Durham University

Durham E-Theses

Web accessibility diagnosis, improvement and maintenance

John Oliver Bailey

How to cite:

Bailey, John Oliver (2007) Web accessibility diagnosis, improvement and maintenance. Doctoral thesis, Durham University.

Use policy

The full-text may be used and/or reproduced, and given to third parties in any format or medium, without prior permission or charge, for personal research or study, educational, or not-for-profit purposes provided that:

- a full bibliographic reference is made to the original source
- a https://etheses.durham.ac.uk/id/eprint/2140/ is made to the metadata record in Durham E-Theses
- the full-text is not changed in any way

The full-text must not be sold in any format or medium without the formal permission of the copyright holders.

Please consult the full Durham E-Theses policy for further details.

Academic Support Office, The Palatine Centre, Durham University, Stockton Road, Durham, DH1 3LE
e-mail: e-theses.admin@durham.ac.uk Tel: +44 0191 334 6107
https://etheses.durham.ac.uk

Web Accessibility: Diagnosis, Improvement and Maintenance

John Oliver Bailey

The copyright of this thesis rests with the author or the university to which it was submitted. No quotation from it, or information derived from it may be published without the prior written consent of the author or university, and any information derived from it should be acknowledged.

Ph.D. Thesis

2007

Department of Computer Science

Durham University

17 OCT 2007

Abstract

Context:

This thesis examines how organisations create and maintain their web pages with particular focus on ensuring pages are accessible. It also investigates the potential for using a Tree-Map based tool to support such web maintenance and process improvement. Novel process improvement recommendations are given and an adaptation of a class web publishing model is presented.

Methods:

To supplement a review of current literature, 20 accessibility specialists and 79 large organisations were surveyed. This identified web accessibility best practices and whether these practices were implemented in the reality. A subsequent assessment of the accessibility of each organisation's web site tested if certain activities could be linked with better accessibility. Finally, a controlled experiment tested the accuracy and efficiency of a Tree-map based tool for web maintenance.

Results:

The survey results suggested a wide variety of web accessibility awareness amongst web developers and accessibility specialists. Best practice appeared to be implemented by many organisations with the exception of training provision. It was found that when the best practices aimed specifically at web accessibility were implemented there was a significant improvement in web accessibility. The Tree-Map based tool was proved to be more efficient than and as accurate as report based tool for web maintenance activities.

Conclusions of the study:

Web accessibility awareness is now reasonably high amongst web developers but the extent to which it is addressed varies. Organisations which take a systematic and mature approach to accessibility have more accessible web sites. As such, accessibility should be integrated into web publishing. Better tools are also required to facilitate this systematic integration.

Copyright Notice

The copyright of this thesis rests with the author. No quotation from it should be published without their prior written consent and information derived from it should be acknowledged.

Declaration

No part of this material has previously been submitted for a higher degree at Durham University or any other University.

The following publications were produced during the course of this thesis:

Tree-Map Visualisation for Web Accessibility. John Bailey and Elizabeth Burd, International Computer Software and Applications Conference, Edinburgh, Scotland, July 26-28 2005.

Web Accessibility Evolution in the United Kingdom. John Bailey and Elizabeth Burd, International Symposium on Web Site Evolution, Budapest, Hungary September 26, 2005

What is the current state of web accessibility? John Bailey and Elizabeth Burd, International Symposium on Web Site Evolution, Philadelphia USA September 23-24, 2006.

Acknowledgments

My respect and deep gratitude are owed to Liz my supervisor, for firstly providing me with the opportunity to do a PhD, and then supporting me with ideas, advice and counsel. Special thanks must also go to her mother who had the unenviable task of proof reading my 1st draft.

I would also like to thank my parents who have inspired and supported me throughout.

Finally I want to acknowledge my Fiancée Marysia who has keenly organised and lived all my trials and tribulations in the last year.

Table of Contents

Abstract ii
Declaration iii
Acknowledgments iv
Table of Contents v
Table of Figures xi
Table of Tables xii
1 Introduction 1
1.1 Accessible Web Publishing 1
1.2 Research Problems 2
1.3 Criteria for Success 3
1.4 Thesis Outline 3
2 Literature Survey 6
2.1 Web Maintenance 6
2.1.1 Nature of the Web 6
2.1.2 Web Publishing and Content Management 8
2.1.3 Content Lifecycle Model 11
2.1.4 Content Management Roles 12
2.1.5 Web Maintenance 15
2.1.6 Differences between web and traditional software development 16
2.2 Web Accessibility 18
2.2.1 Definition 18
2.2.2 General Overview 18
2.2.3 Principles of web accessibility 19
2.2.4 Main beneficiaries of accessible web sites 20
2.2.5 Importance of Web Accessibility 30
2.2.6 Implementing web accessibility (Best Practice) 35
2.3 Web accessibility assessment 36
2.3.1 Accessibility guidelines and standards 36
2.3.2 Assessment tools 38
2.3.3 Quantitative measurement of web accessibility 39
2.3.4 Current State of Web Accessibility 42
2.4 Information Visualisation 43

2.4.1 Goals of Visualisation 44
2.4.2 Potential Techniques for Web Sites 45
2.4.3 Graphs 45
2.4.4 Issues in Graph Visualisation 46
2.4.5 Graph Layout 47
2.4.6 User Interaction 48
2.5 Summary 50
3 Tree-Map Visualisation of Web Accessibility 51
3.1 Introduction 51
3.1.1 Visualising web accessibility within a Tree-Map 51
3.1.2 Benefits of Tree-Maps 51
3.2 Tool Implementation 56
3.2.1 Overview of stages 56
3.2.2 Stage 1: Collection of web pages 56
3.2.3 Stage 2: Automated accessibility assessment of pages 57
3.2.4 Stage 3: Generation of a PageMeasure for each page 62
3.2.5 Stage 4: XML representation of the web site 63
3.2.6 Stage 5: Generation of Tree-Map with maintenance information 65
3.3 Usage Scenarios 68
3.3.1 Effect of web accessibility improvement 68
3.3.2 Effect of new pages 71
3.3.3 Effect of different 3rd party product conversion tools 72
3.4 Summary 73
4 Studies and Experiment 75
4.1 Introduction 75
4.2 Problem Description 75
4.3 Study Outline 76
4.3.1 Introduction 76
4.3.2 Clarification of Terms 77
4.3.3 Aims of Study 77
4.4 ASP Survey 78
4.4.1 Accessibility specialist criteria 79
4.5 WMP Survey 81
4.5.1 Organisations contacted 81

4.5.2 Questionnaire Design 82
4.6 Accessibility Assessments 85
4.6.1 Automated Assessment Method 85
4.6.2 Manual Assessment Method 86
4.7 Tree-Map Experiment 87
4.7.1 Research Question for this study 87
4.8 Experimental Design 88
4.8.1 Web site selection 88
4.8.2 Sample Population 88
4.8.3 Selection and Preparation of Web sites 89
4.8.4 Presentation of Maintenance Information 90
4.9 Data preparation and Collection 94
4.9.1 Data preparation 94
4.9.2 Data Collection 94
4.9.3 Tasks 95
4.9.4 Step by Step Plan 95
4.10 Analysis 97
4.10.1 Hypotheses 97
4.10.2 Validity 97
4.10.3 Limitations of the Study 98
4.10.4 Deviations from Protocol 99
4.11 Summary 100
5 Study and Experiment Results 101
5.1 Introduction 101
5.2 Respondents Background: ASP Survey 101
5.2.1 Country of Residence 101
5.2.2 Experience 102
5.2.3 Respondents Web Accessibility Roles 103
5.2.4 Guidelines and Tools Used 104
5.2.5 Accessibility Issues 105
5.3 Discussion of Qualitative Data 108
5.4 Respondents Background: WMP Survey 113
5.4.1 Type of Organisation 113
5.4.2 Respondents Roles 114

5.4.3 Web Experience 115
5.4.4 Size of Organisations and their Web Teams 115
5.4.5 Web Publishing Practices 117
5.4.6 Accessibility Practices 118
5.4.7 Tools and guidelines used to assess web accessibility 122
5.4.8 3rd Party Software and Training Practice 123
5.4.9 Roles receiving accessibility training 125
5.5 Accessibility Assessments 125
5.5.1 Accessibility Ratings 125
5.5.2 Correlation between automated and manual assessments 127
5.5.3 Effect of Web Publishing Practices on Accessibility 128
Practices specific to Accessibility 132
5.5.4 Affect on accessibility when practices are combined 136
5.5.5 Third Party Content and Training Practices 138
5.6 Tree-Map Experiment Results 140
5.6.1 Group organisation 140
5.6.2 Time to complete tasks 140
5.6.3 Accuracy of responses 141
5.6.4 Analysis of Task Completion Time 141
5.6.5 Analysis of Task Accuracy 143
5.7 Summary 144
6 Evaluation of Results 147
6.1 Increasing sophistication in web publishing 147
6.1.1 More formalised approach to the Web 147
6.1.2 Automated content management 149
6.2 Levels of adaptation for Web Accessibility 150
6.2.1 Working Environment 150
6.2.2 Consultants 151
6.2.3 Training 152
6.2.4 Accessibility Gurus 154
6.2.5 Budget allocation 156
6.2.6 Accessibility tool support 156
6.2.7 User awareness 158
6.2.8 Coping with long term change 160

6.3 Evaluation of the impact of practices on accessibility 162
6.3.1 General accessibility assessment 162
6.3.2 Practices in isolation 163
6.3.3 Practices combined 164
6.3.4 Raising accessibility awareness 165
6.4 Tree-Map Based Tool Support 165
6.4.1 Functionality: Accuracy 165
6.4.2 Usability: Learnability 166
6.4.3 Efficiency: Time Behaviour 167
6.5 Summary 169
7 Process Improvement Recommendations 171
7.1 Introduction 171
7.2 Integrated Accessibility Publishing Model 171
7.2.1 Accessibility Integration Points 171
7.2.2 Tasks at Compile Content Stage (A1) 173
7.2.3 Tasks at the Quality Control stage (A2) 176
7.2.4 Tasks at Authorisation stage (A3) 178
7.2.5 Tasks at Publish stage (A4) 179
7.3 Benefits of the accessibility log 179
7.3.1 Quantitative analysis of web accessibility 180
7.3.2 Record of work completed 181
7.4 Knowledge Repository 182
7.5 Overview and reporting tool support 182
7.5.1 Tool support for multiple roles 183
7.6 Summary 189
8 Conclusion 191
8.1 Criteria for Success 191
8.1.1 Survey current web accessibility best practice 191
8.1.2 Evaluate the effect of current web accessibility best practice and guidelines on accessibility 192
8.1.3 Develop and evaluate a semi-automated web accessibility management tool 192
8.1.4 Develop an integrated web publishing model 193

8.1.5 Create recommendations of tool usage and systematic web accessibility improvement 194
8.2 Further work 194
8.2.1 Short term development 194
8.2.2 Long term development 195
8.3 Thesis Summary 196
8.3.1 Diagnosis 196
8.3.2 Improvement 196
8.3.3 Maintenance 197

Table of Figures

Figure 2-1 Weinstein's Content Lifecycle [112] 11
Figure 2-2 A framework for Web content management [105] 12
Figure 2-3 Example of a table used for layout 23
Figure 2-4 Result of linearization from Figure 2-3 24
Figure 2-5 Table used to layout a menu which will linearise sensibly 24
Figure 2-6 Result of linearisation from Figure 2-5 24
Figure 2-7 W3C Web Content Accessibility Guidelines 1.0 37
Figure 2-8 Zeng's WAB score formula [116] 41
Figure 3-1 Example of Tree-Map with magenta "Inaccessible" web pages 54
Figure 3-2 Example Tree-Map with four web pages highlighted with yellow border 55
Figure 3-3 Potential accessibility barrier within PageValet report 60
Figure 3-4 Transformed Test Result from PageValet 61
Figure 3-5 Page Measure Equation 62
Figure 3-6 URLs were used to form the site hierarchy 63
Figure 3-7 XML DTD for JTreemap 64
Figure 3-8 Partial XML representation of web site 65
Figure 3-9 UML class diagram of PageInfo and related classes 66
Figure 3-10 Tree-Map of Postgraduate section before accessibility improvement 69
Figure 3-11 Postgraduate section after accessibility improvement 70
Figure 3-12 Tree-Map of Computer Science Department web site 71
Figure 3-13 Tree-Map of Computer Science Department web site with new section added 71
Figure 3-14 Slides created by default from MS PowerPoint 72
Figure 3-15 Slides created by using the accessible wizard [103] from MS PowerPoint 73
Figure 4-1 Screen shot of SiteValet report header 90
Figure 4-2 Screen shot of the barriers found by SiteValet 91
Figure 4-3 Meta-data for each Page in Tree-Map Tool 92
Figure 4-4 Screenshot of Tree-Map based tool 93
Figure 5-1 Experience in years of specialists 103
Figure 5-2 Reported web maintainer initial accessibility awareness upon arrival 106
Figure 5-3 Why accessibility specialists believed they were hired. 107

Figure 5-4 Organisation Types 114
Figure 5-5 Web Experience of Respondents in the WMP survey 115
Figure 5-6 Web Publishing Practices Questions and Answers 118
Figure 5-7 Accessibility Practices Questions and Answers 121
Figure 5-8 Web accessibility assessment tools used by respondents to the WMP survey 122
Figure 5-9 3rd Party Requirements and Training Practices Questions and Answers 124
Figure 5-10 Automatic Accessibility Assessment Categories 126
Figure 5-11 Manual Assessment Accessibility Categories 127
Figure 5-12 Time differences in minutes for tasks completion time using Tree-Map based tool. 140
Figure 5-13 Accuracy differences for tasks using Tree-Map based tool. 141
Figure 5-14 ANOVA of task completion time differences for groups 142
Figure 7-1 Adaptation of Weinstein's content lifecycle model [112] with accessibility points labelled 172
Figure 7-2 Incorrect usage of strong tag to mimic heading style. 175
Figure 7-3 Correct use of heading tag to semantically define a heading. 175

Table of Tables

Table 2-1 Web content stages and their associated roles 13
Table 2-2 Differences found in web development and their effect on maintenance ... 17
Table 2-3 Assistive technologies to compensate for blindness 22
Table 2-4 Assistive technologies to assist users with low vision 25
Table 2-5 Summary of countries national accessibility legislations 35
Table 2-6 Checkpoint Priorities [26] 37
Table 2-7 Lazar's web page accessibility categories [66] 40
Table 3-1 Score Range to Category 54
Table 3-2 Stages involved in the generation of the Tree-Map 56
Table 3-3 PageInfo class main attributes 66
Table 3-4 PageAssLog Attributes 67
Table 3-5 AccessTask Attributes 68
Table 4-1 Surveys used in study 77
Table 4-2 Aims of Study 78

Table 4-3 Details of which group performed which task 95
Table 5-1 Country of Residence 102
Table 5-2 Years of Experience working with the Web and Accessibility 102
Table 5-3 Roles of respondents and the percentage who carried them out in Survey One 104
Table 5-4 Activities carried out by specialists 104
Table 5-5 Guidelines used by specialists 105
Table 5-6 Tools used by specialists 105
Table 5-7 Roles of respondents and the percentage who carried them out in the WMP survey 114
Table 5-8 Sizes of respondent organisations 116
Table 5-9 Number of individuals involved in web publishing where respondents work 116
Table 5-10 Web accessibility guidelines used by respondents to the WMP survey.. 122
Table 5-11 Roles which received accessibility training 125
Table 5-12 Explanation of t-test results 129
Table 5-13 Number of practices implemented by organisations which monitor and record access data 136
Table 7-1 Description of each stage of the adapted publishing model. 173

1 Introduction

1.1 Accessible Web Publishing

Web accessibility measures how easily diverse sets of users, regardless of disability or environmental constraints, can access material on a website [75]. Different groups are hindered by different accessibility barriers and so a wide variety of barriers must be taken into account to ensure a document is accessible to the widest possible audience. With the integration of the World Wide Web (web) into everyday life it can no longer be viewed as a "nice-to-have facility" [54] nor then can potential "accessibility barriers" [26] be ignored.

As the web grows in size and technological sophistication, so too have the number of online services provided through the web increased. There has been a change in focus of the web, from an information source for academic research, to a medium through which important daily services are provided. Delivering these services in an inaccessible form undermines the potential value of online services for some individuals. One sector which has embraced online services is banking. Many banks offer internet banking allowing customers to transfer money and pay bills through their web site. On line services such as these offer great benefits to elderly customers, who through diminishing mobility may find visiting their local branch difficult. However, if the web site is designed without accessibility in mind, for example, relying on the ability to use a mouse to navigate, it creates an accessibility barrier. Elderly users who have poor fine motor control, or hand tremors will find navigating using only a mouse extremely difficult.

Accessibility barriers make web sites inaccessible to certain sets of users. The individuals most affected by inaccessible web sites are:

- Individuals with mental and physical disabilities.
- Individuals with age related conditions.
- Individuals using older, slower web technologies.

In recognition of this many governments have introduced legislation protecting the right to information for web site users with disabilities (discussed in section 2.2.5.4). As such, organisations must now consistently produce accessible web content, and for this to be achieved, organisations require:

1. Tools to assess and locate web accessibility barriers.
2. The knowledge to remove these barriers.
3. A web publishing model which ensures that the creation of barriers is the exception rather than the norm.

Without this third requirement, ensuring web accessibility becomes a fire fighting exercise, where after every update, valuable content development time is spent on fixing barriers created in the previous update. An accessible web publishing model should provide the support mechanisms and workflow to ensure that content is not published before it is assessed for accessibility and all relevant barriers have been removed. Rather than viewing accessibility as an optional non-functional requirement, it must be viewed as an integral part of a web publishing strategy. The European Union recognised this in its eAccessibility of public sector services in the European Union report as follows:

"*When inclusion is built-in to public service design from the outset, individual opportunities in education, employment, health and social life are enhanced and this, in turn, has the potential to bring about a significant economic impact in Europe.*" [45]

1.2 Research Problems

Despite recent government legalisation [49] [51] [52] and the existence of internationally recognised web accessibility guidelines for the production of web content [26], there are still many important web sites which remain inaccessible [9]. Therefore, the main research problem to be addressed is - How does an organisation maintain a consistently high level of accessibility throughout a web site? From this the following sub-problems follow:

- Is current accessibility best practice being implemented?
- Is this best practice sufficient to achieve consistently high accessibility?

- How should accessibility integrate into the web publishing model?
- Can tools provide the support needed to manage a web site and in particular its accessibility?

1.3 Criteria for Success

In order to address the research problems outlined in section 1.2 criteria for success have been created. With these criteria, this thesis aims to contribute practical and applicable data to the current state of the art in web accessibility. The criteria for success are as follows:

- **Survey current web accessibility best practice** – it is important to establish whether best practice has been adopted by large organisations. If best practice is being ignored by large organisations, then one conclusion will be that more work on dissemination is needed.
- **Evaluate the effect of current web accessibility best practice and guidelines on accessibility** – where best practice is implemented what effect is this having on the overall accessibility of the web site? If there is no improvement, then either the practice is not being implemented fully or correctly, or the best practice may require changing.
- **Develop and evaluate a semi-automated web accessibility management tool** – this thesis will present a tool which supports web accessibility management. Its aim is to allow the quick identification, diagnosis and correction of pages with accessibility barriers.
- **Develop an integrated web publishing model** – there has been little work on the integration of a web publishing model with accessibility practices. This thesis will adapt an existing web publishing model to take accessibility into account.
- **Create recommendations of tool usage and systematic web accessibility improvement** – finally, based on existing literature and from novel research, recommendations for organisations to systematically improve their web accessibility will be created.

1.4 Thesis Outline

The structure of this thesis is as follows:

Chapter 2 presents a survey of current literature concentrating on three areas of research:

- **Web maintenance** – firstly, the general maintenance of web sites is presented, including the different types of web sites, web publishing models and the roles which are involved in the publication and maintenance of a web site.
- **Web accessibility** – here the focus is on principles behind accessible design and also the technical specifics of web accessibility. There is a discussion on the reasons why web accessibility is important and the specific consequences of poor accessibility.
- **Information visualisation** – this section discusses various methods to visualise information with particular attention paid to the advantages and disadvantages of graphs and Tree-Maps.

Chapter 3 introduces a Tree-Map based tool aimed at supporting web accessibility management. The background and justification for using Tree-Map visualisation as the front-end for the tool is then explained. Following this, general design and implementation detail are outlined providing details of how a model of the web site is developed and how the accessibility of individual pages is measured and then represented as nodes within the Tree-Map visualisation. Finally, usage scenarios are given, showing, for example, how the Tree-Map based tool can be used to identify inaccessible web pages.

Chapter 4 describes a study and controlled experiment carried out as part of this work. The study consists of two surveys: A small scale open ended qualitative survey of expert opinion. The experts surveyed were individuals who worked at either improving accessibility of web sites within their own organisation or as a consultant for others. The results gathered from this first survey and follow up correspondences were then used in combination with current literature to design a second larger scale quantitative survey. The second survey was aimed at those responsible for more general web site maintenance and surveyed how they published their web sites. Following this, the details of a follow up web accessibility assessment are given, which investigated the relationship between best practice and web accessibility. The

second part of Chapter 4 gives the protocol for a controlled experiment using a restricted version of the Tree-Map tool to evaluate if the tool could be used to convey accurate web maintenance information efficiently.

Chapter 5 details the results for the study and controlled experiment described in Chapter 4. It contains discussion of the results from the two surveys within the study and also statistical analysis of the results from the controlled Tree-Map usage experiment.

Chapter 6 gives an evaluation of the results, discussing high level themes and evaluation which can be made from the results. For each theme, specific web site scenarios are given to underline them.

Chapter 7 make process improvement recommendations to allow the integration of accessibility into an organisation's web publishing model. It builds on Weinstein's web content lifecycle [113] to identify which accessibility practices should be implemented and where in the lifecycle they should occur.

Chapter 8 concludes and summarises the work contained within the thesis and provides details of future work which could develop the state of the art.

2 Literature Survey

This chapter presents a survey of literature relevant to this thesis. The topics surveyed are:

- Web maintenance
- Web accessibility
- Information visualisation

2.1 Web Maintenance

The importance of web maintenance has evolved in step with the importance of the web. An increasing world wide audience and cheaper and faster Internet connections have fuelled the transformation of web maintenance into a managed and more automated process.

2.1.1 Nature of the Web

Information is at the core of the web [30]. It has enhanced our way of managing and exchanging it. It has evolved from a simple method for scientists to collaborate remotely, to become the largest information network on the planet. Its audience and user group has grown from hundreds to hundreds of millions [46], so too, the kind of information being exchanged has grown and diversified. The web is now, not just an important tool for science but also for business [81], governance [44], education and society in general. And as usage has increased and diversified, so too has the type and amount of data provided on web sites. Web pages consist now of more than only hypertext and images and can mimic functionality previously seen in desktop software applications. Dart [39] outlined 11 classes of web sites, from purely informational (brochure ware) to interactive web applications [84];

1. **Informational** – delivering information to users about products or services, so called "brochure ware".
2. **Delivery System** – files are delivered for download to the user.

3. **Customised access** – the appearance of the web site can be tailored to specific user preferences. An example might be an online book seller which highlights specific book recommendations to users based on previous purchases.
4. **User-provided information** - information is sought from the user, through web based forms. Examples include surveys or subscriptions to web based services.
5. **Interactive** – the web site acts as a communication tool between organisations, users and resources, for example, an online loan application web site and loan brokers. The web site allows users to search for specific loan products and then passes on the users' information to specific loan brokers.
6. **Transaction oriented** – the web site offers services or products to be bought directly through it.
7. **Service Provider** – the web site provides a service (such as online virus scanning) to users on a pay-per-use basis.
8. **Database access** – the web site acts as user interface for a database. User can formulate queries using a web based form and the results are returned via the web site.
9. **Document access** – similar to database access, but specifically tailored to the access of documents. For example, an online journal allowing users to search for publications matching specific keywords.
10. **Workflow oriented** – the web site follows a specific work flow or process. For example, and online job application, ensuring that all stages are completed in order, before allowing the user to send the application.
11. **Automatic content generator** – web sites which are created automatically from content found on other sites, for example, a news compilation web site.

For each of these web site classes, content generation (the creation and composition of information (content)) differs and so each class will require different processes and content models. Many web sites will contain more than one class of web system, for example, a holiday company may use its web site not only as a brochure (1) listing the package holidays on offer, but also allow users to book their holidays(6) through the web site and register for regular updates on special offers(4).

The evolution of the typical web application from a small scale online brochure to a professional service delivery medium has been rapid. This shift has also meant a rise in the complexity and of volume of work involved in creating an up-to-date and accurate web site. Just as the applications of the web have changed, so too have the models for creating and maintaining a reasonably sized web site. In 1990s a typical model was that of a centralised "web editor" [62] (or small team of web editors depending on the size of the organisation) with almost unique control of the web site. They create the web pages for the whole organisation. To update or add a page, the web editor has to make the change themselves. Such a model obviously leads to bottlenecks [106], and has hence caused the introduction of software and processes based around traditional publishing techniques.

2.1.2 Web Publishing and Content Management

CERN, the organisation which invented the web defines web publishing as "the process or making a web document available to the public" [17]. Web or online publishing aims like, traditional paper based publishing, to put information or content into the public domain. It is, however, subject to very different requirements than its paper based alternative. Cuenca [37] points out that "differences in the characteristics of both types of media... have a profound impact on the production processes that are being employed". These requirements are listed by Cuenca [37] as follows:

- Access to the most up-to-date information.
- Improved functionality over any other system.
- Personalisation.
- User Involvement.

As web pages are a digital media, the time required to update them is limited mainly by human rather than physical conditions. Distribution to the reader over the web is done in seconds rather than hours. Instead of going to a shop to buy a newspaper or book, the reader needs only to reload the web page to view the latest version of a web page. This increased speed has led to a greater expectation on behalf of the user to have up-to-date information. The distributed nature of the internet also allows content to be sourced from multiple geographically remote locations. Vidgen [106] points out that for a web site to 'live and breathe' it must be updated constantly with new content and have expired content removed.

Another important consideration is that with web publishing is that the reader requires both internet enabled hardware and web browser software to access the content. There is a diverse range of internet enabled hardware devices and browsing. These include:

- Standard Web Browser (on a PC)
- Mobile Device Browser (on PDA or Mobile Phone)
- WebTV (through Digital TV)
- Screen Reader (on any Internet enabled device)

All these technologies allow the reader to take advantage of the increased potential functionality offered by the Web. For instance, hyperlinks allow readers to visit related pages within and across publications.

Cuenca [37]also lists several approaches to online publishing as follows:

- Trivial semi-automatic generation of online products
    - This involves the transfer of existing published materials to the Web.
- Ad-hoc development
    - Quickly developed solutions usually on a fairly small scale
- Software (Web) Engineering
    - To address the need for more complicated web systems, software engineering approaches must be applied and adapted to the Web.

Unlike traditional publishing, web sites can include features such as search engines, user feedback forms and personalisation options. This extra functionality requires either the ad-hoc development or the web engineering approaches mentioned by Cuenca [37]. User feedback and preferences can be processed and stored, allowing web sites to offer the reader a view of the web site that meets their specific browsing needs.

All web sites employ some form of content management. The extent of the management depends on the amount of content to be published. A small hobby web site does not require large scale content management software, whereas a large or

even medium sized organisation will require something that allows the creation, management and publication of content to be semi-automated and devolved. By semi-automating these processes, the entire organisation become potential contributors adding content and helping the site grow. Vidgen [106] highlights several web content management issues which must be addressed to ensure an efficient and reliable web publishing:

- **Bottlenecks** – content may be generated by many different sources and in varied formats, hence the task of compiling, formatting and publishing is a potential bottleneck if it is being performed by a small group of web managers.
- **Consistency** – ensuring a consistent "look and feel" and consistent standards of HTML is difficult where web editing is devolved to different organisations. This can be especially true of multi-national, multi-lingual sites. Consistency can also be applied to the information being presented. If a date or address has changed, this change must be reflected throughout the web site.
- **Navigation** – if the web site structure is not controlled the site can grow, becoming difficult to navigate and hindering information searching.
- **Data duplication** – different sections of an organisation might publish the same information in two sections of the web site without realising it.
- **Content audit and control** – procedures and controls must be in place to ensure that
    a) inappropriate material is not published and
    b) if it is, then the content and content authoriser can be identified quickly and steps are taken to ensure it doesn't happen again.
- **Tracking** – tracking uses content meta-data to provide information such as: when content was last updated, the person who created it, authorised it, and is responsible for it, and when it is due for removal.
- **Business processes** – web content management must be linked to business processes to ensure when documents produced for the business are created or updated, this is then reflected on the relevant web pages.

One issue not mentioned above is that of training. Without sufficient training even the most comprehensive content management software is useless in preventing data duplication, inconsistencies and navigation problems.

2.1.3 Content Lifecycle Model

Two of Cuenca's requirements [37] for web publishing, i.e. up-to-date information and personalisation require a more formal control over the content published. These relate to the issues Vidgen [106] refers to as "content audit and control" and "tracking". Meta-data is required to describe the content and a lifecycle model is required to ensure that only relevant and suitable content is published.

Meta-data is often added to content to allow web sites to customise the content delivered to specific users. For example, an online book store may wish to present different special offers to users, who have expressed an interest in "Polish 19th century literature" than to users who have previously bought books on "ancient Greece". Here, the meta-data about the book used is its "subject area" and the lifecycle model should ensure that only current special offers are shown to the user. So offers that were available last month will be stored in the archives rather than published on the current version of the web site.

Figure 2-1 shows Weinstein's content lifecycle model. This is a simple 6 stage model and represents a generally accepted approach to content lifecycle representation.

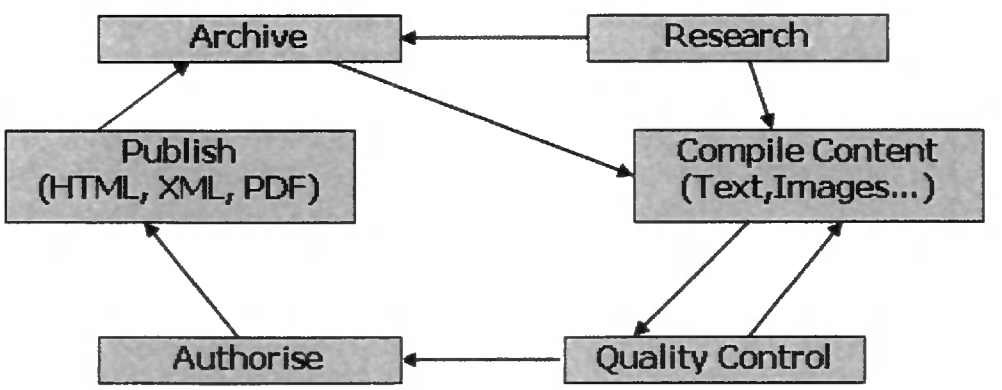

Figure 2-1 Weinstein's Content Lifecycle [113]

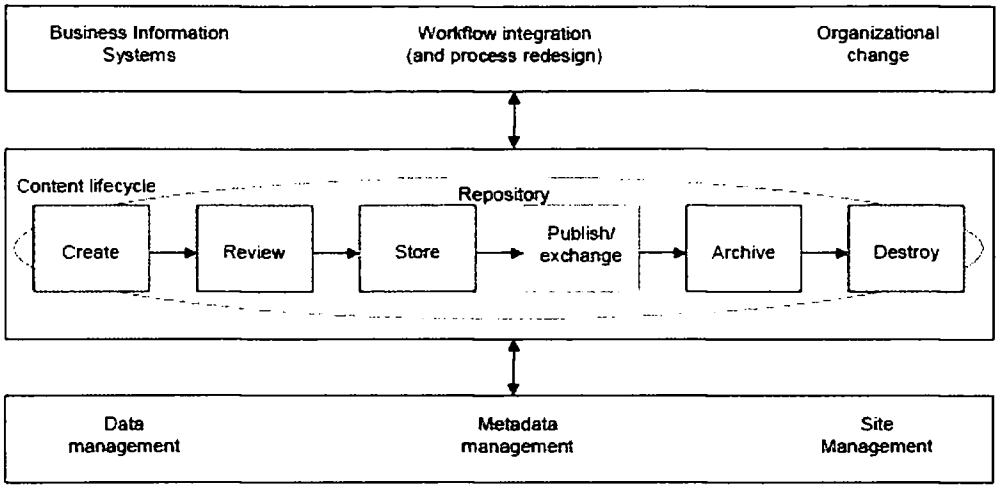

Figure 2-2 A framework for Web content management [106]

Figure 2-2 details Vidgen's framework for web content management [106]. The content lifecycle is similar to Weinstein's content lifecycle with the addition of a final "destroy" stage, where content is removed permanently from storage. Vidgen [106] lists 3 extra requirements specific to publication stage of the content lifecycle:

- *Authentication*: ensuring that only those with the correct authorisation can publish web pages.
- *Personalisation*: giving individuals views of content specific to their needs, or authorisation.
- *Transformation*: constructing the page may require some form of on-demand transformation. Taking content from various sources, collating it and converting it to a form suitable to the reader's agent. This may involve adapting the presentation for a hand-held PDA or other mobile computing device.

2.1.4 Content Management Roles

Just as web sites have evolved from simple static content to incorporate more functional and dynamic content, so too have the roles involved in managing that content. Increased sophistication and devolution of web publishing has meant that new roles have been defined in many organisations to deploy content. Table 2-1 lists common roles adopted used in web publishing. However, the boundaries between certain roles are not always fixed. Schmeiser [91] highlights this stating that "the fluidity between different fields of web design is what appeals to so many people".

Also, there are no officially recognised definitions for content management and the precise nature of each role usually varies on an organisational basis.

Stage	Role
Research	Content author
Compilation of Content	Content author Content coordinator
Quality Control	Content quality manager
Authorise	Business quality manager Content quality manager
Publishing	Content quality manager
Archive	Records manager

Table 2-1 Web content stages and their associated roles

2.1.4.1 Webmaster or Custodian

The webmaster or custodian is a general term assigned to someone responsible for administering or designing a web site. Spainhour [99] defines a webmaster as "a person who works on the content end of the web site" The Oxford English Dictionary defines a webmaster as "The administrator of a web site (who may also be its designer)." This is a general term reflecting the broad nature of the role. The activities typically involve administration of the web server and server side programming, but it might also involve any of the activities covered by the following roles.

2.1.4.2 Web maintainer

Web maintainers undertake a wide range of activities similarly to a webmaster but are not administrators of the whole site. They are involved in technical aspects of web maintenance such as fixing errors in web systems (corrective maintenance) or expanding the functionality of the web applications (perfective maintenance).

2.1.4.3 Content Coordinator

Activities performed by a content coordinator [103] will include:

- Identifying authoritative sources to link to

- Writing content and collaborating with content authors to create, edit and source content.
- Monitoring content expiry and user feedback, and advising the content quality manager.

2.1.4.4 Content author

Content authors [50] are responsible for writing and developing web content. The main activities associated with this are as follows:

- Creating standards compliant (both internal and external standards) content for the web page.
- Generating appropriate meta-data for the content created.
- Submitting content to the content coordinator / content quality manager for publication.

N.B. The two roles of content author and coordinator are also referred to as that of a content provider [91].

2.1.4.5 Content Quality Manager

It is the responsibility of the content quality manager to ensure that a high level of quality is maintained. Typical activities involved in this role are as follows:

- Ensuring corporate standards and branding are maintained throughout the web site.
- Maintaining accurate and timely meta-data for content.
- Approving, rejecting or amending content to be published on the web site.
- Communicating corporate standards to content coordinators and authors.
- Reporting to management about quality issues of the web site.

2.1.4.6 Business Quality/Website Manager

Business quality managers are responsible for establishing and monitoring web publishing processes and strategy. Verblya [105] summarises web site management as "*ensuring that the Web-site satisfies as far as possible the requirements of all stake-holders*" [105]. Their responsibilities include:

- Process improvement and monitoring standards of content, for example, the elimination of bottlenecks or consistently poor content that breaches corporate standards.
- Regular review of web content to ensure that it is timely and covers the needs of the target audience of the web site.
- Liaise with other management to develop web strategies to cope with external influences. For example, accessibility legislation may require extra training or investment in new software tools.
- Organise internal training and software tools for staff.
- Liaise with technical staff to "to escalate issues" [50] to management. For example, if technical staff felt that a web server was underperforming, this issue would be raised with management and if necessary the web server may be upgraded or supplemented.
- Monitoring of usage statistics and user feedback to help identify gaps in the content provision.

2.1.5 Web Maintenance

Web publishing can be categorised as a type of web maintenance. Ensuring that content is updated and accurate is an act of maintenance. However, for web sites which focus on supporting service delivery rather than content delivery (which Pressman refers to as "WebApps" [84]), the maintenance activities are the same types of maintenance as traditional software [107]. In traditional software engineering maintenance changes are classified as corrective, perfective, adaptive or preventative changes. These classes of maintenance are now discussed, followed by a discussion of differences between the development models for traditional software and web applications [84] [70] and how these differences impact on web maintenance.

2.1.5.1 Corrective Maintenance

Swanson [101] states that corrective maintenance is needed in response to a failure in the software's design. Corrective maintenance is the removal of errors which were included in the web application by the developers.

2.1.5.2 Perfective Maintenance

Perfective maintenance is described as "keeping a program up and running at less expense, or up and running so as to better serve the needs of its users" [101]. For the web, this might be removing large and unnecessary bitmap images from web pages, hence reducing the amount of data that must be downloaded. It could also be more system related, for example, switching to a later, more efficient version of a database server.

2.1.5.3 Adaptive Maintenance

Changes in the environment in which the web application operates will require adaptive maintenance. Swanson gives the example of an "installation of a new generation of system hardware" [101], however, in modern web applications this is more likely to be changes in the technologies [71] used by the hardware, such as an upgrade in web server software.

2.1.5.4 Preventative Maintenance

Preventative maintenance is defined by the IEEE as "maintenance performed for the purpose of preventing problems before they occur" [58]. User interface adaptations such as ensuring web accessibility (see section 2.2) could fall under this category.

2.1.6 Differences between web and traditional software development

Although the development and hence maintenance of web applications have similarities with traditional software applications they are not entirely comparable. Table 2-2 contains differences experienced during web development as highlighted by Lowe [71] and Pressman [84] and how these differences affect maintenance.

Difference	Type	Effect
Closer ties to business model	Technical	Maintenance tasks are tied more closely to changes in business models and requirements.
Number of and rapid change in technologies used	Technical	Maintainers must be familiar with a large variety and versions of

		technologies and be aware of interoperability issues. The rapid changes in technology and versions means than preventative maintenance is less likely to be given as much importance as the other 3 types.
Importance of timely and accurate content	Technical	Web applications must fit in with sophisticated content delivery mechanisms. Content updating is one of Lowe's [71] examples of a finer grained maintenance for web applications.
Wider and more demanding user base	Technical	Web site user interfaces must be intuitive, useable and accessible. Hence, perfective maintenance will be an extremely important part of web maintenance.
Shorter developer timeframes	Organisational	Shorter development and maintenance times may result in an increase in corrective maintenance, because the time to properly design and test the web application was not available.
Client uncertainty	Organisational	Maintainers must cope with more drastic adaptive or perfective maintenance tasks as the client's requirements change.

Table 2-2 Differences found in web development and their effect on maintenance

2.2 *Web Accessibility*

2.2.1 Definition

In their introduction to web accessibility, the Worldwide Web Consortium (W3C) in their Web site Accessibility Initiative (WAI) [31] give the following definition, "Web accessibility means that people with disabilities can use the web". There are of course many forms of disability, some less obvious than others, and web accessibility aims to allow individuals to use the web effectively regardless of their physical or mental disability. Stephanidis [100] asserts that "accessibility implies the global requirement for access to information by individuals with different abilities, requirements and preferences...". Web accessibility can be likened to physical accessibility. A building whose entrance is at the top of a flight of steps is inaccessible to those who cannot easily (if at all) climb steps. The absence of a ramp is an **accessibility barrier**. Accessibility barriers also occur in web sites. For example, a web site that relies entirely on graphics for navigation has a serious accessibility barrier to users who cannot easily (if at all) see those graphics.

2.2.2 General Overview

In 2005 91% of small to medium enterprises and 99% of large enterprises within the EU had internet connections [81]. Alongside this, nearly half (48%) of all EU households had access to the internet. The number and range of online services has also increased considerably; banking, education, grocery shopping, and local government administration all now have online presences. Those with disabilities have most to gain from the growth in online services. As seen in the online banking example in the introduction, users with age related disabilities, or any other disability should be able to take advantages of these new services to make their lives a little easier. But to do so, they may need help from assistive technologies. And if a web site is not flexible enough to work well with such technologies it is inaccessible to these users. Despite the efforts of the W3C in promoting and creating guidelines [26] for Web Accessibility and various governments passing legislation (UK SENDA [51] and in the USA Section 508 [52]) requiring Web Accessibility, a large percentage of web sites remain inaccessible [44]. Lazar believes the problem is that web maintainers do not value web accessibility as important [68], and attributes this partly to a lack of

education noting that "Accessibility ... is not a standard part of any national curriculum in Computer Science...or Information Technology" [68].

2.2.3 Principles of web accessibility

The principles of accessible design are defined in version 2 of the W3C's web content accessibility guidelines [36]. These are as follows:

1. Content must be perceivable.
2. Interface components in the content must be operable.
3. Content and controls must be understandable.
4. Content should be robust enough to work with current and future user agents (including assistive technologies).

2.2.3.1 Perceivable content

If users cannot perceive that content is there, it is completely inaccessible. It is therefore important that web pages are designed to ensure that the perception of content does not rely on a single specific ability. For example:

Users with colour blindness – cannot perceive certain colours. Therefore, web pages relying on colour alone to encode information contain accessibility barriers. For example, if a university publishes a list of lectures on its web pages and states: "All lectures highlighted in red are compulsory". If this is the only method of encoding the information, then users who cannot perceive red are at a distinct disadvantage. In this example other stylistic effects, such as text underlining or bordering should be used as well.

