Community Informatics for Developing Countries: Understanding and organising for a participatory future information society

CIDC 2006
Proceedings of the conference

Editors:
Andy Bytheway, Corrie Strümpfer

Main conference sponsor: Department of Communications

Microsoft Corporation

Conference refreshments: Ceres Fruit Juices

CIDC 2006 is endorsed by the Community Informatics Research Network, hosted by the Cape Peninsula University of Technology and organised by The Information Society Institute
CRAWLER VISIBILITY AND HUMAN USABILITY OF A GOVERNMENT SERVICES WEBSITE FROM A TECHNOMUNITY ANGLE

M Weideman

e-Innovation Academy (e-IA)
Cape Peninsula University of Technology
South Africa

Prof Melius Weideman
Cape Peninsula University of Technology
PO Box 652, Cape Town, 8000
meliusw@yahoo.com
ABSTRACT

The primary objective of this research project was to determine how the Cape Gateway Website measures up in the website visibility and usability stakes. The correct use of search engines could assist users in identifying and assessing relevant information in the shortest time possible. For this process to be successful, the search query must be formulated correctly. This involves the nomination of keywords, perceived by the user to be the most accurate in describing the information need. The way in which the website has been designed plays an important role. This is to not only ensure that it is indexed by search engines, but also to ensure that the website is ranked highly during a search. The Cape Gateway Website was analyzed and measured against generally accepted good practice measures in visibility elements, as well as using an industry standard program.

Secondly, the human element in website design involves recognition of the usage patterns and preferences of the human visitor. Much research has been done about user-friendly design; reading patterns, click rates, etc. However, it is also claimed that the implementation of these issues in real-life websites lag behind the theory. A number of usability tests have been done to determine if this claim is true for the chosen website. Results seem to indicate that the website under consideration makes better use of usability than of visibility techniques. As a result, it might not reap the maximum benefit possible in terms of Internet exposure, and to a lesser extent of providing a satisfying user experience.

The primary conclusion reached is that implementation of basic good practice of both issues at hand will improve the service rendered by this website to the community it is serving.
BACKGROUND

It is claimed that ‘Digital government projects are bringing together researchers from the technical and social sciences to participate in interdisciplinary and multidisciplinary collaborations …’ (Delcambre et al 2005, p. 33). It is this movement that has lead to the coining of the term ‘technomunity’: the role of technology in the community. The purpose of this research was to determine how successful a government project is in marrying technical and human user aspects in the production and maintenance of the Cape Gateway Website (CGW).

Shortly after access to the Internet has become commonplace, research started appearing with guidance on the design of websites containing large amounts of data. A number of different genres were identified, based purely on the size of the website (measured in the number of webpages). A total of eight genres were listed, ranging from 1-10 webpages on the small side to ‘more than 5 000 000’ webpages on the large side. However, government websites were not included in the many examples of the kinds of existing websites (Shneiderman 1997, p. 10). This seems to indicate that government services offered via websites is a relatively new offering on the Internet platform.

Government services

It is claimed that: ‘Increasingly, e-government is described as a revolution changing the public sector across the European Union and the rest of the world.’ (Criado et al 2003, p. 192). Focussing on Africa, another study was done to determine to what extent e-government services have been established in Kenya, Tanzania and Uganda (Kaaya 2004, p. 438). These studies confirm the recency of government services being established on an e-platform. The web’s attraction of large numbers of users, hopefully changing into paying customers, was initially noticed and used by the e-commerce world. However, this has started to change as countries moved into e-government, offering more and more web-based services (Barnes et al 2003, p. 297). The recent bold movement of SARS (South African Revenue Services) to enable clients to submit their tax returns online is a case in point.
In summary; although the movement of governments to providing e-services is rather young, it appears to be gaining momentum.

**Website visibility**

Research in the area of making websites visible to search engine crawlers have increased dramatically since the e-commerce bubble of the early 2000’s. One study proposed a model to serve as guidance towards creating websites that are more visible to crawlers (Chambers 2005, p. 128). Sherman (2006) claims that the UK market for search engine marketing grew by 100% in 2005, and could increase to be worth $2.6 billion in 2006. In general it is claimed that special steps have to be taken to ensure that search engine crawlers not only visit but also index a website, if it is hoped that human users should find this website. At the same time, it is claimed that approximately one third only of users are successful in finding relevant information on the Internet during unassisted searching (Voorbij 1999, p. 604, Weideman 2005, p. 13).

