

Towards an Accessible Web through Semantic Web Standards

Clare M. So, Mark Perry and Stephen M. Watt
 Department of Computer Science,
 University of Western Ontario,
 London, Ontario, Canada. N6A 5B7

Abstract—Making information on the Web accessible to all people, including to those having special needs, has become a necessity because of various legal obligations. Information may not be accessible by people with special needs because many Websites are not designed for viewing with alternative technologies. Aesthetic appearance of Websites, as displayed by common technologies, is often more favourable than actual semantic content that is rendered by alternative technologies. To make the Web more accessible, it is necessary to follow the current Web standards. Improving Web accessibility is an ongoing activity. The Semantic Web movement, which suggests that information should be encoded according to its meaning, brings us closer to a truly accessible Web in the future.

Keywords—World Wide Web, Accessibility, Obligations, Disabilities, Web standard

1. Introduction

One of World Wide Web Consortium's long term goals is to "make the Web accessible by all". [1] Much information is readily available on the Web. Unfortunately, the information is not necessarily accessible by all of the target audience because the material is not presented in ways that satisfy all means of access. People with special needs, namely the individuals who are identified as disabled, usually require special hardware or software for surfing the Web. Presently, most Websites' aesthetic appearance is designed specially for the dominant means, namely common hardware and software, to access the information. To fulfil organizations' moral and legal obligations, it is necessary to ensure alternative tools can intervene and present the information in the form needed for different situations.

The Semantic Web movement can further improve the accessibility of the Web in the future. In this framework, information is encoded according to its meaning so that it can be processed and reused by machines. No information should be designed for a limited number of technologies and purposes. When information is encoded this way, it can reach more audience than before because all material is so flexible that it can adapted to all means of access. There exists Web standards, such as Mathematical Markup Language (MathML) [4] [5], that allow one to encode information according to its meaning.

Improving Web accessibility does not only benefit people with special needs: it can also improve the usability, discovery and efficiency of the Web for all users. [2] Markup that conforms to current Web standards enables one to use a variety of devices, such as a PDA or cell phone, to access the content of Websites. Providing clear navigation mechanisms in Websites can make users less frustrated.

2. Methods to Access the Web

Before discussing different issues of the Web, it is important to keep in mind that there is more than one technology to browse the Web. There is more than one Web browser, and not all Web browsers are graphical. Internet Explorer, Netscape, Firefox, Safari and Opera are examples of graphical browsers. Lynx (Figure 1) is an example of text-based browser, which is less popular. Portable devices such as PDA and cell phones have significantly smaller screens than computer monitors. These devices' browsers may render Websites differently than popular browsers because of the limited screen size and resolution.



Fig. 1. Homepage of *University of Western Ontario* rendered by *Lynx* (top) and *Netscape* (bottom). Note that the picture in the headline is replaced by 4211. jpeg when rendered by *Lynx*.

Many problems on Web accessibility are related to not conforming to current Web standards. Many Web designers design pages for one particular set of technologies only. These technologies include specific browsers, plug-ins and screen resolution. It is not difficult to find a Website with a notice of browser requirements such as “This site is best viewed with [a particular browser] with [a particular screen resolution]” or “This site requires [a particular plug-in]”. Certain browsers may have features that are not part of Web standards. As a result, the Websites may not be rendered in the same manner by all browsers.

3. Disabilities and Web Browsing

To understand the potential solutions to improve accessibility of the Web, it is necessary to be aware of how some disabled individuals browse the Web. Different level of cognitive, mobility, hearing and visual impairments can diminish one’s ability to perceive information on Websites.

3.1. Cognitive

Inconsistent layout or obscure organization of Websites makes information difficult to be access by individuals with memory impairment or Attention Deficit Disorder. Sounds and animated graphics at a certain frequency may trigger seizures of individuals with seizure disorders.

