



Web marketing concepts (CMW)

Class 7

Search engine optimization (SEO)

Having a great user interface and an efficient marketing approach would be of no use if nobody gets to see your website. In order for people to visit your website, they need to find you. Of course, your URL appears on your business card, in your Linked-In account and maybe on your car, but that isn't enough. People need to see you in their search results. In fact, they need to find your website in the very first results of the search results or it basically doesn't exist.

What is SEO?

Search engine optimization (SEO) consists in different operations which purpose is to increase the quantity and quality of traffic to a website through organic (non paid) search engine results.

The quality of traffic

The number of visitors to a website is not the most important thing in most cases. Visitors to a website must really be interested in the website they come to. Otherwise, they leave at once.

The quantity of traffic

Once you have quality traffic from search engine result pages (SERP), increasing the traffic should be your goal.

How SEO works

Basically, a search engine is a web page where users enter keywords to generate a list of links to different website that can potentially answer their needs.

In order to generate a list of results, search engines need some sort of database containing references about all existing websites on the Internet. To gather all this information, search engines use crawlers.

A crawler is a program that retrieves the informations of all websites to create a gigantic index. Based on the user-supplied keywords, an algorithm matches all the data with th the query and supplies a results list.

Many factors are taken into consideration by a search engine's algorithm and algorithms are regularly changes in part to avoid unfair and tricky optimization strategies.

What is optimization

Optimization is content-related work that is done to help search engines understand the content while managing to rank as high as possible in the search results lists.

Optimization goes from correctly structuring web documents with proper tags to using meta descriptions, strategically produce text contents, link to pertinent internal or internal pages, etc.



SEO best practice

Importance of SEO

Although traffic can come from different sources such as social medias, the biggest part of traffic on the Internet is generated by major search engines (and mainly Google). They bring the right people to the right websites. So if a site isn't indexed correctly (or not at all), this can lead to dramatic endings. Traffic generated by SEO can be incredibly profitable compared to other types of marketing and promotion.

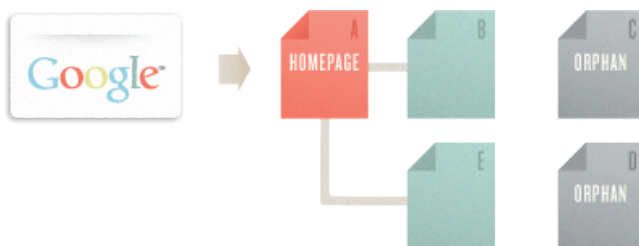
Unfortunately, and although algorithms are getting better and better, search engines cannot extract content from websites with perfect efficiency. They need help. That is where SEO enters the scene. It will help the website being correctly indexed, assuring visibility and boosting ranking so users can find a web site faster and more easily.

SEO is a very complex task that requires continuous analysis of the algorithms being used. Although, minimal knowledge can help you make a great difference. For optimal results, some companies offer extensive SEO services.

How does search engines work?

Link structure indexing

As we have earlier said, search engine use crawlers (spiders) to explore the Internet documents. For each web page, it uses the link structure of websites to retrieve all related documents and save some of the information in an index.



The two operations of search engines

When a search query is made by a users, the search engine first retrieve only relevant results, and then ordered these results based on their relevance (ranking). Relevance is established by the algorithm based on hundreds of *ranking factors*.

Basic SEO guidelines

- Produce the web pages for users and not for search engines. Don't use *cloaking* (presenting different content to users and search engines).
- Make sure your web structure is clear and that all pages and documents are reachable from at least one text link.
- Use semantic markup using proper and relevant tags.
- Create pages with rich contents accurately described.
- Make sure contents and links are not hidden into images or rich medias such as Flash, video, etc.
- Make sure your pages don't provide similar contents.
- Produce new content regularly so the web site is updated frequently.
- In your contents, use terms and expression commonly used by average people.
- Use user-friendly keywords in links.
- Use accurate location targeting.
- The <title> tags and «alt» attributes should be descriptive and accurate.
- Especially when using *301 redirects*, use *rel="canonical"* attribute.
- Make sure there are no errors in the *robot.txt* document.
- Have your content talked about, reviewed, commented, linked to, etc. (marketing).

