
INSIDE YOUR FRIENDLY NEIGHBORHOOD SEARCH ENGINE

JIM CONIGLIARO | DIRECTOR OF ONLINE OPERATIONS & TECHNOLOGY | JOURNAL INTERACTIVE

Search engines are black boxes. You stick a few words in and out come the search results. Exactly what happens inside the black box to transform those words into search results tend to be closely guarded secrets. Unfortunately, this has resulted in copious amounts of disinformation, unfounded suggestions, a constant flood of “how to improve your search ranking” articles, and regular solicitations from people and companies promising to improve search rankings. Without knowing how search engines work it can be difficult to wade through all of these suggestions, solicitations, and products to determine what is or isn’t legitimate.

What I would like to do in this brief paper is shed some light on the inner workings of search engines. Even though the details for individual search engines are classified as trade secrets, you can still get a very solid understanding of the concepts by doing a little research. What follows is a description of how search engines (well, more specifically Google) work. I’ve relied on academic literature and patent information to offer insight into search engines. You will also notice that I have avoided referencing online guides, web sites, and blogs that rely on the claimed, but unfounded, experience of their authors. My intent is to arm with some good, well founded, information that will allow you to have informed discussions about search engine optimization and to equip you to siphon through the ever increasing flood of suggestions and solicitations that promise to improve your search engine rankings.

SOME BACKGROUND

If you want to really gain an understanding of what is going on under the hood of a search engine, you need to understand what it is that the search engine is really trying to do. Yes, return search results, but we are looking for something a little deeper than that. Let’s look at Google. From its inception, the mission of the Google search engine has been to eliminate irrelevant results from searches (Page & Brin, 1998). Aside from the general frustration of seeing junk in your search results, all that junk can camouflage the good results, making them challenging if not impossible to pick out. As living, breathing, thinking human beings, determining if a piece of content is relevant is not altogether difficult. For a computer it is a completely different story. To understand why this is such a complicated mess, consider the following:

Tidbit: Most people know that Google is named for the mathematical term for 10^{100} , what few know is that the company is named after a common misspelling of this term ... correct spelling for the number is googol.

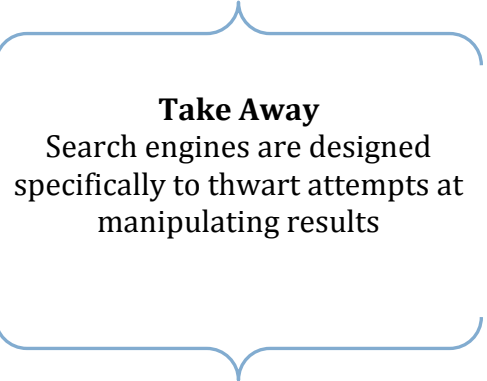
- A. Relevance is a subjective concept – it is difficult to make a computer understand subjective concepts.
- B. How do you keep people from tricking the computer and therefore manipulating the results?

The question of relevancy is the focus of a whole field of academic study called Information Retrieval (IR for short). Formally, information retrieval is the study of the storage, analysis, and retrieval of media (Baeza-Yates & Ribeiro-Neta, 1999). It is a complex, but well studied topic. Keeping people from tricking the computer/manipulating the results? That relates directly to the extent to which a search engine can be trusted and is the focus of researches working for the major search engines.

This whole mess of trying to give you, as the person typing in the search query, good results really has its roots in the fields of cognitive psychology and linguistics (Cilibrasi & Vatanyi, 2007). “Whoa! Time Out.” you say, “Psychology? Linguistics? I thought Google was started by a bunch of computer geeks.” It was. When you combine cognitive psychology, linguistics, and computer science you end up with a field called Computational Linguistics – a fancy way of saying that you are trying to

get computers to understand human languages. I’m not trying to offend the computer engineers of the word, but we’ve never been known for our skill with language and human interaction – so we enlist the help of people that are experts in language and communication. So with the help of their friends in the psychology and linguistics departments the computer engineers can begin to find patterns in how people classify, categorize, and find relevance in content. At the end of the day, this whole process becomes nothing more than an interesting math problem.

