On the Credibility of Wikipedia: an Accessibility Perspective

Rui Lopes  
LaSIGE/University of Lisbon  
Campinho Grande, Edificio C6  
1749-016 Lisboa, Portugal  
rlopes@di.fc.ul.pt

Luís Carriço  
LaSIGE/University of Lisbon  
Campinho Grande, Edificio C6  
1749-016 Lisboa, Portugal  
lmc@di.fc.ul.pt

ABSTRACT
User interfaces play a critical role on the credibility of authoritative information sources on the Web. Citation and referencing mechanisms often provide the required support for the independent verifiability of facts and, consequently, influence the credibility of the conveyed information. Since the quality level of these references has to be verifiable by users without any barriers, user interfaces cannot pose problems on accessing information. This paper presents a study about the influence of accessibility of user interfaces on the credibility of Wikipedia articles. We have analysed the accessibility quality level of the articles and the external Web pages used as authoritative references. This study has shown that there is a discrepancy on the accessibility of referenced Web pages, which can compromise the overall credibility of Wikipedia. Based on these results, we have analysed the article referencing lifecycle (technologies and policies) and propose a set of improvements that can help increasing the accessibility of references within Wikipedia articles.

Categories and Subject Descriptors
H.5.4 [Models and Principles]: User/Machine Systems—Human Factors; I.5.4 [Information Interfaces and Presentation]: Hypertext/Hypermedia—User issues

General Terms
Human Factors, Experimentation.

Keywords
Web Science, Information Credibility, Web Accessibility, Hyperlinking Quality, Wikipedia.

1. INTRODUCTION
The Web’s decentralised architecture [8] has played a major role on its growth and consequent omnipresence in everyone’s lives. It’s scale-free nature has given the opportunity for everyone to contribute independently to the Web with different kinds of information. The hyperlinking capabilities of the Web affords the inter-connection between Web pages and, therefore, related topics, without being necessary any action from the target Web page’s author (or maintainer).

By analysing the creation of such hyperlinks, one can study the notion of authoritative Web pages and Web sites. This is done by measuring the quality and relevance of a Web site by taking into account the quality and relevance of all Web sites that link to it. This perception is akin to citations on academic papers, which lays at the heart of Web search and ranking algorithms such as PageRank [4].

Due to the decentralised nature of the Web’s architecture, hyperlinks can be used as navigable and verifiable forms of making explicit the credibility of a Web page’s content. Readers can follow these hyperlinks and judge by themselves whether the information they are reading is correct, especially since the linked information might have been created by authoritative and reputed sources. The establishment of this type of links is based on the assumption that reputed sources are trusted sources of confirmation from the point-of-view of readers.

Independent verifiability of human-created content is at the core of its credibility. Therefore it is of the utmost importance to ensure that readers can access and understand information, otherwise it will be too cumbersome to verify it. While the content itself is at the core of credibility, the interface between content and the reader must be adequate. In other words, the readers must be able to access reputed sources without any kinds of barriers. For example, in a scenario where a Web page links to a supposedly authoritative source with a Web page that uses images to convey information, blind readers will not be able to grasp that information and, therefore, the credibility associated to this linking scenario is inherently flawed for this group of readers.

In this paper we present a study of how accessibility influences the credibility of Wikipedia articles for readers that are deeply dependent on accessible content. We start by analyzing the ecosystem around the creation and improvement of articles on Wikipedia. Then, we perform an experimentation of verifying the accessibility of the external citations and references provided by contributors, in comparison to the original Wikipedia articles, to verify is there is a discrepancy between both groups of Web pages. Lastly, we discuss some options on how the article edition process can be improved to take into account accessibility issues from the start.
2. BACKGROUND

As thoroughly argued in [3], the Web is a highly dynamic structure with different levels of complexity. One important aspect to be studied concerns how the Web’s technological advancements play a crucial role at a macro scale, such as emergent properties related to how credibility is established and influenced by technology. Along these lines, to study how the credibility of sources are influenced from an accessibility point-of-view, two important topics need to be discussed: the impact of user interface issues on information credibility, and the accessibility assessment of Web pages and hyperlinking structures. Next, each one of this factors is discussed.

