

## Background of this document + what it is

I put this information together in order to have a supplement to the school program in computer science for my son.

This course is a guideline for teachers.

The SEO introduction document describes how to write web pages search engine friendly with the goal that your web pages will be indexed and have a high ranking in search engines.

## Prerequisites

- You have to know the basic concepts mentioned in the [Computer Crash Course](#)
- The [Build a Website easily](#) document explains you which steps you can take to create a Website
- You have to be familiar with the basic HTML elements mentioned in the [HTML introduction](#) document
- You have to know the basics of CSS, explained in the [CSS introduction](#) document

## Computer Crash Course (CCC)

The Computer Crash Course contains a guideline for teachers in basic computer knowledge. It can also be used as an introduction for beginners. You have to follow the mentioned links and read additional information concerning the keywords in the CCC document, otherwise you won't understand this course.

## Build a Website easily

The "Build a Website easily" document explains you which steps you can take to create a Website. This is an umbrella document for the following courses : HTML introduction, CSS introduction and SEO introduction.

## HTML introduction

The HTML introduction gives you an idea how to create HTML pages on a base of examples and a minimum of theoretical background.

SEO introduction by [Marc Oscar Schwager](#)

## **CSS introduction**

The document CSS introduction completes the HTML introduction document.

The CSS introduction course gives you an idea how to use CSS to design the look (style) of your web pages (HTML documents).

## Table of Contents

<b>What is a search engine.....</b>	<b>5</b>
<b>What means SEO.....</b>	<b>5</b>
<b>Google ranking factors.....</b>	<b>6</b>
<b>What is Google looking for.....</b>	<b>7</b>
<b>SEO basics.....</b>	<b>8</b>
<b>Content.....</b>	<b>9</b>
Introduction.....	9
Keyword research.....	9
Technical implementations of relevant keywords.....	10
URL and HTML code.....	10
SILO structure.....	10
Content hub.....	10
Secure content against scrapers.....	11
Ping.....	11
Links.....	11
<b>Spam .....</b>	<b>12</b>
Introduction.....	12
Web page spam.....	12
Domain level spam.....	13
<b>Clean URLs.....</b>	<b>14</b>
<b>Domain URL.....</b>	<b>15</b>
<b>Duplicate content.....</b>	<b>18</b>
<b>Robots.....</b>	<b>19</b>
Introduction.....	19
Robots.txt.....	20
Robots meta tag.....	21
X-Robots.....	22

<b>Technical on-page SEO.....</b>	<b>23</b>
Introduction.....	23
Link structure.....	23
Overview.....	23
Breadcrumbs.....	23
Contextual links.....	24
The canonical meta tag.....	24
The hreflang meta tag.....	25
The title tag.....	26
The meta description tag.....	26
The heading tag.....	27
The rel="nofollow" attribute.....	28
The structured data markup.....	29
Introduction.....	29
Schema markup .....	30
Microformat.....	31
Rich Snippets.....	31
Social media markup.....	32
Mobile optimizing.....	33
Introduction .....	33
The HTML viewport.....	33
The CSS viewport.....	33
The CSS @media rule.....	34
Dimensions.....	34
<b>The sitemap.....</b>	<b>35</b>
<b>SSL.....</b>	<b>35</b>
<b>Speed.....</b>	<b>36</b>
Compression.....	36
The CSS file.....	36
The images.....	37
The browser cache.....	38
Minify.....	39
HTTP headers.....	39
<b>Off-page SEO.....</b>	<b>40</b>
Link building.....	40

<b>Web marketing</b> .....	<b>41</b>
Introduction.....	41
Keys.....	41
<b>Monitoring and analyzing</b> .....	<b>42</b>
<b>Improving and developing</b> .....	<b>42</b>

## What is a search engine

A search engine is a program that crawls the Internet for web pages and builds an index of them. The search engine searches for one or more keywords in their index and then responds with a list of (relevant) links (of web pages, documents, etc) containing one or more keywords the user typed in.

