Content Accessibility of Web documents.
Principles and Recommendations

Elena JITARU, Adriana ALEXANDRU
National Institute for R&D in Informatics, Bucharest, Romania
jitaru@ici.ro, adriana@ici.ro

The paper is an overview of issues related to the accessibility of Web sites, of European initiatives and recommendations in the field, of future solutions, such as developing Web 2.0 accessible applications with WAI-ARIA. As part of the Pro-Inclusiv project was included a set of accessibility recommendations to design web sites, recommendations presented in the paper.
Keywords: *Web accessibility, users with disabilities, standards and recommendations.*

Introduction
The Web is becoming an important part of many people life, being a key source for information gathering, education, employment and civic participation. In this context the Web can have an enormous potential for disabled people: it can make their life easier, it can eliminate many barriers, by allowing them to do things that might have been impossible or very difficult in the past. Unfortunately, people with different abilities must frequently overcome additional obstacles before they can enjoy the
full range of information, services and social interaction offered by the Web.
It is widely known that a site should be user centered in that it is based on users' site knowledge, in particular their technological and physical capacities, their cultural context and their information needs. Moreover, when accessible features are built into Web pages, Web sites become more friendly and more available to everyone, not only to disadvantaged users.
Accessibility of the Web to end users, regardless of ability or browsing environment, is widely accepted as a fundamental requirement if the Web is to reach its true potential as an enabler for the widest possible audience. Indeed, the rights of Web users with disabilities are becoming ever more defined in anti-discrimination policy and legislation around the world.
The next generation of the Web is relying on new technologies to build rich interfaces and applications which enable community, interaction, collaboration, and social networking. This has implication for people with disabilities who have come to rely on the Web to provide more independence, work opportunities, and social interactions. These new interaction models are pushing the limits of the technologies of the Web and the ability of assistive technologies to interpret the changing face of the Web.

1. Definitions of Web accessibility
Web accessibility can be defined as a person's ability to access the Web. A Web site is accessible if it can be used as effectively by people with disabilities as well as by other people. The contents of the site, the facilities and services should be accessible to a wider audience as possible, regardless of age, disability or limitations of the technology or the environment of the end user.
One definition offered by Chuck Letourneau, president of Starling Access Service, is as follows: anyone using any kind of Web browsing technology must be able to visit any site and get a full and complete understanding of the information contained there, as well as have the full and complete ability to interact with the site" [1].
According to Tim Bernes-Lee (director of W3C and inventor of the world wide web), accessibility means "making the Web's benefits available to all people, whatever their hardware, software, network infrastructure, native language, culture, geographical location, or physical or mental ability [2].
One can say that Web accessibility is an as-

pect of quality. Any designer of a Website, concerned about the quality of the Web site, will take into account issues such as navigability, intelligibility, comprehensibility, consistency in structure, which are essential characteristics of Web sites accessible. Accessibility means in broad sense Web for all. Accessibility in the strict sense means taking into consideration of people with disabilities.

2. The WAI accessibility guidelines
The World Wide Web Consortium (W3C) Web Accessibility Initiative (WAI) develops Web accessibility guidelines for the different components [3]:
- Authoring Tool Accessibility Guidelines (**ATAG**) addresses authoring tools;
- Web Content Accessibility Guidelines (**WCAG**) addresses Web content, and is used by developers, authoring tools, and accessibility evaluation tools;
- User Agent Accessibility Guidelines (**UAAG**) addresses Web browsers and media players, including some aspects of assistive technologies.

WAI guidelines are based on the fundamental technical specifications of the Web, and are developed in coordination with W3C technical specifications (HTML, XML, CSS, SVG, SMIL etc.).

The first major responsibility of WAI was to formalize guidelines for Web content developers and designers. WAI introduced Web Content Accessibility Guidelines (**WCAG**) to the public as a draft in 1998 and developed it into a full recommendation in 1999 [4].

WCAG 1.0 has 14 recommendations, which are general principles of accessible design [4]. Each recommendation has one or more checkpoints that explain how the guideline applies in a specific area (total: 64 checkpoints. Each checkpoint is assigned a priority, explained in the "Priorities" section of WCAG 1.0. Under each checkpoint there is a Techniques link that goes to the section within the "TECHNIQUES FOR WCAG 1.0" Gateway document that links to relevant techniques for a specific technology (for example, CSS or HTML).

There are two basic themes reflected in the

WCAG, namely: ensuring graceful transformation (recommendations 1-11) and making content understandable and navigable (recommendations 12-14). Although the implementation of WCAG 1.0 recommendations is not always easy to achieve, their application can help Web designers to make resources accessible. Version 2.0 of WCAG is much improved.