The Google search engine has had major success with its AdWord system, whereby advertisers can bid for and buy keywords, which could give them top rankings (in a separate area marked ‘Sponsored Links’) when users search for those keywords. Yahoo! has just joined the fight for advertiser’s money by offering their ‘Panama’ advertising system (Sullivan 2006). These developments indicate that the search engine user base has enormous potential to generate income for the owners of certain websites. It is a given that a website must be indexed by a search engine before any user can find that website on the search engine.

**Human usability**

As far back as 2002, the UK Cabinet Office issued a lengthy document, prescribing how UK government websites should be designed to ensure universal accessibility. Bottom level detail is provided in this document – down to coding conventions and catering for audiences with hearing, motor, cognitive, visual and other disabilities (Cabinet Office 2002). At the same time, Palmer (2002a, p. 102) suggested that website designers are compelled to maintain its base of users by offering interfaces that address specific needs
Crawler visibility and human usability of a government services website from a technomunity angle and functions. The same author states that the metrics which should be used to measure website effectiveness include download delay, navigability, site content, interactivity and responsiveness (Palmer 2002b, p. 151). Zhang et al (2001/2, p. 9) confirms this claim by stating that customer-centred websites should be created, rather than technology-driven ones.

The mere existence of conferences such as ‘Web Site Usability & Accessibility for Government’ indicates that there is a need for government websites to be easy to find and easy to use (Government Exchange 2003). Other authors proposed a two-factor model to establish website design and evaluation principles. The two factors are ‘hygiene’ and ‘motivators’ (Zhang et al 2000, p. 1253). Those factors that make a website usable are classified as ‘hygiene’ factors – their absence would frustrate users. The ‘motivate’ factors are those that add value to the website.

The Cape Gateway website

This website went online on 30 March 2004, and is a ‘…government service aimed primarily at citizens of the Western Cape, providing information on local, provincial and national government. An initiative of the Provincial Government of the Western Cape (PGWC), in partnership with national and local governments, it gives easy access to government information and services’ (Cape Gateway 2006). As such, it provides citizens with a ‘one-stop-shop’ for all their information needs, instead of these citizens having to know which department to contact for a given service (see Figure 1).

Figure 1 – CGW website homepage (Cape Gateway 2006)
GOVERNMENT SERVICES

In 2003 it was claimed that 68% of US federal agencies offered their services online, while 44% of state agencies did so (West 2003). Steyaert (2004, p. 369) posed the question of whether or not a marketing model could be used to improve the content and value of US state agencies’ electronic services. Thomas et al (2003, p. 83) claims that more and more citizens visited US government websites. However, it appeared as if most of these visits involved only the obtaining of information. Interaction with governmental services was lacking, although many visitors were satisfied with their experience. An alarming finding was that visitor demographics suggested that the digital divide was more pronounced amongst government website visitors than Internet users in general. It appears as if the US has established a solid presence in the offering of government services online, but that full user interaction has not yet been established.

Some authors claim that website design in government websites is a challenge, since it is important to design them in such a way that citizens can find information easily. A model and application instruments are proposed to assist with this process (Wang et al 2005, p. 129b). It is also stated that the websites of e-governments are of the largest in existence, with the largest numbers of users and information providers (Wagner et al 2006, p. 40). This author further suggests the use of Wiki technology to enhance the creation of massive e-government websites.

Political parties around the world have identified a website as a vital campaign and communication tool (Conway et al 2004). Political observers debate the impact of new technology on public-sector service delivery and citizen’s attitudes (West 2004, p. 15). However, the degree to which citizens are willing to adopt the innovation of government initiatives such as online voting and licence renewal, determines the success of these services (Carter et al 2005, p. 5).

CRAWLER VISIBILITY

SEO (Search Engine Optimization) is an established practice that improves the chances of a given webpage to be indexed by search engine crawlers (Wikipedia 2005).
Webpage designers attempt to manipulate the results appearing after a web search by designing the webpage in a certain way (Oppenheim et al., 2000, p. 194). Other authors state that SEO is a complex and sophisticated practice (Zhang et al 2004, p. 666), while Google (n.d.) warns that certain SEO practices are unethical since they attempt to influence search engine result page in ways which short-changes the user. Some of these practices include cloaking, doorway pages, invisible text, artificial link farms, keyword stuffing, hidden links, page redirects, and others (Mbikiwa 2005, p. 38-50).