**Example** of popular search engines : Google, Bing, Yahoo

### Other search "engines"

- A meta search engine sends your query simultaneously to other search engines
- A web directory organizes Websites by subject and it is usually maintained by humans
- An individual Website search engine indexes and retrieves the content of just one Website
- Specialized content search engines

## What means SEO

SEO means Search Engine Optimization. This means optimizing your Website in a way that it will be found optimally when people are searching for a subject related to the Website (high **ranking**). SEO should also help to increase the number of visitors for your Website (**traffic**). SEO helps basically to provide better organic search results.

### SEO overview

- **Technical on-page** SEO concerns all elements on a web page except the content.
- **Off-page** SEO covers all activities outside the web page (link building, social media, etc).
- Of course SEO and the web page **content** have a strong connection!

**Web marketing** (digital marketing) covers basically all the activities and strategies to convert Website visitors to clients. In Web marketing, we distinguish on-page and off-page techniques as well as paid- and non-paid strategies. SEO is a part of Web marketing.

SEO and web marketing techniques are often **interconnected**. The limits between them are not always so clear. For example : great content makes link building easier.

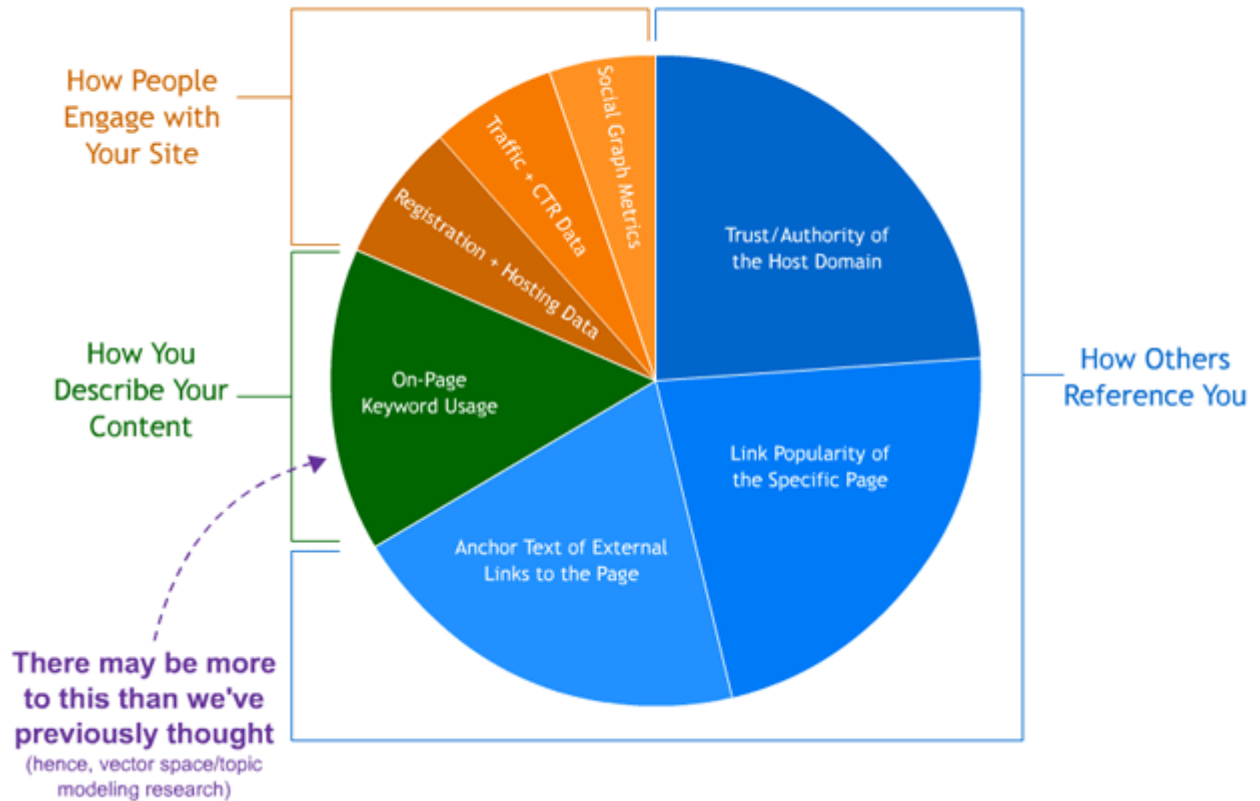
## Google ranking factors

Page-Level Link Features - PLLF (e.g. PageRank, TrustRank, quantity of link links, anchor text distribution, quality of links sources, etc)	19.15 %
Page-Level KW & Content Features - PLKWCF (e.g. TF*IDF, topic-modeling scores on content, content quantity/relevance, etc)	14.94 %
Page-Level, Keyword-Agnostic Features - PLKAF (e.g. Content length, readability, uniqueness, load speed, etc)	9.8 %
Domain-Level Brand Features - DLBF (e.g. offline usage of brand/domain name, mentions of brand.domain in news/ media/press, entry association, etc)	8.59 %
User, Usage, & Traffic/ Query Data - UUTQD (e.g. traffic/ usage signals from browsers/ toolbars/clickstream, quantity/ diversity/ CTR of queries, etc)	8.06 %
Social Metrics - SM (e.g. quantity/quality of tweeted links Facebook shares, Google +1s, etc)	7.24 %
Domain-Level Keyword Usage - DLKU (e.g. exact-match keyword domains, partial-keyword matches, etc)	6.98 %
Domain-Level, Keyword-Agnostic Features - DLKAF (e.g. domain name length, TLD extension, domain HTTP response time, etc)	5.21 %

**What is Google looking for**

### Components of Google's Ranking Algorithm

(According to 72 SEOs Surveyed for SEOMoz's Biennial Search Ranking Factors)





## SEO basics

The Website must be professionally **designed**, easy to **use** and **navigate** as well as **understandable**. Make a unique design - this will be your fingerprint on the Internet !