3.2. Mobility

Websites that limit the mode of interaction makes information hard to access by individuals whose mobility is impaired. Examples of this category of impairments includes Repetitive Stress Injury (RSI) and loss of limbs. These individuals may not be able to use a normal keyboard or mouse. Instead, they may depend on alternative input devices, such as speech recognition or handwriting. If a Web site limits users interaction to only one input device, a mouse for example, these individuals may not able to interact with the Web page at all.

3.3. Hearing

An audio clip is an alternative to deliver information. For those that are aurally impaired, such use of the technology makes it impossible for them to retrieve the information within the audio clips without having a transcript.

3.4. Visual

Low vision or colour blind individuals may have difficulties viewing Websites with fixed presentation details, such as font size and colour scheme. The font enlarging function that is part of most browsers only makes a limited amount of information to be more readable. For example, imaged text cannot be enlarged (Figure 2). Browsers may allow users to specify an alternative colour schemes, such as one suitable for a colour blind individual, but such changes cannot always be effective because the presentation detail is hard-coded in the markup of Websites.

The visually impaired may have limited access to information in Web sites because some Websites depend solely on visual cues, such as images, colours or page layout, to convey the content. These people often use tools, such as text-based or voice browsers, in which case they can only perceive the parts of the

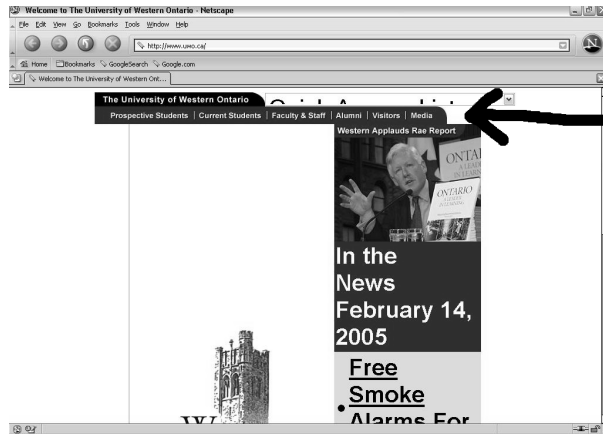


Fig. 2. Homepage of *University of Western Ontario* enlarged by using text zoom function of *Netscape*. Note that the text in the navigation bar near the top (indicated by arrow) is not enlarged because it is embedded as images.

pages that are conveyed by text from pages' source (Figure 1). Some pages' source may be unintelligent when read by these tools because the markup is not meaningful. For example, when description for images or tables are missing or meaningless, it is difficult for the tools to convey such content in Web pages. Often, HTML markup elements are not used as they were intended originally. For example, HTML `<table>`s are usually used for page layout instead of arranging data in many Websites.

4. Legal Implications

There are a growing number of legal obligations being placed on those who mount Websites, in particular services from of government and business. Obligations are dependent on the location of the site, which can be hard to determine in some instances. In the context of access to a Web site under the United States Federal Americans with Disabilities Act, one court has held that a Website is not a place, [12] at least under the terms of that statute, and thus the provisions of that Act do not apply. However, section 508 of the Rehabilitation Act of 1973, as amended in 1998, requires the federal government to meet new standards for accessibility. Any "electronic technology" developed, used or maintained by the government must also meet the new standards, including the millions of publicly accessible Websites.

In the United Kingdom, companies whose sites are not accessible to the disabled could face discrimina-

tion claims under legislation such as the UK Disability Discrimination Act 1995. However, according to a 2003 survey by the Office of the eEnvoy more than three out of every four public sector Websites failed to meet access standards for users with disabilities. The UK Disability Rights Commission tested 1000 Websites with automated tools to see whether these sites complied with voluntary Web access guidelines set by the World Wide Web Consortium: 81% (808) failed to meet minimum standards for disabled Web access. The survey also found that the average home page contains 108 barriers that make it impossible or very difficult for disabled people to use. [16] From October 1st 2004, where a feature makes it impossible or unreasonably difficult for disabled persons to access goods or services, the organization is required to take reasonable steps to:

- Remove the feature; or
- Alter it so it no longer has that effect; or
- Provide a reasonable means of avoiding the feature; or
- Provide a reasonable alternative method of making the service available.