This is exactly what the folks at Google have done. According to Larry Page’s patent (he’s one of the geeks that started Google), the search engine attempts to determine if a web page is relevant in a manner that emulates how a human would make such a determination (Page, 2001). This serves not only to associate relevant content with a search but also to help guard against attempts to manipulate results by misleading the search engines (Page & Brin, 1998).



Take Away
Search engines are designed specifically to thwart attempts at manipulating results

A LOOK UNDER THE HOOD

So how does it work? You can ask the folks at Google, but they'll just give you a vague answer indicating that they take a large number of factors into account. It's a secret, but there are some good clues if you are willing to do some homework.

We have two concepts that need to be explored in relation to ranking a page: *authority* and *content*. Content is all the stuff on the web page – text, pictures, links, etc. Authority is really a measure of how the level of respected a page has earned. Authority and content are both used in conjunction with each other to determine the ranking of a page in a set of search results but what often surprises people is that the authority of a page is treated independently of a page's content (Bianchini, Gori, & Scarselli, 2005). The authority of a page is derived from the link structure of the web (Brinkmeier, 2006), whereas the use of a pages content is derived largely from classic information retrieval techniques that rely on relevance judgments (Mehlitz, Kunegis, & Albayrak, 2007).

MEASURING AUTHORITY

The algorithm that Google uses to represent authority is called the PageRank algorithm. The PageRank algorithm was first described in 1998 in two papers written by Lawrence (Larry) Page and Sergey Brin. In these papers, Page and Brin described a technique for determining the authority of a page based upon the link structure of the web. To understand how this works, let's think of this not in terms of web pages, but in terms of academic writing. If you write a really good paper, there is a good chance that over time a lot of other people will reference your paper. If someone finds your paper referenced over and over again as they are researching a topic, there is a good chance that your paper will be seen as having a high degree of authority. Now, let's take this a step further; this same person that has identified your paper as highly authoritative will subsequently view the references in your paper as more relevant because your highly authoritative paper referenced them.

Tidbit: PageRank isn't named so because it ranks pages, it is named after its inventor – Larry Page.

We can apply this same concept to a web page but rather than citations and references we have hyperlinks. If a lot of other web pages have hyperlinks that point to your page, your page is considered authoritative. The pages that your page links to also gain a little more authority because your page, as an authoritative source, is linking to them. Figure 1 shows how this works – size of the bubble represents the relative authority of a page. The concept is very straight forward.

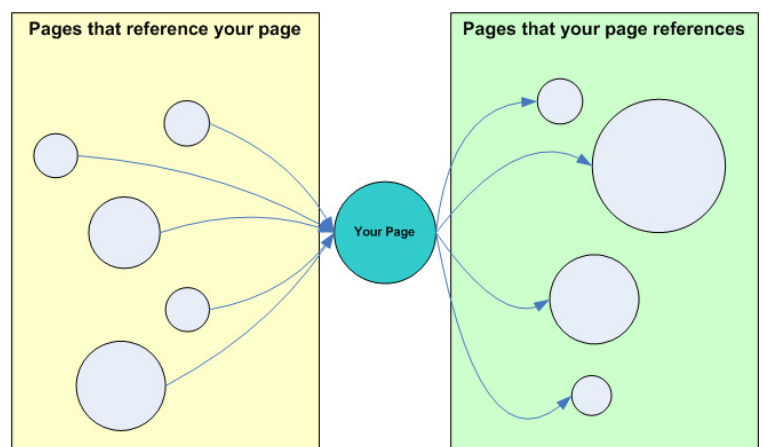


FIGURE 1: EXAMPLE OF USING LINK STRUCTURE TO DETERMINE AUTHORITY

While the concept is remarkably simple, turning the concept into reality is far more difficult. The math behind the algorithm is actually pretty straight forward, just a bunch of multiplication and addition. When you start thinking about what is actually going on, things get a bit nastier. To know how much weight to give a link coming in from another page, I need to know its authority. To know its authority, I need to understand the authority of all the links coming into that page, to know the authority of those pages, I need to know the authority of the pages that link to them, and so on. To make it more complicated, my page may link to a page that links to one of my incoming links.