The keywords

Any content research process is based on keywords. Search engines create databases centered on keywords making it possible to pull out all documents containing one or several terms. You will then understand that a website about books should have the word «book» in its content.

Using keywords

When keywords are typed in for a search, the search engines look into its databases for the words, of course. But it also takes in consideration many other things such as words order, spelling, punctuation, capitalization, in order to give the most possible relevant ranked results.

The best and most basic way to optimize page ranking is to make sure the desired keywords you focus on are used in titles, text, and metadata.

Choosing keywords

As we already have said, the first step is to determine the keywords to be used. It is important to make sure the chosen keywords are the one most of people would use to search for a page like yours.

Since everyone is using more or less the same SEO approach, there is a competition with keywords. This means more general terms will generate much more results than more specific terms.

For instance, just imagine how many possible results you could get with a search using the words «book» and «music». Now, imagine how much narrower would be the results with a search that would use «book about music». There would certainly have less results, but also less competition.

Keywords abuse

In the early age of Internet, it was common for web developers to stuff their pages with a humongous number of keywords to show up in maximum of queries results. But algorithms are smarter now and this would do much more harm than good.

Use your keywords naturally and strategically, and make sure they are relevant with the content of your pages. You don't want a maximum of visitors, you want a maximum of visitors interested in your contents.

Keywords page optimization's best practice

- Keywords should be present in <title> and heading at least once, and 2 or 3 times in the body copy.
- Try to keep keyword phrases as high in the page as possible, as close as possible to the <title> tag.
- Keywords should be used at least once in the alt attribute of an image and once in the meta tags.
- Keywords should be used at least once in URL. Although, avoid using keywords in links to other pages of your website. This is known as *keyword cannibalization*.

<title> tag

A web page's title should be a short and effective description of the content. Here are a few basic recommendations that should be observed :

Make it short

Search engines and social medias results show a maximum of 65 to 75 characters of a title tag in the results. Although not showed, the remaining is still taken into considerations.

The higher the better

Place important keywords as high in the page as possible, as close as possible to the title tag as that is how they will be the more helpful for ranking.

Consider emotional impact

Of course, page titles should be descriptive. But don't forget they are part of the user experience, and an efficient title will grab attention and generate more traffic than a plain description.

The robots.txt file

Located in the root of the web site (at the same level than the main *index.html* file), when present, a robots.txt file is the first one being read by search engines. By default, crawler will index all documents in the link structure from the *index.html* file. The *robots.txt* file is used to tell the crawlers what should and should not be indexed.

The robots.txt structure

The file structure is pretty simple. It consist of indicating the user-agent (the search engine crawler which is targeted) and instructions.

It is important to make sure there are no mistake or the web-site's pages are at risks of not being indexed.

Also, this file is easily readable, which means anyone can know what pages you desire to exclude from indexing, so don't use the robots.txt file to hide content.

Excluding all files for all search engines

```
User-agent: *
Disallow: /
```

In the preceding example, **User-agent: *** means «all crawlers», and **Disallow: /** tells the crawlers it should not visit any of the pages.

Allowing all files to be visited by search engines

You could very well also create an empty robots.txt file or simply not use any.

```
User-agent: *
Disallow:
```

Excluding some files for all search engines

```
User-agent: *
Disallow: /cgi-bin/
Disallow: /temp/
Disallow: /icons/
Disallow: /gallery.html
```

In the preceding example, instructions are given to all crawlers not to visit files located in the folders named «*cgi-bin*», «*temp*» and «*icons*».

Excluding a single search engine

```
User-agent: BadBot
Disallow: /
```

Allowing a single search engine

```
User-agent: Google
Disallow:
```

```
User-agent: *
Disallow: /
```

Common user-agent names

Search engine	Name
Google	Googlebot Googlebot-News Googlebot-Image/1.0 Googlebot-Video/1.0
Bing	Bingbot
Yahoo	Slurp
DuckDuckBot	DuckDuckBot
Baiduspider	Baiduspider
Yandex	YandexBot
Exalead	Exabot
Facebook External Hit	Facebot
Alexa	ia_archiver
Bad Bots*	BadBot

* Bad Bots are crawlers put on a black list

```
User-agent: *

# Sitemaps
Sitemap: http://yourdomain.com/sitemap.xml

# Do not index directories
Disallow: /includes/
Disallow: /wp-admin/
Disallow: /wp-content/
Disallow: /wp-includes/
```

Example of a robots.txt file.

