



# Empirical evaluation of one of the relationships between the user, search engines, metadata and Web sites in three-letter .com Web sites

M. Weideman

Faculty of Business Informatics, Cape Technikon

[meliusw@yahoo.com](mailto:meliusw@yahoo.com); [www.mwe.co.za](http://www.mwe.co.za)

---

## Contents

1. [Introduction](#)
2. [Definitions and other research](#)
3. [Relationships](#)
4. [Methodology](#)
5. [Results and analysis](#)
6. [Conclusion and recommendations](#)
7. [References](#)

---

**Key words:** metatags, search engine, metadata, Web site visibility

---

## 1 Introduction

The Internet offers a plethora of data to the user. The magnitude is impossible to measure, and there is no single categorization scheme in place to enable easy access to this data. Although many programs exist to enable Internet users to explore the landscape (friendly browsers, free search engines), general consensus exists that navigating the Internet is not a straightforward task (Voorbij 1999:598).

There appears to be a link between an Internet user, search engines, metadata and Web sites, as these elements are defined in Section 2. The purpose of this research project was to inspect and report on one relationship between these four elements. Countless reports in research literature seem to suggest that Internet searching is a difficult, error-prone task. For example, Sherman echoes Stoll's sentiments by stating that the Internet has become the world's largest and most complex, chaotic and unstructured search space (Sherman 1999:54). Another author finds that retrieval and organization of materials on the Web are not standardized, while others claim that the WWW was not designed to support organized publishing of data, or the retrieval thereof (Coetzee 1999:9; Robertson 2003). Information

overload has arrived (Wyle 2004). This author regards this situation serious enough to warrant close inspection, experimentation and reporting.

---

[top](#)

## **2 Definitions and other research**

### **2.1 The user**

A skill that eludes many average Internet users is the finding of relevant information on the Internet in a short time (Chun 1999:135). However, many Internet users rely on search engines to find relevant data for a variety of purposes on a daily basis (Wallace, Kupperman and Krajcik 2000:75). Lancaster predicted a change in the searching process, as it was known decades ago, when he described a future scenario in 1978:

‘The scientist of the year 2000 will use a terminal in many different ways: to receive, transmit, compose and search for text, seek answers to factual questions, build files and converse with colleagues. The terminal provides a single entry to a wide range of capabilities that will substitute, wholly or in part, for many activities now handled in different ways’ (Lancaster 1978:13).

Lancaster’s ‘scientist’ can be viewed as today’s average PC user, his ‘terminal’ has become the desktop computer, ‘conversing’ has evolved into e-mail and his ‘different ways’ could be the shift from using an intermediate searching expert to doing the search himself.

However, some studies have shown that early information users were not all eager to become involved in the extraction of knowledge from an electronic source. Over three decades ago Jackson found that engineers are reluctant to use information sources. The United Engineering Information Service failed to elicit financial support from the engineering profession to establish its services (Jackson 1971:431).

To further complicate matters, Large, Tedd and Hartley (1999:5) warn that information seekers must not be treated as a homogenous group – they differ in many aspects, where their information retrieval experience level has a large differentiating effect. Other authors warn that there is a difference in the way that WWW searchers and traditional OPAC searchers work when looking for information (Jansen and Pooch 2001:244). In contrast to this finding, a remarkable similarity was found between the basic methodology of early online searching and modern Internet searching. Lancaster listed a four point sequence of steps to be followed for online retrieval: logging on, negotiating the search process, manipulating the results and logging off (Lancaster 1978:72).

How does the human searcher approach the task of finding information in a collection of information? How does he or she convert the perceived information need into a search strategy? What effect does experience have on the user’s perception of information usefulness? These questions about the user cannot be answered here, but have been addressed in great detail in other research literature. For the purpose of this project, an Internet user is defined as a person who uses search engines to find relevant information on the Internet.