This means that I need to know the authority of my page in order to compute the authority of my page. Not as easy as it first seemed. Now we are into topics like limits, power series, Markov chains, linear algebra, and all sorts of things we tried to avoid learning in college (see sidebar for some more details on the math). In fact, we can't exactly compute the authority of a page. We can guess though.

Take Away
The "Authority" of a page is based on how many pages and which pages link to your page/site.

Let's look at the problem from a different perspective. If we randomly pick some page on the World Wide Web, we can be reasonably confident that this page will have hyperlinks. We can randomly pick one of those links and click on it and then do the same on the subsequent page and so on. If my page has a great many other pages that link to it, then it is more likely that someone could accidentally stumble upon my page just by randomly clicking on links in web pages. Because someone may be more likely to stumble

upon my page, the odds of that person getting to a page that I link to go up. So we can actually view this as a problem of statistics. ***The authority of a page can be directly correlated to the probability that someone would stumble across your page by randomly clicking on links on the World Wide Web.*** This is no longer an exact measure, but rather a probability (Page, Brin, Motwani, & Winograd, 1999). It just so happens that this probability is directly correlated to authority and now we have the ability to compute the incomputable (Page & Brin, 1998).

Now we have an idea about how Google determines the authority of your page. This is fine if we simply want to rank all of the web pages on the Web in order of authority, but that doesn't really give

Crunching the Numbers

More interested in the mathematics behind PageRank? Here's the 10 second overview. We'll define the rank, r , of a page a as r_a . The set of all links that point to page a is $l[a]$, number of links on any given page q is c_q and the total number of pages on the web is N . There is also a damping factor, d . So, the page rank of a page a is:

$$r_a = \frac{d}{N} + (1-d) \sum_{q \in l[a]} \frac{r_q}{c_q}$$

(Page, Method for Node Ranking in a Linked Database, 2001) (Bianchini, Gori, & Scarselli, 2005)

So far, pretty simple - add - multiply - repeat. If you look closer, you will note that we can't directly compute this because of the recursion involved. Instead, we can use a Markov Model to represent this as a probability [Jurafsky & Martin, 2000]. We can then create an transition probability matrix A of size $N \times N$ in which each element of the matrix A_{ij} represents the probability that a user will go to page j from page i . So, if we were to compute the probability that someone ends up on page j after one click is:

$$p_1 = A p_0$$

Where p_i is the probability that the user starts on page i . You can repeat this computation for the second click, third click, and on to infinity:

$$p_\infty = \lim_{n \rightarrow \infty} A^n p_0$$

We are still left with the pesky infinite chain of computation, however, the equation above just happens to be the characteristic vector (eigenvector) of A . (Page, 2001). Eigenvectors are nifty little properties of a matrix that have all sorts of meaning in physics and engineering (Wylie & Louis, 1995). In this case, it happens to be the steady state probability that someone will randomly click on a link to a page (Page, 2001).

an indication as to how relevant a page is to a particular query. To do that, we need to examine content.

MEASURING CONTENT

So how does a search engine determine which pages are relevant to a keyword search? This is where that whole field of information retrieval comes into play. Let me say again, that the specifics of the algorithms involved are closely guarded secrets but by trudging through the literature and patents, we can get very solid understanding of what is going on under the hood.

The patent that was awarded to Larry Page specifically referenced the following factors: font size, font weight, URL, domain, author, institution, time document was last updated, if the document is a home page, and visibility of text (Page, 2001). Other papers written by Page and his colleagues at Stanford refer to features such as title, proximity of terms, position of word, capitalization, and anchor text (Page & Brin, 1998). You may have noted an interesting omission in these lists is anything that can be described as meta-tags, or summary information about the page. *Meta-tag* refers to content that embeds additional information about a web page (such as a list of key words or a description) into the document that isn't displayed the user.

Recall that one of the important characteristics of a good search engine is that it be impervious to attempts of page authors to gain attention for their pages through deliberate attempts to mislead automated search engines. (Page & Brin, 1998). Early search engines relied heavily on meta-data and were therefore easily manipulated by unscrupulous people stuffing their pages full of meta-data and hidden text that had no relevance to the information visible to a user (Noll & Meinel, 2007). For this very reason, modern search engines have all but abandoned the use of meta-tags as a factor search algorithms (Sullivan, 2002).