The meta robots tag

The Meta Robots tag controls search engines crawlers behaviours just like the *robots.txt* file does, but it does it on a given page instead of from the root of a website.

Why use the meta robots tag?

Of course, it is a lot more simple to use the *robotx.txt* file, although, there are some situations when it is preferable to use the meta robots tag instead:

- If you have no control on the robots.txt file (e.g. some blog platform).
- To keep search engines to index *back-of-fice* access to contents (The *robots.txt* files can do it, but it is too easy to hack).
- If a page is restricted by a robots.txt file, it can still be indexed, for instance, when there are *backlinks* pointing to it.

Structure of the meta robots tag

```
<meta name="robots" content="index" />  
<meta name="robots" content="noindex,nofollow" />
```

As any other meta tags, it must be placed in the `<head>` section of the HTML document. The value of the content attribute tells the search engines the way a page should be treated.

Possible values of the content attribute

index/noindex :

Tells the engines whether or not to index the page. By default, search engines are indexing all pages.

follow/nofollow :

Tells the engines whether or not to follow the links of the page. By default, all pages are followed.

none :

Equivalent of *noindex* and *nofollow*.

noarchive :

Tells the engines not to save a cached copy of the page. By default, the engines save all pages they have indexed, accessible through the cached link of the search results.

nocache :

Same as *noarchive*, but for *Internet Explorer* and *Firefox*.

nosnippet :

Tells the engines that it shouldn't display a descriptive block of text next to the page's title and URL in the results. **X-Robots-Tag** also does the same thing and works especially well for content within non-HTML files.

noodp :

In the case of automatically generated Meta Descriptions, *noodp* prevents search engines from using descriptions from the *Open Directory Project*, if it lists the website.

noimageindex :

Tells crawlers not to index any images on a page.

notranslate :

Tells crawlers to not show translations of the page in the search results.

unavailable_after :

Allows crawlers to show a page in their results until a certain date and time (*noindex* is applied after the given date. Must be in RFC850 format).

Using robots.txt file and meta robots tags

Of course, it is possible to use the robots.txt file along with meta robots tags. That means the robots.txt file may allow a crawler to access all files from a given folder, but a meta robots tag may disallow indexation for a given file.

Although, it is important to know that if there are two different instructions for a page, the crawler will use the most restrictive one.

Pages that shouldn't be indexed

Any page which content isn't ready yet, pages with very little content or pages that are too similar to others should use *noindex*, although, you may still want to use *follow*.

The meta tags

Although meta tags do not constitute a magical solution that will boost a website's ranking, it may help a little. One thing is sure, it helps users knowing what the web pages are about, which is quite important. Although, bad practice can have very negative impacts, so it is important to avoid mistakes.

What Are meta Tags?

HTML meta tags are optional data tags placed into the <head> section of HTML pages in order to give browsers and various web services specific information about a page.

Charset meta tag

The charset meta tag tells the browsers the character encoding used in the HTML document. UTF-8 is the default character encoding in HTML5. Using this avoids having to manually code special characters such as accentuated letters.

Although, for various reasons, it is now recommended to avoid using this meta tag and to prefer server side technology.

HTML4

```
<meta http-equiv="Content-Type" content="text/html; charset=UTF-8">
```

HTML5

```
<meta charset="UTF-8">
```

Description meta Tag

```
<meta name="description" content="Your description" />
```

The description meta tag shortly describes what is the page or the website's about. If the description isn't well written or does not reflect the page's content, the search engine replaces it with its own. The same thing would happen if you create an empty description.

The description is what shows up in the search engines results. On top of informing the users on the nature of the page, it should also build interest in order to get more people visiting it. And more clicks a page gets positively influence ranking.

Be careful never to use full quotation marks (") in your description, this would cut it off. Use single quotes instead.