On 8 January 2003, WAI has published a new version WCAG 2.0 [5], a draft version. Thus, WCAG 2.0:
- is more efficiently organized,
- may adjust the priority of some checkpoints,
- modifies, removes, or adds requirements due to changes in Web technologies since the publication of WCAG 1.0,
- incorporates the Errata from WCAG 1.0,
- reflects the experience gained in implementing WCAG 1.0.

WCAG 2.0 applies more broadly to different Web technologies and is designed to apply as technologies develop in the future. The WCAG 2.0 requirements are more testable. WCAG 2.0 is organized around four design principles of Web accessibility. The four main guiding principles of accessibility in WCAG 2.0 are: Perceivable, Operable, Understandable, Robust. Conveniently, these principles spell out an acronym that is relatively easy to remember: POUR.

WCAG 2.0 provides requirements for making Web content more accessible to a wide range of people with disabilities, including blindness and low vision, deafness and hearing loss, learning disabilities, cognitive limitations, limited movement, speech disabilities, and others. However, even content that completely conforms to WCAG may not be fully accessible to every person with a disability.

3. Public Websites accessibility in Romania
The Web is becoming an important part of many people life, being a key source for information gathering, education, employment and civic participation. Recently most of Romanian cities and villages were wired and

the number of people having Internet at home is growing on an exponential scale. That does not mean that accessibility is achieved for all sites.

In the National Strategy for the New Economy and the Implementation of the Information Society developed by the Romanian Ministry of Communications and Information Technology there are including several actions for enabling access for all to Information Society specific services: special measures to adopt the standards for accessibility of information technology products; adoption of the Web Accessibility Initiative (WAI) guidelines for public websites; ensure the establishment and network connection of national "Design-for-all" centers of excellence.

Romanian research activities are contributing to the improvement of the citizens' quality of life in a global Information Society. One of the fields of the recent „Excellence research project", launched by the Romanian Ministry of Education and Research, is ICT meeting societal challenges aiming: to improve equal participation in the information society, to prevent digital divides, to promote assistive technology and design-for-all principles.

The article 70 of the "Law concerning the protection and the promotion of the rights of the persons with handicap" [6] developed by ANPH (National Authority for Persons with Handicap) and adopted by the Romanian Government requires that until December 31, 2007 public administrations act in order to provide the access of the persons with visual and mental disabilities to public web sites and to improve their access to electronic documents.

A study realized by the team of Pro-Inclusiv project underlines the degree to which Romanian public Websites behind with regards to accessibility for persons with disabilities. This is partly due to a lack of information and training for technical resources as well as a lack of commitment and leadership at the political level. The results of the study demonstrated the need to increase the awareness of Web accessibility among the designers of Web sites and at the level of managers. Only 8% of the tested Web sites passed the WCAG guideline priority 1 checkpoints, even though the most violated checkpoints have technically uncomplicated solutions, if designers pay attention to them (for example to provide a text equivalent for every non-text element). None of the web sites analyzed are completely accessible to persons with disabilities, i.e., there were no sites that had no violation of web accessibility rules. Web accessibility, if ever considered, is often an afterthought once the Web content design is finished [7].

The interest of stakeholders in inclusive ICTs was also evaluated by means of a questionnaire. The main finding was that while web designers do know the priorities concerning accessibility, but minimize their importance, managers lack this knowledge. We found out that the most of development specifications did not address accessibility aspects. Moreover, users-group was elitist and did not include people with disabilities.

The costs for accessibility are low when attention is paid since the early design phases, while they are higher after the completion of the work. We call upon decision-makers and Web designers to improve the level of accessibility of the public websites.

4. The Pro-Inclusiv accessibility recommendations

The Pro-Inclusiv project included a set of accessibility recommendations to design web sites. Thus, 19 accessibility recommendations for Web applications were defined, taking into account:

- the W3C Initiatives:
  o Web Content Accessibility Guidelines (WCAG);
  o User Agent Accessibility Guidelines (UAAG);
  o Authoring Tool Accessibility Guidelines (ATAG);
- the standards set forth in paragraph 1194.22 of Section 508 of the US Rehabilitation Act;
- the accessibility standards and technical specifications drawn up by the International Organization for Standardization (ISO);
- the experience gained by the team of Na-

tional Institute for R&D in Informatics concerning the development and the validation of accessible software.