Take Away

Other factors that influence how a page is ranked in search results are largely related to qualities of a page designed for human, not machine, consumption. Qualities such as font size, font weight, title, word choice, etc.

When a search engine interprets a page, it takes the information it finds on the page and creates a set of features. The size of the heading is a feature, the positions of each word are features, the font weights are features, and so on. Each of these features can be represented numerically. When

Turning Words into Numbers

Information retrieval is a large and complicated topic but let's take a stab at some of the basics. Any document (web page) can be represented as a collection of numerically measured features. There is no full list of features of a document – this is one of the proprietary elements of a search engine – but they tend to be things such as the number of times a word appears in the document, the relative position of a word in a document, the density of a keyword, the relative size of the fonts, etc. These are all mechanisms of using the language in a document to create a numeric measurement.

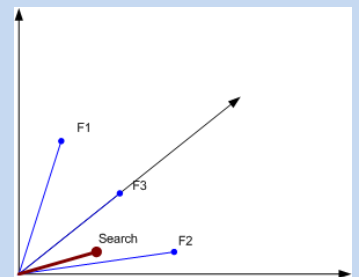
These features are represented mathematically as vectors. So if I'm looking at three documents, each with three features A, B, and C, my vectors are:

$$\begin{aligned} F_1 &= \{A_1 \quad B_1 \quad C_1\} \\ F_2 &= \{A_2 \quad B_2 \quad C_2\} \\ F_3 &= \{A_3 \quad B_3 \quad C_3\} \end{aligned}$$

If we factor in weighting, our documents look something like this:

$$\begin{aligned} F_1 &= \{w_a A_1 \quad w_b B_1 \quad w_c C_1\} \\ F_2 &= \{w_a A_2 \quad w_b B_2 \quad w_c C_2\} \\ F_3 &= \{w_a A_3 \quad w_b B_3 \quad w_c C_3\} \end{aligned}$$

Graphically we can represent these as:



If we normalize these vectors and take the dot product between any two vectors what we actually end up with is the *cosine* of the angle between the vectors. A cosine of 1 represents a perfect match, a cosine of zero represents no match. Easy, right? Well, despite the brackets and subscripts, the math tends to be the easy part. Figuring out how to numerically represent the features, which features, what weights and how to map the search term to those features is the hard part. That's why we work with linguists and cognitive psychologists.

ranking search results, these features determine how close each document matches the search term – a technique called similarity measurement (Jurafsky & Martin, 2000). Some features are considered more important than others and are weighted as such, but Mr. Page points out that in an attempt to prevent manipulation of results, the systems were designed to ensure that no particular factor gets too much weight. To further complicate matters, search engines take contextual factors into account that reside outside the page and web site. For example, the labels of links that point to your pages are used to compute your search rankings. (Page & Brin, 1998).

When Google's search engine performs a search it uses the concept of authority (PageRank algorithm) combined with information about the content to determine if a page is relevant to a search term (Bianchini, Gori, & Scarselli, 2005) and if so, how relevant (Brinkmeier, 2006). This technique is referred to as Rank Merging, which is described by Larry Page and company as "a very difficult problem" (Page, Brin, Motwani, & Winograd, 1999). Now when you've got a bunch of folks with Ph.D.'s in computer science working at a Stanford university describing something as "difficult", you can probably believe them.

That being said, our goal here isn't to try to create our own search engine. It is to understand the search engines. So what can we do with this knowledge?

APPLYING WHAT WE'VE LEARNED

A little working knowledge of search engines can go a long way. This information can help you appropriately structure your content for search engines, wade through all of those offers from consultants promising to improve your search rankings, understand why you see what you see when you perform a search, and just plain be more informed.

USING THE FACTS

In order to appropriately focus effort on improving your site's search rankings the main effort needs to be on writing good content. Recall that the ultimate goal of the search engine is to approximate how a human being would rank the pages (Page, 2001). So first and foremost you should simply concentrate on the quality of your content. Also remember that there are a number of aesthetic and structural elements that are taken into account when ranking your content. These aren't just important concepts for SEO; they are just good writing techniques.