Quite useless meta Tags

Keywords Meta Tag :

```
<meta name="keywords" content="You,keywords" />
```

This meta tag was once very important. Nowadays, only Bing is said to use it in order to detect spam. According to many specialists, it is better not to use this tag anymore. If you decide to use it anyway, use only relevant keywords and avoid using too many of them.

Although, it is a good idea to use this meta tag to include common misspelling of your keywords as well as synonyms.

Author Meta Tag :

```
<meta name="author" content="Your Name" />
```

Copyright Meta Tag :

```
<meta name="copyright" content="Some company" />
```

Tells the crawlers when to revisit :

```
<meta name="revisit-after" content="30 days">
```

Never to be used meta Tags

Automatically refreshes the page :

```
<meta http-equiv="refresh" content="30">
```

Automatically redirects to another url :

```
<meta http-equiv="refresh" content="10; url=http://www.yourdomain.com">
```

These are very annoying to users and it increase your *bounce rate*, which is usually a bad thing.

For permanent or temporary redirections, it is suggested to prefer *301* or *302 redirects*.

URL structure

URL are shown in the search engines results, and therefore they can potentially impact on traffic. They are also used by search engines for ranking. Pages URL containing words used for queries (keywords) are better ranked. Finally, being shown in the address bar, URLs can improve user experience, acting a bit like breadcrumbs.

Here are some guidelines to help your correctly structure your URLs.

Empathy

See things from the users perspective. Manage for the URL to make it possible to predict the content of the page it links to. Without getting into too much specific details, give an overall idea of the content.

Keep it short

Although it must be descriptive, keep your URLs as short and simple as possible so it is fully viewed and easy to copy and paste.

Use keywords

Use relevant folders and files names. As much as possible, use keywords related to the page's content. However, avoid stuffing URLs with keywords, it would get URLs too long and you are at risks of having spam filters blocking the link.

Use hyphens to separate words


Certain symbols (including underscores) are not accurately interpreted by all browsers. To avoid any problems, use hyphens to separate words.

Use canonical link tag

To avoid having quite similar pages being indexed, it is useful to use the canonical link in the <head> section of the page which indicates to the search engines the main page a current page refers to.

For instance, for a product available in three different colors, you wouldn't want the three pages versions to be indexed, but simply the main specific product's page.

```
<link rel="canonical" href="http://site.com/product.html" />
```

 example.com/skinny-jeans.html		use of descriptive keywords, short and sweet
 example.com/skinny_jeans.html		hyphens are generally not preferred and hard to read
 example.com/?p=123		permalinks with non-readable text is not a good idea
 example.com/jeans/skinny		categorization and nesting of pages helps organize the site
 example.com/skinny-jeanz		cute misspellings aren't so cute, and are hard to remember
 example.com/clothing/skinny/denim		deep nesting is also ok as long as it makes sense and is natural

HTML5 structural semantic elements

Just as the web sites and URLs structures are important, the web pages structure also plays a role in SEO. A well coded page will make it a lot easier for the search engines to index contents.

Search engines use standard HTML elements (tags) in trying to understand the content of a page. Optimizing the page structure directly improves the search engines understanding of a page. HTML5 adds new HTML elements to the existing ones called semantic elements that helps optimizing a page's structure.

Semantic element are those carrying a sense, a meaning. Standard HTML elements are simply structural and basically carry no meanings.

<header>

The <header> element can be used as the header of a letter or as a heading such as <h1>. It can identify the header portion of a page as well as the heading of the main, section or article contents, for instance.

<main>

The <main> tag is used to enclose the main content of a page, the dominant content. It may include one or several content of similar nature (e.g. all articles concerning a page's theme). Using such an element will help search engine focus on the important contents of a page.

<section>

The <section> element can be used to specify subsections of the main content, grouping articles in independent groups. Some people uses it the other way around to create subsections of an article

<article>

The <article> tag is used (usually within the <main> section) to indicate independent segments of information. Search engines view these portions of contents as more important and give them more weight. It also cleans up the code and makes it easier to read by reducing the use of neutral <div> elements.

<nav>

The <nav> element is used to indicate the navigation. This is a very important element of any page.



