)
Web sites that triggered 5 distinct diagnostics	3	(3%)	21	(1%)	22	(3.5%)	30	(2%)	76	(1.8%)
Web sites that triggered 6 distinct diagnostics	0	(0%)	0	(0%)	0	(0%)	1	(0.1%)	1	(0%)

4

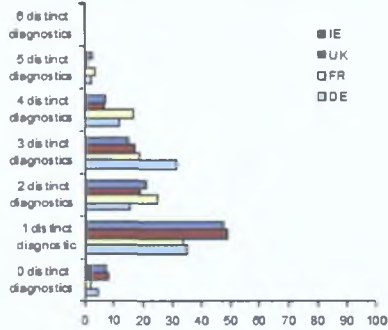Figure E.2: **Density of Bobby Priority 1 diagnostics**

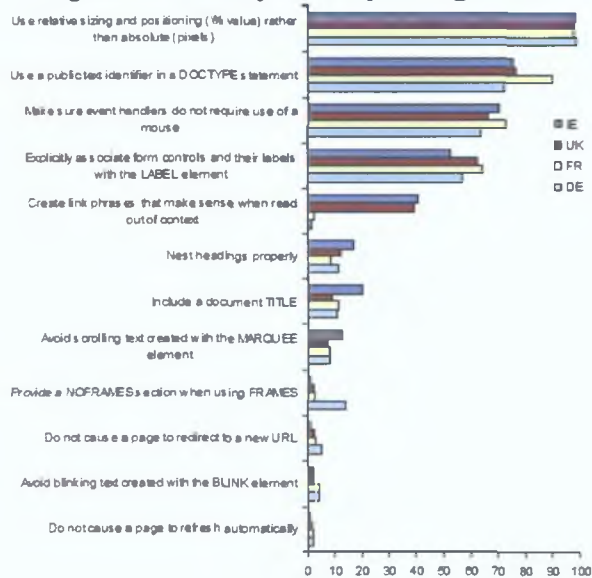Figure E.3: **Bobby Priority 2 diagnostics**
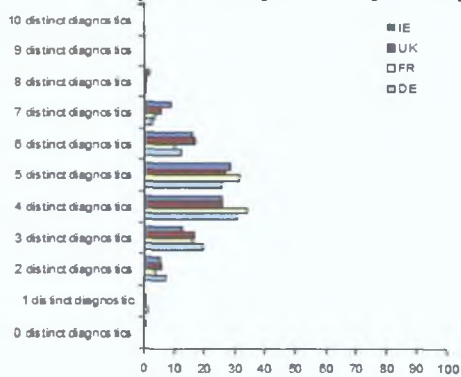

Figure E.4: **Density of Bobby Priority 2 diagnostics**

5

Table E.6: **Bobby Priority 2 diagnostics triggered during the investigation**

ID		IE		UK		FR		DE		Overall	
g104	Use relative sizing and positioning (% value) rather than absolute (pixels)	99	(98%)	1,983	(97.4%)	609	(97.9%)	1,454	(98.4%)	4,145	(97.9%)
g271	Use a public text identifier in a DOCTYPE statement	76	(75.2%)	1,553	(76.3%)	560	(90%)	1,067	(72.2%)	3,256	(76.9%)
g269	Make sure event handlers do not require use of a mouse	71	(70.3%)	1,347	(66.2%)	452	(72.7%)	942	(63.8%)	2,812	(66.4%)
g41	Explicitly associate form controls and their labels with the LABEL element	53	(52.5%)	1,269	(62.3%)	401	(64.5%)	838	(56.7%)	2,561	(60.5%)
g34	Create link phrases that make sense when read out of context	41	(40.6%)	797	(39.1%)	17	(2.7%)	21	(1.4%)	876	(20.7%)
g2	Nest headings properly	17	(16.8%)	243	(11.9%)	54	(8.7%)	167	(11.3%)	481	(11.4%)
g273	Include a document TI-TLE	20	(19.8%)	186	(9.1%)	71	(11.4%)	162	(11%)	439	(10.4%)
g5	Avoid scrolling text created with the MARQUEE element	13	(12.9%)	149	(7.3%)	52	(8.4%)	115	(7.8%)	329	(7.8%)
g37	Provide a NOFRAMES section when using FRAMES	1	(1%)	47	(2.3%)	16	(2.6%)	212	(14.4%)	276	(6.5%)
g254	Do not cause a page to redirect to a new URL	1	(1%)	51	(2.5%)	19	(3.1%)	80	(5.4%)	151	(3.6%)
g4	Avoid blinking text created with the BLINK element	2	(2%)	40	(2%)	27	(4.3%)	63	(4.3%)	132	(3.1%)
g33	Do not cause a page to refresh automatically	1	(1%)	34	(1.7%)	14	(2.3%)	32	(2.2%)	81	(1.9%)

6

Table E.7: **Density of Bobby Priority 2 diagnostics**

	IE		UK		FR		DE		Overall	
Web sites with Bobby completed successfully	101		2,036		622		1,477		4,236	
Web sites that triggered 0 distinct diagnostics	1	(1%)	6	(0.3%)	0	(0%)	1	(0.1%)	8	(0.2%)
Web sites that triggered 1 distinct diagnostic	0	(0%)	24	(1.2%)	5	(0.8%)	20	(1.4%)	49	(1.2%)
Web sites that triggered 2 distinct diagnostics	5	(5%)	113	(5.6%)	21	(3.4%)	104	(7%)	243	(5.7%)
Web sites that triggered 3 distinct diagnostics	13	(12.9%)	336	(16.5%)	98	(15.8%)	285	(19.3%)	732	(17.3%)
Web sites that triggered 4 distinct diagnostics	26	(25.7%)	522	(25.6%)	210	(33.8%)	456	(30.9%)	1,214	(28.7%)
Web sites that triggered 5 distinct diagnostics	29	(28.7%)	548	(26.9%)	196	(31.5%)	379	(25.7%)	1,152	(27.2%)
Web sites that triggered 6 distinct diagnostics	16	(15.8%)	342	(16.8%)	64	(10.3%)	181	(12.3%)	603	(14.2%)
Web sites that triggered 7 distinct diagnostics	9	(8.9%)	115	(5.6%)	23	(3.7%)	42	(2.8%)	189	(4.5%)
Web sites that triggered 8 distinct diagnostics	2	(2%)	24	(1.2%)	5	(0.8%)	9	(0.6%)	40	(0.9%)
Web sites that triggered 9 distinct diagnostics	0	(0%)	5	(0.2%)	0	(0%)	0	(0%)	5	(0.1%)
Web sites that triggered 10 distinct diagnostics	0	(0%)	1	(0%)	0	(0%)	0	(0%)	1	(0%)

7

E.3 HTML technical standards conformance investigation results

Table E.8: **Details regarding OpenSP executions**

	IE		UK		FR		DE		Overall	
Web sites processed by OpenSP	110		2,243		698		1,646		4,697	
Web sites where all the web pages had usable document type information	1	(0.9%)	52	(2.3%)	4	(0.6%)	54	(3.3%)	111	(2.4%)
Web sites considered for HTML defects analysis	45	(40.9%)	792	(35.3%)	175	(25.1%)	706	(42.9%)	1,718	(36.6%)

Table E.9: **Document Type Declaration in web pages considered by OpenSP**

	IE		UK		FR		DE		Overall	
Web pages processed by OpenSP	1,966		40,769		13,461		37,245		93,441	
No document type information	1,310	(66.6%)	28,023	(68.7%)	10,954	(81.4%)	22,382	(60.1%)	62,669	(67.1%)
Incorrect document type information	70	(3.6%)	843	(2.1%)	407	(3%)	2,024	(5.4%)	3,344	(3.6%)
Usable document type information	549	(27.9%)	11,409	(28%)	2,050	(15.2%)	12,502	(33.6%)	26,510	(28.4%)
Considered for HTML defects analysis (in web sites with at least 3 web pages and 100 KB of data with usable document type information)	542	(27.6%)	11,146	(27.3%)	1,934	(14.4%)	12,270	(32.9%)	25,892	(27.7%)

Table E.10: **The 10 most common OpenSP diagnostics triggered in the investigation**

ID		IE		UK		FR		DE		Overall	
108	Undefined attribute	42	(93.3%)	687	(86.7%)	142	(81.1%)	633	(89.7%)	1,504	(87.5%)
127	Required attribute not specified	34	(75.6%)	605	(76.4%)	136	(77.7%)	567	(80.3%)	1,342	(78.1%)
64	Element not allowed by document type	35	(77.8%)	559	(70.6%)	123	(70.3%)	548	(77.6%)	1,265	(73.6%)
79	End tag for not opened element	32	(71.1%)	539	(68.1%)	103	(58.9%)	479	(67.8%)	1,153	(67.1%)
68	Missing required end tag (element DTD)	28	(62.2%)	362	(45.7%)	70	(40%)	358	(50.7%)	818	(47.6%)
25	General entity not defined and no default entity	22	(48.9%)	395	(49.9%)	82	(46.9%)	309	(43.8%)	808	(47%)
139	Non SGML character number	27	(60%)	332	(41.9%)	71	(40.6%)	332	(47%)	762	(44.4%)
73	End tag for element not finished	16	(35.6%)	285	(36%)	44	(25.1%)	302	(42.8%)	647	(37.7%)
131	Attribute value not allowed	21	(46.7%)	285	(36%)	71	(40.6%)	256	(36.3%)	633	(36.8%)
82	Attribute value must be literal	23	(51.1%)	268	(33.8%)	47	(26.9%)	243	(34.4%)	581	(33.8%)

Figure E.5: **The 10 most common OpenSP diagnostics**
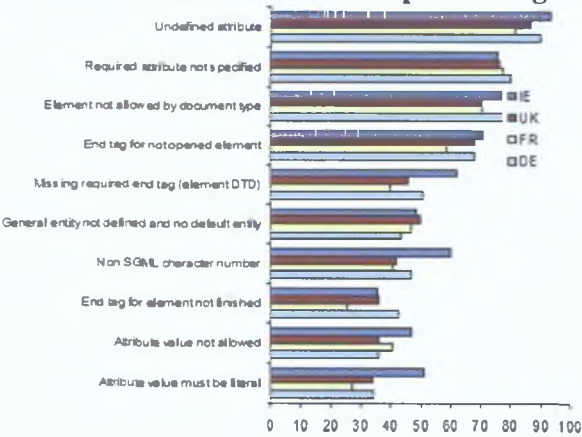

9

Table E.11: **Density of OpenSP diagnostics**

	IE		UK		FR		DE		Overall	
Web sites considered for further analysis	45		792		175		706		1,718	
Web sites that triggered 0 distinct diagnostics	1	(2.2%)	41	(5.2%)	22	(12.6%)	29	(4.1%)	93	(5.4%)
Web sites that triggered 1 distinct diagnostic	1	(2.2%)	11	(1.4%)	1	(0.6%)	6	(0.8%)	19	(1.1%)
Web sites that triggered 2 distinct diagnostics	1	(2.2%)	14	(1.8%)	5	(2.9%)	11	(1.6%)	31	(1.8%)
Web sites that triggered 3 distinct diagnostics	2	(4.4%)	36	(4.5%)	8	(4.6%)	33	(4.7%)	79	(4.6%)
Web sites that triggered 4 distinct diagnostics	0	(0%)	45	(5.7%)	9	(5.1%)	32	(4.5%)	86	(5%)
Web sites that triggered 5 distinct diagnostics	4	(8.9%)	56	(7.1%)	6	(3.4%)	37	(5.2%)	103	(6%)
Web sites that triggered 6 distinct diagnostics	1	(2.2%)	61	(7.7%)	17	(9.7%)	53	(7.5%)	132	(7.7%)
Web sites that triggered 7 distinct diagnostics	3	(6.7%)	73	(9.2%)	10	(5.7%)	63	(8.9%)	149	(8.7%)
Web sites that triggered 8 distinct diagnostics	2	(4.4%)	62	(7.8%)	16	(9.1%)	63	(8.9%)	143	(8.3%)
Web sites that triggered 9 distinct diagnostics	2	(4.4%)	69	(8.7%)	21	(12%)	62	(8.8%)	154	(9%)
Web sites that triggered 10 distinct diagnostics	5	(11.1%)	68	(8.6%)	15	(8.6%)	57	(8.1%)	145	(8.4%)
Web sites that triggered 11 distinct diagnostics	5	(11.1%)	50	(6.3%)	9	(5.1%)	68	(9.6%)	132	(7.7%)
Web sites that triggered 12 distinct diagnostics	5	(11.1%)	47	(5.9%)	5	(2.9%)	39	(5.5%)	96	(5.6%)
Web sites that triggered 13 distinct diagnostics	3	(6.7%)	50	(6.3%)	9	(5.1%)	40	(5.7%)	102	(5.9%)
Web sites that triggered 14 distinct diagnostics	7	(15.6%)	34	(4.3%)	8	(4.6%)	32	(4.5%)	81	(4.7%)
Web sites that triggered 15 distinct diagnostics	1	(2.2%)	35	(4.4%)	3	(1.7%)	21	(3%)	60	(3.5%)
Web sites that triggered 16 distinct diagnostics	1	(2.2%)	19	(2.4%)	3	(1.7%)	16	(2.3%)	39	(2.3%)

Table E.11: **Density of OpenSP diagnostics**

	IE		UK		FR		DE		Overall	
Web sites that triggered 17 distinct diagnostics	0	(0%)	9	(1.1%)	2	(1.1%)	15	(2.1%)	26	(1.5%)
Web sites that triggered 18 distinct diagnostics	0	(0%)	6	(0.8%)	4	(2.3%)	14	(2%)	24	(1.4%)
Web sites that triggered 19 distinct diagnostics	0	(0%)	4	(0.5%)	1	(0.6%)	5	(0.7%)	10	(0.6%)
Web sites that triggered 20 distinct diagnostics	1	(2.2%)	1	(0.1%)	0	(0%)	5	(0.7%)	7	(0.4%)
Web sites that triggered 21 distinct diagnostics	0	(0%)	1	(0.1%)	1	(0.6%)	3	(0.4%)	5	(0.3%)
Web sites that triggered 22 distinct diagnostics	0	(0%)	0	(0%)	0	(0%)	0	(0%)	0	(0%)
Web sites that triggered 23 distinct diagnostics	0	(0%)	0	(0%)	0	(0%)	2	(0.3%)	2	(0.1%)

11

Figure E.6: **Density of OpenSP diagnostics**
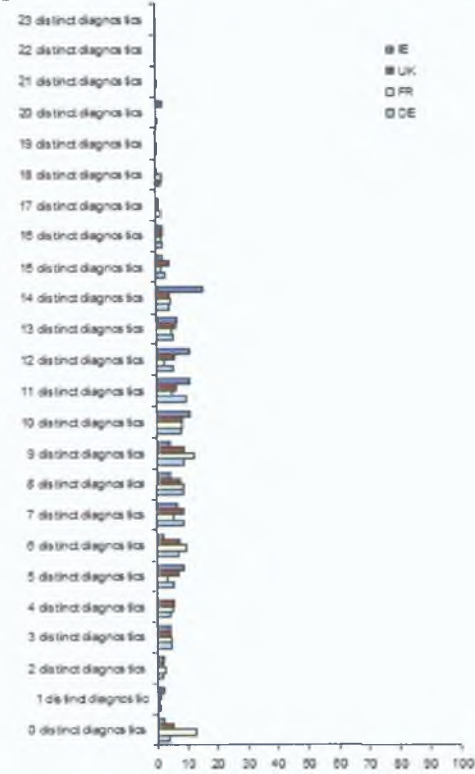
Appendix F

Results of surveys conducted on the web sample of subjectively selected Irish web sites (April 2002–December 2004)

F.1   Web content mirroring process results

Table F.1: **Results of the web content sampling process**

	Apr 2002	Dec 2002	May 2003	Jan 2004	May 2004	Dec 2004	Average
Data amount retrieved	33 MB	35 MB	34 MB	34 MB	36 MB	32 MB	34 MB
Web pages considered for mirroring	3,586	3,552	3,382	3,051	3,052	2,749	3,228.7
Web pages correctly retrieved (size>0 and no retrieval errors)	3,234 (90.2%)	3,131 (88.1%)	2,980 (88.1%)	2,667 (87.4%)	2,690 (88.1%)	2,318 (84.3%)	2,836.7 (87.9%)
Web pages in web sites qualified for conformance investigations	3,001 (83.7%) (83.1%)	2,950 (83.5%)	2,825 (83%)	2,531 (84.6%)	2,582 (79.8%)	2,195 (83%)	2,680.7
Web sites qualified for conformance investigations	167 (78%)	159 (74.3%)	156 (72.9%)	153 (71.5%)	162 (75.7%)	143 (66.8%)	156.7 (73.2%)

Table F.2: **Details regarding the web sites that didn't qualify for investigations**

	Apr 2002	Dec 2002	May 2003	Jan 2004	May 2004	Dec 2004	Common	Average
Web sites not qualified or investigations	47	55	58	61	52	71	18	57.3
No data retrieved	8 (17%)	9 (16.4%)	9 (15.5%)	10 (16.4%)	10 (19.2%)	19 (26.8%)	2 (11.1%)	10.8 (18.9%)
Only the main page retrieved (size>0)	19 (40.4%)	27 (49.1%)	31 (53.4%)	32 (52.5%)	25 (48.1%)	34 (47.9%)	4 (22.2%)	28 (48.8%)
Less than 3 pages (2) and at least 100 KB retrieved	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)
Less than 100 KB and at least 3 pages retrieved	18 (38.3%)	17 (30.9%)	15 (25.9%)	14 (23%)	11 (21.2%)	14 (19.7%)	4 (22.2%)	14.8 (25.9%)
Less than 3 pages (2) and less that 100 KB (but more than 0) retrieved	2 (4.3%)	2 (3.6%)	3 (5.2%)	5 (8.2%)	6 (11.5%)	4 (5.6%)	0 (0%)	3.7 (6.4%)

F.2 WCAG 1.0 conformance investigation results

Table F.3: **Details regarding Bobby executions**

	Apr 2002	Dec 2002	May 2003	Jan 2004	May 2004	Dec 2004	Average
Web sites considered by Bobby	167	159	156	153	162	143	156.7
Bobby abnormally terminated	9 (5.4%)	16 (10.1%)	13 (8.3%)	10 (6.5%)	14 (8.6%)	8 (5.6%)	11.7 (7.4%)
Bobby "locked up"	11 (6.6%)	8 (5%)	9 (5.8%)	7 (4.6%)	8 (4.9%)	5 (3.5%)	8 (5.1%)
Bobby completed successfully	147 (88%)	135 (84.9%)	134 (85.9%)	136 (88.9%)	140 (86.4%)	130 (90.9%)	137 (87.4%)

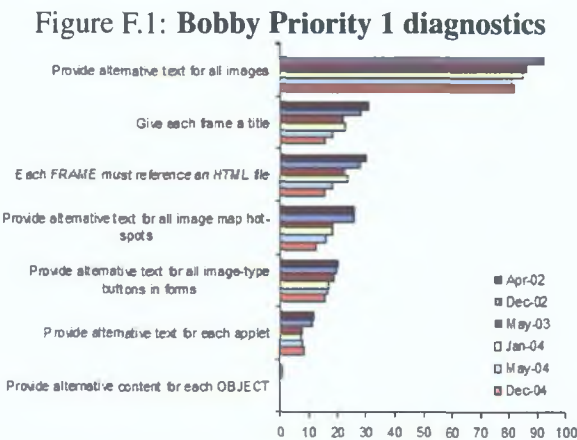Figure F.1: **Bobby Priority 1 diagnostics**

3

Table F.4: **Bobby Priority 1 diagnostics triggered during the investigation**

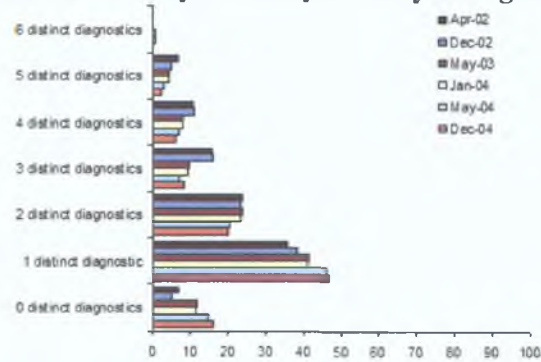| ID		Apr 2002	Dec 2002	May 2003	Jan 2004	May 2004	Dec 2004	Average
g9	Provide alternative text for all images	132 (89.8%)	125 (92.6%)	115 (85.8%)	115 (84.6%)	113 (80.7%)	106 (81.5%)	117.7 (85.8%)
g39	Give each frame a title	45 (30.6%)	38 (28.1%)	29 (21.6%)	31 (22.8%)	25 (17.9%)	20 (15.4%)	31.3 (22.7%)
g38	Each FRAME must reference an HTML file	44 (29.9%)	37 (27.4%)	30 (22.4%)	32 (23.5%)	25 (17.9%)	20 (15.4%)	31.3 (22.7%)
g240	Provide alternative text for all image map hot-spots	38 (25.9%)	35 (25.9%)	24 (17.9%)	25 (18.4%)	22 (15.7%)	16 (12.3%)	26.7 (19.3%)
g10	Provide alternative text for all image-type buttons in forms	29 (19.7%)	26 (19.3%)	25 (18.7%)	23 (16.9%)	23 (16.4%)	20 (15.4%)	24.3 (17.7%)
g21	Provide alternative text for each APPLET	17 (11.6%)	15 (11.1%)	10 (7.5%)	10 (7.4%)	11 (7.9%)	11 (8.5%)	12.3 (9%)
g20	Provide alternative content for each OB-JECT	1 (0.7%)	1 (0.7%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0.3 (0.2%)