### **2.2 The search engine**

Search engines provide the average Internet user with a (mostly) free, apparently easy way to find general information on the Internet. They are programs that offer interaction with the Internet through a front end, where the user can type in a search term, or make successive

selections of relevant directories. The search engine software then compares the search term against an index file, which contains information about many Web sites. Matches found are returned to the user via the front end. The index is updated regularly either by human editors or by automated programs (called spiders, robots or crawlers). Both humans and spiders simply collect information of new Web sites by visiting as many Web sites as possible, and then building them into the index (Weideman 2004b). The three components of a typical search engine (front end, index file and information collectors) have close parallels in the components of a typical information retrieval system, as defined by Lancaster many years before the Internet and search engines became freely accessible. A search engine front-end maps to the 'user system interface', the index file maps to the 'indexing subsystem' and the information collectors map to the 'document selection subsystem' (Lancaster 1978:331). However, one author claims that search engines are complex, trusted without being understood and that users simply deal with their answers without understanding why they receive those answers (Lynch 2001:17). Many authors from both the popular press and other sources have done evaluations, comparisons, measurements and a variety of other tests on a large number of search engines (Chun 1999:137-140; Hoff 2002; Nobles 2003; Oppenheim, Morris, McKnight and Lowley 2000:193-205; Page 2004). A list of commonly used search engines and meta search engines with their URLs and a short evaluation, is available on the WWW (Weideman 2004a). However, in the studies mentioned above, a variety of methods were being used, and it is impossible to compare the results of the different evaluations with each other. Oppenheim and others did a study on this issue, and identified the difficulty of establishing recall on the fast-changing Internet scene as being one of the major reasons why different evaluations were difficult to compare. The creation of a standard set of tools will solve this problem and will also allow the tracking of the variation in performance of a search engine over time (Oppenheim *et al.* 2000:90). The problem of measuring recall could remain unless the search engines are compared on closed databases where prior evaluations of pertinence of the contents have already been made.

Another study confirmed that users make little use of advanced search features. For example, a test on 1 025 910 queries submitted to the Excite search engine revealed that the following features were only used by small percentages of users, as indicated below (Spink, Wolfram, Jansen and Saracevic, 2001:227-229):

Inclusion (+)	2%
Exclusion (-)	0.001%
Boolean operators (NOT/AND NOT)	0.0003%

Yet another study indicated that searchers spent a relatively short amount of time searching for one topic; the average search session seemed to last between five and 10 minutes only (Cooper 2001:137). This could possibly be a reason for the lack of use of advanced operators.

From the work already done on search engines, it has become clear that these programs play an important role in the lives of Internet users. For the purposes of this study, a search engine is defined as a program that is accessible by any average user, and capable of accepting user input which defines the information it produces as output to this user.

### 2.3 The Web page

A large number of authors claim that the relationship between users and Web pages with relevant information, stored on the Internet, is not always a positive one: 'Currently, search is simply bad' (Sherman 1999:54). 'Some respondents seemed confused about what they were to report when asked to list query terms for their search' (Spink, Bateman and Jansen

1999:122). 'The user's ability to specify good search terms and create complex search queries to clearly and precisely capture relevant retrieval seems rather low' (Spink *et al.* 1999:125).

'As the World Wide Web grows exponentially, discovery and retrieval of useful educational material grows more problematic' (Sutton 1999:1191). 'Only 33% of the Internet users agree or strongly agree with the statement: 'It is easy to perform subject searches on the Internet' (Voorbij 1999:604).

In summary it is clear that Web pages also play an important part in the life of the Internet user. For the purposes of this study, a Web page is defined as an entity, stored on the Internet, accessible by any average user, which contains some information.

## 2.4 Metadata

During the Middle Ages, indexing and simple classification of manuscripts were done. The indexers involved in this task were surrounded by an aura of mysticism, resulting from the coding schemes and alphabetical keys (Ingwersen 1992:61). Work on cataloguing started after this period, using written, guardbook and card catalogues, followed by the period of mechanization. This phase saw various card systems being developed, mostly in the form of punched cards: optical coincidence, the Zatocoding system of Mooers, edge-notched coding and 80-column punched cards. The high-powered computer era we currently find ourselves in provided the much-needed technology to empower this development, where document matching is made through inverted indices, string and positional searches. This increasing power of the technology has removed the economic constraint on the searching mechanism; any characteristic of the document could now be matched to a search query. In fact, there is now no technical constraint to prevent an index from including every single term of a given textual document in the index. This has been done some time ago, in the form of typical Bible concordances such as *Strong's Exhaustive Concordance*, first published in 1890. Since all the matching features so far have been based on text only, there remain some areas where matching has not yet been perfected, namely special notations such as music files, photographs and other graphical information.

