

This document is published in:

Online Information Review 37 (2013) 4, pp. 564-580

DOI: 10.1108/OIR-04-2012-0063

© 2013. Emerald Group Publishing Limited

Overlapping factors in search engine optimization and web accessibility

Lourdes Moreno and Paloma Martinez
Department of Computer Science, University Carlos III of Madrid,
Madrid, Spain

Abstract

Purpose – The purpose of this paper is to show that the pursuit of a high search engine relevance ranking for a webpage is not necessarily incompatible with the pursuit of web accessibility.

Design/methodology/approach – The research described arose from an investigation into the observed phenomenon that pages from accessible websites regularly appear near the top of search engine (such as Google) results, without any deliberate effort having been made through the application of search engine optimization (SEO) techniques to achieve this. The reasons for this phenomenon appear to be found in the numerous similarities and overlapping characteristics between SEO factors and web accessibility guidelines. Context is provided through a review of sources including accessibility standards and relevant SEO studies and the relationship between SEO and web accessibility is described. The particular overlapping factors between the two are identified and the precise nature of the overlaps is explained in greater detail.

Findings – The available literature provides firm evidence that the overlapping factors not only serve to ensure the accessibility of a website for all users, but are also useful for the optimization of the website's search engine ranking. The research demonstrates that any SEO project undertaken should include, as a prerequisite, the proper design of accessible web content, inasmuch as search engines will interpret the web accessibility achieved as an indicator of quality and will be able to better access and index the resulting web content.

Originality/value – The present study indicates how developing websites with high visibility in search engine results also makes their content more accessible.

Keywords Search engines, Internet, Web accessibility, Corporate social responsibility, Internet, World Wide Web

Paper type Research paper

Introduction and motivation

Search engine optimization (SEO) is the process of increasing the number of visitors to a website by achieving a high rank in the search results returned by a search engine. The higher a website ranks in the search engine results pages (SERPs), the greater the likelihood that users will visit the site (Enge *et al.*, 2012). In addition to mere presence on the web, it is increasingly important for businesses that their site be ranked more highly than those of their competitors on SERPs.

Web accessibility, on the other hand, is the quality attributed to webpages when their contents can be accessed by an individual regardless of that individual's physical capabilities or the context in which the attempted interaction takes place. Furthermore

This research work has been partially funded by the MA2VICMR (S2009/TIC-1542) and MULTIMEDICA (TIN2010-20644-C03-01) research projects.

web accessibility is understood in this study as that which complies with the Web Content Accessibility Guidelines (WCAG) (W3C WAI, 2012), the definitive reference on the matter in many countries around the world. Not all individuals access the web in the same way. Among individuals with disabilities, for example, assistive technologies (ATs) such as screen readers for the visually-impaired or alternative keyboards for users with reduced mobility are often required. An accessible webpage must guarantee content access to these individuals no matter which of the assistive technologies are used.

Among this population of disabled web users, one particularly important segment to consider is that of elderly individuals with age-related disabilities. With the ageing process, the probability of suffering from a temporary or permanent disability increases (Hanson, 2009). With specific regard to the effects of ageing on internet use, numerous studies can be found citing age-related problems – including, but not limited to, the deterioration of vision and hearing, motor problems (arthritis, joint stiffness and tremors) and cognitive impairment (problems with short-term memory and distractibility) – resulting in reduced access to webpages when these webpages present particular barriers to accessibility (Sangangam and Kurniawan, 2007; Sayago and Blat, 2009). As the mean age of the world population increases, the mean retirement age around the world is also increasing and particularly in developed societies, which are ageing progressively and rapidly. Indeed according to some demographic studies, it is expected that the world population over 60 will have tripled by 2050 (Forrester Research, 2003; UN, 2008). Added to this is the fact that future generations of seniors will be ever more literate with computers and the internet and will probably want to continue using these tools into their later years. As paradigms and technologies change, but the disabilities intrinsic to ageing do not, even the current technologically-proficient adult population can expect to face difficulties with new technologies in the future (Hanson, 2009). The implications of this reality for web accessibility and productivity are clear: the greater number of active IT users with age-related disabilities in the workplace will find themselves at least partially excluded from the many websites and applications that do not guarantee web accessibility.

While individuals with disabilities are the users most frequently affected by the presence of accessibility barriers in websites and web applications, they are by no means alone in their exclusion. Indeed accessibility barriers are currently preventing many other groups of individuals from fully participating in the information society and accessing the wide variety of services available therein. These groups represent users with a great functional diversity – such as those with lifelong physical disabilities or mobility-reducing temporary injuries (e.g. preventing an individual from using a mouse and thus having to navigate solely by means of a keyboard) – and those attempting to access the web in a wide array of contexts – such as those prevented from completing a transaction as a result of a particular web browser being used or others unable to view all available web content from a particular mobile device.

With the goal of designing a universally accessible web, important legislative and non-governmental web standardization and normalization initiatives have been undertaken in many countries. Despite this fact, however, the data indicates that the practices required or suggested by these diverse initiatives to ensure accessibility have, thus far, largely been ignored. Even in government websites required by law to guarantee full accessibility for all users, numerous accessibility barriers have

nonetheless been observed (Gilbertson and Machin, 2012; Jaeger, 2008; Kuzma, 2010; Olsen, 2008; Olalere and Lazar, 2011; Shi, 2007; Yu, 2011).

One of the causes of the current unsatisfactory situation in web accessibility is the general lack of knowledge or training opportunities available for professional web designers and developers (Freire *et al.*, 2008; Lazar *et al.*, 2004; Lopes *et al.*, 2010). As a result the proliferation of misinformation regarding the implications of web accessibility has largely gone unchecked. This generally negative misinformation has generally focused on the problems, rather than the possible solutions, that arise from the attempts of businesses to create webpages that conform to accessibility standards (Kuzma, 2010). It is advisable for governmental policies to raise awareness that web accessibility should be part of corporate social responsibility activities (Andrés and Lorca, 2012).