Figure F.2: **Density of Bobby Priority 1 diagnostics**

Table F.5: **Density of Bobby Priority 1 diagnostics**

	Apr 2002	Dec 2002	May 2003	Jan 2004	May 2004	Dec 2004	Average
Web sites with Bobby completed successfully	147	135	134	136	140	130	137
Web sites that triggered 0 distinct diagnostics	10 (6.8%)	7 (5.2%)	16 (11.9%)	17 (12.5%)	21 (15%)	21 (16.2%)	15.3 (11.2%)
Web sites that triggered 1 distinct diagnostic	53 (36.1%)	52 (38.5%)	56 (41.8%)	56 (41.2%)	65 (46.4%)	61 (46.9%)	57.2 (41.7%)
Web sites that triggered 2 distinct diagnostics	35 (23.8%)	32 (23.7%)	32 (23.9%)	32 (23.5%)	29 (20.7%)	26 (20%)	31 (22.6%)
Web sites that triggered 3 distinct diagnostics	23 (15.6%)	22 (16.3%)	13 (9.7%)	13 (9.6%)	10 (7.1%)	11 (8.5%)	15.3 (11.2%)
Web sites that triggered 4 distinct diagnostics	16 (10.9%)	15 (11.1%)	11 (8.2%)	11 (8.1%)	10 (7.1%)	8 (6.2%)	11.8 (8.6%)
Web sites that triggered 5 distinct diagnostics	10 (6.8%)	7 (5.2%)	6 (4.5%)	6 (4.4%)	4 (2.9%)	3 (2.3%)	6 (4.4%)
Web sites that triggered 6 distinct diagnostics	0 (0%)	0 (0%)	0 (0%)	1 (0.7%)	1 (0.7%)	0 (0%)	0.3 (0.2%)

Table F.6: **Bobby Priority 2 diagnostics triggered during the investigation**

ID		Apr 2002	Dec 2002	May 2003	Jan 2004	May 2004	Dec 2004	Average
g104	Use relative sizing and positioning (% value) rather than absolute (pixels)	145 (98.6%)	133 (98.5%)	129 (96.3%)	130 (95.6%)	132 (94.3%)	122 (93.8%)	131.8 (96.2%)
g41	Explicitly associate form controls and their labels with the LABEL *element*	106 (72.1%)	104 (77%)	100 (74.6%)	101 (74.3%)	106 (75.7%)	100 (76.9%)	102.8 (75.1%)
g271	Use a public text identifier in a DOC-TYPE statement	132 (89.8%)	116 (85.9%)	102 (76.1%)	94 (69.1%)	89 (63.6%)	75 (57.7%)	101.3 (73.7%)
g269	Make sure event handlers do not require use of a mouse	101 (68.7%)	98 (72.6%)	97 (72.4%)	100 (73.5%)	103 (73.6%)	96 (73.8%)	99.2 (72.4%)
g34	Create link phrases that make sense when read out of context	58 (39.5%)	59 (43.7%)	57 (42.5%)	60 (44.1%)	55 (39.3%)	52 (40%)	56.8 (41.5%)
g273	Include a document TITLE	17 (11.6%)	16 (11.9%)	17 (12.7%)	18 (13.2%)	14 (10%)	13 (10%)	15.8 (11.6%)
g2	Nest headings properly	9 (6.1%)	8 (5.9%)	11 (8.2%)	9 (6.6%)	11 (7.9%)	13 (10%)	10.2 (7.5%)
g37	Provide a NOFRAMES section when using FRAMES	6 (4.1%)	6 (4.4%)	6 (4.5%)	6 (4.4%)	4 (2.9%)	4 (3.1%)	5.3 (3.9%)
g254	Do not cause a page to redirect to a new URL	4 (2.7%)	3 (2.2%)	3 (2.2%)	4 (2.9%)	4 (2.9%)	2 (1.5%)	3.3 (2.4%)
g33	Do not cause a page to refresh automatically	5 (3.4%)	2 (1.5%)	2 (1.5%)	3 (2.2%)	3 (2.1%)	4 (3.1%)	3.2 (2.3%)
g5	Avoid scrolling text created with the MARQUEE element	3 (2%)	3 (2.2%)	2 (1.5%)	2 (1.5%)	3 (2.1%)	2 (1.5%)	2.5 (1.8%)
g4	Avoid blinking text	2	1	1	0	0	0	0.7

Table F.6: **Bobby Priority 2 diagnostics triggered during the investigation**

ID		Apr 2002	Dec 2002	May 2003	Jan 2004	May 2004	Dec 2004	Average
	created with the BLINK element	(1.4%)	(0.7%)	(0.7%)	(0%)	(0%)	(0%)	(0.5%)

Figure F.3: **Bobby Priority 2 diagnostics**

7
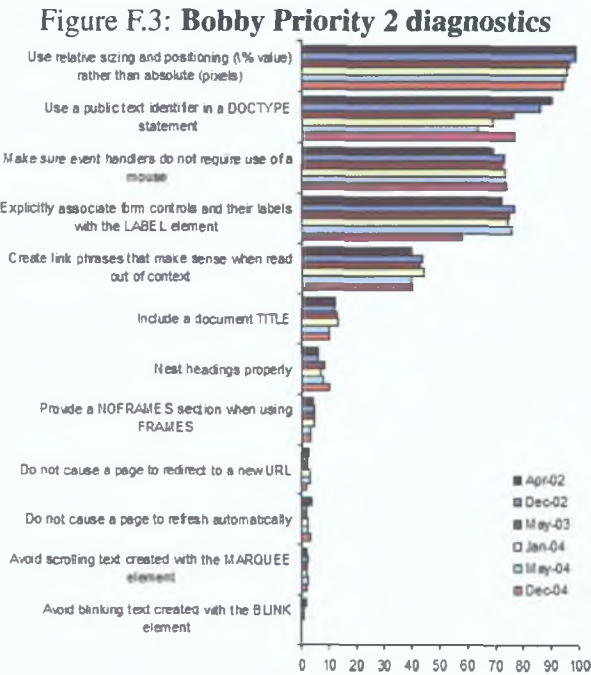
Table F.7: **Density of Bobby Priority 2 diagnostics**

	Apr 2002	Dec 2002	May 2003	Jan 2004	May 2004	Dec 2004	Average
Web sites with Bobby completed successfully	147	135	134	136	140	130	137
Web sites that triggered 0 distinct diagnostics	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)
Web sites that triggered 1 distinct diagnostic	0 (0%)	0 (0%)	1 (0.7%)	1 (0.7%)	2 (1.4%)	1 (0.8%)	0.8 (0.6%)
Web sites that triggered 2 distinct diagnostics	4 (2.7%)	4 (3%)	3 (2.2%)	6 (4.4%)	7 (5%)	4 (3.1%)	4.7 (3.4%)
Web sites that triggered 3 distinct diagnostics	17 (11.6%)	9 (6.7%)	12 (9%)	17 (12.5%)	15 (10.7%)	17 (13.1%)	14.5 (10.6%)
Web sites that triggered 4 distinct diagnostics	46 (31.3%)	43 (31.9%)	40 (29.9%)	34 (25%)	40 (28.6%)	41 (31.5%)	40.7 (29.7%)
Web sites that triggered 5 distinct diagnostics	36 (24.5%)	32 (23.7%)	42 (31.3%)	40 (29.4%)	45 (32.1%)	34 (26.2%)	38.2 (27.9%)
Web sites that triggered 6 distinct diagnostics	32 (21.8%)	37 (27.4%)	24 (17.9%)	25 (18.4%)	17 (12.1%)	24 (18.5%)	26.5 (19.3%)
Web sites that triggered 7 distinct diagnostics	7 (4.8%)	7 (5.2%)	8 (6%)	8 (5.9%)	12 (8.6%)	9 (6.9%)	8.5 (6.2%)
Web sites that triggered 8 distinct diagnostics	4 (2.7%)	2 (1.5%)	4 (3%)	4 (2.9%)	1 (0.7%)	0 (0%)	2.5 (1.8%)
Web sites that triggered 9 distinct diagnostics	1 (0.7%)	0 (0%)	0 (0%)	1 (0.7%)	1 (0.7%)	0 (0%)	0.5 (0.4%)
Web sites that triggered 10 distinct diagnostics	0 (0%)	1 (0.7%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0.2 (0.1%)

8

Figure F.4: **Density of Bobby Priority 2 diagnostics**
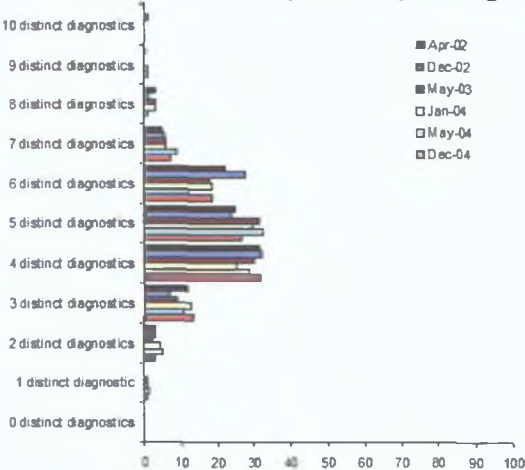
F.3 HTML technical standards conformance investigation results

Table F.8: **Details regarding OpenSP executions**

	Apr 2002	Dec 2002	May 2003	Jan 2004	May 2004	Dec 2004	Average
Web sites processed by OpenSP	167	159	156	153	162	143	156.7
Web sites where all the web pages had usable document type information	0 (0%)	0 (0%)	1 (0.6%)	0 (0%)	0 (0%)	0 (0%)	0.2 (0.1%)
Web sites considered for HTML defects analysis	54 (32.3%)	57 (35.8%)	67 (42.9%)	74 (48.4%)	80 (49.4%)	79 (55.2%)	68.5 (43.7%)

Table F.9: **Document Type Declaration in web pages considered by OpenSP**

	Apr 2002	Dec 2002	May 2003	Jan 2004	May 2004	Dec 2004	Average
Web pages processed by OpenSP	3,295	2,950	2,825	2,531	2,582	2,195	2729.7
No document type information	2,222 (67.4%)	2,109 (71.5%)	1,866 (66.1%)	1,547 (61.1%)	1,497 (58%)	1,206 (54.9%)	1741.2 (63.8%)
Incorrect document type information	46 (1.4%)	57 (1.9%)	23 (0.8%)	8 (0.3%)	16 (0.6%)	8 (0.4%)	26.3 (1%)
Usable document type information	704 (21.4%)	751 (25.5%)	932 (33%)	974 (38.5%)	1,068 (41.4%)	980 (44.6%)	901.5 (33%)
Considered for HTML defects analysis (in web sites with at least 3 web pages and 100 KB of data with usable document type information)	671 (20.4%)	726 (24.6%)	910 (32.2%)	952 (37.6%)	1,037 (40.2%)	958 (43.6%)	875.7 (32.1%)

10

Table F.10: **The 10 most common OpenSP diagnostics triggered in the investigation**

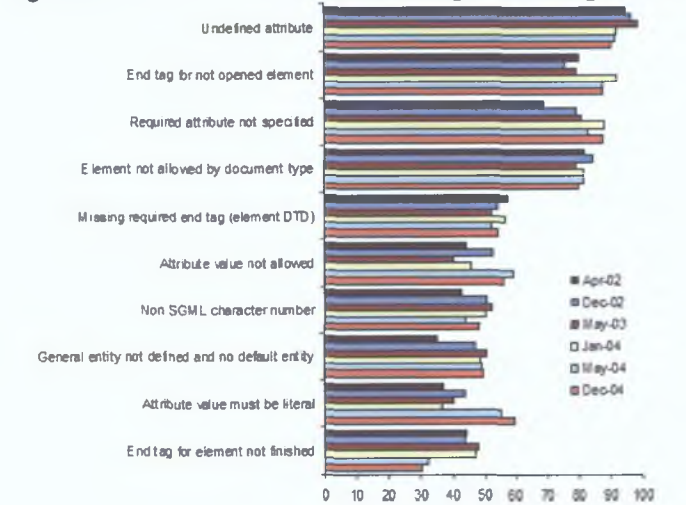| ID		Apr 2002	Dec 2002	May 2003	Jan 2004	May 2004	Dec 2004	Average
108	Undefined attribute	51 (94.4%)	55 (96.5%)	66 (98.5%)	68 (91.9%)	73 (91.3%)	71 (89.9%)	64 (93.7%)
64	Element not allowed by document type	44 (81.5%)	48 (84.2%)	53 (79.1%)	60 (81.1%)	65 (81.3%)	63 (79.7%)	55.5 (81.1%)
79	End tag for not opened element	43 (79.6%)	43 (75.4%)	53 (79.1%)	68 (91.9%)	70 (87.5%)	69 (87.3%)	57.7 (83.5%)
127	Required attribute not specified	37 (68.5%)	45 (78.9%)	54 (80.6%)	65 (87.8%)	66 (82.5%)	69 (87.3%)	56 (81%)
68	Missing required end tag (element DTD)	31 (57.4%)	31 (54.4%)	35 (52.2%)	42 (56.8%)	42 (52.5%)	43 (54.4%)	37.3 (54.6%)
131	Attribute value not allowed	24 (44.4%)	30 (52.6%)	27 (40.3%)	34 (45.9%)	35 (43.8%)	38 (48.1%)	31.3 (45.9%)
73	End tag for element not finished	24 (44.4%)	25 (43.9%)	32 (47.8%)	35 (47.3%)	39 (48.8%)	39 (49.4%)	32.3 (46.9%)
139	Non SGML character number	23 (42.6%)	29 (50.9%)	35 (52.2%)	37 (50%)	47 (58.8%)	44 (55.7%)	35.8 (51.7%)
82	Attribute value must be literal	20 (37%)	25 (43.9%)	27 (40.3%)	27 (36.5%)	26 (32.5%)	24 (30.4%)	24.8 (36.8%)
25	General entity not defined and no default entity	19 (35.2%)	27 (47.4%)	34 (50.7%)	36 (48.6%)	44 (55%)	47 (59.5%)	34.5 (49.4%)

11

Figure F.5: **The 10 most common OpenSP diagnostics**

Table F.11: **Density of OpenSP diagnostics**

	Apr 2002	Dec 2002	May 2003	Jan 2004	May 2004	Dec 2004	Average
Web sites considered for HTML defects analysis	54	57	67	74	80	79	68.5
Web sites that triggered 0 distinct diagnostics	0 (0%)	0 (0%)	0 (0%)	1 (1.4%)	1 (1.3%)	0 (0%)	0.3 (0.5%)
Web sites that triggered 1 distinct diagnostic	1 (1.9%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	1 (1.3%)	0.3 (0.5%)
Web sites that triggered 2 distinct diagnostics	1 (1.9%)	1 (1.8%)	0 (0%)	0 (0%)	1 (1.3%)	1 (1.3%)	0.7 (1%)
Web sites that triggered 3 distinct diagnostics	2 (3.7%)	2 (3.5%)	2 (3%)	1 (1.4%)	0 (0%)	1 (1.3%)	1.3 (1.9%)
Web sites that triggered 4 distinct diagnostics	4 (7.4%)	3 (5.3%)	6 (9%)	3 (4.1%)	3 (3.8%)	1 (1.3%)	3.3 (4.9%)
Web sites that triggered 5 distinct diagnostics	4 (7.4%)	5 (8.8%)	6 (9%)	5 (6.8%)	6 (7.5%)	6 (7.6%)	5.3 (7.8%)
Web sites that triggered 6 distinct diagnostics	1 (1.9%)	3 (5.3%)	2 (3%)	2 (2.7%)	2 (2.5%)	1 (1.3%)	1.8 (2.7%)
Web sites that triggered 7 distinct diagnostics	7 (13%)	3 (5.3%)	6 (9%)	5 (6.8%)	4 (5%)	6 (7.6%)	5.2 (7.5%)
Web sites that triggered 8 distinct diagnostics	2 (3.7%)	1 (1.8%)	5 (7.5%)	6 (8.1%)	8 (10%)	7 (8.9%)	4.8 (7.1%)
Web sites that triggered 9 distinct diagnostics	5 (9.3%)	6 (10.5%)	7 (10.4%)	8 (10.8%)	10 (12.5%)	7 (8.9%)	7.2 (10.5%)
Web sites that triggered 10 distinct diagnostics	4 (7.4%)	7 (12.3%)	3 (4.5%)	8 (10.8%)	8 (10%)	5 (6.3%)	5.8 (8.5%)
Web sites that triggered 11 distinct diagnostics	9 (16.7%)	7 (12.3%)	5 (7.5%)	6 (8.1%)	4 (5%)	5 (6.3%)	6 (8.8%)
Web sites that triggered 12 distinct diagnostics	3 (5.6%)	5 (8.8%)	7 (10.4%)	8 (10.8%)	11 (13.8%)	12 (15.2%)	7.7 (11.2%)
Web sites that triggered 13 distinct diagnostics	4 (7.4%)	2 (3.5%)	2 (3%)	6 (8.1%)	3 (3.8%)	5 (6.3%)	3.7 (5.4%)
Web sites that triggered 14 distinct diagnostics	1 (1.9%)	3 (5.3%)	4 (6%)	3 (4.1%)	7 (8.8%)	9 (11.4%)	4.5 (6.6%)
Web sites that triggered 15 distinct diagnostics	3 (5.6%)	5 (8.8%)	4 (6%)	5 (6.8%)	8 (10%)	7 (8.9%)	5.3 (7.8%)
Web sites that triggered 16 distinct diagnostics	1 (1.9%)	0 (0%)	4 (6%)	2 (2.7%)	0 (0%)	1 (1.3%)	1.3 (1.9%)
Web sites that triggered	0	4	1	3	1	2	1.8

13

Table F.11: **Density of OpenSP diagnostics**

	Apr 2002	Dec 2002	May 2003	Jan 2004	May 2004	Dec 2004	Average
17 distinct diagnostics	(0%)	(7%)	(1.5%)	(4.1%)	(1.3%)	(2.5%)	(2.7%)
Web sites that triggered 18 distinct diagnostics	0 (0%)	0 (0%)	2 (3%)	0 (0%)	1 (1.3%)	1 (1.3%)	0.7 (1%)
Web sites that triggered 19 distinct diagnostics	0 (0%)	0 (0%)	0 (0%)	2 (2.7%)	1 (1.3%)	0 (0%)	0.5 (0.7%)
Web sites that triggered 20 distinct diagnostics	0 (0%)	0 (0%)	1 (1.5%)	0 (0%)	1 (1.3%)	1 (1.3%)	0.5 (0.7%)
Web sites that triggered 21 distinct diagnostics	1 (1.9%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0.2 (0.2%)
Web sites that triggered 22 distinct diagnostics	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)
Web sites that triggered 23 distinct diagnostics	1 (1.9%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0.2 (0.2%)

14

Figure F.6: **Density of OpenSP diagnostics**
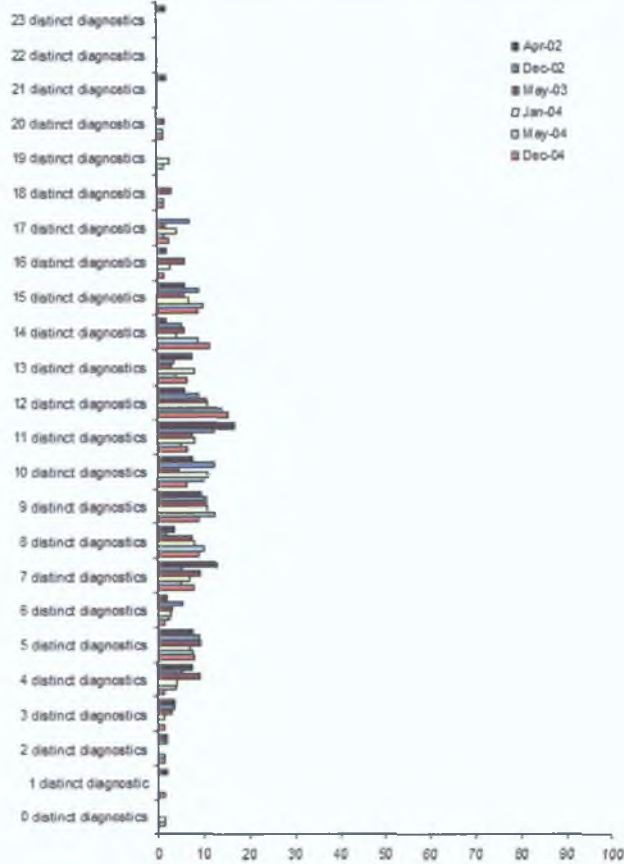
Appendix G

Results of surveys conducted on the web sample of randomly selected Irish web sites (May 2003–December 2004)

G.1   Web content mirroring process results

Table G.1: **Results of the web content sampling process**

	May 2003	Dec 2003	May 2004	Dec 2004	Average
Data amount retrieved	29 MB	26 MB	26 MB	27 MB	26 MB
Web pages considered for mirroring	3,319	3,043	2,901	2,542	2,951.3
Web pages correctly retrieved (size>0 and no retrieval errors)	2,817 (84.9%)	2,541 (83.5%)	2,400 (82.7%)	2,223 (87.5%)	2,495.3 (84.5%)
Web pages qualified for conformance investigation	2,288 (68.9%)	2,064 (67.8%)	1,966 (67.8%)	1,769 (69.6%)	2,021.8 (68.5%)
Web sites qualified for conformance investigation	123 (45.2%)	111 (40.8%)	110 (40.4%)	100 (36.8%)	111 (40.8%)

Table G.2: **Details regarding the web sites that didn't qualify for investigations**

	May 2003	Dec 2003	May 2004	Dec 2004	Common	Average
Web sites not qualified for investigations	149	161	162	172	132	161
No data retrieved	9 (6%)	15 (9.3%)	21 (13%)	24 (14%)	2 (1.5%)	17.3 (10.7%)
Only the main page retrieved (size>0)	91 (61.1%)	102 (63.4%)	103 (63.6%)	110 (64%)	78 (59.1%)	101.5 (63%)
Less than 3 pages (2) and at least 100 KB retrieved	1 (0.7%)	1 (0.6%)	0 (0%)	0 (0%)	0 (0%)	0.5 (0.3%)
Less than 100 KB and at least 3 pages retrieved	46 (30.9%)	40 (24.8%)	34 (21%)	31 (18%)	24 (18.2%)	37.8 (23.4%)
Less than 3 pages (2) and less that 100 KB (but more than 0) retrieved	2 (1.3%)	3 (1.9%)	4 (2.5%)	7 (4.1%)	1 (0.8%)	4 (2.5%)