Example of a page structure using HTML5 semantic elements.

<aside>

The <aside> element is used to indicate secondary content or complementary information.

<footer>

The <footer> element also offers SEO benefits as it is known by search engines to provide various typical elements such as links to social medias, site map, press link, etc.

Output elements

Just like there is a difference between neutral and semantic structural elements, there are standard and semantic elements used for contents of various nature (text, image, audio and video). These also may greatly help search engines retrieving the important content of your web pages.

<address>

The <address> element is used to specify the contact information for the author/owner of a page.

<figure>

The <figure> element specifies visual content, such as illustrations, diagrams, photos, etc.

```
<figure>

</figure>
```

<figcaption>

The <figcaption> element is associated to a <figure> element to describe its content (just like a photograph's legend).

```
<figure>

<figcaption>President of USA</figcaption>
</figure>
```

<h1> to <h6> headings

Using standard headings elements tells more about the content hierarchy of a page than neutral <div> or . It also make these words more important than the others.

, <q>, <cite> and <blockquote>

Generating italic text like the neutral <i> element, is used to put emphasis on a word, <q> for a short quotation not requiring line break, and <cite> to indicate a quotation. Note that the standard element <blockquote> may still be used for an extended quotation.

Generating bold text like the neutral element, specifies an important term.

<picture>

Defines a container for multiple image resources using <source> to define the images to be used.

```
<picture>
<source media="(min-width: 650px)" srcset="red.jpg">
<source media="(min-width: 465px)" srcset="pink.jpg">

</picture>
```

<video> and <audio>

The <video> element is a very useful one as it allows for cross-browser compatibility to display videos without having to use third party plug-ins such as flash. <source> specify the audio or video file to be used.

```
<audio controls>
<source src="scream.ogg" type="audio/ogg">
<source src="scream.mp3" type="audio/mpeg">
Your browser does not support the audio tag.
</audio>
```

```
<video width="320" height="240" controls>
<source src="movie.mp4" type="video/mp4">
<source src="movie.ogg" type="video/ogg">
Your browser does not support the video tag.
</video>
```

See online documentation for all optional attributes such as *autoplay*, *loop*, *poster*, *preload*, etc.

NOTE:

HTML5 makes it possible to add information about the video such as closed captions and subtitles.

What affects ranking

As you now know, the structure of the web site, the page and the links structure, the content markup, the keywords and the meta tags plays a role in search engines ranking of a web page. But that's not all. Although there aren't magical solution and that a good SEO can sometimes take years to build the wanted results, many other aspects can improve ranking.

Usability and user experience

Those two aspects give indirect benefit which search engines can measure.

A well designed web site providing quality contents will make people read content, follow the internal links, bookmark the page, share the link and revisit the website. All of these are positive signals contributing to a higher ranking.

The traffic

When a user clicks a link in a search engine results list to a web site, two things can happen: the user starts reading the content or comes back to the search engine results list to click another link.

These two possibilities generate a short or a long click. A short click is called a «bounce»; people see the page, consider this isn't what they were looking for and leave. What we are looking for are long clicks, long visits.

A high **bounce rate** indicates to the search engines that users are not satisfied with the web site and will then give it a bad mark which will affect the ranking negatively. So, it makes sense mobile redirects, redirecting to another domain and 404 errors are to be avoided.

The number of clicks from the search engine results list to a page as well as within a web site also has a positive impact on ranking.

Quality of content

In 2011, Google introduced the Panda algorithm which task was to evaluate the quality of content of web pages. It is now part of the core of the search engines algorithm, but the many factors used to identify low quality content websites are still applied.

Here are some of the factors negatively impacting ranking:

- **Keyword stuffing**: using irrelevant keywords or hidden texts within content in order to boost search ranking.
- **Short content**: Providing very little or useless content on a page.
- **Irrelevant content**: Content not matching the keywords, titles and other elements used in SEO.
- **Broken links**: Non-functional links or links not taking users where it is supposed to.
- **On-site and off-site duplicated content**: Duplicated or similar content across the pages of a given website or compared to other websites.
- **Deceptive content**: Content purposely designed to deceive users (e.g. ads that look like regular content to generate clicks).
- **Content farms**: Websites producing great volumes of content solely for SEO.
- **Broad topics**: Websites providing content not related to a focused topic or about too many different topics sometimes not relevant to each other
- **Machine-generated content**: Low-quality content produced by softwares.
- **Poor spelling and grammar**: Errors making content difficult to read or to understand.
- **Non-optimized pages**: Pages that are not optimized for search.