Users within the Deaf community – cannot perceive aural content and so all multi-media content with sound should have either captioning (for video) or transcripts containing equivalent information should be provided.

2.2.3.2 Operable interface

Not all users view web sites using standard computing hardware (i.e. a mouse, keyboard and monitor) and standard visual browsers (i.e. Mozilla, Firefox or Internet Explorer). Users of assistive technologies may be required to interact with the web site in a restricted manner (see section 2.2.4). For example, a web page should not

rely on the use of a mouse and should provide alternative mechanisms to "to help users find content, orient themselves within it, and navigate" [36].

2.2.3.3 Understandable content and controls

Web content which is overly complicated and unclear is difficult to understand. This is especially true for novice web users and users who may have difficulties in understanding the language used (for example, children, non-native speakers and users with cognitive disabilities). Donkin argues that plain English is required improve web accessibility [42]. This should be generalised to plain language as plain English will be incomprehensible to users who speak no English; however, it will surely be an improvement to users who have at least some knowledge of the language. Bolchini [12] states that users interacting aurally with web sites have different control requirements that those using visual interaction.

2.2.3.4 Robust content

Web content should be robust enough to cope with current and future versions of browsing software including assistive technologies. Content must have accessible alternatives if the original requires specific senses or bespoke software. For example, if a video clip is encoded to function only in specific software then a transcript of its contents would meet both these criteria.

2.2.3.5 Separation of presentation and structure

Ensuring compatibility with these principles can be partly achieved by following the W3C's recommendations for authoring HTML documents [25]. It specifies the need for the separation of presentation and structure of a HTML document. This facilitates device independence by allowing the web browser to use its own style sheet to determine how content should be presented based on its semantics. Centralised style sheets can be used to control the default presentation of multiple web pages.

2.2.4 Main beneficiaries of accessible web sites

As well as users with disabilities there are other groups of users who benefit from good accessibility. The groups of users who have the most to gain from accessible web content fall into the following categories:

- Users with disabilities may require assistive technologies to view web content.
- Users with low speed internet access (for example, users in developing nations) may use text-only browsers or at least restrict the display of images to lower the amount of bandwidth used when browsing.
- Users of mobile devices to browse the web such as mobile phones or PDA devices.

One of the challenges in creating accessible web sites is creating content which is accessible to users with diverse sets of needs, some of which might conflict. For example, users with visual impairments may require content to be delivered aurally, whereas users with hearing impairments would find this method of content delivery inaccessible.

There now follows a discussion of various impairments, the assistive technologies available to overcome them and the issues which must be addressed if content is to be accessible to these technologies and their users. The issues discussed come from four sources. They are as follows:

- W3C Web Content Accessibility Guidelines (versions 1 [26] and 2 [36])
- Web Accessibility In Mind (WebAIM) articles [110].
- Royal National Institute for the Blind (RNIB) Web Access Centre [88].

2.2.4.1 Visual Impairments

There are several classification models of visual impairment but it is beyond the scope of this work to discuss these in detail. The types of visual impairment now discussed were taken from the World Health Organisation's (WHO) International Statistical Classification of Diseases and Health Related Problems [80] and with additional references from the WebAIM Visual Disabilities introduction [110]. The three types of visual impairments discussed here are:

- Blindness.
- Low Vision.
- Colour Blindness.

For each type of impairment any assistive technologies available are described, followed by a discussion of the implications for web page design.

Blindness

Addressing the issues associated with blindness has been a main focus of web accessibility research. Blind users have an obvious disadvantage in that they cannot see what is on the screen. As such the two main assistive technology streams have adapted existing technologies: Braille and vocalisation of text.

Assistive technology	Description
Electronic Braille displays	Electronic Braille displays are tactile devices which dynamically refresh allowing the user to read what is displayed on the computer screen.
Screen reading software or aural browsers	Both aural browsers and screen readers read out to the user what appears on the screen. Aural browsers are explicitly designed for web pages navigation.

Table 2-3 Assistive technologies to compensate for blindness

Table 2-3 shows the assistive technologies available to users with blindness wishing to browse web pages. They present information to users in a linear form with no awareness of layout and as such, require web pages to be created in a way that will "linearise" sensibly. Issues which must be considered when designing accessible web pages are:

Images - Information contained within images is not accessible, and hence require a textual equivalent. However, purely decorative images require blank textual equivalents rather than pointless descriptions, such as "a picture of a leaf falling from a tree" or "the company's logo: a cow standing in a field". These contain no informational value and slow screen readers down unnecessarily.

Navigation – Providing consistent navigation throughout the web site is important to ensure that users can easily move around the web site. Users

should not have to search for navigation links on the page; rather they should be grouped together in a logical order. However, an unwelcome side affect caused by the restriction that all content is presented linearly is that aural browsers are forced to present the same group navigation hyperlinks again and again. In this case, before the user can access content specific to the page they are presented with all navigation links which both frustrates and wastes time. The solution is to provide "skip navigation links" (recognised by the Section 508 checklist see section 2.3.1), these are links to internal anchors within the page that when followed take the browser to the main content of the web page.

Layout – Users of either Braille displays or aural browsers do not have the visual layout of the web page available to them. Hence, it is important that the page linearises properly. A common technique to organise visual layout employed is to use the `<table>`, this is not recommended by the W3C as the purpose of the tag is to mark-up data tables. The problem is that by using tables to devise visual layout, there is a risk that it will not linearise properly for screen readers or text only browsers (see Figure 2-3, Figure 2-4, Figure 2-5 and Figure 2-6)

Choose a module from either Software Engineering or Artificial Intelligence	
Software Engineering	**Artificial Intelligence**
Software Design	AI Languages
Configuration Management	Planning

Figure 2-3 Example of a table used for layout

Figure 2-3 shows a table used to layout menu options which will not linearise properly.

Choose a module from either Software Engineering or Artificial Intelligence

Software Engineering

Artificial Intelligence

Software Design

AI languages

Configuration Management

Planning

Figure 2-4 Result of linearization from Figure 2-3

Figure 2-4 shows how the table is displayed using either an aural browser or text-based browser. The course headings are read out or listed first and then each module is read out or listed. It is not possible from using only the linearised version of this table to discern to which course a module belongs.

Choose a module from either Software Engineering or Artificial Intelligence		
Software Engineering	Software Design	Configuration Management
Artificial Intelligence	AI languages	Planning

Figure 2-5 Table used to layout a menu which will linearise sensibly

Figure 2-5 shows the more sensible table structure for layout. As is shown in Figure 2-6, when the table is linearised the result is much more understandable.

Choose a module from either Software Engineering or Artificial Intelligence

Software Engineering

Software Design

Configuration Management

Artificial Intelligence

AI languages

Planning

Figure 2-6 Result of linearisation from Figure 2-5

Figure 2-6 shows how the menu will be linearised more sensibly in an aural or text-based browser. The headings of the courses are read out before each group or modules, which allows the user to determine to which courses each module belongs.

Web pages specifically designed for blind users

Research has also been carried out to develop alternatives to standard HTML based web pages, both in developing specific requirements for the production of aural web pages [12] and in developing transformation tools to generate aural web pages [73]. These recognise the differences between the needs of users of aural and users of visual browsers and that the linear sequencing restriction imposed on aural browsers "completely changes the interaction paradigm" [12].

Low Vision

Low vision users have a partial view of a web page and hence are not restricted to accessing information in a strictly linear form. There are assistive technologies and design techniques aimed at enabling web browsing for those with low vision. They are as follows in Table 2-4.

Assistive technology	Description
Screen Magnifiers	Areas of the screen are magnified using software, simulating a magnification lens.
Customisable presentation	Users with low vision may wish to change the size of text and also background colour and text colours to make pages easier to read.

Table 2-4 Assistive technologies to assist users with low vision

Table 2-4 shows the assistive technologies available to users with low vision. Issues which must be considered when creating accessible web pages are:

Layout – as the magnification increases, the amount of information which can be displayed on the screen decreases.

Images and Text – text contained within images does not enlarge as plain text does when the user increases the font size. Image quality degrades with magnification appearing "pixelated" and unclear. As such text contained within images should be avoided where possible. Some users may wish to override the background and font colours used to suit their own visual needs.

Therefore, web page presentation should be controlled by a style sheet rather than presentational HTML which can be overridden by the user's own preferences.

Colour Blindness

There are no mainstream tools to aid users with colour blindness. There are several issues that must be addressed in order to ensure web content is accessible. Colour blindness only becomes problematic when colour is used to convey information. Examples include the following:

- Public transport route maps where different colours are used to represent routes.
- Highlighting keywords in a sentence using colour alone.

These design considerations are also useful for users with no colour blindness, as they may be viewing the web page for the first time and unaware of any colouring conventions being used.

2.2.4.2 Hearing Impairments

The Royal National Institute for Deaf People (RNID) [89] describe the ranges of individuals with hearing impairments as:

- **Hard of hearing people** – people with mild to severe hearing loss which has developed gradually during their lifetime.
- **Deafened** – people with profound hearing loss which has occurred after they have learnt to speak.
- **Deaf community** – people belonging to *the deaf community* have sign language as their first language.

For the first two categories of users, those who have no inherent difficulty in understanding spoken language in written form, the only issue that must be considered is to ensure any audio content is provided with a corresponding textual equivalent. For example, if the chancellor of a university delivers a speech and a video is made available on the web site, a transcript of the speech should also be published.

An important difference between those belonging to the deaf community and those who have suffered some form of hearing loss after learning to speak is that the natural spoken language of their country may not be their first language. Hence, these signers approach the written word as a foreign language. Donkin [42] suggests that written language should be plain and simple, avoiding slang and jargon, benefiting both human understanding and also facilitating the possibility of converting the text to sign language. Some specialist organisations provide signing video clips to provide extra information for the deaf community; however, this is a time consuming and expensive activity. To overcome the shortage of available signers and the cost and time delays in producing the video clips there has been some research into producing "virtual signers" [8]. These are computer generated video clips of sign language using an XML based Gestured Modelling Language (GML).

2.2.4.3 Cogitative Disabilities

There has been relatively little research into web accessibility for people with cognitive disabilities [98]. Cognitive disabilities cover a broad range of disabilities which affect an individual's mental processes. They affect the perception, memory, judgment, and reasoning of individuals and hence can interfere with their understanding of web content. Some of the main difficulties users with cognitive disabilities face are listed by the WebAIM [111] organisation:

1. **Memory** – short term memory issues may result in users not remembering how they reached a web page and so becoming lost in the navigation. Users may also forget what they have just read previously on the web page.
2. **Problem-solving** – errors such as broken links or complicated instructions could confuse and frustrate users.
3. **Attention** – users may be unable to focus their attention for extended periods of time or become distracted easily.
4. **Reading, linguistic, and verbal comprehension** – users will have varying levels of comprehension and may become confused if the language is too complex or if the presentation does not aid comprehension.

Addressing these issues requires extra attention to how the information contained on the web site is structured and presented to the user. Although it is impossible to cater for every possible cognitive disability there is a valid case for ensuring that

unnecessary barriers are removed from web pages. Hudson argues that "if someone who could understand web content is unable to, because of the design choices of the web author, then that web content is inaccessible" [57]. This argument corresponds with Donkin's [42] argument for the use of plain natural language usage within web content. WebAIM [111] also state that using plain and simple language cannot guarantee understanding but it can at least increase the chances of content being understood. There are, as yet, no specific assistive technologies available for users with cognitive disabilities, there are however a number of design issues which must be considered if content to support the creation of accessibility web sites.

Navigational Support

Providing accessible navigational support is difficult because of the memory and problem-solving deficits experience of users. Neerincx identified four "fundamental cognitive determinants of navigational performance" [78]. They are as follows:

- **Situation Awareness** – recognition that the user is looking at a web page and that they are able to navigate around the current page and to other pages. Neerincx [78] recommends the provision for supplementary and multi-media information to enhance the perception and comprehension of users' goals when using the website.
- **Spatial Ability** – the user's ability to navigate the web site using the navigational features of the web site. Neerincx's recommendation for "spatial representations" [78] coincides with the WebAIM's recommendations for providing more detailed hyperlink texts [111]. For example, rather than having links using the text "Previous Page", more information should be provided to help the user recall where the previous page link will take them.
- **Task-set switching** – the user's ability to move between specific tasks (for example, from activating a "contact us" hyperlink to filling out a web based form). Neerincx [78] calls for the introduction of a task scheduler to assist users to focus on new tasks. Many e-commerce web sites have already recognised the need for such staged additional navigation aids to facilitate online purchasing. Rossi [90] defines 2 specific design patterns which could support task-set switching for the general population:
    - ***Advising*** – created to "help the user find a product in the store" [90].

- ***Explicit process*** – created to "help the user understand the buying process when it is not atomic" [90]

- **Capacity and preference differences** – recognising that different users will have individual preferences and support needs.

2.2.4.4 Low-bandwidth Input

To clarify, low bandwidth users discussed in this section are those users "with such limited motor and speech that they can only produce one or two signals when communicating with a computer." [74]. It does not refer to users who have low-bandwidth / low-speed internet connections. Because of the limited number of signals low-bandwidth users are capable of producing, they take longer to complete tasks and, therefore, for content to be accessible users must be able to achieve their goals with as few input signals as possible. Mankoff stresses not enough work has been carried out to develop solutions for this user group and has produced 7 new low-bandwidth accessibility requirements [74]:

1. **The currently selected link is visible** – this can be achieved with a breadcrumb trail showing users where they are in the web site hierarchy.
2. **The user can read and navigate text that contains no links** – extra navigational links at the end of paragraphs should be provided.
3. **The user can traverse the history list** – links which access the user's browser links can be added to a web page.
4. **The user can access bookmarks and add to them** – an extra link at the top of each page allowing the user to bookmark it should be added.
5. **The user can quickly access text and links of interest** - providing skip navigation links can allow low-bandwidth users to avoid repetitive and time consuming navigation.
6. **The user is given alternatives for form elements** – for example provide a soft keyboard in the screen.
7. **The user is given information about link targets** – this is achievable through the use of the "title" element of anchor hyperlink elements. By providing a brief description users can decide whether or not the page is worth visiting.

2.2.5 Importance of Web Accessibility

The importance of web accessibility is directly linked to the importance of the web. As the web becomes further integrated into everyday life, the more vital it is that this important resource is accessible to everyone [13]. Users with physical disabilities should be able to take advantage of the remote access to services offered by the web. Potential physical barriers are removed by offering users the chance to access services such as shopping, bank and education via the web. But whereas access barriers in the geographic sense have been addressed (a user in the USA can email a user in the UK as easily as someone living in the next house) access barriers may still exist for users with non-standard browsers or disabilities. In the physical world, this would be the equivalent of a wheelchair user travelling to an art gallery using accessible public transport, but because the gallery was on the 4th floor of a building and the lift was out of order, they are unable to access the exhibition.

One formal investigation [21] accessed 1000 home pages and found that 81% "failed to satisfy the most basic WAI category. In addition... they have characteristics that make it very difficult, if not impossible, for people with certain impairments... to make use of the services provided" [21]. Governments have for some time recognised the need to ensure web accessibility. In 1999 then United States Secretary of Commerce William M. Daley wrote in the introductory letter to the US Department for Commerce report that it is vital that all US citizens "have the information tools and skills that are critical to their participation. Access to such tools is an important step to ensure that our economy grows strongly and that in the future no one is left behind." [38]. Similar sentiments were expressed in an EU report concerning the accessibility of public sector web sites. It stated that if accessibility is built into public sector web sites "individual opportunities in education, employment, health and social life are enhanced and this, in turn, has the potential to bring about a significant economic impact in Europe [44]".

WebAIM [109] offers 3 main motivations for producing accessible web content:

"To improve the lives of people with disabilities

To capitalize on the a wider audience or consumer base

To avoid lawsuits and/or bad press" [109].

These follow closely with the four main factors cited by the W3C of which social and financial factors are two main factors for justifying a business case for organisations to ensure web accessibility [32]. These are now described in detail.

2.2.5.1 Social Factors

Social factors include ethical issues such as not discriminating against those with disabilities, and recognising that organisations have a moral responsibility to ensure this includes web content. It is also necessary to be aware that not just those with disabilities are affected but also people with low literacy and age related conditions [64]. Richards [87] argues that until the benefits of accessibility to a wider audience are recognised, not just those with disabilities, large scale adoption of accessibility will not occur.

2.2.5.2 Technical Factors

By developing web sites that are accessible from the start, this saves time later in the web maintenance to comply with legislation or corporate policy. Several techniques required for accessible design such as clear and consistent design result in fewer pages and thus reduce the load on the server. Further techniques, such as the separation of content and presentation make changing how web sites look much easier. Providing web sites that have text based alternatives to graphics and multimedia content enables users with low bandwidth not to load images and hence reduce download time. Kirkpatrick [64] lists the following technical benefits gained by ensuring accessibility:

- **Reduced Site Development and Maintenance Time** – The benefit of using CSS style sheets (see section 2.2.3) for the organisation is that if they wish to change the style of their web sites they need only make changes to the style sheet rather than repeated changes to every page on the site. If organisations used style sheets to format headings instead of creating bitmap graphics the amount of workload to localise or internationalise their web site is also greatly reduced. Textual information contained within bitmap images is not machine readable. However, if no textual alternative is provided, it is not only non-visual browsers that are affected. Indexing robots used by search engines to index web pages and rank them will not be able to access the information.

Hence the web page's ranking in a search engine will be affected negatively if the search engine's robot can not access heading information.

- **Reduced Server Load** – By using centralised style sheets to control presentation the amount of HTML needed to create pages is reduced and for large web sites this can amount to a considerable reduction in server load.
- **Improved Interoperability** – Users have many more browser agents at their disposal now than when the web was first introduced in the late 1990s. The increasing usage of mobile agents such as mobile phones and PDAs means that web sites must be designed to work on a more diverse range of platforms.
- **Prepared for Advanced Technologies** – by conforming to standards devised by international bodies such as the W3C, organisations can better prepare themselves for technologies built on top of existing standards.

2.2.5.3 Financial Factors

Increasing the accessibility of a web site has several side effects which can increase its usage and hence increase either direct or indirect revenue. Accessible web sites are well organised and use HTML appropriately. This has a positive impact on how search engine robots index and then rank a site. A better ranking means that a web site will be "found" by more potential customers. Accessibility can be considered as a sub-branch of usability [13] [69] [87] and an improvement in accessibility results in a more usable web site. Improvements in usability and accessibility also benefit older users who may suffer from age related conditions [10]. Internet usage amongst older users is growing. In 2003 Nielsen//Netratings reported the percentage of US citizens over the age of 65 using the Internet at work and home increased by 25%. In 2005 in the European Union 11.8% of 55-64 year olds and 4% of users aged 65 and over used the Internet to obtain information from public authority web sites [86]. In the commercial sector within the EU 10% of 55-74 year olds used the Internet for the purpose of travel and accommodation [65]. In Sweden where Internet usage is very high, 64% of men and 48% of women between the ages of 55-74 had used the Internet [102]. As the population ages and as Internet usage grows, the number of users with age related conditions (and hence accessibility needs) will increase [116]. Older users may have a significant amount of disposable income. However, if an organisation's web site contains accessibility barriers (causing older users difficulty) they are less likely to complete the transaction and will visit other sites or revert to using more

traditional methods of accessing services. For example, if a supermarket's online store is accessible and usable it will be easier for older users to purchase more items as they do not have to physically carry the items. Hence, they will be inclined to buy more and spend more money at the supermarket.

2.2.5.4 Legal and Policy Factors

Legal and policy factors are very strong driving forces for organisations [7]. By implementing accessible web sites and installing best practice now, organisations are ensuring against changes in policy or the law which may mean making costly redevelopments in the future. Governments have started to recognise the importance of the web as a source of commerce, communication and information. Consequently since the late 1990s many countries have introduced legislation governing the accessibility of web sites providing public information. Table 2-5 summarises some of the countries which have introduced legislation or guidelines concerning web accessibility.

Country	Date	Description
Australia	August 2002	The Australian Government have produced advisory notes for the Disability Discrimination Act [22]. Web pages and services are considered a service by the DDA. Maintainers must provide equal access to people with disabilities as a legal requirement "where it can reasonably be provided" [22]
Israel	August 2005	In 2005 the Israeli Government amended the Equal Rights for Persons with Disabilities Law [41] to ensure that all public buildings and services are accessible. Public organisations must make accommodation unless they cause "undue burden".
Italy	January 2004	Public agencies are required in the case of similar bids for contracts to select the bid that conforms best to accessibility guidelines. Public agencies must also ensure that any modification or creation of web pages conform to accessibility guidelines. [3]

Portugal	August 1999	Portuguese law requires all information published on public administration web sites be accessible. Specifically that: a) "Reading can be performed without resorting to sight, precision movements, simultaneous actions or pointing devices, namely mice. b) Information retrieval and searching can be performed via auditory, visual or tactile interfaces." [53]
Spain	July 2002	From 31st December 2005 public administrations must have adopted the necessary means to ensure that information on their web sites is accessible to old people and people with disabilities. They must also ensure that web pages of organisations they support financially are also accessible. [19]
United Kingdom	1995	Part III of the Disability Discrimination Act [51] prohibits any organisation from discriminating against people with disabilities by failing to provide equal access to "Goods, facilities and services". In Code of Practice associated with part III of the act the following example is given, "An airline company provides a flight reservation and booking service to the public on its website. This is a provision of a service and is subject to the Act." [20]. The Act states that organisations must make "reasonable adjustments" to ensure equal access.
Unite States of America	1998	Section 508 [52] is an amendment to the Rehabilitation Act which requires federal departments and agencies to ensure their electronic information technology is accessible. Exceptions occur if undue burden would be imposed on the department or agency in making the necessary changes.

Table 2-5 Summary of countries national accessibility legislations

2.2.6 Implementing web accessibility (Best Practice)

Section 2.2.5 lists why accessibility is important and why organisations should be motivated to implement high levels of accessibility for their web sites. However, there are obviously costs associated with its implementation. Richards [87] points out that for web sites with large numbers of legacy web pages the costs for even minor changes quickly escalates. The W3C cites the following steps in establishing best practice in its implementation plan for web accessibility [29].

- **Establish Responsibilities** – identify key personnel who have responsibility for accessibility, such as a champion or guru who will proactively encourage and support others within the organisation.
- **Conduct Initial Assessment** – determine the current level or accessibility (See section 2.3).
- **Develop Organisational Policy** – an organisational policy is required because the web site is a representation of the organisation and minimum levels of accessibility should be set.
- **Select Software** – accessibility should be considered when procuring software. If software is used that produces inaccessible web content then it will cost the organisation in lost time to remove the barriers later.
- **Provide Training** – the people who are responsible for creating and maintaining web content must have the right level of training to create accessible content.
- **Develop Accessible Web Site** – the web site must be developed with respect to the organisational policy using the tools and training developed.
- **Promote Organisational Awareness** – without awareness of the issues surrounding accessibility the likelihood of a web site being accessible is very small.
- **Monitor Web Site Accessibility** – once the web site's accessibility has been improved, it must remain consistently accessible, otherwise after time the accessibility could degrade leading to an inaccessible web site.

2.3 *Web accessibility assessment*

The first step to improving web accessibility is the ability to assess the current level of web accessibility. Assessment requires criteria of what is and what is not considered as an accessibility barrier.

2.3.1 Accessibility guidelines and standards

Content authors and web developers require guidelines and standards if they are to recognise potential accessibility barriers and create accessible web content. Guidelines provide the checkpoints and techniques needed to create accessible web content. Standards define the minimum requirements for web content to be considered accessible. Many organisations specify their own standards for web content but rather than creating their own accessibility guidelines from scratch, augment or refer to one of the guidelines or standards now described.

2.3.1.1 W3C Web Content Accessibility Guidelines

The W3C published version 1 of their Web Content Accessibility Guidelines [26] in 1999. These are internationally recognised guidelines and are used by the international web community. There are 14 general guidelines each addressing a different aspect of web content design. They are listed in Figure 2-1.

1. Provide equivalent alternatives to auditory and visual content.
2. Don't rely on colour alone.
3. Use mark-up and style sheets and do so properly.
4. Clarify natural language usage.
5. Create tables that transform gracefully.
6. Ensure that pages featuring new technologies transform gracefully.
7. Ensure user control of time-sensitive content changes.
8. Ensure direct accessibility of embedded user interfaces.
9. Design for device-independence.
10. Use interim solutions.
11. Use W3C technologies and guidelines.
12. Provide context and orientation information.
13. Provide clear navigation mechanisms.
14. Ensure that documents are clear and simple.

Figure 2-7 W3C Web Content Accessibility Guidelines 1.0

Within each guideline shown in Figure 2-7 are prioritised checkpoints. There are 3 levels of priority.

Priority	Description
1	A Web content developer **must** satisfy this checkpoint.
2	A Web content developer **should** satisfy this checkpoint.
3	A Web content developer **may** address this checkpoint.

Table 2-6 Checkpoint Priorities [26]

Table 2-6 contains the priority rankings and importance for the checkpoints. As is implied, priority 1 checkpoints are considered the most important to implement, followed by the priority 2 and 3 checkpoints. For example, the European Union (EU) recommends that all EU public sector web sites meet all priority 2 checkpoints [44]. Because of the historical importance and international recognition of the W3C and their work in creating standards and guidelines for the web, many organisations (including governments) have based their web accessibility guidelines on the W3C recommendations.

At the time of writing this thesis, the W3C are updating their guidelines and have drafted version 2 [36] of the web content accessibility guidelines. This is as yet a draft publication and has been under development during the course of this thesis. In contradiction to guideline 11 of version 1 (see Figure 2-7); the version 2 guidelines do not require the use of W3C technologies. Its focus has shifted from primarily HTML based web content to recognise that "the web is used in hundreds of ways that were not possible in 1999" [31]. As such it has developed the concept of **baselines.**

Technology Baselines

Baselines are groups of technologies that the content authors can assume are available and running in an accessible user agent. As such, they must ensure that content conforms to all the guidelines even if *only* the baseline technologies are available. These baselines then can be viewed as minimum specification for user agents. They can be set by various types of organisations; "government body, client, organization, author, or combination of these" [31]. For example, a government might define a

baseline specifying that all public sector organisation web sites must conform to the version 2 accessibility guidelines assuming on browsing software supporting HTML 4.01 Transitional, JPEG and GIF images.

2.3.1.2 Section 508 Standards

The legal aspect of America's Section 508 legislation is discussed in section 2.2.5. The legislation itself contains a prescriptive checklist that must be met by web sites for organisations wishing to comply with the Section 508 legislation. The checklist overlaps the W3C guidelines covering:

- Providing textual equivalents for non-textual content.
- Ensuring the web page can be read using assistive technologies.
- Documents should be readable without a style sheet.
- Providing skip navigation links to allow the user to bypass navigation links to the main content of the page.
- If a timed response is required from the user, this should be clear and methods to increase this limit should be made available.
- Pages shall be designed to avoid causing the screen to flicker with a frequency greater than 2 Hz and lower than 55 Hz.
- Correct mark-up should be used for tables. For example, table heading should be marked-up explicitly as headings and to which column or row they relate.

2.3.2 Assessment tools

Currently, there are no tools which are able to fully assess web accessibility without human judgement. Lazar [66] mentions that during his study "human intervention is therefore required on some guidelines to determine if there is in fact an accessibility problem". Assessing whether a web page meets the guidelines specified by the W3C or conforms to all the checkpoints of the Section 508 checklist requires human judgement. Many of the W3C guidelines (Figure 2-7) are entirely subjective; for example, guideline 11, "Ensure that documents are clear and simple" cannot at present be checked by automated methods. Also what is clear and simple to one person, may not be considered clear and simple to another. As such, some form of manual assessment is both time consuming and necessary. There are, however, tools which provide support for semi-automated assessment and can help alert users to

potential accessibility barriers. These tools will now be described in the following subsections:

2.3.2.1 Watchfire (formerly Bobby) [108]

Watchfire provides a multifunctional web page analysis as both a free but limited use online service or as a standalone tool. The tool was originally called Bobby and only assessed web accessibility based on version 1.0 of the W3C Guidelines. Results from the online version are presented as a web page report, highlighting warnings and "errors" (barriers) detected on the web page. Warnings are potential barriers found on the page but cannot be confirmed by automated techniques, whereas errors are accessibility barriers which the tool is able to confirm as an accessibility barrier.

2.3.2.2 AccessValet (also known as PageValet) [112]

PageValet is a stand alone accessibility assessment tool which assesses a web page against either version 1.0 of the W3C Guidelines or Section 508. As PageValet was used as part of the experimental work for this thesis, further details can be found in section 3.2.3.1.

2.3.2.3 W3C Markup Validation Service [34]

The W3C offer a free online service to validate HTML, XHTML against the standards defined. They can ensure, for example, that XHTML documents conform to the correct version of the DTD (Document Type Definition).

2.3.2.4 A-Prompt [2]

A-Prompt is a semi-automated accessibility assessment and repair tool. It assesses a web page for accessibility barriers and then uses a wizard style support tool to assist web maintainers in the removal of each barrier detected.

2.3.3 Quantitative measurement of web accessibility

Manually assessing the accessibility of a large web site is not feasible. The manual process takes too long and is hence very expensive. Research has therefore focused on how automated and semi automated methods can be used to minimise assessment time and maximise accuracy. There is also a need for some method to objectively classify and compare web pages, in order to determine which is less accessible than

others. If this is achievable then prioritisation of effort becomes possible, and the web pages with the most serious accessibility barriers can be addressed first. Quantitative measures provide the means to monitor changes in accessibility and provide empirical data for maintenance decisions. However, since many aspects of web accessibility are subjective, quantitative measures are only part of the solution and are only as reliable as the assessment tools which produce them, to be totally confident of a classification some human judgement is required. There have been several attempts to measure accessibility, Lazar [66] has created a classification scheme which is applied after both automated and manual checks have been performed. Zeng's WAB score [117] is entirely automated and hence uses a subset of automatically verifiable checkpoints, as is Arrue's metrics [5]. These are described in greater detail below.

2.3.3.1 Lazar categories

Table 2-7 shows Lazar's 4 categories of web page accessibility [66] and how this categorisation is achieved. The rules (accessibility checkpoints) are an amalgamation of W3C priority 1 checkpoints and Section 508 checkpoints. The number of instances of a rule violation is not taken into account, Lazar explains this decision "was based on vigorous debate by the researchers" [66]. The justification is that removing several instances of the same accessibility barrier requires less effort that removing several distinct types of barriers. Manual checks were performed to ensure that rule violations detected by the two semi-automated assessment tools were actually breaches of the checkpoints.

Category	Accessibility Level
0 rules violated	Accessible
1 - 3 rules violated	Marginally inaccessible
4 - 6 rules violated	Moderately inaccessible
7+ rules violated	Substantially inaccessible

Table 2-7 Lazar's web page accessibility categories [66]

2.3.3.2 Zeng's Web Accessibility Barrier (WAB) score [117]

Zeng's WAB score [117] differs from Lazar's approach in four ways. Firstly, Zeng includes the number of times (instances) a barrier has occurred (a rule has been violated), only the W3C checkpoints are considered but instead of restricting the

assessment to priority 1 checkpoints, all 3 priorities are included. Finally, only those checkpoints which do not require human interpretation are included. Each barrier's importance is determined by the adding a weighting to it (the inverse of the checkpoint's priority level). Zeng [117] highlights that "The WAB score is a proxy of web accessibility". To establish the reliability of the WAB score, Zeng applied a gold standard trial which looked for a correlation between the score and the W3C accessibility ratings web site's have awarded themselves (for example, that a web site complied with all the W3C priority 1 checkpoints). As Zeng points out, this method has threats to validity; specifically the trial depended on the self assessment of web site compliance with W3C guidelines. A web developer's assessment of their own web site's accessibility is biased and is therefore must be considered as a serious threat to validity. A further threat is that even if the assessment was accurate once, there was no check to see how recently the assessment was carried out.

However, the principles behind the WAB score are sound. Only barriers which are automatically verifiable and both the severity of the barriers and the size of the whole web page are used to give an indication of accessibility.

$$WABScore = \frac{\sum_{p}\sum_{v} (\frac{n_v}{N_v})(w_v)}{Np}$$

p: Total pages of a website.
v: Total violations of a Web page.
n_v: Number of true violations
N_v: Number of potential violations
Wv:Weight of violations in inverse proportion to WCAG priority level.
Np: Total number of pages checked.

Figure 2-8 Zeng's WAB score formula [117]

2.3.3.3 Arrue's Quantitative Metrics [5]

Arrue's metrics [5] use the principles set out in version 2.0 the W3C's accessibility guidelines [36]. It maps accessibility checkpoints from version 1.0 [26] to the principles in version 2.0 [36] to which they relate (i.e. Perceivable, Operable,

Understandable and Robust)(see section 2.2.3). By classifying each checkpoint in this way, each principle is given an accessibility score based on the number and priority of the checkpoints it covers. For example, if a principle contains more high priority checkpoints, its associated accessibility score will be greater. Once each principle's score is worked out, it is used to calculate the "global" accessibility score of a web page. This is the mean of each principle's associated score multiplied by the number of automatically verifiable barriers related to its checkpoints detected in the web page.

2.3.4 Current State of Web Accessibility

There have been many studies which have supported the claim by Lazar that "the majority of websites today remain inaccessible, even by rudimentary standards" [67]. Lazar bases this statement on results from a previous study of 50 homepages in the Mid-Atlantic United States, where depending on the type of organisation 70% - 98% of the web sites assessed were found to be inaccessible. Previous work by the author [6] confirmed these results showing that although there has been some improvement in the accessibility of UK based web sites; 80% of commercial web sites, 92% of university web sites and 98% public organisation web sites contained inaccessibility barriers. Sloan [97] assessed the accessibility of 11 UK universities and found that only 2 did not contain any of the most basic priority 1 W3C accessibility barriers and that all contained priority 2 and 3 accessibility barriers. Alexander [4] assessed a selection of "key pages" from 45 Australian university web sites and found that 98% failed to comply with Australian anti-discrimination legislation (i.e. they contained W3C priority 1 accessibility barriers). Jackson-Sanborn [60] carried out an accessibility study of 6 different genres within the USA and found that 66.1% failed accessibility assessment. These studies are already several years old, and the nature of web site publishing is that content and structure change frequently (see section 2.1.6). However, a recent article by BBC Online reported that "Ninety seven percent of websites did not provide even minimum levels of accessibility" [9].

2.4 Information Visualisation

In order to display the accessibility data a practical and simple technique was required. As such, the area of information visualisation was investigated, there now follows an introduction to the subject, focussing on Tree-Maps and Graph visualisations. Too much information can be overwhelming. A mass of data is meaningless unless it is filtered, sorted and presented in a understandable fashion. One such fashion is visualisation. Visualisation is the practice of graphically representing data or as Brodlie defines it, "a transformation of numbers into pictures" [14]. McCormick describes visualisation as "a method for seeing the unseen" [76]. The principle behind this method is that most humans are better suited to interpreting visual imagery than long lists of figures. Shneiderman [83] explained in an interview (pg. 209) why he favoured the use of human analysis of visual information over computer processing raw data. He states "We're ... capitalizing on the remarkable human skill to see patterns. It would be hard to write a computer program to find 'interesting patterns' but you can see them right away" [83]. Shneiderman [93] follows this up with the statement that "abstract-information visualization has the power to reveal patterns, clusters, gaps or outliers in statistical data, stock-market trades, computer directories, or document collections." [93]. Fekete [47]agrees with Shneiderman [47] asserting that "to be effective, visualization techniques should rely as much as possible on preattentive graphical features". Preattentive features are those that are taken in humans "at a glance" without having to be too attentive to the specific low level details of a visualisation. Healey [18] defines preattentive processing as "an initial organization of the visual field based on cognitive operations believed to be rapid, automatic, and spatially parallel.". Examples of features which can be detected in this way are: "hue, intensity, orientation, size, and motion" [18]. An example might be a rock garden made up of 19 white stones and one black stone. The stone which stands out will be the one with a different hue (colour) than all the others. The black stone is almost immediately obvious and because of its different colour.

2.4.1 Goals of Visualisation

The Special Interest Group on Computer Graphics [1] lists the goals of visualisation in science as:

Goal 1: exploration/exploitation of data and information

Goal 2: enhancing understanding of concepts and processes

Goal 3: gaining new (unexpected, profound) insights

Goal 4: making invisible visible

Goal 5: effective presentation of significant features

Goal 6: quality control of simulations, measurements

Goal 7: increasing scientific productivity

Goal 8: medium of communication/collaboration

These goals reflect the aim of visualisation to gain greater value or use out of existing raw data. Visualisation also provides easier communication of information, and through abstraction, allows more information to be displayed in less space. Shneiderman refers to this as the *bandwidth of information* and notes that "the bandwidth of information presentation is potentially higher in the visual domain than it is for media reaching any of the other senses" [95].

Goals 3, 5 and 8 are the most significant to this thesis. **Goal 3** is to gain new or profound insights. These might be insights into a failing web strategy or the poor quality management within certain sections of a web site. Specific methods of publishing information on the web might prove to be more problematic than others. For example, documents converted from proprietary formats (such as Microsoft Word) may prove less accessible than HTML originals. **Goal 5** is very relevant as a large part of the thesis work will be to evaluate how the presentation of significant accessibility features assists managers and web maintainers. As such there are two points to evaluate, firstly whether or not the tool presents features effectively and secondly what impact this presentation has on overall accessibility of a web site. **Goal 8** is important because any change in policy or practice based upon findings or insights provided by the tool will have to be discussed. It is crucial then that any visualisation used communicates these insights simply and efficiently in order to support the argument being made.

2.4.2 Potential Techniques for Web Sites

When considering techniques for visualising a web site there is one option that is immediately obvious, a graph. The concept of a collection of related nodes and edges maps easily to that of web pages and hyperlinks. Herman [56] offers a simple method to determine whether to use a graph to visualise a particular kind of data. One must ask, "is there an inherent relationship amongst the data elements to be visualized?" [56]. When visualising web sites the main data elements are the web pages and they are linked by either a hierarchical link (i.e. their physical file location or logical location in the information hierarchy) or a hyperlink.

There are specialised modelling languages available, such as WebML [16], these are in essence graphs with extra semantics included. They are not investigated in this thesis as they are considered to be better suited to representing data and control flow than the structure and accessibility of web sites.

This chapter will focus mainly on graphs and a specific type of graph a Tree-Map; a more detailed rationale for this choice is given in chapter 3.

2.4.3 Graphs

Graphs have a number of attributes that lend themselves to visualising an overview of a web site. This section discusses those reasons. They are as follows:

- **High level representation** - allows users to gain an overview. Algorithms can alter layout to make the presentation more user-friendly and hence facilitate pattern spotting.
- **Extra information** - can be encoded within graphs by changing the size, colour or shape of nodes.
- **Filtering** - can be easily incorporated, sub-sites can be hidden or expanded.
- **Animation** - provides a method of either showing how the represented data changes over time or to reflect user triggered changes in the graph's structure. Yee [114] uses a novel animation technique for "animating the transitions from one view to the next, in a smooth, appealing manner" [114]. Yee's technique allows users to focus on one node, this then becomes the central point in the graph and all other nodes are relocated around this focal node.
- **Extensibility** - allows graphs to specialise. Many UML diagrams are simply graphs with added restrictions and set rules, for example, class and activity

diagrams. UriGraph [96] is a graph based modelling language that visualises both a web site's structure and the transitions linking one page to another. Rather than just showing only hyperlinks it is able to model dynamic pages by representing query or structural (directory paths) driven transitions encoded within the Uniform Resource Identifier (URI).

- **Easily understood** - graphs and trees are found in many guises from project management charts to the London Underground tube map and family trees.

Chen [115] has created a system (WebCiao) which "analyses web pages of selected websites, stores their structural information in a database, and then allows users to query and visualize that database with graphs or HTML" [115]. The graph based user interface is used to display the results of queries executed. The graph is not strictly tied to the web site structure. If a query returns disjoint sections of a web site then these sections are displayed side by side. Graphs are also used to highlight changes to a web site. WebCiao [115] colours nodes (representing web pages) that have been modified or deleted, whilst also displaying hyperlinks as labelled edges between nodes. The labels used are taken from the text or image name used in the hyperlink.

2.4.4 Issues in Graph Visualisation

One of the main problems when presenting large amounts of information is the danger of *information overload.* Websters dictionary defines it as "an overwhelming feeling upon the receipt or collection of an indigestible or incomprehensible amount of information". As the idea behind visualisation is to increase the understandability of data this is an issue that must be addressed. Herman [56] also highlights the "issue of viewability or usability" [56]. If too many nodes and edges are squashed into a small space it becomes impossible to distinguish between them. This impacts on the usability of the interface as it makes selecting specific nodes or edges very difficult. Another problem to be addressed is that of context. It is also difficult for users to keep track of their relative position in a large graph and still be able to gain detailed information about nodes and node clusters. This issue is discussed further in section 2.4.6.1.

2.4.5 Graph Layout

Layout is an aesthetic property that impacts greatly on the "preattentive" [48] quality of the graph. A good layout can aid the discovery of new and interesting patterns whereas a poor layout can obscure them. Herman [56] notes that different layouts can cause users to have different perceptions for the same graphs. Another important quality of a layout algorithm is predictability meaning that an algorithm applied to the same or similar graphs will have consistent layouts. Some popular graph layouts will now be reviewed.

2.4.5.1 Layout examples

- **Tree Layout** - places children beneath their common ancestor (like a genealogy tree). The classic example of such a layout is by Riengold and Tilford [85] who have collected aesthetical requirements to produce four criteria to which a tree layout must conform, as follows:
    - Nodes at the same level of the tree should lie along a straight line, and the straight lines defining the levels should be parallel.
    - A left son should be positioned to the left of its father and a right son to the right.
    - A father should be centred over its sons.
    - A tree and its mirror image should produce drawings that are reflections of one another.

These requirements ensure a balanced and symmetric layout forming triangles with the top point being the parent node.

