Setting up a Website

Getting found by Search Engines

Conor Gilmer BSc
Dublin
Ireland

December 20, 2009

Contents

1 Search Engines

Since most people use search engines as their first point of contact with the internet and indeed first point of call to find some service it is imperative that your website get found by the search engine. Search Engine Optimization is a "dark art", the rules are always changing and getting more sophisticated since websites get tailored to cheat the system and get higher rankings than perhaps they deserve on their actual content.

Below is a kind of action plan of what should be done to get your website onto the search engines and found, and in this document these will be discussed.

- Meta Tag your website
- Robots.txt file
- Register with Google
- Register with Yahoo
- Register with MSN
- Register with DMOZ - Open Directory
- Sitemaps.xml file
- Register with google sitemaps(beta)
- Link Building
- Content Tailored for Internet Search Engine Optimization

2 Meta Tags

While the relevance of Meta tags have diminished over time, since people were using them purely and simply to get ranked and not reflecting the actual webpages to which they were on. Google hadly uses them now for ranking, however some other search engines still use them.

It is still important to have relevant words. Particularly keywords which match content inside the page and linked content as well as header text title text.

```
<META HTTP-EQUIV="TITLE" CONTENT="My Companys Name">
<!-- Limit to 10 keywords, do not repeat words, no spaces after comma -->
<META NAME="KEYWORDS" CONTENT="Companyname, typeofbusiness,
businessrelevantword1, businessrelevantword2, businessrelevantword3....">
<META NAME="DESCRIPTION" CONTENT="My Company Name">
<!-- Do not edit below this line -->
<META NAME="ROBOTS" CONTENT="index, follow">
<META NAME="AUTHOR" CONTENT="Myself">
....
```

2.1 Keywords

The keywords you choose for Metatags should be relevant to your website. ou should limit it to the 10 most salient keywords, it is a good idea to have certain key keywords repeated in the Headings, Title, Keywords, Links, Filenames and in the Text on your webpages. By reinforcement it the "keywords" relevence to your website therefore gets increase

- Meta Tag
- Title Text
- Headings
- Italics/Bold
- ALT Text behind pictures
- File names of pictures
- Multimedia Content?

3   Robots.txt

To have your website crawled better by the search engine robots (googlebots, etc) you need a robots.txt file in your flat directory i.e. Http://www.mycompanyname.ie/robots.txt this allows tells the search engine robot to check sub directories and pages and not just the home/index page.

```
# /robots.txt

# no exclusion at all
Agent: *
Disallow:
```

4   Register with Search Engines and Directories

By far the most popular way people use the Internet to find information is via Search engines so getting on their databases is a priority. Another vehicle for searching the internet is browsing directories, where information is classified and organised into directories of websites of a similar topic.

4.1   Google

Google is the market leader of search engines, it has even won the battle to be the verb for searching on the internet. When your website is metatagged and tailored for search engines you should register your website with google. It is free to register with them and with in at the most six weeks your website should have been trawled by a *googlebot*.

> http://www.google.com/addurl/?continue=/addurl

4.2   Yahoo

Yahoo is narrowly behind google in search engine charts. Yahoos software is also used by a pletora of smaller search engines so it is very imortant to be registered with Yahoo and get your webname into their database.

> http://search.yahoo.com/info/submit.html

Select the submit for free, this takes some time to get you registered and on their database, as with many yahoo products they ask that you set up and use a yahoo account before you do this

4.2.1   Yahoo Directories

While Yahoo Directory costs to get immediately list
https://ecom.yahoo.com/dir/submit/intro/
you can get into these directories since they do use the Open Directory Project.

4.3   MSN

While MSN is the new kid on the block of search engines it only takes a few moments to register with them.

4.4 DMOZ

DMOZ is an open source directory service, you can log on an select the category you wish to classify your site in and they should. DMOZ is used by a number of directory services, like Google Directory (http://www.google.com/dirhp)

4.5 Google, Yahoo and MSN/Live Images

If your site has been trawled by search engine robots and there is Images on the website they will be on the Search Engine Images listings.

4.6 Sitemaps

Another way to get quicker and better evaluation by search engines it to have a sitemap file (either xml/html) on your website even if it is not generally navigable from the site this is a quick map of the site to the search engine to trawl through your website. As you see in the Sitemap.xml file in Appendix 1, the sitemap just lists the pages you have.