Responsive design

Note that web pages well adapted to mobile devices benefit from a higher ranking. Since over 55% of visits are made using mobile devices, it wouldn't be very smart to ignore this.

Popularity

One very important aspect taken into consideration by search engines is the level of popularity of a page. If a lot of people visits a page for long periods and are revisiting regularly, it sends a positive signal to the engine. Although it is difficult to make a website popular, on top of what we have seen, there are many things that can be done to make web pages more popular or, at least, make them seem more popular.

Back linking

If there are many links pointing to a specific website, this sends a positive signal to the search engines telling it the site must be important and popular. Having external links pointing to a website is then quite important.

Back linking is even more powerful if the links pointing at a website are «local», treating of the same topic and/or considered a trusted or highly trusted domain (e.g. universities, governments, non-profit organizations, etc.).

On the opposite, back linking from a bad website of any kind affects the trust search engines can have in a given website.

Also, since linked pages may lose popularity over a period of time, make sure to add new links regularly (and remove those that aren't relevant anymore).

Getting linked

The first way for getting linked from other websites is to publish great content. It is also possible to compile useful informations such as different types of lists in order to become a quick reference for people.

You can also ask websites and bloggers to talk about you and share your URL and register to all sorts of lists and directories. One good idea is to register to a few popular relevant blogs and participate regularly (posting articles or commenting existing ones), making sure to add the website's URL in your signature. Be careful with these, though, as they often can be considered spammy.

Anchor text

Any anchor text in a web page is of great importance for search engines. Although, the anchored text must be relevant to the URL it is pointing to, whether it is external or internal. So, it is a good idea to link keywords in the body copy.

Social medias

It isn't a secret that social medias are more and more important. Although search engines don't treat social media linking the same way they do for regular links, these are noticed and taken into accounts.

More important, even if good ranking in search engines is the main goal here, most of the time, the greatest portion of your traffic will be generated by social medias unless you are a huge well known multinational company. It is important to maintain a presence and activities on social medias, bringing people to visit your website.

Although, traffic being mainly oriented from the social medias to a website, think of making it possible for the users to add your social medias accounts from your website so they can get kept up to date with your communications.

Most popular social networks (2018)

Social network	Monthly Active Users
Facebook	2,230,000,000
YouTube	1,900,000,000
Instagram	1,000,000,000
Qzone	563,000,000
Weibo	376,000,000
Twitter	336,000,000
Reddit	330,000,000
Pinterest	200,000,000
Ask.fm	160,000,000
Tumblr	115,000,000
Flickr	112,000,000
Google+	111,000,000
LinkedIn	106,000,000
VK	97,000,000
Odnoklassniki	71,000,000
Meetup	35,300,000

Useful tools

Robots.txt and robots meta tags

You can get all informations regarding robots.txt and robots meta tags here :

<http://www.robotstxt.org/>

Keywords help

moz.com/explorer

adwords.google.com/keywordplanner

trends.google.com

www.wordtracker.com

moz.com/blog/category/keyword-research

Google analytics

Free web analytics tool offered by Google to help you analyze your website traffic. It allows to know how many users visited a website with great details such as age, gender, interests, device, browser, country, operator, bounce rate, users behaviours, etc.

analytics.google.com

Site maps

A site map helps search engines finding and classifying contents. Although XML is the most recommended, they can be produced RSS, XML or TXT.

You can get the full information at this address :

Sitemaps.org

Or you can easily build your own XML site map here :

XML-Sitemaps.com

Moz Open Site Explorer

Provides valuable insight into your website and links :

moz.com/researchtools/ose/

Domain Authority

A query-independent measure of how likely a domain is to rank for any given query.

moz.com/learn/seo/domain-authority