Know your target. Before you can even think about tuning your content for search engines, you really need to know what your targets are. In other words, what search terms/words do you want your content associated with. You want to make sure that these are the words you are keeping in mind when you are working through the rest of your content.

THE BIG ONES

The best way to attain good search rankings is to have good, well written content (Bianchini, Gori, & Scarselli, 2005). That being said, there are certain clues we can pull from the literature that can point us to certain elements of your page that will be weighted more heavily by the search engines. Recall that the following elements have all been specifically referenced in literature as factors: font size, weight, visibility, title, and capitalization. Just based upon these factors we can be reasonable certain that the headings (or headlines) and title of a page are all considered important elements of the page. So if you are going to focus attention on doing anything, focus efforts on writing good headings/titles to your page that include important terms. Figure 2 shows an example of a page with a headline and a sub-headline with words that are relevant to the content. Without reading the rest of the page the reader can determine what this page is about. Given that the aim of the search engine is to emulate how humans make such distinctions, it follows that these would be important elements of a page in the eyes of a search engine.

Grafton may move its Village Hall

Officials are looking at Brooks Stevens Design building as possible new location for offices

By LAWRENCE SUSSMAN
lsussman@journalsentinel.com

Posted: May 10, 2008

Grafton - Village officials are considering whether to move Village Hall about a half-mile north to the Brooks Stevens Design building, 860 Badger Circle, in the Grafton Business Park.

Advertisement

The Village Board on Monday night approved a resolution authorizing village staff to begin negotiations to buy the building, which is for sale for \$1.25 million.

If the negotiations are successful, the resolution authorizes the

FIGURE 2: THE HEADLINE AND SUB-HEADLINE INCLUDE IMPORTANT INFORMATION ABOUT THE PAGE

The title that appears in the web browser is also an important feature of a page that you should be paying attention to. Aside from the fact that the page title is specifically referenced in literature, it is the one piece of information that is visible outside the context of your page. Figure 3 shows the page titles of three web pages with similar subject matter as they are displayed on the task bar on a Windows computer. These same titles are used in bookmarks and also appear in the title bar of your web browser. Making these titles relevant to the content only serves to tell your users, and the search engines, what your article is about.

Take Away
If you are going to focus on anything, focus those elements of your page that jump out to the reader. Headlines, titles, lead paragraphs, and links.

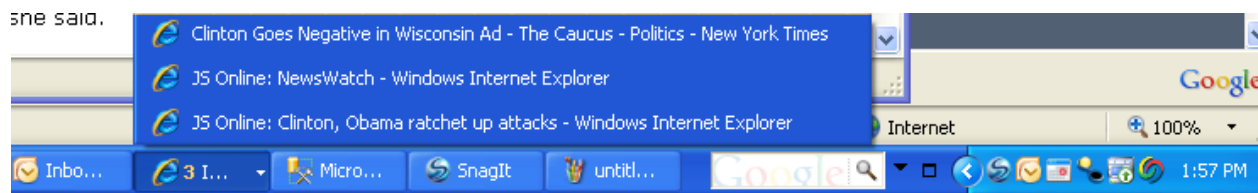


FIGURE 3: WHICH OF THESE PAGE TITLES CLEARLY INDICATE NEGATIVE AD CAMPAIGNS BY CLINTON IN WISCONSIN?

A good rule of thumb is to write content such that a human user scanning a page can quickly determine what the page is about. There is one exception to the “write for the human reader: rule of thumb: despite the best attempts to approximate human reasoning, concepts such as sarcasm and word play are completely lost on a computer. The clever title and summary might be good for a print publication, but such forms of creativity on your web page can hurt your rankings.

LESS OBVIOUS FACTORS

We can glean other factors affecting rankings from literature as well. These are less obvious but still factors in search rankings.