There are a number of scripts in php and perl which would generate a sitemap.html or xml file. There are also a number of online sitemap generators which create a sitemap file for you : -
e.g. http://www.xml-sitemaps.com/

Every time a new page is added, for that page to be be rated/ranked quickly the sitemap file should be regenerated!

4.6.1 Google Sitemaps

When this is created register this file with search enging sitemaps
(http://www.google.com/webmasters/sitemaps/login there was a beta yahoo version)

5 Link Building

By association it is good to be linked to from other websites which have high search engine rankings. It is desireable to be linked to from the websites of your previous clients, this is a search engine benefit, but it is also beneficial when viewed by a prospective client.

Salient links are always beneficial to getting a site ranked (try and avoid link farms but directories are ususally), so if your business or service is in a particular area, the local area online directory should be browse and email asking them to include you in their list, some would have online link submittal forms where you would complete a little paragraph(Your sales pitch).

6   Content

Content can be tailored to be "searchengine clever", by using salient words in page titles, Headings <H1>, text in actual links, <ALT> Text for images when mouse hovers over the image (displayed in Internet Explorer)

Alternative text for Images - ALT

```
<img src=http://www.mywebname.ie/images/mylogo_sm.gif alt="My Company Logo"
 width=143 height=59 hspace=5 vspace=5 border=0></a>
```

Headings

```
<h1>The Greatest Company in Rahney</h1>
```

7 Social Networking, Blogging and Web 2.0

The explosion on to the scene in recent years of social networking and web 2.0 channels of communication provide different ways of attracting people to view your website or delivering your online material to them.

- Yahoo Groups/360

- Google Groups

- RSS News Feeds/ATOM

- MySpace

- Bebo

- Facebook

- LinkedIn

- Hi5

- YouTube (video sharing owned by google now)

- Vimeo (video Sharing)

- MetaCafe (Video Sharing)

- Orkut (google)

- nimble (Irish Based)

- Message Boards e.g. boards.ie

- Blogs e.g. Live Journal, blogs.ie, blogspot, eblogger etc.

- Micro Blogs - twitter

The benefits of many of these networks and web facilities, is that in many cases such as groups and social networks it allows for potential clients to communicate with you, it also allows them to "opt in" to recieve information from you, and in the cases of the groups, message boards, blogs and social networking allows them to provide potential content for you.

The potential content that users can generate is a massive double edged sword since you are opening up to the anarchy of the Internet and while positive information may be produced so to can negative content, and this must be addressed.

References

[1] A. Weckler, *Social Networking Tools*, Computers in Business : Sunday Business Post (September 2007), 8-11 available at http://www.thepost.ie.

A Appendix 1

A.1 sitemap.xml

Here is the Format of a Sitemap XML file for google-sitemaps. There are a number of ways to create these files, there are a number of scripts available, there are also a number of websites which will generate a sitemap.xml for your website.

```xml
<?xml version="1.0" encoding="UTF-8"?>
<urlset
  xmlns="http://www.google.com/schemas/sitemap/0.84"
  xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
  xsi:schemaLocation="http://www.google.com/schemas/sitemap/0.84
                      http://www.google.com/schemas/sitemap/0.84/sitemap.xsd">

<url>
  <loc>http://www.mycompanyname.ie/</loc>
  <priority>0.5</priority>
  <changefreq>daily</changefreq>
</url>
<url>
  <loc>http://www.mycompanyname.ie/index.html</loc>
  <priority>0.5</priority>
  <changefreq>daily</changefreq>
</url>
<url>
  <loc>http://www.mycompanyname.ie/about.html</loc>
  <priority>0.5</priority>
  <changefreq>daily</changefreq>
</url>
<url>
  <loc>http://www.mycompanyname.ie/services.html</loc>
  <priority>0.5</priority>
  <changefreq>daily</changefreq>
</url>
<url>
  <loc>http://www.mycompanyname.ie/clients.html</loc>
  <priority>0.5</priority>
  <changefreq>daily</changefreq>
</url>
<url>
  <loc>http://www.mycompanyname.ie/testimonials.html</loc>
  <priority>0.5</priority>
```

```
  <changefreq>daily</changefreq>
</url>
<url>
  <loc>http://www.mycompanyname.ie/contact.html</loc>
  <priority>0.5</priority>
  <changefreq>daily</changefreq>
</url>
</urlset>
```