In general, it can be stated that any website containing information and services of importance to an audience of reasonable size, should be indexed by the most important search engines. This warrants some time and money to be spent on implementing acceptable and ethical SEO practices, while attempting to not raise the spam and other red flags from search engines (Mbikiwa 2005, p. 48).

**HUMAN USABILITY**

Many research projects have been completed in the general field of website usability. For example, one study determined that an increased level of activity on a website has a positive effect on users’ satisfaction and general attitude towards a website (Hock-Hai et al 2003, p. 281). The highly competitive e-Commerce market is used as a benchmark to claim that websites should be easy to use. In this case, the alternative could spell financial failure, since if potential customers find it difficult to for example navigate a website, they simply migrate to another one which is not (Huei Huang et al 2005, p. 174a).

Research has established that a human visitor forms an opinion about a new website within the first 50 ms (Lindgaard et al 2006). Visual appeal plays a role here - if a website does not grab the user’s attention virtually on first sight, the chances are that a potential customer is lost. Other authors claim that no fewer than 91 website components play a role in the functions offered by a website. The major categories for this collection were: navigability, security, respect, accreditation, identification, image, expertise, products, services, external marketing, contract, transaction, customer service,
customer loyalty, investors, partners, internal relations and monitoring and leadership (Boisvert et al 2006, p. 178).

In summary it can safely be said that a website should be easy to navigate, appeal to the general user and satisfy the information or service need that drove the customer to the website in the first place. If not, the user/potential customer will simply find another website that does satisfy this need.

METHODOLOGY

An independent industry expert in SEO claims that ‘The clear leader in the perfect page numbers game is WebPosition…’ – referring to the program Web Position Gold, used to determine a webpages’ visibility to search engine crawlers (Sullivan 2000). After extensive previous experience on similar projects with this program, the author decided to use this tool to measure CGW’s visibility and other issues. An extensive test session produced a multitude of statistics, from which the results listed below were drawn.

A total of five relevant keywords were chosen to evaluate the visibility of the CGW homepage using Web Position Gold: government information, services, culture, red door and cape gateway. Seven search engines were identified as being worth using by Sullivan (2004); AltaVista, EntireWeb, Google, HotBot, Looksmart, MSN and Yahoo!, and the top local search engine was added to the list: Ananzi. For each search engine, the first 30 results yielded were considered. Consequently, a total of 5 (keywords) x 8 (search engines) x 30 (results) = 1200 listings were possible. Finally, a manual inspection of the HTML code of CGW’s homepage was done to determine to what extent basic good practice of website visibility was implemented.

Secondly, the guidance of an academic and world-leader in industry was used to measure CGW’s human usability (Nielsen 2006a). Nielsen has proven that webpage readers scan a webpage using the “F” pattern (see Figure 2) – top left to right, down the left side, to the right again (but not as far as the first time), then further down the left. Exceptions to this pattern include users reading search engine result pages (right third of
Crawler visibility and human usability of a government services website from a technomunity angle

Figure 2), where many shorter sidebars were evident. Red areas in Figure 2 indicate the most often read sections, yellow lesser often-read and blue least-read sections. The weighting of the relevant areas on the CGW homepage were evaluated by superimposing the F-pattern on the first visible area of this webpage.

![Figure 2 – The F-Pattern heat map in reading webpages (Nielsen 2006b).](image)

Thirdly, some of Nielsen’s real-life tests were used to evaluate CGW: ‘Top Ten Web Design Mistakes of 2005’ (Nielsen 2005).

In an attempt to quantify these measurements, a basic scoring system was designed. Points were allocated to each measure, totalled and expressed as a percentage at the end. The allocation of these points is discussed at each point of measurement below.

**Limitations**

Only the homepage of CGW was inspected, due to overhead limitations. Some of the measurements were subjective, and were based on personal experience and perceptions. Ideally the Web Position Gold statistics could be drawn over a period of time, to determine if a pattern is present.

**RESULTS AND ANALYSIS**

**Website visibility**
Out of the possible 1200 listings, CGW achieved only 12 hits, which is a success rate of 1%. These hits include webpages from the CGW website other than the homepage. Further analysis of these 12 results is provided in Table 1.