- **Radial Layout** - "Nodes are arranged on concentric rings around the focus node." [114]. Immediately neighbouring nodes are plotted along the inner most circle and their neighbours lay on the next circle out and so on. Radial layouts create very aesthetically pleasing graphs and the use of focal points within a graph simplifies cluster spotting.
- **Tree-map** - makes "100 per cent use of the available display space" [61]. Tree-maps present hierarchical data using nested rectangles. Colour coding and size are used to convey information and text labelling can be added to aid comprehension. The advantages of this visualisation technique are that it is

possible to see the data density possible and the ability to give an overview of the entire data set and still provide enough detail about individual datum. According to Schneiderman, "Tree-Maps provide an overall view of the entire hierarchy, making the navigation of large hierarchies much easier" [61].

2.4.6 User Interaction

Dreyfuss [43] made the point to his designers that "what we are working on is going to be... in some way used by people individually or en masse. If the point of contact between the product and the people becomes a point of friction, then the industrial designer has failed" [43]. This section will focus on the challenges in designing graph based user interfaces. The problems of navigation, information overload and presenting detailed data without losing context (getting lost in a sea of data) has been commented on in section 2.4.4.

To tackle the problem Shneiderman [95] cites the information visualisation mantra of "Overview first, zoom and filter, then details on demand" [95]. Providing a birds-eye view to the user as it allows preattentive detection of patterns and clusters. Users can zoom in on these areas of interest and also remove useless data to hone in on specific data. This hiding of detail until the user requests it solves the problem of information overload but can create another one in the form of loss of context. A good example of loss of context might be a roadmap of the UK. If a users zooms to a small scale on a town in the north west of Scotland how can they relate this to the rest of the UK and navigate to the capital London (in the South East of England).

2.4.6.1 Focus Context

Herman [56] points out that, "A well known problem with zooming is that if one zooms on a focus, all contextual information is lost" [56]. Maintaining a balance between the need for detailed information about specific portions of the visualization, whilst remaining aware of the overall picture is a difficult problem. The fisheye distortion is one of the most popular techniques aimed at solving it. With the fisheye, areas under the user's focus are distorted to magnify the data under it. Bjoerk [11] states that "the basic idea with focus+context visualizations is to enable users to have the object of primary interest presented in detail while at the same time having an overview or a context available" [11].

2.4.6.2 Heterogeneous context

Another technique is to provide "user interface widgets" [79]. These widgets can add information or functionality to "select a (new) focus, filter out extraneous information... ...and create query criteria for finding particular information." [79]. The mixing of different types of information (for example providing crime or health statistics alongside the map of a certain location) provide, as Bjoerk [11] discusses "heterogeneous context". Merging different types of data within the same presentation can lead to a better perception of data and the relationships between them. Allowing users to specify what data to merge allows them to customise the visualization to suit their needs and hence get the most value out of it.

Filtering out unnecessary information allows a simplified and customised version of the visualization to be presented to the user. This simplification places less cognitive load on the user and improves the "clarity and simultaneously increases performance of layout and rendering" [56]. Querying can be added to the interface within a widget to enable users to focus in on specific regions. With the previous example of a map of the UK users might know the post code of the location they wish to view. Querying the application with this postal code can instantly set the focus to the correct location at an appropriate zoom level.

Kimelman [63] uses a variety of operations that filter out or "dispose" of nodes

- **ghosting** - relegating nodes to the background with the use of "ghostly" grey colouring
- **hiding** - completely removing nodes from the display
- **grouping** - grouping up of nodes into a new meta-node or cluster. One example might be using sub-site nodes to represent all the pages within that site.

2.4.6.3 Clustering

The final operation of "grouping" [63] relates closely to clustering. "Clustering is the process of discovering groupings or classes in data based on a chosen semantics." [56]. Clustering creates new groups that share particular attributes or are linked

together. If web pages are clustered by content (this could be achieved by counting key words or relying on meta-data) rather than by their position within a structure (either physical file or information hierarchy) it would be a simple task to spot all pages relating to software engineering or chemistry.

2.5 Summary

This chapter has presented a review current state of the art in web accessibility and information visualisation. It examined the nature of web publishing and reviewed the following issues:

- Content Management,
- Content lifecycle models and how this can be implemented,
- Web maintenance with a focus on the differences between traditional software maintenance and web software.

Following this there was a discussion of web accessibility including the principles behind accessible web design, why it is important and how it can be achieved and assessed. Finally information visualisation in relation to web sites was discussed with examples of graphs and other techniques given. Particular emphasis is given to the benefits and costs of using graphs to represent web pages. The following chapter presents the novel adaptation of the Tree-Map visualisation for representing web sites.

3 Tree-Map Visualisation of Web Accessibility

3.1 Introduction

This chapter describes the creation of a prototype tool used to investigate whether a Tree-Map based tool could be used to support the creation, evolution and maintenance of an accessible web site. Tree-Maps are a fixed space two dimensional visualisation technique (see section 2.4) which are designed primarily to visualise hierarchies. Web sites are often organised into hierarchical structures to aid information retrieval and comprehension and hence the Tree-Map was chosen as the basis of the Graphical User Interface for the tool. There now follows a more detailed explanation of why a Tree-Map was selected as the basis for the tool, how it was implemented with examples of usage scenarios given. Finally a summary is given followed by discussion of future work. It should be noted that the adaptation of the Tree-Map to represent web accessibility is novel and forms part of this thesis's original contribution to the state of the art.

3.1.1 Visualising web accessibility within a Tree-Map

There are many web accessibility assessment tools available (see section 2.3.2). These tools are available as either web based services or as stand alone applications and are focussed on the repair of individual pages highlighting individual web accessibility barriers within each page. The main focus of the Tree-Map based tool is not to highlight the errors within individual pages, but to highlight which pages within the web site are the most inaccessible. By providing a visualisation of the entire web site (or if necessary a sub-section of the web site) web pages which require the most urgent attention can be quickly identified and repaired no matter where they are in the web site hierarchy. This reduces greatly the risk that pages are forgotten about or ignored and because interpreting the visualisation does not require a large amount of technical expertise it can be used by both managerial and technical personnel.

3.1.2 Benefits of Tree-Maps

There are several key benefits of using Tree-Maps as the visualisation basis for this tool; they are described in the section that follows:

3.1.2.1 Scalability for larger web sites

One of the main benefits of using Tree-Maps to visualise a website is the visualisation is fixed in size. Each node in the Tree-Map represents a web page and when new pages are added to the web site; the visualisation itself does not become larger. Instead, the nodes representing individual pages are resized to allow the nodes representing the new pages to be inserted. This means that a web site with 10,000 web pages can be represented in the same screen area as a web site with 100 web pages. This is in contrast to many graph based visualisations, where as size and connectivity increases, so too does the required screen area [55]. If the graph becomes too large to be shown on one screen, the user is required to navigate around the graph using either scrolling or zooming techniques. This introduces the risk that the user could become lost within the hierarchy. Herman points out "If the number of elements (in a graph) is large it can compromise performance or even reach the limits of the viewing platform." [55]

3.1.2.2 Web page context and focus

Keeping track and control of hundreds of web pages or indeed a smaller set within a large web site can be difficult. As mentioned in section 3.1.2.1, there is a scalability issue which must be addressed. Content authors might "lose" or "forget" about certain pages. To improve efficiency a method is needed to locate, navigate to and focus in on their web pages. The Tree-Map tool facilitates this through a zoom function. In the higher level view details of web pages are hidden, hence allowing more pages to be displayed together on the same page. This provides the overview required to locate pages and gather the contextual information (such as what other pages are contained in the same section of the site). Once a web page has been located and more detail is required, the content author can "drill down" the hierarchy to discover more details and information about the web page and even perform accessibility assessments. Zooming increases the individual sizes of nodes but the relationship between the sizes of nodes within the zoom is not affected and hence visualisation consistency is maintained.

3.1.2.3 Accessibility information encoding

It is important to encode as much information as possible within the Tree-Map; this allows content authors to gain a greater understanding of the state of a web site in less time. Tree-Maps encode information in the four ways originally specified by Schneiderman [92]:

1. **Location** – each node is located into an appropriate position according to the pages position in the web site information hierarchy.
2. **Colour** – each node is coloured to represent its accessibility.
3. **Size** – each node's size is determined by the size of the web page.
4. **Text labels** – if under the current view the node is large enough, then the filename of the web is superimposed as text onto it.

In addition **coloured highlighting** was also used to identify pages with specific properties. Highlighting is achieved by adding a coloured border around the node, the colour was determined by which property is of interest. For example, the maintainers of specific pages could be quickly shown by highlighting the pages with colours corresponding to certain maintainers. Each encoding technique is now described in more detail.

Node colour

The node colour was based on PageMeasure for the web page (this is described in section 3.2.4). The PageMeasure is a positive real number and from this number an RGB value (Red-Green-Blue) Java 'Color' class was generated. The decision was taken to use the red and green values at varying intensities to create a two stage colour scheme to represent the PageMeasure's colour. The RGB value for each node's 'Color' object was determined as follows:

- If the PageMeasure = 0 then the RGB was (0, 0, 0) was black.
- As the PageMeasure increases the green value increases in intensity in steps. The increase is decided by the PageMeasure which is multiplied by a modifier value (1.5).
- If the PageMeasure is above a threshold (set at 5.5), the intensity of the red value is increased. This created a magenta which grows in intensity as the PageMeasure increases.

Category	Rank	PageMeasure Range
Very Accessible	1 or 2	<= 3.5
Accessible	3	> 3.5 <= 4.5
Requires Attention	4	> 4.5 <= 5.5
Inaccessible	5	> 5.5

Table 3-1 Score Range to Category

Table 3-1 shows the categorisation of web page's accessibility based solely on their PageMeasure. The PageMeasure ranges are a modified version of those used in previous work [6] and were also used in the automated accessibility assessments described in Section 5.5.1.1. Introducing the increase in the red alpha intensity differentiates clearly web pages which are considered "Inaccessible" and hence need attention. Examples are shown in the areas pointed to by arrows in Figure 3-1.

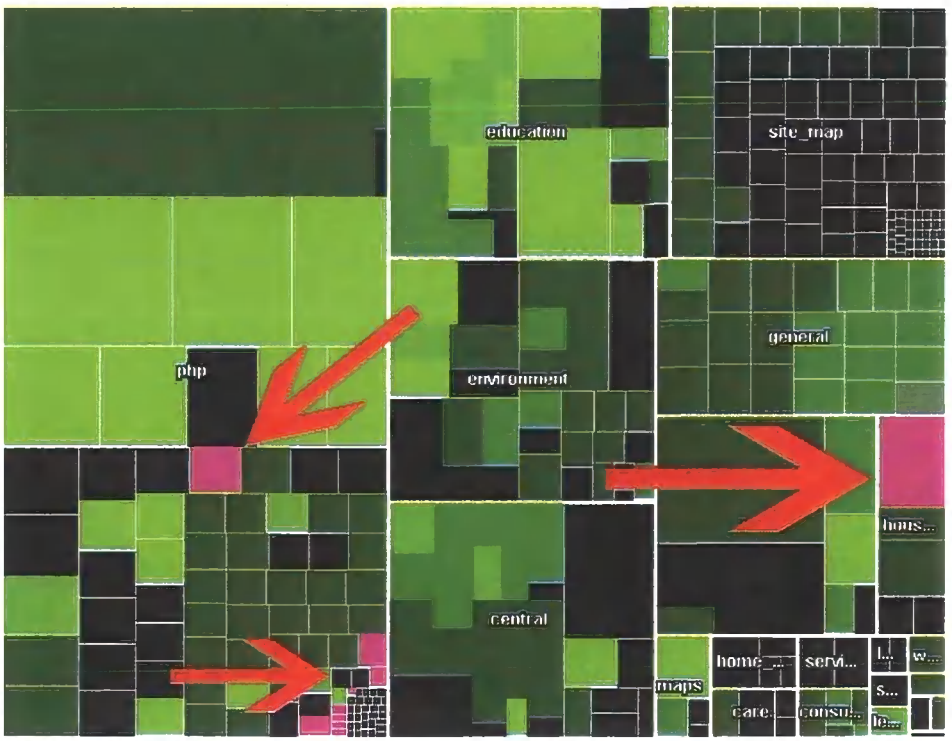

Figure 3-1 Example of Tree-Map with magenta "Inaccessible" web pages

Node size

Each node represents a single web page. The size of a node is determined by the number of HTML elements contained in the web page. This means that pages which

have more HTML will be represented by larger nodes. The size of the web page is important for maintenance as larger web pages may require more maintenance than smaller ones as the barriers might be spread throughout the page.

Node highlighting

Highlighting specific nodes is useful as it allows the results of a search to be immediately visualised with context and frequency information encoded. A manager may wish to highlight all pages maintained by a specific content author. By highlighting the results of this search within the Tree-Map the manager will establish immediately how many pages the content author maintains and where those pages are within the hierarchy of the web site. Highlighting was achieved through drawing a coloured border around each node of interest (see Figure 3-2). This, rather than a colour change of the actual node, was chosen so that it would not interfere with the accessibility information encoded in the node's colour. It may be that the accessibility of pages maintained by a content author is of particular interest and so any change in the node colour is not desirable. Within the tool each content author or web maintainer was assigned a specific colour this meant that the results from multiple searches could be combined on the Tree-Map, making visual comparison of results possible.

Figure 3-2 Example Tree-Map with four web pages highlighted with yellow border

3.2 Tool Implementation

The Tree-Map based tool was implemented using a combination of existing software tools, Java and PHP based scripting.

3.2.1 Overview of stages

Table 3-2 lists the steps involved in generating a Tree-Map.

	Action	Performed by	Output
1	Spider / walk through website, collecting web pages.	Webbot [27]	Text file containing URLs
2	Evaluate each web page against W3C guidelines.	PageValet using customised XSLT style sheet inside Java wrapper.	Directory containing one text file per web page
3	Process each text file generating and collecting accessibility metrics for each web page	New Java application	CSV summary file containing each URL and its corresponding metrics
4	Conversion from CSV summary file to XML document representing a hierarchical view of the Website	New Java application	XML file containing of URLS their metrics, and other maintenance data
5	Generation of Tree-Map with added accessibility and maintenance data	Modified JTreemap Java classes	Interactive Tree-Map user Interface

Table 3-2 Stages involved in the generation of the Tree-Map

3.2.2 Stage 1: Collection of web pages

In order to build an accurate representation of a web site it was necessary to "visit" each page from the client side rather than using the file system. This means all pages including those generated dynamically by the web server will also be collected. In order to find other pages the Webbot [27] visits the homepage of a web site, parses it retrieving all hyperlinks and then repeats this process for each web page exhaustively. This allowed the tool to gain the same view of a web site as a user. The spider can only locate pages connected (via hyperlinks) to the homepage of a web site. One

advantage of this approach is that it makes assessing web sites whose file systems are not available possible. Webbot [27] was chosen to collect web pages for the following reasons:

- It is lightweight and fast.
- It offers extensive customisation. This means it is possible to restrict the set of web pages visited.
- It is available for use free of charge.

Because the Tree-Map based tool was only a prototype, the following restrictions were made on the Webbot's search:

- Only HTML based web pages were collected – this was because the assessment tool PageValet [112] can only parse HTML and so to download non-HTML pages (such as word processor documents or other binary formats) would have been very time consuming and offer no benefit.
- The Webbot's search was fixed to the internal web pages only – any pages which did not share the same domain name as the home page were considered external and hence not relevant to the assessment.
- The depth of the search was set to 3 hyperlinks from the original home page – this was to limit the search space and save time. It was felt that the most important and representative web page would be reachable within 3 links from the home page.

3.2.3 Stage 2: Automated accessibility assessment of pages

The standalone command line version of PageValet [112] was used to assess each web page collected by the Webbot. PageValet offers several parameters to alter the output:

- Different assessment guidelines can be used. There are two sets of guidelines against which the page can be assessed; W3C and Section 508 [52]. For the Tree-Map based tool only W3C checkpoints were used. This was because the W3C guidelines are internationally recognised and its checkpoints are very similar to the Section 508 checkpoints.
- As each W3C checkpoint is prioritised, it is also possible to restrict the tool to identify checkpoints which have a specific priority. This extra level of

specificity is useful for organisations which have decided to limit the removal of barriers to those considered more serious. PageValet was set to check checkpoints from all 3 priorities.

- A XSL style sheet was created to convert PageValet's XML output to a comma delimited text file.

3.2.3.1 PageValet Output

By default PageValet generates a report in an XML based Evaluation and Report Language (EARL) [23]. This report contains:

- An XML representation of the assessed web page's HTML. This XML version of the web page HTML is annotated with potential accessibility barriers identified by PageValet.
- A summary of the accessibility barriers which have potentially been breached, and references to individual instances of the barrier. Each barrier is qualified by a confidence rating: Certain, High, Medium or Low. (see section 3.2.3.4)
- The settings used to configuration PageValet and a final result (whether the web page has passed or failed).

From this original EARL report an XSLT style sheet was used to transform this output into a comma delimited text file. This transformation reduced the amount of storage required and focused solely on elements of the PageValet report that were used in generating the accessibility measures. These included:

- HTML elements (tags) (including bogus and deprecated elements) used within the document. It was important to keep this data to allow normalisation of any accessibility metrics derived from the reports. For example, larger pages are more likely to have more barriers simply because they contain more marked up content.
- Each W3C checkpoint (potentially) breached alongside a corresponding confidence rating. Confidence ratings indicate how likely an accessibility barrier detected is a true positive. Ratings provided were "No Confidence", "Low Confidence", "Medium Confidence", "High Confidence" and finally for those checkpoints that can be assessed completely automatically (machine readable barriers, such as syntactical errors), "Certain".

- A final result from PageValet stating whether the page passed or failed the automated assessment. To reflect the lack of absolute certainty (e.g. if a document has been written clearly) in many cases the qualifier "Probable" is used. In this case it is necessary for a human to verify PageValet's results.
- Meta-data provided within the HTML document header. These are Meta-data provided by the content author or software which generated the HTML of the web page under assessment. Depending on the web page it may be possible to retrieve who created the web page, which department the web page belongs to and when the page was last updated. Some web sites also produce meta-data stating which accessibility ratings the page has passed.

3.2.3.2 Conversion to plain text and measures

Once each web page has been parsed, it was transformed using the XSLT style sheet to a plain text file. These files are used to establish specific measures. The measures collected at this stage are as follows:

- Number of bogus, deprecated and valid HTML attributes (see section 3.2.3.3).
- Number of bogus, deprecated and valid HTML elements (see section 3.2.3.3).
- Potential breaches of W3C checkpoints, including W3C assigned priorities and confidence indicators provided by PageValet (see section 3.2.3.4).

3.2.3.3 HTML attributes and elements

HTML elements are the building blocks of a HTML based web page. Their purpose is to add structure to a document and to mark up content semantically. HTML attributes contain the specific properties of an element included in a page and add more information about how it should be interpreted by the user agent. Therefore, elements with the same names might be interpreted differently depending on their attributes. For example,

```
<img src="welcomeimage.jpg" alt="Welcome to Durham University's Web
site" />
```

The `<img>` element is used within HTML to include a graphic in a web page. The `src` attribute specifies the name of the file which contains the image to be displayed. The `alt` text contains the alternative textual version of the image. This will be displayed

if, for example, the image file has been deleted or if the user agent is not capable of displaying images.

All elements and attributes were retrieved from the web pages. Both elements and attributes fell into one of three categories: *valid*, *bogus* and *deprecated*. It was decided that for accessibility, greater weight should be given to HTML elements over attributes as bogus elements would not be rendered properly by a web client. Attributes are only used to provide extra information about elements. For example, if the "**table**" elements were mistakenly written as "**stable**" this would cause a much greater rendering problem than if a valid "**table**" element had a bogus attribute "**address**" as this would simply be ignored by the rendering device. Similarly a distinction was made between bogus and deprecated HTML. The W3C define deprecated as, "element or attribute is one that has been outdated by newer constructs." [25]. This means that previously the element or attribute was valid and therefore rendering devices should be aware that legacy code may contain such mark-up and make adequate adjustments. Hence such out of date mark-up was considered less of a barrier compared with HTML that has never been valid.

3.2.3.4 W3C Checkpoint Breaches

PageValet produces a list of potential breaches of the accessibility checkpoints. These are represented as EARL `<testref>` elements. For the Tree-Map based tool these were the W3C 1.0 checkpoints. Each checkpoint contains a hyperlink to an explanation on the W3C WAI Guidelines webpage. For example

```
<testref href="http://www.w3.org/TR/WCAG10/#tech-tab-order"
confidence="Low" message="Create keyboard shortcuts and/or a logical
tab order between controls."
id="E33"><node>x65</node><node>x63</node></testref>
```

Figure 3-3 Potential accessibility barrier within PageValet report

Figure 3-3 shows the XML element produced for each potential breach detected by PageValet. Each breach is reported by a "`testref`" element and instances of where the breach occurs are referenced within the child "node" elements within the "`testref`". In this example, the hyperlink in the "`href`" attribute links to the

definition of the checkpoint (9.4 in this case) on the W3C web site [26]. Although the priority of the checkpoint is not contained within the "`testref`" element, the priority was determined by using the hyperlink in the "`href`" to identify the specific W3C checkpoint breached and from this its priority. For example, the URL http://www.w3.org/TR/WCAG10/#tech-redundant-client-links points to a description of a specific checkpoint and contains it priority level. Rather than continually parsing the W3C guidelines a database containing all W3C guidelines and their checkpoints was created. This greatly improved the speed of the Tree-Map based tool. The confidence attribute in Figure 3-1 indicates that PageValet only can only offer a low certainty that the controls identified are in need of modification. Hence human confirmation is required. The specific HTML elements that have triggered this barrier warning are listed within the child "node" elements.

Once transformed through an XSLT the equivalent data shown in Figure 3-3 appears as follows:

```
4^E33^http://www.w3.org/TR/WCAG10/#tech-tab-order^^Create keyboard
shortcuts and/or a logical tab order between controls.^Low^x65^x63
```

Figure 3-4 Transformed Test Result from PageValet

Figure 3-4 contains the barrier test information but is represented as a comma separated file. Subsequently, each line in the text file contains the following data:

1. A number indicating the type of data, for example, 4 represents a potential barrier warning
2. PageValet's internal ID for each accessibility barrier (`testref`).
3. The hyperlink to the specific checkpoint within the W3C guidelines web page. This was used as a unique key in the database.
4. Description of the accessibility barrier.
5. Confidence rating generated by the PageValet tool. This was eventually ignored during the course of the tools development and only barriers which could be automatically assessed were used in the generation of the PageMeasure.
6. References to individual accessibility barrier breaches within the web page.

3.2.4 Stage 3: Generation of a PageMeasure for each page

Although the PageMeasure was developed independently of Zeng's WAB score [117] (see Figure 2-8) it uses a similar method to measure accessibility. The main difference between Zeng's score and the PageMeasure was that originally the PageMeasure formula included barriers which required manual assessment; it used the confidence rating provided by the PageValet tool to weight certain barriers. However, investigation into the reliability of this confidence rating and after consultation with the tool's author it was determined that only barriers which could be automatically verified as true positives were included. Hence the two methods are now almost identical. They both measure and classify the number of potential barriers contained in a web page, based on the W3C accessibility guidelines. The only difference is that the PageMeasure includes attributes of elements when normalising the measure. Figure 3-5 shows the equation used to work out each web page's PageMeasure.

$$\frac{\sum \forall_{i \in \text{issues}} (N_i . P_i^{-1})}{\text{Attribs} + \text{Elements}}$$

Figure 3-5 Page Measure Equation

Each accessibility barrier or issue detected has a related W3C checkpoint with its W3C priority ranging from one to three. Priority one issues are considered the most severe, followed by priority two and then three. The number of instances (N_i) found is then multiplied by the inverse of the checkpoint's priority. For example, priority one issues have a weight of 3 whereas Priority two issues are weighted as 2. Ensuring that the PageMeasure of web pages of different sizes can be compared to the accessibility score is normalised. The sum of multiplying the number of instances by the inverse of their barrier's priority is divided by the total number of HTML attributes (Attribs) and elements (Elements) contained within the webpage. This normalisation ensures that pages of different sizes can be compared fairly. Both attributes and elements are included in the normalisation because they can both cause accessibility barriers. For example, an element, such as the `<img>` or `<a>` can cause more than one barrier if one or more of its attributes are bogus or incorrectly used.

3.2.5 Stage 4: XML representation of the web site

Once each web page has been collected, assessed and has a PageMeasure calculated, the next stage is to rebuild a hierarchical representation of the web site. In order for this representation to be readable by the software which generates the Tree-Maps (JTreemap) an XML file was created which organised the site into branches and leaves (see Figure 3-7 for a DTD of JTreemap's XML format). The web site hierarchy was built by parsing the URL of each web page in the site. This created a tree based on the directory structure of the URLs. For example, Figure 3-6 shows how a site with 4 pages and 6 directories is parsed using the '/' character to form a new level within the tree structure.

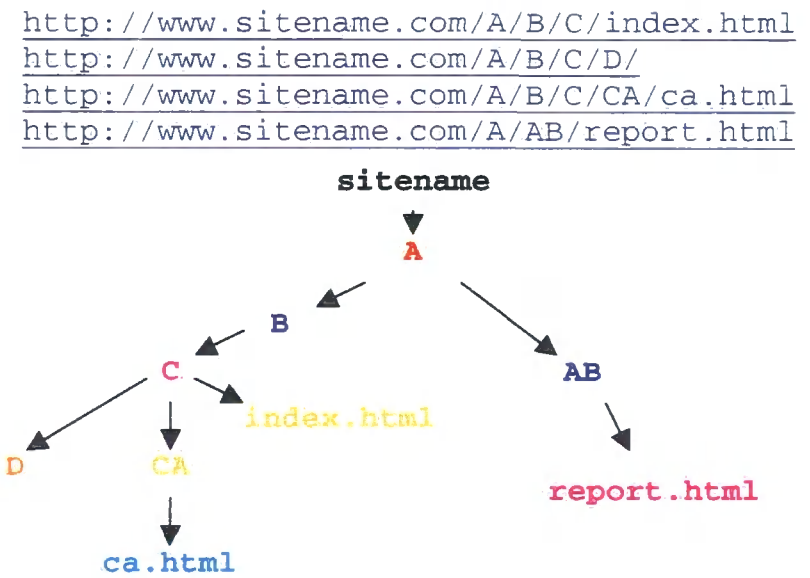

Figure 3-6 URLs were used to form the site hierarchy.

Figure 3-6 shows how the tree structure was generated from the URL of each web page. Web pages (ca.html, report.html and index.html) are represented as leaves while directories (A, B, C, D, AB and CA) are branches (or sections).

For each directory if no index page is explicitly stated in the URL then one is added. This is because (dependent on the web server settings) the URL http://www.dur.ac.uk/ is equivalent to http://www.dur.ac.uk/index.html. The server will present the browser

with either an explicitly created index page or a default index page created by the web server.

This hierarchy was stored as a `java.util.Hashtable` data structure and from that written out as XML (see section 3.2.5.1).

3.2.5.1 JTreemap XML Format

The XML language used to represent Tree-Maps is relatively simple. In the original DTD there were three basic elements that were used to build the Tree-Map: *root*, *branch* and *item*. These represent respectively the root, branch and leaf nodes within the tree structure. Each of the three elements must contain exactly one *label* element. This is used to help distinguish between the different sections of the visualised Tree-Map.

Each Tree-Map is made up of branches and item nodes. The branches can be nested and represent branches of the hierarchy tree. Hence, whereas a branch node represents a section of the web site, item nodes represent a web page. An item element contains two child elements, *weight* and *value*. Weight corresponds to the relative size of the node and value to its colour.

```
<?xml version='1.0' encoding='ISO-8859-1'?>
<!ELEMENT root (label,(branch |item)*)>

<!ELEMENT branch (label,(branch|item)*)>

<!ELEMENT item (label,weight,value,url*, attribute*) >

<!ELEMENT attribute EMPTY>

<!ATTLIST attribute name CDATA "">

<!ATTLIST attribute value CDATA "">

<!ELEMENT label (#PCDATA) >

<!ELEMENT weight (#PCDATA) >

<!ELEMENT value (#PCDATA) >

<!ELEMENT url (#PCDATA) >
```

Figure 3-7 XML DTD for JTreemap

Figure 3-7 shows the DTD for the JTreemap XML Parser. The DTD was modified slightly to allow the inclusion of specific attributes. These modifications are marked in Figure 3-7 **as bold and underlined**.

```
<root>
  <label>sitename</label>
  <branch>
    <label>A</label>
    <branch>
      <label>B</label>
        <branch>
          <label>C</label>
            <branch>
              <label>D</label>
            </branch>
            <branch>
              <label>CA</label>
              <item>
                <label>c.html</label>
                <url>http://www.sitename.com/A/B/C/CA/ca.html</url>
                <attribute name="PageMeasure">10.0</attribute>
                <attribute name="Maintainer">j.o.bailey</attribute>

                <weight>50</weight>
                <value>10.0</value>
              </item>
          </branch>
```

Figure 3-8 Partial XML representation of web site

Figure 3-8 shows part of the resulting XML document which JTreemap uses to generate the Tree-Map. The modification to the DTD allowing "attribute" elements to be added allows the addition of maintenance and potentially any other measures (such as the Lazar rating (see Section 2.3.3.1) to be included in the Tree-Map based tool.

3.2.6 Stage 5: Generation of Tree-Map with maintenance information

As well as structural and accessibility information, maintenance information was also added to the XML representation. This was either gathered from parsing the meta-data of the web page (collected during the automated accessibility assessment in stage 2) or added via the Tree-Map based tool itself. To represent this extra information a PageInfo class was created and held the following data about the page.

Attribute	Description
Maintainer	The Person or Team object who/which is responsible for maintaining the web page
Creator	The Person object representing whoever created the web page
URL	The URL of the web page
Creation Date	Date the web page was created
Assessments	Mapping from Date -> PageLogAss object. This provides a history of assessment carried out for this web page
Last Assessed Date	Date the of the last assessment
Attributes	List of extra attributes. These allowed any extra data to be stored. For example, if an organisation has bespoke meta-data which it uses for its web site.

Table 3-3 PageInfo class main attributes

Table 3-3 contains the attributes of the PageInfo class. It can hold maintenance, accessibility and miscellaneous information (Attributes).

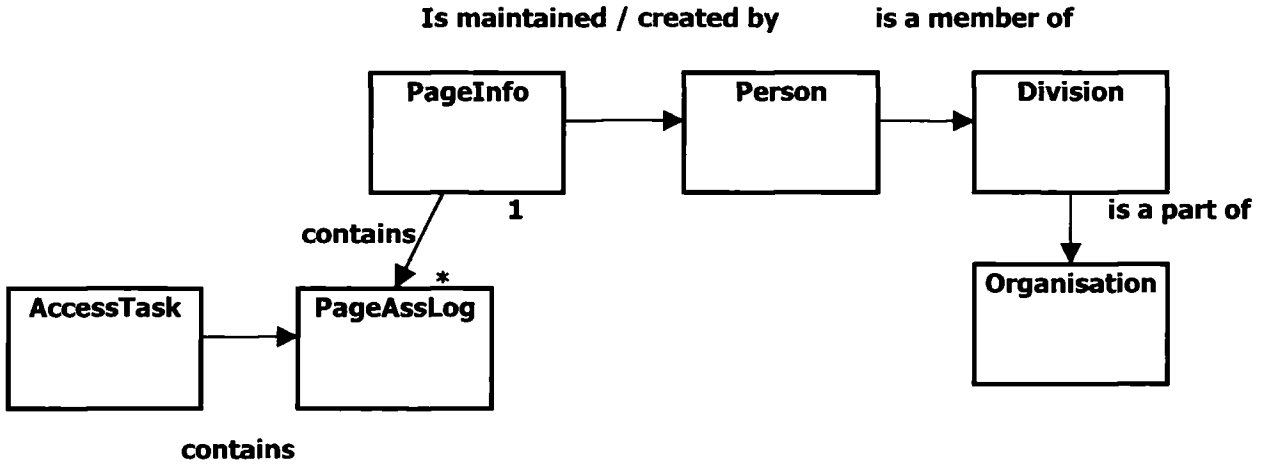

Figure 3-9 UML class diagram of PageInfo and related classes

3.2.6.1 Web page data classes

Figure 3-9 shows the relationship between PageInfo and the classes which hold information about who maintains the web page and their position within the web

site's organisation. It also shows how accessibility log information is captured. Table 3-4 and Table 3-5 show the main attributes contained within the PageAssLog and AccessTask.

Attribute	Description
PageID	Identifier of the webpage held in the database
Maintainer Email	Email address of maintainer that carried out the assessment (if different from the registered maintainer of the page)
CheckpointID	Identifier of the W3C checkpoint stored in the database
Confidence	Confidence indicator given by PageValet during the automated assessment.
IsActualBarrier	Records whether the barrier is a true or false positive.
Date	Date of assessment
Status	Current status of the assessment
NextDate	Date the next assessment is due to be carried out
Notes	Notes made during assessment
Nodes	Number of instances where potential barrier occurs

Table 3-4 PageAssLog Attributes

Table 3-4 shows the attributes of the PageAssLog. Its purpose is to store information about web page assessments. For each potential accessibility barrier detected by PageValet, a PageAssLog object is created. This ensures that every barrier in the page is recorded for auditing or review purposes. Some barriers detected by PageValet might be immediately recognised as false positives and as such the IsActualBarrier attribute will be set to false. Recording the occurrences of false positives provides precision and recall data for the automated tool being used (further discussion is made in section 7.2). The status attribute tracks the current status of each barrier for the page, possible values are: unseen, open, closed. Web pages containing unseen or open PageAssLogs can be displayed through the Tree-Map interface.

Attribute	Description
Maintainer Email	Email address of maintainer that carried out the task (if different from the registered maintainer of the page)
TechniqueID	Reference to a W3C recommended technique for the removal of the barrier identified by the checkpoint identifier of the page assessment log.
Status	Current status of the task (unseen, open, closed)
Notes	Notes made during task
Name	AccessTask's label

Table 3-5 AccessTask Attributes

Table 3-5 shows the attributes of the AccessTask class. Each AccessTask object is associated with a specific task required to remove an accessibility barrier. They are designed to be compatible with the recommended techniques published by the W3C [28]. For example, checkpoint 1.1 of the W3C guidelines states "Provide a text equivalent for every non-text element". If PageValet identifies a non-text element without such a text (i.e. an accessibility barrier) it generates a PageAssLog corresponding to that barrier for the page under assessment. It then presents the user with a list of potential AccessTasks based on the specific techniques created by the W3C to remove this particular barrier e.g. using the alt attributes for images.

3.3 *Usage Scenarios*

In order to show how the Tree-Map tool can support the identification of inaccessible web pages and the removal of accessibility barriers, several usage scenarios will now be discussed. The tool is designed to be used to monitor the current state of a web site and to highlight changes in and facilitate the improvement of web accessibility. As such, it is important to show the effects of commonly occurring web maintenance tasks.

3.3.1 Effect of web accessibility improvement

A common maintenance activity is to monitor web pages for accessibility barriers. In this scenario, a content quality manager responsible for the Postgraduate section of the web site (pages providing information for the department's postgraduate students)

uses the Tree-Map based tool to identify inaccessible web pages and remove accessibility.

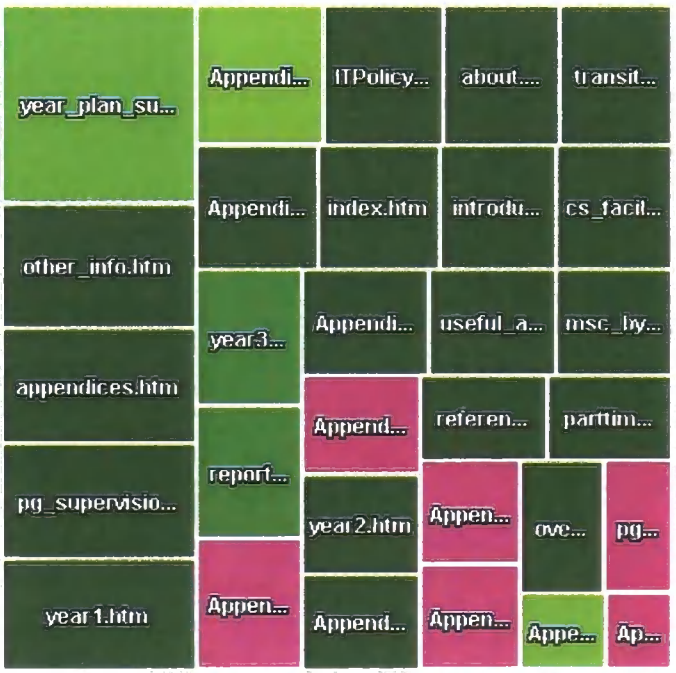

Figure 3-10 Tree-Map of Postgraduate section before accessibility improvement

Figure 3-10 is a zoomed in Tree-Map of the Postgraduate section. There are 6 pages which stand out in this Tree-Map as being inaccessible. These 6 pages are displayed in bright magenta. In this scenario these would be the pages that require accessibility improvement.

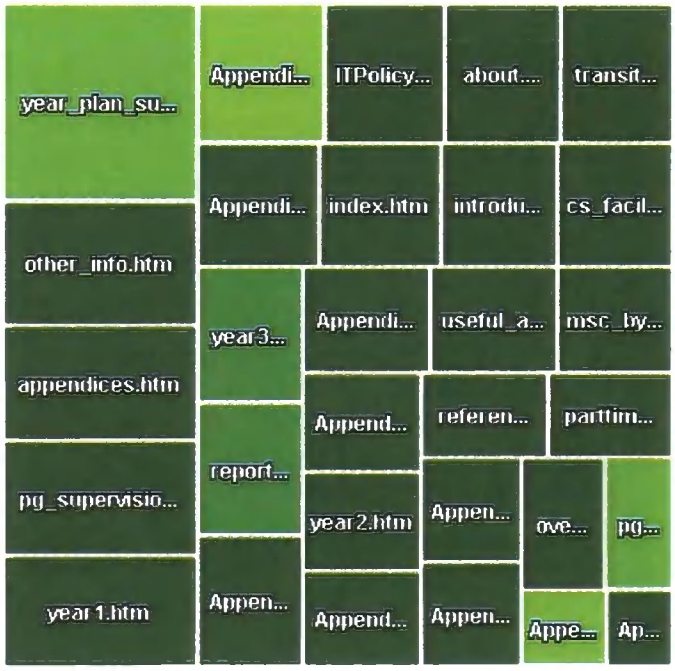

Figure 3-11 Postgraduate section after accessibility improvement

Figure 3-11 shows the same Postgraduate section of the web site as Figure 3-10. However, the pages that were identified as being inaccessible have been improved by addressing the accessibility issues detected by PageValet.

In this scenario the content quality manager uses the Tree-Map tool to identify pages which need attention, and used the tool to assign / notify the content authors responsible for those pages. The content author has then used the reports generated by PageValet to determine what needs to be changed. Once the change has been implemented both the content author and quality manager can confirm that the pages have been corrected through the Tree-Map tool, rather than inspecting the HTML of the web page.

3.3.2 Effect of new pages

Figure 3-12 Tree-Map of Computer Science Department web site

Figure 3-12 shows an overview of the Computer Science Department's web site.

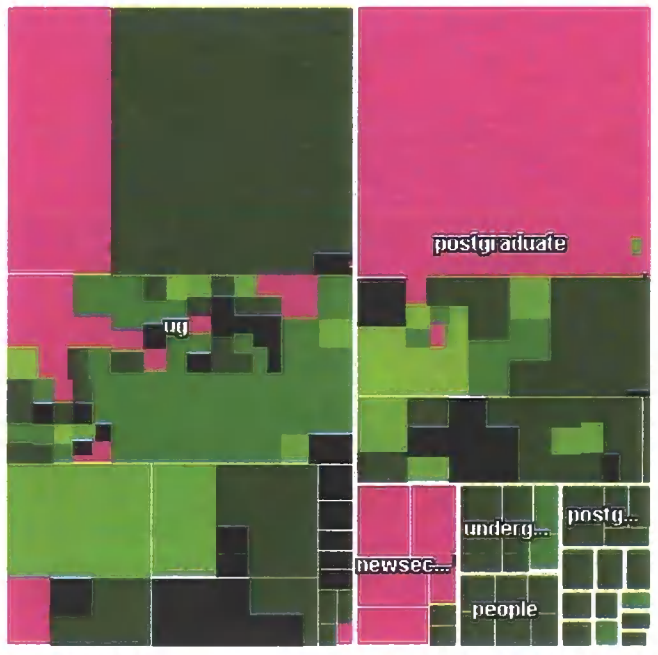

Figure 3-13 Tree-Map of Computer Science Department web site with new section added

Figure 3-13 shows the effect of adding section "newsection" to the web site.

When new sections are added to a web site, it is important to ensure that they comply with the organisations web publishing guidelines. It is obvious from Figure 3-13, that

the new section added (bottom-centre) is inaccessible. The pages added were taken from standard templates available freely. In this case the content author responsible for adding the pages will have to ensure that their accessibility is improved. It may be the case that the accessibility problems are such that different templates are sought for the production of new pages in the future.

3.3.3 Effect of different 3rd party product conversion tools

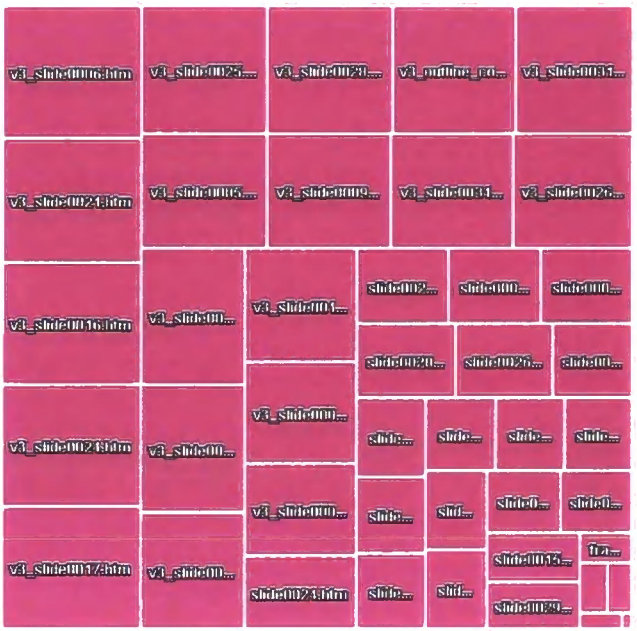

Figure 3-14 Slides created by default from MS PowerPoint

Figure 3-14 shows the result of creating a HTML version of a PowerPoint presentation on the web site using the default "saves as web page" function.