Despite some ambiguities in legislation, or reluctance to enforce there provisions, there have been several high profile cases. The Bank of America and the Californian Council of the Blind settled an agreement in March 2000 under which the bank agreed to make its Website accessible and to provide universal ATM machines in California and Florida. A case between AOL and the National Federation of the Blind was recently settled out of court under with AOL enhancing its browser. The Sydney Organising Committee for the Olympic Games (SOCOG) was forced to make its Olympics Website accessible to a blind user who made a complaint pursuant to Australian Disability Discrimination Act 1992 The Human Rights and Equal Opportunities Commission determined that SOCOG both directly and indirectly discriminated against the complainant in failing to provide a Website which was accessible to him because of his disability. It ordered SOCOG to do all that was necessary to render its site accessible to blind Internet users and awarded the complainant A\$20,000 compensation. [13]

In the Supreme Court of Canada, dignity has become a basic element of section 15 interpretation of the Canadian Charter of Rights and Freedoms

(Every individual is equal before and under the law and has the right to the equal protection and equal benefit of the law without discrimination), and disability has been seen by the court as a unique type of disadvantage where discrimination has arisen from stereotypes about a disability as well as features of the disability itself. Some provincial legislation has taken on-board Website issues, for example the Ontario Human Rights Code, by enacting provisions for making internet sites accessible. [14] [15]

5. Web Design Recommendations with Current Technologies

Standards, guidelines and tools are available for developing and evaluating Websites. The World Wide Web Consortium (W3C) develops standards (“Recommendations” of W3C), guidelines and some of the tools. Improving accessibility of Websites do not have to be expensive: the majority of these suggestions are free. Developing Web standards is an ongoing process of W3C.

Some people may argue that providing plain text version of Websites is the solution to improve accessibility the Web. This is not necessary true because maintaining two set of pages, namely plain text and graphical version, involves extra time and energy to do so. As the result, the plain text version of the Websites, which is less popular, would be poorly maintained. Having the same version of a particular Website that can be read by a variety of technologies can provide the most recent version of information for all.

5.1. Adhere to Web Standards

Some technologies are developed with accessibility as one of the goals. Cascading Style Sheets (CSS) [3] is a process for applying presentation details, such as spacing, colour scheme and font style, to Web pages. By using CSS, presentation details are not hard-coded in the markup and it is possible to specify alternative style for people with visual disabilities. Mathematical Markup Language (MathML) [4] [5] specifies how mathematics should be included in Web pages. Traditionally, mathematical expressions are embedded in Web pages as images, which may make the expressions unreadable by individuals with visual disabilities. When mathematical expressions

are described in MathML in Web pages, their appearance can be intervene by other means.

5.2. Evaluate Websites for Accessibility

W3C also defines concepts of accessible Web design through Web Content Accessibility Guidelines. [8] This set of guidelines emphasizes on using W3C technologies properly, creating Websites independent on users’ hardware and software, and providing clear navigation clues. For example, auditory and visual content of Web sites should have text equivalent because text can be output in multiple ways such as having text enlarged.

Automated tools are available for checking some Web accessibility criteria. W3C Markup Validation Service [7] checks if the Web pages conform to the Web standards. HTML Tidy [11] attempts to fix markup that does not conform to the standards. Bobby Online Free Portal [10] is a service to check Web pages to see if they conform to Web Content Accessibility Guidelines or U. S. Section 508 Guidelines. Note that automated tools may not cover all criteria of Web accessibility. Some criteria, such as natural language usage, are subjective and must be checked manually.

6. A Look into the Future: Semantic Web

The need to improve the accessibility of the Web motivates us to examine all avenues to achieve that goal. One direction that is often overlooked is the way the information of the Web is encoded.