2

G.2 WCAG 1.0 conformance investigation results

Table G.3: **Details regarding Bobby executions**

	May 2003		Dec 2003		May 2004		Dec 2004		Average	
Web sites considered by Bobby	123		111		110		100		111	
Bobby abnormally terminated	3	(2.4%)	3	(2.7%)	3	(2.7%)	1	(1%)	2.5	(2.3%)
Bobby "locked up"	6	(4.9%)	4	(3.6%)	6	(5.5%)	3	(3%)	4.8	(4.3%)
Bobby completed successfully	114	(92.7%)	104	(93.7%)	101	(91.8%)	96	(96%)	103.8	(93.5%)

Table G.4: **Bobby Priority 1 diagnostics triggered during the investigation**

ID		May 2003		Dec 2003		May 2004		Dec 2004		Average	
g9	Provide alternative text for all images	106	(93%)	94	(90.4%)	92	(91.1%)	88	(91.7%)	95	(91.5%)
g38	Each FRAME must reference an HTML file	24	(21.1%)	23	(22.1%)	23	(22.8%)	20	(20.8%)	22.5	(21.7%)
g39	Give each frame a title	25	(21.9%)	23	(22.1%)	22	(21.8%)	20	(20.8%)	22.5	(21.7%)
g240	Provide alternative text for all image map hot-spots	25	(21.9%)	20	(19.2%)	22	(21.8%)	18	(18.8%)	21.3	(20.4%)
g10	Provide alternative text for all image-type buttons in forms	9	(7.9%)	10	(9.6%)	12	(11.9%)	11	(11.5%)	10.5	(10.2%)
g21	Provide alternative text for each APPLET	10	(8.8%)	8	(7.7%)	7	(6.9%)	7	(7.3%)	8	(7.7%)

3

Figure G.1: **Bobby Priority 1 diagnostics**
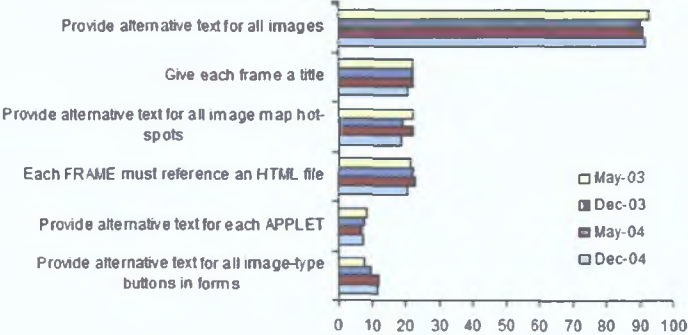
Table G.5: **Density of Bobby Priority 1 diagnostics**

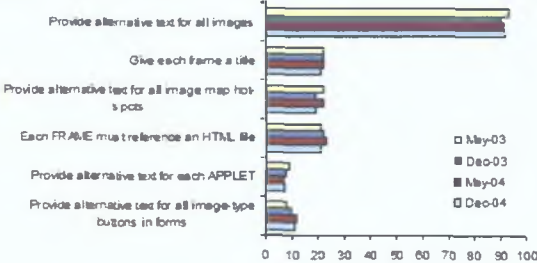|	May 2003	Dec 2003	May 2004	Dec 2004	Average
Web sites with Bobby completed successfully	114	104	101	96	103.8
Web sites that triggered 0 distinct diagnostics	7 (6.1%)	7 (6.7%)	7 (6.9%)	7 (7.3%)	7 (6.7%)
Web sites that triggered 1 distinct diagnostic	55 (48.2%)	50 (48.1%)	48 (47.5%)	46 (47.9%)	49.8 (48%)
Web sites that triggered 2 distinct diagnostics	25 (21.9%)	25 (24.0%)	21 (20.8%)	22 (22.9%)	23.3 (22.4%)
Web sites that triggered 3 distinct diagnostics	16 (14%)	12 (11.5%)	15 (14.9%)	12 (12.5%)	13.8 (13.3%)
Web sites that triggered 4 distinct diagnostics	9 (7.9%)	8 (7.7%)	7 (6.9%)	7 (7.3%)	7.8 (7.5%)
Web sites that triggered 5 distinct diagnostics	2 (1.8%)	2 (1.9%)	3 (3%)	2 (2.1%)	2.3 (2.2%)

Figure G.2: **Density of Bobby Priority 1 diagnostics**

Table G.6: **Bobby Priority 2 diagnostics triggered during the investigation**

ID		May 2003		Dec 2003		May 2004		Dec 2004		Average	
g104	Use relative sizing and positioning (% value) rather than absolute (pixels)	112	(98.2%)	102	(98.1%)	99	(98%)	94	(97.9%)	101.8	(98.1%)
g271	Use a public text identifier in a DOCTYPE statement	100	(87.7%)	83	(79.8%)	76	(75.2%)	74	(77.1%)	83.3	(80%)
g269	Make sure event handlers do not require use of a mouse	85	(74.6%)	76	(73.1%)	71	(70.3%)	66	(68.8%)	74.5	(71.7%)
g41	Explicitly associate form controls and their labels with the LABEL element	71	(62.3%)	58	(55.8%)	53	(52.5%)	47	(49%)	57.3	(54.9%)
g34	Create link phrases that make sense when read out of context	41	(36%)	43	(41.3%)	41	(40.6%)	35	(36.5%)	40	(38.6%)
g273	Include a document TITLE	18	(15.8%)	20	(19.2%)	20	(19.8%)	17	(17.7%)	18.8	(18.1%)
g2	Nest headings properly	17	(14.9%)	20	(19.2%)	17	(16.8%)	13	(13.5%)	16.8	(16.1%)
g5	Avoid scrolling text created with the MARQUEE element	9	(7.9%)	12	(11.5%)	13	(12.9%)	11	(11.5%)	11.3	(10.9%)
g4	Avoid blinking text created with the BLINK element	2	(1.8%)	3	(2.9%)	2	(2%)	1	(1%)	2	(1.9%)
g37	Provide a NOFRAMES section when using FRAMES	3	(2.6%)	3	(2.9%)	1	(1%)	1	(1%)	2	(1.9%)
g33	Do not cause a page to refresh automatically	1	(0.9%)	2	(1.9%)	1	(1%)	1	(1%)	1.3	(1.2%)
g254	Do not cause a page to redirect to a new URL	4	(3.5%)	0	(0%)	1	(1%)	0	(0%)	1.3	(1.1%)

Figure G.3: **Bobby Priority 2 diagnostics**

Categories (top to bottom):
- Use relative sizing and positioning (% value) rather than absolute (pixels)
- Use a public text identifier in a DOCTYPE statement
- Make sure event handlers do not require use of a mouse
- Explicitly associate form controls and their labels with the LABEL element
- Create link phrases that make sense when read out of context
- Include a document TITLE
- Nest headings properly
- Avoid scrolling text created with the MARQUEE element
- Do not cause a page to redirect to a new URL
- Provide a NOFRAMES section when using FRAMES
- Avoid blinking text created with the BLINK element
- Do not cause a page to refresh automatically

Legend: □ May-03 ■ Dec-03 ■ May-04 □ Dec-04

Figure G.4: **Density of Bobby Priority 2 diagnostics**

Categories (top to bottom):
- 8 distinct diagnostics
- 7 distinct diagnostics
- 6 distinct diagnostics
- 5 distinct diagnostics
- 4 distinct diagnostics
- 3 distinct diagnostics
- 2 distinct diagnostics
- 1 distinct diagnostic
- 0 distinct diagnostics

Legend: □ May-03 ■ Dec-03 ■ May-04 □ Dec-04

6
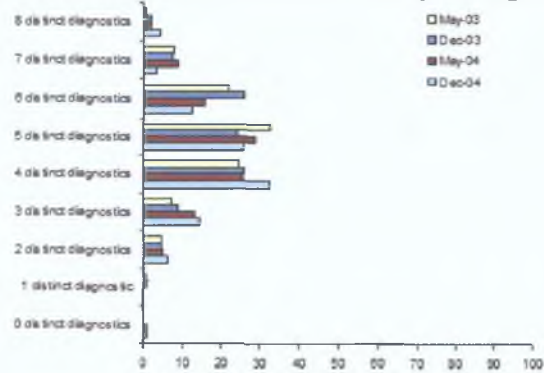
Table G.7: **Density of Bobby Priority 2 diagnostics**

	May 2003		Dec 2003		May 2004		Dec 2004		Average	
Web sites for which Bobby completed successfully	114		104		101		96		103.8	
Web sites that triggered 0 distinct diagnostics	0	(0%)	0	(0%)	1	(1%)	1	(1%)	0.5	(0.5%)
Web sites that triggered 1 distinct diagnostic	1	(0.9%)	1	(1%)	0	(0%)	0	(0%)	0.5	(0.5%)
Web sites that triggered 2 distinct diagnostics	5	(4.4%)	5	(4.8%)	5	(5%)	6	(6.3%)	5.3	(5.1%)
Web sites that triggered 3 distinct diagnostics	8	(7%)	9	(8.7%)	13	(12.9%)	14	(14.6%)	11	(10.6%)
Web sites that triggered 4 distinct diagnostics	28	(24.6%)	27	(26%)	26	(25.7%)	31	(32.3%)	28	(27%)
Web sites that triggered 5 distinct diagnostics	37	(32.5%)	25	(24%)	29	(28.7%)	25	(26%)	29	(28%)
Web sites that triggered 6 distinct diagnostics	25	(21.9%)	27	(26%)	16	(15.8%)	12	(12.5%)	20	(19.3%)
Web sites that triggered 7 distinct diagnostics	9	(7.9%)	8	(7.7%)	9	(8.9%)	3	(3.1%)	7.3	(7%)
Web sites that triggered 8 distinct diagnostics	1	(0.9%)	2	(1.9%)	2	(2%)	4	(4.2%)	2.3	(2.2%)
Web sites that triggered 9 distinct diagnostics	0	(0%)	0	(0%)	0	(0%)	0	(0%)	0	(0%)
Web sites that triggered 10 distinct diagnostics	0	(0%)	0	(0%)	0	(0%)	0	(0%)	0	(0%)

G.3 HTML technical standards conformance investigation results

Table G.8: **Details regarding OpenSP executions**

	May 2003	Dec 2003	May 2004	Dec 2004	Average
Web sites processed by OpenSP	123	111	110	100	111
Web sites where all the web pages had usable document type information	2 (1.6%)	2 (1.8%)	1 (0.9%)	1 (1%)	1.5 (1.4%)
Web sites considered for HTML defects analysis	36 (29.3%)	39 (35.1%)	45 (40.9%)	49 (49%)	42.3 (38.1%)

Table G.9: **Document Type Declaration in web pages considered by OpenSP**

	May 2003	Dec 2003	May 2004	Dec 2004	Average
Web pages processed by OpenSP	2,288	2,064	1,966	1,769	2021.8
No document type information	1,732 (75.7%)	1,435 (69.5%)	1,310 (66.6%)	1,071 (60.5%)	1387 (68.6%)
Incorrect document type information	42 (1.8%)	78 (3.8%)	70 (3.6%)	29 (1.6%)	54.8 (2.7%)
Usable document type information	465 (20.3%)	509 (24.7%)	549 (27.9%)	632 (35.7%)	538.8 (26.6%)
Considered for HTML defects analysis (in web sites with at least 3 web pages and 100 KB of data with usable document type information)	454 (19.8%)	500 (24.2%)	542 (27.6%)	623 (35.2%)	529.8 (26.2%)

Table G.10: **The 10 most common OpenSP diagnostics triggered in the investigation**

ID		May 2003		Dec 2003		May 2004		Dec 2004		Average	
108	Undefined attribute	33	(91.7%)	36	(92.3%)	42	(93.3%)	44	(89.8%)	38.8	(91.8%)
64	Element not allowed by document type	29	(80.6%)	31	(79.5%)	35	(77.8%)	36	(73.5%)	32.8	(77.8%)
79	End tag for not opened element	27	(75%)	28	(71.8%)	32	(71.1%)	36	(73.5%)	30.8	(72.8%)
127	Required attribute not specified	19	(52.8%)	28	(71.8%)	34	(75.6%)	39	(79.6%)	30	(69.9%)
139	Non SGML character number	20	(55.6%)	24	(61.5%)	27	(60%)	33	(67.3%)	26	(61.1%)
68	Missing required end tag (element DTD)	20	(55.6%)	22	(56.4%)	28	(62.2%)	24	(49%)	23.5	(55.8%)
25	General entity not defined and no default entity	17	(47.2%)	20	(51.3%)	22	(48.9%)	25	(51%)	21	(49.6%)
82	Attribute value must be literal	16	(44.4%)	18	(46.2%)	23	(51.1%)	24	(49%)	20.3	(47.7%)
131	Attribute value not allowed	10	(27.8%)	15	(38.5%)	21	(46.7%)	21	(42.9%)	16.8	(38.9%)
73	End tag for element not finished	11	(30.6%)	16	(41%)	16	(35.6%)	17	(34.7%)	15	(35.5%)

Figure G.5: **The 10 most common OpenSP diagnostics**
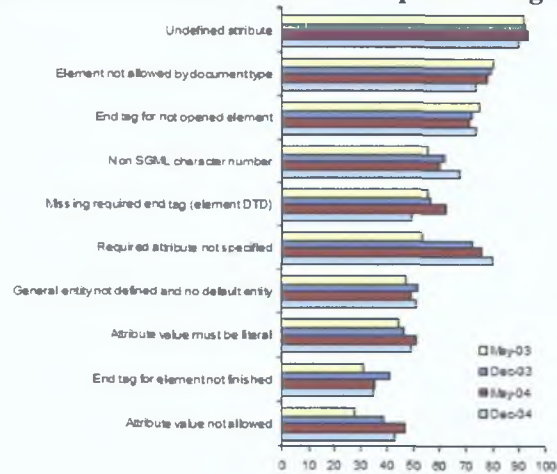
9

Table G.11: **Density of OpenSP diagnostics**

	May 2003		Dec 2003		May 2004		Dec 2004		Average	
Web sites considered for HTML defects analysis	36		39		45		49		42.3	
Web sites that triggered 0 distinct diagnostics	1	(2.8%)	1	(2.6%)	1	(2.2%)	1	(2%)	1	(2.4%)
Web sites that triggered 1 distinct diagnostic	1	(2.8%)	1	(2.6%)	1	(2.2%)	2	(4.1%)	1.3	(3%)
Web sites that triggered 2 distinct diagnostics	2	(5.6%)	1	(2.6%)	1	(2.2%)	2	(4.1%)	1.5	(3.6%)
Web sites that triggered 3 distinct diagnostics	1	(2.8%)	1	(2.6%)	2	(4.4%)	2	(4.1%)	1.5	(3.6%)
Web sites that triggered 4 distinct diagnostics	2	(5.6%)	0	(0%)	0	(0%)	1	(2%)	0.8	(1.8%)
Web sites that triggered 5 distinct diagnostics	1	(2.8%)	3	(7.7%)	4	(8.9%)	2	(4.1%)	2.5	(5.9%)
Web sites that triggered 6 distinct diagnostics	1	(2.8%)	1	(2.6%)	1	(2.2%)	3	(6.1%)	1.5	(3.6%)
Web sites that triggered 7 distinct diagnostics	3	(8.3%)	3	(7.7%)	3	(6.7%)	0	(0%)	2.3	(5.3%)
Web sites that triggered 8 distinct diagnostics	3	(8.3%)	1	(2.6%)	2	(4.4%)	4	(8.2%)	2.5	(5.9%)
Web sites that triggered 9 distinct diagnostics	1	(2.8%)	2	(5.1%)	2	(4.4%)	1	(2%)	1.5	(3.6%)
Web sites that triggered 10 distinct diagnostics	5	(13.9%)	6	(15.4%)	5	(11.1%)	7	(14.3%)	5.8	(13.6%)
Web sites that triggered 11 distinct diagnostics	5	(13.9%)	6	(15.4%)	5	(11.1%)	5	(10.2%)	5.3	(12.4%)
Web sites that triggered 12 distinct diagnostics	4	(11.1%)	3	(7.7%)	5	(11.1%)	6	(12.2%)	4.5	(10.7%)
Web sites that triggered 13 distinct diagnostics	2	(5.6%)	1	(2.6%)	3	(6.7%)	3	(6.1%)	2.3	(5.3%)
Web sites that triggered 14 distinct diagnostics	2	(5.6%)	2	(5.1%)	7	(15.6%)	3	(6.1%)	3.5	(8.3%)
Web sites that triggered 15 distinct diagnostics	0	(0%)	4	(10.3%)	1	(2.2%)	4	(8.2%)	2.3	(5.3%)
Web sites that triggered 16 distinct diagnostics	1	(2.8%)	0	(0%)	1	(2.2%)	2	(4.1%)	1	(2.4%)

10

Table G.11: **Density of OpenSP diagnostics**

	May 2003		Dec 2003		May 2004		Dec 2004		Average	
Web sites that triggered 17 distinct diagnostics	0	(0%)	1	(2.6%)	0	(0%)	0	(0%)	0.3	(0.6%)
Web sites that triggered 18 distinct diagnostics	0	(0%)	1	(2.6%)	0	(0%)	1	(2%)	0.5	(1.2%)
Web sites that triggered 19 distinct diagnostics	0	(0%)	1	(2.6%)	0	(0%)	0	(0%)	0.3	(0.6%)
Web sites that triggered 20 distinct diagnostics	0	(0%)	0	(0%)	1	(2.2%)	0	(0%)	0.3	(0.6%)
Web sites that triggered 21 distinct diagnostics	0	(0%)	0	(0%)	0	(0%)	0	(0%)	0	(0%)
Web sites that triggered 22 distinct diagnostics	1	(2.8%)	0	(0%)	0	(0%)	0	(0%)	0.3	(0.6%)

Figure G.6: **Density of OpenSP diagnostics**

Appendix H

Results of surveys conducted on the web sample of UK web sites (May 2003–December 2004)

H.1   Web content mirroring process results

Table H.1: **Results of the web content sampling process**

	May 2003	**Dec 2003**	**May 2004**	**Dec 2004**	**Average**
Data amount retrieved	552 MB	527 MB	519 MB	500 MB	524 MB
Web pages considered for mirroring	67,598	62,491	60,440	57,022	61,887.8
Web pages correctly retrieved (size>0 and no retrieval errors)	59,248 (87.6%)	54,135 (86.6%)	52,260 (86.5%)	48,726 (85.5%)	53,592.3 (86.6%)
Web pages qualified for conformance investigation	45,857 (67.8%)	42,189 (67.5%)	40,769 (67.5%)	38,132 (66.9%)	41,736.8 (67.4%)
Web sites qualified for conformance investigation	2,380 (41.7%)	2,269 (39.8%)	2,243 (39.3%)	2,142 (37.6%)	2,258.5 (39.6%)

I

Table H.2: **Details regarding the web sites that didn't qualify for investigations**

	May 2003	Dec 2003	May 2004	Dec 2004	Common	Average
Web sites not qualified for investigations	3,322	3,434	3,459	3,560	2,915	3,443.8
No data retrieved	270 (8.1%)	417 (12.1%)	534 (15.4%)	668 (18.8%)	138 (4.7%)	472.3 (13.7%)
Only the main page retrieved (size>0)	1,669 (50.2%)	1,770 (51.5%)	1,749 (50.6%)	1,815 (51%)	1,260 (43.2%)	1,750.8 (50.8%)
Less than 3 pages (2) and at least 100 KB retrieved	4 (0.1%)	3 (0.1%)	5 (0.1%)	4 (0.1%)	1 (0%)	4 (0.1%)
Less than 100 KB and at least 3 pages retrieved	1,211 (36.5%)	1,084 (31.6%)	1,004 (29%)	926 (26%)	675 (23.2%)	1,056.3 (30.7%)
Less than 3 pages (2) and less that 100 KB (but more than 0) retrieved	168 (5.1%)	158 (4.6%)	167 (4.8%)	147 (4.1%)	73 (2.5%)	160 (4.6%)

2

H.2   WCAG 1.0 conformance investigation results

Table H.3: **Details regarding Bobby executions**

	May 2003	Dec 2003	May 2004	Dec 2004	Average
Web sites considered by Bobby	2,380	2,269	2,243	2,142	2,258.5
Bobby abnormally terminated	137 (5.8%)	133 (5.9%)	123 (5.5%)	113 (5.3%)	126.5 (5.6%)
Bobby "locked up"	92 (3.9%)	76 (3.3%)	83 (3.7%)	101 (4.7%)	88 (3.9%)
Bobby completed successfully	2,151 (90.4%)	2,059 (90.7%)	2,036 (90.8%)	1,928 (90%)	2,043.5 (90.5%)

Table H.4: **Bobby Priority 1 diagnostics triggered during the investigation**

ID		May 2003	Dec 2003	May 2004	Dec 2004	Average
g9	Provide alternative text for all images	1,989 (92.5%)	1,896 (92.1%)	1,840 (90.4%)	1,685 (87.4%)	1,852.5 (90.6%)
g39	Give each frame a title	579 (26.9%)	511 (24.8%)	491 (24.1%)	421 (21.8%)	500.5 (24.4%)
g38	Each FRAME must reference an HTML file	548 (25.5%)	477 (23.2%)	452 (22.2%)	378 (19.6%)	463.8 (22.6%)
g240	Provide alternative text for all image map hot-spots	429 (19.9%)	401 (19.5%)	372 (18.3%)	307 (15.9%)	377.3 (18.4%)
g10	Provide alternative text for all image-type buttons in forms	183 (8.5%)	170 (8.3%)	171 (8.4%)	149 (7.7%)	168.3 (8.2%)
g21	Provide alternative text for each applet	158 (7.3%)	134 (6.5%)	116 (5.7%)	98 (5.1%)	126.5 (6.2%)
g20	Provide alternative content for each OBJECT	5 (0.2%)	3 (0.1%)	2 (0.1%)	4 (0.2%)	3.5 (0.2%)