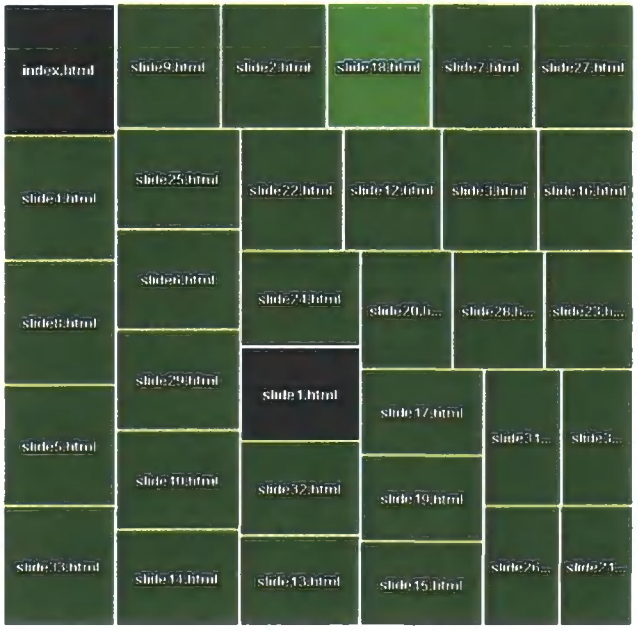

Figure 3-15 Slides created by using the accessible wizard [104] from MS PowerPoint

Figure 3-15 shows the HTML slides produced using an accessible conversion wizard [104]. The presentation from which the sliders were created is the same as in Figure 3-14.

The contrast in the Tree-Maps shown in Figure 3-14 and Figure 3-15 demonstrates the difference an accessible 3rd part product can make to the accessibility of a web site. The web pages produced through the default conversion method were much more inaccessible. Since many web sites publish web pages which have been converted directly from standard desktop publishing software this will prove to be a significant problem. The creation of such a tool by University of Illinois-Champaign [104] points to a shortfall in the way some desktop publishing software packages produce web pages.

3.4 Summary

This chapter has introduced a novel approach to visualising web accessibility through a Tree-Map. As part of this it presented the implementational details of a Tree-Map based tool and the theory behind its use. This included:

- Discussion of the benefits of using Tree-Maps as opposed to other forms of visualisation.
- How each page on the web site was collected and assessed.
- How the web site's were stored internally within the Tree-Map based tool.

Its applicability was tested in part through the use of several web maintenance scenarios. These scenarios showed:

- Changes in accessibility through a specific accessibility improvement can be spotted immediately.
- Additional web pages being added to a web site are obvious.
- 3rd party software being installed on the web site can cause major changes to the structure and accessibility of a web site.

The next chapter outlines a study of web maintenance practice alongside an experiment into the applicability of a Tree-Map

4 Studies and Experiment

4.1 Introduction

This chapter explains why it is important for web sites to have mature and systematic processes that ensure accessibility is considered at all stages of the web publishing model. It also presents a study aimed at establishing to what extent best practice has been adopted within organisations and whether this has had an effect on the overall accessibility of their web sites. Finally, a protocol for an experiment is presented. It is designed to evaluate a Tree-Map based tool, and to test whether it is more efficient at presenting web maintenance information than current report based tools.

4.2 Problem Description

The design and purpose of web sites have undergone a number of changes within a short period of time. Organisations no longer use web sites as a collection of hypertext documents giving for example, information about services or opening times [37]. Web sites have developed beyond just information systems and now include many forms of web applications [40]. With the added functionality and promise of convenience, comes the added responsibility on their creators and maintainers to ensure that disabled users are not discriminated against. This responsibility has now for many organisations become not just a moral one but also a legal one. Hence it is important that they develop better practices and a more mature approach to web development in order to fulfil it.

To date, very little research has been carried out into how organisations are coping with the added responsibility of the accessibility of their web sites. How organisations react to these changes can provide an important indication of how prepared organisations might be for future developments affecting the publication of web sites. The following research questions are important to answer if an understanding is to be gained of current trends and standards in web accessibility;

- How is accessibility assessed?
- Which organisational traits lead to better accessibility?
- How much awareness of accessibility is there amongst web maintainers?

- What web processes are currently implemented within organisations?
- How can tools support process improvement?
- How accurate are automatically derived accessibility measures?
- Can a centralised visualisation tool provide the necessary information required to support maintenance activities efficiently?

To answer these questions, a study consisting of two surveys: Accessibility Specialists Survey (ASP) and Web Maintenance Practices Survey (WMP Survey). The ASP survey was designed to qualitatively explore the perceptions of accessibility specialists and to use these to design the questions for a second quantitative survey of web maintainers. The WMP survey established current publishing practices amongst organisations and then investigated links between these findings and the accessibility of web sites.

Once the WMP survey was carried out, the results were compared to both manual and automated accessibility assessments of their web sites. There now follows a description of the study, followed by a discussion of the protocol for the Tree-Map tool experiment.

4.3 Study Outline

4.3.1 Introduction

This section describes a study carried out involving two surveys. The first of which, the ASP survey, was to establish what accessibility specialists consider the current state of web accessibility and its best practice. This is then compared to what web maintainers report for their own organisations. To achieve this, two surveys were conducted.

Survey	Type of data collected	Nature
Accessibility Specialist Perceptions Survey (**ASP Survey**)	Qualitative	Small scale but detailed, applicable to a narrow audience.
Web Maintenance Practices (**WMP Survey**)	Quantitative	Requires less in-depth knowledge of accessibility, questions are closed, applicable to a wider audience.

Table 4-1 Surveys used in study

Both surveys were carried out using online questionnaires. Further data was taken from follow-up email correspondences with participants from both surveys.

4.3.2 Clarification of Terms

Because the precise roles of the respondents in the ASP and WMP surveys differ, the general terms of "accessibility specialist" or "specialist", and "web maintainer" or "maintainer" has been used to describe respondents to the ASP and WMP surveys respectively.

Following these surveys, an experiment was conducted to evaluate whether a Tree-Map based tool offers time saving and accuracy benefits to the web development community.

4.3.3 Aims of Study

Table 4-2 lists the aims of the study and how each aim was achieved. To complement the two surveys, information from the literature review was used to determine "best practice". The relationship between web maintenance practices and web accessibility was investigated by both manually assessing web accessibility and also with the use of semi-automatic assessment tools. Details of these tools are given in the literature review. Finally, the design of an evaluation for the Tree-Map based tool will be given.

Aim	Method
Establish how the specialists evaluate and improve web accessibility.	ASP Survey
Establish current web publishing practices.	ASP & WMP Surveys
Establish best web accessibility best practice.	ASP Survey & Literature Review
Establish which best practice is implemented by maintainers.	WMP Survey
Investigate how web practices impact on web accessibility.	WMP Survey and Accessibility Assessment
Evaluate Tree-Map tool prototype efficiency and effectiveness for web accessibility and maintenance tasks	Tool Evaluation

Table 4-2 Aims of Study

Table 4-2 lists the aims of the study in the left hand column and how this work addresses them in the right hand column.

4.4 ASP Survey

The ASP survey addresses the following research questions;

- Which tools and guidelines are currently used by accessibility specialists?
- How much is spent on ensuring web accessibility?
- What are the attitudes of non-specialist web maintainers towards accessibility?
- Why do organisations employ accessibility specialists?
- How are current guidelines and best practice viewed by specialists?

There has been very little research into how those charged with web accessibility carry out assessments, and where necessary, changes to web sites to comply with standards (such as W3C [26] or US Section 508 guidelines [52]). Because the ASP survey targeted a specific group of individuals (i.e. web accessibility specialists, see Table 4-1), achieving a large enough response was difficult. The following approach

was taken; firstly 60 accessibility consultants were contacted. These accessibility specialists were identified using a search engine, and then afterwards their web site was assessed to establish whether they met the criteria expected of an accessibility specialist (see section 4.4.1), they were contacted via a personally written email. Establishing contact in this way was considered to have a better chance of getting a reply and generating a follow-up correspondence.

4.4.1 Accessibility specialist criteria

The web site assessment criteria used for selecting potential specialist was as follows:

1. Was the web site accessible? A genuine accessibility specialist is highly unlikely to have an inaccessible web site. Accessibility was judged using the SiteValet [112] semi automated assessment tool.
2. Were web sites worked on by the specialist available?
3. Were these web sites considered accessible (following a brief viewing and semi-automated assessment)?

If the answer to those 3 points were all yes, then the specialist was contacted.

Following this, invitations to participate were posted on several web accessibility forums. Participants were asked to fill out an online questionnaire. Any personal responses made to the author were also retained and used later as reference material. This second approach proved much less successful.

Due to the nature of the participants an online approach, rather than a paper based questionnaire or interviews was chosen. Online forms were also considered the most efficient and effective way to get cooperation from the most participants. As this survey was explorative, a mainly qualitative approach was selected for its design. The results of which were used to help formulate a better understanding of how accessibility specialists perform their job and also to help structure the design of the questionnaire used in the WMP survey (see below). The ASP survey's questionnaire had 14 questions grouped into 4 sections, there now follows a description of each section and its purpose.

- **Background Details** - Questions were asked to determine what their respondents' role was and how long they had been employed in this role. This gave information about what the kind of specialists had completed the survey.

- **Accessibility Assessment** - Specialists were asked which tools they used to evaluate web pages. The list of tools widely used by the community was provided (see W3C [35]) to minimise the effort required by respondents and so increase the chance of a response. A free-text field was also provided in case the specialists used a tool that did not appear in the list. The accessibility activities performed by specialists (such as assessment or training) were surveyed, as well as the guidelines against which web sites were assessed. By establishing which guidelines and tools are being used to assess web sites it was possible to identify which tools and guidelines were popular amongst specialists.

- **Accessibility Experiences** - This section gauged how much of a web budget was typically allocated to web accessibility projects. It also asked specialists to describe web maintainer perceptions towards web accessibility and why they were hired to address web accessibility by organisations. The aim of this section was to gather data that could be compared with the perceptions of non-specialist web maintainers. It is reasonable to assume that those charged specifically with web accessibility would be aware of maintainer attitudes to their responsibilities and their area of expertise. The questions aimed to determine whether specialists believe the problem is being properly addressed within the organisations or if there is a negative attitude towards accessibility within the organisation.

- **Accessibility Views** – To investigate whether current recommendations for best practice were accepted by the specialists, the final section of the survey asked specialists to explain the following:
  - Whether or not they believed integrating accessibility into the workflow is a good idea.
  - If the advantages of retro-fitting accessibility, (repairing an existing web site without a major re-design), into a web site outweighed the disadvantages.

- Whether they employed any form of measurement or process for tracking changes / improvement.
- If they have recognised any organisational traits that directly impact either positively or negatively on accessibility.

The results from the ASP survey are presented in section 5.2. The WMP survey Two will now be described.

4.5 WMP Survey

The WMP survey addresses the following research questions:

- What impact does content management software have on web accessibility?
- Do more mature general web publishing practices result in better web accessibility?
- To what extent is best practice followed? And what effect has this on accessibility?
- What impact does an organisations approach to 3rd party content accessibility have on the overall accessibility of a web site?
- Does training have a positive influence on accessibility?

There is little value in guidelines and standards unless they are implemented. Hence, it is important to establish to what degree accessibility standards and best practice have been implemented for widely used and service orientated web sites. If these types of web sites are not accessible, it will affect not only a large number of potential users, but also negatively influence other web developers. For instance, popular and high profile web sites (such as those of large organisations with real world customers) will have a large influence on the state of the art in web site design.

4.5.1 Organisations contacted

At the start of the WMP survey, private sector UK companies from the FTSE 100 index were invited to participate. However, after extremely low participation this was strategy was changed. The low participation level may have been due to the fact many private sector organisations are reluctant to reveal internal business processes. This theory was backed up by several private companies responding that due to issues of

privacy or the organisational policy they do not participate in surveys. After this negative response, organisations within the public sector were targeted. Over 500 local government agencies (town and county councils) were contacted via a general email. Due to the large number of organisations contacted, it was infeasible to write individually to each organisation. Email addresses were generated by collecting the domain names of local government agencies and adding either webmaster or postmaster. As with the ASP survey, participants were asked to complete an online questionnaire.

All respondents were assured that responses to the survey would be kept private and informed that any result published would not reveal the names of their organisation. The questionnaire for the WMP survey was longer than that used for the ASP survey, with 35 questions (in contrast to 14 in the ASP survey) grouped into 5 sections. Since the questions were of a closed nature and require few textual responses, it was considered an acceptable length. As an incentive to participate those who responded were offered a manual accessibility assessment of their web site. This was not only useful for encouraging participation, but also for investigating links between best practice within organisations and web accessibility. Results from this manual assessment were later combined with a semi-automated assessment and used to investigate the relationship between organisational practices and web accessibility.

4.5.2 Questionnaire Design

There now follows a description of each section of the questionnaire and its purpose. The sections are as follows:

- **Respondent / Organisation Details** - Respondents were asked to list their position(s) and level of experience within the organisations as well as the number of individuals who worked there. This information is important as it reveals the roles of web maintainers responding and what size of organisation they work in.

- **Web Publishing** - Organisations have adopted different approaches to web publishing. As such, it is necessary to establish the following:
    - The number of individuals involved in the publishing process.

- Whether they have used content management software.
- Whether specific roles and a content publishing model are used.
- If feedback is tracked.

Content management software (CMS) is becoming increasingly popular for larger web sites and so it is important to investigate the percentage of the organisations in the survey that use a CMS. Also, the link between the use of content management software and accessibility will be investigated. The presence of specific roles, a content publishing model and user feedback tracking are an indication that the organisation has a mature approach to web publishing. By isolating organisations with this mature approach a comparison of accessibility can be made.

- **Accessibility Practices** - The W3C recommendations for managing accessibility [33] provide a baseline upon which best practice can be drawn. The issue to be examined is to what extent this best practice is being implemented by organisations. In the largest section of the questionnaire the following questions, based on W3C recommendations, are asked:
  - Does your organisation have a clear and comprehensive policy addressing web accessibility?
  - Have management given genuine backing to web accessibility?
  - Has web accessibility been promoted within your organisation? And if so how?
  - How do you rate awareness of accessibility issues within your organisation?
  - Are accessibility experts or 'gurus' available who can be consulted for assistance?
  - Are guidelines set which govern the accessibility of organisational web pages? (If so, which guidelines are used?)
  - Do Web projects within your organisation have a specific budget for accessibility testing and assessment?
  - Which methods do you use to evaluate web accessibility?
  - Does your organisation actively monitor and record accessibility data for your website?

- Is accessibility integrated into the web publishing process as part of quality control?
- Are there measures in place to monitor the latest web accessibility Recommendations and Guidelines? (And if so, what are they?)
- Are there procedures in place to procure and evaluate the latest accessibility tools?

- **3rd Party Content** - It is important that **all** web content is accessible; many web sites contain pages that are generated by 3rd party software, for example, a section of a web site which is controlled by content management system software. As such, organisations should ensure that these pages are of the same standard as the in those produced in-house. The effect of 3rd party content policies will be examined to determine whether companies with stricter controls on 3rd party content have more accessible web sites. This section determined the following:
  - When selecting 3rd party software for the web, accessibility was considered.
  - If the same guidelines and standards were applied to both in-house and 3rd party generated web pages.

- **Training Practices** - The aim of training is to improve the quality of work, and, hence, accessibility training should raise both awareness and the quality of web pages within the organisation. Questions asked in this section were as follows:
  - How many organisations offer staff accessibility training?
  - To whom are they offering it?

  Legislation and regulations governing accessibility such as SENDA [49] within education organisations in the UK and Section508 [52] in the USA require organisations to be proactive to ensure accessibility. Without training their staff to ensure they conform, organisations are unlikely to meet their legal obligations. It is important to establish, to what extent training improves the accessibility of web sites.

4.6 Accessibility Assessments

In order to identify which practices of the respondents of the WMP survey influenced the accessibility of web sites, two different assessments were conducted. The methods behind these two assessments are now described.

4.6.1 Automated Assessment Method

Three stages were used to assess each organisation's web site using the automated method and are described below.

4.6.1.1 Stage One: Gathering Page Sample

A set of pages from each site was taken using the W3C Web Robot [27]. The following restrictions were placed on Web Robot:

- Pages that were more than three hyperlinks distance from the initial page (homepage) were ignored. This cut off was chosen because of time and resource limitations.
- All web pages were collected, excluding binary attachments such as executable files.
- Only local pages were collected. External pages are not of interest as they are assumed not to be under the maintenance control of the organisation.

4.6.1.2 Stage Two: Evaluate Sample of Web Pages and Measure Accessibility

Once the web robot had finished its search, a subset of the total pages found were evaluated using the automated evaluation tool called SiteValet [112]. As the evaluation process is extremely processor and time intensive only 10% of pages of each site were evaluated. In total 13,660 web pages were assessed. By taking a percentage of the site rather than a fixed size sample it ensured that the proportions of each site was comparable and therefore prevented a potential bias being introduced of smaller sites having a higher sampling rate.

The first step of the analysis process used the SiteValet report to calculate a PageMeasure (see Section 3.2) to provide a quantitative estimate of a page's accessibility. A Site Measure was then generated from the mean of the Page Measures

for each web site. Since the measure is based on the detection of barriers, a page with a higher Page Measure is considered less accessible than one with a lower Page Measure. This rule also follows for Site Measures; web sites with a higher Site Measure are considered less accessible.

4.6.1.3 Step Three: Compare Site Measures with Survey Results

Once the Site Measures for each organisation's web site were calculated, the results were compared with the responses given in the questionnaire for the related web site.

4.6.2 Manual Assessment Method

The manual assessment of accessibility is a time consuming process. As such, it was only possible to carry out an assessment of a small sample of web pages from each web site. Home pages plus nine random pages from each web site were assessed. These web pages may or may not have been used in the previous automated assessment of the web site. Each page assessment involved the following activities:

- Reviewing Watchfire and SiteValet reports.
- Reviewing the Automated Assessment page measure.
- Cross browser viewing of the page using the following browsers:
    - Lynx (text-only browser)
    - Internet Explorer
    - Mozilla Firefox
    - Opera

4.6.2.1 Assessment Report

For each web site a report was produced which was then sent to the respondent. The following points were included in the report.

- **Overall Rating:** a categorisation of the web site based on the pages sampled was phrased carefully to ensure that respondents did not feel it was overly critical or discouraging. The categories used differed from Lazar [66]. Since the standard of accessibility was generally fairly good, it was felt that the categories should differentiate between web sites at the higher end of the accessibility scale. Also, whereas Lazar uses only one category called *Accessible*, to classify accessible web sites, this was considered too broad.

Therefore this category was expanded to include a range of accessibility categories. The **Issues** and **Manual Sample** sections (see below) of the report contained specific details of inaccessibility.

  - Very Accessible
  - Accessible
  - Requires Attention
  - Inaccessible

- **Issues:** In this section general accessibility barriers were listed. These are general and specific problems which should be addressed to make the web site more accessible.
- **Manual Sample:** This section reported specific problems, where applicable, on a page by page basis, including: cross browser compatibility display problems and areas of false or deprecated HTML that needed corrective maintenance.
- **General comments:** This section included results of the site's semi-automated assessment and included specific points of interest detected during the manual assessment of the sample pages.
- **Positive Points:** It was felt that the assessment should not only criticise where mistakes / errors were present, but that it should also offer praise for well designed pages and good accessibility. Thus in this section positive attributes discovered during the assessment were highlighted.

There now follows the protocol for the Tree-Map experiment.

4.7 Tree-Map Experiment

The purpose of the Tree-Map experiment was to establish whether or not Tree-Maps could be used as a base to provide an efficient and accurate web maintenance tool. There are currently few web maintenance tools which have been designed to full address web accessibility and as such this experiment aims to provide justification for the development of such a tool.

4.7.1 Research Question for this study

The study attempts to answer the question:

Does a Tree-Map based tool make the retrieval of accurate web maintenance information quicker?

The investigation will not only test whether or not participants complete information retrieval tasks in less time, but also whether the information retrieved is accurate.

4.8 Experimental Design

This section outlines how the experiment was carried out. Participants are asked to complete a set of maintenance related information retrieval tasks using a Tree-Map tool and HTML based reports.

4.8.1 Web site selection

Two live web sites will be selected. Using real sites provides a better test for the practicability of a Tree-Map based tool. The Tree-Maps visualise a partial view of the web sites; this was to ensure participant's tasks were manageable and also saved a considerable amount of time when setting up the experiments. In order to minimise any learning effects on task performance, participants will perform two different sets of tasks on two different web sites. However, the tasks chosen for each web site should be very similar. The following attributes will be required for the two web sites used:

- Similar size.
- Similar nature / purpose (in this case local government web sites will be selected).
- Similar provision of meta-data. This allows the tasks set to be similar. The data provided was as follows:
    - Creation date of a web page.
    - Contributor / Creator or Maintainer of a web page.
    - Last modified date.
    - Keywords relating to the content of a page.
    - Description or title of a web page.

4.8.2 Sample Population

Since this experiment focuses on information retrieval using tools rather than analysis or verification of the SiteValet's assessment, the population does not need any

specialist web knowledge. The only skill prerequisite is a reasonable degree of computer literacy. Hence, participants were selected from individuals easily accessible to the author. Those selected were:

- Students (both Computer Science and Earth Sciences)
- Researchers

There are several important issues to be considered, they are as follows:

- Colour-blindness.
- Participant familiarity with Windows File Explorer.

To encourage participation, a muffin / chocolate bar and tea / coffee was offered.

4.8.2.1 Colour Blindness

Unfortunately, participants with colour blindness must be excluded from this study. Screening for colour blindness will be done using the Ishihara Test for Colour Blindness. The irony of including this accessibility barrier is not lost on the author. However, making adjustments required can compensate for the effects of colour blindness. For example, changing the range of colours used in the Tree-Map or switching to a pattern based rather than solid colouring of nodes adds an extra variable and therefore is not included in this experiment.

4.8.2.2 Windows File Explorer

Working without the visualisation tool is expected to be more time consuming. Participants must open and read individual reports to gain information about a specific web page. In order to compensate for this lack of global perspective, participants must be aware of the search and sorting functions provided within Windows File Explorer. To ensure that participants are able to perform basic search and navigation, training will be provided that enables the participants to complete the tasks set.

4.8.3 Selection and Preparation of Web sites

To ensure that fair comparisons can be made, web sites from the same domain were chosen. Actual web sites, rather than artificially generated ones were used as the basis for this study. The sites were chosen from UK local government organisations. It was

also important to choose sites with similar, meaningful meta-data and a definite hierarchy; this enables more interesting Tree-Maps and maintenance tasks to be set during the experiment.

4.8.4 Presentation of Maintenance Information

The two groups of tasks will be completed using HTML based reports and Tree-Maps, each uses information generated from the SiteValet tool.

4.8.4.1 HTML based Reports

The HTML based reports are now described.

URL	http://www.hillingdon.gov.uk/jobs/index.php
Assessed by	John Bailey
Guidelines	WCAG3
Date	Wed Jun 14 14:49:42 2006
Level 1 Result	Fail

Figure 4-1 Screen shot of SiteValet report header

Figure 4-1 shows the header of the HTML report produced by SiteValet. There are five fields; reading from the top to bottom:

- The URL of the web page which was assessed. This is also a hyperlink to the URL shown.
- The name of the individual who ran the assessment (with an email hyperlink)
- The Guidelines against which the page was assessed. In this case it was the WAI guidelines including checkpoints from all 3 priorities.
- The date the assessment took place.
- Summary pass/fail based only on the automated assessment.

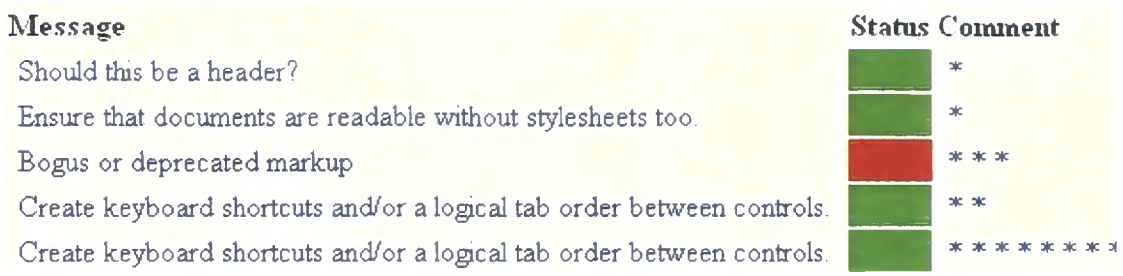

Figure 4-2 Screen shot of the barriers found by SiteValet

Figure 4-2 shows the main reporting element of the HTML report. The report contains a summary table and also the full HTML source of the page which has been assessed. The HTML source is syntax highlighted with the addition of barrier information. Specific barriers are reachable from the third column in the summary table.
The three columns of the summary table are as follows;

- **Message** - contains natural language descriptions of the type of barrier found by SiteValet. This is hyperlinked to the relevant section of the W3C guidelines for further information
- **Status** – shows a colour representing the level of confidence the tool has that the barrier is an actual barrier. Status uses a Green-Red scaling. Bright green represents a very low confidence rating (which reflects the subjective nature of a particular barrier) and deep red represents a barrier which the tool was able to confirm automatically as a certain barrier. For example, where bogus HTML has been used.
- **Comment** – contains internal hyperlinks to the HTML source of the assessed page where instances of the barriers described in the message column are located. This allows the assessor to click directly to the potential barriers HTML and verify the barrier manually.

4.8.4.2 Tree-Map Tool

Tree-Maps were generated for two web sites based on the hierarchical structure of each site. The Tree-Map is split into a series of rectangular nodes. Each node in the Tree-Map represents a web page; each node's size is determined by the size of that page and the amount of space available in the Tree-Map view. Its colour is determined by the Page Measure of the web page. This version of the Tree-Map tool was generated from the original Java based application (see Chapter 3) using a

combination of HTML, images, CSS and JavaScript. This tool had less functionality than the Java application but was more portable and the simplicity prevented the need for extensive training. Participants could move up and down the hierarchy with the Tree-Map visualisation adjusting the focus to specific sections of the web site. This allowed them to get a more detailed view of pages contained in sections further down the hierarchy. This made comparing the size of nodes (page) within the same section of a web site easier.

For the experiment, users have access to a variety of meta-data and functionality, this is now described.

4.8.4.3 Page specific meta-data and functions

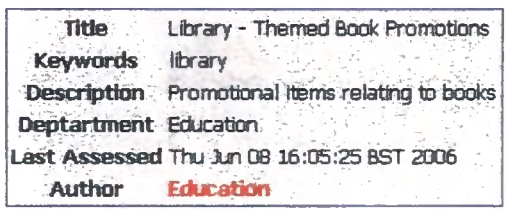

Figure 4-3 Meta-data for each Page in Tree-Map Tool

The meta-data provided to the user was as follows:

- **Title** – the title of the web page (defined in the web page's HTML).
- **Keywords** – list of words provided by a content author which should relate to the web page's content. These are often used to aid web search engines.
- **Description** – a description of the web page's content
- **Department** – the name of the department responsible for creating the web page.
- **Last assessed** – the date the page was last assessed for web accessibility.
- **Author** – an identifier of who created the web page.

Figure 4-3 shows how the information listed above was presented to the user. This information was taken directly from the web page's HTML, so whether or not it is available or accurate is dependent on whether the content author / coordinator has entered the data. Due to limited screen space and the diversity and amount of meta-data associated with some pages, only a limited number of meta-data items were selected. To ensure that participants were not confused some standardisation was

required. For example, one site used the meta-tag "content author" and the other "content creator" these represented the same information and were displayed as "Author" in the Tree-Map. Keywords and Description are commonly used meta-tags and appeared in the specification of HTML 2.0 [24]. As well as the meta-data shown in Figure 4-3, the URL of the web page represented by a node was displayed at the bottom of the screen. This changed as the participant moved the mouse over nodes.

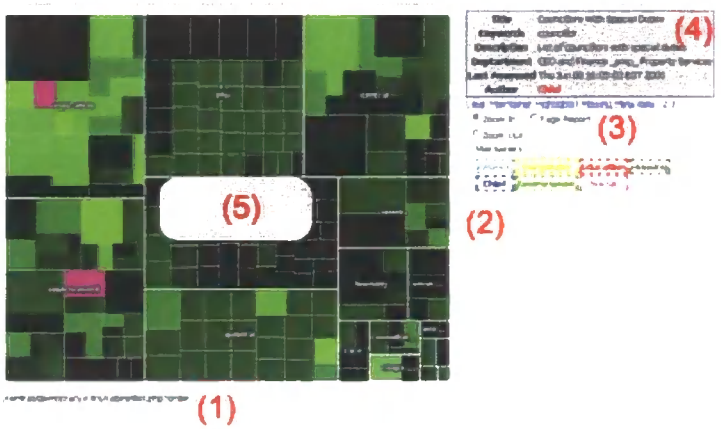

Figure 4-4 Screenshot of Tree-Map based tool

Figure 4-4 is a screenshot of the Tree-Map based tool used in the experiment. The following sections have been numbered:

1. URL of web page represented by each Tree-Map node.
2. List of content authors / web maintainers – participants click on these to highlight which pages they created / maintain. Each content author was displayed using a specific colour; pages which they created were highlighted using a dashed border in the same colour.
3. Zoom functions and page report – participants select a function and then click on a node of interest to either zoom in or out or view the SiteValet report generated for the web page. Additional functions were:
    a. Highlighting pages which had meta-data missing.
    b. Clearing all highlights from the Tree-Map.
4. Meta-data generated from the web page (see Figure 4-3).
5. Tree-Map of the web site or section of the web site (depending of the level of zoom).

4.9 Data preparation and Collection

This section describes how the accessibility and structural data used in the experiment was collected and prepared in order to run the experiment.

4.9.1 Data preparation

Both the Tree-Map and HTML reports had to be prepared for the experiment.

4.9.1.1 Tree-Maps

Pre-processing of the web sites was required to reduce the overhead of the accessibility assessment tool and to simplify the structure. Firstly all non-HTML based web pages were removed (XML files, PDFs, Microsoft Office documents, and other binary pages). SiteValet can only evaluate HTML pages and so these non-HTML pages were not used in the study. Following this, all "cloned" web pages were removed. Cloned pages can be defined here as pages with the same URL but, through a variable passed to the web page, varied presentational features such as colour scheme / font size. There was no difference, however, in the information provided on each "clone". Since these modifications are generated by software rather than created manually; problems should be detected in the software generating the clones, rather than manually on a page by page basis.

The two web sites selected are:

A. www.hillingdon.gov.uk
B. www.chichester.gov.uk

4.9.1.2 HTML Reports

Each report was generated into a directory structure that directly mirrored the structure of the URL of the page. This simplified the process of referencing page reports from the Tree-Map tool and removed the need for a lookup table. For example, the report for the web page http://www.website.ac.uk/department/postgraduate/fees.html will be stored in the following directory c:\www.website.ac.uk\department\postgraduate\fees.html.

4.9.2 Data Collection

Data collection was achieved in individual experimental sessions under controlled conditions in a laboratory. Since one of the dependent variables measured was task

completion time it is important to remove all possible distractions. After consultation with Professors David Budgen (Durham University) and Professor Barbara Kitchenham (Keele University) and Dr. Alfred Kobsa (University of California, Irvine) a sample size of 24 has been chosen. Each group therefore had 6 members. Table 4-3 shows the specific details of which group performed which tasks and crucially in which order. The first column shows the group number, the next two columns show the combination of tool and web site for the Task One and Task Two.

Groups	Task One (A)	Task Two (B)	No. of Participants
1	Treemap, Site A	HTML, Site B	6
2	HTML, Site A	Treemap, Site B	6
3	Treemap, Site B	HTML, Site A	6
4	HTML, Site B	Treemap, Site A	6
		Total	**24**

Table 4-3 Details of which group performed which task

4.9.3 Tasks

The tasks were similar simple information retrieval activities for both sites; the only difference was the site specific data. The tasks are listed below:

- Identify the author of a specific web page.
- Report specific meta-data (description / title) for a specific web page.
- Count the number of web pages contained within sections of a web site.
- Identify the number of web pages maintained by a maintainer.
- Find out when a web page was last assessed for accessibility.
- Locate the largest web page in a certain section of the web site.
- List the potential accessibility problems for a specific web page
- Compare two web pages and estimate which is the least accessible.

4.9.4 Step by Step Plan

The plan for the experiment was as follows:

1. At the beginning of each individual session, each participant was assigned a random participant identifier. This makes blind analysis possible; and reduces the chance of bias based of specific individuals during the analysis of results. The use of an incremental candidate numbering system was considered but

this might have allowed the experimenter to work out when the participant carried out the experiment and make it easier to recall their assigned group. Each participant was assigned a different group from the previous participant in sequential order. The group (1 - 4) for the first participant was chosen at random. The tasks were completed in the following order:

   a. The participant was screened for colour blindness using the Ishihara Test for Colour Blindness. This was done using an on screen display of the Ishihara Tests taken from http://www.toledo-bend.com/colorblind/Ishihara.html.
   b. The participant was given two task sheets, one for Sites A and B and asked to write their participant identifier on each sheet.

2. Approx. 20 minutes of training will be provided covering:
   a. The concept behind the Tree-Map and how the nodes are drawn.
   b. A general description of web accessibility with specific examples of barriers given. The specific examples were:
      i. Images missing alternate text and the problems they present to non-sighted or low bandwidth users.
      ii. Ensuring clear and simple documents, highlighting the difficulties of creating such documents and of automated assessment of this guideline.
   c. How to work with the Tree-Map based tool - a description of the tool's functionality with a demonstration of usage.
   d. How meta-data is stored in HTML pages and how to find this information using both Tree-Map based tools and the HTML based reports was explained and demonstrated to participants.
   e. How to search for specific meta-data within the HTML based reports using Windows Explorer and the search options provided.
3. Any questions raised by the participant were answered.
4. The participant was asked to begin Task One, and was instructed not to spend too much time on individual question (bearing in mind the 20 minute time limit). Participants were told that if they felt a question was infeasible after investigation to mark "n/a" and move onto the next question. The experimenter timed each task.

5. Originally, a distracter activity was designed to reduce the learning effect carried over between each task, however after considering timing and the potential for learning effect, this was abandoned.
6. Task Two was carried out.
7. The task sheets were collected from the participant and they will be given their thank-you muffin / chocolate bar as a reward.

4.10 Analysis

This section outlines the analysis carried out on the data collected. The purpose of the experiment was to test whether a Tree-Map based tool provided any advantages over text based accessibility assessment tools. To test this theory, the null hypothesis tests whether there is no significant different between how long it takes to retrieve accurate web maintenance information using the Tree-Map based tool and the HTML report based tool. The alternate hypothesis asserts that the Tree-Map based tool will enable the retrieval of maintenance information in less time that the HTML report based tool.

4.10.1 Hypotheses

- **Null Hypothesis (H_0):**
  - A Tree-Map based tool does not make the retrieval of accurate web maintenance information quicker than HTML based tool.
- **Alternative Hypothesis (H_a):**
  - A Tree-Map based tool does make the retrieval of accurate web maintenance information quicker than HTML based reporting

4.10.2 Validity

In order to ensure the experiment was as fair and unbiased as possible the following measures were taken.

4.10.2.1 Cross over design

For a within-group (AB/BA) cross over design there were two treatments, A and B; each participant undergoes treatment A and treatment B and this reduces the risk that variability between participants has affected the results. To address the carry-over effect (i.e. that undergoing the first treatment affects the results for the second

treatment), half the participants undergo treatment A first, followed by treatment B and half undergo treatment B first, and then treatment A.

Specifically for the Tree-Map experiment, the advantage of using a within-group, cross over design is that the effect of using the Tree-Map tool is judged on an individual basis, each participant acts as his/her own control. Each participant carries out a task using both the Tree-Map based tool and the HTML report based tool. By using 4 groups in an AB/BA cross over design [94] (see Table 4-3) it is possible to balance the risk that a web site might be better suited to use with a particular tool, since half the participants will use the Tree-Map tool with Site A and half with Site B, and also that the order in which a task is performed has an effect. Using different web sites for each task ensures that results are not affected by any learning effect. Information learned about the web site in Task One could not assist in the completion of Task Two. By selecting web sites which had very similar characteristics (see section 4.8.3), this variable was not expected to have a significant effect on the results. The order in which the participants carried out the tasks could also have had an effect on task competition time and accuracy. To test whether task order had an effect on timings an ANOVA analysis of the differences in the timings for the four groups was carried out. Once the effects of order on timing has been analysed the significance of the timing differences were analysed using statistical t-tests.

4.10.2.2 Blind Analysis

Each participant will be given an anonymous identifier to write on their answer sheets. This ensures that the experimenter cannot know to which group a participant was assigned and guards against any potential bias in the analysis.

4.10.3 Limitations of the Study

Limited resources meant that it was necessary to keep the experiment as short as possible, as the likelihood that a person will volunteer is greatly influenced by the amount of time they are required to commit. To ensure this, each task was given a maximum completion time of 20 minutes. If the participant fails to answer all the questions for a task within 20 minutes, they were asked to stop and it is noted. Any tasks left incomplete were marked as (n/a) or "not able to complete within timescale".

The use of Windows File Explorer to interact with the HTML reports was used as an alternative to the Tree-Map based tool. It provides appropriate hierarchical views to the participants alongside text based searching and provides a graphical user interface.

Because the author of the tool was also the experimenter this may have introduced a bias based known as the "Hawthorne Effect". Participants may have attempted to please or help the experimenter by increasing effort during their performance of the Tree-Map based tasks. To counter this however, participants were assured that this was supposed to be an exploration of both tools and they should perform both tasks to the best of their abilities.

4.10.4 Deviations from Protocol

Unfortunately it was not possible to record the time taken for each individual question, preliminary runs of the experiment showed that tasks were not always completed in a linear fashion and it was very difficult to determine when a task had been completed. Participants were asked to inform the experimenter when each task was completed, but often forgot to do so.

As a result the 5 minute limit per task proved meaningless, instead participants were asked to attempt each task, but if they felt a task was not feasible given the overall time limit of 20 minutes to mark it as n/a ("not able to complete within timescale") and move on.

A laboratory was reserved to conduct the experiments. However there were several instances when participants were unable to make appointments during the time booked and the experiment was moved to another location. The exact same equipment was used in all cases.

During one experiment a fire alarm went off. Fortunately, this was during the introduction rather than the experiment itself. One participant was colour blind, however their particular type of colour blindness did not prevent them from differentiating between the colours used in the experiment.

To save time, the distracter task was removed. It was deemed that the time taken to change from the first to the second period with preparation and questions provided enough of a distraction for the participant.

The reward was made more flexible as many participants expressed a fondness for doughnuts and cheese cake rather than muffins!

4.11 Summary

This chapter has outlined the study containing the following surveys:

- ASP Survey – open questionned survey aimed at collecting the experiences and opinions of accessibility specialists. A description is given of how these specialists were selected, vetted and what was asked of them.
- WMP Survey – quantitative survey looking at which best practice was performed within organisations. It also tested whether the opinions of the specialists about non-specialist web maintainers were accurate. Details are also given of how the organisations were selected and of the worked carried out by the author in assessing their web sites. The purpose of this assessment was to compare best practice with accessibility.

Following the surveys a protocol for a controlled laboratory experiment is given. The aim of the experiment is to determine whether the Tree-Map based tool supports more efficient and accurate web maintenance. The protocol describes in detail the design of the experiment including the following:

- The selection and preparation of web sites,
- Participant selection and data collection

The hypotheses of greater accuracy and efficiency used and the statistics which will be applied were also given.

5 Study and Experiment Results

5.1 Introduction

Chapter 4 outlined the study and related experiment. The results from the study and the Tree-Map tool experiment are now presented. Following the introduction, information about respondents from the ASP Survey and then the WMP Survey is given. In addition, data from the WMP Survey about web practices is presented. After this, analysis of their responses to accessibility questions is also presented, and where possible, compared. For example, the awareness of accessibility was one point of interest that was examined. In both the ASP and WMP surveys the responses for each question are discussed using the following format:

- **Question Statement(s)** – gives the actual text of the question used in the survey.
- **Rationale for question** – describes the reasoning for asking the question.
- **Question response(s)** – describes the results and responses for the question(s).

5.2 Respondents Background: ASP Survey

The results discussed here reference section 4.4

The ASP survey was devised to explore the experiences and views of accessibility experts. Results from this survey were then used to design the WMP survey. The ASP survey was a small, mainly qualitative survey aimed at exploring the experiences and perceptions of individuals who specialised in web accessibility. The following sections now report the results of this survey.

5.2.1 Country of Residence

65 accessibility specialists were contacted to complete the questionnaire. 20 completed the questionnaire. It should be noted that there was a diverse range of countries represented, although most (90%) were from English speaking countries.

Country	Respondents
USA	4
UK	5
Canada	5
Australia	2
New Zealand	1
Ireland	1
Germany	1
Denmark	1
Total	20

Table 5-1 Country of Residence

5.2.2 Experience

Question Statements - Approximately how long have you been working in web development? Of this how long have you been working with web accessibility?

Rationale for question – It was important to ensure that the specialists were experienced enough to contribute experiences and findings to the survey. A person who had just started working with the web or in web accessibility will be unable to offer as much as someone with years of experience.

Question Response - The specialists had on average just less than 7 years of experience working within web development and just over 5 years of that working with web accessibility, both had a standard deviation of approximately 2.4 years (see Table 5-1 and Figure 5-1).