- **Links** – if the situation warrants it, don’t be afraid to link to other sites on the web that your reader may find helpful. The quality and relevance of the pages to which you link are considered in your rankings.
- **Link names** – naming your link “click here” really doesn’t say much about the link. A link does tend to stand out in the content of a page, so naming it with appropriately descriptive text is important. Oddly enough this also affects the rankings of the site you link to.
- **Age of site** – sites that are well established tend to have more authority than those that are new. Even if the content is better in a new site, older sites can carry more weight. Respect for elders I guess.
- **Freshness of page** – a page that is more frequently updated tends to get better rankings. This is less important when looking at a news story but when looking at landing pages for sites, site sections, blogs, forums, etc, having fresh content relevant to the theme or topic of the site/section is important.

THINGS THAT ARE DIFFICULT TO CONTROL

There are other factors that can be challenging to control. This shouldn't be a surprise. Remember that search engines have been designed and are constantly being updated to thwart attempts at manipulating results. However, through good marketing, PR, and community involvement, you can affect these factors.

- **Number of incoming links** – the number of people that link to your page has a direct affect on the ranking of your site.
- **Text of incoming links** – just like you should name your own links things like “click here”, folks linking to your site shouldn't name the links “click here” either. Inappropriately named incoming links won't hurt your rankings, but appropriately named links can certainly help.
- **Quality of referring sites** – it isn't just the number, but the quality of the sites that are linking to you. Well ranked sites and pages have more weight than link farms, aggregators, and personal pages.
- **Authority of site owner** – some organizations have more authority than others. Newspapers, major corporations, media companies, etc. tend to carry more weight.

THE MYTHS

Because of the closely guarded nature of search ranking algorithms there are a lot of myths about what will help propel your rankings in search results.

Meta data – Meta data is one element that I see most often as a supposed factor in search rankings. Meta data, particularly keywords and descriptions, are elements of a page that are used to describe the content on the page but are not directly displayed to the user. There have been studies that focused on the statistical correlation among pages that use meta data and the resultant rankings of the pages in search results and have shown no correlation (Noll & Meinel, 2007).

Numbers in the URLs – at one time in history search engines missed things like ID numbers in pages when building their databases and as such, large numbers of pages were simply dropped. This isn't the case any more. While using URLs and page names with relevant text may help your search rankings, not having them will not hurt.

Visitor statistics – I've had a number of folks tell me that the number of people that visit your site affects the rankings. This is data that search engine firms simply don't have.

OF SILVER BULLETS, CONSULTANTS, AND MAGIC PRODUCTS

Given all these precautions against result manipulation, one might ask what a search engine optimization (SEO) consultant brings to the table. In many cases ... not much. I've been known to describe many SEO consultants as akin to 13th century apothecaries attempting to sell snake oil to cure malaria. There are a lot of unqualified SEO consultants and separating the good from the bad isn't always easy. I receive a constant flood of offers from consultants claiming to be able to improve search engine rankings. Some of these offers are clearly misleading (offering to "guarantee top rankings") and are relatively simple to weed out. What a good SEO consultant will tell you is that there is no silver bullet. The key to SEO is to promote sites and pages by working on well crafted content and making content easy to find (Bianchini, Gori, & Scarselli, 2005) and that it takes time. A good SEO consultant sticks with you for the better part of the year (if not longer) and is involved in your content management activities. They come armed with tools and utilities – not to improve rankings but to perform advanced searches and analytics to monitor the effectiveness of the content management teams. A good SEO consultant is, and needs to be, in it for the long haul.

If you are considering hiring a consultant to improve your rankings, consider the following:

- Do they have the academic credentials to back their claims of expertise? Attending a "search engine boot camp" or non-accredited certifications is of little value.
- Do they have empirical evidence of improvement based upon their recommendations?
- Can they back their recommendations with literature? That is, if they make a recommendation is it backed by research or is it backed by references to web sites and blogs from others that claim to be experts as well?
- How long is the engagement? Someone coming in to fix your site is usually there for a quick buck, someone coming in to work side by side with your content editors is there to educate and not offer silver bullets.