Figure H.1: **Bobby Priority 1 diagnostics**
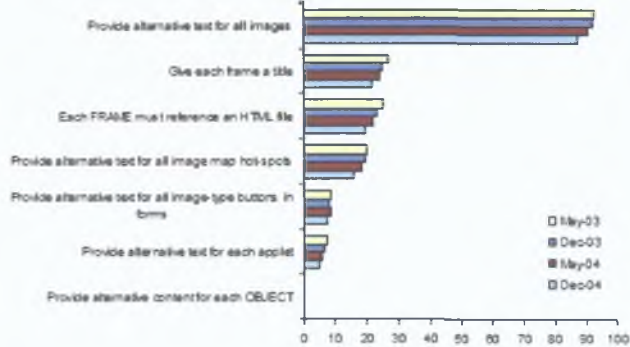
Table H.5: **Density of Bobby Priority 1 diagnostics**

	May 2003	Dec 2003	May 2004	Dec 2004	Average
Web sites for which Bobby completed successfully	2,151	2,059	2,036	1,928	2,043.5
Web sites for which were triggered 0 distinct diagnostic	119 (5.5%)	123 (6%)	157 (7.7%)	191 (9.9%)	147.5 (7.3%)
Web sites for which were triggered 1 distinct diagnostic	1,021 (47.5%)	1,008 (49%)	991 (48.7%)	972 (50.4%)	998 (48.9%)
Web sites for which were triggered 2 distinct diagnostic	415 (19.3%)	400 (19.4%)	388 (19.1%)	354 (18.4%)	389.3 (19%)
Web sites for which were triggered 3 distinct diagnostics	377 (17.5%)	352 (17.1%)	344 (16.9%)	294 (15.2%)	341.8 (16.7%)
Web sites for which were triggered 4 distinct diagnostics	186 (8.6%)	152 (7.4%)	135 (6.6%)	105 (5.4%)	144.5 (7%)
Web sites for which were triggered 5 distinct diagnostics	33 (1.5%)	24 (1.2%)	21 (1%)	12 (0.6%)	22.5 (1.1%)

Figure H.2: **Density of Bobby Priority 1 diagnostics**

4
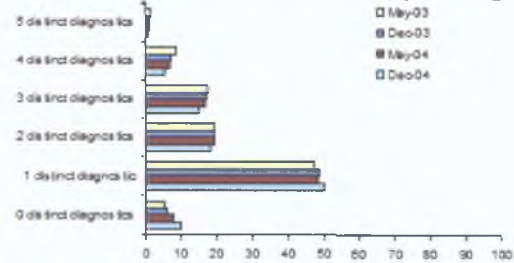
Table H.6: **Bobby Priority 2 diagnostics triggered during the investigation**

ID		May 2003	Dec 2003	May 2004	Dec 2004	Average	
g104	Use relative sizing and po-sitioning (% value) rather than absolute (pixels)	2,113 (98.2%)	2,020 (98.1%)	1,983 (97.4%)	1,852 (96.1%)	1,992	(97.4%)
g271	Use a public text identifier in a DOCTYPE statement	1,775 (82.5%)	1,631 (79.2%)	1,553 (76.3%)	1,390 (72.1%)	1,587.3 (77.5%)	
g269	Make sure event handlers do not require use of a mouse	1,404 (65.3%)	1,356 (65.9%)	1,347 (66.2%)	1,225 (63.5%)	1,333	(65.2%)
g41	Explicitly associate form controls and their labels with the LABEL element	1,368 (63.6%)	1,303 (63.3%)	1,269 (62.3%)	1,192 (61.8%)	1,283	(62.8%)
g34	Create link phrases that make sense when read out of context	840 (39.1%)	822 (39.9%)	797 (39.1%)	733 (38%)	798	(39%)
g2	Nest headings properly	298 (13.9%)	263 (12.8%)	243 (11.9%)	255 (13.2%)	264.8	(12.9%)
g273	Include a document TI-TLE	191 (8.9%)	185 (9%)	186 (9.1%)	203 (10.5%)	191.3	(9.4%)
g5	Avoid scrolling text cre-ated with the MARQUEE element	178 (8.3%)	158 (7.7%)	149 (7.3%)	127 (6.6%)	153	(7.5%)
g254	Do not cause a page to redirect to a new URL	62 (2.9%)	53 (2.6%)	51 (2.5%)	33 (1.7%)	49.8	(2.4%)
g37	Provide a NOFRAMES section when using FRAMES	46 (2.1%)	48 (2.3%)	47 (2.3%)	40 (2.1%)	45.3	(2.2%)
g4	Avoid blinking text cre-ated with the BLINK ele-ment	55 (2.6%)	43 (2.1%)	40 (2%)	38 (2%)	44	(2.1%)
g33	Do not cause a page to re-fresh automatically	35 (1.6%)	29 (1.4%)	34 (1.7%)	29 (1.5%)	31.8	(1.6%)

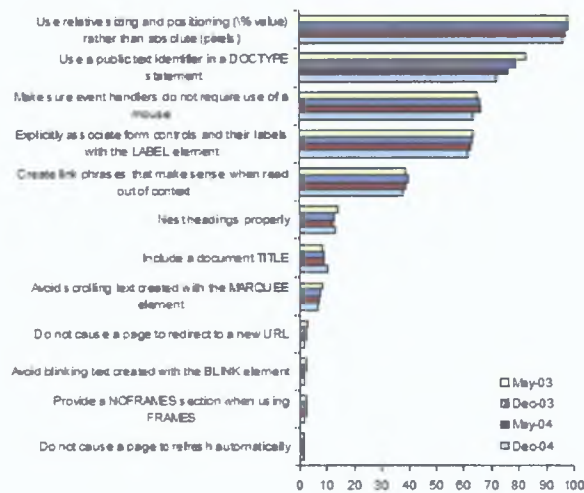## Figure H.3: **Bobby Priority 2 diagnostics**
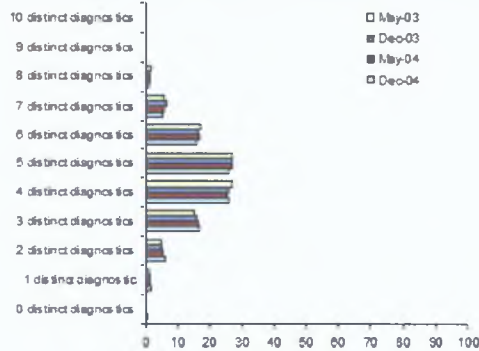

Figure H.4: **Density of Bobby Priority 2 diagnostics**

Table H.7: **Density of Bobby Priority 2 diagnostics**

	May 2003		Dec 2003		May 2004		Dec 2004		Average	
Web sites for which Bobby completed successfully	2,151		2,059		2,036		1,928		2,044	
Web sites that triggered 0 distinct diagnostics	4	(0.2%)	4	(0.2%)	6	(0.3%)	10	(0.5%)	6	(0.3%)
Web sites that triggered 1 distinct diagnostic	22	(1%)	24	(1.2%)	24	(1.2%)	33	(1.7%)	26	(1.3%)
Web sites that triggered 2 distinct diagnostics	100	(4.6%)	99	(4.8%)	113	(5.6%)	115	(6%)	107	(5.2%)
Web sites that triggered 3 distinct diagnostics	329	(15.3%)	328	(15.9%)	336	(16.5%)	324	(16.8%)	329	(16.1%)
Web sites that triggered 4 distinct diagnostics	576	(26.8%)	535	(26%)	522	(25.6%)	501	(26%)	534	(26.1%)
Web sites that triggered 5 distinct diagnostics	577	(26.8%)	562	(27.3%)	548	(26.9%)	502	(26%)	547	(26.8%)
Web sites that triggered 6 distinct diagnostics	374	(17.4%)	339	(16.5%)	342	(16.8%)	312	(16.2%)	342	(16.7%)
Web sites that triggered 7 distinct diagnostics	125	(5.8%)	134	(6.5%)	115	(5.6%)	105	(5.4%)	120	(5.9%)
Web sites that triggered 8 distinct diagnostics	34	(1.6%)	28	(1.4%)	24	(1.2%)	20	(1%)	27	(1.3%)
Web sites that triggered 9 distinct diagnostics	9	(0.4%)	5	(0.2%)	5	(0.2%)	5	(0.3%)	6	(0.3%)
Web sites that triggered 10 distinct diagnostics	1	(0%)	1	(0%)	1	(0%)	1	(0.1%)	1	(0%)

H.3   HTML technical standards conformance investigation results

Table H.8: **Details regarding OpenSP executions**

	May 2003	Dec 2003	May 2004	Dec 2004	Average
Web sites processed by OpenSP	2,380	2,269	2,243	2,142	2,258.5
Web sites where all the web pages had usable document type information	35 (1.5%)	50 (2.2%)	52 (2.3%)	60 (2.8%)	49.3 (2.2%)
Web sites considered for HTML defects analysis	663 (27.9%)	720 (31.7%)	792 (35.3%)	904 (42.2%)	769.8 (34.1%)

Table H.9: **Document Type Declaration in web pages considered by OpenSP**

	May 2003	Dec 2003	May 2004	Dec 2004	Average
Web pages processed by OpenSP	45,857	42,189	40,769	38,132	41,736.8
No document type information	34,564 (75.4%)	30,528 (72.4%)	28,023 (68.7%)	23,637 (62%)	29,188 (69.9%)
Incorrect document type information	965 (2.1%)	873 (2.1%)	843 (2.1%)	796 (2.1%)	869.3 (2.1%)
Usable document type information	9,562 (20.9%)	10,192 (24.2%)	11,409 (28%)	13,209 (34.6%)	11,093 (26.6%)
Considered for HTML defects analysis (in web sites with at least 3 web pages and 100 KB of data with usable document type information)	9,291 (20.3%)	9,916 (23.5%)	11,146 (27.3%)	12,980 (34%)	10,833.3(26%)

Table H.10: **The 10 most common OpenSP diagnostics triggered in the investigation**

ID		May 2003		Dec 2003		May 2004		Dec 2004		Average	
108	Undefined attribute	600	(90.5%)	631	(87.6%)	687	(86.7%)	713	(78.9%)	658	(85.9%)
127	Required attribute not specified	496	(74.8%)	560	(77.8%)	605	(76.4%)	634	(70.1%)	574	(74.8%)
64	Element not allowed by document type	509	(76.8%)	520	(72.2%)	559	(70.6%)	584	(64.6%)	543	(71%)
79	End tag for not opened element	487	(73.5%)	506	(70.3%)	539	(68.1%)	563	(62.3%)	524	(68.5%)
25	General entity not defined and no default entity	326	(49.2%)	349	(48.5%)	395	(49.9%)	428	(47.3%)	375	(48.7%)
68	Missing required end tag (element DTD)	351	(52.9%)	342	(47.5%)	362	(45.7%)	366	(40.5%)	355	(46.7%)
139	Non SGML character number	270	(40.7%)	285	(39.6%)	332	(41.9%)	362	(40%)	312	(40.6%)
131	Attribute value not allowed	261	(39.4%)	282	(39.2%)	285	(36%)	290	(32.1%)	280	(36.6%)
73	End tag for element not finished	254	(38.3%)	270	(37.5%)	285	(36%)	304	(33.6%)	278	(36.4%)
82	Attribute value must be literal	251	(37.9%)	261	(36.3%)	268	(33.8%)	255	(28.2%)	259	(34%)

Figure H.5: **The 10 most common OpenSP diagnostics**

9
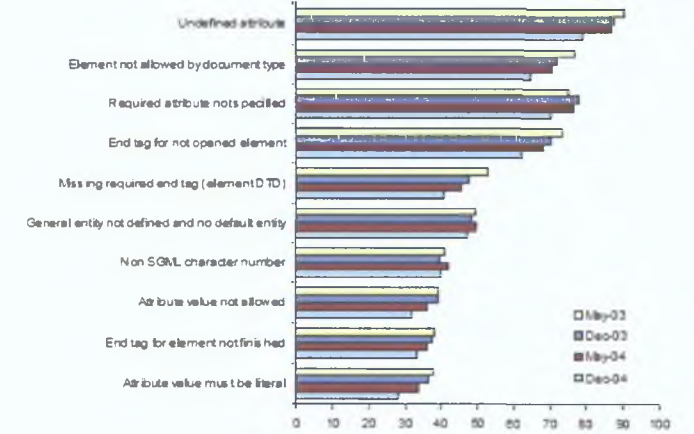
Table H.11: **Density of OpenSP diagnostics**

	May 2003		Dec 2003		May 2004		Dec 2004		Average	
Web sites considered for further analysis	663		720		792		904		769.8	
Web sites that triggered 0 distinct diagnostics	29	(4.4%)	37	(5.1%)	41	(5.2%)	104	(11.5%)	52.8	(6.9%)
Web sites that triggered 1 distinct diagnostic	8	(1.2%)	8	(1.1%)	11	(1.4%)	10	(1.1%)	9.3	(1.2%)
Web sites that triggered 2 distinct diagnostics	20	(3%)	17	(2.4%)	14	(1.8%)	34	(3.8%)	21.3	(2.8%)
Web sites that triggered 3 distinct diagnostics	19	(2.9%)	29	(4%)	36	(4.5%)	39	(4.3%)	30.8	(4%)
Web sites that triggered 4 distinct diagnostics	37	(5.6%)	32	(4.4%)	45	(5.7%)	43	(4.8%)	39.3	(5.1%)
Web sites that triggered 5 distinct diagnostics	46	(6.9%)	48	(6.7%)	56	(7.1%)	50	(5.5%)	50	(6.5%)
Web sites that triggered 6 distinct diagnostics	40	(6%)	55	(7.6%)	61	(7.7%)	60	(6.6%)	54	(7%)
Web sites that triggered 7 distinct diagnostics	58	(8.7%)	67	(9.3%)	73	(9.2%)	65	(7.2%)	65.8	(8.5%)
Web sites that triggered 8 distinct diagnostics	51	(7.7%)	59	(8.2%)	62	(7.8%)	73	(8.1%)	61.3	(8%)
Web sites that triggered 9 distinct diagnostics	56	(8.4%)	60	(8.3%)	69	(8.7%)	76	(8.4%)	65.3	(8.5%)
Web sites that triggered 10 distinct diagnostics	54	(8.1%)	63	(8.8%)	68	(8.6%)	72	(8%)	64.3	(8.3%)
Web sites that triggered 11 distinct diagnostics	60	(9%)	55	(7.6%)	50	(6.3%)	44	(4.9%)	52.3	(6.8%)
Web sites that triggered 12 distinct diagnostics	52	(7.8%)	54	(7.5%)	47	(5.9%)	64	(7.1%)	54.3	(7%)
Web sites that triggered 13 distinct diagnostics	37	(5.6%)	44	(6.1%)	50	(6.3%)	50	(5.5%)	45.3	(5.9%)
Web sites that triggered 14 distinct diagnostics	38	(5.7%)	32	(4.4%)	34	(4.3%)	42	(4.6%)	36.5	(4.7%)
Web sites that triggered 15 distinct diagnostics	19	(2.9%)	24	(3.3%)	35	(4.4%)	23	(2.5%)	25.3	(3.3%)
Web sites that triggered 16 distinct diagnostics	20	(3%)	17	(2.4%)	19	(2.4%)	25	(2.8%)	20.3	(2.6%)

10

Web sites that triggered 17 distinct diagnostics	8	(1.2%)	8	(1.1%)	9	(1.1%)	11	(1.2%)	9	(1.2%)
Web sites that triggered 18 distinct diagnostics	7	(1.1%)	4	(0.6%)	6	(0.8%)	7	(0.8%)	6	(0.8%)
Web sites that triggered 19 distinct diagnostics	2	(0.3%)	5	(0.7%)	4	(0.5%)	5	(0.6%)	4	(0.5%)
Web sites that triggered 20 distinct diagnostics	1	(0.2%)	2	(0.3%)	1	(0.1%)	5	(0.6%)	2.3	(0.3%)
Web sites that triggered 21 distinct diagnostics	0	(0%)	0	(0%)	1	(0.1%)	0	(0%)	0.3	(0%)
Web sites that triggered 22 distinct diagnostics	1	(0.2%)	0	(0%)	0	(0%)	0	(0%)	0.3	(0%)
Web sites that triggered 23 distinct diagnostics	0	(0%)	0	(0%)	0	(0%)	2	(0.2%)	0.5	(0.1%)
Web sites that triggered 24 distinct diagnostics	0	(0%)	0	(0%)	0	(0%)	0	(0%)	0	(0%)

Figure H.6: **Density of OpenSP diagnostics**

24 distinct diagnostics
23 distinct diagnostics
22 distinct diagnostics
21 distinct diagnostics
20 distinct diagnostics
19 distinct diagnostics
18 distinct diagnostics
17 distinct diagnostics
16 distinct diagnostics
15 distinct diagnostics
14 distinct diagnostics
13 distinct diagnostics
12 distinct diagnostics
11 distinct diagnostics
10 distinct diagnostics
9 distinct diagnostics
8 distinct diagnostics
7 distinct diagnostics
6 distinct diagnostics
5 distinct diagnostics
4 distinct diagnostics
3 distinct diagnostics
2 distinct diagnostics
1 distinct diagnostic
0 distinct diagnostics

☐ May-03
■ Dec-03
■ May-04
☐ Dec-04

0 10 20 30 40 50 60 70 80 90 100

12

Appendix I

Results of surveys conducted on the web sample of French web sites (May 2003–December 2004)

I.1   Web content mirroring process results

Table I.1: **Results of the web content sampling process**

	May 2003	Dec 2003	May 2004	Dec 2004	Average
Data amount retrieved	167 MB	162 MB	157 MB	145 MB	158 MB
Web pages considered for mirroring	22,319	20,406	19,148	17,056	19,732.3
Web pages correctly retrieved (size>0 and no retrieval errors)	19,727 (88.4%)	17,956 (88%)	16,388 (85.6%)	14,364 (84.2%)	17,108.8 (86.7%)
Web pages in web sites qualified for conformance investigations	16,312 (73.1%)	14,874 (72.9%)	13,461 (70.3%)	11,644 (68.3%)	14,072.8 (71.3%)
Web sites qualified for conformance investigations	747 (48.3%)	722 (46.7%)	698 (45.2%)	640 (41.4%)	701.8 (45.4%)

Table I.2: **Details regarding the web sites that didn't qualify for investigations**

	May 2003	Dec 2003	May 2004	Dec 2004	Common	Average
Web sites not qualified for investigations	798	823	847	905	638	843
No data retrieved	106 (13.3%)	124 (15.1%)	178 (21%)	270 (29.8%)	51 (8%)	170 (20.1%)
Only the main page retrieved (size>0)	328 (41.1%)	358 (43.5%)	361 (42.6%)	362 (40%)	214 (33.5%)	352 (41.8%)
Less than 3 pages (2) and at least 100 KB retrieved	4 (0.5%)	3 (0.4%)	1 (0.1%)	0 (0%)	0 (0%)	2 (0.2%)
Less than 100 KB and at least 3 pages retrieved	301 (37.7%)	285 (34.6%)	258 (30.5%)	236 (26.1%)	144 (22.6%)	270 (32%)
Less than 3 pages (2) and less that 100 KB (but more than 0) retrieved	59 (7.4%)	53 (6.4%)	45 (5.3%)	37 (4.1%)	14 (2.2%)	49 (5.8%)

I.2 WCAG 1.0 conformance investigation results

Table I.3: **Details regarding Bobby executions**

	May 2003	Dec 2003	May 2004	Dec 2004	Average
Web sites considered by Bobby	747	722	698	640	702
Bobby abnormally terminated	63 (8.4%)	51 (7.1%)	51 (7.3%)	40 (6.3%)	51 (7.3%)
Bobby "locked up"	45 (6%)	28 (3.9%)	25 (3.6%)	52 (8.1%)	38 (5.3%)
Bobby completed successfully	639 (85.5%)	643 (89.1%)	622 (89.1%)	548 (85.6%)	613 (87.4%)

Table I.4: **Bobby Priority 1 diagnostics triggered during the investigation**

ID		May 2003	Dec 2003	May 2004	Dec 2004	Average
g9	Provide alternative text for all images	625 (97.8%)	627 (97.5%)	603 (96.9%)	530 (96.7%)	596.3 (97.2%)
g39	Give each frame a title	263 (41.2%)	257 (40%)	235 (37.8%)	211 (38.5%)	241.5 (39.4%)
g240	Provide alternative text for all image map hot-spots	241 (37.7%)	235 (36.5%)	227 (36.5%)	209 (38.1%)	228 (37.2%)
g38	Each FRAME must reference an HTML file	251 (39.3%)	238 (37%)	217 (34.9%)	194 (35.4%)	225 (36.6%)
g10	Provide alternative text for all image-type buttons in forms	82 (12.8%)	85 (13.2%)	79 (12.7%)	74 (13.5%)	80 (13.1%)
g21	Provide alternative text for each APPLET	45 (7%)	38 (5.9%)	34 (5.5%)	32 (5.8%)	37.3 (6.1%)
g20	Provide alternative content for each OBJECT	3 (0.5%)	3 (0.5%)	3 (0.5%)	2 (0.4%)	2.8 (0.4%)

Figure I.1: **Bobby Priority 1 diagnostics**
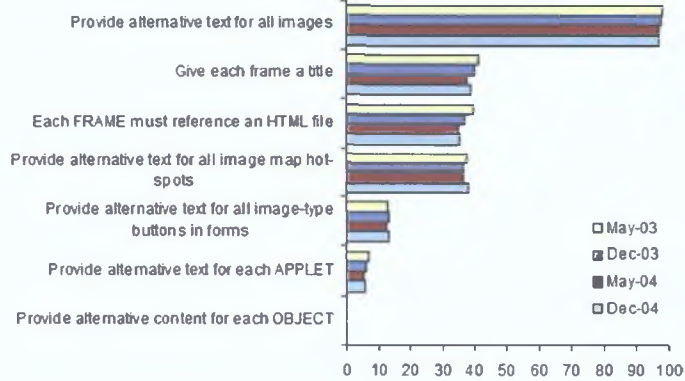
Table I.5: **Density of Bobby Priority 1 diagnostics**