The manner in which Web-based information is encoded can make a huge difference in the effort required to make it more accessible. Improving this can make a difference for all users, not just those who have different levels of ability. For example, accessibility for sightless Web use is of interest to those accessing the Web through a voice-only telephone, as well as for visually impaired users.

One choice in information encoding for the Web is the semantic level of the markup. For example, one could equally well start a new paragraph in an HTML Web page using a <p> element or two consecutive line breaks

. Even though these two encodings would present pages that look the same, the first captures the *semantic notion of a*

paragraph, while the second captures only that there is to be a blank line between to parts of the text. In other words, using the markup `

` does not capture the *semantics* of the encoded objects.

The more semantic information that is used to mark up material for the Web, the easier is the job to make this material accessible. There are a number of existing and evolving standards to encode various types of material in a semantically rich way. For example, MathML provides an encoding for mathematics that is much more useful than simply including a picture of an equation (as was done previously). The same argument applies to musical notation, chemical formulas, organizational charts, numerical models or any other well-defined type of object. A number of organizations have developed or are developing standards for rich markup of different data types and for the composition of documents from these sub-objects. The World Wide Web Consortium is a leader in this area, with a focus on generally applicable standards. Corporations, professional organizations and other interest groups are also providing specialized standards for particular areas.

The movement toward higher-level encoding of Web-based information is known as the “Semantic Web.” As its principal proponents describe it,

“The Semantic Web is an extension of the current Web in which information is given well-defined meaning, better enabling computers and people to work in cooperation.”
— Tim Berners-Lee, James Hendler, Ora Lassila [17]

Some very simple examples can be given even from the use of HTML. Style guides typically recommend HTML usage that is aimed at keeping the textual markup as semantic as possible, recommending, for example, that authors use `` rather than `` (bold) and `<emph>` (emphasis) rather than `<i>` (italics) to emphasize text. Likewise, the proper use of heading markers `<h1>`, `<h2>` *etc.* is superior to recommended over using emboldened text of different sizes.

These examples, however, might not be totally convincing, since one could imagine, say, a screen-reader program that knew about different forms for emphasis and typical textual layout. To provide a more convincing example of the utility of semantic markup, we use mathematics.

Mathematics as a Case Study for Semantic Markup

Mathematics is a typical example of source material that is awkward to display in Web pages. The first Web pages included mathematical formulae by displaying .gif files with pictures of the equations. This approach has a number of problems

- the content of the formulae is not available to software tools (e.g. for search, spell-checking, screen-reading)
- the appearance of the formulae does not properly follow changes of visual style (e.g. change of font size, colours, text base-line and center-line,...)
- the formulae cannot respond to changes in window geometry (e.g. to properly re-flow as the line width changes)
- cut and paste of formulae to mathematical applications (e.g. Maple, Mathematica) cannot work.

To address the first problem, some authors have attached alternative textual forms of formulae to their images. The following typical example for the equation $x = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}$ is taken from the MathML 2.0 (2nd edition) standard:

```

```

The alternative description uses the syntax of the $\text{T}_{\text{E}}\text{X}$ language, commonly used by mathematicians. This approach provides only a weak response, however, to the first point, and leaves the remaining points unsatisfied. The textual form of the formula is in a syntax that is completely different from other Web-based information so software tools will ignore it.

To address these issues, a the Math Working Group of the World Wide Web Consortium (W3C) introduced MathML in 1998. For comparison, we show how the simple formula $(ad - bc)^2$ is represented in $\text{T}_{\text{E}}\text{X}$ and MathML:

- 1) $\text{T}_{\text{E}}\text{X}$: $\$(ad-bc)^2\$$

At first sight, this looks simple and sufficient. However $\text{T}_{\text{E}}\text{X}$ is a complicated language with full programming capability and there have been numerous failed attempts at well-defined partial implementations. Even if $\text{T}_{\text{E}}\text{X}$ would be properly implemented in multiple software tools, problems remain because the mathematical text is solely visually oriented. Among

these is the problem of automatic line breaking for large formulae. In this example, \TeX has no knowledge that multiplication occurs before subtraction so automatic text wrapping would just as soon split the equation between b and c as between $-$ and b . Another grouping related issue is that sub-expression information is not available. In this example, it is only the right parenthesis that is squared with $)^2$, not the entire expression $ad - bc$.