	Years involved in the Web	Years involved in Accessibility
MEAN	6.76	5.14
STDEV	2.4	2.43

Table 5-2 Years of Experience working with the Web and Accessibility

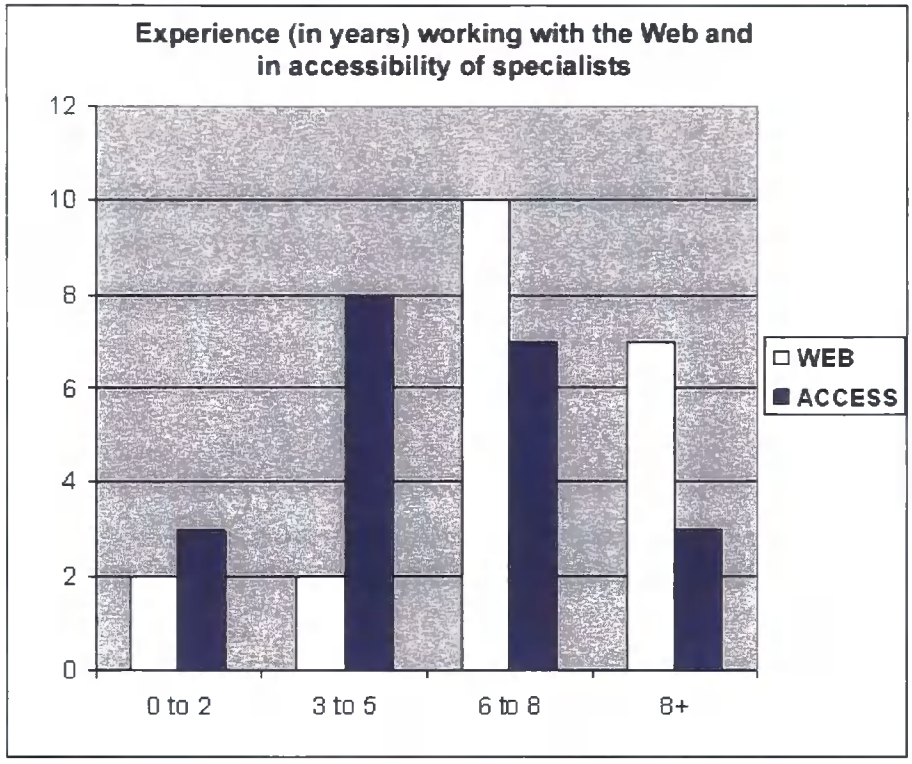

Figure 5-1 Experience in years of specialists

5.2.3 Respondents Web Accessibility Roles

Question Statements - What role(s) do you have in relation to web accessibility? And what are your main activities within web accessibility work?

Rationale for question – Accessibility specialists can take on a number of roles and it was important to gauge the typical roles carried out by the specialists surveyed. The second question surveyed the main activities performed by the specialists.

Question Response - It is clear from Table 5-3 that web accessibility consultancy was by far the most popular role amongst the specialists. 95% of the respondents acted as accessibility consultants. The one respondent who did not list consultancy had previously worked as a consultant. The most popular activities performed (see Table 5-4) were assessment, training and raising awareness. This would seem logical as a consultant would normally have to assess a site for accessibility (assessment), and where needed raise awareness of the problem to web maintainers and then provide training to members of the organisation.

Role	Percentage of specialists

	who performed this role
Web Programmer	55%
Web Interface Designer	45%
Web Content Provider	20%
Web accessibility Consultant	95%
Web accessibility Researcher	30%

Table 5-3 Roles of respondents and the percentage who carried them out in Survey One

Table 5-3 contains the roles performed by the specialists. The Role column gives the name of the role and the right hand column shows the percentage of respondents who listed this as one of their roles.

Activity	Percentage of specialists who performed each task
Assessment	95%
Repair / Barrier Removal	45%
Raising Awareness	80%
Training	85%
Process Improvement / Management	65%
Content Creation	35%

Table 5-4 Activities carried out by specialists

Table 5-4 shows accessibility related activities and the percentage of specialists which performed them.

5.2.4 Guidelines and Tools Used

Question Statements - Which guidelines / standards do you use as reference? And which tools do you use to evaluate web accessibility?

Rationale for question – Which standards are used to evaluate web accessibility and the tools used is important to reveal trends within the specialist community. It can reveal whether certain guidelines or standards are more popular and which tools are being used by the community.

Question Response – 90% of the specialists referred to version 1.0 of the W3C Accessibility Guidelines [26]. This would be expected because of the wide acceptance of the W3C within the international web development community and that version 1.0 is the official version (version 2.0 is still draft). Just over half of the specialists stated they used either Section 508 [52] or W3C version 2.0 [36]. The results showed that

the most popular choice of tools was a combination of stand alone (that is tools which are executed on the user's machine) and online tools (provided via a web interface) (58%). 42% of specialists used exclusively standalone assessment tools. Online tools are accessed via the internet, for these types of tools accessibility assessment is carried out on the server side and the results are then returned remotely to the user. The results of the ASP survey highlighted that none of the specialists used online tools exclusively.

Guidelines	Percentage of specialists who used each guideline
W3C Accessibility Guidelines 1.0	90%
W3C Accessibility Guidelines 2.0	55%
Section 508	55%
IBM Web Development Guidelines	15%

Table 5-5 Guidelines used by specialists

Table 5-5 shows the guidelines referenced by the specialists in their work.

Tool Type	Usage %
Stand alone	**42%**
Online	**0%**
Both	**58%**
Disabled Users	**n/a**
Disability Simulation	**n/a**

Table 5-6 Tools used by specialists

Table 5-6 shows the outcome of the ASP survey analysis of the tools used by specialists to assess web sites. Statistics for assessment through disability simulation software or disabled users are not available because these were only considered as a result of feedback with one of the specialists.

5.2.5 Accessibility Issues

The ASP survey attempted to determine how the specialists perceived the attitudes and levels of awareness of organisations and web maintainers.

Question Statements - Initially, in the organisations you have worked in, was there awareness of web accessibility? Followed by a series of options (Non-Existent, Little, Acceptable, Good and Excellent)

Rationale for question – The purpose of this question was to survey how specialists perceived the initial levels of accessibility awareness within organisations.

Question Response - A large majority of specialists reported the initial level of accessibility awareness (see Figure 5-2) within the organisations they had worked with / in as either little or non-existent. Based on this assessment of maintainer awareness, it appears that the specialists believe that many organisations are ignorant of web accessibility and hence more work in promotion is required. Even those who offered more positive experiences in the survey qualified their responses with additional statements such as "amongst standard based designers", (i.e. web page designers who follow W3C recommendations) and "it's getting better". Furthermore, one specialist after stating that most of their experiences were with maintainers who were "passionate" about web accessibility went on to add "but I think that is unusual".

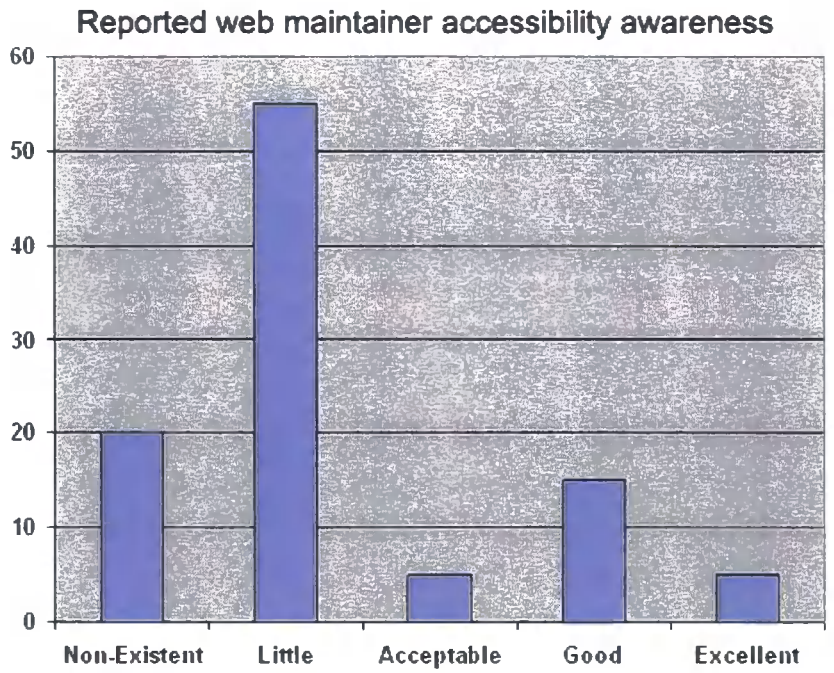

Figure 5-2 Reported web maintainer initial accessibility awareness upon arrival

Figure 5-2 shows the perceived initial level of accessibility awareness within organisations the specialists worked.

Question Statements - Why were you really hired / charged with accessibility evaluation?

Rationale for question - This question investigates the motivation behind organisations' realisation that web accessibility has to be addressed. Respondents could select more than one motivation from a list of options, a free text box for other motivations was also provided.

Question Response - As can be seen in Figure 5-3 changes in the law are the most significant factor for organisations. Aside from this, internal guidelines (which may well be linked to changes, internal branding changes or indirectly to legal requirements) were the other main motivating factor for change. One interesting result was that very few respondents believed they were hired in response to customer feedback. Specialists reported that organisational accessibility awareness was low, this combined with a lack of customer feedback about accessibility indicates a low awareness of the problem in the general public. Within the free text "other, please describe" field, respondents expressed that although all five options were factors in motivating organisations, but the main motivation was to meet legal requirements (75%).

Figure 5-3 Why accessibility specialists believed they were hired.

Figure 5-3 shows the reasons why specialists believed they were hired by organisations to perform accessibility assessments.

5.3 Discussion of Qualitative Data

The remaining questions were open ended and required longer descriptive answers. These questions explored the viewpoints of the specialist in respect to best practice recommended by organisations such as the W3C and the RNIB other and whether such best practice is widely accepted within organisations and especially by the web maintainers. For each question the rationale behind the it is given, followed by the question statement itself and finally the response will be examined. The number of questions is to aid referencing rather than a reflection of the order they were presented to specialists.

Question 1: In your experience, how is accessibility perceived by web designers?
Rationale for question – The attitudes of those working within web page design is crucial to the implementation of the pages. If accessibility is not valued or seen as unnecessary then it is less likely that web pages will be designed with accessibility in mind.
Question Response - Most specialists (65%) reported that web designers had expressed or held a negative opinion on accessibility. Terms such as "annoyance", and "roadblock" were used frequently. The respondents felt that many web designers were unmotivated by the social factors of accessible design and see it as an imposition of benefit to a small group of users. 20% of the respondents reflected an improvement in the overall accessibility perceptions amongst web designers. Finally, 15% of the respondents felt that web designers they had worked with had positive perceptions of web accessibility. Although one respondent added that this was probably unusual. Overall, there was a general agreement that more experienced and competent web designers were more positive about web accessibility.

Question 2: Do the advantages of integrating accessibility into the workflow outweigh the disadvantages? If so, explain why.
Rationale for question – Integrating accessibility into the web publishing model ensures that all content published on a web site is subject to the accessibility checks and guidelines used by an organisation. Establishing this may increase the publication

delay and costs of publishing web content. But managers must consider whether these disadvantages are outweighed by the benefits of producing more accessible web content.

Question Response - The answer to this question was a unanimous and clear "yes" (100%). The points raised were:

- Saves time and money overall, by reducing the need for "rewrites" or modifications.
- Accessibility should be seen as part of an overall quality issue, and workflow should ensure better quality code.
- Better quality code should equate to better web sites and lower maintenance costs.
- High enough level of accessibility awareness amongst developers.

Rationale for question – Retro-fitting in this case can be considered as corrective or perfective maintenance, as the web site might not work for some users and hence the change is corrective. For example, if the web site navigation menu relies on a specific web browser than this is an accessibility barrier to other browser users. Corrective maintenance is applied by changing the navigation menu to remove the reliance on a specific web browser. For other users the site might work, but not very efficiently and hence the "retro-fit" is perfective. Retro-fitting might be considered as it does not require a total re-design of the web site and can be selectively applied to the most popular or visible sections of a web site.

Question 3: Do the advantages of retro-fitting accessibility outweigh the disadvantages?

Question Response - The results show a mixed response to this question. The points against retro-fitting were:

- Many specialists highlighted the fact that many web sites are too large to completely renovate.
- Even if content management system templates are fixed, maintainers can either break these templates, or insert other accessibility errors.

The points in favour for retro-fitting were:

- Where the problem is small enough to be correct simply (the example of a poorly designed form was given).
- Where time and/or financial budget do not allow a complete re-think or re-design of the web site.

Question 4: Do you measure accessibility? If yes, please describe your criteria for measuring accessibility and, please briefly describe how you measure accessibility (i.e. tools, techniques).

Question Rationale - Determining whether or not specialists measure accessibility and the varying methods used is important. This information informs any recommendations for a common measurement method. One point to investigate is whether or not a common method of accessibility measurement can be agreed upon to allow comparison across projects, independent of whoever carried out the assessments. This would allow organisations to monitor improvement even if the original accessibility specialist is unavailable or too expensive.

Question Response - Measures used by respondents were as follows:

- W3C accessibility guideline checkpoints (various priorities). This was the most used method for measurement. Some respondents expressed that they disagreed strongly with certain checkpoints and hence ignored them.
- Section 508 checkpoints.
- http://www.bitvtest.de guidelines.
- General user feedback.
- RNIB See It Right checkpoints.

The second part of the question asked for more detail about tools and techniques a number of methods were uncovered:

- Expert Review.
- Task completion trials.
- Disabled user testing.
- Semi-automated tools (see Table 5-6).
- Cross browser testing – i.e. operating the site in many different browsers.

Two specialists stated they didn't measure accessibility. The reasons given were:

- They had never been asked to by a client.
- Accessibility is too subjective to measure and compare.

Question 5: Do you track how changes made to the site have affected the site's accessibility? If yes, please describe how you track these changes. And, please describe how this change tracking data is employed.

Question Rationale - Without a method for tracking improvement, it is impossible for organisations to know whether the work carried out has been effective. Such a method could make use of the measurements described in Question 4. By comparing current measurements with those taken previously, the organisation can estimate whether improvement has been made.

Question Response - Surprisingly, despite most of the specialists stating they measured the accessibility of their sites only 4 specialists stated this formed part of a formal measurement process including comparison with previous measurements. Several stated that they would only track improvements if asked by the client. Very little was said about the how data was used. However, the following was reported:

- Regressions in accessibility are reported to management.
- Data is mostly used by management to justify accessibility budgeting.
- Data is used to monitor how much improvement has been achieved.

Reasons for not tracking changes were as follows:

- Lack of any tool available.
- Lack of money / time to perform tracking
- Outside the remit of their work.

Question 6: - Have you recognised organisational traits that influence accessibility in a positive / negative way? If so, what are these organisational traits?

Question Rationale - The specialists were questioned as to their impression of crucial traits possessed by an organisation which affects accessibility. The outcome of this question would be used to help design the questions for organisations involved in the WMP survey. The WMP survey can be used to establish whether such traits exist in

general and whether there is an actual link between their presence and improvements in accessibility.

Question Response - This question gained the most detailed answers with varied observations and opinions expressed. These observations and opinions are now summarised. Specialists identified the following organisational traits:

- Knowledgeable web developers.
- Accessibility integrated into organisational culture and web policy.
- Openness to learn and exchange best practice with other organisations.
- Proactive adaptation to new legislation.
- The "Level of exposure and understanding to people with disabilities".
- Availability / presence of accessibility "champions".
- Appreciation of the value to business of accessibility.

Amongst the positive traits described by specialists were also traits that impact negatively on accessibility. These included:

- Ignorance of social responsibility.
- Singular focus on marketing and highly graphical designs.

5.4 Respondents Background: WMP Survey

The results discussed relate to section 4.5

The WMP survey had a wider, more general audience. Therefore a mass-email approach to contacting respondents was chosen. This approach yielded a much higher response. In total there were 84 responses, of which 79 successfully completed the online questionnaire. There now follows the results of that survey. After all the results from the WMP survey are presented, the results of the accessibility assessment of the respondent's web sites are given.

5.4.1 Type of Organisation

Different types of organisations have different uses for their web sites. Hence, the type of organisation has an important influence on the design and structure of the web site. Within the WMP survey the vast majority of respondents came from UK local government organisations. This was due to the much higher response rate amongst these organisations. For further details see section 4.5.1. Web sites of local government organisations contain large amounts of textual information and due to the growth in e-governance must be timely, accurate and offer a high degree of interactivity.

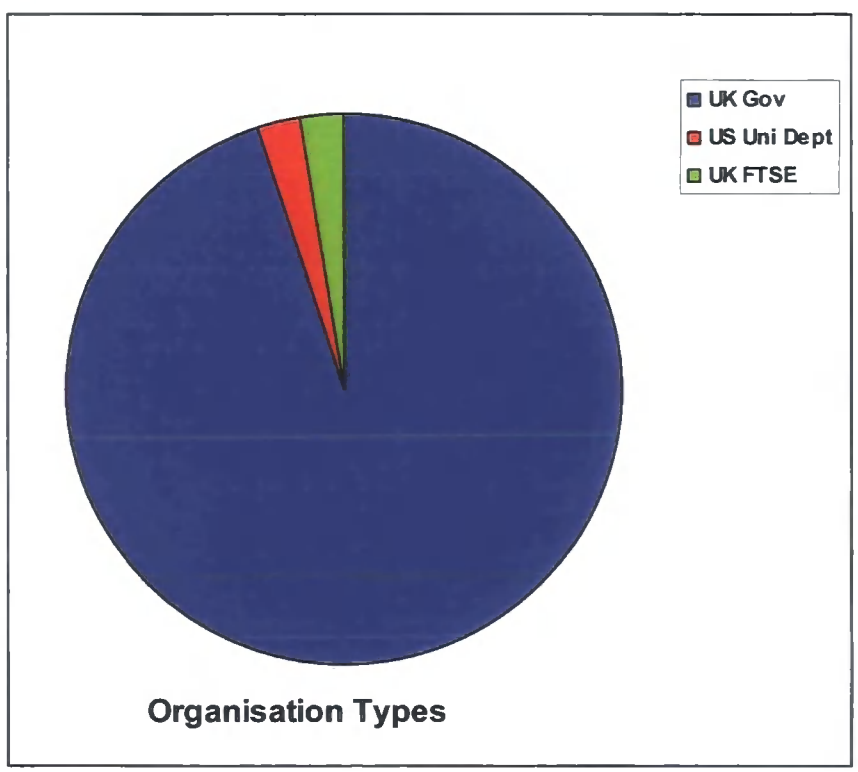

Figure 5-4 Organisation Types

Figure 5-4 shows that 94% of the maintainers worked for UK government organisations, the remaining 6% are split equally amongst UK FTSE companies and US university departments. As with the specialist questions the question numbering is used to aid reference rather than representing order.

5.4.2 Respondents Roles

Question 1: - Which best describes your current position?

Rationale for question - As in the ASP survey, the respondents gave the roles / web responsibilities. Since the survey targeted a more general web maintainer audience, it was natural that more roles are carried out by the respondents and hence there were more choices presented to respondents.

Question Response - Table 5-7 shows that most respondents had more than one role in their organisations. The most popular role was as a content developer (56%).

Role	%
Project Management	47.44
Content Expertise	44.87
Accessibility Testing	44.87
Information Architect	38.46
Information Design	20.51
User Interface Design	35.9
Content Development	56.41
Programming	46.15
Graphical Design	33.33

Table 5-7 Roles of respondents and the percentage who carried them out in the WMP survey

In Table 5-7 each row is out of 100. For example, 47.44% of all the respondents' questions were involved in project management. The roles listed in Table 5-7 are the result of grouping together the original options given to respondents and those listed in the free text "others" option. In contrast to the corresponding results in the ASP survey (see Table 5-3), there is no clear single role carried out by all web maintainers. Content development is the most popular role, with 56.4% of respondents listing it as one of their roles. The least carried out role is information design.

5.4.3 Web Experience

Question 2: What is your experience working in this/these role(s) at your current organisation? And what is your experience working in this/these role(s) in total?
Rationale for question – The total number of years of experience working in web development and more specifically in their current organisation will influence their ability to answer the questions accurately. Ideally respondents should have been working at their current organisation for at least one year.
Question Response - As shown in Figure 5-5, most of the respondents in the WMP survey (approximately 66%) had in total over 4 years of web maintenance experience. Just over half the respondents had been working in web maintenance for their current organisation.

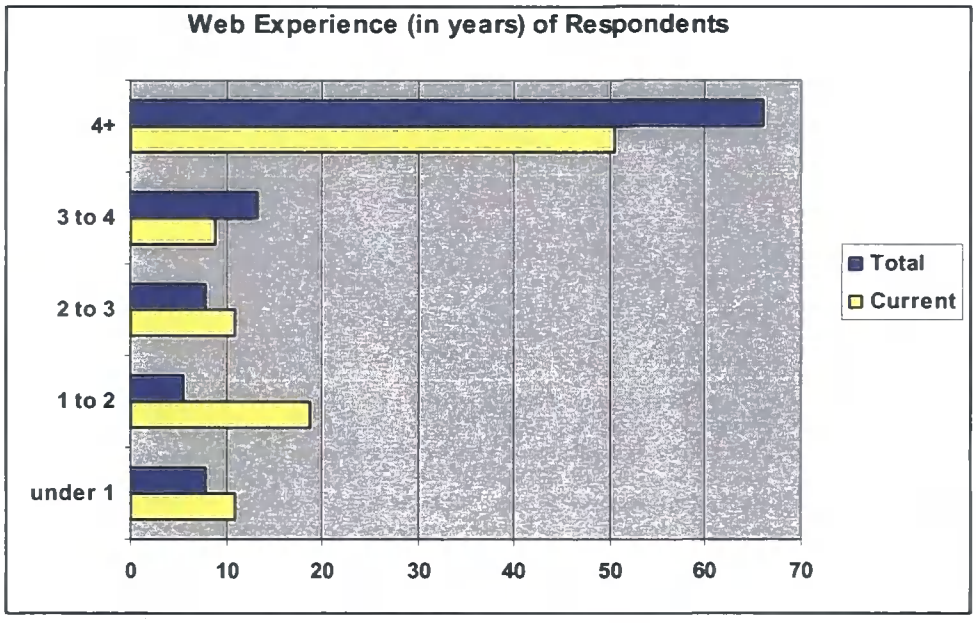

Figure 5-5 Web Experience of Respondents in the WMP survey

Figure 5-5 shows that most of the respondents had over 4 years web development experience in total and also in their current organisation.

5.4.4 Size of Organisations and their Web Teams

Question 3: - Approximately how many individuals work for your organisation? And approximately how many individuals are involved in Web Publishing for your organisation?

Rationale for question – Some of the practices discussed in this thesis will be particularly relevant to larger organisations which have enough information to publish thus justifying the cost involved. Hence it was important to ascertain the average size of the organisations involved in the WMP survey. If they are mainly small organisations then the responses in the survey will be skewed. It is also interesting to monitor the trends in web development team size. This is an indication as to the amount of web work deemed necessary by the organisations, as paying for such a team requires a budget allocation.

Question Response - A large majority (approximately 92%) of the organisations surveyed had a minimum of 100 employees.

Table 5-8 shows also that a large proportion (40%) of respondents worked in organisations with over 1000 employees. Table 5-9 shows that the size of web publishing teams varied greatly. Approximately 30% of the organisations had fewer than 5 people responsible for web publishing.

Organisation Size	%
0 to 10	8
11 to 50	0
51 to 100	0
100 to 500	28.7
501 to 1000	22.99
1000+	40.2

Table 5-8 Sizes of respondent organisations

Publishing Size	%
0 to 5	29.9
6 to 10	6.9
11 to 20	11.5
21 to 50	20.7
51 to 100	18.4
100+	12.6

Table 5-9 Number of individuals involved in web publishing where respondents work

Table 5-9 shows the number of individuals involved in web publishing within respondents' organisations.

5.4.5 Web Publishing Practices

The questions addressing the web publishing practices of the organisations are now discussed. At then end of each question statement the shorthand label used in Figure 5-6 is given.

Question 4: Does your organisation use Content Management Software? (*CMS Installed*)

Rationale for question – Content Management software usage was mentioned during the ASP survey (see section 5.3). Some specialists commented that some content management software produced web pages which contained serious accessible barriers. As such, it was important to see what level of content management usage was amongst the organisations surveyed.

Question Response - 90% of organisations in the WMP survey made use of content management systems (CMS installed). Such a large percentage reflects the increasing importance of and growth in web sites. To cope with this, many organisations have installed CMSs and are publishing more information and services via the web. Several web maintainers (in follow up correspondences) echoed the comments made by specialists that at times efforts to ensure web accessibility were frustrated by content management systems.

Question 5: Does your organisation have a clearly defined and documented model for publishing web content? (*Web Publishing Model*)

Rationale for question – The presence of a clearly defined and documented model for web publishing shows a commitment to a more formalised process of web publishing.

Question Response – 77% of organisations had adopted a web publishing model.

Question 6: Are roles and responsibilities clearly defined and attributed to content throughout the site? (*Roles & Responsibilities*)

Rationale for question – assigning roles and responsibilities shows that organisations have recognised the need to control who publishes and updates web content within the organisation.

Question Response - 79% of organisations assigned web roles and responsibilities and so is further evidence of a more organised approach to web publishing.

Question 7: Is there a process for tracking user feedback? (*User Feedback Tracked*)
Rationale for question – If feedback is tracked then an organisation is committed to accepting user feedback and indicates a willingness to respond to comments made by their users.
Question Response - 67% of the organisations implemented user feedback tracking.

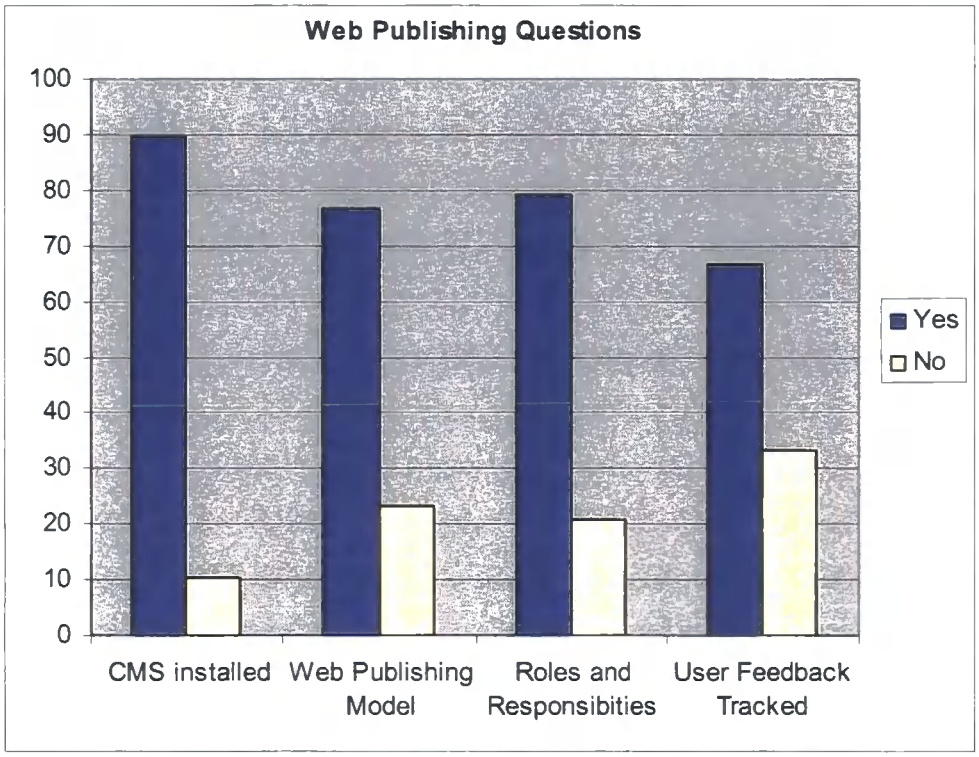

Figure 5-6 Web Publishing Practices Questions and Answers

Figure 5-6 shows the responses to the questions about web publishing practices within the organisations the respondents worked.

5.4.6 Accessibility Practices

The questions addressing the accessibility practices are now discussed. At the end of each question statement the shorthand label used in Figure 5-7 is given.

Question 8: Does your organisation have a clear and comprehensive policy addressing web accessibility? (*Clear policy*)

Rationale for question – Establishing a policy specifically to address accessibility shows that the organisation's management is at least aware of web accessibility.
Question Response – 75% of organisations had specified a clear and comprehensive accessibility policy.

Question 9: Have Management given genuine backing to web accessibility? (*Man Backing*)
Rationale for question – Discovering whether there is genuine backing from management was important to identify to what extent accessibility has been accepted by management.
Question Response - 68% of respondents reported genuine support from management.

Question 10: Has web accessibility been promoted within your organisation? (*Promotion*)
Rationale for question – Promotion is crucial to raising awareness and shows that the organisation is willing to spend time and effort to address the issue.
Question Response - 76% of organisations promoted accessibility issues.

Question 11: Are Accessibility experts or 'gurus' available who can be consulted for assistance? (*Gurus*)
Rationale for question – One of the recommendations on implementing an accessibility plan from the W3C [29] is the installation of Gurus or champions. To what extent this advice is taken is another indication of the importance placed on accessibility.
Question Response - 71% of organisations provided accessibility gurus.

Question 12: Are guidelines set which govern the accessibility of organisational web pages? (*Guidelines set*)
Rationale for question – Improving accessibility requires guidelines to follow. If no guidelines are set then the accessibility efforts may yield little improvements.
Question Response - Nearly all organisations (95%) had set specific standards / guidelines which must be met.

Question 13: Do Web projects within your organisation have a specific budget for accessibility testing and evaluation? (*Budget*)
Rationale for question – A specific budget shows that management are concerned enough to spend money on the issue.
Question Response - 94% of organisations failed to assign a specific budget to meeting accessibility standards.

Question 14: Does your organisation actively monitor and record accessibility data for your website? (*Monitor*)
Rationale for question – Monitoring and recording accessibility is crucial to systematic improvement. It indicates a mature attitude towards process improvement. Monitoring and recording data allows organisations to determine whether accessibility is improving.
Question Response - Only 43% of the organisations surveyed specifically monitored accessibility data.

Question 15: Is accessibility integrated into the web publishing process as part of quality control? (*Integrated*)
Rationale for question – The majority of specialists surveyed in the ASP survey agreed that integrating accessibility into a web publishing process was beneficial.
Question Response - 79% integrated accessibility into the web publishing process.

Question 16: Are there measures in place to monitor the latest web accessibility Recommendations and Guidelines? (*Latest standards*)
Rationale for question – Monitoring the latest recommendations and guidelines demonstrates that organisations are aware that changes may affect how they have to create their web site in the future.
Question Response - 56% of organisations had procedures to monitor the latest accessibility standards.

Question 17: Are there procedures in place to procure and evaluate the latest accessibility tools? (Latest tools)

Rationale for question – Utilising tools to assist in accessible web publishing can save organisations time and money, so it is important whether organisations are actively monitoring the latest accessibility tools available.

Question Response – Only 19% of the organisations had procedures in place for the procurement and evaluation of accessibility tools.

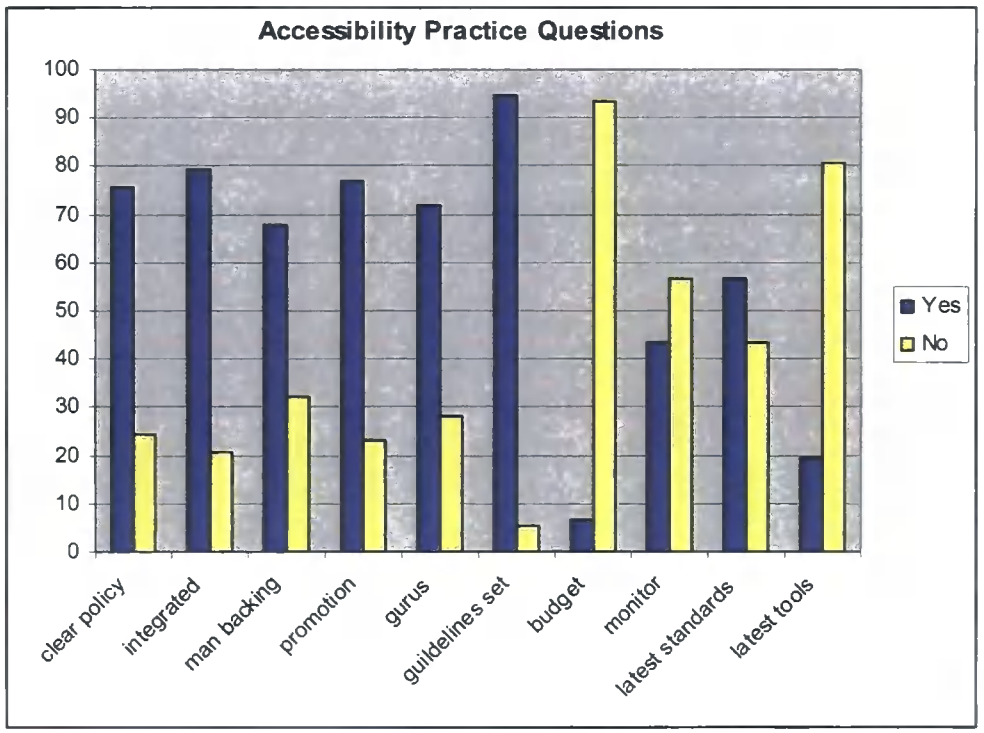

Figure 5-7 Accessibility Practices Questions and Answers

Responses for the accessibility practices carried out by organisations are given in Figure 5-7.

5.4.7 Tools and guidelines used to assess web accessibility

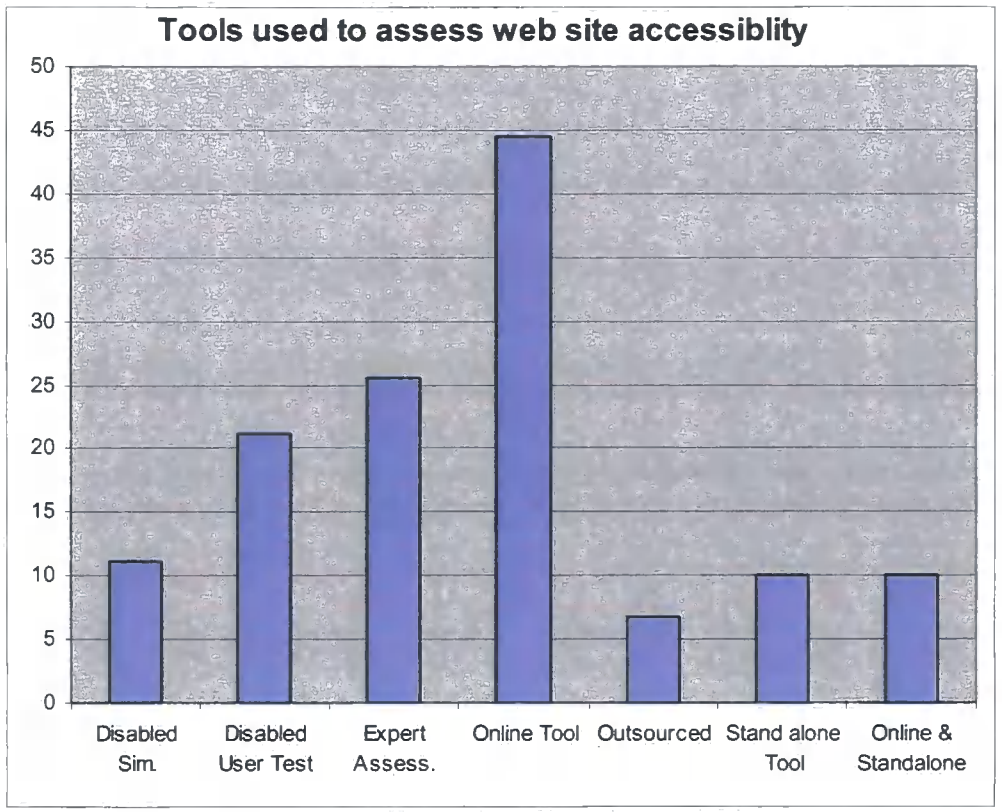

Figure 5-8 Web accessibility assessment tools used by respondents to the WMP survey

Figure 5-8 shows the tools or methods used by respondents to assess web accessibility. The majority of respondents used online tools as part of their assessment (44%). 26% of the respondents used some form of expert assessment and 21% conducted testing with disabled users as part of the assessment.

Accessibility Guidelines	**%**
W3C Accessibility Guidelines 1.0	52
W3C Accessibility Guidelines 2.0	53
Section 508	9
IBM Guidelines for Web Development	8
RNIB	5
internal	33
Misc.	8

Table 5-10 Web accessibility guidelines used by respondents to the WMP survey

Table 5-10 shows the percentage of respondents who made use of the various guidelines. Some organisations may list more than one guideline and the percentage shown is the percentage of respondents who responded as using a particular guideline. Some organisations, for example, state they followed both internal guidelines and version 1.0 of the W3C guidelines.

5.4.8 3rd Party Software and Training Practice

The rationale behind the 3rd party software and training practice questions was to survey how organisations managed accessibility training and how they viewed the accessibility of the web content produced by 3rd party software. The graph label column gives the labels used in Figure 5-9.

Question 18: Does your organisation have a specific written training policy to address accessibility? (*Trainingpolicy*)
Rationale for question – A written training policy is a clear sign that the organisation considers it important. A written policy makes it easier for individuals to access training.
Question Response – Only 30% of the organisations had a policy on training.

Question 19: Is accessibility training provided available for all relevant staff? (*Trainingavailable*)
Rationale for question – This question tried to establish whether organisations actually provided training to people who needed it.
Question Response – 58% of the respondents stated that there was training available to the relevant staff.

Question 20: Please list the current roles that receive such training (Please check as many boxes as apply) (*See Table 5-11*)
Rationale for question – Discovering which roles receive training for web accessibility is important to learn whether the right people are receiving training. For example, are management being trained?
Question Response – see section 5.4.9.

Question 21: Is accessibility awareness a requirement when selecting 3rd party software / services? (*3rdpartyrequire*)

Rationale for question – The user does not make any distinction between web content developed in house by an organisations and content generated by 3rd party software, hence they should be subject to the same standards.

Question Response – 87% of organisations specified accessibility as a requirement when selecting 3rd party products.

Question 22: Are organisational accessibility guidelines / requirements applied to 3rd part content / services (*Access3rdsamestandards*)

Rationale for question – If the same standards are not applied to content generated by 3rd party software, there is a risk that the site may develop a divergence in accessibility standards.

Question Response – 73% of organisations applied the same internal standards to content generated by 3rd party software.

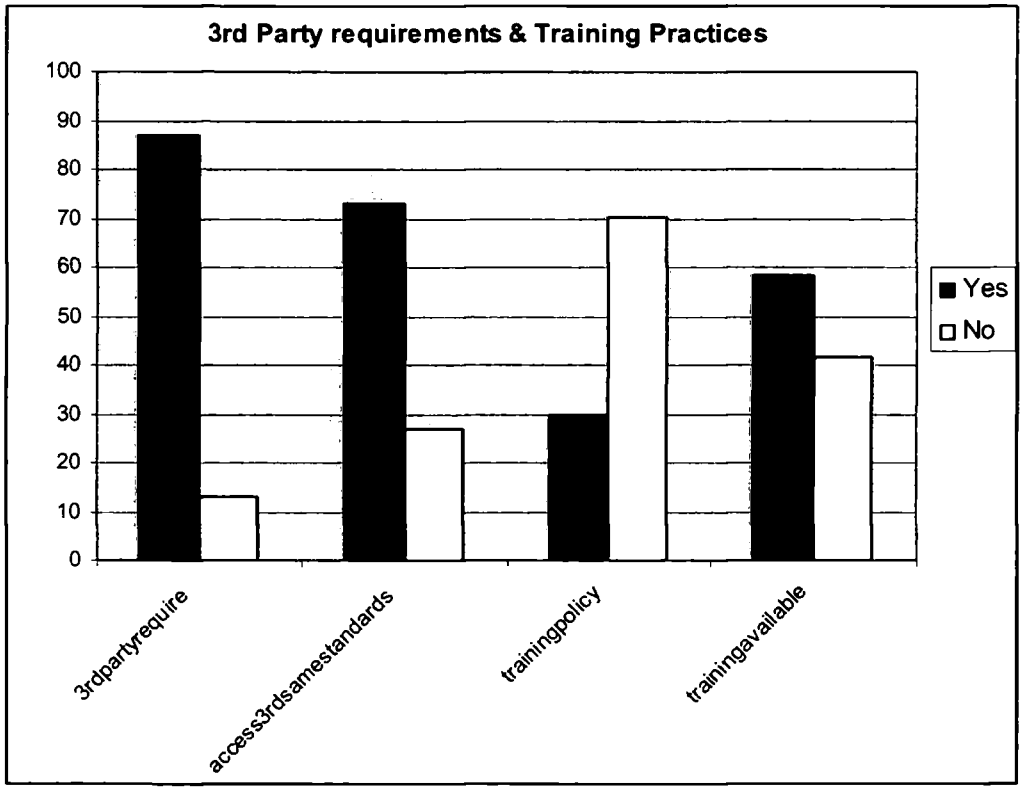

Figure 5-9 3rd Party Requirements and Training Practices Questions and Answers

5.4.9 Roles receiving accessibility training

Table 5-11 shows which roles within the organisations were offered accessibility training (if training was offered at all). There appears to be a great diversity in the roles which are provided with accessibility training. Almost all (98%) of the organisations provided content developers with accessibility training. Web programmers received training in 60% of the organisations surveyed. Fewer than 25% of organisations offered accessibility training to project managers and information designers.

Role	%
Project Manager	22
Accessibility Tester	22
Information Designer	20
User Interface Designer	44
Content Developer	98
Web Programmer	60

Table 5-11 Roles which received accessibility training

5.5 Accessibility Assessments

Following completion of the WMP Survey, manual assessments of the respondents' web sites were carried out. Only the web sites of organisations with respondents who stated in the questionnaire that they desired a manual assessment were assessed. In total 66 web sites were assessed. Results for each were recorded for use in this thesis and also sent to the respondents. Respondents only saw the assessments for their organisation's web site.

5.5.1 Accessibility Ratings

The results for the automated assessments, followed by the manual assessments now follow. The assessment categories were based on the average Page Measure (defined in section 4.6.1) for each site (see Table 3-1).