The most ideal document representation is simply to include the whole document as the index, but the initial absence and later the limitations of technological tools (such as storage space) made this ideal impossible to achieve. In the absence of computer-based full-text retrieval systems, two main lines of thought were initially followed to provide a summary of the document: assigned and derived indexing. Assigned indexing depended on an indexer's perception of the document contents, the matching of that perception to a term list or classification and the selection of appropriate descriptors to be assigned to the document. Derived indexing depended on the indexer (later the computer program) to extract terms as used in the document – key word in context is the most notable early example.

During the late 1950s and 1960s, noted authors in this area of document content presentation did landmark work. The controversial Uniterm system sparked interest in the UK and the USA, which led to the Cranfield tests (Cleverdon, Mills and Keen 1966).

Documents were indexed via a single term in the Uniterm system (hence the name), having been extracted from the document title or abstract. After some structured tests, Uniterm results were compared to those using more traditional indexing methods. The test apparently broke down owing to the disagreement over relevance judgement, and the results were inconclusive. One group of testers claimed that the Uniterm system worked well, while the other claimed the exact opposite (Tonta 2004).

The actual Cranfield series of tests was done at the College of Aeronautics, Cranfield, UK. It involved another comparison of performance, this time between the Uniterm system and a modified version of the Universal Decimal Classification (UDC) system. A subset of 200 documents on aeronautics was extracted and used as the master collection of documents. A further extract of 40 documents from this set was made, and these documents were used to generate a set of 40 artificial requests. The assumption was that if query number one was submitted, document number one should be returned as being the most relevant document from the collection of 200 documents. Although certain inherent limitations of the study were evident, it did prove the effectiveness of the Uniterm system above the UDC classification. Modern day versions of metadata include HTML metatags and the DCMI (Dublin Core Metadata Initiative). More recent work indicates that the use of these metadata elements plays an important part in making Web pages more visible. A large number of authors make a case to urge Web site authors to make use of HTML and the DCMI metatags respectively. The HTML TITLE, KEYWORD and DESCRIPTION metatags have been singled out as being most relevant to Web site visibility (Altavista 2002; Washington University 2004).

Again it is clear that metadata play an important part in making Web pages more visible to the user. For the purposes of this study, metadata are defined as data about a Web page, contained in the Web page, used to describe the contents and other features of it.

---

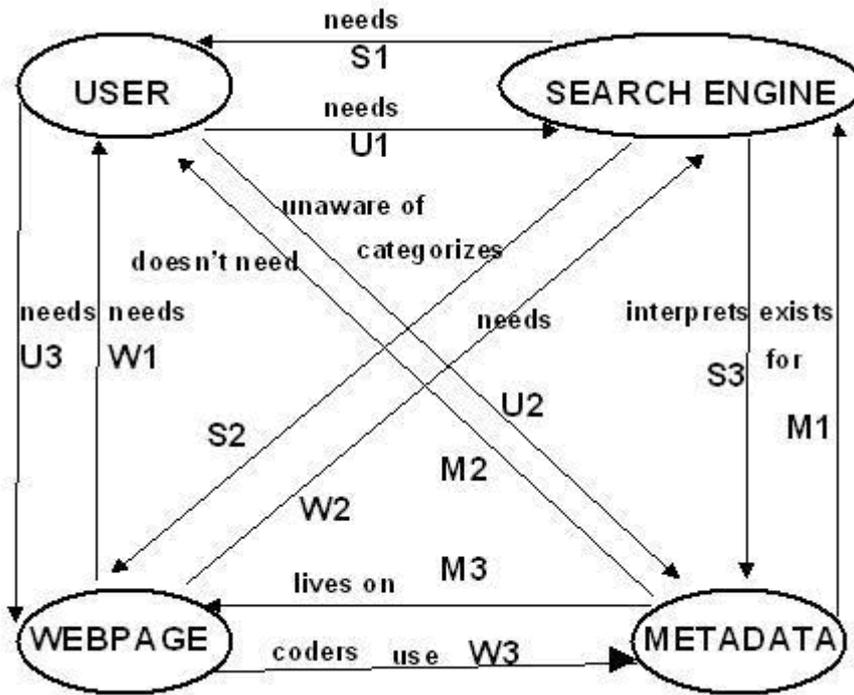
[top](#)