<table>
<thead>
<tr>
<th>KEYWORD</th>
<th>SEARCH ENGINE</th>
<th>government information</th>
<th>services</th>
<th>red door</th>
<th>culture</th>
<th>cape gateway</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>AltaVista</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1, 2</td>
</tr>
<tr>
<td></td>
<td>Ananzi</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>EntireWeb</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>Google</td>
<td>0</td>
<td>0</td>
<td>17</td>
<td>0</td>
<td>1, 2</td>
</tr>
<tr>
<td></td>
<td>HotBot</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1, 2</td>
</tr>
<tr>
<td></td>
<td>Looksmart</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>MSN</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1, 2</td>
</tr>
<tr>
<td></td>
<td>Yahoo!</td>
<td>0</td>
<td>0</td>
<td>28</td>
<td>0</td>
<td>1, 2</td>
</tr>
</tbody>
</table>

Table 1 – Keyword hits in eight search engines, using five keywords

The fact that the keyword ‘cape gateway’ registered a first and second position in five cases is considered to be of little value. A typical user looking for information on government services is unlikely to know the name of the central body that provides this service. The remaining two hits, both under the ‘red door’ keyword, are more useful, although both will appear below the screen limit on a typical result page due to their relative low ranking on the first page. If one considers only these two to be of value, the hit rate decreases from 1% to 0.17%.

One point was allocated for each potential hit in Table 1, yielding a maximum of 45 points. Therefore the actual points earned were 12 out of 45.

Furthermore, the manual inspection was done, in an attempt to evaluate certain well-known factors associated with high website visibility. The following positive factors became evident:
Crawler visibility and human usability of a government services website from a technomunity angle

- the title tag is used, and is meaningful (‘Cape Gateway - easy access to government information and services’) (Chambers 2005, p. 128)
- hyperlinks contain mostly sensible English text (Thurow 2003, p. 95)

A total of two out of two points were allocated for these two issues.

However, some negative ones were also present:

- no keyword metatag was used (Chambers 2005, p. 128)
- no description metatag was used (Chambers 2005, p. 128)
- Javascript was being used (Ngindana 2005, p. 26)
- a crucial piece of text (the title) is coded as a .gif file instead of crawlable English text (Weideman et al 2005)
- no site map was provided (Chambers 2005, p. 41)

A total of zero out of five points was earned here, producing an overall total for website visibility of: 12/45 + 2/2 + 0/5 = 14/52. This equates to a success percentage of 26.9%.

In summary, it appears as if CGW is not highly visible to search engine crawlers. The average user, looking for a website that will provide the kind of answers provided by this service, is unlikely to find it through Internet searching.

**Human usability**

Regular visual inspection of the test site revealed that the homepage was updated every few days, by rotating ‘latest news’ type items in the top bar of the webpage. This top bar fits neatly into the top of the F-shaped reading pattern, and seems to be the first item a reader would see when facing this webpage (see Figure 3). A point value of one out of one was allocated here. However, some identity confusion was evident in this area: three different ways of presenting the title are used within close proximity of each other. In the URL the title appears to be ‘capegateway’, the very first sentence that would be read by a crawler at the top contains ‘Cape Gateway’, while a large logo at the top left
Crawler visibility and human usability of a government services website from a technomunity angle announces that the site is that of ‘cape>gateway’, using two different colours. A point value of zero out of one was allocated here.

![Image](image_url)

**Figure 3 – The top bar of the CGW homepage (Cape Gateway 2006).**

Thirdly, the results from the ‘Top Ten Web Design Mistakes of 2005’ list (Nielsen 2005) follow. In each case, the heading for each ‘mistake’ was taken verbatim from the source. Since more than one sub-issue was evaluated under each heading, a point value between zero and two (indicated in square brackets) was allocated for each of the ‘mistakes’ Nielsen identified.

- **Legibility problems:** CGW has an easily readable font, the size was considered to be legible for most readers, and the contrast was chosen to be the maximum possible (black ink on a white background, except for hyperlinks). [2/2]

- **Non-standard links:** Hyperlinks are not clearly identified - links are in light green text and are not underlined, in contrast to the accepted standard of dark blue underlined text. An indication of visited links is provided, but in a similar light green colour – the contrast between the two greens was considered to be inadequate. The task of explaining to the user what to expect at the other end of each link was done properly- easy to read English text was used as anchors. Javascript is unfortunately used, which has a positive effect on the user experience, but detracts from the website visibility. Pages are not opened in new windows, which tie in with the human expectation. [1/2]
Crawler visibility and human usability of a government services website from a technomunity angle