5.5.1.1 Automated assessment results

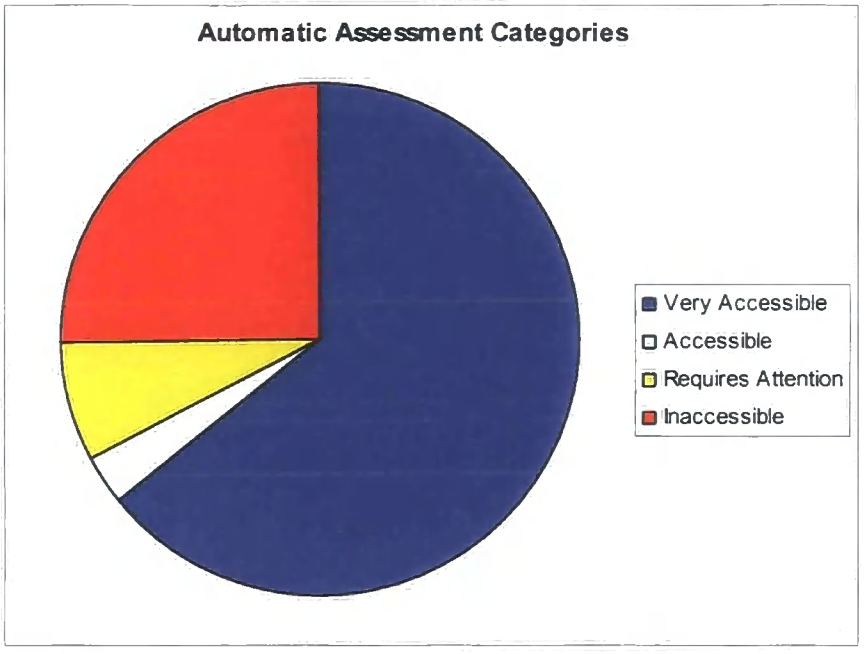

Figure 5-10 Automatic Accessibility Assessment Categories

The assessment process is described in section 4.6.1. Figure 5-10 shows mixed results from the automated assessments. A large percentage (64%) of sites assessed was rated as *Very Accessible*. Very few sites fell into the *Accessible* or *Requires Attention* categories, but approximately 25% of sites were deemed *Inaccessible*.

5.5.1.2 Manual assessment results

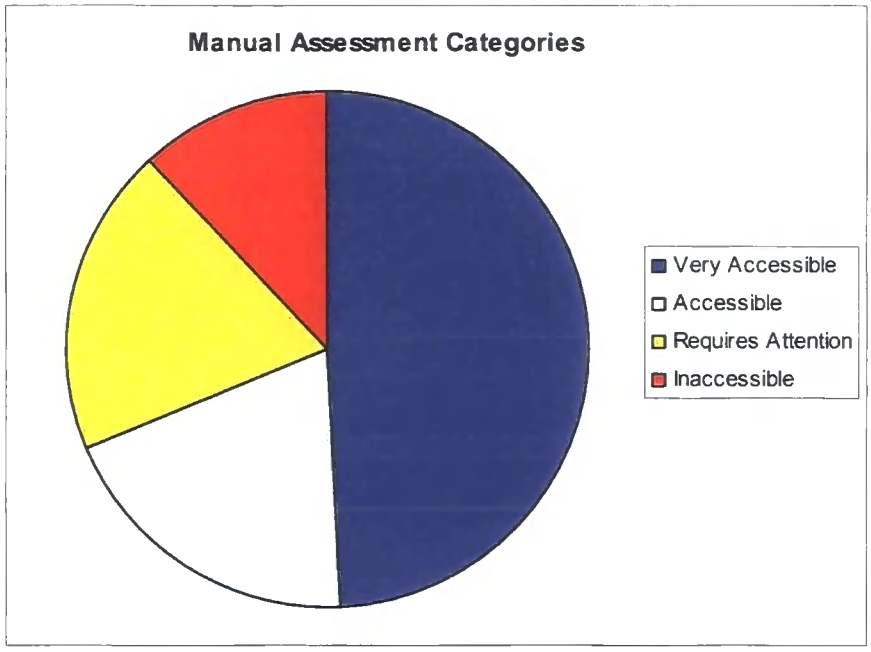

Figure 5-11 Manual Assessment Accessibility Categories

The method used to manually assess each web site is given in section 4.6.2. Results from the manual assessment are shown in Figure 5-11. They reveal that the number of web sites categorised as Very Accessible (49%) is lower than with the automated assessments. Only 11% of the web sites were found to be Inaccessible, this is explored further in the evaluation of results in section 6.3.1. During the manual assessment a greater proportion of web site were categorised as Accessible or Requires Attention (both 19%).

5.5.2 Correlation between automated and manual assessments

To test whether the automated assessments correlate to the manual assessments a Pearson Product Moment correlation was used. A mapping was used to convert the real number automated assessments scores to the integer representations for each of the accessibility rating categories used in the manual assessments (see Table 3-1). Because the Page Measure measures the number and severity of accessibility barriers a page with a higher Page Measure is considered *less accessible* than a page with a lower Page Measure.

From the 66 sites that were assessed using both manual and automated methods the Pearson correlation coefficient was **0.68+**. This is a medium correlation and indicates

that the manual assessment gives a similar estimate of accessibility to the automated assessments.

5.5.3 Effect of Web Publishing Practices on Accessibility

The focus of this section is to compare how different practices affect the accessibility of web sites (represented by the web site's average PageMeasure). One of the main aims of the WMP survey was to determine if certain web publishing practices had a significant effect on the accessibility of the web sites produced. In order to investigate this, the average Page Measures of web sites whose organisations implemented a practice were analysed against the average Page Measures of the web sites whose organisation did not implement the practice.

Statistical analysis of the results is now presented. The affect of each practice will be analysed, followed by combinations of practices to investigate whether certain practices have more impact in conjunction with others.

5.5.3.1 Summary of analysis

Further discussion of these results can be found in section 6.3.2

The analysis showed that there was no significant link between the accessibility of a web site and the implementation of the web publishing practices surveyed in the WMP survey, either in isolation or in combination with other practices. This in itself is to be expected as it is perfectly feasible to have a well organised web publishing practices but without specific attention paid to the accessibility of what is being published. In this case, accessibility will be achieved rather by chance or individual effort, rather than as the result of organisational strategy. However, this thesis will argue that without such fundamental web publishing practices in place, systematically improving web accessibility is not possible.

Similarly, no significant link was found between the implementation of training and third party software practices and web accessibility.

However, there was a significant link found between two of the specific web accessibility improvement practices (*monitoring and recording of accessibility data* and *specifying of accessibility guidelines*) and an improvement in web accessibility.

There was also a significant link found between organisations that implemented all the web accessibility practices and better web accessibility. It is therefore the standpoint of this thesis that in order to ensure accessibility, organisations must implement as many as the accessibility practices as possible.

5.5.3.2 Effect on accessibility of publishing practices in isolation

T-tests were performed on the average PageMeasures (SiteMeasure) for organisation's web site surveyed in the WMP survey. Table 5-12 shows how to interpret the results presented below. The t-tests tested if the average SiteMeasure score for organisations that implemented a certain practice was significantly different to those that did not. For example, do organisations that use Mac PCs have significantly more accessible web sites than organisations using IBM compatible PCs.

Label	*Represents*
Mean	In the No column is the average SiteMeasure for organisations which did not implement the practice. In the yes the same for organisations that did implement it.
Variance	Variance in the Site measures for those that did (Yes) and did not (No) implement the practice.
Observations	Number of organisations who did (Yes) and did not (No) implement the practice.
df	Degrees of Freedom.
t Stat	t statistic.
P(T<=t) one-tail	Probability the result occurred by chance.
t Critical one-tail	Critical value in the t-statistics table.

Table 5-12 Explanation of t-test results

Content Management System (CMS) used

	No	Yes
Mean	2.122814	3.397959
Variance	5.027328	18.79193
Observations	8	58
df	16	
t Stat	**-1.30661**	
P(T<=t) one-tail	0.104907	
t Critical one-tail	1.745884	

The t-test shows there is no significant difference in the accessibility of web sites from organisations that used a CMS and those organisations which did not use a CMS.

Clearly defined and documented model for web publishing

	No	Yes
Mean	2.141616	3.595966
Variance	6.098334	20.42142
Observations	16	50
df	48	
t Stat	**-1.63671**	
P(T<=t) one-tail	0.054116	
t Critical one-tail	1.677224	

The t-test shows that the accessibility of web sites from organisations with a documented web publishing model was *not significantly different* from organisations without one.

Roles and responsibilities clearly defined and attributed to content throughout the site

	No	Yes
Mean	3.243667	3.24331
Variance	7.982312	20.36791
Observations	16	50
df	41	
t Stat	**0.000375**	
P(T<=t) one-tail	0.499851	
t Critical one-tail	1.682878	

The t-test shows the accessibility of web sites from organisations that clearly defined roles and responsibility for content was *not significantly different* from organisations without clearly defined roles.

Is there a process for tracking user feedback?

	No	Yes
Mean	3.017948	3.35612
Variance	15.77807	18.2499
Observations	22	44
df	45	
t Stat	**-0.31785**	
P(T<=t) one-tail	0.376034	
t Critical one-tail	1.679427	

The t-test shows that the accessibility of web sites from organisations that track user feedback was *not significantly different* from organisations that do not track feedback.

Since no significant differences in accessibility were found in the individual web publishing practices, another test was performed on organisations that perform all of the above practices against organisations that carried out none.

5.5.3.3 Effect on accessibility of publishing practices combine

All Web Publishing Practices combined

	No	Yes
Mean	3.104767	3.39069
Variance	13.31613	21.83815
Observations	34	32
df	59	
t Stat	**-0.27589**	
P(T<=t) one-tail	0.3918	
t Critical one-tail	1.671093	

The t-test shows that the accessibility of web sites from organisations that perform all the web publishing practices was *not significantly different* from organisations that perform none.

Practices specific to Accessibility

There now follows analysis of the effect of the practices which are directly related to accessibility.

Clear and comprehensive policy addressing web accessibility

	No	Yes
Mean	4.539645	2.828597
Variance	23.4249	14.91636
Observations	16	50
df	21	
t Stat	**1.288879**	
P(T<=t) one-tail	0.105729	
t Critical one-tail	1.720743	

Despite the large decrease in mean SiteMeasure, the t-test shows that it is not possible to prove that a clear and comprehensive accessibility policy alone is enough to produce more accessible web sites.

Genuine support from management for accessibility

	No	Yes
Mean	3.033068	3.34856
Variance	14.27686	18.98811
Observations	22	44
df	48	
t Stat	**-0.30351**	
P(T<=t) one-tail	0.381406	
t Critical one-tail	1.677224	

There is *no significant difference* between the mean accessibility of the web sites which management support for accessibility and the web sites in organisations without such backing.

Accessibility promotion within the organisation

	No	Yes
Mean	4.061582	2.981576
Variance	31.57163	12.85816
Observations	16	50
df	19	
t Stat	**0.723162**	
P(T<=t) one-tail	0.239194	
t Critical one-tail	1.729133	

Although there is a considerable reduction in the mean SiteMeasure for sites whose organisations promote accessibility, however, the t-test shows this *not to be significant.*

Use of "gurus" to assist in improving accessibility

	No	Yes
Mean	4.126152	2.859589
Variance	19.6633	16.03997
Observations	20	46
df	33	
t Stat	**1.097482**	
P(T<=t) one-tail	0.140187	
t Critical one-tail	1.69236	

Although there is a considerable reduction in the mean SiteMeasure for sites whose organisations employ accessibility gurus, the t-test shows that this reduction is *not significant.*

Guidelines governing accessibility specified

	No	*Yes*
Mean	8.497039	2.904451
Variance	18.72306	15.47651
Observations	4	62
df	3	
t Stat	**2.518676**	
P(T<=t) one-tail	0.043139	
t Critical one-tail	2.353363	

There is a large reduction in the mean SiteMeasure for those web sites with accessibility guidelines specified. However, there were *only 4 organisations* that did not specify guidelines which do not produce reliable t-tests results.

Budget

Since there were only two organisations that had a dedicated budget for accessibility, it was not possible to analyse the affects of budget on accessibility.

Monitor and record accessibility data

	No	**Yes**
Mean	4.215565	1.924023
Variance	24.47611	4.721896
Observations	38	28
df	54	
t Stat	**2.541851**	
P(T<=t) one-tail	0.006965	
t Critical one-tail	1.673565	

There is a large reduction in the mean SiteMeasure of those sites whose organisations monitor and record accessibility (4.22) data and those that do not (1.92). The t-test gives a t Stat of 2.17 which is above the critical value of 1.67 required to be

considered significant. The probability of this occurring by chance is 0.007 which is below the threshold 0.01 to be 99% confident that this did not occur by chance and hence *is significant.*

Accessibility integrated into the web publishing process

	No	Yes
Mean	4.865092	2.791254
Variance	28.50017	14.21646
Observations	13	52
df	15	
t Stat	**1.320698**	
P(T<=t) one-tail	0.103199	
t Critical one-tail	1.75305	

Despite the considerable reduction in the mean SiteMeasure for sites whose organisations integrate accessibility, the t-test shows that this reduction is *not significant.*

Measures in place to monitor the latest accessibility recommendations and guidelines

	No	Yes
Mean	3.601683	2.926057
Variance	24.14768	11.34788
Observations	31	35
df	52	
t Stat	**0.643254**	
P(T<=t) one-tail	0.261443	
t Critical one-tail	1.674689	

There is a reduction in the mean SiteMeasure for sites whose organisations monitor the latest accessibility recommendations and guidelines; however, the t-test clearly shows that this reduction is *not significant.*

Procedures in place to procure and evaluate the latest accessibility tools

	No	Yes
Mean	3.498721	2.094432
Variance	19.55405	5.639637
Observations	54	12
df	31	
t Stat	**1.539478**	
P(T<=t) one-tail	0.066917	
t Critical one-tail	1.695519	

There is a reduction in the mean SiteMeasure for sites whose organisations have procedures in place to procure the latest accessibility tools; however, the t-test shows that this reduction is *not significant.*

5.5.4 Affect on accessibility when practices are combined

Only two of the accessibility practices by themselves influenced accessibility:

- Monitor and record accessibility data
- Guidelines governing accessibility specified

However, it is an over simplification to state that all an organisation needs do is monitor and record accessibility data or specify which guidelines they will use to ensure their web site is accessible. Rather these practices are symptomatic of how an organisation addresses the problem. Therefore the relationship between these two practices and the other publishing practices is now examined.

No Practices	% of organisations
2	3.85
3	3.85
5	11.53
6	19.23
7	19.23
8	34.62
9	7.69
10	3.85

Table 5-13 Number of practices implemented by organisations which monitor and record access data

There are 10 practices related to accessibility. Table 5-13 shows that the majority of organisations which monitored and recorded accessibility data combined this with most (84%) of the other accessibility practices.

Due to the small number of organisations (4) which did not specify guidelines further analysis will not be made. However, section 6.3.2 discusses the significance of not specifying guidelines and what this indicates about an organisation.

Accessibility practices combined

	No	Yes
Mean	3.8335	1.39932
Variance	20.26032	3.545709
Observations	50	16
df	59	
t Stat	**3.074565**	
P(T<=t) one-tail	0.001596	
t Critical one-tail	1.671093	

Finally, the 16 organisations which implemented the accessibility practices (with the exception of specifying a budget, the procurement of the latest tools and research into the latest standards) were found to have web pages with a very low mean SiteMeasure of 1.40 compared to those that implemented none. The t-test shows this improvement in accessibility *was significant and the 0.05 level.*

5.5.4.1 Omission of accessibility practices from combination

3 practices were omitted from the combination of accessibility practices the reason behind this is now explained. The practices were as follows:

- **Measures in place to monitor the latest accessibility recommendations and guidelines** - The tool used to evaluate the web sites is based on current recommendations.
- **Specific accessibility budget available** - There were not enough organisations with a specific budget available.
- **Tools** – currents tools may be sufficient.

5.5.5 Third Party Content and Training Practices

Accessibility awareness requirement set for selecting 3rd party software or services

	No	Yes
Mean	2.727395	3.298902
Variance	13.44489	18.27775
Observations	9	56
df	12	
t Stat	**-0.4236**	
P(T<=t) one-tail	0.339676	
t Critical one-tail	1.782288	

Sites whose organisations state accessibility awareness as a requirement for 3rd party software or services had a higher mean SiteMeasure. However, the t-test shows that this increase *was not significant.*

Same organisational guidelines set for 3rd party software

	No	Yes
Mean	4.189653	2.965085
Variance	27.47424	14.31483
Observations	15	51
df	18	
t Stat	**0.842569**	
P(T<=t) one-tail	0.205263	
t Critical one-tail	1.734064	

Although there is a decrease in the mean SiteMeasure for sites whose organisations set the same organisational guidelines for 3rd party software or services, the t-test shows that this reduction is *not significant.*

Specific written accessibility training policy in place

	No	Yes
Mean	3.056787	3.826548
Variance	15.48494	23.45455
Observations	50	16
df	22	
t Stat	**-0.57767**	
P(T<=t) one-tail	0.284675	
t Critical one-tail	1.717144	

Although there is small increase in the mean SiteMeasure for sites whose organisations have a written accessibility training policy in place, the t-test shows that this reduction is *not significant.*

Accessibility training provided for relevant staff

	No	Yes
Mean	2.729732	3.716178
Variance	11.11062	22.01231
Observations	28	37
df	63	
t Stat	**-0.99055**	
P(T<=t) one-tail	0.162848	
t Critical one-tail	1.669402	

Although there is an increase in the mean SiteMeasure for sites whose organisations provide accessibility training for relevant staff, the t-test shows that this reduction is *not significant.*

Third party and training practices combined

	No	Yes
Mean	3.390861	2.579804
Variance	19.25038	8.275808
Observations	54	12
df	24	
t Stat	**0.79297**	
P(T<=t) one-tail	0.217782	
t Critical one-tail	1.710882	

There is a small decrease in the mean SiteMeasure for sites whose organisations implement the training and 3rd party software practices; however the t-test shows that this reduction is *not significant* and could be the result of chance.

5.6 Tree-Map Experiment Results

The Tree-Map experiment aimed to establish whether a Tree-Map based tool could provide more accurate web maintenance information more efficiently than using a report based tool. The timing and accuracy results from this experiment (see section 4.7) are now presented.

5.6.1 Group organisation

As specified in the experiment protocol in section 4.8, participants were randomly assigned to one of four groups.

Group	Task One	Task Two
1	Tree-Map, Hillingdon	HTML report, Chichester
2	HTML report, Hillingdon	Tree-Map, Chichester
3	Tree-Map, Chichester	HTML report, Hillingdon
4	HTML report / Chichester	Tree-Map, Hillingdon

5.6.2 Time to complete tasks

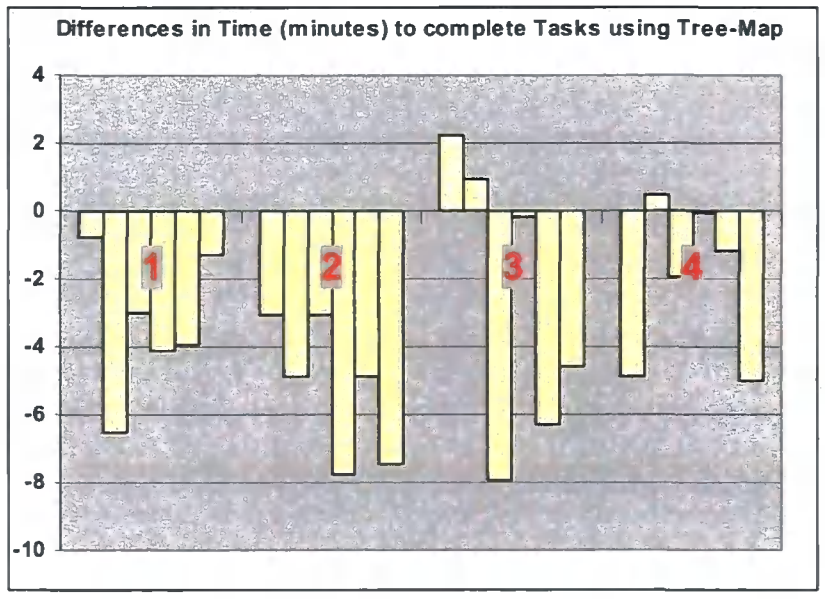

Figure 5-12 Time differences in minutes for tasks completion time using Tree-Map based tool.

Figure 5-12 shows the time differences for the two tasks for the four experimental groups. All four groups show a clear trend, that the tasks were completed in less time using the Tree-Map based tool. Two participants in group 3 and one in group 4 required more time to complete the tasks but the extra time was less than the average time savings offered by the Tree-Map tool.

5.6.3 Accuracy of responses

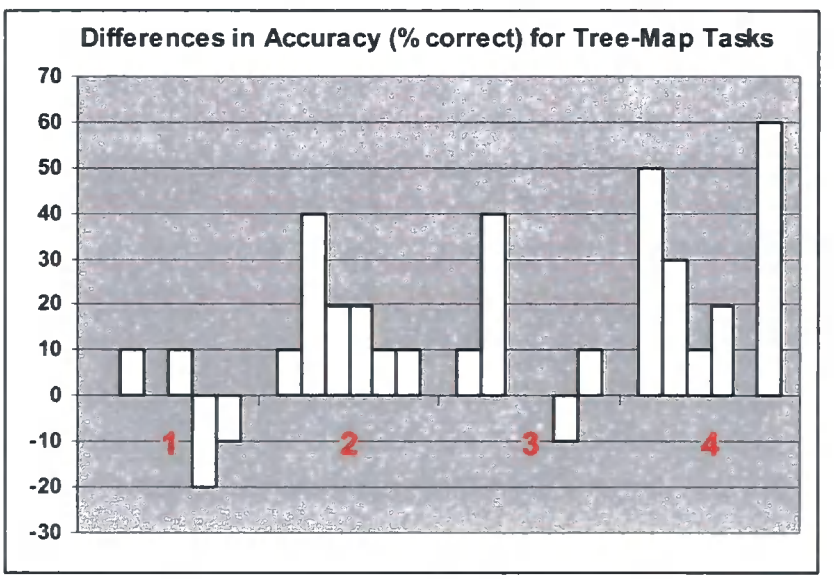

Figure 5-13 Accuracy differences for tasks using Tree-Map based tool.

Figure 5-13 shows the differences in percentage of correct answers provided by participants using the Tree-Map based tool. The majority of subjects within each of the four groups achieved either the same or greater accuracy using the Tree-Map tool than using the report based tool. There now follows statistical analysis of both the timings and accuracy results.

5.6.4 Analysis of Task Completion Time

The overall mean difference in task completion time when using the Tree-Map based tool was **-3.2983 minutes**.

SUMMARY

Groups	Count	Sum	Average	Variance
1	6	-19.63	-3.27167	4.443617
2	6	-31.18	-5.19667	4.163747
3	6	-15.78	-2.63	17.70844
4	6	-12.57	-2.095	5.60303

ANOVA

Source of Variation	SS	df	MS	F	P-value	F crit
Between Groups	32.99437	3	10.99812	1.378261	0.278236	3.098391
Within Groups	159.5942	20	7.979708			
Total	192.5885	23				

Figure 5-14 ANOVA of task completion time differences for groups

Figure 5-14 is the ANOVA of the competition time differences and shows that there are ***no significant effects due to the order*** in which participants applied each treatment (tool) or the site they used. The F ratio is **1.378**, which is below the F critical value of **3.098**. Therefore the mean competition times differences between Tree-Map and HTML based tool tasks were not significantly different amongst the four groups.

Testing the null hypothesis "*that there is no difference between the time taken to perform the tasks using Tree-Map based tool and the time taken using the HTML reports*" requires testing whether the observed difference between sample means (-3.2983) is significantly different from 0.

Standard Error (SE) is calculated by using the within-group standard deviation (sd) from an ANOVA of the differences with 20 degrees of freedom (df).

Within-group standard deviation (sd) = ***7.98.***

SE = Square root (sd = 7.979708333 / N = 24) = ***0.58.***

*t-statistic = (-3.2983/0.5766) = **5.72** with 20 df.*

*The 0.01 critical t-value from the t-distribution with 20 df is **2.53.***

*This t statistic of **5.72** (>2.53) confirms that the mean value is significantly less than 0, (p<0.01) using a two-tailed test.* Thus participants took less time to complete their tasks using Tree-Map based tool than when they used the HTML report based tool.

5.6.5 Analysis of Task Accuracy

The mean percentage of correct answers for participants using the Tree-Map based tool was **77.5%**. This was **13.3%** greater than those using the HTML report based tool (**64.2%** correct).

The null hypothesis "*that there is no difference between the accuracy of the tasks carried out using Tree-Map based tool and the accuracy of tasks carried out using the HTML reports*" is now tested.

Within-group standard deviation (sd) = **276.67**
SE = Square root (sd = 276.67/ N = 24) = **3.39**
t statistic = (-13.33/3.39) = **3.93** with 20 df

The 0.01 critical value from the t-distribution with 20 df is **2.53**.

However, the order of tasks had a significant effect on the task accuracy results. The difference in accuracy for participants in groups 1 and 3 who used the Tree-Map based tool first only was **3.33%**, whereas for those in groups 2 and 4 the difference was **23.33%**. Participants who carried out the Tree-Map task first did not significantly improve their accuracy in the second task, whereas participants using the HTML report based tool first improved significantly in the second task.

To test this, the standard error is calculated by taking the square root of within-group standard deviation of 276.67 divided by 12 (number of participants in groups 1 & 2 (and also in 3 & 4)) = SQRT(276.67 / 12) = **4.8**.

Testing whether using the HTML report based tool followed by Tree-Map based tool leads to no improvement in correctness gives a t value = 23.33/4.80 = **4.86** which is significant ($p<0.01$, 2-tailed test). The 0.01 critical value from the t-distribution with 20 df is **2.53** and so the null hypothesis is rejected.

Testing whether using the Tree-Map based tool followed by HTML report based tool leads to no improvement in correctness gives a t value 3.33/4.80 = **0.69**, which is not significant. The null hypothesis cannot be rejected and it cannot be concluded that the Tree-Map tool alone offers a significant improvement in accuracy for web maintenance tasks. It appears that once participants know how to do the task, they perform at a standard level of competence irrespective of the tool support. However, whereas the HTML based tool doesn't help participants learn how to complete tasks, the Tree-Map based tool does support learning. Participants performed significantly better in terms of accuracy when their second task was the Tree-Map based task.

5.7 Summary

The analysis of task completion time shows that participants using the Tree-Map based tool took on average less time (just over 3 minutes) to complete the tasks. Also it should be noted that 8 HTML report based tasks (and only 1 Tree-Map based report) exceeded the time limit of 20 minutes and their timings were record as 20 minutes. This means that the timing savings offered by the Tree-Map tool may in reality be greater. The Tree-Map based tool used in the experiment offered less functionality and lower usability than the full application developed in Java. It is believed this too will have affected the completion time negatively. In terms of accuracy improvement, it is not possible to establish whether the Tree-Map based tool produces more accurate results. Regardless of tool support the accuracy levels achieved remained fairly similar, however the results show that using the HTML report based tool first leads to a significant increase in accuracy in the second group of tasks.

Certain trends were identified from the ASP and WMP surveys and the Tree-Map experiments results. They are listed below and will be evaluated in more detail in the evaluation of results (Chapter 6):

- **Increasing sophistication in web publishing** – it is evident from the results of the WMP survey (see section 5.4.5) that organisations are implementing practices which enable a more sophisticated approach to web publishing:
    - **More formalised approach to the web** – the majority (77%) of organisations had a web publishing model, and had roles and responsibilities for content publishing (78%).
    - **Automated content management** – 90% of organisations used content management software. This reflects both the importance of efficient web content management. Content management software is seen by organisations as a tool to support formalised web publishing without significantly increasing the workload or skills demand on employees.
- **Levels of adaptation for web accessibility** – one of the main aims of the ASP and WMP surveys was to establish how far organisations have adapted to provide accessible web sites. This is achieved by evaluating the number of organisations implementing accessibility best practice (see Literature survey accessibility section). In particular the following aspects of web accessibility are discussed:
    - **Working Environment** – the WMP and ASP surveys revealed there was a mixed attitude towards web accessibility. The specialists reported a negative and under valuation of accessibility. Whereas, the majority of respondents in the WMP survey (75%) reported that an accessibility policy had been implemented in their organisation and 95% had specific accessibility guidelines (95%) in place.
    - **Consultants** – the specialists in the ASP survey reported they were employed for a variety of roles and carried out a diverse range of activities.
    - **Training** – the levels of training reported were low. Only 30% of organisations had a formal training policy.
    - **Gurus** – appointing in-house accessibility champions or gurus is intended to act as a source of know-how and motivation. 70% of the organisations had installed gurus.

- **Budget allocation** – implementing these adaptations will cost time and money. Budget allocation is a good indication an organisation is earnest about improving a situation. With only 6% of organisations having a specific budget there appears to be no shift towards providing specific monies to ensuring web accessibility.
- **Accessibility tool support** – just as content management software is required to assist organisations in web publishing, so too are accessibility tools important in assisting organisations to create accessible web sites.
- **Coping with long term change** – there were a relatively small percentage of organisations which implemented practices aimed at anticipating new guidelines (55%) or identifying new and potentially beneficial accessibility tools (20%).

6 Evaluation of Results

6.1 *Increasing sophistication in web publishing*

This chapter contains an evaluation of the results of the study outlined in Chapter 5. In order to highlight specific problems and potential solutions in web publishing scenarios are used. The scenarios are based partly on the author's own professional experience and are centred upon around a fictitious university. Universities publish large amounts of diverse content and have a devolved management structure. A university web site typically includes:

- Teaching material for students.
- Administration information for both staff and students.
- Promotional material for industry, media and prospective students.
- Community focussed pages.

Universities face the problem of balancing the need for devolved content authoring, (e.g. lecturers must be able to create and publish their own online teaching material), whilst still ensuring pages conform to the accessibility guidelines and internal "*branding*". The web site is a resource that supports communication and the daily running of a university. It is crucial that information on the web site is current, consistent and accessible. The scenarios demonstrate how the practices evaluated can contribute to achieving this goal.

6.1.1 More formalised approach to the Web

The results discussed in this section relate to section 5.4.5.

A more formalised approach to web publishing is represented by the presence of practices aimed at controlling the authorisation, publication and updating of content. Results from the WMP survey show that a large percentage of organisations had taken a more formalised approach to publishing web content. The majority (77%) of the organisations had a formalised model for web publication and had defined roles and responsibilities (79%) within that model. Hence, content publication has moved beyond a hobby for technically minded individuals or the IT support office, where individuals upload content to the web server as and when they like, and has matured

into a delivery service in need of the same degree of management and infrastructure as any other mass media. Just as an organisation would not allow individuals to use organisational resources to publish paper based media without some form of authorisation, it is also now less likely that they would allow web publication without similar levels of authorisation. Introducing more control can also reduce inconsistency and reduce the amount of out-of-date or inaccurate information published on their web site. Such qualities as accuracy and timeliness are crucial to public information web sites. The scenario below gives an example of the potential problems posed by ad-hoc web publishing and the solution offered by formalising the web publishing model.

The following scenario demonstrates the benefits of a more formalised approach to web publishing for a university health centre.

University Health Centre Scenario

A university health centre's publishes a web site containing dates for open surgery sessions. The web pages are maintained by the office secretary and are updated casually.

Problem:

- Inconsistent surgery dates are published. Last year's diary is not removed from the web site. Some students have bookmarked the old page and so receive the wrong day for open surgery.
- Without a formalised approach, the maintenance of meta-data such as who is responsible for updating the page, and when it needs to be updated is unknown. There is also no formal method of authorisation and hence the accuracy of the information published is dependent on one person.

Consequence:

- Students make a wasted journey to the surgery. This causes unnecessary frustration and physical strain, and harms the reputation of the web site. Patients might stop using the site.

Recommendations:

- Implement web publishing model where all content has an expiry date (see section 7.2). Content expiry maintenance is a manual and semi-automated task

within the Compile Content stage (A1) and also at the Authorisation stage (A3).

- Assign the responsibility for maintaining this content to a relevant member of the surgery. This is also addressed at the Compile Content stage (A1) meta-data such as date of creation can be generated automatically, whereas others are either semi-automated or manually created.

6.1.2 Automated content management

The results discussed in this section relate to section 5.4.5.

One of the most interesting observations from the WMP survey is the uptake of content management systems (90%). As most of the participating organisations were local government agencies this may be in response to government recommended practice [15]. The popularity of content management systems is recognition of increasing web usage [46] and the demand for information and services to be provided online. This demand comes from both from central government who wish to inform, and the general public who wish to be informed, via web publications. Content management systems provide the automated support required for a more formalised approach (described in section 6.1.1) required to manage greater involvement in web publishing. They allow the assignment of roles and responsibilities, the enforcement of content verification and expiry dates as well as providing basic configuration management. Such systems are also frequently used to de-skill and standardise web content creation and publication. Their usage shows a clear commitment by organisations to include content from individuals who previously would have had no direct contact with web publishing. Individuals who would not feel comfortable learning and creating web pages from scratch instead use pre-defined page templates which restrict the scope for accessibility barriers or other breaches in the organisation's web standards. The template can also ensure that meta-data such as page expiry, content- author identification is added at the time of creation. Once this has been completed, the newly created or updated content will be automatically passed to a content quality manager for verification and authorisation.

The following scenario demonstrates the benefits of an automated content management for an academic department within a higher education institution

University Teaching Material Scenario

- University academic departments publish teaching material online.

Problem Scenario:

- A disabled student fails their final year exams and fails to gain a degree. Upon leaving take legal action against the university, stating the online teaching material published during their time was inaccessible and hence they were at substantial disadvantage to other non-disabled students.

Consequence:

- Without a content management system in place it is impossible for the administrators of the web site to retrieve the precise teaching material that was published, as it might have been subsequently removed or altered. Teaching staff might have since left the university. Retrieving each published version of the article is obviously important to verify the claims made by the student.

Recommendations:

- The integrated accessibility publishing model introduced in Section 7.2 aims to improve accessibility and to catch any severe barriers before they are published.
- One of the benefits of an accessibility log discussed in Section 7.3 is that an organisation has an audit trail of all accessibility work carried out on the site. The log can be used to prove that reasonable adjustments were made to online teaching material.

6.2 *Levels of adaptation for Web Accessibility*

6.2.1 Working Environment

The results discussed in this section relate to sections 5.3 and 5.4.6.

The ASP survey showed that many web maintainers saw accessibility policies as a *roadblock* or restriction on their creativity (see section 5.3). This perception will obviously negatively affect the acceptance of the policy. For any restrictive policy to fully succeed there must be both awareness and acceptance of it within the organisation. In the WMP survey only 68% of respondents reported genuine support for web accessibility from management. This means that almost a third of the organisations did not have support from management, support which is vital in

convincing individuals to participate in implementation of accessibility throughout the web site.

6.2.2 Consultants

The discussion in this section relates to section 5.3.

Some organisations which find themselves legally obligated to ensure web accessibility may not have the required skills in-house to implement an accessible web policy. As such accessibility consultants will be employed to improve accessibility. As the ASP survey (see section 5.2.3) showed consultants were mainly employed to:

- Assess the accessibility of the organisation's web site.
- Raise awareness of accessibility.
- Train content authors and Content Quality Managers.
- Process improvement and strategy.

The three most popular roles carried out by the accessibility specialists (raising awareness, training and process improvement) are aimed at ensuring that an organisation has staff members that are aware of the importance of accessibility, capable of producing accessible web pages and have the processes and strategy in place to support them. There is little value (and great cost) in consultants simply assessing a web site and removing accessibility barriers, if the next update to the web site introduces new barriers or even re-introduces barriers which had been previously removed. Without some action taken within the organisation to improve the process of web publication organisations are reliant on external consultants which is costly in terms of time and money.

University Prospectus Scenario

- University admissions department publishes their prospectus and other information on their web site.

Problem Scenario:

- The prospectus was designed by the IT officer and does not comply with accessibility guidelines. There is no one within the department with the skills needed to improve accessibility.

- The IT officer is unfamiliar with web accessibility does not have the time required to research the area fully to gain the necessary skills.

Consequence:

- The web site remains inaccessible.

Recommendations:

- The department hires an accessibility consultant who trains the IT officer, making him / her aware of accessibility, how to assess web pages and remove barriers.
- The IT officer is then sent on regular training courses to ensure they maintain the skills necessary, and eventually, they take on the role of accessibility guru, offering advice and training to others within the department. This means there is a consistent and ever present source of expertise within the department rather than incurring the expense of bringing in outside consultants.

6.2.3 Training

The results discussed in this section relate to sections 5.4.8 and 5.4.6

6.2.3.1 General training issues

Ensuring accessible web content requires training. Even if an organisation uses a content management system with accessible templates to create pages it can be nullified if the individual content author does not realise how to use them correctly. The WMP survey revealed that only 30% of organisations had a specific written training policy addressing accessibility but 58% made training available to those who required it. When considering training it is important to look at which roles within the organisation receive accessibility training. The WMP survey showed that amongst organisations that provided training, almost all of them provided it for those involved in content development. However, only a minority of organisations (22%) provided project managers or information architects with accessibility training. The lack of training for project managers may explain the relatively low percentage of respondents (68%) in the WMP survey who reported genuine web accessibility support from management within their organisation.

One possible explanation for the poor provision of training provided for management is the perception that accessibility is a purely technical issue. This viewpoint can lead

to unaware staff underestimating the time, tools and tactics required to ensure a high standard of accessibility. It is better to have staff members that are well trained in producing accessible web content from the offset rather than having to employ staff or external specialists to remove serious accessibility barriers later on. The costs incurred by the selection of inadequate content management systems (i.e. those which do not produce accessible web pages) is substantial as enterprise class content management systems are expensive. If one is installed at considerable cost, management are even less likely to pay for either corrective maintenance to enable accessible web content or invest in a new system. Therefore, the extra work created for content authors to fix the problems will increase and costs will increase further.

6.2.3.2 Management training

Management passing the burden of responsibility onto content developers without explanation may lead to the negative web maintainer attitudes reported by the specialists in the ASP survey. Management must be able to make the moral and business case to their team and they can only do this if they themselves fully grasp the issues. Training provided to management should not be the same as that provided to programmers or content developers. In recognition of this fact the Royal National Institute for the Blind provides two forms of accessibility training [88]. The first is a non-technical workshop aimed at those responsible for the management and marketing of web sites and concentrates on introducing accessibility and the surrounding issues. The second is a practical based workshop aimed at providing technical staff with both background and technical skills required to implement accessibility.

New member of staff scenario

A new member of staff joins the departmental support team and is required to author and update external web pages.

Problem Scenario:

- The staff member is unfamiliar with the university's policy and requirements for accessibility.
- They have to start authoring and updating immediately and there is neither the budget nor time to offer training before they start.

Consequence:

- The web pages they update do not conform to university accessibility requirements. Hence the accessibility of the department's web site is reduced.

Recommendations:

- During the Compile Content stage (see 7.2.2) accessibility is automatically assessed.
- The new staff member is made aware of potential accessibility barriers in content by the automated assessment after each major update.
- If the staff member is unsure of the action to take for either removing or ignoring the reported barrier, they can consult the accessibility log. This shows them actions taken and comments made by other content authors within their organisation. This form of semi-automated knowledge transfer allows staff members to learn at their own speed and is specific to the needs of the organisation.
- If the accessibility log does not contain the information required, then further research is required. Once a solution has been found, the staff member can add it to the accessibility log and hence the knowledge gained by each staff member is retained within the organisations for use later.

6.2.4 Accessibility Gurus

The results discussed in this section relate to sections 5.4.6 and 5.4.8.

Employing gurus is a practical solution for organisations with restricted budgets. Hiring or training accessibility gurus gives staff continual access to a source of information and accessibility expertise. This combined with the costs and current relative lack of formal accessibility training available perhaps explains why the provision of accessibility gurus is practiced on a much wider scale (71% had access to Gurus) than that of formal training (30% had a formal training programme). Gurus can provide a method of skills transfer specifically related to the organisation and its structures and processes. They can also be a good driver for progress and improvement and it is reasonable to assume that part of a guru's role will be to stay informed of the latest standards and guidelines, and also familiarise themselves with potentially useful and cost saving tools.

University Web site development

- The university web development team wish to redevelop the university web site.

Problem Scenario:

- The university's central web site is an important advertising mechanism and source of information for members of the public, media and in particular prospective students.
- The web site must be accessible to ensure that students using assistive technologies can use with the web site unhindered.
- The web development team do not have any specific expertise in accessible design and other commitments and budget restrictions which mean that providing training for everyone in the team is unfeasible.

Consequence:

- Despite the best efforts of the web development team, they fail to improve accessibility to a high level.

Recommendations:

- The web development team nominate an accessibility guru. This individual attends external training courses and is given a budget to purchase accessibility related literature.
- The accessibility guru returns to the web development team and reports back on what they have learned. They act as a source of advice, and promoter for accessible design. The guru will be called on to help decide on any major accessibility issues.
- The integrated accessibility publishing model supports the work of Gurus by providing controlled stages (such as the Quality Control stage described in section 7.2.3) in the content lifecycle where a Guru can check the accessibility of content before it is published.
- Gurus can monitor and add to the accessibility log to ensure that the work being carried out by content authors is correct and effective. For example, if the automated assessment logs show that a particular content author's is ignoring or not properly addressing accessibility barriers, the Guru can contact that content author using the log as evidence.

6.2.5 Budget allocation

The results discussed in this section relate to section 5.4.6.

Only a very small (6%) number of organisations had a specific budget for web accessibility. This reveals a low estimation of the value of accessibility and the costs incurred ensuring continual accessibility. If a web site is to be made accessible then some extra work will be required and this will incur costs. Assessment, training, process improvement and tool acquisition all require consideration in any budget planning. If specific time and money is not set aside for this work then it is likely to be rushed and corners cut. It is difficult for a web developer to justify a delay or cost increase in a project to a project manager solely on the grounds of accessibility, especially if the organisation in which a web developer works is not supportive of web accessibility. However, it might be the case that even though an organisation does not have a specific budget available it incorporates accessibility into the overall web budget.