With SEO companies are given techniques by which to attain these objectives, increase useful traffic to their webpages and introduce potential clients to their content, services and products. Among these diverse techniques of SEO are best practices (white hat) and worst practices (black hat). Certain SEO techniques directly infringe the guidelines published by the search engines. While the specific guidelines vary slightly, they can all be summed up as: show the same content to search engines as you show to users (Malaga, 2008). Black hat SEO is the practice of using optimization tactics that cause a site to rank more highly than its content would otherwise justify, or making changes specifically for search engines that do not improve the user's experience of the site. In other words black hat optimization goes against search engine guidelines. If you step too far over the mark, your site may be penalized, or even removed from the index (Enge *et al.*, 2012).

As will be shown in this paper, the SEO techniques and practices used by a website directly affect the accessibility of that website. More specifically, as web accessibility is intrinsically bound to the quality of a webpage's content, the accessibility of a page improves with the use of white hat SEO techniques. Despite requiring more time to implement, white hat techniques – as opposed to black hat techniques whose speed in producing results, despite the poor quality of the web content produced, explains their popularity among certain developers and businesses – achieve their objective of good SERP listings through the production of genuine, quality content (Malaga, 2008). In this study distinct concepts are presented and guidelines offered for the production of accessible, high quality web content with good search engine access.

It is the aim of the present paper to show that the pursuit of a high search engine relevance ranking for a webpage is not necessarily incompatible with the pursuit of accessibility. On the contrary it will be argued that significant overlaps exist between the two. Having presented the motivation for the study, the next section offers more detailed information about SEO and web accessibility, with attention paid to related studies and publications. Then the particular overlaps between SEO and web accessibility are introduced and analyzed, followed by some concluding remarks.

Review of the two concepts

This section discusses the concepts of SEO and web accessibility and presents relevant works from the literature related to each. The overlaps between the two concepts will be discussed in depth in the subsequent section.

Search engine optimization

The analysis of commonly-used SEO techniques offers useful information on how webpages are designed and implemented with the aim of securing the highest possible position among results retrieved by search engines. Such an analysis points to the importance, for example, of the placement of particular information in the webpage, including keywords and hyperlinks.

While many different search engines are in current use, the most interesting with regard to results rankings are general search engines such as Google – currently the most widely used at 83 percent (Netmarketshare, 2013) – Yahoo, Bing, Baidu and Ask. In these large general search engines, search results are ordered and ranked by relevance, a calculation made according to factors which – in order to prevent black hat practices including automatic mechanisms to subvert the system, that is, deliberately confusing the search engine into believing that a particular webpage is more relevant than it actually is – are generally kept secret from the public.

A survey of the literature finds many studies that analyze the factors with which a website may obtain greater search engine visibility (Gandour and Regolini, 2011; Grappone and Couzin, 2010; Zhang and Dimitroff, 2005). Worth particular mention despite its date of publication is the study by Pringle *et al.* (1998) examining responses given by different search engines which, through the use of decision trees and regression analysis, concludes that a high ranking requires “[an] informative title, headings, meta fields, important keywords in the title, headings and meta fields, but do not use excessive repetition which will be caught out”. Given the age of the publication, however, a number of search engines currently used are not included in the study. Other studies used a linear regression model to approximate the dynamics underlying Google’s PageRank algorithm, and predict the PageRank of a webpage (Fortunato *et al.*, 2006; Moran and Hunt, 2006). The study by Bifet *et al.* (2005) used many different factors in an estimation function derived for the ranking function of a search engine; with this function they compare their own predicted rankings with the actual rankings of Google. Although these studies obtain a number of interesting results, they are in themselves not decisive for many existing SEO factors.

Indeed SEO factors are notoriously difficult to enumerate, as search engines do not reveal the specific factors used when determining the ranking of a website. To complicate matters further, search engines constantly work to improve their ranking calculations (SEOMoz, 2011; Evans, 2007). As a result the calculated relevance of webpages has been observed to have varied over time, due to changes in the relative weights assigned to individual SEO factors, as well as the incorporation of new factors and the elimination or modification of others. In the case of Google, for instance, the identification of more than 200 factors which have varied over the lifespan of the search engine has, thus far, prevented the definition of a precise method by which the high ranking of a website on the search engine could be guaranteed.

Therefore, without a large dataset of millions of SERPs and extremely sophisticated data-mining techniques, the identification of the factors involved in a search engine’s ranking algorithm becomes extremely difficult. Due to the secrecy maintained by the most important search engines, the only information currently available about the potential factors or groups of factors involved in the determination of results rankings is that offered by experts with no official affiliation with the search engines in question. Thus, apart from the research field, the SEO industry emerged to determine the most

important factors for achieving a high ranking. However, despite the large proliferation of such companies (e.g. Bruce Clay Inc., HighRankings.com, SearchEngineWorld.com, SearchEngineWatch.com, searchengineland.com and SEOMoz.org), each possesses only partial information on a search engine's heuristics. Nevertheless it must be noted that many of these professionals often provide interesting advice through personal or professional websites, including Google's, and have been known to issue recommendations to webmasters on topics such as, in the case of Google, avoiding the use of black hat SEO techniques (Google, 2010; Google Webmaster Central, 2011).

These numerous and diverse SEO factors can generally be categorized into two groups (Clay and Esparza, 2009; Enge *et al.*, 2009; Evans, 2007; Moran and Hunt, 2006; Fortunato *et al.*, 2006; SEOMoz, 2009). The first group, known as on-page factors (or query-factors), includes those factors related to information that can be gathered directly from the pages of a website whose relevance is to be optimized, such as the existence and frequency of keywords. The second group, known as off-page factors (or query-independent factors), includes those factors related to information about the website being optimized that can be culled from other, external websites. As discussed above, each search engine uses different factors and accords different weights to each when determining the relevance of a particular website. The present study focuses on on-page factors.