2.1 Credibility and User Interfaces

Online credibility is a perceptual experience, as thoroughly discussed in the literature [11]. Therefore, the interface used to present information to the user has a strong influence on credibility through its case of use. This concept is one of the keystones of usability [17, 12]. It is important to notice that, for information to be credible, it must be easily accessed (i.e., no barriers to get the information) and used, from the perspective of each user’s needs and requirements [18].

In [10], the authors have defined a model for trust where linking to other sites plays an important role on triggering trust-related behaviours on the Internet. Therefore, link navigation is in fact explored as a way to convey credibility based on trusted sources, as studied by others [13].

In fact, in a large survey conducted on Web site credibility [6], it was revealed that trustworthiness plays a major role on increasing information credibility. More specifically, users have stated that linking to outside materials and sources is of high relevance. This study emphasises too that expertise is also a factor that contributes to credibility, particularly based on information citation and referencing mechanisms.

Summing up, the quality of a Web page’s user interface dictates whether it is perceived as a credible source of information. A reputable source of information can only be trusted by users on citation and referencing scenarios if credibility is perceived.

2.2 Web Accessibility

Universal usability, as detailed in [18], states that each user has an optimal user experience when accessing information if the user interface takes into account her/his particular characteristics and preferences. One important factor inherent of this concept, side-by-side with usability, is accessibility [16]. On the domain of the Web, several methods exist to verify if a given Web page is accessible. The most important method has been published by W3C, the Web Content Accessibility Guidelines (WCAG) [5]. Verifying Web accessibility is deeply centred on inspecting Web pages’ HTML structure and inherent styling and behavioural aspects (i.e., CSS and JavaScript), a process known as checkpoint verification.

However, WCAG and other methods have been specified to yield a qualitative mark on the accessibility of a Web page (i.e., in the case of WCAG, 0, A, AA, and AAA quality). While these qualitative marks provide a valuable information to users to help increase the credibility of the Web page, they convey a specific accessibility semantics that is not sufficient to users that depend on them [22]. Furthermore, Web site creators often do not take them into account, since it often involves expert and manual verification procedures [14].

To mitigate such issues, several authors propose quantification methods to assess the accessibility of a Web page based on WCAG [20, 21, 9]. This way, accessibility can be controlled without requiring particular expertise from Web site creators, and the results yielded become more close to reality (by being quantifiable). These quantification procedures open the way to the application of Web page ranking algorithms that take into account accessibility as a credibility factor [2]. Google, through a “Google Accessible Search” initiative is starting to use these procedures to filter search results with automated accessibility verifications. It has been shown that people who dependent on accessibility awareness from Web pages often dismiss the relevance of an information source if it is inaccessible, despite ranking algorithms yield relevant results [1].

3. CREDIBILITY IN WIKIPEDIA

Wikipedia is a free, editable-by-all online encyclopaedia. It has been a great phenomenon of massive collaboration of users, where exponential growth of the number of articles available [23] (around 2.5 million articles in English only, as of July 2008) is making it a highly used source of information.

The quality of the articles within Wikipedia is, therefore, the main aspect that drives users to browse and interact with it as a credible source for information [7]. Furthermore, Wikipedia’s software platform, MediaWiki, delivers a collaborative environment which motivates contributors to iterate constantly on articles’ contents, in order to improve their quality [29].

As discussed in the previous Section, the perceived credibility on an information source (e.g., an article on Wikipedia) is often related to the verifiability of the content on external trusted sources. Accordingly, Wikipedia contributors have self-organised to devise a set of best practices to improve the perceived credibility of the articles [27].