### **3 Relationships**

The four entities identified above are related to each other in a true quadrilateral fashion. The author intuitively defined a set of relationships between these four. No claim is made in this project as to the accuracy of these definitions, as listed below. However, the author had reason to believe that at least two of them could be invalid as initially defined: S3 and W3. The purpose of this project was to inspect and report on the accuracy of one of the definitions of these two relationships. The remaining eleven could be inspected and criticized in detail in future research.

The nature of each relationship is given by some action term(s) in Figure 1, and these terms is used to complete an implied English sentence. A brief discussion clarifies each relationship.

**Figure 1** Relationships between the user, search engine, Web page and metadata



### 3.1 Relationship U1

‘The user NEEDS the search engine.’

As discussed above, many users rely on search engines to provide them with timely, relevant information. Thus the user appears to be dependent on the search engine for timely and relevant information.

### 3.2 Relationship U2

‘The user is UNAWARE OF the existence of metadata.’

The average user, as defined earlier, is not aware (and does not have to be) of the existence of metadata. The creation of metadata falls in the domain of the Web page designer and the search engine software.

### 3.3 Relationship U3

‘The user NEEDS the Web page.’

The Internet searcher is prepared to spend time and money (for access, specialized search services, individual responses, printing, etc.) to obtain important information from the Internet. This ‘relevant data’ manifest as the content of one or more Web sites, as coded by the author of one or more Web pages. Therefore the user is dependent on one or more Web pages to store this information.

### 3.4 Relationship S1

‘The search engine NEEDS the user.’

Most search engines provide their basic services free to users. In an attempt to survive financially, they use advertising through pop-ups, key word sales, free e-mail accounts with

advertising and other services to generate income. As a result, search engines need users to view their pages and respond to the variety of prompts offered to them.

### **3.5 Relationship S2**

‘Search engines CATEGORIZES Web pages.’

Before a search engine can provide information in an understandable format to its users, it has to find these Web pages, somehow summarize each one, store it in a database and intelligently extract it. Correct categorization of the information on a Web page is to the advantage of the search engine and the user.

### **3.6 Relationship S3**

‘Search engines INTERPRETS metadata during the categorization of Web pages.’

If metadata are provided in the coding of a Web page, some search engines could use them to categorize the contents of the Web page in its own database. If the Web page author used the metadata correctly, the search engine would ‘know’ what the Web page contains.

### **3.7 Relationship M1**

‘Metadata EXIST FOR use by search engines.’

The purpose of putting metadata into Web pages, apart from serving the occasional human source code reader, is to enhance search engines’ understanding of Web page contents.

### **3.8 Relationship M2**

‘Metadata DO NOT NEED TO BE UNDERSTOOD by the user.’

Since a designer creates metadata, to be used by a search engine, the average user does not need to understand or even know about their existence.

### **3.9 Relationship M3**

‘Metadata LIVE ON Web pages.’

The metadata systems considered here are coded as part of the header section of a typical HTML Web page.

### **3.10 Relationship W1**

‘Web pages NEED Internet users.’

The main reason why most Web pages are created is to serve users with information. Some of these pages are actually instruments of income generation. In both cases, the [owner(s) of] Web pages need users to view and respond to their Web site content.

### **3.11 Relationship W2**

‘Web pages NEED search engines.’