	May 2003	Dec 2003	May 2004	Dec 2004	Average
Web sites for which Bobby completed successfully	639	643	622	548	613
Web sites for which were triggered 0 distinct diagnostic	9 (1.4%)	10 (1.6%)	12 (1.9%)	8 (1.5%)	9.8 (1.6%)
Web sites for which were triggered 1 distinct diagnostic	199 (31.1%)	204 (31.7%)	211 (33.9%)	171 (31.2%)	196.3 (32%)
Web sites for which were triggered 2 distinct diagnostic	155 (24.3%)	166 (25.8%)	156 (25.1%)	152 (27.7%)	157.3 (25.7%)
Web sites for which were triggered 3 distinct diagnostics	135 (21.1%)	131 (20.4%)	119 (19.1%)	112 (20.4%)	124.3 (20.3%)
Web sites for which were triggered 4 distinct diagnostics	110 (17.2%)	107 (16.6%)	102 (16.4%)	84 (15.3%)	100.8 (16.4%)
Web sites for which were triggered 5 distinct diagnostics	30 (4.7%)	24 (3.7%)	22 (3.5%)	21 (3.8%)	24.3 (4%)
Web sites for which were triggered 6 distinct diagnostics	1 (0.2%)	1 (0.2%)	0 (0%)	0 (0%)	0.5 (0.1%)

4

Figure I.2: **Density of Bobby Priority 1 diagnostics**
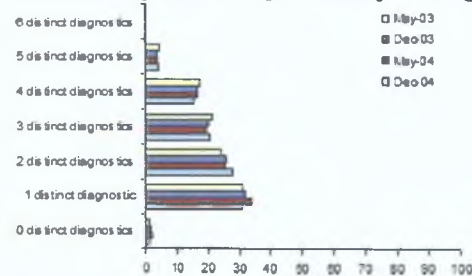
Figure I.3: **Bobby Priority 2 diagnostics**
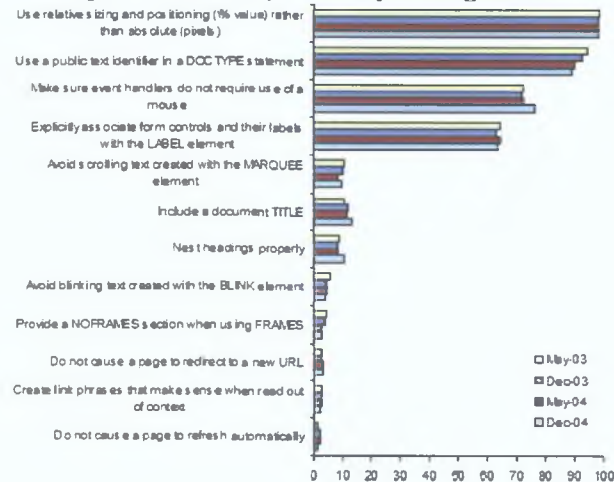
Figure I.4: **Density of Bobby Priority 2 diagnostics**
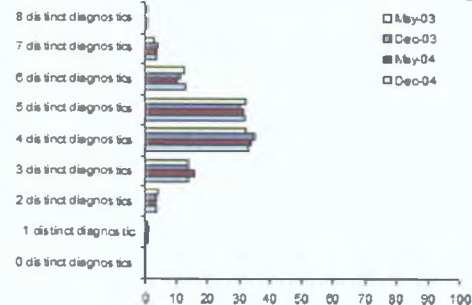
Table I.6: **Bobby Priority 2 diagnostics triggered during the investigation**

ID		May 2003		Dec 2003		May 2004		Dec 2004		Average	
g104	Use relative sizing and positioning (% value) rather than absolute (pixels)	627	(98.1%)	630	(98%)	609	(97.9%)	537	(98%)	600.8 (98%)	
g271	Use a public text identifier in a DOCTYPE statement	602	(94.2%)	596	(92.7%)	560	(90%)	490	(89.4%)	562	(91.6%)
g269	Make sure event handlers do not require use of a mouse	464	(72.6%)	464	(72.2%)	452	(72.7%)	419	(76.5%)	449.8 (73.5%)	
g41	Explicitly associate form controls and their labels with the LABEL element	412	(64.5%)	407	(63.3%)	401	(64.5%)	349	(63.7%)	392.3 (64%)	
g4	Avoid blinking text created with the BLINK element	37	(5.8%)	29	(4.5%)	27	(4.3%)	349	(63.7%)	110.5 (19.6%)	
g273	Include a document TITLE	68	(10.6%)	77	(12%)	71	(11.4%)	72	(13.1%)	72	(11.8%)
g5	Avoid scrolling text created with the MARQUEE element	69	(10.8%)	65	(10.1%)	52	(8.4%)	52	(9.5%)	59.5 (9.7%)	
g2	Nest headings properly	56	(8.8%)	55	(8.6%)	54	(8.7%)	59	(10.8%)	56	(9.2%)
g37	Provide a NOFRAMES section when using FRAMES	30	(4.7%)	25	(3.9%)	16	(2.6%)	14	(2.6%)	21.3 (3.4%)	
g254	Do not cause a page to redirect to a new URL	19	(3%)	16	(2.5%)	19	(3.1%)	17	(3.1%)	17.8 (2.9%)	
g34	Create link phrases that make sense when read out of context	15	(2.3%)	15	(2.3%)	17	(2.7%)	12	(2.2%)	14.8 (2.4%)	
g33	Do not cause a page to refresh automatically	9	(1.4%)	13	(2%)	14	(2.3%)	7	(1.3%)	10.8 (1.7%)	

Table I.7: **Density of Bobby Priority 2 diagnostics**

	May 2003		Dec 2003		May 2004		Dec 2004		Average	
Web sites with Bobby completed successfully	639		643		622		548		613	
Web sites that triggered 0 distinct diagnostics	0	(0%)	1	(0.2%)	0	(0%)	0	(0%)	0.3	(0%)
Web sites that triggered 1 distinct diagnostic	7	(1.1%)	8	(1.2%)	5	(0.8%)	3	(0.5%)	5.8	(0.9%)
Web sites that triggered 2 distinct diagnostics	26	(4.1%)	23	(3.6%)	21	(3.4%)	19	(3.5%)	22.3	(3.6%)
Web sites that triggered 3 distinct diagnostics	87	(13.6%)	89	(13.8%)	98	(15.8%)	76	(13.9%)	87.5	(14.3%)
Web sites that triggered 4 distinct diagnostics	207	(32.4%)	226	(35.1%)	210	(33.8%)	182	(33.2%)	206.3	(33.6%)
Web sites that triggered 5 distinct diagnostics	207	(32.4%)	194	(30.2%)	196	(31.5%)	175	(31.9%)	193	(31.5%)
Web sites that triggered 6 distinct diagnostics	80	(12.5%)	74	(11.5%)	64	(10.3%)	70	(12.8%)	72	(11.7%)
Web sites that triggered 7 distinct diagnostics	19	(3%)	25	(3.9%)	23	(3.7%)	20	(3.6%)	21.8	(3.5%)
Web sites that triggered 8 distinct diagnostics	6	(0.9%)	3	(0.5%)	5	(0.8%)	3	(0.5%)	4.3	(0.7%)

I.3 HTML technical standards conformance investigation results

Table I.8: **Details regarding OpenSP executions**

	May 2003	Dec 2003	May 2004	Dec 2004	Average
Web sites processed by OpenSP	747	722	698	640	701.8
Web sites where all the web pages had usable document type information	1 (0.1%)	2 (0.3%)	4 (0.6%)	3 (0.5%)	2.5 (0.4%)
Web sites considered for HTML defects analysis	140 (18.7%)	142 (19.7%)	175 (25.1%)	196 (30.6%)	163.3 (23.3%)

Table I.9: **Document Type Declaration in web pages considered by OpenSP**

	May 2003	Dec 2003	May 2004	Dec 2004	Average
Web pages processed by OpenSP	16,312	14,874	13,461	11,644	14,073
No document type information	13,837 (84.8%)	12,578 (84.6%)	10,954 (81.4%)	9,116 (78.3%)	11,621 (82.6%)
Incorrect document type information	427 (2.6%)	457 (3.1%)	407 (3%)	175 (1.5%)	367 (2.6%)
Usable document type information	1,788 (11%)	1,758 (11.8%)	2,050 (15.2%)	2,324 (20%)	1,980 (14.1%)
Considered for HTML defects analysis (in web sites with at least 3 web pages and 100 KB of data with usable document type information)	1,672 (10.3%)	1,646 (11.1%)	1,934 (14.4%)	2,217 (19.0%)	1,867 (13.3%)

Table I.10: **The 10 most common OpenSP diagnostics triggered in the investigation**

ID		May 2003	Dec 2003	May 2004	Dec 2004	Average
108	Undefined attribute	116 (82.9%)	117 (82.4%)	142 (81.1%)	171 (87.2%)	136.5 (83.4%)
127	Required attribute not specified	102 (72.9%)	104 (73.2%)	136 (77.7%)	160 (81.6%)	125.5 (76.4%)
64	Element not allowed by document type	101 (72.1%)	99 (69.7%)	123 (70.3%)	151 (77%)	118.5 (72.3%)
79	End tag for not opened element	84 (60%)	86 (60.6%)	103 (58.9%)	118 (60.2%)	97.8 (59.9%)
25	General entity not defined and no default entity	57 (40.7%)	63 (44.4%)	82 (46.9%)	105 (53.6%)	76.8 (46.4%)
139	Non SGML character number	66 (47.1%)	64 (45.1%)	71 (40.6%)	77 (39.3%)	69.5 (43%)
131	Attribute value not allowed	59 (42.1%)	59 (41.5%)	71 (40.6%)	79 (40.3%)	67 (41.1%)
68	Missing required end tag (element DTD)	65 (46.4%)	53 (37.3%)	70 (40%)	77 (39.3%)	66.3 (40.8%)
73	End tag for element not finished	52 (37.1%)	52 (36.6%)	44 (25.1%)	68 (34.7%)	54 (33.4%)
82	Attribute value must be literal	54 (38.6%)	51 (35.9%)	47 (26.9%)	59 (30.1%)	52.8 (32.9%)

Figure I.5: **The 10 most common OpenSP diagnostics**
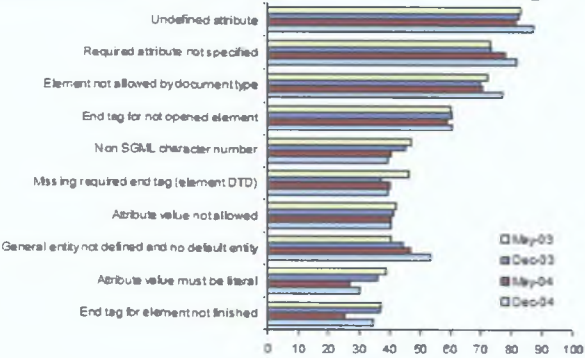
9

Table I.11: **Density of OpenSP diagnostics**

	May 2003		Dec 2003		May 2004		Dec 2004		Average	
Web sites considered for further analysis	140		142		175		196		163.3	
Web sites that triggered 0 distinct diagnostics	16	(11.4%)	17	(12%)	22	(12.6%)	18	(9.2%)	18.3	(11.2%)
Web sites that triggered 1 distinct diagnostic	2	(1.4%)	3	(2.1%)	1	(0.6%)	0	(0%)	1.5	(0.9%)
Web sites that triggered 2 distinct diagnostics	3	(2.1%)	1	(0.7%)	5	(2.9%)	5	(2.6%)	3.5	(2.1%)
Web sites that triggered 3 distinct diagnostics	3	(2.1%)	4	(2.8%)	8	(4.6%)	6	(3.1%)	5.3	(3.2%)
Web sites that triggered 4 distinct diagnostics	11	(7.9%)	12	(8.5%)	9	(5.1%)	11	(5.6%)	10.8	(6.6%)
Web sites that triggered 5 distinct diagnostics	6	(4.3%)	6	(4.2%)	6	(3.4%)	5	(2.6%)	5.8	(3.5%)
Web sites that triggered 6 distinct diagnostics	10	(7.1%)	13	(9.2%)	17	(9.7%)	14	(7.1%)	13.5	(8.3%)
Web sites that triggered 7 distinct diagnostics	13	(9.3%)	8	(5.6%)	10	(5.7%)	14	(7.1%)	11.3	(6.9%)
Web sites that triggered 8 distinct diagnostics	9	(6.4%)	16	(11.3%)	16	(9.1%)	21	(10.7%)	15.5	(9.5%)
Web sites that triggered 9 distinct diagnostics	7	(5%)	7	(4.9%)	21	(12%)	11	(5.6%)	11.5	(7%)
Web sites that triggered 10 distinct diagnostics	11	(7.9%)	10	(7%)	15	(8.6%)	22	(11.2%)	14.5	(8.9%)
Web sites that triggered 11 distinct diagnostics	10	(7.1%)	10	(7%)	9	(5.1%)	14	(7.1%)	10.8	(6.6%)
Web sites that triggered 12 distinct diagnostics	11	(7.9%)	7	(4.9%)	5	(2.9%)	18	(9.2%)	10.3	(6.3%)
Web sites that triggered 13 distinct diagnostics	10	(7.1%)	11	(7.7%)	9	(5.1%)	15	(7.7%)	11.3	(6.9%)
Web sites that triggered 14 distinct diagnostics	3	(2.1%)	1	(0.7%)	8	(4.6%)	7	(3.6%)	4.8	(2.9%)
Web sites that triggered 15 distinct diagnostics	3	(2.1%)	3	(2.1%)	3	(1.7%)	5	(2.6%)	3.5	(2.1%)
Web sites that triggered 16 distinct diagnostics	6	(4.3%)	5	(3.5%)	3	(1.7%)	4	(2%)	4.5	(2.8%)

10

Table I.11: **Density of OpenSP diagnostics**

	May 2003		Dec 2003		May 2004		Dec 2004		Average	
Web sites that triggered 17 distinct diagnostics	3	(2.1%)	1	(0.7%)	2	(1.1%)	0	(0%)	1.5	(0.9%)
Web sites that triggered 18 distinct diagnostics	2	(1.4%)	3	(2.1%)	4	(2.3%)	4	(2%)	3.3	(2%)
Web sites that triggered 19 distinct diagnostics		(0%)	3	(2.1%)	1	(0.6%)	1	(0.5%)	1.3	(0.8%)
Web sites that triggered 20 distinct diagnostics	1	(0.7%)	1	(0.7%)	0	(0%)	0	(0%)	0.5	(0.3%)
Web sites that triggered 20 distinct diagnostics	0	(0%)	0	(0%)	1	(0.6%)	1	(0.5%)	0.5	(0.3%)

Figure I.6: **Density of OpenSP diagnostics**
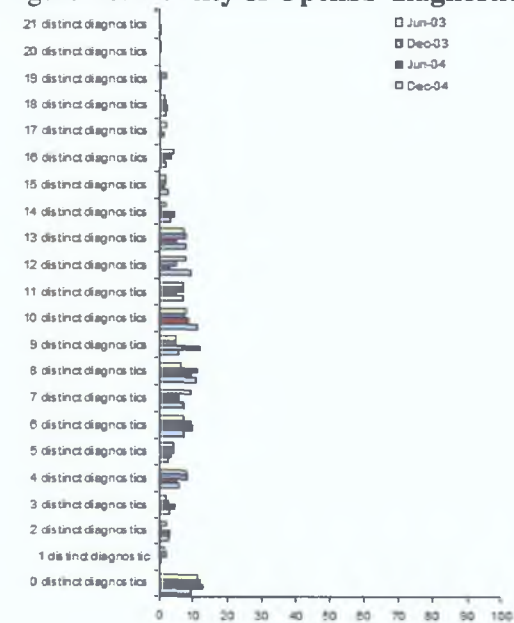
Appendix J

Results of surveys conducted on the web sample of German web sites (May 2003–December 2004)

J.1 Web content mirroring process results

Table J.1: **Results of the web content sampling process**

	May 2003	Dec 2003	May 2004	Dec 2004	Average
Data amount retrieved	377 MB	378 MB	380 MB	371 MB	376 MB
Web pages considered for mirroring	58,278	57,598	56,641	55,037	56,888.5
Web pages correctly retrieved (size>0 and no retrieval errors)	52,825 (90.6%)	51,473 (89.4%)	50,461 (89.1%)	47,806 (86.9%)	50,641.3 (89%)
Web pages in web sites qualified for conformance investigations	38,062 (65.3%)	37,974 (65.9%)	37,245 (65.8%)	35,749 (65%)	37,257.5 (65.5%)
Web sites qualified for conformance investigations	1,636 (38.5%)	1,649 (38.8%)	1,646 (38.7%)	1,609 (37.9%)	1,635 (38.5%)

Table J.2: **Details regarding the web sites that didn't qualify for investigations**

	May 2003	Dec 2003	May 2004	Dec 2004	Common	Average
Web sites not qualified for investigations	2,615	2,602	2,605	2,641	2,248	2,615.8
No data retrieved	147 (5.6%)	178 (6.8%)	205 (7.9%)	275 (10.4%)	79 (3.5%)	201.3 (7.7%)
Only the main page retrieved (size>0)	1,159 (44.3%)	1,209 (46.5%)	1,226 (47.1%)	1,277 (48.4%)	914 (40.7%)	1,217.8 (46.6%)
Less than 3 pages (2) and at least 100 KB retrieved	1 (0%)	4 (0.2%)	4 (0.2%)	3 (0.1%)	1 (0%)	3 (0.1%)
Less than 100 KB and at least 3 pages retrieved	1,177 (45%)	1,082 (41.6%)	1,054 (40.5%)	976 (37%)	734 (32.7%)	1,072.3 (41%)
Less than 3 pages (2) and less that 100 KB (but more than 0) retrieved	130 (5%)	128 (4.9%)	115 (4.4%)	110 (4.2%)	55 (2.4%)	120.8 (4.6%)

J.2 WCAG 1.0 conformance investigation results

Table J.3: **Details regarding Bobby executions**

	May 2003	Dec 2003	May 2004	Dec 2004	Average
Web sites considered by Bobby	1,636	1,649	1,646	1,609	1,635
Bobby abnormally terminated	101 (6.2%)	102 (6.2%)	102 (6.2%)	94 (5.8%)	99.8 (6.1%)
Bobby "locked up"	90 (5.5%)	75 (4.5%)	66 (4%)	60 (3.7%)	72.8 (4.4%)
Bobby completed successfully	1,444 (88.3%)	1,471 (89.2%)	1,477 (89.7%)	1,455 (90.4%)	1,461.8 (89.4%)

Table J.4: **Bobby Priority 1 diagnostics triggered during the investigation**

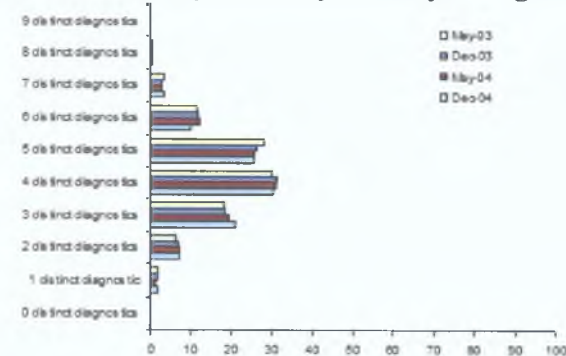| ID		May 2003	Dec 2003	May 2004	Dec 2004	Average
g9	Provide alternative text for all images	1,336 (92.5%)	1,343 (91.3%)	1,339 (90.7%)	1,287 (88.5%)	1326.3 (90.7%)
g39	Give each frame a title	735 (50.9%)	730 (49.6%)	721 (48.8%)	664 (45.6%)	712.5 (48.7%)
g38	Each FRAME must reference an HTML file	726 (50.3%)	715 (48.6%)	706 (47.8%)	659 (45.3%)	701.5 (48.0%)
g240	Provide alternative text for all image map hot-spots	265 (18.4%)	243 (16.5%)	233 (15.8%)	218 (15%)	239.8 (16.4%)
g10	Provide alternative text for all image-type buttons in forms	126 (8.7%)	135 (9.2%)	141 (9.5%)	132 (9.1%)	133.5 (9.1%)
g21	Provide alternative text for each applet	110 (7.6%)	92 (6.3%)	72 (4.9%)	67 (4.6%)	85.3 (5.8%)
g20	Provide alternative content for each OBJECT	3 (0.2%)	3 (0.2%)	4 (0.3%)	5 (0.3%)	3.8 (0.3%)

Figure J.1: **Bobby Priority 1 diagnostics**
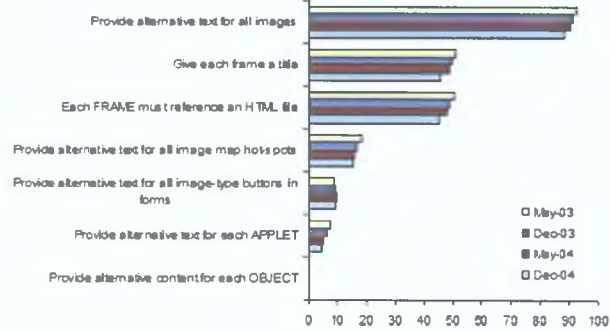
Table J.5: **Density of Bobby Priority 1 diagnostics**

	May 2003	Dec 2003	May 2004	Dec 2004	Average
Web sites with Bobby completed successfully	1,444	1,471	1,477	1,455	1,461.8
Web sites that triggered 0 distinct diagnostics	66 (4.6%)	69 (4.7%)	69 (4.7%)	101 (6.9%)	76.3 (5.2%)
Web sites that triggered 1 distinct diagnostic	461 (31.9%)	496 (33.7%)	513 (34.7%)	513 (35.3%)	495.8 (33.9%)
Web sites that triggered 2 distinct diagnostics	200 (13.9%)	207 (14.1%)	222 (15%)	214 (14.7%)	210.8 (14.4%)
Web sites that triggered 3 distinct diagnostics	463 (32.1%)	474 (32.2%)	465 (31.5%)	440 (30.2%)	460.5 (31.5%)
Web sites that triggered 4 distinct diagnostics	220 (15.2%)	197 (13.4%)	177 (12%)	165 (11.3%)	189.8 (13%)
Web sites that triggered 5 distinct diagnostics	33 (2.3%)	27 (1.8%)	30 (2%)	21 (1.4%)	27.8 (1.9%)
Web sites that triggered 6 distinct diagnostics	1 (0.1%)	1 (0.1%)	1 (0.1%)	1 (0.1%)	1 (0.1%)