For each requirement the following are indicated: the requirement number, the statement of the requirement, the justification for the need of applying the recommendation; accessibility techniques; possible validation of the recommendation, comments, references.

Where they exist, references to the WCAG 1.0 checkpoints and Section 508 paragraph 1194.22 standards have been indicated.

We present below, the statement of the 19 recommendations.

1. Build pages (including objects contained in page) using technologies recommended by formal grammars, using the most recent version available when these are supported by user programs. It is recommended to use elements and attributes so as to comply with the specifications and respects their semantic aspect.

2. Avoid the use of frames in the development of the Web sites.

3. Provide a text equivalent for every non-text element. The text equivalent of a non-text object must communicate essentially the same content as well as the object in that specific context. Ensure that when the contents of the object changes dynamically the text equivalent is also updated.

4. Web pages must be designed so that all information transmitted through the color can be available without color. Ensure that the information and functions transmitted with the help of color are also available without color. Make sure the text and graphics can be understood if viewed without color.

5. Avoid using blinking or moving objects and scripts, especially those whose frequency may cause disturbances, difficulty with concentration or malfunctioning of assistive technologies. If you use such objects for information, alerting users of the risk and provide ways to avoid such objects.

6. Colors of foreground and background should create a contrast high enough to be understood by someone who has deficiencies in the reception of color or when using a black and white monitor.

7. Use CSS - Cascading Style Sheets to control content presentation and page layout to ensure readability even when CSS are disabled or not supported.

8. Avoid the use of the tables for page layout. If you use tables for layout:
• make sure that they are comprehensible when linearized;
• when you use the elements and attributes of the tables, follow the specifications laid down in semantic markup language used.

9. When electronic forms are designed to be completed on-line, the form shall allow people using assistive technology to access the information, field elements, and functionality required for completion and submission of the form, including all directions and cues

10. Provide clear mechanisms for navigation.

11. Ensure that links can be selected and activated through keyboard commands or keyboard emulation technology or pointer systems other than the mouse.

12. Use a clean and simple language, to promote effective communication.

13. Use markup to facilitate proper reading or interpretation of texts in different languages or abbreviations.

14. When films or multimedia presentations are indispensable to the information or service, provide an equivalent synchronized textual alternative in form of caption and/or auditory description. Otherwise provide a summary or a simple label for every video and multimedia element depending on the importance and production difficulties in the case of real-time presentations.

15. The content of the pages must be usable when applets, scripts and other programming objects are disabled or not supported. If this is not possible, provide an equivalent text in an accessible page.

16. For scripts and applets ensure that the event handlers are independent of devices.

17. If after all the efforts you can not create an accessible page, provide a link to another page that uses W3C technologies, providing the same content and updated as often as the main page that is inaccessible

18. Use client-side rather than server-side

image maps, except where sensitive regions cannot be defined with one of the valid geometric shape defined by the DTD in use.

19. Provide information on the general organization of the site (for example, using a map of the site or using an index). If there are search functions activate various search types for various levels of knowledge and preferences.

5. Architecture principles

When designing or evaluating a Web site one should consider architecture of the site, then validate each part of the structure according to current standards in force and ultimately use the recommendations of accessibility by order of priority.

Separation text-layout. One of the first achievements of the W3C was to propose a good solution to separate the text content from the presentation. This is important because managing content and presentation involves different technologies that can be validated better when they are separated. It offers the possibility to have different structures for presenting the same content, adapted to different devices or used in different environments. Furthermore, a separate description of the layout can be reused on multiple pages, ensuring a greater consistency. This principle of separation is associated with a better semantic, with beneficial effects for accessibility and vital for certain software, such as search engines. The main standard that allows the separation of content and presentation is CSS (Cascading Style Sheets).

The set of international character. Along time have been developed different character sets generally incompatible: in 2005 the European Union was using five major codifications from family ISO-8859-X. Fortunately, since 1991 the standard Unicode has been recommended as unification solution: it can handle most of the characters (over 96000 already in version 4) and is widely supported.

Web persistent addresses. Web addresses should be made in a scalable way, so as not to be necessary to be renamed when the Website evolves. If the original address is poorly chosen and should be altered or deleted, this should be done with HTTP own mechanisms so that the user can be redirected to the new address instantly and links can be updated quickly and automatically.

Comprehensible Web addresses. Web addresses short, but understandable for users are appreciated. This approach is beneficial for the quality of search engines indexation. The set of characters should be Unicode UTF-8. This is important if there are other than ASCII characters. The addresses are an important factor of trust and, in particular the name of the site (DNS - Domain Name Server). Extension (.Org,.Com,.Info) should be chosen with care.