From the perspective of contributors, the lifecycle of improving the credibility of a Wikipedia article can be synthesised into three steps: (1) a contributor challenges some part of the article's content; (2) the same contributor triggers a discussion over the challenge, which yields research tasks to find appropriate references; and (3) those references are added to the article’s content. This iterative process is depicted in Figure 1.

Next, each one of the steps in the lifecycle is discussed.

3.1 Challenge Material

Several Wikipedia contributors are actual curators of the articles they follow. Wikipedia itself, as an entity, supports and fosters this curating spirit as a way to improve the quality of the articles through verifiability [28]. The credibility of Wikipedia is only ensured if readers are able to discern what articles convey information that can be trusted or not. For this purpose, Wikipedia provides a set of templates that can be placed by contributors on articles where appropriate [26]. For the purposes of this paper, the most relevant templates are presented in Figure 2.

http://www.mediawiki.org

http://www.wikipedia.org

http://labs.google.com/accessible/
These templates focus on two main aspects of stating the insufficient accuracy and veracity of a Wikipedia article. By adding one of these templates to an article, contributors challenge the information conveyed on it with different semantics, as follows:

- **No references**: supported by the `{{unreferenced}}` template, contributors can mark an article as an unreliable source of information, as no references (i.e., external trusted sources) can corroborate the article’s content;

- **Additional references**: through the `{{refimprove}}` template, contributors state that some references verify that the article’s content can be partially trusted, but not on its entire piece.

Afterwards, contributors are encouraged to discuss the addition of these templates (on the article’s discussion page), in order to reach consensus about how to improve the issues raised on the quality of the article.

Furthermore, to better pinpoint which statements within the article’s content are susceptible of requiring the addition of external references, contributors have a set of footnote-oriented templates available to be embedded in situ, as depicted on Figure 3.

These footnote templates have the following semantics (ordered according to Figure 3):

- When some statement within the article’s content appears to be true, `{{fact}}` signals that a citation is needed to verify the stated information;

- In other cases when some statement might be challenged by contributors, `{{verify_source}}` further emphasizes that it is not consensual (therefore requires it to be verified externally);

- Lastly, when a given statement is supported by an existing citation or reference from an external information source, `{{verify_credibility}}` represents a challenge to its credibility.

As detailed in this Section, citing and referencing external sources that are credible, verifiable, and trusted, is the main credibility aspect of Wikipedia. By tagging an article’s content with the templates described above, readers can discern which information can be trusted or not. The capability of navigating to the external references puts on readers the onus of trusting the article’s content.

### 3.2 Research for References

The second step in the lifecycle relates to researching for references. As explained, Wikipedia contributors thrive for providing the best external information that can corroborate the different pieces of information that compose an article.

Through the discussion pages of each article, contributors start to find out which external references can help improving the credibility of the information conveyed in the article. Hence, contributors perform their own research, typically by navigating through different Web sites. Their own knowledge, experience, and intrinsic characteristics (i.e., abilities, disabilities) play a critical role on the set of references that are obtained through this process.

### 3.3 Adding References to Article

Finally, the third step of the lifecycle is centred on adding references to the article. Contributors collaboratively find appropriate references as explained above. Afterwards, referencing footnotes are created by substituting the placeholders defined through the footnote templates discussed in the first step of the lifecycle.

Contributors have several referencing templates at their disposal (with similar semantics to those of BibTeX [15]), including `{{cite book}}, `{{cite journal}}, and `{{cite web}}. Readers can follow the links provided in the footnotes to verify the credibility of the sources provided and, consequently, have their own opinion about the whole credibility of the Wikipedia article they are consulting.

Although the lifecycle is defined just through these three steps, the collaborative and always-improving nature of Wikipedia develops in an iterative way. For example, after a
certain reference is added to an article, collaborators can challenge the same part of the information the reference is attached to, or even challenge the nature of the reference itself. This will trigger another set of discussions, new research to be performed and, ultimately, new references to be added to the article.