More and more frequently, I find myself reviewing offers and sales pitches in which *search engine optimization* is considered an added benefit or feature of some product or service. This kind of sales pitch is preying on the mystery surrounding search engines to sell a product. While not necessarily done maliciously, it is misleading. If you are seriously considering a product that claims to improve search rankings, make sure you are asking the following questions:

- Will the product genuinely improve the quality of the content?
- Will it make the writing better?
- Will it make your headings and titles better?
- Will it make your content easier to find?
- Can the vendor explain, in great detail, exactly how the product improves rankings?

Most products cannot stand up to the claims that they can improve your rankings, so it is important to challenge any claims that a product is good for SEO.

CONCLUSIONS

If you understand the concepts behind search engines you can gain an appreciation for what types of content impacts your search rankings. While no one outside the walls of the search engines' research facilities can be positive of exactly what factors influence rankings and by how much, published literature can provide some valuable insight. All indications point to search engines working to emulate the human concept of relevance to determine how well a page relates to a search term or phrase. Search engine algorithms use features that are important to a human reader such as titles, headlines, and fonts into account while using measures outside the control of a site owner to determine the credibility of a page. Factors that can easily be manipulated by the content owners that do not directly contribute to the quality of content are ignored.

For these reasons it is best to focus efforts on writing good, high quality content. Avoid taking measures that are designed to trick or manipulate search engines. This will have no effect.

REFERENCES

- Baeza-Yates, R., & Ribeiro-Neta, B. (1999). *Modern Information Retrieval*. New York: ACM Press.
- Bianchini, M., Gori, M., & Scarselli, F. (2005). Inside PageRank. *ACM Transactions on Internet Technology*, 5 (1), 92-128.
- Brinkmeier, M. (2006). PageRank Revisited. *ACM Transactions on Internet Technology*, 6 (3), 282-301.
- Cilibrasi, R., & Vatanyi, P. (2007). The Google Similarity Distance. *IEEE Transactions on Knowledge and Data Engineering*, 19 (3), 370-383.
- Conigliaro, J. (2003). *Author Identification Using Naïve Bayes Classification*. Milwaukee: Marquette University.
- Conigliaro, J., Haifeng, L., & Kim, J. E. (2004). *Mining the United States Patent and Trademark Office Database*. Milwaukee: Marquette University.
- Jansen, B., & Eastman, C. (2003). The Effects of Search Engines and Query Operators on Top Ranked Results. *Proceedings of the International Conference on Information Technology: Computers and Communications* (pp. 135-140). IEEE Computer Society.
- Jurafsky, D., & Martin, J. (2000). *Speech and Language Processing*. Upper Saddle River, New Jersey: Prentice Hall.
- Mehlitz, M., Kunegis, J., & Albayrak, S. (2007). Using Novel IR Measures to Learn Optimal Cluster Structures for Web Information Retrieval. *IEEE/WIC/ACM International Conference on Web Intelligence* (pp. 312-316). IEEE.
- Noll, M., & Meinel, C. (2007). Authors vs. Readers - A Comparative Study of Document Metadata and Content in the WWW. *Proceedings of the 2007 ACM symposium on Document engineering* (pp. 177-186). Manitoba, Canada : Association of Computing Machinery.
- Page, L. (2001). *Patent No. 6,285,999*. United States.
- Page, L., & Brin, S. (1998). The Anatomy of a Large-Scale Hypertextual Web Search Engine. *Computer Networks and ISDN Systems*, 30 (1-7), 107-117.
- Page, L., Brin, S., Motwani, R., & Winograd, T. (1999). *The PageRank Citation Ranking: Bringing Order to the Web*. Retrieved 2008, from Stanford University Web Site: <http://dbpubs.stanford.edu:8090/pub/1999-66>
- Sullivan, D. (2002). *Search Engine Features for Webmasters*. Retrieved 2008, from Search Engine World: <http://searchenginewatch.com/showPage.html?page=2167891>
- Wylie, C. R., & Louis, B. (1995). *Advanced Engineering Mathematics*. New York: McGraw Hill.
- Yaltaghian, B., & Chignell, M. (2002). Re-ranking Search Results using Network Analysis: A Case Study with Google. *Proceedings of the 2002 Conference of the Center for Advanced Studies on Collaborative Research*, (pp. 14-24). Toronto, Ontario, Canada.