Metadata. Searching the Web and finding some relevant information is not always easy. This process should be enhanced through: pertinent addresses for the documents, titles and subtitles, but also by the inclusion of extra information, "metadata" such as keywords, authors, classification, etc. In addition to basic metadata defined in the HTML, there are some attempts of standardization of metadata, as Dublin Core Metadata Initiative.

Semantics. A real bonus for accessibility is to provide users with rich semantic alternative formats, such as a description of the latest news from a Website in a RSS file or other ontologies, because this content is designed to be understood automatically and to be easily transformed for the end user.

6. Future solutions

HTML improvements. The current HTML defines a very limited set of widget. These are very simple, with few possibilities for customization. To extend the capabilities of interaction of the HTML there are some proposed formats such as "Web Forms 2.0 "[8] and W3C Xforms [9] trying to improve accessibility and semantics of Web forms.

W3C has already released the first draft of HTML5 aimed of creating Semantic HTML. This means the creation of HTML documents devoid of any reference relative to the presentation. This means that Semantic HTML contains only the implied meaning of data through the use of appropriate tags. Separation of content and presentation was pushed

into HTML5 closer to reality. Semantic HTML introduce a number of new features and exclude a number of old items. The new additions of HTML5 emphasizes on Semantic HTML, allowing the Web developer to assign, easily meaning to content via tags. The list below shows some of these elements [10]:

Table 1. New additions in HTML5

article	Represents an independent piece of content within a page
audio	Provides multimedia support
canvas	Renders dynamic bitmap graphics
dialog	Marks up a conversation
embed	Provides support for plug-in content
footer	Represents the footer of a section
header	Represents the header of a section
nav	Represents a section of a document whose purpose is navigation
section	Represents a generic section of a document
video	Provides multimedia support

This subset of new elements shows that presentational HTML disappeared. Web developers should focus on CSS when they submit content to the community of users.

Device independence. Accessibility rules are useful to achieve a single good and accessible version of the Web document, and although this approach is important, it has a series of limitations. Therefore, outside the accessibility rules there is a series of actions to facilitate dialogue between the user devices and the server of documents, to provide more personalized versions of the documents. The Device Independence working group is in charge of this topical at the W3C and has published recommendations like CC/PP (Composite Capabilities/Preference Profiles) [7].

7. Web 2.0 applications accessible with WAI-ARIA

Web 2.0 is a new approach to Web content, making it more interactive and allowing sites to combine features in new ways. This change in paradigm brings new challenges to people with disabilities. New types of services are made possible by the ability of sites to share or aggregate data, and for users to be part of the authoring process.

Although Web 2.0 applications behave increasingly like desktop applications, they continue to use Web technologies for content transmission, encoding, and presentation. These technologies are used or combined in novel ways that lack the accessibility features that have been built over years into their platform-specific counterparts and are just becoming effectively mainstream in "Web 1.0" content. Web 2.0 thus presents significant risks to accessibility [12].

Web 2.0 applications often have accessibility and usability problems because of the limitations of XHTML. Accessible Rich Internet Applications (ARIA) is a specification being brought forward by the W3C Web Accessibility Initiative's (WAI) Protocols and Formats Working Group. The goal of ARIA is to add additional semantic data into HTML and XHTML to allow assistive technologies to better represent user interface components and dynamic interactions to the user. The specification also addresses providing input focus and full keyboard navigation within the components of an application [13].

The technical specifications are provided in two parts: Roles for Accessible Rich Internet Applications [14] and States and Properties Module for Accessible Rich Internet Applications [15] WAI-ARIA Roles identifies the types of widgets that are recognized by accessibility products, and provides an ontology of roles for these that can be attached to content. WAI-ARIA States and Properties allows authors to declare important properties of an element that affect and describe interaction.

ARIA provides authors with the means to makes the different elements in a Web application semantically rich. User agents use the role semantics to understand how to handle each element. Complex web applications become inaccessible when assistive technologies cannot determine the semantics behind portions of a document or when the user is unable to effectively navigate to all parts of it in a usable way. ARIA divides the semantics into roles (the type defining a user interface

element) and states and properties supported by the roles. Roles conveys missing information that the assistive technology needs to anticipate the behavior of the elements inside the application such as how to present the corresponding ARIA states and properties to the user. The user agent will use the accessibility semantics from the host language and ARIA accessibility semantics which may override those of the host language and present them to an assistive technology through the Document Object Model or the platform accessibility API. When supporting the platform accessibility API the user agent will create accessible objects containing the accessibility semantics for each visual element on your page. It will use the chosen API to notify the assistive of changes to the semantics as well [13].