4. EXPERIMENT

By detailing the process of adding credible references to Wikipedia articles, it becomes visible that it lacks the mechanisms and tools to take into account accessibility issues for the start. It is important to stress that the verifiability of credibility is a core concept of credibility itself, thus accessibility issues contribute to its verifiability property.

However, it is crucial to analyse whether in fact these accessibility issues arise on Wikipedia articles. It is not known whether contributors by themselves do consider accessibility issues when creating references, despite not existing a proper support in the article referencing lifecycle.

Therefore, we have devised an experiment over the accessibility of Wikipedia articles and the external links embedded in them (i.e., references and citations). This experiment was formulated to verify if such accessibility problems arise.

4.1 Method

We created an accessibility evaluation framework based on the UWEM methodology [20], which provides directions towards the automation of the evaluation process (which becomes relevant due to scalability issues). Our framework implemented a subset of checkpoints from WCAG (the Web Content Accessibility Guidelines) that can assess particular accessibility issues of Web pages’ HTML structures. Each checkpoint is responsible for analysing a specific detail on these structures (e.g., image tags without alternate text attributes). Afterwards, the outcome of each checkpoint contributes to a final, (quantitative) score representative of the accessibility level of the evaluated Web page.

We have based this experiment on previous results from large scale Web accessibility assessments [9]. Therefore, this experiment follows closely the setup devised for these assessments:

- $P = \{p_1, \ldots, p_{100}\}$, a set of 100 seed articles from Wikipedia, was randomly selected through a special purpose URL provided by Wikipedia$^4$;
- Each seed article’s Web page, $p_i$, was analysed in order to extract all hyperlinks that point to Web pages outside Wikipedia’s domains (i.e., we excluded all pages from Wikipedia sub-domains, such as hyperlinks to the article’s equivalent in other languages). A set $H_i = \{h_1, \ldots, h_n\}$ of Web pages was constructed based on this criterion;
- A pre-processing task was applied to each Web page (more specifically JTidy$^5$), which allows the verification of standard-conformance (i.e., whether it respects HTML specifications). This task also included healing processes on ill-formed Web pages, for structural analysis purposes. This task yields two important metrics from these verification and healing processes – errors (e.g., non-parseable markup) and warnings (e.g., non-fixable markup) – about the correctness level of Web pages according to standards, which is important for accessibility purposes due to the way assistive technologies (e.g., screen readers) interpret HTML structures;
- Finally, each Web page (both the initial set of pages and the external pages) is evaluated against a set of accessibility checkpoints $C = \{c_1, \ldots, c_n\}$ that can be automated, according to the WCAG methodology. When combining the results yielded from this process into a quantifiable result $A$, we opted to apply a simple evaluation metric adapted from [19]. While more sophisticated metrics have been proposed, e.g., in [21], it is out of the scope of this paper to detail on this aspect. The metric we used for this evaluation is presented in Equation 1:

$$A = \frac{1}{n} \sum_{i=1}^{n} c_i, c_i \in C$$ (1)

Here, each checkpoint $c_i$ returns 1 if its corresponding assessment was successful when applied to the Web page that is being analysed, or 0 if not. Consequently, the overall evaluation result $A$ averages all checkpoint values in the $[0, 1]$ range (i.e., a percentage), stating the accessibility quality level for the analysed Web page.

Next, we present the results of the conducted experiment.

4.2 Results

This experiment has resulted on the analysis of 365 Web pages on total, 100 of which (as defined by the experiment) come from Wikipedia, whereas 265 are related to external sources. Thus, on average, each Wikipedia article references 2.65 external Web pages.