If an Internet user knows the URL of a given Web page, he or she can simply type it in the browser and visit the site. In many cases, however, Web pages are dependent on search engines to find and categorize them, before users have access to the information they provide.

### **3.12 Relationship W3**

‘Web page coders USE metadata to enhance the electronic visibility of their Web pages.’

The designer and/or coder of a Web page should ensure that metadata elements are present in a Web page to increase its level of electronic visibility.

---

[top](#)

## **4 Methodology**

The next step in the research was to inspect one of the two relationships identified earlier. Relationship W3 was singled out as being the topic of this experiment, since the literature seemed to indicate that this relationship is suspect.

### **4.1 Random URL set**

According to a landmark study done on Internet metrics, a total of 2564, or about 4,3 billion Internet servers could exist. Furthermore, each server could store one or more Web pages, typically an average of 289 (Lawrence and Giles 1999:107). As a result of these large values, all Web pages could not be inspected. A subset had to be identified, from which a sample would be drawn. This concept is based on the extraction of documents used in the Cranfield tests. The subset was defined as being all Web sites in the .com domain, with a three-letter midfield (eg. [www.aol.com](http://www.aol.com)). The possible URLs thus ranged from [www.aaa.com](http://www.aaa.com) to [www.zzz.com](http://www.zzz.com). A total of 17576 (263) different URLs of this format are in existence.

A list was created containing these 17576 URLs in alphabetical order, with a random number allocated to each one. Entries in order of appearance included: [www.aaa.com](http://www.aaa.com), [www.aab.com](http://www.aab.com), ... [www.aba.com](http://www.aba.com), [www.abb.com](http://www.abb.com) ... [www.zzy.com](http://www.zzy.com) and [www.zzz.com](http://www.zzz.com). The list was then sorted, using the random number as key. The entries in the resultant list were then loaded one by one, starting at the lowest random number. This process provided a random sequence of Web sites within the specified subset. A total of 238 URLs were loaded and inspected. Since 37 of them were invalid (all generated time-out errors), a total of 201 produced a Web page, of which the HTML coding was then inspected.

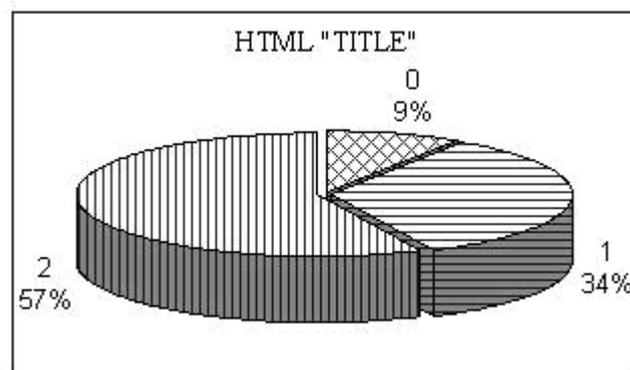
### **4.2 Scoring scheme**

A decision was taken to inspect the occurrence and usage of those metatags considered to be relevant to Web site visibility. This includes the HTML TITLE, KEYWORD and DESCRIPTION metatags, as well as the following DCMI tags: TITLE, SUBJECT, DESCRIPTION, TYPE and SOURCE. A Web page would earn a score of 0 for each one of these metatags if it were not present at all. A score of 1 would be allocated if the tag was present, but not used effectively (e.g. title totally irrelevant, too many keywords, no text inside tag, text contains excessive forms of spam, spelling mistakes in important key words, etc). A score of 2 was allocated if the element was present, and if it was used effectively. The Web page score therefore ranges between 0 and a maximum of  $8 \times 2 = 16$ . The author used the judgement of Sullivan (2004) plus personal experience based on many years of exposure in the field for the subjective issues of effective use of metatags just described.

## 5 Results and analysis

Inspection of the 201 Web sites' HTML coding produced not one instance of the usage of any one of the DCMI metatags. However, a number of HTML tags were found. As a result, it was decided to omit the five DCMI metatags from the scoring system, since it would simply produce a score of 0/10 for all 201 Web pages. That left the three HTML metatags only to produce scores. The maximum score for a Web site was therefore altered to  $2+2+2 = 6$ . The header section of every valid URL in the sample was inspected, a judgement was made on the effective usage of each one of the three metatag elements, the scoring system was applied, and the results summarized.