It must be mentioned that the number of SEO factor classification groups has been updated since 2011 with relevance given to groups of factors such as Linkscape URL metrics, Linkscape anchor text, social media signals, on-page factors and domain/URL factors among others. However the group of on-page factors studied here is still recognized as a separate category (SEOMoz, 2011).

Web accessibility

When designing accessible webpages, professionals must take into account the diverse ways – determined by various personal, technological and contextual factors – in which individuals' access web content. Understood in this way, an accessible webpage is one in which all content therein can be accessed by everyone independently of the way in which those individuals interact with the webpage. The so-called digital divide refers to any inequalities between groups, broadly construed, in terms of access to, use of or knowledge of information and communication technologies. Among the causes of this digital divide are barriers to web accessibility. In light of the great diversity of factors influencing how individuals access websites, as well as the rapid advance of technology in general, particular efforts must be made both to prevent the widening of this divide and to work to diminish it.

The principal organization working to promote accessibility on the web is the World Wide Web Consortium (W3C). Through its Web Accessibility Initiative (WAI) (W3C WAI, 2012), the W3C has elaborated various series of guidelines, including the Web Content Accessibility Guidelines (WCAG), aimed at eliminating accessibility barriers for individuals with disabilities (Kelly *et al.*, 2007). The WCAG 2.0, currently the latest version (W3C WAI, 2012) and on the way to becoming an ISO standard (ISO, 2012), has been developed to be applied not only to existing W3C technologies, but also to other current and emerging technologies. Furthermore, and to address accessibility in currently-used rich internet applications (RIAs) such as Flash and AJAX, additional guidelines and techniques such as WAI-ARIA have been created (Moreno *et al.*, 2011a). This latter series of guidelines still in the draft stage is not discussed in the present study.

WCAG 2.0 defines how to make web content more accessible to people with a diverse array of disabilities (visual, auditory, physical, speech-related, cognitive, language-related, learning, and neurological). Although these guidelines cover a wide range of issues, they are nevertheless unable to address the needs of people with all types, degrees and combinations of disability. The objective of WCAG 2.0 is to offer accessibility with the use of assistive technologies such as screen magnifiers, screen readers, text-to-speech software, speech recognition software, alternative keyboards and alternative pointing devices.

The WCAG are considered the official standard in the European Union and are referenced in most legislation, worldwide. Other important initiatives also exist, such as Barrierefreie Informationstechnik-Verordnung (BITV 2) (Bundesministerium der Justiz, 2011), Référentiel Général d'Accessibilité pour les Administrations (RGAA, 2009), Accessibility for Ontarians with Disabilities Act (Ontario, 2011), UNE 139803:2012 web content accessibility requirements (AENOR, 2012) and Section 508 of the Rehabilitation Act, 29 USC § 794d (US Access Board, 1998) technical standards (G3ict, 2012). Although less extensive, these standards are nevertheless very similar to the WCAG. In the texts of many of these standards and laws (in which the standards are referenced), it is often said that WCAG 2.0 should be followed directly. When the standards have their own requirements, appendices or accompanying documentation are often provided indicating the correspondence of these requirements with WCAG 2.0 (Bundesministerium der Justiz, 2011; RGAA, 2009; Ontario, 2011; AENOR, 2012; US Access Board, 1998). Given the position of WCAG 2.0 as the agreed-upon standard for governments, as well as the business and academic communities, the standard has been used here for the present study.

With the aim of obtaining the most up-to-date results possible from the comparison of web accessibility and SEO, the guidelines reviewed in the present study were taken solely from WCAG 2.0. With regard to its structure WCAG 2.0 is organized into separate layers differentiated by the specificity of the guidance offered for the design and implementation of accessible websites. The most general of these layers is represented by four fundamental principles (perceivable, operable, understandable, and robust) which, in turn, are articulated and supported by a more specific layer including a total of 12 particular guidelines. To ensure conformance with each of these specific guidelines, a third layer of even greater specificity is offered, providing the developer with testable success criteria, with each being classified under one of three levels of conformance: A (lowest), AA or AAA (highest). Depending on the website's ability to satisfy the different success criteria at the different levels of conformance, an overall level of conformance of A (lowest), AA or AAA (highest) is then determined for the webpage as a whole. Following the application of the WCAG 2.0 – a process facilitated through the provision of supporting documents – in which the level of conformance of the website is determined, developers may indicate this overall level of accessibility through the use of an official, corresponding conformance logo (W3C, 2011).

Study of SEO and web accessibility overlap

As explained above, there are numerous studies in which the visibility of a website and the potential for it to improve its ranking in SERPs are analyzed. Among the studies, considerable overlap exists between the distinct factors proposed and analyzed and web accessibility guidelines (Pemberton, 2005; Enge *et al.*, 2012). For instance, just as

bad navigability might be to blame for a webpage's low ranking by a search engine, so too is it often a barrier to user accessibility. Additionally, just as the presence of certain scripts prohibits the evaluation of particular content by the search engine, so too might it be an accessibility barrier for visually-impaired individuals attempting to interact with web content via a screen reader.

Despite this, however, and apart from a small WAI document (W3C WAI, 2009) and a positive related case study (W3C WAI, 2010), very little has been written about the overlaps between SEO techniques and web accessibility guidelines. The present section attempts to fill this gap in the literature, identifying some of the overlaps present between SEO on-page factors and WCAG 2.0.

Perhaps due to the fact that web accessibility and SEO projects have generally been carried out by experts from different fields and promoted in different ways and to different sectors, each type of project has been understood as distinct and undertaken independently from the other. Nevertheless, a deeper understanding of both web accessibility guidelines and SEO factors reveals an intimate relationship and significant overlap between the two. Furthermore, conformance with web accessibility guidelines seems to offer inherent advantages for a website with respect to SEO, inasmuch as many of the additional semantics required under the former imply techniques used in the latter. To offer an illustrative analogy, one may imagine a web crawler as accessing the web in a manner similar to that of a visually-impaired user via a screen reader. Similarly, one may imagine a web crawler as accessing the web in a manner similar to how a hearing-impaired user does with audio content provided via video. In the latter example the information will be accessed by a web crawler using captioning for the hearing impaired.