University department budget review

A department is reviewing its annual budget and the question of how much money to assign to web development is discussed.

Problem Scenario:

- The department has limited budget and the web team have to justify extra resources to provide training and tools to improve accessibility.

Consequence:

- The budget for the web team is not increased and accessibility is not properly implemented in order to add the new content needed within budget.

Recommendations:

- Quantitative accessibility data produced at the Compile Content (see section 7.2.2), Quality Control (see 7.2.3) stages combined with the workload represented by the accessibility log can be used to justify an increased budget to employ Gurus or run extra training courses.

6.2.6 Accessibility tool support

The results discussed in this section relate to sections 5.2.4 and 5.4.7

Providing tool support for accessibility to those involved in web development can help improve the efficiency and accuracy of accessibility assessments but it is not a complete solution to the problem. Currently, the majority of the guidelines specified by the W3C are only verifiable with human judgement. However, accessibility tools can serve to remind content authors about potential barriers such as missing meta-data or the correct use of HTML. In recognition of web accessibility as a legal and moral obligation, the interest and investment in tools to facilitate improvements will increase. In section 6.2.5 the lack of a specific budget was discussed and Figure 5-8 shows one of the consequences of this lack of budget. 54% of respondents from the WMP survey used an online tool as part of accessibility assessment (44% relied entirely on online tools for accessibility assessment). This is more than double the 20% who used only standalone tools in assessments. One reason for this could be the financial costs of purchasing a standalone tool. Online tools are predominantly free of charge and are therefore an obvious choice when there is little or no budget to purchase standalone tools. But online tools have many drawbacks, they contain less functionality than a standalone tool, they are operated over a network on a shared processor and are subject to network and processing latency. Because they are free and have limited resources, most restrict the number of pages which can be assessed within a certain time frame. All these problems mean that using an online tool for a larger scale assessment is impractical. It is for this reason that none of the specialists in the ASP survey relied solely on online tools (see section 5.2.4) and all had access to standalone tools. This is understandable as an important part of their role is to assess many web pages (see Table 5-3) and as such the time delay imposed on them by online tools is impractical. In addition to this, accessibility specialist will have more need for customisation in terms of which guidelines they use to assess web pages and in the reporting facilities. Since they are offering a professional service they may want to produce reports with their own branding. 58% of the accessibility specialists stated they used a combination of online and standalone tools. Online tools might be used to show potential customers the kinds of barriers that exist in their web site. There are well known and respected online services available which can be accessed via a web site. For example, in an initial meeting the specialist might use an online tool to assess a client's homepage to highlight some of the accessibility barriers contained on the web page.

University department accessibility improvement

- A university department wishes to carry out an accessibility assessment of their web site before the new academic year begins.

Problem Scenario:

- The department has limited time and only one member of staff available to perform the assessment.

Consequence:

- The staff member doesn't know where to start with the assessment and which pages to address first, as all are considered of equal importance.

Recommendations:

- The department invests in a semi-automated accessibility tool which analyses the whole site and ranks pages based on the number of potential accessibility barriers detected.
- The staff member now knows which pages potentially contain the most serious barriers for users and has assistance from the tool to view the source code and identify the potential accessibility barriers.
- By prioritising the repair of pages, the staff member can make the best use of their limited time.
- The staff can consult and amend the accessibility log so that the knowledge gained from the assessment is retained within the organisation.

6.2.7 User awareness

The results discussed in this section relate to sections 5.4.5 and 5.3

Web sites are created to provide information and services to users. Without users there is no reason for the web site to exist. Hence, it is important that a web site fulfils the needs of its user base. This is especially true for commercial organisations' web sites as users are also potential customers. If users are dissatisfied with a web site either because it has poor usability or accessibility then they are unlikely to purchase goods or gain a positive impression of the organisation. For public organisations, although there is no financial incentive to ensure usable and accessible web services, there is a public duty and (as mentioned in section 2.2.5) legal factors in the implementation of accessible web sites. User feedback has become an important driver for accessibility and usability improvement [82]. Even after organisations

establish policy and practices to ensure web accessibility for the majority of cases, there will always be exceptional cases wherein some users are faced with barriers unforeseen by the web development team. 67% of the organisations in the WMP survey (see Figure 5-7) tracked user feedback. This is recognition that it is an important tool in the continual improvement of a web site. Without user feedback tracking an organisation may be unaware of potential problems within their web site or miss opportunities to improve their site. One of the traits of accessible organisations put forward in section 5.3 was their level of exposure to individuals with disabilities. Organisations may lack either the resources or the foresight to carry out testing with individuals with disabilities. But by tracking user feedback, any problems raised by users with disabilities in using the site will contribute to this exposure and help raise accessibility awareness amongst content authors and web maintainer s. The following scenario is based on an actual event experienced by the author. It was one of the motivating factors in pursing this research topic.

University academic office publishes anonymous marking codes

- The academic office of the university decides to publish anonymous marking codes (used to identify students on exam papers in place of names).

Problem Scenario:

- The academic office publishes the anonymous marking code using HTML presentational tags `<i>` to italicise characters.
- Students with certain visual disabilities find it difficult to separate the characters. Since each character in the marking code is important, they cannot just guess their code.
- There is no explicit form of user feedback tracking available on the web site.

Consequence:

- Some students find it impossible to read the code and do not know how to resolve the problem or report the issues.
- The academic office receives numerous telephone complaints which waste clerical time.

Recommendations:

- The academic office ensures that there is a proper system of user feedback tracking so users know how to submit feedback and also that the content author is made aware of feedback as soon as it is submitted.
- The accessibility log is updated stating the problem and the solution found by the content author. In this particular scenario the solution was to remove the presentational tags and stylise the text using more accessible style sheet controlled techniques.

6.2.8 Coping with long term change

The results discussed in this section relate to section 5.4.6

It is in the best interest of an organisation (which publishes large amounts of information and services on their web site) to actively seek out new tools and methods to assist them. They should also be aware of changes in best practice or industry standards to ensure they keep pace with the evolution of the Web. This long term view was missing from the organisations in the WMP survey. Only 20% of the organisations involved in the WMP survey had procedures in place to monitor the latest accessibility tools. The percentage monitoring the latest requirements and guidelines is higher (55%) but still small in comparison to the percentage of organisations which have established a clear policy on accessibility or specified which guidelines against which their web site will be assessed. Such a low level of implementation reveals a passive attitude towards accessibility; organisations have reacted to their legal obligation but have failed take a proactive stance and exceed their legal obligations. This would appear to agree with the 75% of accessibility specialists (see Figure 5-3) who cited that legal requirements were amongst the main reasons for an organisation to address accessibility. It might be argued that this is, for some organisations, all that can be expected as it is not their remit to develop and investigate the latest accessibility goals and technologies as they do not have the skill set required to evaluate tools properly or digest the technical specifics of a W3C publication. In this case the only goal of the organisation is to react to changes after they have reached a certain threshold that they appear in print or in legislation. Where there is an interest in taking a long term approach, and exceeding current recommendations (which in the case of the W3C guidelines [26] where published in 1999 and can be considered quite old in terms of the Web) it is important that

organisations monitor and utilise the latest tools and guidelines. This could partly be addressed by ensuring that accessibility gurus are proactive in monitoring trends in accessibility standards and tools. Such gurus can help organisations prepare for any changes that need to be made to conform to the latest accessibility guidelines.

As well as enabling organisations to anticipate changes in accessibility technology and policies necessary to conform to new guidelines, by paying attention to the best practice either through formal scientific publications or (and perhaps more likely) through industrial press, organisations can gain new techniques and approaches to providing accessible web sites.

University Oriental Museum Online Tour Scenario

- The University Oriental Museum's web site provides an on-line tour of latest exhibitions.

Problem Scenario:

- The museum wishes to create an on line **aural** tour of their latest exhibition for users with visual difficulties. This aural tour will vocalise interesting exhibitions, providing descriptions of pieces and giving supplementary background information.
- The Oriental Museum is not sure how best to design the aural tour to ensure accessibility with screen reader software. They are unfamiliar with designing non-visual based web sites.
- There requirements for aural web pages are poorly addressed by current accessibility guidelines [12].

Consequence:

- The aural tour web site is designed based on the traditional visual requirements and does not provide an adequate aural tour for those using alternate browser technologies.

Recommendations:

- Content authors raise the issue of aural tour provision with management (see section 7.5.1.6)

- Management appoint Gurus to focus on current web accessibility research and discovers recommendations for the design of aural web sites [12] (see Section 6)
- These findings are reported back to management and the new recommendations are adopted.

6.3 *Evaluation of the impact of practices on accessibility*

6.3.1 General accessibility assessment

The results discussed in this section relate to section 5.5.1

Both the manual and automated assessments revealed that a majority of web sites were classified as Very Accessible (automated 64% and manual 49%). The discrepancy between the manual and automated assessments reveals the risks in relying solely on automated assessment methods. In order for a web site to be considered Very Accessible (rather than the lower category of Accessible) it had to address more than the removal of barriers specified by the W3C guidelines [12]. The presence and appropriateness of these enhancements required human judgement. For example, during the manual cross-browser assessments of the web sites from the WMP survey, each page was viewed using a text-only browser. The accessibility of the page was greatly increased by the inclusion of an internal hyperlink allowing the user to skip past the navigation links (this appears in the Section 508 checklist see section 2.3.1) at the top of the page to the main content of the web page. Without this link the user is forced to navigate by pressing the "tab" button on the keyboard through all the navigation links one by one for every web page. Hence, if this link was missing it was considered as a negative aspect during the manual assessment. However, it was ignored in the automated assessments, as there is neither specific HTML for this internal hyperlink, nor a standard text which must be used it is impossible for a tool to confirm the presence or absence of the skip navigation link.

However, there may be circumstances where accessibility work has to be prioritised due to budget or time constraints. These PageMeasures can be used to rank pages in the order of which have the most serious accessibility barriers. The Tree-Map tool can be used to visually identify pages which require the most urgent attention and work can be assigned accordingly.

Another difference between the manual and automated assessments was the proportion of web sites categorised as Inaccessible. The automated assessment categorised 25% of the web sites as Inaccessible which was considerably higher than the 11% in the manual assessment. The explanation for this is the presence of deprecated or bogus HTML elements or attributes. Deprecated elements and attributes are those which are being phased out by the W3C and bogus are those that were never part of any HTML specification. During the automated assessment each instance of a deprecated attribute or element contributed to the Page Measure and as the Page Measure increases so the accessibility category is lowered. However, after careful consideration and cross browser testing deprecated HTML was not considered to present a real accessibility barrier as although not ideal, most browsers will be able to cope and ignore deprecated HTML. Also it possible that for some pages the HTML wasn't deprecated when the page was created. Since the results of the manual assessment were to be revealed to the respondents it seemed unfair to consider this as part of the assessment. Instead it was included in the general comments of the assessment report (see section 4.6.2) with a reference to the W3C recommendation to avoid its usage. However, this was not the case for bogus HTML; bogus HTML represents a clear error in the page creation or ignorance on the part of the content author.

6.3.2 Practices in isolation

The results discussed in this section relate to section 5.5.3.2.

Only two practices which when considered in isolation were found to have a significant impact on accessibility. They were as follows:

- The specification of accessibility guidelines.
- The monitoring and recording of accessibility data.

These two practices represent maturity extremes in organisations. Specifying which accessibility guidelines are to be used is one of the first steps an organisation must take when implementing accessibility improvement. The absence of specified guidelines is symptomatic of an organisation with little awareness of, and little

maturity in, accessibility. This lack of maturity is likely to lead to an inaccessible web site as many accessibility barriers will be missed or ignored. Only a small minority (5%) of organisations had failed to specify accessibility guidelines and these organisations had web sites which were substantially less accessible than the majority of the organisations which had specified guidelines. Organisations without specified guidelines were substantially less accessible than any other group. At the other end of the accessibility maturity scale are those organisations which have set up procedures to monitor and record accessibility data. Table 5-13 shows this practice is never implemented in isolation but rather as part of a combination of practices. Monitoring and recording can only occur when there is accessibility data available. Therefore, if an organisation monitors and records accessibility data it must first have implemented other practices. In itself the practice indicates a heightened appreciation on the part of an organisation to monitor accessibility continually and a commitment to install the tools and procedures to record the data. By implementing this practice an organisation can easily monitor where accessibility barriers occur and record any improvement or decline in accessibility. With this information readily available management can gain a better picture of areas of the web site which are failing or succeeding. This information then dictates where and which resources are allocated. Management are also more willing to commit resources if reliable quantitative data is at hand. Records of accessibility enable management to review how changes within the organisation have affected the accessibility of the web site. One other interesting result was where accessibility was integrated into the web publishing process; web sites were much more accessible. However, despite this improvement not proving statistically significant it is reasonable to assume that given a larger sample size this difference would prove significant.

6.3.3 Practices combined

The results discussed in this section relate to section 5.5.4.

Other than the two extreme situations discussed in 6.3.2 accessibility was only found to have been improved when organisations implemented a combination of the practices. The practices surveyed have inter-dependencies. Without a clear and comprehensive policy and specific guidelines against which accessibility is measured there would be nothing to measure and monitor. Neither is the genuine support of management effective if there is not the expertise, or training provided to gain the

expertise present within the organisation to create accessible web content. Therefore no single practice by itself can improve accessibility, rather the practices shown in section 5.5.4 must be implemented together as part of an overall process improvement strategy. Chapter 7 gives recommendations for an accessibility process improvement strategy and highlights how it can be integrated into traditional web publishing cycle.

6.3.4 Raising accessibility awareness

Whilst it is possible to create and maintain accessible web pages without being aware of accessibility, it is very unlikely. Even with software (such as a content management system) that supports the creation of accessible content, it is always possible that if web developers are ignorant of potential web barriers they are more likely to include them in a web page. Because only one person per organisation completed the WMP survey it was not possible to accurately measure the levels of awareness within each organisation. However, it is reasonable to assume that greater awareness of accessibility should lead to the creation of more accessible web sites. The practices which are considered to have the most affect on accessibility awareness are:

- Availability of training.
- Promotion of accessibility.
- Availability of accessibility gurus

6.4 Tree-Map Based Tool Support

This section evaluates the results of the Tree-Map experiment. Its purpose was to establish whether such a tool would provide an accurate and efficient support for web maintenance tasks. Three ISO 9126 software quality sub-characteristics are now evaluated: Accuracy (sub-characteristic of Functionality), Learnability (sub-characteristic of Usability) and Time Behaviour (sub-characteristic of Functionality). There is now a description of how the Tree-Map tool was assessed in terms of each of these software quality sub-characteristics.

6.4.1 Functionality: Accuracy

One of practices covered in the WMP survey was the integration of accessibility into web publishing, monitoring and reporting of accessibility data. Hence, the experiment covered a wide range of web maintenance activities to test whether it could be used as part of an integrated approach to web accessibility. As such, the accuracy of the tool

for all activities including accessibility was measured; this also reflects that although important, accessibility is not the sole focus of web publishing. Accurately measuring accessibility is difficult because of the number of subjective barriers i.e. there is no definite wrong or right answer. However, for the purposes of the experiment, the automated assessment tool used to generate the reports and Tree-Maps was considered as 100% accurate since the participants were unable to perform any supplementary assessments to confirm or reject its results. The accuracy results of the experiment do not show a significant difference based purely on the tool support available. Participants who performed the HTML based report tasks first achieved a mean accuracy of 54%, this score significantly improved in the second Tree-Map based tool tasks by 23% to a mean accuracy 77%. However, participants who carried out the Tree-Map based tool tasks first had the same level of accuracy as those who carried it out second (77%) and then achieved a similar level of accuracy for the HTML report based tasks (74%). This suggests that once participants achieved a certain level of competence they performed consistently regardless of tool support. However, it appears that the Tree-Map based tool supported participants better in learning how to complete the tasks accurately.

6.4.2 Usability: Learnability

Overall, there is no significant difference in accuracy. However the ease of learning of the Tree-Map based tool is a significant improvement on the HTML report based tool and so could be used by less technical staff. Organisations require a tool which is both accurate and easy to learn so that it can be used by staff (especially management), who may not be comfortable with accessibility guidelines or technical reports. HTML based reports alone requires consideration of accessibility guidelines, the confidence factor given by the tool and the degree of severity of each barrier detected. The Tree-Map based tool abstracts away from this detail by presenting the PageMeasure within a Tree-Map visualisation. Hence it does not require such specialist knowledge for general maintenance information or for the detection of barriers. The user requires a basic knowledge of the tool's functionality (such as locating and highlighting pages) and how to interpret the size and colour of the Tree-Map nodes. One of the main advantages of the Tree-Map based tool is that its visualisation can be exported to reports easily. With a small amount of explanation it can be used to accurately convey

accessibility information to non-technical individuals. It can provide answers to the following questions:

- Does the web site have inaccessible web pages?
- How many pages are there on the web site and how are these organised?
- Which pages require the most attention to improve accessibility?
- How often are pages updated or assessed for accessibility?
- Who maintains which web pages?
- What are the accessibility barriers found in the web site?

The results reported by the Tree-Map based tool as well as the HTML based reports required a degree of interpretation by the participants. Using the Tree-Maps, participants compared the colour of the web pages (represented by nodes) using the colour legend provided. For the HTML based reports participants compared the list of barriers detected, the confidence factor given by the tool and the number of instances of the barrier found on the page. Other maintenance data such as the date of the last accessibility assessment, the name of the content author for specific pages or the number of pages contained with a section of the web site were objective and hence reliable.

6.4.3 Efficiency: Time Behaviour

Results from the Tree-Map experiment show that Tree-Map based tool gave a significant time saving of over 3 minutes. When the mean time to complete the tasks is examined (approx. 16 minutes) this shows an overall improvement of 19% in the time taken to complete tasks. There was also no significant loss of accuracy as a result of this time saving. Web maintenance is a continuous and time intensive activity [40] [72] and so any time savings made without reducing the quality of maintenance will have substantial long term benefits for an organisation. Reducing the amount of time web maintainers and content quality managers spend on corrective maintenance such as accessibility improvements, allows more opportunities to add new functionality or update the web site. Maintaining an up-to-date web site is one of the requirements set by Cuenca [37]. This requirement is further supported by the addition of web page meta-data to the Tree-Map visualisation. Displaying meta-data within the visualisation means that users do not have to open individual web pages and locate the

specific information in which they are interested. Adding meta-data to the Tree-Map based tool also provides users with the ability to view all pages within the site which contained specific meta-data. For example, it is possible to highlight ***all*** pages maintained by a specific content coordinator, assessed by a content quality manager or contain content which has passed its expiry date. This global overview allows patterns and trends to be detected, especially when combined with searches. Several examples of how this might be applied are now given.

Example 1: Failing content coordinator

If management suspect that a content coordinator is not updating their content regularly enough, they can highlight all pages maintained by that individual and then highlight all web pages containing expired content. The overlap between these two sets of pages shows quickly and clearly *all* pages which have expired content belonging to that content coordinator. The advantages are that detecting this information is quicker than searching through the web site and shows pages from the entire web site in one report.

Example 2: Identifying sections of the web site with poor accessibility

The tool allows management to quickly spot sections of the web site that have potentially poor accessibility. To confirm this, management can use the tool to focus on the more detailed accessibility reports and if necessary contact the content coordinator or quality manager responsible for that section of the web site. Also, by comparing the Tree-Map visualisations generated either in the past or from older versions of the web site management it is also possible to identify sections of the web site with potentially improving or worsening accessibility. If a section is getting worse then more training might be required, whereas improvement is indicative that current practices and staff training are sufficient and best practice could be promoted to less advanced content coordinators.

Example 3: Verifying implementation of new guidelines

An organisation might decide to implement new guidelines governing the provision of meta-data for web pages. For example, an organisation which has adopted the Dublin Core Metadata Terms [59] can highlight all web pages which do not contain the required meta-data terms and contact each of the content coordinators responsible.

6.4.3.1 Features of the Tree-Map tool that allow shorter task completion times

The features of the Tree-Map based tool which supports all these examples and are those which allow users to complete their maintenance tasks correctly and in less time are:

- ***No requirement for searching*** - Common maintenance information has already been acquired by the tool and is immediately accessible.
- ***Fixed global overview of the web site*** - All web pages are viewable and sections which the user might not even know about (e.g. recently added sections) will be viewable. Since Tree-Maps are a fixed space visualisation there is no requirement for scrolling.
- ***Focus on demand through zooming*** - Should the visualisation of the whole web site provide too much information; the user can zoom in on the sections of interest.
- ***Visualisation makes use of preattentive features*** - Viewing the results of the maintenance queries can be spotted at a glance with little effort.

6.5 Summary

This chapter has evaluated the results from the ASP and WMP surveys and the effect of publishing practices on the accessibility of the WMP respondent's organisations. It has also evaluated the results of the Tree-Map based tool experiments. It found the following:

- There is an increasingly sophisticated approach to web maintenance and content management with over 90% of the organisations surveyed using or planning to use a CMS.
- According to those who specialise in web accessibility, there are still organisations which do not value the worth of web accessibility and web maintainers who view it as a roadblock.
- Evidence of poor training provision was also found with only 30% of the organisations surveyed in the WMP survey having a formal written traning policy.

- There was however, a note worthy trend towards providing gurus for accessibility (70% of the organisations used accessibility gurus), which may denote a less formal and more practical form of training / knowledge transfer.
- The lack of formalised training and use of gurus might be explained by the lack of a specific budget for accessibility work made available within 94% of the organisations surveyed in the WMP survey.

From the results of the WMP survey and the follow up accessibility assessment the following conclusions can be made:

- There was a generally good level of accessibility amongst the web sites of the organisations surveyed (64% were assessed as Very Accessible).
- No meaningful stastical link can be made between one specific web maintenance practice in isolation and better accessibility. Although the lack of any specific guidelines was found to produce poor accessibility and those organisations who recorded and monitored accessibility data had more accessible web sites. However, to have accessibility data to monitor and record other practices must first be in place.
- There was, however, a statistically significant link between the web sites who implemented all the web maintenance practices related to accessibility surveyed and better accessibility.
- The Tree-Map based tool was proved to offer significant gains in the efficiency of web maintenance tasks without any effect in accuracy. And results also indicated that if offered an easier to learn interface than report based web maintenance tasks.

7 Process Improvement Recommendations

7.1 Introduction

This thesis has focussed on how specialists approach and view web accessibility. The web publishing practices implemented by organisations have been surveyed. Following this, the impact of these practices on overall web accessibility has been evaluated. Finally, the feasibility of a Tree-Map based tool for web site maintenance was tested. This chapter will now present the following:

- A novel adaptation to an existing web publishing model.
- Process recommendations based on the results from the WMP and ASP surveys and the Tree-Map based tool evaluation.
- Discussion of how tool support can be used within an organisation to improve web accessibility.

7.2 Integrated Accessibility Publishing Model

The majority of organisations (79%) in the WMP survey have already integrated web accessibility into their publishing practices. This integration then should be as unobtrusive as possible. Accessibility tasks should be a logical extension of existing procedures. For this reason, the recommendations contained in this chapter are set around a generic and traditional web publishing model. The various tasks have been assigned at points in the lifecycle where they are most appropriate; this reduces the effort and training needed to make the transition from a traditional publishing model.

7.2.1 Accessibility Integration Points

The approach taken to integrate accessibility into a web publishing model is to adapt Weinstein's content lifecycle [113]. Accessibility integration points have been added to the lifecycle. These represent points where accessibility can be assessed and if necessary improved. Adapting a current and simple publishing model has the advantage that many organisations will already recognise it and are more likely to accept the adaptations. The main aim is to include accessibility related tasks in an unobtrusive and logical way. There are no additional stages or changes in the data

flow of the model. Figure 7-1 shows this adapted model with the accessibility integration points labelled with the letter 'A' and a number.

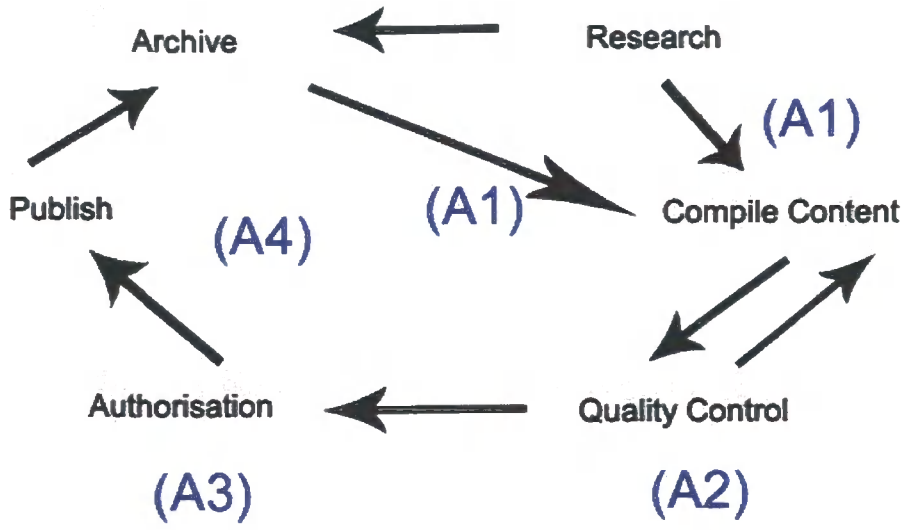

Figure 7-1 Adaptation of Weinstein's content lifecycle model [113] with accessibility points labelled

A brief description of each stage is given in Table 7-1 followed by a description of each group of tasks. For each stage, details of which tasks can be achieved through automated, semi-automated or manual means are given. Tasks are organised under the headings: Automated tasks, Semi-automated tasks and Manual tasks.

Stage	Description
Research	The first stage of content creation is to gather the content that will form the web page. This content can be obtained from a variety of sources. Information might, for example, come from archived material. For example, for news item content from previous articles about related stories might be reused.
Compile Content	Once all the content has been gathered, it must then be compiled into a web page. This might be performed using a text editor, or through content management software.
Quality Control	After the web page has been compiled it has to be assessed before it can be authorised for publication. For example, compliance with corporate standards, accuracy, quality of language and spelling are all quality issues which are addressed at this stage.
Authorisation	Once quality control has taken place, a member of the organisation with the correct level of authorisation can authorise publication of the web page.
Publish	The publish stage is automated where web page(s) are moved to the live web server.
Archive	By incorporating an archiving stage, organisations maintain a history of their web site which can be referred to if there are queries about older web content.

Table 7-1 Description of each stage of the adapted publishing model.

7.2.2 Tasks at Compile Content Stage (A1)

Content compilation is the most important part of creating accessible content. If content is created without considering accessibility, then it will create a number of accessibility barriers which must be addressed later on in the content lifecycle. The WMP and ASP surveys revealed that some content creation tools hindered, and in some cases prevented, the creation of accessible content (see question 4 of section 5.4.5). Preventing the inclusion of accessibility barriers during content compilation stage is therefore critical. The WMP survey showed that a majority of organisations (90%) used content management systems (and their associated content authoring

tools) to author their web site. Hence, it is essential that this software enables and encourages the creation of accessible web content. By producing web pages with fewer accessibility barriers the amount of times required to assess accessibility is reduced. Accessibility assessment is both time consuming and laborious and is often prone to mistakes. This in some cases has led to the perception (described in question 1 of Section 5.3) that accessibility work is seen as a roadblock. Accessible content authoring tools enable content authors to create content without having to constantly check if accessibility barriers have been inserted into the content. One issue arising from the ASP survey is that these tools must be used appropriately. If content authors do not have sufficient training they can inadvertently create accessibility barriers regardless of the quality of the authoring tool.

7.2.2.1 Automated tasks

Some tasks at the Compile Content Stage can be automated:

- HTML validity can be assessed automatically. Any syntax errors found can be corrected using an automated validation tool. This then reduces the amount of effort in the Quality Control stage (A2).
- An automated accessibility assessment should be executed at the end of each significant change to the content. The assessment reports barriers detected in the content to and allows content authors or coordinators to either ignore or edit the content. This is of particular use for older content, created before an accessibility policy was introduced and is more likely to contain accessibility barriers. Once the content has been updated it is archived and hence has the side-effect of improving the accessibility of the content gradually, avoiding the need for a complete overhaul.
- Quantitative accessibility data should be collected and logged. This thesis recommends the following original measures:
    - Number and type of barriers detected in the content by which tool.
    - Number and type of barriers ignored by content author / coordinator.
    - Number and type of barriers which are removed by content author / coordinator and the time spent on correcting them.
    - Whether the content came from the archive and what changes were made to it.

- Number of times and types of invalid HTML produced by authoring tool as consistently poor HTML indicates an error in the authoring tool.

Recording this data allows organisations to monitor the types of barriers that are being produced by their content authoring software. Barriers which occur regularly reveal either a problem with the tool used to create them, or training provided to content authors. If a specific type of barrier is continually ignored (i.e. the content author deems it not to be a barrier), this indicates that either the automated assessment tool returns a high proportion of false-positives, or the content authors are not fully aware of its consequences and require further training. The following example shows how the ignorance of content authors can lead to accessibility barriers being inserted into web pages even if assessment tools detect them.

Example of an actual accessibility barrier is detected, but incorrectly ignored by a content author.

During the compilation of content, the automated assessment detects cases of inappropriate HTML tag usage (such as tags being used to achieve the appearance of a heading (see Figure 7-2), instead of the h(1-6) (see Figure 7-3) tags which are specifically designed to denote a heading in HTML.

```
<strong>Summer Programme</strong>
<p>The summer programme is an exciting one...</p>
```

Figure 7-2 Incorrect usage of strong tag to mimic heading style.

```
<h1>Summer Programme</h1>
<p>The summer programme is an exciting one...</p>
```

Figure 7-3 Correct use of heading tag to semantically define a heading.

In this example the content authors may not be aware that they are misusing the `<strong>` tag (which is designed to add stronger emphasis to a word or phrase) and so ignore the barrier flagged up by the assessment. During a review of the automated

assessment logs it would be discovered that the barrier was reported to the content author but they chose to ignore it. Hence two issues are raised:

1. The content authoring tool should discourage the use of strong tags for headings.
2. Training is required to ensure that content authors understand the difference between misusing strong tags to mimic the appearance of a heading and using the semantic heading tags which adds structure to the document.

The WMP survey showed that only 30% of organisations had a specific training policy for accessibility and only 59% provided any training at all for accessibility. If organisations are to take full advantage of assessment tools then training must be provided to ensure their results are interpreted properly.

7.2.2.2 Semi-Automated tasks

The semi-automated tasks for the Compile Content Stage are now described:

- Meta-data for the content must be created, or in the case where content is reused, validated. This is semi-automated as some meta-data can be automatically completed (e.g. the date of creation and content author identification). The meta-data which must be included should be specified by the organisation's accessibility policy and will be presented to the content author after each major change to the content. Some fields may be optional whereas others will be mandatory.
- In addition to the quantitative data recorded, an accessibility log should be completed by the content author or coordinator describing any points of interest or unresolved issues. This provides a record of the content's lifetime and forces the content author to consider accessibility and the actions they have taken. Section 7.3 gives a description of the usage and benefits of the accessibility log.

7.2.3 Tasks at the Quality Control stage (A2)

The Quality Control stage addresses aspects of content, such as language style, accuracy and readability. These are all related to web accessibility and it is therefore sensible that the main accessibility assessment takes place at this stage. Many of the

manual quality control checks carried out at this stage will be a simple extension of existing checks.

7.2.3.1 Manual tasks

By ensuring that many of the easily avoidable accessibility barriers (such as misuse of HTML tags) are filtered out in the content creation stage, the quality control stage focuses on the less obvious and more subjective accessibility barriers such as:

- **Clear and simple language**. Whether the language used within content is clear and simple is highly subjective and requires human interpretation [77].
- **Appropriate use of alternative text for graphical images**. Text used as an alternative for an image should be appropriate to the meaning of the image. If the image is therefore decoration then there is no value in adding text to describe it. For example, many web sites use images to create a visual effect on a web page (e.g. rounded corners for rectangles). These images should have a blank alternate text (i.e. white space), which allows non-visual browsers to ignore the image rather than present the user with a meaningless description such as "right corner of main logo" or in the case where no alternate text is provided, the physical name of the file.

Assessing these barriers requires not only technical expertise but also an understanding and appreciation of how an organisation wishes to communicate with the outside world. Accessibility is not the only requirement placed on the web site; as the WMP survey discovered it must also conform to internal guidelines (see Section 5.4.7). Content may need to be checked to ensure the tone, language and terminology used in the page meets is appropriate for the target audience of the web site. Content quality managers should be trained to ensure that they possess the knowledge and experience required to assess content against branding and accessibility guidelines.

Quality Control must also cover aesthetic aspects, such as how the content appears on the page. For accessibility, it is important that when content is included in a web page it is previewed using several different web browsers to ensure that presentation and usability is consistent across web browsers. Users can access web sites using different browsers and significant changes in the appearance of the site might cause confusion and distrust amongst users. If a user accesses their online banking web site using a

different browser and the page is displayed inconsistently it does not promote trust in the online service.

7.2.3.2 Semi-automated tasks

Semi-automated support is provided to the content quality managers by assisting with the classification and logging of accessibility barriers. Any accessibility barriers found should be logged. If more work is required on the part of the content author, this accessibility log will be used by them to locate and remove the barrier; else it will be logged and used for later analysis. Logging of information with time stamping is useful for future cost estimation. By tracking specific types of barriers, such as inappropriate alternate text for a decorative image or unclear language, the organisation can estimate how frequently these barriers occur and how long they take to be identified and removed. Logging also provides a method to check whether barriers have been resolved or not. Once a barrier has been resolved this can be recorded in the log. This then allows automated checking that all barriers have been resolved in the Authorisation stage.

7.2.4 Tasks at Authorisation stage (A3)

The Authorisation stage is used to check the content has been subjected to the accessibility checks required by the organisation's accessibility policy. Once the content has passed the Quality Control stage it is ready to be authorised for publication.

7.2.4.1 Automated tasks

The Accessibility log is checked to ensure that all the barriers identified during assessments have been inspected and have either been removed or marked as irrelevant.

7.2.4.2 Manual tasks

All web pages containing content must have *valid* meta-data associated with them. The Authorisation stage is used as a final check to ensure that the meta-data such as the content expiry (i.e. when the content should be published until) is correct. If a web page contains information about undergraduate courses for the 2007 academic year, then the expiry date should be set to an appropriate date. So too should meta-data such

as description and keywords, giving an accurate representation of the page's content. This extra meta-data provides more information for search engines and is particularly useful for the users of assistive technologies for whom reading web pages is extremely time consuming. A short and accurate description can help the user decide if they wish to visit the web page or not and hence time is saved.

7.2.5 Tasks at Publish stage (A4)

The final stage of the content lifecycle is the Publish stage. At this point all the accessibility tasks should have been completed.

7.2.5.1 Manual tasks

There is only one task at this level which must be done manually. Content quality managers must ensure that for non-textual content, that textual alternatives have also been published and that they relate to the original. For example, if a video clip of a speech by the Chancellor of a university is published, then a full transcript of that speech should also be available.

7.2.5.2 Semi-Automated tasks

One semi-automated task is to verify every page has been published with the correct meta-data and accessibility credential certification. Web sites maintainers may wish to display the logos of which accessibility guidelines to which they conform.

7.3 Benefits of the accessibility log

The accessibility log is designed to keep a record of activities throughout the content lifecycle and serves three main purposes:

- To enable quantitative analysis of web accessibility.
- As a record of the work completed on a web site.

Such a log can be implemented using a simple SQL database. Stored queries can be created to fulfil the needs of the various roles, depending on the information they regularly need. It should be noted that while the accessibility log use is serparate from that of the know

7.3.1 Quantitative analysis of web accessibility

It is impossible to objectively assess the state of an organisation's web accessibility maturity without some form of objective and quantitative measurement. This thesis has shown how an automated PageMeasure can be calculated for a web page to give an estimate of its accessibility. Such a measure is not entirely reliable because there is no fully automated method to measure accessibility. However, this study has shown that accessibility assessments using a PageMeasure based only on the barriers that can be checked automatically has a strong correlation with manual assessments. In conjunction with the PageMeasure there are several other measures both by the web page and web site level that can be used by organisations to monitor accessibility. The novel web page level measures (apart from Web page size which is not novel) recommended by this thesis are as follows:

- **Number of different confirmed accessibility barriers found per page** – fixing different accessibility barriers will require different actions and hence more effort than repeatedly fixing the same barrier.
- **Web page size** – the size of the web page will have an impact on accessibility. If a web page is growing too large, it might be feasible to split it up into separated pages.
- **Date and number of updates in the last 3, 6 or 12 months** – shows how frequently the web page is updated. If a web page hasn't been update within the last year it may require reviewing.
- **Date and number of assessments in the last 3, 6 or 12 months** – if the web page has been updated 6 times in the last 12 months but only assessed after the first update then if could have changed substantially and hence should be reassessed.
- **Mean assessment time** – by recording assessment times a mean can be taken which can be used for cost estimation. It can also be used to detect deviations from that mean. If assessment times exceed expected time limits, this may indicate a problem with the assessment process or individual assessment. A resolution might be more training or more investment or in tools. If assessments are being completed in too short a period it may be evidence that assessments are not being carried out properly. Changes in the mean assessment time indicate the effect of a new assessment tool or policy.

- **Number of different and frequency of problems/barriers reported by users** – a large number of different problems reported for a single web page is an indication that it was created by either a person or tool without accessible design experience. If the same barrier is continually reported even after modification, then the assessment and correction techniques being used are insufficient.

If recording these metrics is automated and integrated it into the web publishing process then it will cause less disruption, and ensure that the data is recorded systematically, rather than on an ad hoc basis. If content authors and coordinators are required to record accessibility data explicitly it will add an unwanted overhead to each update and could lead to either content updating not being performed or accessibility assessments being abandoned through lack of time.

If management wish to monitor changes in web page accessibility then reporting on an individual web page or a small group of pages can be achieved through the creation of reports. However, if management want to maintain an overview of the whole site's accessibility then a diagnostic tool which gives a broader viewpoint of the site and its accessibility measures is required. Such a tool would support content quality managers who are unlikely (even with training) to be accessibility experts. A diagnostic tool which gives web maintainers or management a global view provides an extra level of management of the web site (see section 7.5). It allows changes in accessibility anywhere in the web site (either positive or negative) to be detected. When combined with details from the web publishing model for that organisation it can also be used to highlight trends and coordinate maintenance efforts. Question 5 of the ASP survey revealed that the specialists were reporting the quantitative results of accessibility assessments for the purposes of web site budgeting. Hence, the easier it is for them to retrieve this information the better their chances of presenting a convincing case for extra budgeting to management.

7.3.2 Record of work completed

The accessibility log provides a record of the work carried out on the web site. With database query support all measurements and details can be retrieved. A webmaster may wish to learn how many web pages were assessed by a specific content author in

a certain week. If there is a problem with the work from a content author it is possible to locate all content they have worked on and view the actions they have taken. This may also be used should any legal actions be taken against an organisation. Under British Law organisations are required to make "reasonable adjustments" [51] to their web sites to ensure accessibility. Organisations can use the log to check or prove that reasonable measures have been taken.

7.4 Knowledge Repository

As an aside to the accessibility log, a repository of knowledge should be kept. This repository should contain human readable descriptions of actions taken in addressing accessibility barriers form the basis of a knowledge repository. Documenting the actions carried out to remove different types of accessibility barriers provides a means for organisational learning. It will enable new content authors to learn from the experiences and mistakes of others as successful and unsuccessful approaches will be recorded.

7.5 Overview and reporting tool support

Maintaining an overview of the accessibility tasks and the content stages for the whole organisation requires tool based support. The Tree-Map tool developed as part of this thesis provides the following features needed to organise and monitor the integrated accessibility publishing model and the effect it has on accessibility:

- **Roles and responsibilities** – The content authors and quality managers assigned to content are stored and can be visualised as part of the tool.
- **Meta-data maintenance** – The tool can display the meta-data created for each web page without the need for viewing the HTML source of the page.
- **Accessibility logs** – Logs can be stored for each web page. This was not fully implemented in the prototype but could be easily achieved using an extension of the existing task list table in the database. Logs of assessments are stored in the database and provide a record of previous work carried out on the web site.
- **Quantitative accessibility measures** – Accessibility data are stored and used to generate the colour and dimensions of the Tree-Map. The tool already records some of and should be extended to record all the measures discussed in sections 7.2.2 and 7.3.

- **Overview of accessibility** – by generating a fixed space visualisation of a whole web site, an overview is given to a web master. (see section 7.2.4).
- **Integration with automated assessment tools** – The tool allows on-demand assessments of web pages. Content authors or web masters may be interested in re-assessing an older web page. This assessment is recorded and the results should trigger the appropriate actions. For example, if certain accessibility barriers (such as missing alternative text) are found, then either the person running the assessment or the content author should be made aware of the barrier and given a chance to remove it.

7.5.1 Tool support for multiple roles

Web teams are often small and members are usually tasked with a number of varied web maintenance roles (see Table 5-7). Approximately 30% of the organisations surveyed in the WMP survey had web teams with less than 6 members (see section 5.4.4). As such, any tool aimed at supporting the web maintenance tasks of these individuals has to provide the functionality, information and external tool integration such a varied job requires. This thesis has shown (section 5.6) that a Tree-Map based tool has the potential to offer significant time savings over current report based methods. By extending the kind of data visualised by the Tree-Map tool from accessibility data to more general analytical statistics the tool can be used to support many aspects of web site maintenance and analysis. There now follows a discussion of how such a tool currently supports, and also how it could be extended to support, the typical roles defined in section 2.1.4.

7.5.1.1 Webmaster / Web maintainer

The webmaster needs to have a general overview of the web site. The fixed space view provided by the Tree-Map visualisation enables the webmaster to gain this overview instantly. The following data of interest to a webmaster is provided:

Changes in the size and structure of the web site

New pages or groups of pages (for example a new sub-section of a web site) will change how the Tree-Map is sub-divided. File modification timestamps could be visualised using colour. For example, all pages updated on that day are shown in blue. With this information the webmaster can monitor how the

site's structure is evolving and identify sub-sections of the site which have not been updated for long periods of time. The Tree-Map based tool could also be used to create new pages or sub-sections within the web site.