Going into more detail on the analysis of the experiment data, 13 Wikipedia articles fully complied with the evaluation process (i.e., where 100% of checkpoints passed), making that 13% of the seed articles. In opposition, only 5 external Web pages fully complied, which is roughly 2% of the whole set of external Web pages used for citation and referencing purposes. Table 1 presents some statistics from the evaluation:

<table>
<thead>
<tr>
<th></th>
<th>Seed</th>
<th>External</th>
</tr>
</thead>
<tbody>
<tr>
<td>Average</td>
<td>84.14%</td>
<td>60.75%</td>
</tr>
<tr>
<td>Maximum</td>
<td>100.00%</td>
<td>100.00%</td>
</tr>
<tr>
<td>Minimum</td>
<td>64.29%</td>
<td>0.00%</td>
</tr>
<tr>
<td>Standard deviation</td>
<td>9.17%</td>
<td>36.16%</td>
</tr>
</tbody>
</table>

Table 1: Seed vs. external evaluation

The results presented on this Table show evidence of the disparity on accessibility quality between Wikipedia articles and referenced Web pages. On average, Wikipedia articles are 23% more accessible than referenced Web pages, and more stable (lower standard deviation). We believe that this stability is a by-product of the templating mechanisms of the MediaWiki software, as well as a strong emphasis on providing off-the-shelf accessibility for readers.

While the in both cases the maximum accessibility quality was attained, there is a great discrepancy between the
minimum quality. While the lower bound for seed articles is 20% below their average, external Web pages hit the bottom value, where no accessibility quality can be found. Furthermore, the highly variability on the accessibility quality of external Web pages (a standard deviation of 36%) does not help enforcing a stable accessibility quality on the references presented on Wikipedia articles.

On Figures 4 and 5 these findings are further detailed. In both cases, the X-axis represents the accessibility quality level (in percentage) for Web pages, whereas the Y-axis represents the number of Web pages meet those quality levels.

The shape of the graph for seed Web pages (i.e., Wikipedia articles) depicts the aforementioned stability of the accessibility quality. The “spike” centres around 80%, with a sharp decrease towards lower quality Web page count, and a more stable decrease on higher quality Web page count.

Regarding the shape of the graph for external Web pages (i.e., those used for referencing purposes within Wikipedia articles), the higher variation of quality is perceived. First, the initial spike shows that there is a high number of Web pages that do not meet the minimum requirements. This typically happens when pages are either totally inaccessible (e.g., HTTP 404 errors) or have severe parse errors. It is important to mention that the number of Web pages in this situation is 1.5 times more than the average Web page count for the large chunk of external Web pages analysed. From around 50%, the quality level distribution steadily increases until stabilising between 70% and 90%. Finally, the quality level decreases abruptly once again for the highest percentile (between 90% and 100%), which contrasts to the results of seed Wikipedia articles.

Another aspect that has been analysed in this experiment concerns the number of errors and warnings yielded by the parsing and pre-processing task described in the experiment setup. No Wikipedia article processing has resulted on a parsing error, which satisfies the high availability and good baseline quality of the MediaWiki software platform and its instantiation into Wikipedia. On the other hand, the high variability and uncontrolled nature of the Web has influenced the referencing mechanism used by Wikipedia article contributors. 66 external Web pages have resulted on errors, which accounts for nearly 25% of all external Web pages analysed. This aspect is manifested as the initial spike of the accessibility quality distribution presented in Figure 5.

As explained earlier, parse warnings are directly related to accessibility issues. Typical examples include unfixable markup, such as non-standard HTML tags. In such cases, accessibility-dependent readers are faced with a different information presentation which negatively influences the credibility given to the referenced Web page and, consequently, to the referral Wikipedia article.

The analysis of the seed Wikipedia articles has yielded 2323 warnings total, which accounts for an average of 23 warning per article. In contrast, external Web pages have summed up to 19523 warnings total, on an average of almost 74 parse warnings per Web page. This 31.5% increase on warnings is another indicator of the accessibility discrepancies between Wikipedia articles and the Web pages used for citation and referencing purposes.

A deeper analysis of this part of the experiment is presented in the graphs depicted in Figures 6 and 7. Each of these graphs present on the X-axis the number of warnings on a Web page parse process (represented in a logarithmic scale), and, on the Y-axis, the number of Web pages that have those warnings.