4

Figure J.2: **Density of Bobby Priority 1 diagnostics**
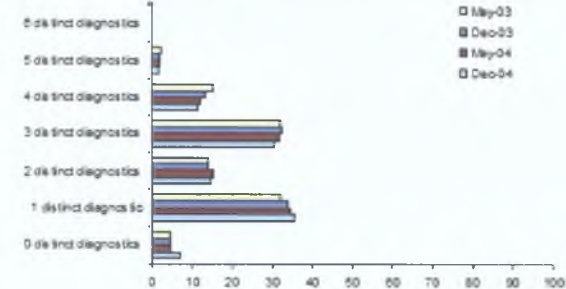
Figure J.3: **Bobby Priority 2 diagnostics**
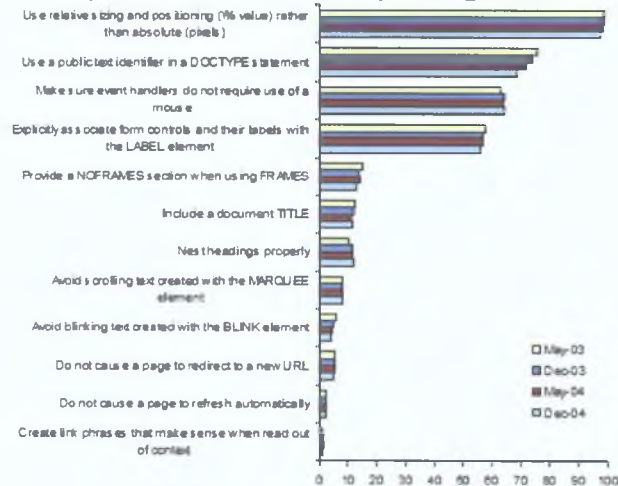
Figure J.4: **Density of Bobby Priority 2 diagnostics**

5

Table J.6: **Bobby Priority 2 diagnostics triggered during the investigation**

ID		May 2003	Dec 2003	May 2004	Dec 2004	Average
g104	Use relative sizing and positioning (% value) rather than absolute (pixels)	1,429 (99%)	1,454 (98.8%)	1,454 (98.4%)	1,418 (96%)	1438.8 (98.1%)
g271	Use a public text identifier in a DOCTYPE statement	1,093 (75.7%)	1,085 (73.8%)	1,067 (72.2%)	1,001 (67.8%)	1061.5 (72.4%)
g269	Make sure event handlers do not require use of a mouse	915 (63.4%)	948 (64.4%)	942 (63.8%)	938 (63.5%)	935.8 (63.8%)
g41	Explicitly associate form controls and their labels with the LABEL element	836 (57.9%)	836 (56.8%)	838 (56.7%)	814 (55.1%)	831 (56.6%)
g37	Provide a NOFRAMES section when using FRAMES	219 (15.2%)	207 (14.1%)	212 (14.4%)	187 (12.7%)	206.3 (14.1%)
g273	Include a document TI-TLE	182 (12.6%)	177 (12%)	162 (11%)	165 (11.2%)	171.5 (11.7%)
g2	Nest headings properly	150 (10.4%)	167 (11.4%)	167 (11.3%)	169 (11.4%)	163.3 (11.1%)
g5	Avoid scrolling text created with the MARQUEE element	119 (8.2%)	114 (7.7%)	115 (7.8%)	116 (7.9%)	116 (7.9%)
g4	Avoid blinking text created with the BLINK element	80 (5.5%)	73 (5%)	63 (4.3%)	58 (3.9%)	68.5 (4.7%)
g254	Do not cause a page to redirect to a new URL	78 (5.4%)	79 (5.4%)	80 (5.4%)	75 (5.1%)	78 (5.3%)
g33	Do not cause a page to refresh automatically	37 (2.6%)	32 (2.2%)	32 (2.2%)	29 (2%)	32.5 (2.2%)
g34	Create link phrases that make sense when read out of context	18 (1.2%)	19 (1.3%)	21 (1.4%)	15 (1%)	18.3 (1.2%)

Table J.7: **Density of Bobby Priority 2 diagnostics**

	May 2003	Dec 2003	May 2004	Dec 2004	Average
Web sites for which Bobby completed successfully	1,444	1,471	1,477	1,455	1,461.8
Web sites for which were triggered 0 distinct diagnostic	0 (0%)	0 (0%)	1 (0.1%)	2 (0.1%)	0.8 (0.1%)
Web sites that triggered 1 distinct diagnostic	25 (1.7%)	25 (1.7%)	20 (1.4%)	29 (2%)	24.8 (1.7%)
Web sites that triggered 2 distinct diagnostics	92 (6.4%)	99 (6.7%)	104 (7%)	104 (7.1%)	99.8 (6.8%)
Web sites that triggered 3 distinct diagnostics	265 (18.4%)	273 (18.6%)	285 (19.3%)	301 (20.7%)	281 (19.2%)
Web sites that triggered 4 distinct diagnostics	431 (29.8%)	459 (31.2%)	456 (30.9%)	442 (30.4%)	447 (30.6%)
Web sites that triggered 5 distinct diagnostics	406 (28.1%)	387 (26.3%)	379 (25.7%)	372 (25.6%)	386 (26.4%)
Web sites that triggered 6 distinct diagnostics	167 (11.6%)	176 (12%)	181 (12.3%)	148 (10.2%)	168 (11.5%)
Web sites that triggered 7 distinct diagnostics	47 (3.3%)	41 (2.8%)	42 (2.8%)	49 (3.4%)	44.8 (3.1%)
Web sites that triggered 8 distinct diagnostics	9 (0.6%)	10 (0.7%)	9 (0.6%)	8 (0.5%)	9 (0.6%)
Web sites that triggered 9 distinct diagnostics	2 (0.1%)	1 (0.1%)	0 (0%)	0 (0%)	0.8 (0.1%)

J.3 HTML technical standards conformance investigation results

Table J.8: **Details regarding OpenSP executions**

	May 2003	Dec 2003	May 2004	Dec 2004	Average
Web sites processed by OpenSP	1,636	1,649	1,646	1,609	1,635
Web sites where all the web pages had usable document type information	52 (3.2%)	52 (3.2%)	54 (3.3%)	55 (3.4%)	53 (3.3%)
Web sites considered for HTML defects analysis	603 (36.9%)	647 (39.2%)	706 (42.9%)	734 (45.6%)	673 (41.1%)

Table J.9: **Document Type Declaration in web pages considered by OpenSP**

	May 2003	Dec 2003	May 2004	Dec 2004	Average
Web pages processed by OpenSP	38,062	37,974	37,245	35,749	37,257.5
No document type information	24,258 (63.7%)	23,422 (61.7%)	22,382 (60.1%)	20,456 (57.2%)	22,629.5 (60.7%)
Incorrect document type information	2,692 (7.1%)	2,411 (6.3%)	2,024 (5.4%)	1,903 (5.3%)	2,257.5 (6.1%)
Usable document type information	10,681 (28.1%)	11,726 (30.9%)	12,502 (33.6%)	13,087 (36.6%)	11,999 (32.2%)
Considered for HTML defects analysis (in web sites with at least 3 web pages and 100 KB of data with usable document type information)	10,431 (27.4%)	11,474 (30.2%)	12,270 (32.9%)	12,853 (36%)	11,757 (31.6%)

Table J.10: **The 10 most common OpenSP diagnostics triggered in the investigation**

ID		May 2003	Dec 2003	May 2004	Dec 2004	Average
108	Undefined attribute	549 (91%)	588 (90.9%)	633 (89.7%)	645 (87.9%)	603.8 (89.9%)
64	Element not allowed by document type	486 (80.6%)	511 (79%)	548 (77.6%)	532 (72.5%)	519.3 (77.4%)
127	Required attribute not specified	463 (76.8%)	515 (79.6%)	567 (80.3%)	589 (80.2%)	533.5 (79.2%)
79	End tag for not opened element	416 (69%)	446 (68.9%)	479 (67.8%)	487 (66.3%)	457 (68%)
68	Missing required end tag (element DTD)	327 (54.2%)	344 (53.2%)	359 (50.8%)	350 (47.7%)	345 (51.5%)
73	End tag for element not finished	276 (45.8%)	294 (45.4%)	302 (42.8%)	300 (40.9%)	293 (43.7%)
25	General entity not defined and no default entity	269 (44.6%)	287 (44.4%)	309 (43.8%)	319 (43.5%)	296 (44%)
139	Non SGML character number	268 (44.4%)	313 (48.4%)	332 (47%)	347 (47.3%)	315 (46.8%)
131	Attribute value not allowed	217 (36%)	230 (35.5%)	256 (36.3%)	246 (33.5%)	237.3 (35.3%)
82	Attribute value must be literal	212 (35.2%)	222 (34.3%)	243 (34.4%)	245 (33.4%)	230.5 (34.3%)

Figure J.5: **The 10 most common OpenSP diagnostics**
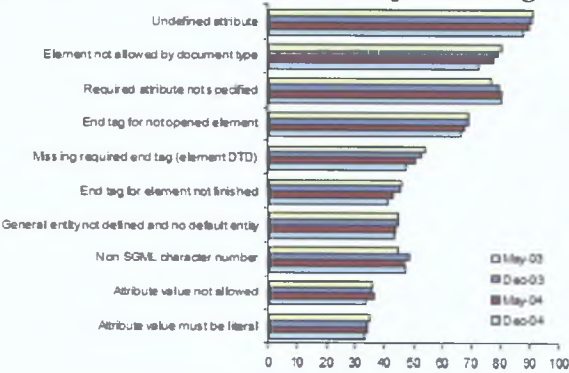
9

Table J.11: **Density of OpenSP diagnostics**

	May 2003		Dec 2003		May 2004		Dec 2004		Average	
Web sites considered for further analysis	603		647		706		734		672.5	
Web sites that triggered 0 distinct diagnostics	29	(4.8%)	26	(4%)	29	(4.1%)	33	(4.5%)	29.3	(4.3%)
Web sites that triggered 1 distinct diagnostic	9	(1.5%)	6	(0.9%)	6	(0.8%)	12	(1.6%)	8.3	(1.2%)
Web sites that triggered 2 distinct diagnostics	6	(1%)	10	(1.5%)	11	(1.6%)	21	(2.9%)	12	(1.8%)
Web sites that triggered 3 distinct diagnostics	19	(3.2%)	30	(4.6%)	33	(4.7%)	27	(3.7%)	27.3	(4.1%)
Web sites that triggered 4 distinct diagnostics	21	(3.5%)	28	(4.3%)	32	(4.5%)	37	(5%)	29.5	(4.4%)
Web sites that triggered 5 distinct diagnostics	29	(4.8%)	29	(4.5%)	37	(5.2%)	42	(5.7%)	34.3	(5.1%)
Web sites that triggered 6 distinct diagnostics	44	(7.3%)	47	(7.3%)	53	(7.5%)	52	(7.1%)	49	(7.3%)
Web sites that triggered 7 distinct diagnostics	56	(9.3%)	53	(8.2%)	63	(8.9%)	66	(9%)	59.5	(8.8%)
Web sites that triggered 8 distinct diagnostics	63	(10.4%)	60	(9.3%)	63	(8.9%)	68	(9.3%)	63.5	(9.4%)
Web sites that triggered 9 distinct diagnostics	52	(8.6%)	59	(9.1%)	62	(8.8%)	61	(8.3%)	58.5	(8.7%)
Web sites that triggered 10 distinct diagnostics	49	(8.1%)	63	(9.7%)	57	(8.1%)	45	(6.1%)	53.5	(8%)
Web sites that triggered 11 distinct diagnostics	48	(8%)	53	(8.2%)	68	(9.6%)	73	(9.9%)	60.5	(9%)
Web sites that triggered 12 distinct diagnostics	39	(6.5%)	33	(5.1%)	39	(5.5%)	48	(6.5%)	39.8	(5.9%)
Web sites that triggered 13 distinct diagnostics	43	(7.1%)	46	(7.1%)	40	(5.7%)	45	(6.1%)	43.5	(6.5%)
Web sites that triggered 14 distinct diagnostics	20	(3.3%)	22	(3.4%)	32	(4.5%)	29	(4%)	25.8	(3.8%)
Web sites that triggered 15 distinct diagnostics	22	(3.6%)	20	(3.1%)	21	(3%)	25	(3.4%)	22	(3.3%)
Web sites that triggered 16 distinct diagnostics	22	(3.6%)	26	(4%)	16	(2.3%)	15	(2%)	19.8	(2.9%)

10

Table J.11: **Density of OpenSP diagnostics**

	May 2003		Dec 2003		May 2004		Dec 2004		Average	
Web sites that triggered 17 distinct diagnostics	7	(1.2%)	11	(1.7%)	15	(2.1%)	16	(2.2%)	12.3	(1.8%)
Web sites that triggered 18 distinct diagnostics	10	(1.7%)	6	(0.9%)	14	(2%)	10	(1.4%)	10	(1.5%)
Web sites that triggered 19 distinct diagnostics	8	(1.3%)	8	(1.2%)	5	(0.7%)	2	(0.3%)	5.8	(0.9%)
Web sites that triggered 20 distinct diagnostics	5	(0.8%)	5	(0.8%)	5	(0.7%)	5	(0.7%)	5	(0.7%)
Web sites that triggered 21 distinct diagnostics	1	(0.2%)	5	(0.8%)	3	(0.4%)	1	(0.1%)	2.5	(0.4%)
Web sites that triggered 22 distinct diagnostics	0	(0%)	1	(0.2%)	0	(0%)	0	(0%)	0.3	(0%)
Web sites that triggered 23 distinct diagnostics	1	(0.2%)	0	(0%)	2	(0.3%)	0	(0%)	0.8	(0.1%)
Web sites that triggered 24 distinct diagnostics	0	(0%)	0	(0%)	0	(0%)	1	(0.1%)	0.3	(0%)

11

Figure J.6: **Density of OpenSP diagnostics**

Appendix K

Results of surveys conducted on the subjective and random web samples of Irish web sites (May 2003)

K.1   Web content mirroring process results

Table K.1: **Results of the web content sampling process**

	Subjective	Random
Data amount retrieved	34 MB	29 MB
Web pages considered for mirroring	3,382	3,319
Web pages correctly retrieved (size>0 and no retrieval errors)	2,980 (88.1%)	2,817 (84.9%)
Web pages qualified for conformance investigation	2,825 (83.5%)	2,288 (68.9%)
Web sites qualified for conformance investigation	156 (72.9%)	123 (45.2%)

Table K.2: **Details regarding the web sites that didn't qualify for investigations**

	Subjective		Random	
Web sites not qualified for investigations	58		149	
No data retrieved	9	(15.5%)	9	(6%)
Only the main page retrieved (size>0)	31	(53.4%)	91	(61.1%)
Less than 3 pages (2) and at least 100 KB retrieved	0	(0%)	1	(0.7%)
Less than 100 KB and at least 3 pages retrieved	15	(25.9%)	46	(30.9%)
Less than 3 pages (2) and less that 100 KB (but more than 0) retrieved	3	(5.2%)	2	(1.3%)

K.2 WCAG 1.0 conformance investigation results

Table K.3: **Details regarding Bobby executions**

	Subjective		Random	
Web sites considered by Bobby	156		123	
Bobby abnormally terminated	13	(8.3%)	3	(2.4%)
Bobby "locked up"	9	(5.8%)	6	(4.9%)
Bobby completed successfully	134	(85.9%)	114	(92.7%)

Table K.4: **Bobby Priority 1 diagnostics triggered during the investigation**

ID		Subjective		Random	
g9	Provide alternative text for all images	115	(85.8%)	106	(93%)
g38	Each FRAME must reference an HTML file	30	(22.4%)	24	(21.1%)
g39	Give each frame a title	29	(21.6%)	25	(21.9%)
g10	Provide alternative text for all image-type buttons in forms	25	(18.7%)	9	(7.9%)
g240	Provide alternative text for all image map hot-spots	24	(17.9%)	25	(21.9%)
g21	Provide alternative text for each APPLET	10	(7.5%)	10	(8.8%)

3

Figure K.1: **Bobby Priority 1 diagnostics**
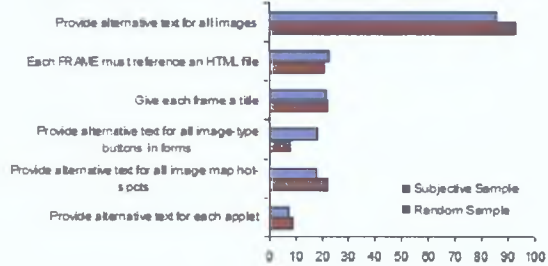
Table K.5: **Density of Bobby Priority 1 diagnostics**

	Subjective	Random
Web sites with Bobby completed successfully	134	114
Web sites that triggered 0 distinct diagnostics	16 (11.9%)	7 (6.1%)
Web sites that triggered 1 distinct diagnostic	56 (41.8%)	55 (48.2%)
Web sites that triggered 2 distinct diagnostic	32 (23.9%)	25 (21.9%)
Web sites that triggered 3 distinct diagnostics	13 (9.7%)	16 (14%)
Web sites that triggered 4 distinct diagnostics	11 (8.2%)	9 (7.9%)
Web sites that triggered 5 distinct diagnostics	6 (4.5%)	2 (1.8%)

Figure K.2: **Density of Bobby Priority 1 diagnostics**

4
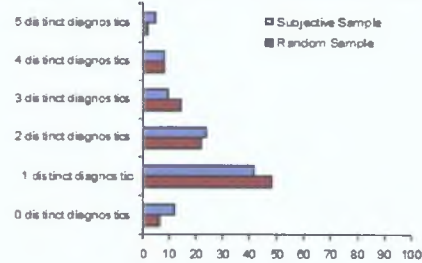
Table K.6: **Bobby Priority 2 diagnostics triggered during the investigation**

ID		Subjective		Random	
g104	Use relative sizing and positioning (% value) rather than absolute (pixels)	129	(96.3%)	112	(98.2%)
g271	Use a public text identifier in a DOCTYPE statement	102	(76.1%)	100	(87.7%)
g41	Explicitly associate form controls and their labels with the LABEL element	100	(74.6%)	71	(62.3%)
g269	Make sure event handlers do not require use of a mouse	97	(72.4%)	85	(74.6%)
g34	Create link phrases that make sense when read out of context	57	(42.5%)	41	(36%)
g273	Include a document TITLE	17	(12.7%)	18	(15.8%)
g2	Nest headings properly	11	(8.2%)	17	(14.9%)
g37	Provide a NOFRAMES section when using FRAMES	6	(4.5%)	3	(2.6%)
g254	Do not cause a page to redirect to a new URL	3	(2.2%)	4	(3.5%)
g5	Avoid scrolling text created with the MARQUEE element	2	(1.5%)	9	(7.9%)
g33	Do not cause a page to refresh automatically	2	(1.5%)	1	(0.9%)
g4	Avoid blinking text created with the BLINK element	1	(0.7%)	2	(1.8%)

Figure K.3: **Bobby Priority 2 diagnostics**
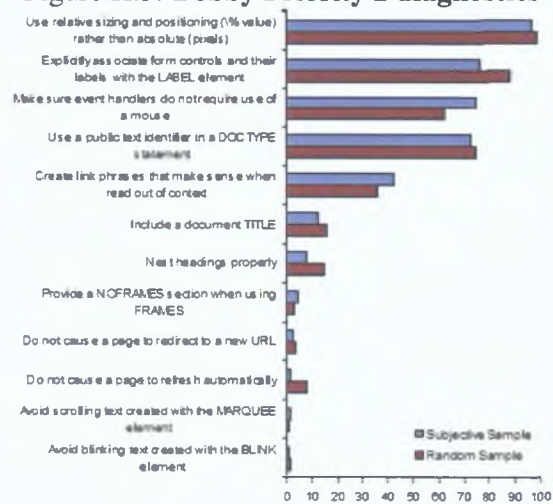
Table K.7: **Density of Bobby Priority 2 diagnostics**

	Subjective		Random	
Web sites with Bobby completed successfully	134		114	
Web sites that triggered 0 distinct diagnostics	0	(0%)	0	(0%)
Web sites that triggered 1 distinct diagnostic	1	(0.7%)	1	(0.9%)
Web sites that triggered 2 distinct diagnostics	3	(2.2%)	5	(4.4%)
Web sites that triggered 3 distinct diagnostics	12	(9%)	8	(7%)
Web sites that triggered 4 distinct diagnostics	40	(29.9%)	28	(24.6%)
Web sites that triggered 5 distinct diagnostics	42	(31.3%)	37	(32.5%)
Web sites that triggered 6 distinct diagnostics	24	(17.9%)	25	(21.9%)
Web sites that triggered 7 distinct diagnostics	8	(6%)	9	(7.9%)
Web sites that triggered 8 distinct diagnostics	4	(3%)	1	(0.9%)

Figure K.4: **Density of Bobby Priority 2 diagnostics**
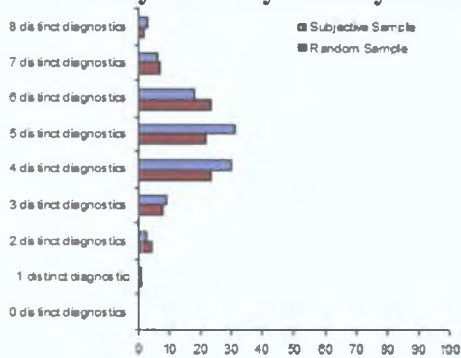
K.3   HTML technical standards conformance investigation results