**Figure 2** Usage of HTML TITLE tag



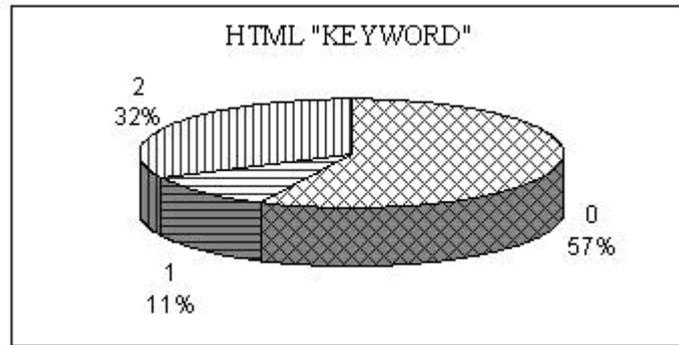
### 5.1 TITLE metatag

From Figure 2 it can be seen that 57% of Web pages made effective use of the TITLE metatag and 34% had this tag in the coding of the Web page, but did not use it correctly or effectively. Only 9% did not use it at all.

In summary, only approximately 10% of the sample did not use the TITLE metatag, while over half of the Web pages used it effectively. Since this feature is used to identify a Web page in the result listing produced by most search engines, users would at least be assured that the title of the Web page, as described by the coder, would be represented accurately in 57% of all cases. However, the value of having a correct title displayed could be diminished by the lack of other, more useful indicators like the contents. It remains unknown why ALL Web page do not make use of this simple feature to enhance Web page visibility.

### 5.2 KEYWORD metatag

**Figure 3** Usage of HTML KEYWORD tag

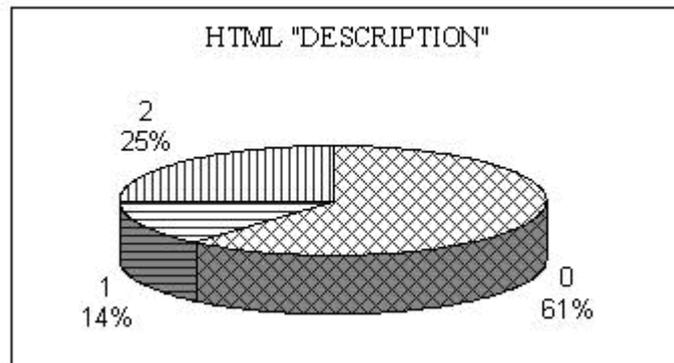


From Figure 3 it can be seen that 32% of Web pages in the sample made effective use of the KEYWORD metatag and 11% had this tag in the coding of the Web page, but did not use it correctly or effectively. A total of 57% did not use it at all.

In summary, only 33% of the sample Web sites made effective use of key words to draw visitors to their site. A total of 68% made no or ineffective use of this metatag. This figure is lower than the equivalent figure for the TITLE metatag, indicating apparent indifference by Web page authors.