This study focuses on factors applied following best practices such as on-page white hat techniques. With respect to on-page black hat techniques, black hat optimizers use a variety of on-page methods, most of which aim to provide certain content only to spiders, while actual users see content that is completely different. The reason for this is that the content used to achieve high rankings may not be conducive to good site design, an element which directly and negatively affects its accessibility. The methods falling into this category include cloaking, doorway pages and invisible elements, among others. Of all of these, the method which most adversely affects web accessibility is invisible content, resulting from the optimizer's use of cascading style sheets (CSS) to hide elements. These hidden elements are placed within hidden div tags. Google, for one, removes content contained within hidden div tags from its index. From the perspective of web accessibility, users who access a webpage with a screen reader will encounter unwanted content that may be a great source of confusion. Black hat optimizers also make use of tools allowing them to automatically generate thousands of webpages very quickly. Thus these so-called content generators essentially copy content from other sites and include it in the new site. Content generators present a problem for legitimate web owners as their original content may be extensively copied. Furthermore, since some search engines penalize duplicate content, legitimate sites may also be penalized (Malaga, 2008). These low-quality websites with irrelevant, copied content, that are often the product of black hat techniques, result in a reduction of web accessibility. This is particularly the case with web users with cognitive disabilities, including many elderly users, who often become frustrated when accessing a website of this type.

In the present paper no attempt was made to study the overlaps present between web accessibility guidelines and off-page factors such as external links (closely related to the generation of web traffic). As the web has been evolving more towards new ways to generate traffic rather than towards new forms of web content, these current developments have made the analysis of off-page factors a highly complex undertaking. Indeed, the SEO techniques previously employed with the websites and portals of Web 1.0, or even with sites such as blogs that marked the beginning of Web 2.0, are a far cry from those currently being employed in the era of social networks. With these current SEO techniques applied to social networks, not only is greater web traffic generated, but links to other websites are also obtained. This explains the current proliferation of websites including the scripting of social networks such as Twitter or Facebook. In response to these changes, search engine algorithms are currently being updated to prevent a high relevance ranking from being conferred on a website as a result of bad internet practices such as black hat techniques – including, for example, the acquisition of numerous links via social networks by low-quality websites lacking in original content (Google Webmaster Central, 2011; SEOMoz, 2011). If anything may be concluded from current trends, therefore, it is that in addition to a high relevance ranking that reflects off-page factors, this ranking should be supported by good quality web content and services and follow white hat techniques – hence the importance of content, for example, that is accessible to users with diverse functional capabilities and who interact with the website from diverse contexts – that make the SEO strategy employed ultimately effective (Schwartz, 2012).

Summary of web accessibility and SEO overlaps

Table I presents the overlaps between WCAG 2.0 and SEO on-page factors. To generate the table, related WCAG documentation (W3C WAI, 2012) was used, reinforced by additional correspondences identified by the authors (experts in the field of web accessibility). We have used a non-academic SEO resource (SEOMoz.org, 2009). This source has been chosen here due to its comprehensive consideration of all the most frequent on-page factors found both in the academic literature, as well as in the professional sites discussed earlier.

Table I is structured according to the organization of guidelines in WCAG 2.0 in order to facilitate the easy presentation of overlaps with SEO factors. Some such groups from WCAG 2.0 focus on the provision of alternate text, simple navigation, the logical structure of the web document and its tags, as well as the inaccessible use of script.

Finally, it is important to mention that a new standard, HTML5, is being developed. Similar to HTML 4.01 the new HTML5 standard will provide semantics, facilitating web accessibility and allowing SEO techniques to continue to inform the development of webpages with an eye to their ranking in search engine results. Furthermore, search engines continue to have information at their disposal from websites that allows the former to locate relevant content in the latter. One example of this are the tags from the new standard such as `< figure >` and `< figcaption >` for the association of images with a descriptive text; `< nav >`, `< header >`, `< footer >`, `< article >` and `< aside >` for the identification of important sections of a webpage; and `< track >`, `< audio >` and `< video >` permitting the inclusion and association of alternative content (e.g. captioning) with multimedia (W3C, 2011; Moreno *et al.*, 2011b).

Table I.
Overlaps between SEO factors and web accessibility guidelines

Web accessibility guidelines	SEO factors	Accessibility/SEO relationship justification
WCAG 2.0 (W3C, 2012; W3C WAI, 2008) success criteria	(SEOMoz, 2009)	
1.1.1. Non-text content (Level A): All non-text content that is presented to the user should have a text alternative/equivalent	Keyword use in image alt text Keyword use/number of repetitions in the HTML text on the page Keyword use in the first 50-100 words in HTML on the page	Currently, search engines are not able to truthfully interpret the semantics transmitted by images, audio or video; nor are they able to understand the contents displayed in text images. However the inclusion of the alternative text descriptions indicated in WCAG 2.0 provides a mechanism whereby the corresponding multimedia content can not only be understood by search engines, but also – when this content is considered relevant – lead to the attainment of a higher ranking by the search engines. Following the success criteria, the web content is semantically annotated; for example, images can be labelled using the “alt” attribute of the < IMG > tag
1.2.1. Audio-only and video-only (Prerecorded) (Level A): For prerecorded audio-only and video-only content, an alternative for the time-based media is provided presenting equivalent information		
1.2.2. Captions (Prerecorded) (Level A): Captions are provided for all prerecorded audio included in synchronised media		
1.2.8. Media alternative (Prerecorded) (Level AAA): An alternative for time-based media content is provided both for all prerecorded synchronised media and for all prerecorded video-only media		
1.4.5. Images of text (Level AAA): 1.4.9. Images of text (No exception) (Level AAA): Only text, rather than images of text, is used to convey information, except when a particular presentation of text is essential to the information being conveyed (e.g. logotypes) or when the image of text can be visually customised to the user’s requirements (1.4.5.) Likewise, apart from the last exception from the previous paragraph which is not present (1.4.9)		