2) Presentation MathML:

```
<math>
  <msup>
    <mfenced>
      <mrow>
        <mi>a</mi>
        <mo>&InvisibleTimes;</mo>
        <mi>d</mi>
      </mrow>
      <mo>-</mo>
      <mrow>
        <mi>b</mi>
        <mo>&InvisibleTimes;</mo>
        <mi>c</mi>
      </mrow>
    </mfenced>
  </msup>
</math>
```

This form is verbose, but contains all the information necessary to properly render the expression in various contexts, both visual and non-visual. The content is searchable and transformable by all the standard Web-tools. The meaning of the expression is still not captured entirely. The use of the `<msup>` element states solely that the “2” is a superscript of the $(ad - bc)$. That it is a second power, as opposed, e.g., to the second component of a vector, is left to the reader to infer. To convey more meaning, MathML content markup is used.

3) Content MathML:

```
<math>
  <apply>
    <power/>
    <apply>
      <minus/>
      <apply>
        <times/>
        <ci>a</ci> <ci>d</ci>
      </apply>
      <times/>
      <ci>b</ci> <ci>c</ci>
    </apply>
  </math>
```

```
</apply>
<cn>2</cn>
</apply>
</math>
```

Content MathML describes only the meaning of the formula, and not the way it should appear. A browser would have complete freedom to display `<apply> <times/> <ci>a</ci> <ci>d</ci> </apply>` as ad , $a \times d$ or $a \cdot d$, for example.

4) Mixed MathML markup:

```
<math>
  <apply>
    <power/>
    <apply>
      <minus/>
      <apply>
        <times/>
        <ci>a</ci> <ci>d</ci>
      </apply>
      <times/>
      <ci>b</ci> <ci>c</ci>
    </apply>
  </apply>
  <cn>2</cn>
</math>
<semantics>
  <annotation-xml
    encoding="MathML-Presentation">
    <msup>
      <mfenced>
        <mrow>
          <mi>a</mi>
          <mo>&InvisibleTimes;</mo>
          <mi>d</mi>
        </mrow>
        <mo>-</mo>
        <mrow>
          <mi>b</mi>
          <mo>&InvisibleTimes;</mo>
          <mi>c</mi>
        </mrow>
      </mfenced>
    </msup>
  </annotation-xml>
</semantics>
</math>
```

The last form of MathML provides information about *both* the *form* and the *meaning* of the mathematical expression. This allows finely-tuned display as well as proper access to the semantics for cut-and-paste to other applications, searching and indexing, transformation, and accessibility.

7. Conclusion

We see that there are growing legal obligations for accessibility that certain Website providers face, in addition to moral obligations and good business sense. A number of avenues provide partial solutions to accessibility, including various alternative technologies to access Web-based information. Ultimately, the greatest strides in accessibility are to be achieved through more appropriate encoding of the Web-based information, using semantic markup that captures the meaning of the Web-based data. Mathematics serves as an interesting example for this point.

In principle, any movement toward a Semantic Web will allow people and computer programs to work together in a richer way and increased accessibility will be a natural consequence. In practice, we see that the current trend of appearance consciousness of Websites is pulling in the *opposite* direction: semantic content is being dropped in favour of aesthetic appearance. This is not a triviality: considerable marketing resources are deployed in tuning the appearance aspects of Websites.

At present, only a fraction of Websites are legally obligated to provide a truly generally accessible interface. Among those only some fraction are in compliance. The majority of commercial Websites across various jurisdictions will provide general accessibility only if it requires minimal cost, minimal effort, and does not interfere with the visual design.