Regarding the seed Wikipedia articles, the quick decay depicted on the graph in Figure 6 is directly correlated to the templating and stable mechanism of the MediaWiki software. No Wikipedia article that has been analysed had no parse warnings. Nevertheless, from 10 warnings onwards, the number of articles is close to zero. This decay follows a logarithm, thus making the correlation between the number of Web pages and related parse warnings to follow the Power law for Wikipedia articles.

Regarding the distribution of warnings for external Web pages, the figure changes dramatically. The wilderness of the Web pays a significant role on the different distribution of warnings, which might be smoothed only by the influence of contributors on choosing the reference Web pages on Wikipedia articles.

The first interesting fact on this analysis is that there is a significant number of Web pages that have little to no parse warnings (in opposition to Wikipedia articles). How-
ever, this figure tends to shape differently afterwards. The biggest chunk of external Web pages orbit around 100 parse warnings, an order of magnitude above to those of Wikipedia articles. This means that, on average, readers are faced with 10 times more warnings from Wikipedia articles to external Web pages. This might pose severe accessibility problems to readers and, once again, negatively influence the credibility of Wikipedia articles.

5. DISCUSSION

The results presented in the previous Section have an impact on the credibility model for Wikipedia. In the light of these results, we discuss in this Section how contributors are unaware of accessibility issues, as well as how this knowledge can have design implications on Wikipedia and on its referencing lifecycle. Finally, we discuss the limitations of the conducted experiment, pointing out which directions can help leveraging more accurate results.

5.1 Contributor’s Unawareness

The first outcome of the conducted experiment concerns the high variability of the accessibility quality of external Web pages, i.e., ranging from no accessibility (or even unavailability) to an average distribution between 50% and 100%. Wikipedia provides a way to track unavailable links through the {{dead link}} template, resulting on more than 400,000 invalid links as of April 2007 [25]. However, in contrast to this situation, none is known about the accessibility of the available external links. This leads to the conclusion that Wikipedia contributors are unaware of the accessibility quality of the external Web pages they use for credible citation and referencing purposes.

These problems contrast with the results obtained from existing accessibility features within Wikipedia. Contributors have at their disposal a “higher order” markup which enforces several accessibility aspects (e.g., obligatory image captions). Furthermore, Wikipedia is also known for its self-organisation to support other inclusive practices, such as the existence of spoken articles6.

We believe that, with small changes to the lifecycle of Wikipedia article referencing, the increase of the accessibility quality of references can be easily supported by self-organisation practices, as long as supportive tools exist. We present in the next Section some possible solutions to this problem.

5.2 Design Implications

The introduction of new features in the article referencing lifecycle cannot be taken lightly. First, changes to the MediaWiki software can have repercussions on the stability of Wikipedia itself. Furthermore, this type of changes require an adaptation of policies and processes devised by Wikipedia contributors. The disruption of their corresponding lifecycles (e.g., such as the one described in this paper) can pose barriers to adoption of new features.

Therefore, we propose an incremental improvement to Wikipedia’s article referencing lifecycle. This improvement implies both technological and policy aspects on each one of the steps described earlier in this paper. It is important to notice that the addition of such aspects to the lifecycle does not disrupt current practices of improving the credibility of Wikipedia articles. It exploits the existing article referencing lifecycle to ease the burden of contributors to take into account accessibility issues from the start. Next, we revisit this lifecycle in the light of the findings from the conducted experiment.

Challenge Material. As detailed earlier, an important process of improving the credibility of a Wikipedia article concerns challenging the veracity of its contents. This policy allows readers to discern which facts stated in an article can be perceived as credible or not, and motivates contributors to find citations and references in authoritative and credible sources in external Web sites.

In the light of the available templates, we propose the introduction of new templates that challenge the accessibility of an article’s citations and references. First, we propose the addition of a {{refinaccessible}} template, which adds to an article the statement This article has inaccessible references. Please help improving the article by adding accessible references. Inaccessible references might be challenged and removed. This template will help readers to have a quick grasp that they might have problems on verifying the credibility of the facts presented in the article’s content.