(continued)

Web accessibility guidelines	SEO factors	
<p>4.1.1. Parsing (Level A)</p> <p>4.1.2. Name, role, value (Level A): In content implemented using markup languages, the language specification is followed (4.1.1.). All user interface component elements can be determined by software, including assistive technologies (4.1.2.). In other words the types of components used must be differentiated, for example, standard markup language components, markup language components modified with script, or standard programming technology interface components</p> <p>3.1.1. Language of page (Level A)</p> <p>3.1.2. Language of parts (Level AA)</p> <p>The default language of each webpage can be determined via software (3.1.1.) The language of each passage or phrase in the content can be determined via software if it differs from the general content of the webpage (3.1.2.)</p>	<p>HTML validation to W3C standards</p> <p>Language of the content used on the site Existence of a meta description tag</p>	<p>Standards must be followed. When search engines interpret content which does not follow the model of a standard, they can often become confused and fail to locate particular content. If this guideline is respected and standards are followed, the search engine will enjoy simpler and more efficient access to web content, since the content being accessed will be in X(HTML) and will be separated from styles defined in CSS and from script code. Furthermore the following of standards produces webpages with a lower loading speed, a Google SEO factor (Google Webmaster Central, 2010) Apart from being indispensable for assistive technologies such as screen readers, the determination of the language of a website or web content allows search engines to locate relevant information from the site or content in the search language. As the majority of search engines maintain indexes of country and language data, the determination of the language for web content aids in the selection of the correct index or indexes</p>
<p>2.4.4. Link purpose (In context) (Level A)</p> <p>2.4.9. Link purpose (Link only) (Level AAA): The purpose of a link should be clearly identified. This purpose may be determined from the link text alone or with link context (2.4.4), or, more strictly, must always be identifiable in the link text (2.4.9)</p> <p>2.4.5. Multiple ways (Level AA): More than one way is provided to locate a webpage within a set of webpages (e.g. lists of related pages, tables of contents, site maps and webpage search features)</p>	<p>Keyword use in internal link anchor text on the page Keyword use in external link anchor text on the page Location in information architecture of the site</p>	<p>Following the indicated success criteria, the purpose of a link should be clear for the user and identifiable through information included in the link text. If this text is relevant with regard to the webpage content and the particular search being carried out, it may help improve the webpage's positioning among the search engine results. To offer one example, the use of keywords in navigation labels, along with navigation menus (nested lists in (X)HTML), is an important resource which assists Googlebot. Furthermore the provision of a site map also helps search engines to access each page of the website, facilitating a higher ranking in the search results, so long as the pages contain relevant content</p>

(continued)

Table I.

Table I.

Web accessibility guidelines	SEO factors
<p>1.3.1. Info and relationships (Level A): Information, structure and relationships conveyed through presentation can be determined via software or are available as text</p>	<p>HTML validation to W3C standards Keyword use in list items < li > on the page Keyword use as the first word(s) of the title tag Keyword use anywhere in the title tag</p>
<p>2.4.2. Page titled (Level A): Webpages have titles that describe their topic or purpose</p>	<p>Keyword use anywhere in the H1 headline tag Keyword use in other headline tags (< h2 > – < h6 >)</p>
<p>2.4.6. Headings and labels (Level AA)</p>	<p>As indicated in the success criteria (1.3.1), information, structure and relationships conveyed through presentation can be separated and determined. WCAG 2.0 advises the use of CSS to define presentation. This separation of content and presentation leads to lighter webpages and a lower loading speed, a Google SEO factor</p>
<p>2.4.10. Section headings (Level AAA) Headings and labels describe their topic or purpose (2.4.6.)</p>	<p>The structure is determined with the use of labels with keyword tags (< TITLE > < H1 > , < H2 > , ...) allowing for the identification by search engines of relevant and descriptive section headings in a website.</p>
<p>Section headings are used to organise the content (2.4.10)</p>	<p>That said, while many search engines positively value the presence of headings and labels in websites, their excessive use is often penalised by search engines</p>
<p>2.1.1. Keyboard (Level A):</p>	<p><i>Negative:</i> Cloaking by JavaScript/rich media</p>
<p>All functionality of the content is operable through a keyboard without requiring specific timings for individual keystrokes</p>	<p>Scripts presenting barriers to accessibility should not be used in a webpage. Just as certain visually-impaired users who, interacting with web content via a screen reader, must deactivate JavaScript, search engines cannot access content generated with scripts. Some examples of such content include those accessible only by mouse (and not by keyboard), pop-up windows opened through JavaScript technology such as AJAX (Google Webmaster Tools, 2012), and redirect through the meta refresh tag</p>
<p>3.2.5. Change on request (Level AAA):</p>	<p><i>Negative:</i> Cloaking by JavaScript/rich media support detection</p>
<p>Changes of context (important changes in webpage content that, if enacted without prior warning, could disorient the user if he/she is not able to see the complete page in a single viewing) are initiated only by user request or a mechanism is provided to turn off such changes</p>	<p>Scripts presenting barriers to accessibility should not be used in a webpage. Just as certain visually-impaired users who, interacting with web content via a screen reader, must deactivate JavaScript, search engines cannot access content generated with scripts. Some examples of such content include those accessible only by mouse (and not by keyboard), pop-up windows opened through JavaScript technology such as AJAX (Google Webmaster Tools, 2012), and redirect through the meta refresh tag</p>

Guidelines for web professionals

From the analysis of the resources summarized in Table I, the following guidelines are offered for professionals working in web design projects and for whom the development of a good SEO strategy is an important objective. These recommendations indicate which on-page SEO factors are fundamental to include among SEO strategies, inasmuch as they promote accessible content and, therefore, equal access for all web users:

- Incorporate on-page factor, “Keyword use in image alt text”, in the SEO strategy. This is due to the fact that search engines cannot truthfully interpret the semantics transmitted by images, nor are they able to understand the contents displayed in text images. Through the inclusion of alternative text descriptions as indicated in WCAG 2.0 (1.1.1, 1.4.5, 1.4.9 success criteria [SC]), a mechanism is provided whereby the corresponding image content may not only be understood by search engines, but also by ATs such as screen readers.
- Derived from the former recommendation (not included in SEOmoz, 2009), incorporate the on-page factor, “Keyword use in video subtitles and transcription”, in the SEO strategy if video content in the website provides relevant content. Insofar as: search engines are not able to truthfully interpret the semantics transmitted by audio and video; there has been a continued growth of video content on the web; and video search engines (e.g. that provided by Google) are proliferating, if subtitles and transcriptions are incorporated as indicated in WCAG 2.0 (1.1.1, 1.2.1, 1.2.2 y 1.2.8 SC), a mechanism is provided whereby the corresponding video content may not only be understood by search engines, but also by hearing-impaired users accessing web content through subtitles.
- Incorporate the on-page factor, “HTML validation to W3C standards”, in the SEO strategy as indicated by WCAG 2.0 (4.1.1, 4.1.2 SC). In this way the search engine may benefit from simpler and more efficient access to web content, besides accessing webpages with a lower loading speed. Additionally following valid HTML schemes is fundamental for proper access by ATs and the resulting access will be more usable and reliable in case of a lower loading speed.
- Incorporate the on-page factor, “Existence of a meta description tag”, in the SEO strategy for the language of the content used on the site, as indicated in WCAG 2.0 (3.1.1, 3.1.2 SC). Apart from being indispensable for ATs including screen readers, text-to-speech software and speech recognition software, the determination of the language of a website or web content allows search engines to locate relevant information in the search language.
- Incorporate the on-page factor, “Keyword use in internal/external link anchor text on the page”, in the SEO strategy. Following WCAG 2.0 (2.4.4, 2.4.9 SC), the purpose of a link should be clear for the user and identifiable through information included in the link text. If this text is relevant with regard to the webpage content, it may help improve the webpage’s positioning on SERPs.
- Incorporate the on-page factor, “Location in information architecture of the site”, in the SEO strategy, in concert with WCAG 2.0 (2.4.5 SC) for the provision of HTML site maps. This resource offers a help mechanism to users for the location of web content and also helps search engines access each page of the website,

often resulting in a higher ranking in the search results, as long as the pages contain relevant content.

- Incorporate the on-page factor, “Keyword use as the first word(s) of the title tag”, in the SEO strategy as indicated in WCAG 2.0 (2.4.2 SC). In this way users receive help to find content and orient themselves within the content by ensuring that each webpage has a descriptive title. Additionally, with this factor, the search engine may index information related to the principal content of the page.
- Incorporate the on-page factor, “Keyword use as anywhere in the H1 headline tag and use in other headline tags (< h2 > – < h6 >)”, in the SEO strategy. The factor is also indicated in WCAG 2.0 (1.3.1, 2.4.6, 2.4.10 SC). Structure is determined by the use of labels with keyword tags (< h1 > – < h6 >) allowing for the identification by search engines of relevant and descriptive section headings in a website. In this way content is logically organized using the heading and level elements. Furthermore, users are shown what information is contained in webpages and how that information is organized. When headings are clear and descriptive, users can more easily find the information they seek, as well as understand the relationships between different parts of the content. Finally, some individuals accessing the web with screen readers navigate contents through the use of headings.
- Incorporate quality content that includes relevant content. Avoid using black hat techniques including cloaking with JavaScript and invisible content.

Conclusion

The present study shows that a clear relationship exists between SEO on-page factors and web accessibility guidelines according to WCAG 2.0. As a result of this relationship, the presence of accessibility barriers in a website threatens to negatively influence the relevance ranking generated for that website by search engines, just as the use of SEO techniques may improve accessibility. The authors of the study therefore maintain that any SEO project undertaken should include as a prerequisite the proper design of accessible web content, inasmuch as search engines will interpret the web accessibility achieved as an indicator of quality and will be able to better access and index the resulting web content.

In sum, it is important to remember that access to web content is the central link between SEO and accessibility. It is not sufficient to consider SEO simply a matter of generating more traffic to a given website following black hat techniques; rather, for SEO to be effective, actual benefits for the website must also be generated. These benefits can only be obtained when the website comprises high-quality content that supports the white hat SEO strategies employed. In addition to supporting the SEO factors applied in this way, website conformance with WCAG ensures access to web content by the greatest possible number of users without discrimination and regardless of any functional limitations they may have.

Guidelines for web professionals were provided. This resource can help to make sites more accessible and provide a good SEO strategy within the white hat practices.

Among interesting areas for future research is the continued analysis of newly appearing SEO factors and indicators with respect to their relation to web accessibility. Such a study should also include the WAI-ARIA guides once these have become

official recommendations. In this way help documentation for professionals would be continually generated so that overlaps between accessibility guidelines and other important business marketing strategies would be made clear. Given the regulative framework existing in most countries regarding accessibility and the increased number of disabled individuals as a result of the ageing population, it is hoped that this documentation may serve as a guide for professionals. By following this documentation and following SEO strategies that also generate accessible web content, companies and organizations would thereby make important advances in their accessibility policies and level of ethical responsibility.