- Flash: No evidence could be found of Flash being used, which is to the advantage of the user. [2/2]

- Content: The text on the homepage was well written in easy to follow English, while an Afrikaans and a Xhosa option are also offered. The Afrikaans homepage, however, mixes Afrikaans with English text, and some hyperlinks use identical English text to those on the English homepage. In general however, it was considered that the homepage answer potential users questions reasonably well. [1/2]

- Bad search: A sample of five (considered to be) typical searches was done on both the English and Afrikaans homepages, and the first page of the results rated. A score of ‘1’ was allocated if the result page contained at least one satisfactory and relevant answer to the perceived information need, or ‘0’ if no relevant answers were found – see Table 2.

<table>
<thead>
<tr>
<th>Keywords</th>
<th>English Results</th>
<th>Afrikaans Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>driving license</td>
<td>0</td>
<td>bestuurs lisensie</td>
</tr>
<tr>
<td>AIDS</td>
<td>1</td>
<td>VIGS</td>
</tr>
<tr>
<td>contact Mbeki</td>
<td>0</td>
<td>kontak Mbeki</td>
</tr>
<tr>
<td>crime prevention</td>
<td>1</td>
<td>misdaad voorkoming</td>
</tr>
<tr>
<td>housing Western Cape</td>
<td>1</td>
<td>behuising Weskaap</td>
</tr>
</tbody>
</table>

*Table 2 – CGW homepage search results in two languages*

In this case, a value between zero and two was allocated per row in Table 2, yielding a total of [6/10]. An interesting observation here is that the same keyword expressed in two languages does not always produce the same result (success or failure).

- Browser incompatibility: For technical reasons, this feature could not be tested.

- Cumbersome Forms: No user forms were found on the homepage, or after drilling down one level for each menu option from the homepage.

- Contact detail: More detail than the minimum expected was supplied – even a physical street address. This fact increases the credibility of the site. [2/2]
Crawler visibility and human usability of a government services website from a technomunity angle

- Frozen layouts: Resizing the homepage produced no unpleasant surprises, even after testing on a variety of different monitors. [2/2]
- Inadequate photo enlargement: Photo enlargement is not offered, so it could not be evaluated.

A total of $\frac{1}{1} + 0/1 + 2/2 + 1/2 + 2/2 + 1/2 + 6/10 + 2/2 + 2/2 = 17/24$ was earned for website usability. This equates to a success percentage of 70.8%. This figure seems to indicate that website usability did receive some attention during the design of this webpage.

CONCLUSION AND CHALLENGES

Whether or not the crawler visibility of a given website is important, depends on the purpose, the target audience and the webmaster of that site. If the purpose is to generate income, the target audience is large and the webmaster (for some reason, i.e. information dissemination) wants the site to be found, then website visibility is important, and it should be vigorously pursued. If the purpose is to list one’s hobbies, the target audience is undefined and the webmaster does not really care or know about visibility, then effort and money should not wasted on achieving a high website visibility.

In the case of CGW, it is the author’s belief that this website deserves to be visible. Although it does not generate income directly, the target audience is very large (probably every South African citizen with Internet access, even though only at times), and the webmaster/owner of the website should want the information contained thereon to be widely distributed. It should also be targeted for a collection of relevant keywords, making it easy for the average user to find.

However, it appears as if the CGW website does not manage to use technology to the full benefit of the community. The figure produced by the scoring system was 26.9%, indicating low website visibility to search engine crawlers. Most of the search engines
Crawler visibility and human usability of a government services website from a technomunity angle (including the major ones) do not even have CGW listed in their index. The challenge here is to lift the website from obscurity to become electronically visible to the community of South African citizens.

Secondly, where visibility can be considered an ‘option’, usability is a ‘must’. No website that will be read by a few users or more should be difficult to use, whether it be navigation, use of colours, download speed, or any of the other factors forming part of website usability. The CGW website measures up reasonably well in the usability area, with a usability index of 70.8%. It does appear as if many of the usability issues were addressed in the design, although there is room for some improvement. The challenge here is to increase the usability of the CGW even more by applying some of the best practice guidelines to the benefit of the community of South African citizens.

It is suggested that some basic changes be made to this website, and both crawler visibility and human usability will be increased. Technomunity has materialized to some extent on the Cape Gateway website, but can easily be improved upon.

REFERENCES


Crawler visibility and human usability of a government services website from a technomunity angle