We also propose $\text{\{verify\_accessibility\}}$, a footnote template to be inlined next to references within Wikipedia article's contents. This will tag a reference with an inaccessible source message, thus conveying the semantics of being an inappropriate reference for accessibility dependant readers.

**Research for References.** The second step in the lifecycle puts the onus on Wikipedia contributors the role of finding verifiable citations and references for all challenges raised in the previous step. While in these cases the credibility of a reference is thoroughly discussed in the article's discussion page, the same process has to be augmented in the case of accessibility issues. Since contributors are not required to be expert on accessibility evaluation procedures, there must exist an easy way for them to verify such aspects on external Web sites.

We propose that this step should be augmented with automated accessibility evaluation tools (such as those used in the experiment presented in this paper) that can be easily used by contributors. With quantifiable results as the outcome of such tools, contributors can discuss and develop new policies on the accessibility of citations and references. Examples of such policies include the definition of minimum quality thresholds for accepted references.

**Adding References to Articles.** Finally, the third step of the lifecycle concerns the addition of references to articles. Contributors perform this action by using citation templates such as $\text{\{cite web\}}$ or $\text{\{cite book\}}$.

By being the visible side of references within Wikipedia articles, these templates should also have a role on accessibility issues. Readers should be aware of the inherent accessibility quality level of the external Web pages of references. Hence, this awareness can exploit the same accessibility evaluation tools used by contributors in reference researching tasks. The quantitative output of these tools can be presented side-by-side with the references themselves, e.g., through star-ratings (1-5 stars).

The implementation of such features can be done by the MediaWiki software itself (e.g., adding the accessibility quality level to all external links, on all article edits), through bots [24] which automate the task of adding the information to references, or added by contributors by hand (based on the results from the previous step). In this case, citation templates should be adapted accordingly.

## 5.3 Limitations of the Experiment

The experiment that has been described earlier has its limitations. First, there is a small amount of accessibility checkpoints that can be performed automatically. This leads to a less accurate accessibility assessment. Nevertheless, it has been shown that automated accessibility evaluation tests such as those performed in the experiment follow the same pattern of results when more complete tests are performed by experts [21].

Another limitation concerns the relationship between each Wikipedia article and the external Web sites that are referenced within it. This study only analyses the two sets of Web pages as a whole, thus not taking into account the possible existing details of the hyperlinking structure. More research should be done to understand if there is a correlation between the accessibility quality of each Wikipedia article and external Web sites. This will allow studying, e.g., which contributors are more aware of accessibility issues.

Finally, the experiment does not make any kind of distinction between the content quality and ranking of external Web pages. No correlation has been studied between the accessibility quality level of an external Web page and its ranking on the Web (e.g., its PageRank value), especially concerning whether a certain reference conveys relevant information about the subject of the Wikipedia article.

## 6. CONCLUSIONS

This paper presented a study on the influence of accessibility on the credibility of Wikipedia articles. We have defined the lifecycle of references creation and, in the light of the results obtained, further improved the model with new actions and policies to improve the credibility of articles. While in Wikipedia, the accessibility quality is stable, the high variability of the Web is reflected in the citations and references of the analysed data. We believe that, with the improvements suggested in this paper, Wikipedia articles can have their credibility increased based on improved inclusive practices.

Ongoing work is currently being done in two fronts: (1) the analysis on the role of accessibility in the credibility of other authoritative Web sites (e.g., online newspapers, e-commerce), and (2) improving the analysis methodology to better differentiate the results between different accessibility and usability issues (e.g., blind users, mobile devices, etc.).

## 7. ACKNOWLEDGEMENTS

This work is funded by FCT through SFRH/BD/29150/2006 and the Multiannual Funding Programme.

## 8. REFERENCES