References

- AENOR (Spanish Association for Standardization and Certification) (2012), "UNE 139803:2012 Web content accessibility requirements on the web", available at: www.aenor.es/aenor/actualidad/actualidad/noticias.asp?campo=4&codigo=23655#.UA-mmWm28QI (accessed 1 December 2012).
- Andrés, J. and Lorca, P. (2012), "Factors influencing web accessibility of big listed firms: an international study", *Online Information Review*, Vol. 34 No. 1, pp. 75-97.
- Bifet, A., Castillo, C., Chirita, P-A. and Weber, I. (2005), "An analysis of factors used in search engine ranking", *Proceedings of the Workshop on Adversarial IR on the Web, Chiba, Japan, 10-14 May*, available at: <http://airweb.cse.lehigh.edu/2005/bifet.pdf> (accessed 1 December 2012).
- Bundesministerium der Justiz (2011), *Barrierefreie Informationstechnik-Verordnung (BITV) 2.0*, available at: www.gesetze-im-internet.de/bitv_2_0/index.html (accessed 1 December 2012).
- Clay, B. and Esparza, S. (2009), "Building an SEO friendly site", in Clay, B. and Esparza, S. (Eds), *Search Engine Optimization All-in-One for Dummies*, John Wiley & Sons, Hoboken, NJ, pp. 215-240.
- Enge, E., Spencer, S., Fishkin, R. and Stricchiola, J. (2009), *The Art of SEO: Mastering Search Engine Optimization*, O'Reilly Media, Sebastopol, CA.
- Enge, E., Spencer, S., Stricchiola, J. and Fishkin, R. (2012), *The Art of SEO: Mastering Search Engine Optimization*, 2nd ed., O'Reilly Media, Sebastopol, CA.
- Evans, M.P. (2007), "Analyzing Google rankings through search engine optimization data", *Internet Research Journal*, Vol. 17 No. 1, pp. 21-37.
- Forrester Research (2003), "The wide range of abilities and its impact on computer technology", available at: www.microsoft.com/enable/research/phase1.aspx (accessed 1 December 2012).
- Fortunato, S., Boguna, M., Flammini, A. and Menczer, F. (2006), "How to make the top ten: approximating PageRank from In-degree", paper presented at the 14th International World Wide Conference, Edinburgh, 22-26 May.
- Freire, A.P., Russo, C.M. and Fortes, R. (2008), "The perception of accessibility in web development by academy, industry and government: a survey of the Brazilian scenario", *New Review Hypermedia Multimedia*, Vol. 14 No. 2, pp. 149-175.
- Gandour, A. and Regolini, A. (2011), "Web site search engine optimization: a case study of Fragnet", *Library Hi Tech News*, Vol. 28 No. 6, pp. 6-13.
- Gilbertson, T.D. and Machin, C.H.C. (2012), "Guidelines, icons and marketable skills: an accessibility evaluation of 100 web development company homepages", *International Cross-Disciplinary Conference on Web Accessibility (W4A '12)*, ACM, New York, NY, 4 pages.

- Google (2010), *Search Engine Optimization (SEO), Search Engine Optimization Starter Guide*, available at: <http://support.google.com/webmasters/> (accessed 1 December 2012).
- Google Webmaster Central (2010), "Using site speed in web search ranking", available at: <http://googlewebmastercentral.blogspot.co.nz/2010/04/using-site-speed-in-web-search-ranking.html> (accessed 2 January 2013).
- Google Webmaster Central (2011), "More guidance on building high-quality sites", available at: <http://googlewebmastercentral.blogspot.com.es/2011/05/more-guidance-on-building-high-quality.html> (accessed 1 December 2012).
- Google Webmaster Tools (2012), "AJAX crawling: guide for webmasters and developers", available at: www.google.com/support/webmasters/bin/answer.py?answer=174992 (accessed 2 January 2013).
- Grappone, J. and Couzin, G. (2010), *Search Engine Optimization (SEO): An Hour a Day*, Wiley, Indianapolis, IN.
- G3ict, The Centre for Internet & Society and The Hans Foundation (2012), *Web Accessibility Policy Making: An International Perspective* (Revised Edition 2012), Narasimhan, N. (Ed.), available at: http://g3ict.org/resource_center/publications_and_reports/p/product_Category_whitepapers/subCat_0/id_150 (accessed 1 December 2012).
- Hanson, V.L. (2009), "Age and web access: the next generation" in *International Cross-Disciplinary Conference on Web Accessibility (W4A '09)*, ACM, New York, pp. 7-15.
- ISO (2012), *ISO/IEC DIS 40500 Information technology – W3C Web Content Accessibility Guidelines (WCAG) 2.0 (TC/SC: JTC 1)*, available at: www.iso.org/iso/catalogue_detail.htm?csnumber=58625 (accessed 1 December 2012).
- Jaeger, P.T. (2008), "User-centered policy evaluations of Section 508 of the Rehabilitation Act evaluating e-government web sites for accessibility for persons with disabilities", *Journal of Disability Policy Studies*, Vol. 19 No. 1, pp. 24-33.
- Kelly, B., Sloan, D., Brown, S., Seale, J., Petrie, H., Lauke, P. and Ball, S. (2007), "Accessibility 2.0: people, policies and processes" in *International Cross-Disciplinary Conference on Web Accessibility (W4A '07)*, ACM, New York, NY, pp. 138-47.
- Kuzma, J.M. (2010), "Accessibility design issues with UK e-government sites", *Government Information Quarterly*, Vol. 27 No. 2, pp. 141-146.
- Lazar, J., Dudley-Sponaugle, A. and Greenidge, K. (2004), "Improving web accessibility: a study of webmaster perceptions", *Computers and Human Behavior*, Vol. 20 No. 2, pp. 269-288.
- Lopes, R., Van Isacker, K. and Carricco, L. (2010), "Redefining assumptions: accessibility and its stakeholders", *Proceedings of the 12th International Conference on Computers Helping People with Special Needs: Part I, Vienna, Austria*, Springer-Verlag, Berlin, pp. 561-8.
- Malaga, R.A. (2008), "Worst practices in search engine optimization", *Communications of the ACM*, Vol. 51 No. 12, pp. 147-150.
- Moran, M. and Hunt, B. (2006), *Search Engine Marketing, Inc. – Driving Search Traffic to your Company's Web Site*, IBM Press, Armonk, NY.
- Moreno, L., Martínez, P., Ruiz-Mezcua, B. and Iglesias, A. (2011a), "Toward an equal opportunity web: applications, standards, and tools that increase accessibility", *IEEE Computer*, Vol. 44 No. 5, pp. 18-26.
- Moreno, L., Martínez, P., Iglesias, A. and González, M. (2011b), *HTML 5 support for an accessible user-video-interaction on the web*, *INTERACT 2011, Lecture Notes in Computer Science*, 6949, pp. 535-539.
- Netmarketshare (2013), "Search engine market share", available at: <http://marketshare.hitslink.com/search-engine-market-share.aspx?spider=1&qprid=4> (accessed 2 January 2013).