Table K.8: **Details regarding OpenSP executions**

	Subjective	Random
Web sites processed by OpenSP	156	123
Web sites where all the web pages had usable document type information	1 (0.6%)	2 (1.6%)
Web sites considered for HTML defects analysis	67 (42.9%)	36 (29.3%)

Table K.9: **Document Type Declaration in web pages considered by OpenSP**

	Subjective	Random
Web pages processed by OpenSP	2,825	2,288
No document type information	1,866 (66.1%)	1,732 (75.7%)
Incorrect document type information	23 (0.8%)	42 (1.8%)
Usable document type information	932 (33%)	465 (20.3%)
Considered for HTML defects analysis (in web sites with at least 3 web pages and 100 KB of data with usable document type information)	910 (32.2%)	454 (19.8%)

Table K.10: **The 10 most common OpenSP diagnostics triggered in the investigation**

ID		Subjective		Random	
108	Undefined attribute	66	(98.5%)	33	(91.7%)
127	Required attribute not specified	54	(80.6%)	19	(52.8%)
64	Element not allowed by document type	53	(79.1%)	29	(80.6%)
79	End tag for not opened element	53	(79.1%)	27	(75%)
139	Non SGML character number	35	(52.2%)	20	(55.6%)
68	Missing required end tag (element DTD)	35	(52.2%)	20	(55.6%)
25	General entity not defined and no default entity	34	(50.7%)	17	(47.2%)
73	End tag for element not finished	32	(47.8%)	11	(30.6%)
82	Attribute value must be literal	27	(40.3%)	16	(44.4%)
131	Attribute value not allowed	27	(40.3%)	10	(27.8%)

Figure K.5: **The 10 most common OpenSP diagnostics**
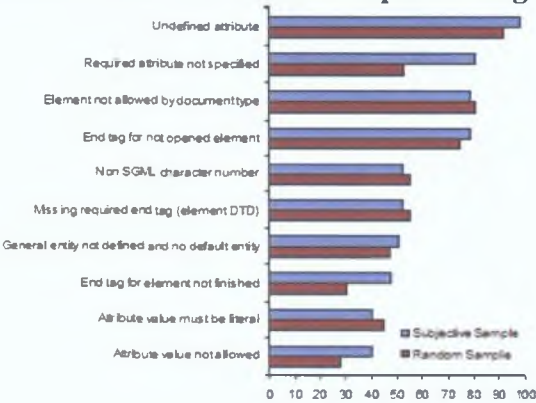
9

Table K.11: **Density of OpenSP diagnostics**

	Subjective		Random	
Web sites considered for HTML defects analysis	67		36	
Web sites that triggered 0 distinct diagnostics	0	(0%)	1	(2.8%)
Web sites that triggered 1 distinct diagnostic	0	(0%)	1	(2.8%)
Web sites that triggered 2 distinct diagnostics	0	(0%)	2	(5.6%)
Web sites that triggered 3 distinct diagnostics	2	(3%)	1	(2.8%)
Web sites that triggered 4 distinct diagnostics	6	(9%)	2	(5.6%)
Web sites that triggered 5 distinct diagnostics	6	(9%)	1	(2.8%)
Web sites that triggered 6 distinct diagnostics	2	(3%)	1	(2.8%)
Web sites that triggered 7 distinct diagnostics	6	(9%)	3	(8.3%)
Web sites that triggered 8 distinct diagnostics	5	(7.5%)	3	(8.3%)
Web sites that triggered 9 distinct diagnostics	7	(10.4%)	1	(2.8%)
Web sites that triggered 10 distinct diagnostics	3	(4.5%)	5	(13.9%)
Web sites that triggered 11 distinct diagnostics	5	(7.5%)	5	(13.9%)
Web sites that triggered 12 distinct diagnostics	7	(10.4%)	4	(11.1%)
Web sites that triggered 13 distinct diagnostics	2	(3%)	2	(5.6%)
Web sites that triggered 14 distinct diagnostics	4	(6%)	2	(5.6%)
Web sites that triggered 15 distinct diagnostics	4	(6%)	0	(0%)
Web sites that triggered 16 distinct diagnostics	4	(6%)	1	(2.8%)

Table K.11: **Density of OpenSP diagnostics**

	Subjective		Random	
Web sites that triggered 17 distinct diagnostics	1	(1.5%)	0	(0%)
Web sites that triggered 18 distinct diagnostics	2	(3%)	0	(0%)
Web sites that triggered 19 distinct diagnostics	0	(0%)	0	(0%)
Web sites that triggered 20 distinct diagnostics	1	(1.5%)	0	(0%)
Web sites that triggered 21 distinct diagnostics	0	(0%)	0	(0%)
Web sites that triggered 22 distinct diagnostics	0	(0%)	1	(2.8%)

Figure K.6: **Density of OpenSP diagnostics**
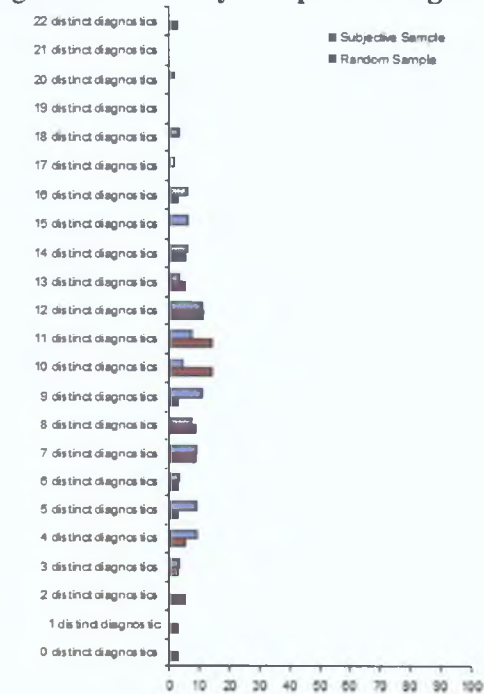
Appendix L

Results of surveys conducted on the subjective and random web samples of Irish web sites (December 2003)

L.1   Web content mirroring process results

Table L.1: **Results of the web content sampling process**

	Subjective	Random
Data amount retrieved	34 MB	26 MB
Web pages considered for mirroring	3,051	3,043
Web pages correctly retrieved (size>0 and no retrieval errors)	2,667 (87.4%)	2,541 (83.5%)
Web pages in web sites qualified for conformance investigations	2,531 (83%)	2,064 (67.8%)
Web sites qualified for conformance investigations	153 (71.5%)	111 (40.8%)

Table L.2: **Details regarding the web sites that didn't qualify for investigations**

	Subjective	Random
Web sites not qualified for investigations	61	161
No data retrieved	10 (16.4%)	15 (9.3%)
Only the main page retrieved (size>0)	32 (52.5%)	102 (63.4%)
Less than 3 pages (2) and at least 100 KB retrieved	0 (0%)	1 (0.6%)
Less than 100 KB and at least 3 pages retrieved	14 (23%)	40 (24.8%)
Less than 3 pages (2) and less that 100 KB (but more than 0) retrieved	5 (8.2%)	3 (1.9%)

L.2   WCAG 1.0 conformance investigation results

Table L.3: **Details regarding Bobby executions**

	Subjective		Random	
Web sites considered by Bobby	153		111	
Bobby abnormally terminated	10	(6.5%)	3	(2.7%)
Bobby "locked up"	7	(4.6%)	4	(3.6%)
Bobby completed successfully	136	(88.9%)	104	(93.7%)

Table L.4: **Bobby Priority 1 diagnostics triggered during the investigation**

ID		Subjective		Random	
g9	Provide alternative text for all images	115	(84.6%)	94	(90.4%)
g38	Each FRAME must reference an HTML file	32	(23.5%)	23	(22.1%)
g39	Give each frame a title	31	(22.8%)	23	(22.1%)
g240	Provide alternative text for all image map hot-spots	25	(18.4%)	20	(19.2%)
g10	Provide alternative text for all image-type buttons in forms	23	(16.9%)	10	(9.6%)
g21	Provide alternative text for each APPLET	10	(7.4%)	8	(7.7%)

3

Figure L.1: **Bobby Priority 1 diagnostics**
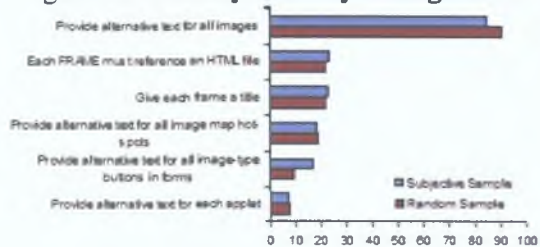
Table L.5: **Density of Bobby Priority 1 diagnostics**

	Subjective		Random	
Web sites with Bobby completed successfully	136		104	
Web sites that triggered 0 distinct diagnostics	17	(12.5%)	7	(6.7%)
Web sites that triggered 1 distinct diagnostic	58	(42.6%)	50	(48.1%)
Web sites that triggered 2 distinct diagnostic	28	(20.6%)	25	(24%)
Web sites that triggered 3 distinct diagnostics	18	(13.2%)	12	(11.5%)
Web sites that triggered 4 distinct diagnostics	8	(5.9%)	8	(7.7%)
Web sites that triggered 5 distinct diagnostics	6	(4.4%)	2	(1.9%)
Web sites that triggered 6 distinct diagnostics	1	(0.7%)	0	(0%)

Figure L.2: **Density of Bobby Priority 1 diagnostics**

4
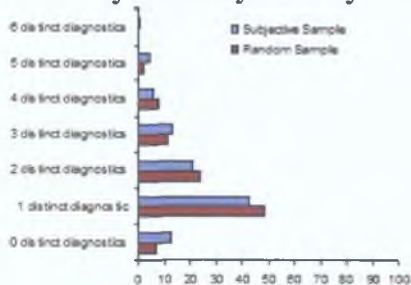
Table L.6: **Bobby Priority 2 diagnostics triggered during the investigation**

ID		Subjective		Random	
g104	Use relative sizing and positioning (% value) rather than absolute (pixels)	130	(95.6%)	102	(98.1%)
g41	Explicitly associate form controls and their labels with the LABEL element	101	(74.3%)	58	(55.8%)
g269	Make sure event handlers do not require use of a mouse	100	(73.5%)	76	(73.1%)
g271	Use a public text identifier in a DOCTYPE statement	94	(69.1%)	83	(79.8%)
g34	Create link phrases that make sense when read out of context	60	(44.1%)	43	(41.3%)
g273	Include a document TITLE	18	(13.2%)	20	(19.2%)
g2	Nest headings properly	9	(6.6%)	20	(19.2%)
g37	Provide a NOFRAMES section when using FRAMES	6	(4.4%)	3	(2.9%)
g254	Do not cause a page to redirect to a new URL	4	(2.9%)	0	(0%)
g33	Do not cause a page to refresh automatically	3	(2.2%)	2	(1.9%)
g5	Avoid scrolling text created with the MARQUEE element	2	(1.5%)	12	(11.5%)
g4	Avoid blinking text created with the BLINK element	0	(0%)	3	(2.9%)

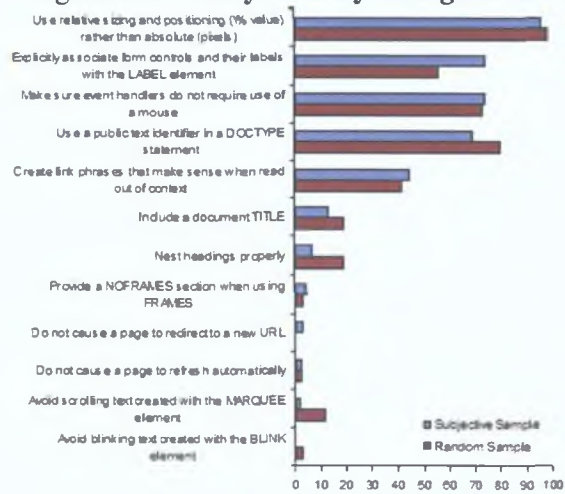Figure L.3: **Bobby Priority 2 diagnostics**

Use relative sizing and positioning (% value) rather than absolute (pixels)	
Explicitly associate form controls and their labels with the LABEL element	
Make sure event handlers do not require use of a mouse	
Use a public text identifier in a DOCTYPE statement	
Create link phrases that make sense when read out of context	
Include a document TITLE	
Nest headings properly	
Provide a NOFRAMES section when using FRAMES	
Do not cause a page to redirect to a new URL	
Do not cause a page to refresh automatically	
Avoid scrolling text created with the MARQUEE element	
Avoid blinking text created with the BLINK element	

□ Subjective Sample
■ Random Sample

0 10 20 30 40 50 60 70 80 90 100

Table L.7: **Density of Bobby Priority 2 diagnostics**

	Subjective		Random	
Web sites for which Bobby completed successfully	136		104	
Web sites that triggered 0 distinct diagnostic	0	(0%)	0	(0%)
Web sites that triggered 1 distinct diagnostic	1	(0.7%)	1	(1%)
Web sites that triggered 2 distinct diagnostics	6	(4.4%)	5	(4.8%)
Web sites that triggered 3 distinct diagnostics	17	(12.5%)	9	(8.7%)
Web sites that triggered 4 distinct diagnostics	34	(25%)	27	(26%)
Web sites that triggered 5 distinct diagnostics	40	(29.4%)	25	(24%)
Web sites that triggered 6 distinct diagnostics	25	(18.4%)	27	(26%)
Web sites that triggered 7 distinct diagnostics	8	(5.9%)	8	(7.7%)
Web sites that triggered 8 distinct diagnostics	4	(2.9%)	2	(1.9%)
Web sites that triggered 9 distinct diagnostics	1	(0.7%)	0	(0%)

Figure L.4: **Density of Bobby Priority 2 diagnostics**

7
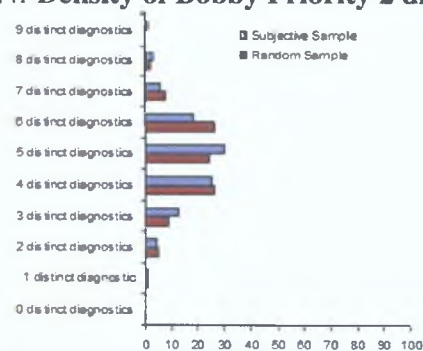
L.3 HTML technical standards conformance investigation results

Table L.8: **Details regarding OpenSP executions**

	Subjective	Random
Web sites considered by OpenSP	153	111
Web sites where all the web pages had correct DOCTYPE	0 (0%)	2 (1.8%)
Web sites considered for further analysis	74 (48.4%)	39 (35.1%)

Table L.9: **Document Type Declaration in web pages considered by OpenSP**

	Subjective	Random
Web pages processed by OpenSP	2,531	2,064
No document type information	1,547 (61.1%)	1,435 (69.5%)
Incorrect document type information	8 (0.3%)	78 (3.8%)
Usable document type information	974 (38.5%)	509 (24.7%)
Considered for HTML defects analysis (in web sites with at least 3 web pages and 100 KB of data with usable document type information)	952 (37.6%)	500 (24.2%)

Table L.10: **The 10 most common OpenSP diagnostics triggered in the investigation**

ID		Subjective		Random	
108	Undefined attribute	68	(91.9%)	36	(92.3%)
79	End tag for not opened element	68	(91.9%)	28	(71.8%)
127	Required attribute not specified	65	(87.8%)	28	(71.8%)
64	Element not allowed by document type	60	(81.1%)	31	(79.5%)
68	Missing required end tag (element DTD)	42	(56.8%)	22	(56.4%)
139	Non SGML character number	37	(50%)	24	(61.5%)
25	General entity not defined and no default entity	36	(48.6%)	20	(51.3%)
73	End tag for element not finished	35	(47.3%)	16	(41%)
131	Attribute value not allowed	34	(45.9%)	15	(38.5%)
82	Attribute value must be literal	27	(36.5%)	18	(46.2%)

Figure L.5: **The 10 most common OpenSP diagnostics**
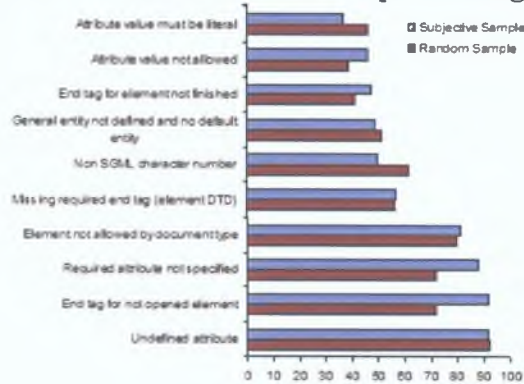
Table L.11: **Density of OpenSP diagnostics**

	Subjective		Random	
Web sites considered for further investigation	74		39	
Web sites that triggered 0 distinct diagnostics	1	(1.4%)	1	(2.6%)
Web sites that triggered 1 distinct diagnostic	0	(0%)	1	(2.6%)
Web sites that triggered 2 distinct diagnostics	0	(0%)	1	(2.6%)
Web sites that triggered 3 distinct diagnostics	1	(1.4%)	1	(2.6%)
Web sites that triggered 4 distinct diagnostics	3	(4.1%)	0	(0%)
Web sites that triggered 5 distinct diagnostics	5	(6.8%)	3	(7.7%)
Web sites that triggered 6 distinct diagnostics	2	(2.7%)	1	(2.6%)
Web sites that triggered 7 distinct diagnostics	5	(6.8%)	3	(7.7%)
Web sites that triggered 8 distinct diagnostics	6	(8.1%)	1	(2.6%)
Web sites that triggered 9 distinct diagnostics	8	(10.8%)	2	(5.1%)
Web sites that triggered 10 distinct diagnostics	8	(10.8%)	6	(15.4%)
Web sites that triggered 11 distinct diagnostics	6	(8.1%)	6	(15.4%)
Web sites that triggered 12 distinct diagnostics	8	(10.8%)	3	(7.7%)
Web sites that triggered 13 distinct diagnostics	6	(8.1%)	1	(2.6%)
Web sites that triggered 14 distinct diagnostics	3	(4.1%)	2	(5.1%)
Web sites that triggered 15 distinct diagnostics	5	(6.8%)	4	(10.3%)
Web sites that triggered 16 distinct diagnostics	2	(2.7%)	0	(0%)

10

Table L.11: **Density of OpenSP diagnostics**

	Subjective		Random	
Web sites that triggered 17 distinct diagnostics	3	(4.1%)	1	(2.6%)
Web sites that triggered 18 distinct diagnostics	0	(0%)	1	(2.6%)
Web sites that triggered 19 distinct diagnostics	2	(2.7%)	1	(2.6%)
Web sites that triggered 20 distinct diagnostics	0	(0%)	0	(0%)
Web sites that triggered 21 distinct diagnostics	0	(0%)	0	(0%)
Web sites that triggered 22 distinct diagnostics	1	(1.4%)	0	(0%)

Figure L.6: **Density of OpenSP diagnostics**
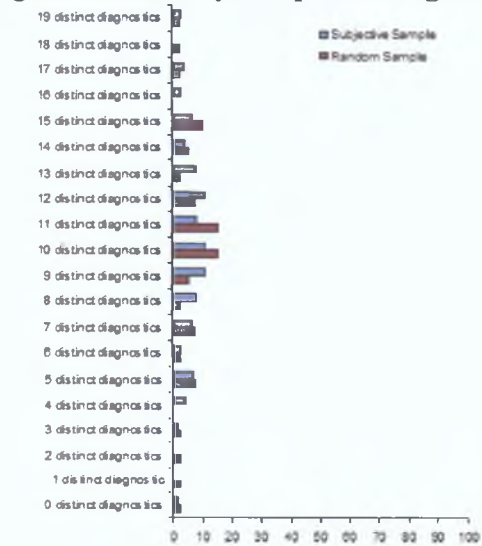
Appendix M

Results of surveys conducted on the subjective and random web samples of Irish web sites (May 2004)

M.1 Web content mirroring process results

Table M.1: **Results of the web content sampling process**

	Subjective	Random
Data amount retrieved	36 MB	26 MB
Web pages considered for mirroring	3,052	2,901
Web pages correctly retrieved (size>0 and no retrieval errors)	2,690 (88.1%)	2,400 (82.7%)
Web pages in web sites qualified for conformance investigations	2,582 (84.6%)	1,966 (67.8%)
Web sites qualified for conformance investigations	162 (75.7%)	110 (40.4%)

Table M.2: **Details regarding the web sites that didn't qualify for investigations**

	Subjective		Random	
Web sites not qualified for investigations	52		162	
No data retrieved	10	(19.2%)	21	(13%)
Only the main page retrieved (size>0)	25	(48.1%)	103	(63.6%)
Less than 3 pages (2) and at least 100 KB retrieved	0	(0%)	0	(0%)
Less than 100 KB and at least 3 pages retrieved	11	(21.2%)	29	(17.9%)
Less than 3 pages (2) and less that 100 KB (but more than 0) retrieved	6	(11.5%)	1	(0.6%)

2

M.2 WCAG 1.0 conformance investigation results

Table M.3: **Details regarding Bobby executions**

	Subjective	Random
Web sites considered by Bobby	162	110
Bobby abnormally terminated	14 (8.6%)	3 (2.7%)
Bobby "locked up"	8 (4.9%)	6 (5.5%)
Bobby completed successfully	140 (86.4%)	101 (91.8%)

Table M.4: **Bobby Priority 1 diagnostics triggered during the investigation**

ID		Subjective	Random
g9	Provide alternative text for all images	113 (80.7%)	92 (91.1%)
g38	Each FRAME must reference an HTML file	25 (17.9%)	23 (22.8%)
g39	Give each frame a title	25 (17.9%)	22 (21.8%)
g10	Provide alternative text for all image-type buttons in forms	23 (16.4%)	12 (11.9%)
g240	Provide alternative text for all image map hot-spots	22 (15.7%)	22 (21.8%)
g21	Provide alternative text for each APPLET	11 (7.9%)	7 (6.9%)