There is no technical reason that a Semantic Web cannot also be an aesthetic Web. We have shown examples of how this can be achieved, e.g. through parallel markup for mathematics. More broadly, this allows the wider use of Web-based information, necessary for accessibility and rich data exchange. It also acknowledges the reality that the appearance of Websites is big business.

References

- [1] I. Herman, I. Jacobs, and J.-F. Abvramatic. *About the World Wide Web Consortium*. <http://www.w3.org/Consortium/>, February 4 2004.
- [2] A. Arch and C. Letourneau. Auxiliary benefits of accessible web design. <http://www.w3.org/WAI/bcase/benefits.html>, March 24 2002.
- [3] B. Bos, H. W. Lie, C. Lilley, and I. Jacobs. Cascading Style Sheets, level 2. CSS2 Specification. W3C Recommendation.
- [4] S. Buswell, S. Devitt, A. Diaz, P. Ion, R. Miner, N. Poppelier, B. Smith, N. Soiffer, R. Sutor, S. Watt. Mathematical Markup Language (MathML) Version 1.01. W3C Recommendation <http://www.w3.org/TR/REC-MathML/>, July 7 1999.
- [5] R. Ausbrooks, S. Buswell, D. Carlisle, S. Dalmas, S. Devitt, A. Diaz, M. Froumentin, R. Hunter, P. Ion, M. Kohlhasse, R. Miner, N. Poppelier, B. Smith, N. Soiffer, R. Sutor, and S. Watt. Mathematical Markup Language (MathML) Version 2.0 (Second Edition). W3C Recommendation. <http://www.w3.org/TR/2003/REC-MathML2-20031021/>, October 21 2003.
- [6] Judy Brewer et al. How People with Disabilities Use the Web. <http://www.w3.org/WAI/EO/Drafts/PWD-Use-Web/>, January 2001.
- [7] W3C Markup Validation Service. <http://validator.w3.org/>.
- [8] Web Content Accessibility Guidelines 1.0. W3C Recommendation. <http://www.w3.org/TR/WCAG10/>, May 5, 1999.
- [9] W3C Semantic Web Homepage <http://www.w3.org/2001/sw/>
- [10] Bobby Online Free Portal <http://bobby.watchfire.com/>
- [11] HTML Tidy <http://www.w3.org/People/Raggett/tidy/>
- [12] Access Now, Inc. v. Southwest Airlines, Co., 227 F. Supp. 2d 1312, 1317 (S.D. Fla. 2002)
- [13] Maguire v SOCOG H 99/115 http://www.hreoc.gov.au/disability_rights/decisions/comdec/2000/DD000120.htm
- [14] The Ontario Code Human Rights Code, R.S.O. 1990, c. H.19
- [15] Ontarians with Disabilities Act 2001, S.O. 2001, c. 32. It should be noted that a new Act, Bill 118, has passed its second reading in December 2004, and if this becomes law it will apply to the private sector as well.
- [16] Disabilities Rights Commission <http://www.drc-gb.org/>
- [17] Tim Berners-Lee, James Hendler, and Ora Lassila. The Semantic Web. In *Scientific American*. May 2001.
- [18] A. Chadwick-Dias, M. McNutty, and T. Tullis. Web Usability and Age: How Design Changes Can Improve Performance. In *ACM Conference on Universal Usability*, pages 30–37, Vancouver, British Columbia, Canada, 2003. ACM Press.
- [19] S. Harper, C. Goble, and R. Stevens. A Pilot Study to Examine the Mobility Problems of Visually Impaired Users Travelling the Web. *ACM SIGCAPH Computers and the Physically Handicapped*, pages 10–19, September 2000.
- [20] A. I. Karshmer and K. Kaugars. Equal Access to Information for All: Making the World of Electronic Information More Accessible to the Handicapped in our Society. *ACM SIGCAPH Computers and the Physically Handicapped*, pages 11–23, November 1995.
- [21] Mathspeak. <http://www.mathspeak.org/>.