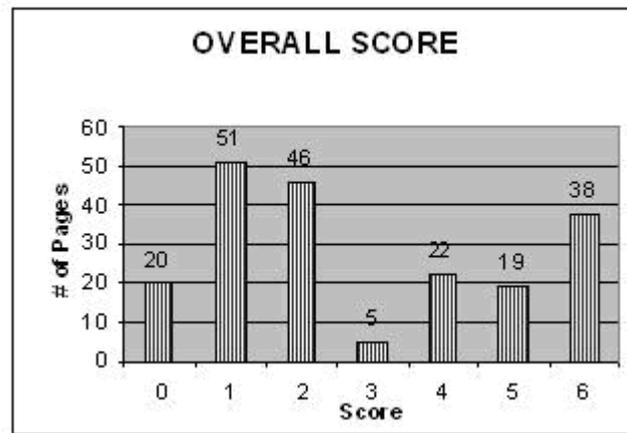
### 5.3 DESCRIPTION metatag

**Figure 4** Usage of HTML DESCRIPTION tag



From Figure 4 it is clear that ONLY 25% of Web pages in the sample made effective use of the DESCRIPTION metatag and 14% had this tag in the coding of the Web page, but did not use it correctly or effectively. A total of 61% did not use it at all. Once again there appears to be an even lower amount of effective usage of this metatag, compared to the previous two. Only 25% of the sampled Web pages used it effectively, leaving the other 75% without it, or using it ineffectively.

**Figure 5** Overall scoring



#### 5.4 Overall score

From Figure 5 it can be seen that 117 out of 201 Web pages (58.2%) earned a score of two out of six and lower, while only 38 out of 201 (18.9%) earned the maximum of six out of six. A total of 20 Web pages, or close to 10%, did not use any one of the three metatags in any way.

[top](#)

### 6 Conclusion and recommendations

To summarize some of the basic results listed above:

- Average of Web pages which used none of the three metatags at all: 42.3%
- Average of Web pages which used the three metatags effectively: 38%
- Web pages which earned the minimum score: 9.95%
- Web pages which earned the maximum score: 18.9%
- Average overall score was 3,75 out of 6, or 62.5%.

The KEYWORD metatag seems to be ignored by all search engines except the Inktomi database (used by a number of search engines). This could explain the relatively low usage of this metatag.

It is therefore concluded that the relationship W3, as defined in Section 3, is invalid. Web page authors do not use metadata as a rule to enhance the electronic visibility of Web pages.

The figures above indicate a lack of application of a basic methodology that could increase Web page visibility. It is recommended that company management insist on the implementation of these metatags on mission critical Web sites. If this potential visibility enhancing mechanism is ignored, potential income opportunities are neglected.

At the same time, the other relationship identified earlier as being suspect (S3), should be investigated in future research. If search engines fail to recognize metatags on a large scale, Web page authors have little reason to include them in their coding efforts. However, this can only be determined by an empirical investigation into the recognition of metatags by search engines. At the time of writing, a pilot literature study seemed to indicate that search engines were not attaching as much value to metatags as in the past.

Finally, the total absence of Dublin Core metatags indicates either ignorance on the part of designers, or lack of usefulness of these tags. Further research could shed light on this topic.

## 7 References

- Altavista, 2002. Search help. [Online]. Available WWW: [http://help.altavista.com/adv\\_search/ast\\_haw\\_metatags](http://help.altavista.com/adv_search/ast_haw_metatags). [Accessed 1 July 2002].
- Chun, T.Y. 1999. World Wide Web robots: an overview. *Online & CD-ROM Review* 23 (3):135-142.
- Cleverdon, C.W., Mills, J. and Keen, E.M. 1966. *Factors determining the performance of indexing systems, Vol 1. Design, Part 1*. Cranfield: Aslib Cranfield Research Project.
- Coetzee, H.S. 1999. The Dublin core initiative as a metadata standard for information retrieval on the Internet. *Meta-info Bulletin*8(4):9-10.
- Cooper, M.D. 2001. Usage patterns of a Web-based library catalog. *Journal of the American Society for Information Science and Technology* 52(2):137.
- Hoff, S.R.C. 2002. *AltaVista vs. Excite vs. Hotbot vs. Infoseek: which is the one to rely on?* [Online]. Available WWW: <http://www4.zdnet.com/pccomp/srchoff/srchoff.html>. [Accessed 1 October 2003].
- Ingwersen, P. 1992. *Information retrieval interaction*. London: Taylor Graham.
- Jackson, E.B. 1971. The engineer as reluctant information user – a remedial plan. *Proceedings of the International Conference on Training for Information Work*. Rome. November 15-19:430.
- Jansen, B.J. and Pooch, U. 2001. A review of Web searching studies and a framework for future research. *Journal of the American Society for Information Science and Technology* 52 (3):244.
- Lancaster, F.W. 1978. *Information retrieval systems: characteristics, testing and evaluation*. New York, NY: John Wiley.
- Large, A., Tedd, L.A. and Hartley, R.J. 1999. *Information seeking in the online age: principles and practice*. London: Bowker-Saur.
- Lawrence, S. and Giles, L. 1999. Accessibility of information on the Web. *Nature* 400:107.
- Lynch, C.A. 2001. When documents deceive: trust and provenance as new factors for information retrieval in a tangled Web. *Journal of the American Society for Information Science and Technology* 52(1):17.
- Nobles, T. 2003. *What will it be, directory or search engine?* [Online]. Available WWW: <http://www.phoenixmlm.com/sengine.html>. [Accessed 1 December 2003].
- Oppenheim, C., Morris, A., McKnight, C. and Lowley, S. 2000. The evaluation of WWW search engines. *Journal of Documentation* 56(2):190.
- Page, A. 2004. *The search is over*. [Online]. Available WWW: <http://www4.zdnet.com/pccomp/features/fea1096/sub2.html>. [Accessed 16 May 2004].