Figure M.1: **Bobby Priority 1 diagnostics**
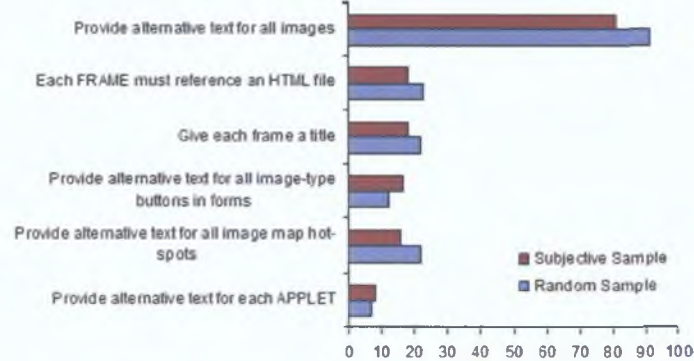
Table M.5: **Density of Bobby Priority 1 diagnostics**

	Subjective		Random	
Web sites with Bobby completed successfully	140		101	
Web sites that triggered 0 distinct diagnostics	21	(15%)	7	(6.9%)
Web sites that triggered 1 distinct diagnostic	65	(46.4%)	48	(47.5%)
Web sites that triggered 2 distinct diagnostic	29	(20.7%)	21	(20.8%)
Web sites that triggered 3 distinct diagnostics	10	(7.1%)	15	(14.9%)
Web sites that triggered 4 distinct diagnostics	10	(7.1%)	7	(6.9%)
Web sites that triggered 5 distinct diagnostics	4	(2.9%)	3	(3%)
Web sites that triggered 6 distinct diagnostics	1	(0.7%)	0	(0%)

4

Figure M.2: **Density of Bobby Priority 1 diagnostics**
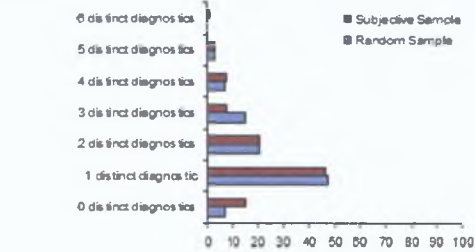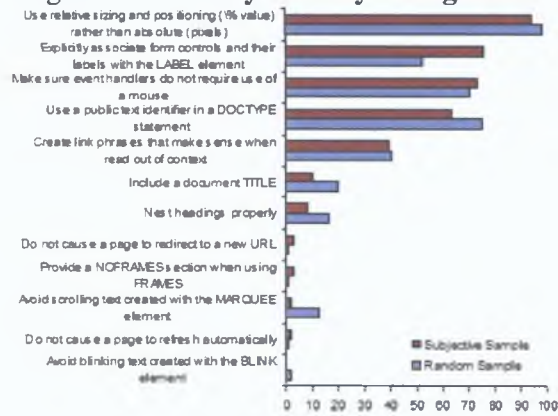

Figure M.3: **Bobby Priority 2 diagnostics**

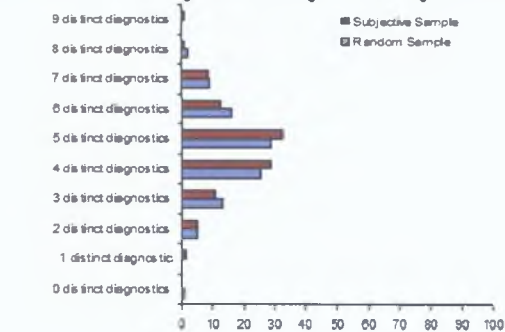Figure M.4: **Density of Bobby Priority 2 diagnostics**

Table M.6: **Bobby Priority 2 diagnostics triggered during the investigation**

ID		Subjective		Random	
g104	Use relative sizing and positioning (% value) rather than absolute (pixels)	132	(94.3%)	99	(98%)
g41	Explicitly associate form controls and their labels with the LABEL element	106	(75.7%)	53	(52.5%)
g269	Make sure event handlers do not require use of a mouse	103	(73.6%)	71	(70.3%)
g271	Use a public text identifier in a DOCTYPE statement	89	(63.6%)	76	(75.2%)
g34	Create link phrases that make sense when read out of context	55	(39.3%)	41	(40.6%)
g273	Include a document TITLE	14	(10%)	20	(19.8%)
g2	Nest headings properly	11	(7.9%)	17	(16.8%)
g37	Provide a NOFRAMES section when using FRAMES	4	(2.9%)	1	(1%)
g254	Do not cause a page to redirect to a new URL	4	(2.9%)	1	(1%)
g5	Avoid scrolling text created with the MARQUEE element	3	(2.1%)	13	(12.9%)
g33	Do not cause a page to refresh automatically	3	(2.1%)	1	(1%)
g4	Avoid blinking text created with the BLINK element	0	(0%)	2	(2%)

6

Table M.7: **Density of Bobby Priority 2 diagnostics**

	Subjective		Random	
Web sites for which Bobby completed successfully	140		101	
Web sites that triggered 0 distinct diagnostics	0	(0%)	1	(1%)
Web sites that triggered 1 distinct diagnostic	2	(1.4%)	0	(0%)
Web sites that triggered 2 distinct diagnostics	7	(5%)	5	(5%)
Web sites that triggered 3 distinct diagnostics	15	(10.7%)	13	(12.9%)
Web sites that triggered 4 distinct diagnostics	40	(28.6%)	26	(25.7%)
Web sites that triggered 5 distinct diagnostics	45	(32.1%)	29	(28.7%)
Web sites that triggered 6 distinct diagnostics	17	(12.1%)	16	(15.8%)
Web sites that triggered 7 distinct diagnostics	12	(8.6%)	9	(8.9%)
Web sites that triggered 8 distinct diagnostics	1	(0.7%)	2	(2%)
Web sites that triggered 9 distinct diagnostics	1	(0.7%)	0	(0%)

7

M.3   HTML technical standards conformance investigation results

Table M.8: **Details regarding OpenSP executions**

	Subjective	Random
Web sites processed by OpenSP	162	110
Web sites where all the web pages had usable document type information	0 (0%)	1 (0.9%)
Web sites considered for HTML defects analysis	80 (49.4%)	45 (40.9%)

Table M.9: **Document Type Declaration in web pages considered by OpenSP**

	Subjective	Random
Web pages processed by OpenSP	2,582	1,966
No document type information	1,497 (58%)	1,310 (66.6%)
Incorrect document type information	16 (0.6%)	70 (3.6%)
Usable document type information	1,068 (41.4%)	549 (27.9%)
Considered for HTML defects analysis (in web sites with at least 3 web pages and 100 KB of data with usable document type information)	1,037 (40.2%)	542 (27.6%)

Table M.10: **The 10 most common OpenSP diagnostics triggered in the investigation**

ID		Subjective		Random	
108	Undefined attribute	73	(91.3%)	42	(93.3%)
79	End tag for not opened element	70	(87.5%)	32	(71.1%)
127	Required attribute not specified	66	(82.5%)	34	(75.6%)
64	Element not allowed by document type	65	(81.3%)	35	(77.8%)
139	Non SGML character number	47	(58.8%)	27	(60%)
25	General entity not defined and no default entity	44	(55%)	22	(48.9%)
68	Missing required end tag (element DTD)	42	(52.5%)	28	(62.2%)
73	End tag for element not finished	39	(48.8%)	16	(35.6%)
131	Attribute value not allowed	35	(43.8%)	21	(46.7%)
82	Attribute value must be literal	26	(32.5%)	23	(51.1%)

Figure M.5: **The 10 most common OpenSP diagnostics**

9
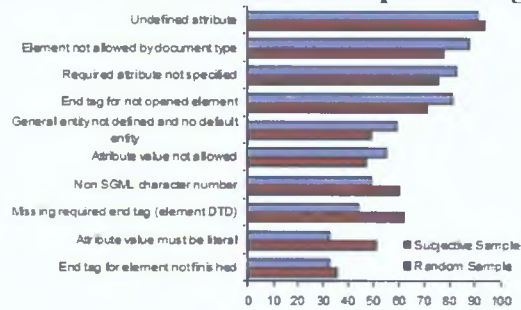
Table M.11: **Density of OpenSP diagnostics**

	Subjective		Random	
Web sites considered for HTML defects analysis	80		45	
Web sites that triggered 0 distinct diagnostics	1	(1.3%)	1	(2.2%)
Web sites that triggered 1 distinct diagnostic	0	(0%)	1	(2.2%)
Web sites that triggered 2 distinct diagnostics	1	(1.3%)	1	(2.2%)
Web sites that triggered 3 distinct diagnostics	0	(0%)	2	(4.4%)
Web sites that triggered 4 distinct diagnostics	3	(3.8%)	0	(0%)
Web sites that triggered 5 distinct diagnostics	6	(7.5%)	4	(8.9%)
Web sites that triggered 6 distinct diagnostics	2	(2.5%)	1	(2.2%)
Web sites that triggered 7 distinct diagnostics	4	(5%)	3	(6.7%)
Web sites that triggered 8 distinct diagnostics	8	(10%)	2	(4.4%)
Web sites that triggered 9 distinct diagnostics	10	(12.5%)	2	(4.4%)
Web sites that triggered 10 distinct diagnostics	8	(10%)	5	(11.1%)
Web sites that triggered 11 distinct diagnostics	4	(5%)	5	(11.1%)
Web sites that triggered 12 distinct diagnostics	11	(13.8%)	5	(11.1%)
Web sites that triggered 13 distinct diagnostics	3	(3.8%)	3	(6.7%)
Web sites that triggered 14 distinct diagnostics	7	(8.8%)	7	(15.6%)
Web sites that triggered 15 distinct diagnostics	8	(10%)	1	(2.2%)
Web sites that triggered 16 distinct diagnostics	0	(0%)	1	(2.2%)

	Subjective		Random	
Web sites that triggered 17 distinct diagnostics	1	(1.3%)	0	(0%)
Web sites that triggered 18 distinct diagnostics	1	(1.3%)	0	(0%)
Web sites that triggered 19 distinct diagnostics	1	(1.3%)	0	(0%)
Web sites that triggered 20 distinct diagnostics	1	(1.3%)	1	(2.2%)

Figure M.6: **Density of OpenSP diagnostics**
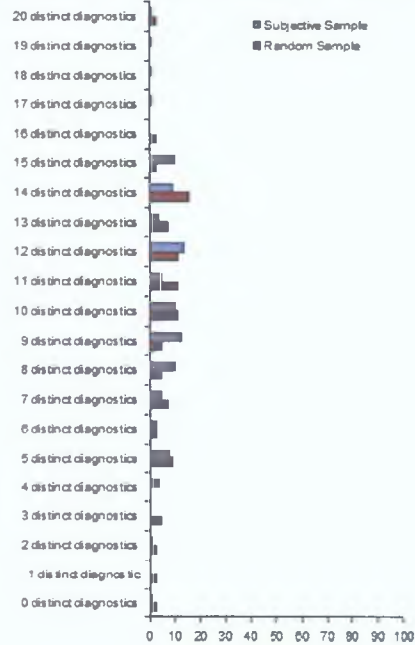
Appendix N

Results of surveys conducted on the subjective and random web samples of Irish web sites (December 2004)

N.1   Web content mirroring process results

Table N.1: **Results of the web content sampling process**

	Subjective	Random
Data amount retrieved	32 MB	24 MB
Web pages considered for mirroring	2,749	2,542
Web pages correctly retrieved (size>0 and no retrieval errors)	2,318 (84.3%)	2,223 (87.5%)
Web pages in web sites qualified for conformance investigations	2,195 (79.8%)	1,769 (69.6%)
Web sites qualified for conformance investigations	143 (66.8%)	100 (36.8%)

Table N.2: **Details regarding the web sites that didn't qualify for investigations**

	Subjective		Random	
Web sites not qualified for investigations	71		172	
No data retrieved	19	(26.8%)	24	(14%)
Only the main page retrieved (size>0)	34	(47.9%)	110	(64%)
Less than 3 pages (2) and at least 100 KB retrieved	0	(0%)	0	(0%)
Less than 100 KB and at least 3 pages retrieved	14	(19.7%)	31	(18%)
Less than 3 pages (2) and less that 100 KB (but more than 0) retrieved	4	(5.6%)	7	(4.1%)

N.2 WCAG 1.0 conformance investigation results

Table N.3: **Details regarding Bobby executions**

	Subjective		Random	
Web sites considered by Bobby	143		100	
Bobby abnormally terminated	8	(5.6%)	1	(1%)
Bobby "locked up"	5	(3.5%)	3	(3%)
Bobby completed successfully	130	(90.9%)	96	(96%)

Table N.4: **Bobby Priority 1 diagnostics triggered during the investigation**

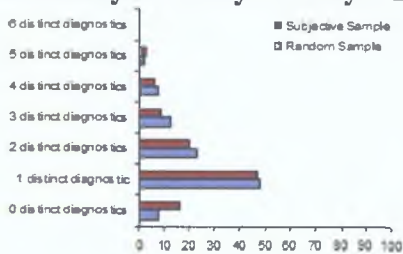| ID		Subjective		Random	
g9	Provide alternative text for all images	106	(81.5%)	88	(91.7%)
g38	Each FRAME must reference an HTML file	20	(15.4%)	20	(20.8%)
g39	Give each frame a title	20	(15.4%)	20	(20.8%)
g10	Provide alternative text for all image-type buttons in forms	20	(15.4%)	11	(11.5%)
g240	Provide alternative text for all image map hot-spots	16	(12.3%)	18	(18.8%)
g21	Provide alternative text for each APPLET	11	(8.5%)	7	(7.3%)

Figure N.1: **Bobby Priority 1 diagnostics**

Table N.5: **Density of Bobby Priority 1 diagnostics**

	Subjective		Random	
Web sites with Bobby completed successfully	130		96	
Web sites that triggered 0 distinct diagnostics	21	(16.2%)	7	(7.3%)
Web sites that triggered 1 distinct diagnostic	61	(46.9%)	46	(47.9%)
Web sites that triggered 2 distinct diagnostic	26	(20%)	22	(22.9%)
Web sites that triggered 3 distinct diagnostics	11	(8.5%)	12	(12.5%)
Web sites that triggered 4 distinct diagnostics	8	(6.2%)	7	(7.3%)
Web sites that triggered 5 distinct diagnostics	3	(2.3%)	2	(2.1%)
Web sites that triggered 6 distinct diagnostics	0	(0%)	0	(0%)

Figure N.2: **Density of Bobby Priority 1 diagnostics**

ID		Subjective		Random	
g104	Use relative sizing and positioning (% value) rather than absolute (pixels)	122	(93.8%)	94	(97.9%)
g41	Explicitly associate form controls and their labels with the LABEL element	100	(76.9%)	47	(49%)
g269	Make sure event handlers do not require use of a mouse	96	(73.8%)	66	(68.8%)
g271	Use a public text identifier in a DOCTYPE statement	75	(57.7%)	74	(77.1%)
g34	Create link phrases that make sense when read out of context	52	(40%)	35	(36.5%)
g273	Include a document TITLE	13	(10%)	17	(17.7%)
g2	Nest headings properly	13	(10%)	13	(13.5%)
g33	Do not cause a page to refresh automatically	4	(3.1%)	1	(1%)
g37	Provide a NOFRAMES section when using FRAMES	4	(3.1%)	1	(1%)
g5	Avoid scrolling text created with the MARQUEE element	2	(1.5%)	11	(11.5%)
g254	Do not cause a page to redirect to a new URL	2	(1.5%)	0	(0%)
g4	Avoid blinking text created with the BLINK element	0	(0%)	1	(1%)

5

Figure N.3: **Bobby Priority 2 diagnostics**
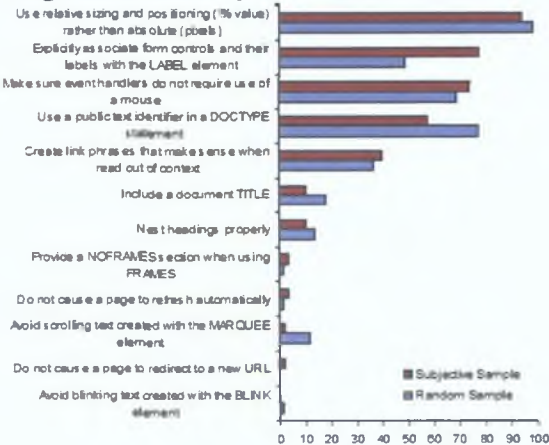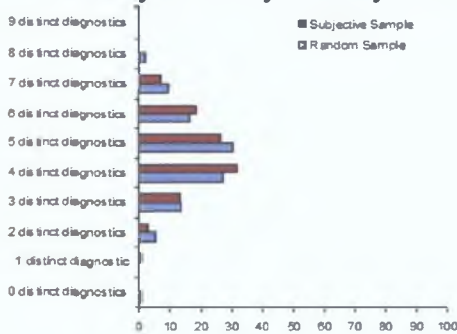

Figure N.4: **Density of Bobby Priority 2 diagnostics**

6

Table N.7: **Density of Bobby Priority 2 diagnostics**

	Subjective		Random	
Web sites for which Bobby completed successfully	130		96	
Web sites that triggered 0 distinct diagnostics	0	(0%)	1	(1%)
Web sites that triggered 1 distinct diagnostic	1	(0.8%)	0	(0%)
Web sites that triggered 2 distinct diagnostics	4	(3.1%)	5	(5%)
Web sites that triggered 3 distinct diagnostics	17	(13.1%)	13	(12.9%)
Web sites that triggered 4 distinct diagnostics	41	(31.5%)	26	(25.7%)
Web sites that triggered 5 distinct diagnostics	34	(26.2%)	29	(28.7%)
Web sites that triggered 6 distinct diagnostics	24	(18.5%)	16	(15.8%)
Web sites that triggered 7 distinct diagnostics	9	(6.9%)	9	(8.9%)
Web sites that triggered 8 distinct diagnostics	0	(0%)	2	(2%)
Web sites that triggered 9 distinct diagnostics	0	(0%)	0	(0%)

N.3 HTML technical standards conformance investigation results

Table N.8: **Details regarding OpenSP executions**

	Subjective	Random
Web sites processed by OpenSP	143	100
Web sites where all the web pages had usable document type information	0 (0%)	1 (1%)
Web sites considered for HTML defects analysis	79 (55.2%)	49 (49%)

Table N.9: **Document Type Declaration in web pages considered by OpenSP**

	Subjective	Random
Web pages processed by OpenSP	2,195	1,769
No document type information	1,206 (54.9%)	1,071 (60.5%)
Incorrect document type information	8 (0.4%)	29 (1.6%)
Usable document type information	980 (44.6%)	632 (35.7%)
Considered for HTML defects analysis (in web sites with at least 3 web pages and 100 KB of data with usable document type information)	958 (43.6%)	623 (35.2%)

Table N.10: **The 10 most common OpenSP diagnostics triggered in the investigation**

ID		Subjective		Random	
108	Undefined attribute	71	(89.9%)	44	(89.8%)
127	Required attribute not specified	69	(87.3%)	39	(79.6%)
79	End tag for not opened element	69	(87.3%)	36	(73.5%)
64	Element not allowed by document type	63	(79.7%)	36	(73.5%)
25	General entity not defined and no default entity	47	(59.5%)	25	(51%)
139	Non SGML character number	44	(55.7%)	33	(67.3%)
68	Missing required end tag (element DTD)	43	(54.4%)	24	(49%)
73	End tag for element not finished	39	(49.4%)	17	(34.7%)
131	Attribute value not allowed	38	(48.1%)	21	(42.9%)
82	Attribute value must be literal	24	(30.4%)	24	(49%)

Figure N.5: **The 10 most common OpenSP diagnostics**
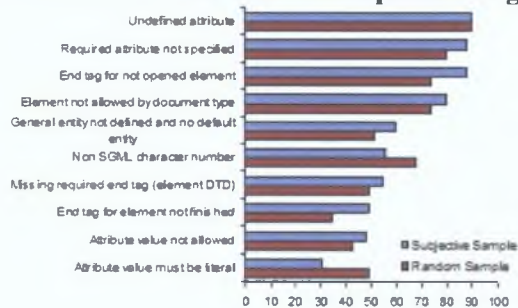
9

Table N.11: **Density of OpenSP diagnostics**

	Subjective		Random	
Web sites considered for HTML defects analysis	79		49	
Web sites that triggered 0 distinct diagnostics	0	(0%)	1	(2.2%)
Web sites that triggered 1 distinct diagnostic	1	(1.3%)	2	(4.4%)
Web sites that triggered 2 distinct diagnostics	1	(1.3%)	2	(4.4%)
Web sites that triggered 3 distinct diagnostics	1	(1.3%)	2	(4.4%)
Web sites that triggered 4 distinct diagnostics	1	(1.3%)	1	(2.2%)
Web sites that triggered 5 distinct diagnostics	6	(7.5%)	2	(4.4%)
Web sites that triggered 6 distinct diagnostics	1	(1.3%)	3	(6.7%)
Web sites that triggered 7 distinct diagnostics	6	(7.5%)	0	(0%)
Web sites that triggered 8 distinct diagnostics	7	(8.8%)	4	(8.9%)
Web sites that triggered 9 distinct diagnostics	7	(8.8%)	1	(2.2%)
Web sites that triggered 10 distinct diagnostics	5	(6.3%)	7	(15.6%)
Web sites that triggered 11 distinct diagnostics	5	(6.3%)	5	(11.1%)
Web sites that triggered 12 distinct diagnostics	12	(15%)	6	(13.3%)
Web sites that triggered 13 distinct diagnostics	5	(6.3%)	3	(6.7%)
Web sites that triggered 14 distinct diagnostics	9	(11.3%)	3	(6.7%)
Web sites that triggered 15 distinct diagnostics	7	(8.8%)	4	(8.9%)
Web sites that triggered 16 distinct diagnostics	1	(1.3%)	2	(4.4%)

10

Table N.11: **Density of OpenSP diagnostics**

	Subjective		Random	
Web sites that triggered 17 distinct diagnostics	2	(2.5%)	0	(0%)
Web sites that triggered 18 distinct diagnostics	1	(1.3%)	1	(2.2%)
Web sites that triggered 19 distinct diagnostics	0	(0%)	0	(0%)
Web sites that triggered 20 distinct diagnostics	1	(1.3%)	0	(0%)

Figure N.6: **Density of OpenSP diagnostics**