- Olalere, A. and Lazar, J. (2011), "Accessibility of U.S. federal government home pages: Section 508 compliance and site accessibility statements", *Government Information Quarterly*, Vol. 28 No. 3, pp. 303-309.
- Olsen, M.G. (2008), "How accessible is the public European web", available at: www.mortengoodwin.net/publicationfiles/how_accessible_is_the_european_web.pdf (accessed 1 December 2012).
- Ontario (2011), "Integrated Accessibility Standards made under the 'Ontario Regulation 191/11' (Accessibility for Ontarians with Disabilities Act, 2005) (AODA), 7 June, available at: www.e-laws.gov.on.ca/html/source/regs/english/2011/elaws_src_regs_r11191_e.htm#BK15 (accessed 1 December 2012).
- Pemberton, S. (2005), "The future of web interfaces", *Human-Computer Interaction - INTERACT 2005, Lecture Notes in Computer Science*, Vol. 3585, pp. 4-5.
- Pringle, G., Allison, L. and Dowe, D.L. (1998), "What is a tall poppy among web pages?", *Proceedings of the 7th International World Wide Web Conference, Brisbane, April*, p. 369-77, available at: www.csse.monash.edu.au/~lloyd/tilde/InterNet/Search/1998_WWW7.html (accessed 1 December 2012).
- Référentiel Général d'Accessibilité pour les Administrations (RGAA) (2009), "Le portail de la modernisation de l'Etat", available at: www.references.modernisation.gouv.fr/rgaa-accessibilite (accessed 1 December 2012).
- Sangangam, P. and Kurniawan, S. (2007), "An investigation of older persons' browser usage: International - Universal Access in HCI 2007", *Lecture Notes in Computer Science*, Vol. 4554, pp. 1000-1009.
- Sayago, S. and Blat, J. (2009), > "About the relevance of accessibility barriers in the everyday interactions of older people with the web", *International Cross-Disciplinary Conference on Web Accessibility (W4A '09)*, ACM, New York, NY, pp. 104-13.
- Schwartz, B. (2012), "Is on page SEO dead with Google?", available at: www.seroundtable.com/google-on-page-seo-14626.html (accessed 2 January 2013).
- SEOMoz (2009), "Search engine 2009: ranking factors", available at: www.seomoz.org/article/search-ranking-factors/2009 (accessed 1 December 2012).
- SEOMoz (2011), "2011 edition of the Search Engine Ranking Factors", available at: www.seomoz.org/article/search-ranking-factors (accessed 1 December 2012).
- Shi, Y. (2007), "The accessibility of Chinese local government web sites: an exploratory study", *Government Information Quarterly*, Vol. 24 No. 2, pp. 377-403.
- United Nations (2008), "World population prospects: the 2008 revision", available at: www.un.org/esa/population/publications/wpp2008/wpp2008_text_tables.pdf (accessed 1 December 2012).
- United States Access Board (1998), "The Rehabilitation Act Amendments (Section 508)" available at: www.access-board.gov/sec508/guide/act.htm (accessed 8 January 2013).
- W3C (2011), "HTML5. A vocabulary and associated APIs for HTML and XHTML", W3C Working Draft 25 May 2011, available at: www.w3.org/TR/2011/WD-html5-20110525/ (accessed 1 December 2012).
- W3C (2012), Understanding WCAG 2.0, available at: www.w3.org/TR/UNDERSTANDING-WCAG20/Overview.html#contents (accessed 2 January 2013).
- W3C Web Accessibility Initiative (2008), *Web Content Accessibility Guidelines (WCAG) 2.0. W3C Recommendation 11 December 2008*, Caldwell, B., Cooper, M., Guarino Reid, L., Vanderheiden, G., available at: www.w3.org/TR/WCAG20/ (accessed 1 December 2012).

- W3C Web Accessibility Initiative (2009), “Financial factors in developing a web accessibility business case for your organization”, available at: www.w3.org/WAI/bcase/fin.html (accessed 1 December 2012).
- W3C Web Accessibility Initiative (2010), “Resources for developing a web accessibility business case for your organization”, available at: www.w3.org/WAI/bcase/resources (accessed 1 December 2012).
- W3C Web Accessibility Initiative (2012), available at: www.w3.org (accessed 1 December 2012).
- Yu, D.X. (2011), “US state government websites demonstrate better in terms of accessibility compared to federal government and commercial websites”, *Government Information Quarterly*, Vol. 28 No. 4, pp. 484-490.
- Zhang, J. and Dimitroff, A. (2005), “The impact of webpage content characteristics on webpage visibility in search engine results (Part I, Part II)”, *Information Processing and Management*, Vol. 41 No. 3, pp. 665-690.

About the authors

Lourdes Moreno has a PhD in Computer Science (focusing on the design and development of accessible web applications) from the University Carlos III de Madrid and a degree in Mathematics – Computation Sciences from the Universidad Complutense of Madrid. She has worked with several IT companies in their R&D departments, focusing on infometrics (information measurement) in web channels. Since 2002 she has been at the Department of Computer Science at the Universidad Carlos III de Madrid, currently as an Associate Professor. She is involved in several international research projects and is the co-author of numerous research papers. Lourdes Moreno is the corresponding author and can be contacted at: lmoreno@inf.uc3m.es

Paloma Martinez received a degree and a PhD in Computer Science from the Polytechnic University of Madrid in 1992 and 1998, respectively. Since 1992 she has been with the Advanced Databases Group in the Department of Computer Science, Universidad Carlos III of Madrid, where she teaches Database Design and Management. Her research areas are human language technologies (multilingual information extraction and retrieval in several domains, question answering, name entity recognition and temporal information management) as well as web accessibility.