- Robertson, S.E. 2003. *Conflicting philosophies*. [Online]. Available WWW: <http://www soi.city.ac.uk/research/cisr/ser/ucla/node3.html>. [Accessed 10 April 2003].
- Sherman, C. 1999. The future of Web search. *Online* May/June:54-57.
- Spink, A., Bateman, J. and Jansen, B.J. 1999. Searching the Web : a survey of Excite users. *Internet Research: Electronic Networking Applications and Policy* 9(2):122-125.
- Spink, A., Wolfram, D., Jansen, B.J. and Saracevic, T. 2001. Searching the Web : the public and their queries. *Journal of the American Society for Information Science and Technology* 52(3):229.
- Sullivan, D. 2004. *More about meta tags*. [Online]. Available WWW: <http://searchenginewatch.com/subscribers/more/metatags.html>. [Accessed 01 June 2004].
- Sutton, S.A. 1999. Conceptual design and development of a metadata framework for educational resources on the Internet. *Journal of the American Society for Information Science* 50(13):1191.
- Tonta, Y.A. 2004. *Failure analysis in document retrieval systems: a critical review of studies*. [Online]. Available WWW: <http://yunus.hacettepe.edu.tr/~tonta/yayinlar/phd/bolum-3.htm>. [Accessed 18 April 2004].
- Voorbij, H.J. 1999. Searching scientific information on the Internet: a Dutch academic user survey. *Journal of the American Society for Information Science* 50(7):604-605.
- Wallace, R.M., Kupperman, J. and Krajcik, J. 2000. Science on the Web : students online in a sixth-grade classroom. *The Journal of the Learning Sciences* 9(1):75.
- Washington University . 2004. [Online]. Available: <http://depts.washington.edu/trio/comp/howto/site/design/metatags.shtml>. [Accessed 1 June 2004].
- Weideman, M. 2004a. *Search engines*. [Online]. Available WWW: <http://www.mwe.co.za/home/seaengin.htm>. [Accessed 10 June 2004].
- Weideman, M. 2004b. *Newspaper articles*. [Online]. Available WWW: <http://www.mwe.co.za/home/searticles.htm>. [Accessed 8 June 2004].
- Wyle, M.F. 2004. *Information overload*. [Online]. Available WWW: <http://eiger.wyle.org/~mfw/diss/node12.html>. [Accessed 10 April 2004].

### **Disclaimer**

Articles published in SAJIM are the opinions of the authors and do not necessarily reflect the opinion of the Editor, Board, Publisher, Webmaster or the Rand Afrikaans University. The user hereby waives any claim he/she/they may have or acquire against the publisher, its suppliers, licensees and sub licensees and indemnifies all said persons from any claims, lawsuits, proceedings, costs, special, incidental, consequential or indirect damages, including damages for loss of profits, loss of business or downtime arising out of or relating to the user's use of the Website.



ISSN 1560-683X

Published by [InterWord Communications](#) for the Centre for Research in Web-based Applications,  
Rand Afrikaans University