Overall Accessibility

One of the responsibilities of the webmaster is to ensure that the site's overall accessibility is maintained to a high standard. As presented in Chapter 3 and Section 6.4, Tree-Maps can be used to accurately and efficiently represent accessibility. In line with the integrated accessibility publishing model quantitative accessibility data is automatically logged (see Section 7.2.2), so this can be visualised in a similar way to the PageMeasure used in the prototype. Integrating reports produced for manual accessibility checks such as those described in Section 7.2.4 allows webmasters and content quality managers to ensure tasks have been completed properly.

Maintainer and content author information and workload assignment

Webmasters often re-assign work to other members of the organisation and this tool allows them to see which members of their organisation are responsible for which web pages. This overview allows them to decide who in the organisation should be assigned the work. This decision could be based on how many pages the content author maintains already and whether the new pages are in the scope of that content author. Work can also be assigned based on the accessibility of certain web pages. For example, when inaccessible pages are identified by the Tree-Map, the webmaster can contact the content author responsible through the tool.

Web technology usage

To ensure interoperability and consistency the webmaster should be able to view a summary of the technologies used to create every web page. This information is useful for example if the web server configuration settings change. For example, in a university environment many departments may use a server side scripting language such as PHP or ASP on their web pages. These pages will be affected by any changes in configuration to the central

web server. As such, the web master will be able to highlight all pages that will be affected by the change and contact the appropriate content authors.

7.5.1.2 Content Coordinator, Authors and Content Quality Manager

Content coordinators and authors will typically be responsible for more than one page and these might be distributed throughout the web site. As such the Tree-Map based tool provides a useful method of organising the pages and should be integrated with the user's web page editing tool. This gives the content author a better idea of how their pages relate to the others in their website. If the web site is large, then the view presented could be customised to show only the sections of the site to which they contribute. Tree-Maps make this simple as they provide a zoom facility. The tool should also be integrated with the publishing model used (see section 7.2). One of the main features is to support communication within the content lifecycle between content author / coordinators and content quality managers. In particular the following activities should be supported by the tool:

Accessibility Maintenance

The Tree-Map tool can be used by another member of the organisation to inform content authors that an update has taken place. It can also highlight changes in accessibility. Shared access to the accessibility log (see section 7.3) will allow the content author to log information easily and therefore encourage usage. After major updates have been made to a page the tool should highlight pages where tasks which must be completed manually are still outstanding. This provides a reminder to the content author to complete the manual checks detailed in section 7.2.4. Content quality managers can also use this to ensure the manual checks are completed.

Meta-data Maintenance

Part of the integrated accessibility publishing model (see 7.2) is the creation of accurate and timely meta-data. By providing a preview of meta-data contained within the web page the tool ensures that these are completed properly and are kept up to date. Highlighting pages that contain specific meta-data (such as

keywords or description) would be a useful by enabling content authors to identify other potentially related web pages. Hence making content authors aware of other pages to which a hyperlink could be created. It may be a content author finds that a web page already contains the information they wish to publish and therefore saving time creating and reducing redundancy on the site.

Workflow Support

Ensuring that content authors and content quality managers are aware of each others' work is important to the efficiency of the publishing model. By including workflow support into the tool, content authors and coordinators can be made aware immediately of any comments or problems found by the content quality manager in stage A2 or A3 (see sections 7.2.3 and 7.2.4). By integrating this with a web page editing tool they can be sure that content is published efficiently.

7.5.1.3 Business Quality / Website Manager

The Tree-Map based tool can be used to present general web analytical data to the less technical management roles. These roles are primarily management based, focussing on issues such as: budgeting, achieving a bigger audience and possibly generating revenue from the web site. As such they are mainly concerned with how many visitors the site receives and how they interact with the site. The support the tool can offer them in their work is as follows:

Usage statistics

The purpose of most web sites is to disseminate information so the success of any web site is dependent on it gaining a large enough audience. The Tree-Map based tool should allow webmasters to visualise how frequently pages on their site are visited. Having this information presented on one screen will allow them to quickly identify pages which are attracting very little attention and hence may need moving or modification.

User Feedback Tracking

The importance of user feedback is discussed in section 6.2.7. The tool should highlight web pages which have received user feedback. The W3C's Implementation Plan for Web Accessibility [29] recommends that organisations:

- "Provide feedback mechanisms for users both within and external to the organization." [29] And that:
- "Feedback pertaining to accessibility must be made available to responsible team member." [29]

Integrating the user feedback mechanism with the Tree-Map based tool can be achieved through the use of an online form which is generated dynamically for each page. Each form is created for a specific web page so when user feedback is submitted through this form the Tree-Map based tool is notified and hence it can incorporate this information into the visualisation. This functionality is initially aimed at content authors allowing them to address any minor issues promptly. However, if the problem is more serious or requires attention from management then this can be passed to business quality managers. Managers should also be able to track the progress of user feedback through the tool. For example, if user feedback is not addressed within a specified time period, they may wish to contact the content author / coordinator responsible for that page.

7.5.1.4 Planning for the future

The web and the technologies that drive it, change and evolve rapidly. Many of the accessibility guidelines described in the first version of the W3C's Guidelines [26] have been out of date for some time. The challenge then for organisations, once they have established a method of consistently creating accessible web content, is to plan for future developments of the web and ensure all web technologies used remain accessible. To achieve this, awareness of the latest accessibility publications is important. Publications such as the guidelines and recommendations produced by the W3C and RNIB are used as reference materials by legislative bodies. It is therefore in the best interests of organisations to be aware of guidelines to which they may be obliged to conform in the near future. Monitoring web accessibility publications allows an organisation time to prepare for any changes in recommended practice. In this way, organisation practices can be changed over a longer period of time. The

extra time allows for better planning and internal promotion so that everyone within the organisation is aware of the changes and how they affect their roles. To monitor standards effectively, the latest tools and technologies aimed at supporting web accessibility should also be reviewed. Very few (19%) of the organisations in the WMP survey carried out reviews of new tools aimed at improving accessibility. By not implementing this practice they are neglecting potential time and cost savings. Accessibility assessment is a time consuming activity and more reliable semi-automated assessment tools or accessible authoring tools will reduce the amount of human assessment needed and also the number of accessibility barriers created by authors.

7.5.1.5 Tool review and selection

Accessibility support tools can save an organisation time and money. Therefore it is in an organisation's interests to have detailed knowledge of the tools available on the market, their features and their costs. In order to select the right tools for their needs the quantitative measurements discussed in section 7.2.2 and 7.3 provide the basis of selection criteria. Tailoring the selection criteria to tools that address the most frequently occurring and costly accessibility barriers gives management the confidence that the cost of purchasing the tool is justified and will yield measurable benefits. Once the tool has been installed and established, the effects can be observed by comparing measurements before and after the installation. This is ideal for organisations with restrictive budgets as they can perhaps trial the software and only if the effects lead to a significant improvement, purchase the tool. For example, if an organisation identifies that the HTML produced by their content management system is consistently incorrect; they can monitor the accessibility log and the measurements for the time taken to correct all web pages. Using this data they can estimate the total cost of fixing the incorrect HTML for the web site and compare it with the cost of either purchasing a tool to automatically repair HTML once it has been created, or (if possible) to commission corrective maintenance on the content management system.

7.5.1.6 Communication between developers and management

There must be communication between those applying technical solutions for a web site and those setting policy and budgets for the project. The presence of an accessibility guru was cited by the specialists in Section 5.3 as one of the factors they

noted about companies with good accessibility. 71% of organisations in the WMP employed an accessibility guru (see Section 5.4.6) and so it is relevant that these experts should have a responsibility to both improve accessibility and to communicate these issues to management. They then must inform management about any serious accessibility problems within the organisation and about developments within the web community which might affect future policies. For new accessibility guidelines to be established as standard practice or a budget increase to be sanctioned, management support is required. In order to get that support, management must be regularly informed about the state of accessibility. There are two factors which affect organisational policy:

- **Internal Factors** – the current state of the web site, costs incurred maintaining the current levels of accessibility.
- **External Factors** – changes in disability legislation, legal precedence, introduction of new guidelines and tools.

Management must be aware of both of these factors and it is important that the method used to inform them is at a suitable level of formality and regularity. A cross section of the organisations web development team should attend regular minuted meetings to discuss the organisation's web strategy. It is important that any actions agreed upon are noted and progress is checked. It is at these meetings that decisions regarding accessibility policy should be discussed and made. This ensures that both technical and non-technical factors are considered when policy changes are considered.

7.6 Summary

This chapter has presented the adaptation of an existing web publishing model to include points at which accessibility can be integrated. It has made a series of recommendations including:

- The adaptation of an existing and simple web publishing model with accessibility integration points were accessibility tasks can be inserted. Details of these tasks and how they can be achieved were also provided.
- Any adaptations to improve accessibility must be simple and not interfere with content management.

- The use of the accessibility log to record and monitor quantitative accessibility data. This will provide a method to monitor accessibility improvements and serve as an audit trail, should it be necessary to prove that the organisation addressed accessibility issues.
- Finally, recomendations were given for tool support such as:
    - Knowledge Repository – allowing maintainers to share the solutions they had developed with others within the organisation.
    - Multiple-role tool support – support for maintainers who may (as shown in both ASP and WMP surveys) embody more than one role within web maintenance.

8 Conclusion

8.1 Criteria for Success

In section 1.3, the criteria for success are outlined. This section now discusses each criterion and how the thesis has addressed it.

8.1.1 Survey current web accessibility best practice

The study described in Chapter 4 consisted of two surveys investigating specialist opinions on best practice and take up of best practice amongst web maintainers. The ASP survey took a qualitative approach and examined the opinions of 20 accessibility specialists from 8 different countries. The WMP survey reviewed the web publishing practices of 79 organisations. The ASP survey found a mixed opinion amongst the specialists as to web maintainer acceptance of web accessibility guidelines. The most important roles performed by the specialists were: accessibility assessment, training and raising awareness within organisations. 90% of the specialists based their assessments on version 1.0 of the W3C Guidelines, which indicates that these relatively old set of guidelines (1999) are still considered best practice amongst the specialists. Within the web maintenance community, best practice focussed on the updated version (2.0) of the W3C Guidelines, along with internal guidelines. The WMP survey revealed that there is an acceptance of best practice; however, those practices which required long term commitment or financial investment (recording and monitoring of accessibility data and budget allocation) were less popular. This leads to the conclusion that although there is a superficial acceptance of the need to produce accessible web sites, the commitment to invest significant time and money is less prevalent. Raising the profile of web accessibility and ensuring that organisations are willing to commit resources requires more than legislation. Although a legal requirement was seen as a motivation factor, this appears to have only had a limited influenced on web publishing practices. Better dissemination of the moral and business case for improved accessibility is required. The WMP survey revealed that relatively few organisations were actively researching the latest tools and standards so it is important that more is done to raise the awareness of web accessibility developments outside the web maintenance community.

8.1.2 Evaluate the effect of current web accessibility best practice and guidelines on accessibility

The analysis of the relationship between the implementation of the best practice surveyed in the WMP survey and the resulting accessibility of the web site revealed that there was no significant relationship between the implementation of most individual practices in isolation and an increase in accessibility. However, when all best practices specific to accessibility were implemented together, there was a significant increase in the accessibility of that organisation's web site. This justifies the conclusion that organisations must attempt to implement as many of the practices as possible to ensure an improvement in accessibility.

Integration of accessibility into the web publishing process did not significantly improve the accessibility of an organisation's web site. On reason for this is that perhaps what the organisation was integrating into the web publishing process was not sufficient. For example, if an organisation made just one small adjustment to improve accessibility but this was integrated into the publishing process, this is unlikely to have a large impact on the web site's accessibility. Individual practices which did have a statistically significant improvement on accessibility and hence can give an indication of maturity were as follows:

- **Guidelines governing accessibility specified** – this was revealed as a basic requirement for accessibility. Organisations where no guidelines were specified had by far least accessible web sites.
- **Monitor and record accessibility data** – by monitoring and recording accessibility data an organisation demonstrates a long term commitment to accessibility improvement. Monitoring allows changes in accessibility to be detected quickly and keeping a record of this provides historical information about whether the web site is improving or not. In order to implement such a practice a number of other accessibility practices must first be in place, and hence this practice is a good indicator for accessibility maturity.

8.1.3 Develop and evaluate a semi-automated web accessibility management tool

The Tree-Map based tool was developed to provide support for web accessibility management. It was evaluated through a controlled experiment. The results revealed

that although no improvement in accuracy (compared to a report based tool) was achieved, efficiency in the completion of maintenance tasks and in the ease of learning how to perform tasks was significantly increased. These are now discussed in detail:

8.1.3.1 Efficiency: Time behaviour

The Tree-Map based tool had a time saving of approximately 198 seconds (3 1/3 minutes) on tasks which were estimated to take a maximum of 20 minutes. It is believed that with further training and extra functionality (as the experiment used a limited functionality version of the tool for simplicity) this time saving can be further increased. If web accessibility is to be fully integrated into web publishing then it is important that usable and efficient tools are developed. The Tree-Map based tool is therefore a good candidate for further development from academic prototype into a commercially available tool.

8.1.3.2 Usability: Learnability

The experiment indicated that the Tree-Map based tool assisted in task learning and will therefore mean that new web maintainers will take less time to become proficient in their web maintenance tasks.

The features which allow shorter completion times and the ease of learning are:

- No requirement for searching to find web pages and assessing them.
- A fixed global overview of the web site.
- Support for focus on demand through zooming.
- Making use of preattentive features.

8.1.4 Develop an integrated web publishing model

This thesis has adapted the Weinstein's web content lifecycle [113] to show where accessibility practices can be implemented within an integrated accessibility publishing model. At each point in the content's lifecycle, recommendations have been made classifying which tasks can be achieved manually, semi-automatically and automatically. By providing this classification of tasks it provides scope for better tool selection and where human efforts must be more focussed.

The WMP survey revealed that Content Management Systems (CMS) are now used by a high percentage of larger organisations to create web sites and control content. One of the central roles of a CMS is to ensure that a specific publishing model is followed. Since the integrated accessibility publishing model is based on a standard content lifecycle it is reasonable to assume that given that the CMS software is sufficiently extensible, it can be modified to support this new publishing model.

8.1.5 Create recommendations of tool usage and systematic web accessibility improvement

This thesis has provided recommendations about how a tool can support multiple roles within web development. Management need different sets of data from those involved in technical aspects of web site creation. Any tool should support the communication between technical and managerial staff by allowing the presentation of accessibility trends at a higher level of abstraction as well as supporting more technical views. A Tree-Map can be used to both present high level information (through integration with automated accessibility assessment tools) and as the front end of semi-automated assessments. Chapter 6 presents further recommendations regarding an accessibility log should be used for knowledge transfer, allowing the skills and knowledge obtained during accessibility improvement to be captured and reused by others in the organisations, even if the author is no longer part of the organisation. By creating logs of completed work it is possible to ensure that best practice is being implemented. If measures such as time taken and the type of task are recorded then there is also scope for more accurate future cost estimations.

8.2 Further work

8.2.1 Short term development

8.2.1.1 Better empirical evidence for automated assessment

In the short term, more research is needed to empirically assess the automated PageMeasure presented in Section 3.2.4. The correlation between the results of automated web site assessments and the manual assessments (see sections 4.6.1 and

4.6.2) showed promise, but this was only carried out on a fairly limited sample size of 79 web sites with 10 pages per site.

8.2.1.2 Cost estimation

There has, to date, been very little overlap between the research topics of cost estimation and web accessibility. However, one thing organisations really need to know is how long accessibility work will take and therefore how much it will cost them. This thesis has proposed the possible use of an accessibility log as the basis for a cost estimation model.

8.2.1.3 Industrial case study for Tree-Map based tool

To authenticate the suitability of the Tree-Map based tool there should also be an industrial based case study. Such a case study should evaluate how well it supports the entire web publishing process, rather than only the selected web maintenance tasks conducted for this thesis.

8.2.1.4 Follow up study for web publishing practices

The WMP survey focussed mainly on UK based local government organisations with a small number of other domains. Further research into practices across a variety of domains is needed to gain better insight in their differences.

8.2.2 Long term development

So far web accessibility has been defined for users with a broad scope of needs and disabilities. However, further development is needed in customised accessibility. Just as some cars allow passengers to save their favourite car seat settings, web enabled devices and sites should allow the same customisation. For example, if a disabled user is at an airport and wishes to use one of the pay-per-use internet terminals it should automatically display web sites in the users preferred mode. To achieve this, a method is needed to inform the browser software of user display and language preferences. This might be achieved locally, via a USB device or remotely, via a specific URL containing the data needed. Related to this is the issue of language. There is no doubt that English is dominant on the web, however with the growth of China and other non-English speaking economies this might not always be the case. It is also arguably that relying on the fact that someone has to learn English to use the web is

discrimination. Hence, one important development will be the automated and reliable translation of web pages from any one language into another. Presently most automatic translation services are not very sophisticated and inconsistent. One step towards more reliable translations is the use of correct mark-up. Once this has been achieved, a service which could simplify and summarise information reliably would be of great value to users with slower internet connections and / or reading difficulties. Video is increasingly popular; without doubt one of the greatest successes in recent web history is that of *YouTube* (a video sharing web site) and a tool which could automatically add captions (or sub-titles in British English) would allow users with hearing disabilities to gain better access to online videos.

8.3 Thesis Summary

This thesis has attempted to address the diagnosis, improvement and maintenance of web accessibility.

8.3.1 Diagnosis

The thesis explored the use of automatic assessment measures of web accessibility and presented a novel application of Tree-Maps to visual web sites and highlight varying levels of accessibility. A survey of 79 large organisations was carried out with an automatic and manual accessibility assessment provided as feedback to those organisations. This information was used both to benefit those organisations and also to investigate the implementation of best practice and accessibility.

8.3.2 Improvement

An existing web publishing mode was extended to include accessibility points. This aimed to provide a simple and gradual transition to an accessible web publishing model. Accessibility specialists were surveyed for their experiences and perceptions of working on web accessibility projects. This was followed up the survey of 79 organisations web practices mentioned in the previous paragraph. Process improvement recommendations based on literature, author experience and results from the survey were presented and elaborated through scenarios. As part of this, the novel idea of an accessibility solutions log was developed. This solutions log is aimed at allowing individuals within an organisation to share information and experiences gained when addressing accessibility problems.

8.3.3 Maintenance

Developed Tree-Map based maintenance tool to located and repair poorly accessible web pages. The tool was evaluated through a controlled laboratory experiment aimed at simulating web maintenance task. The tool proved to significantly improve efficiency and maintain accuracy as opposed to report based methods.

This thesis introduced a novel usage of Tree-Maps to visualise web sites and developed a tool to support maintenance. Through experimentation this tool was shown to improve maintenance efficiency significantly.
A link has been established between a set of recommended accessibility practices and good accessibility. Finally, the thesis developed a novel process improvement model and introduced the idea of a knowledge repository as a method to transfer accessibility knowledge within an organisation.

References

1. (SIGCGEC), Special Interest Group on Computer Graphics Education Committee, *Definitions, History, and Goals of Visualization.* 1999.
2. Adaptive Technology Resource Centre, University of Toronto, *A-Prompt Web Accessibility Verifier Website*, in *Secondary A-Prompt Web Accessibility Verifier Website* 2004 http://aprompt.snow.utoronto.ca/, Last accessed: 12th July 2004
3. Centro Nazionale per l'Informatica della Pubblica Amministrazione 2004 Provisions to support the access to information technologies for the disabled http://www.pubbliaccesso.it/normative/law_20040109_n4.htm
4. Alexander, Dey, *How Accessible Are Australian University Web Sites?* in *Australian World Wide Web Conference.* 2003: Gold Coast, Australia.
5. Arrue, Myriam and Markel Vigo, Julio Abascal. *Quantitative Metrics for Web Accessibility Evaluation.* in *1st Workshop on Web Measurement and Metrics.* 2005. Sydney, Australia.
6. Bailey, John and Burd, Elizabeth, *Web Accessibility Evolution in the United Kingdom*, in *Web Site Evolution.* 2005: Budapest.
7. Bailey, John and Burd, Elizabeth, *What is the current state of Web Accessibility?* in *Workshop on Website Evoluton.* 2006: Philadelphia, USA.
8. Bangham, J.A., et al., *Virtual Signing: Capture, Animation, Storage and Transmission: an Overview of the ViSiCAST Project*, in *IEE Seminar on "Speech and language processing for disabled and elderly people".* 2000: London.
9. BBC, *'Most websites' failing disabled*, in *Secondary 'Most websites' failing disabled* 2006 http://news.bbc.co.uk/1/hi/technology/6210068.stm, Last accessed: 10th December 2006
10. Bergel, Marguerite,Chadwick-Dias, Ann, and Tullis, Tom, *Leveraging Universal Design in a Financial Services Company*, in *ACM Special Interest Group on Accessible Computing.* 2005: ACM Special Interest Group on Accessible Computing. p. 18 -24.
11. Bjoerk, Staffan, *Redefining the Focus and Context of Focus+Context Visualizations*, in *IEEE Symposium on Information Visualization.* 2000: Salt Lake City, Utah.

12. Bolchini, D.,Colazzo, S., and Paolini, P., *Requirements for Aural Web Sites*, in *8th IEEE International Symposium on Web Site Evolution*. 2006: Philadelphia.
13. Boldyreff, Cornelia, *Determination and Evaluation of Web Accessibility.* Proceedings of IEEE 11th Intl. Workshops on Enabling Technologies: Infrastructure for Collaborative Enterprises, 2002: p. 36 - 41.
14. Brodlie, Ken. *Visualization Systems and the Internet.* in *Proceedings of Electronic Media in Mathematics*. 2001. Coimbra, Portugal.
15. Cabinet-Office, British Government, *Guidelines for Government Websites*, in *Secondary Guidelines for Government Websites* 2002 http://www.cabinetoffice.gov.uk/e-government/resources/handbook/html/1-10-4.asp, Last accessed: 7th June 2005
16. Ceri, Stefano,Fraternali, Piero, and Bongio, Aldo, *Web Modeling Language (WebML): a modeling language for designing Web sites.* Computer Networks (Amsterdam, Netherlands: 1999), 2000. **33**(1-6): p. 137-157.
17. CERN, *Web Authoring Introduction*, in *Secondary Web Authoring Introduction* 2006 http://webservices.web.cern.ch/WebServices/Help/?kbid=090001, Last accessed: 4th November 2006
18. Christopher G. Healey, Kellogg S. Booth, James T. Enns, *High-speed visual estimation using preattentive processing.* ACM Transactions on Computer-Human Interaction (TOCHI), 1996. **3**(2): p. 107 - 135.
19. Comercio, Ministerio de Industria Turismo y, *LAW 34/2002 11th July, Services of the Information Society and Electronic Commerce*. 2002. p. 279 - 298. http://www.congreso.es/public_oficiales/L7/CONG/BOCG/A/A_068-13.PDF
20. Disability Rights Commission 2002 Code of Practice Rights of Access Goods, Facilities, Services and Premises
21. Commission, Disability Rights, *The Web: Access and Inclusion for Disabled People*, in *Secondary The Web: Access and Inclusion for Disabled People* 2004, Last accessed: 7th October 2004
22. Disability Rights Unit 2002 World Wide Web Access: Disability Discrimination Act Advisory Notes http://www.hreoc.gov.au/disability_rights/standards/www_3/www_3.html

23. Consortium, Word Wide Web, *Evaluation and Report Language (EARL) 1.0*, in *Secondary Evaluation and Report Language (EARL) 1.0* 2002 http://www.w3.org/TR/EARL10/, Last accessed: 3rd May 2005
24. Consortium, World Wide Web, *Hypertext Markup Language - 2.0*, in *Secondary Hypertext Markup Language - 2.0* 1995 http://ftp.ics.uci.edu/pub/ietf/html/rfc1866.txt, Last accessed: 16th September 2006
25. Consortium, World Wide Web, *HTML 4.01 Specification*, in *Secondary HTML 4.01 Specification* 1999 http://www.w3.org/TR/REC-html40/, Last accessed: 9th December 2005
26. Consortium, World Wide Web, *Web Content Accessibility Guidelines 1.0*, in *Secondary Web Content Accessibility Guidelines 1.0* 1999 http://www.w3.org/TR/WCAG10/, Last accessed: 19th August 2006
27. Consortium, World Wide Web, *Webbot - the Libwww Robot*, in *Secondary Webbot - the Libwww Robot* 1999 http://www.w3.org/Robot/, Last accessed: 1st December 2005
28. Consortium, World Wide Web, *Techniques For Accessibility Evaluation And Repair Tools*, in *Secondary Techniques For Accessibility Evaluation And Repair Tools* 2000 http://www.w3.org/TR/AERT, Last accessed: 16th July 2004
29. Consortium, World Wide Web, *Implementation Plan for Web Accessibility*, in *Secondary Implementation Plan for Web Accessibility* 2002 http://www.w3.org/WAI/impl/Overview.html, Last accessed: 27th June 2006
30. Consortium, World Wide Web, *Architecture of the World Wide Web*, in *Secondary Architecture of the World Wide Web* 2004 http://www.w3.org/TR/webarch/, Last accessed: 30th June 2005
31. Consortium, World Wide Web, *Web Accessibility Initiative*, in *Secondary Web Accessibility Initiative* 2004 http://www.w3.org/WAI/, Last accessed: 19th July 2004
32. Consortium, World Wide Web, *Developing a Web Accessibility Business Case for Your Organization*, in *Secondary Developing a Web Accessibility Business Case for Your Organization* 2005 http://www.w3.org/WAI/bcase/Overview.html, Last accessed: 22nd August 2006

33. Consortium, World Wide Web, *WAI Resources on Managing Accessibility*, in *Secondary WAI Resources on Managing Accessibility* 2005 http://www.w3.org/WAI/managing.html, Last accessed: 24th January 2006
34. Consortium, World Wide Web, *W3C Markup Validation Service*, in *Secondary W3C Markup Validation Service* 2006 http://validator.w3.org/, Last accessed: 29th November 2006
35. Consortium, World Wide Web, *Web Accessibility Evaluation Tools: Overview*, in *Secondary Web Accessibility Evaluation Tools: Overview* 2006 http://www.w3.org/WAI/ER/tools/, Last accessed: 20th July 2006
36. Consortium, World Wide Web, *Web Content Accessibility Guidelines 2.0*, in *Secondary Web Content Accessibility Guidelines 2.0* 2006 http://www.w3.org/TR/WCAG20/, Last accessed: 19th September 2006
37. Cuenca, Pedro and Sosa, Vicente. *Experiences in the Use of Metadata for Web Publishing*. in *Electronic Publishing*. 1999. Ronneby, Sweden.
38. US Department of Commerce 1999 Falling Through the Net: Defining the Digital Divide
39. Dart, Susan. *Change Management:Containing the Web Crisis*. in *Software Configuration Management Symposium*. 1999. Toulouse, France.
40. Deshpande, Yogesh, et al., *Web Engineering*. Journal of Web Engineering, 2002. **1**(1): p. 3 - 17.
41. Israeli Ministry for Justice 2005 Israel's new accessibility law: An executive summary
42. Donkin, Joanna, et al., *The Case for the Use of Plain English to Increase Web Accessibility.*, in *3rd International Workshop on Web Site Evolution (WSE'01)*. 2001: Florence, Italy.
43. Dreyfuss, John, *Designing for People*. 1955: Allworth Press.
44. e-Government-Unit-(UK), *eAccessibility of public sector services in the European Union*, in *Secondary eAccessibility of public sector services in the European Union* 2005 http://www.cabinetoffice.gov.uk/e-government/resources/eaccessibility/exec_brief/index.asp, Last accessed: 13/04/2006
45. e-Government Unit, Cabinet Office (UK), *eAccessibility of public sector services in the European Union - Executive briefing*. 2006.

http://www.cabinetoffice.gov.uk/e-government/resources/eaccessibility/exec_brief/foreword.asp

46. Eurostat, *Level of Internet access households*, in *Secondary Level of Internet access households* 2005 http://epp.eurostat.ec.europa.eu/, Last accessed: 3rd February 2006
47. Fekete, Jean-Daniel and Plaisant, Catherine, *Interactive Information Visualization of a Million Items.* Proceedings of IEEE Symposium on Information Visualization, 2002: p. 117 - 125.
48. Fekete, Jean-Daniel and Plaisant, Catherine, *Interactive Information Visualization of a Million Items.* Proceedings of IEEE Symposium on Information Visualization (InfoVis'02), 2002: p. 117 - 125.
49. Government, British, *Special Educational Needs and Disability Act 2001.* 2001. http://www.hmso.gov.uk/acts/acts2001/20010010.htm
50. Government of Victoria, Australian State, *Whole of Victorian Government Web Content Lifecycle and Content Management Roles*, in *Secondary Whole of Victorian Government Web Content Lifecycle and Content Management Roles* 2003 http://egov.vic.gov.au/pdfs/WebContentLifecycleReport-v1-Oct2003.pdf, Last accessed: 18th August 2006
51. Government, United Kingdom, *Disability Discrimination Act 1995.* 1995. http://www.opsi.gov.uk/acts/acts1995/1995050.htm
52. Government, US, *Section 508 Website*, in *Secondary Section 508 Website* 2004 http://www.section508.gov/, Last accessed: 31st October 2004
53. National Initiative for Citizens with Special Needs in the Information Society 1999 Resolution of the council of ministers concerning the accessibility of public adminstration web sites for citizens with special needs
54. Gybels, Guido, *Deaf and Hard of Hearing Users and Web Accessibility.* 2004, The Royal National Institute for Deaf People.
55. Herman, Ivan,Melancon, Guy, and Marshall, M. Scott, *Graph Visualization and Navigation in Information Visualization: a Survey.* IEEE Transactions on Visualization and Computer Graphics, 2000. **6**: p. 24 - 43.
56. Herman, Ivan,Melançon, Guy, and Marshall, M. Scott, *Graph Visualization and Navigation in Information Visualization: a Survey.* Proceedings of IEEE Transactions on Visualization and Computer Graphics, 2000. **6**(1): p. 24-43.

57. Hudson, Roger,Weakley, Russ, and Firminger, Peter, *An Accessibility Frontier: Cognitive disabilities and learning difficulties*, in *Secondary An Accessibility Frontier: Cognitive disabilities and learning difficulties* 2005 http://www.usability.com.au/resources/cognitive.cfm, Last accessed: 27th October 2006

58. IEEE, *Standard on Software Maintenance*, in *IEEE 1219 Standard for Software Maintenance*. 1993, IEEE Computer Society Press: Los Alomitos CA.

59. Initiative, Dublin Core Metadata, *The Dublin Core Metadata Initiative*, in *Secondary The Dublin Core Metadata Initiative* 2006 http://dublincore.org/, Last accessed: 13th September 2006

60. Jackson-Sanborn, Emily,Odess-Harnish, Kerri, and Warren, Nikki, *Web site accessibility: a study of six genres.* Library Hi Tech, 2002. **20**: p. 308-317.

61. Johnson, Brian and Shneiderman, Ben, *Tree-Maps: A Space-Filling Approach to the Visualization of Hierarchical Information Structures.* Proceedings of IEEE Visualization 91, 1991: p. 284 - 291.

62. Kelly, Brian, *The Role Of A Web Editor*, in *Secondary The Role Of A Web Editor* 1998 http://www.ariadne.ac.uk/issue18/web-focus/, Last accessed: 2nd December 2004

63. Kimelman, D., et al., *Reduction of Visual Complexity in Dynamic Graphs*, in *Proceedings of 2nd International Symposium on Graph Drawing*. 1994. p. 218 - 225.

64. Kirkpatrick, Andrew, et al., *Web Accessibility: Web Standards and Regulatory Compliance*. 1 ed. 2006: friends of ED. 696.

65. Knauth, Bettina, *Tourism and the Internet in the European Union*, in *Focus on Statistics*. 2006, European Union: Luxembourg.

66. Lazar, J., Beere, P., Greenidge, K., and Nagappa, Y., *Web Accessibility in the Mid-Atlantic United States: A Study of 50 Home Pages.* Universal Access in the Information Society Journal, 2003. **2**: p. 331-341.

67. Lazar, J., et al. *Detour Ahead: Current Roadblocks to Web Accessibility.* in *Conference on Universal Access in Human-Computer Interaction*. 2003.

68. Lazar, Jonathon,Dudley-Sponaugle, Alfreda, and Greenidge, Kisha-Dawn, *Improving Web Accessibility: A Study of Webmaster Perceptions.* Computers and Human Behavior, 2004. **20**(2): p. 269-288.

69. Leporini, Barbara and Paternò, Fabio, *Testing the effects of web usability criteria for vision impaired users*. 2004, Institute for Information Science and Technology (Italy).

70. Lientz, Bennett P. and Swanson, E. Burton, *Software Maintenance Management*. 1980, Boston, MA, USA: Addison-Wesley Longman Publishing Co., Inc. 160.

71. Lowe, D. and Henderson-Sellers, B, *Characteristics of Web Development Processes*, in *International Conference on Advances in Infrastructure for Electronic Business, Science*. 2001: Italy.

72. Lowe, David, *Engineering the Web - Web Engineering or Web Gardening?* WebNet Journal, 1999. **1**(1).

73. Macias, Mercedes and Sanchez, Fernando, *Improving Web Accessibility for Visually Handicapped People Using KAI*, in *Proceedings of the 3rd International Workshop on Web Site Evolution*. 2001. p. 49 - 55.

74. Mankoff, Jennifer, et al., *Web Accessibility for Low Bandwidth Input*, in *ACM conference on Assistive technologies*. 2002, ACM Press: Edinburgh, Scotland.

75. Mankoff, Jennifer,Fait, Holly, and Tran, Tu, *Is your web page accessible? A comparative study of methods for assessing web page accessibility for the blind*, in *Proceeding of the SIGCHI conference on Human factors in computing systems*. 2005, ACM Press: Portland, Oregon, USA.

76. McCormick, Bruce and DeFanti, T.A., *Visualization in Scientific Computing and Computer Graphics.* ACM Special Interest Group on Computer Graphics, 1997. **21**.

77. Morkes, John and Nielsen, Jakob, *Concise, SCANNABLE, and Objective: How to Write for the Web*, in *Secondary Concise, SCANNABLE, and Objective: How to Write for the Web* 1997 http://www.useit.com/papers/webwriting/writing.html, Last accessed: 13th October 2006

78. Neerincx, Mark A.,Lindenberg, Jasper, and Pemberton, Steven. *Support Concepts for Web Navigation: A Cognitive Engineering Approach.* in *World Wide Web Conference*. 2001. Hong Kong.

79. Nguyen, Quang Vinh and Huang, Mao Lin, *A Fast Focus + Context Viewing Technique for the Navigation of Classical Hierarchical Layout*, in

Proceedings of Seventh International Conference on Information Visualization (IV'03). 2003. p. 42 - 47.

80. Organisation, World Health, *International Statistical Classification of Diseases and Related Health Problems*. 2nd ed. 2006: World Health Organisation.

81. Ottens, Morag, *Use of the Internet among individuals and enterprises*. 2006, Statistical Office of the European Communities: Luxembourg.

82. Perkins, Annuska, *Usability Research and User Feedback are Keys to Accessible Product Planning at Microsoft*, in *Secondary Usability Research and User Feedback are Keys to Accessible Product Planning at Microsoft* 2006 http://www.microsoft.com/enable/microsoft/perkins.aspx, Last accessed: 24th October 2006

83. Preece, Jenny, *Human-Computer Interaction*. 1994: Addison-Wesley Publishing Company.

84. Pressman, Roger, *What a tangled Web we weave.* IEEE Software, 2000. **17**(1).

85. Reingold, E.M. and Tilford, J.S., *Tidier Drawing of Trees.* Proceedings of IEEE Transactions on Software Engineering, 1981. **2**: p. 223 - 228.

86. Reis, Fernando, *e-Government: Internet based interaction with the European businesses and citizens*, in *Statistics in Focus*. 2005, European Union: Luxembourg.

87. Richards, John and Hanson, Vicki L, *Web Accessibility: A Broader View*, in *Proceedings International WWW Conference*. 2004: New York, USA.

88. RNIB, *RNIB Web Access Centre*, in *Secondary RNIB Web Access Centre* 2006 http://www.rnib.org.uk/xpedio/groups/public/documents/PublicWebsite/public_webacctraining.hcsp, Last accessed: 29th September 2006

89. RNID, *Deaf and hard of hearing people (leaflet)*, in *Secondary Deaf and hard of hearing people (leaflet)* 2004 http://www.rnid.org.uk/information_resources/factsheets/deaf_awareness/factsheets_leaflets/?ciid=290319, Last accessed: 2610/06

90. Rossi, G.,Schwabe, D., and Lyardet, F., *Patterns for E-commerce applications*, in *Fifth European Conference on Pattern Languages of Programs*. 2000: Irsee, Germany.

91. Schmeiser, Lisa, *The Complete Website Upgrade & Maintenance Guide*. 1999: Sybex. 1086.

92. Schneiderman, Ben, *Tree visualization with tree-maps: 2-d space-filling approach.* ACM Transactions on Graphics (TOG), 1992. **11** (1): p. 92 - 99.
93. Schneiderman, Ben, *Designing the User Interface: Strategies for Effective Human-Computer Interaction.* 1998: Addison-Wesley Publishing Company. 511 - 549.
94. Senn, Stephen, *Cross-over Trials in Clinical Research.* 2002: Wiley. 364.
95. Shneiderman, Ben, *Designing the User Interface: Strategies for Effective Human-Computer Interaction.* 1998: Addison-Wesley Publishing Company. 511 - 549.
96. Simic, Hrvoje, *Modeling Web Site Structure in UriGraph*, in *Proceedings of The Twelfth International World Wide Web Conference.* 2003.
97. Sloan, David, et al., *Auditing Accessibility of UK Higher Education Web Sites.* Interacting with Computers, 2002. **12**: p. 313-325.
98. Small, Jeon, et al., *Web accessibility for people with cognitive disabilities*, in *Human Factors in Computing Systems.* 2005, ACM Press: Portland, OR, USA.
99. Spainhour, Stephen and Quercia, Valerie, *WebMaster in a Nutshell, Deluxe Edition.* 1st ed. 1997: O'Reilly Media, Inc. 374.
100. Stephanidis, Constantine, et al., *Toward an Information Society for All: An International R and D Agenda.* International Journal of Human-Computer Interaction, 1998. **10**(2): p. 107-134.
101. Swanson, E. Burton, *The dimensions of maintenance*, in *International Conference on Software Engineering.* 1976, IEEE Computer Society Press: San Francisco, California, United States.
102. Sweden, Statistics, *Use of computers end of the Internet*, in *Secondary Use of computers end of the Internet* 2006 http://www.scb.se/templates/tableOrChart____154541.asp, Last accessed: 4th March 2005
103. University, Griffith, *Web Publishing Roles and Responsibilities*, in *Secondary Web Publishing Roles and Responsibilities* 2006 http://www.griffith.edu.au/webdevsupport/intro/roles/content.html, Last accessed: 18th June 2006
104. Urbana-Champaign, University of Illinois at, *Illinois Accessible Web Publishing Wizard*, in *Secondary Illinois Accessible Web Publishing Wizard*

2006 http://www.accessiblewizards.uiuc.edu/, Last accessed: 26th November 2006

105. Verbyla, Janet, *The Seven Habits of Effective Web Managers*, in *3rd Australian World Wide Web Conference*. 1997: Lismore, Australia. http://ausweb.scu.edu.au/proceedings/verbyla/

106. Vidgen, RIchard,Goodwin, Steve, and Barnes, Stuart. *Web Content Management.* in *e-Everything: e-Commerce, e-Government, e-Household, e-Democracy, 14th Bled Electronic Commerce Conference*. 2001. Bled, Slovenia.

107. Warren, Paul,Boldyreff, Cornelia, and Munro, Malcolm, *The Evolution of Websites*, in *Workshop on Program Comprehension*. 1999, IEEE: Pittsburgh, PA.

108. Watchfire, *Bobby Online Free Portal*, in *Secondary Bobby Online Free Portal* 2004 http://bobby.watchfire.com/bobby/html/en/index.jsp, Last accessed: 10th July 2004

109. WebAIM, *Constructing a POUR (Perceivable, Operable, Usable and Robust) Website*, in *Secondary Constructing a POUR (Perceivable, Operable, Usable and Robust) Website* 2006 http://www.webaim.org/articles/pour/, Last accessed: 22nd August 2006

110. WebAIM, *Introduction to Web Accessibility*, in *Secondary Introduction to Web Accessibility* 2006 http://www.webaim.org/intro/, Last accessed: 23rd August 2006

111. WebAIM, *WebAIM Articles on Web Accessiblity*, in *Secondary WebAIM Articles on Web Accessiblity* 2006 http://www.webaim.org/articles/, Last accessed: 27th October 2006

112. WebThing, *Accessibility Valet*, in *Secondary Accessibility Valet* 2004 http://valet.webthing.com/access/, Last accessed: 25th October 2004

113. Weinstein, Anja, *Content Management.* Internet Professionell, 2000: p. 38 - 43.

114. Yee, Ka-Ping, et al., *Animated Exploration of Graphs with Radial Layout.* Proceedings of IEEE Information Visualization Symposium, 2001: p. 43 - 49.

115. Yih-Farn Chen and Koutsofios, Eleftherios, *WebCiao: A Website Visualization and Tracking System.* Proceedings of WebNet97, 1997. citeseer.ist.psu.edu/chen97webciao.html

116. Zajicek, Mary,Lee, Andrew, and Wales, Richard, *Older adults and the usability of speech interaction*, in *Proceedings of the Latin American conference on Human-computer interaction*. 2003, ACM Press: Rio de Janeiro, Brazil. p. 209-215.

117. Zeng, Xiaoming and Parmanto, Bambang. *Evaluation of web accessibility of consumer health information websites*. in *Proceedings of the American Medical Informatics Association (AMIA) Symposium 2003*. 2003. Washington.