A renewed approach to accessibility, which builds on previous work but priorities the importance of the user is 'Accessibility 2.0'. This term builds on the wide penetration of the 'Web 2.0' term and related terms such as e-learning 2.0, library 2.0, etc. which aim to communicate a step change in approaches. B. Kelly summarize in [16] the characteristics of Accessibility 2.0:

- User-focused;
- Rich set of stakeholders;
- Sustainability;
- Always beta;
- Flexibility;
- Diversity;
- Blended, aggregated solutions;
- Accessibility as a bazaar, not a cathedral.

Conclusions
Building a list of Web best practices from a technical and an accessibility viewpoint is a huge work. Therefore the approach of the W3C recommendations is pragmatic and effective. Emerging technologies with multimodal interaction are both a challenge and an opportunity for the future of Web accessibility. After Yakov Fain in "Rich Internet Applications with Adobe Flex and Java: Secrets of the Masters" 2007 was a year of Rich Internet Applications and Web 2.0. YouTube became a part of the lives of millions of people

around the world". Jeremy Geelan considers that 2007 was undoubtedly the year of Social Networking, but what of 2008? Will '08 be the year of "Unified Communications" or the year when CMS comes to stand for "Community Management System" - or even "Collaboration Management System"? Or will it be the year of a giga-merger, to beat the mere mega-mergers of 2007? [17].

It is clear that current approaches to accessibility must adapt in order to respond to changes in the way Web content is created, provided and accessed and it is necessary to continue the development of the Web as a way in which social exclusion can be minimized.

References
[1] SAS (2003) - *Accessible Web design - a definition*,
http://www.starlingweb.com/webac.htm
[2] Berners-Lee, T. (n.d.) *Web Accessibility Initiative Home Page*.
http://www.w3.org/WAI/.
[3] Essential Components of Web Accessibility, http://www.w3.org/WAI/intro/ components.php
[4] WAI (1999) - Web Content Accessibility Guidelines 1.0, W3C 1999 http://www.w3.org/TR/WCAG10/
[5] WAI (2004b) - Web Content Accessibility Guidelines 2.0. W3C Working Draft 27 April 2006 http://www.w3.org/TR/WCAG20/
[6] ANPH (2005) Legea privind protectia si promovarea drepturilor persoanelor cu handicap, http://www.anph.ro/LEGI/ LEGI2005/Noiembrie/Proiect3noiembrie.doc
[7] Jitaru E., A. Alexandru, I. Moisil, E. Tudora: Information Center for Persons with Disabilities, Proceedings of Med-e-Tel 2006 - The International Trade Event and Conference for eHealth, Telemedicine and Health ICT, Luxembourg, April 5-7, 2006, pag. 46-48, ISSN 1818-9334.
[8] Web Forms 2.0, http://www.whatwg.org/specs/web-forms/current-work/
[9] W3C XForms, http://www.w3.org/MarkUp/Forms/

[10] Patton, Tony, HTML 5 aims to formalize Semantic HTML, http://www.zdnetasia.com/techguide/webdev/0,39044903,62037976,00.htm

[11] CC/PP, Composite Capabilities/Preference Profiles, http://www.w3.org/Mobile/CCPP

[12] Cooper, Michael Accessibility of Emerging Rich Web Technologies: Web. 2.0 and the Semantic Web, http://www.w4a.info/2007/prog/k3-cooper.pdf

[13] Roadmap for Accessible Rich Internet Applications (WAI-ARIA Roadmap). W3C Working Draft 4 February 2008, http://www.w3.org/TR/wai-aria-roadmap/

[14] SEEMAN, L. Roles for Accessible Rich Internet Applications. W3C Working Draft, December 20 2006. http://www.w3.org/TR/2006/WD-aria-role-20061220/

[15] SEEMAN, L. States and Properties Module for Accessible Rich Internet Applications. W3C Working Draft, December 20 2006. http://www.w3.org/TR/2006/WD-aria-state-20061220/

[16] Kelly, B., Sloan, D., Brown, S., Seale, J, Petrie, H., Lauke, P. and Ball, *Accessibility 2.0: People, Policies and Processes,* http://www.w4a.info/2007/prog/15-kelly.pdf

[17] Geelan, Jeremy: i-Technology 2008 Predictions: Where's RIAs, AJAX, SOA and Virtualization Headed in 2008? SYS-CON's Annual i-Technology Predictions Round-Up, http://java.sys-con.com/node/478303