Design web pages basically **for users**, not for search engines. Describe what the user can learn, solve, buy, fix, treat or understand from your web page, use "call to action". Crafting a thoughtful, empathetic user experience helps ensure that visitors of your Website perceive it positively. Encourage also sharing and bookmarking. This will generate returning visits and inbound links. All this signals trickle down to the search engines and contribute to **high rankings**.

**Interlink** the web pages hierarchically, every pages should be reachable by a link.

The Website must be **accessible** for modern browsers, including mobiles, as well as search engines.

A search engine will index your web page easier if the **code** is correct. The code (HTML, CSS, JS, PHP) can easily be checked by on line tools.

## Content

### Introduction

Your content is the most important factor that affects the ranking in the search engine result pages (SERP) for your Website ! **As from the perspective of SEO, 90% of your effort should go into creating great content !** Write basically your content for your audience, then optimize it for the search engines (keyword research). The content should be **information rich** and **engaging**. Aside from the text include images, multimedia and interactive elements. Add fresh content on a **regular basis**. Longer in-depth content normally sends more relevancy signals to the search engine than a little text.

Don't put relevant content in rich media like flash, etc and images. Relevant content must be (HTML) text.

Avoid duplicate content ! Avoid keyword stuffing ! Avoid spamming !

Think also about research concerning your **target audience**, where they are and what they want !

### Keyword research

Keyword research means the process of selecting the best keywords for your Website. Keyword stands for : keyword(s), keyword phrases, long tail keywords.

You can use Google's free [keyword planner](#) to do keyword research.

**Most people never take the time to do enough keyword research before creating their Websites. So they often target keywords that are either too competitive or not searched enough.**

Do keyword research and content creation "together". It is important the selected relevant keywords meet your marketing criteria, because using specific keywords associated with your content will result in a higher ranking for your Website. Good keywords for your Website are the ones with a good amount of searches and minimal competition.

## Technical implementations of relevant keywords

### URL and HTML code

Use descriptive and accurate keywords in :

- page URLs
- `<description>`
- `<title>`
- Text after the `<body>` tag :
  - the title tags `<h1>` , `<h2>` , etc
  - the visible text for your audience
- anchor text for links (navigation, etc)
- ALT attributes for images
- If it makes sense, add a title attribute for images and links

### SILO structure

SILO is the overall directory structure and organization of your Website itself.

Use a specific directory name for each subject.

For example, tell the world about your hobbies : martial arts / biking / swimming / dancing / IT / foreign countries. In this case, you have to structure your site into specific subdirectories in order to have a good ranking in search engines!

### Content hub

Create one or more content hubs (also called center web pages), then interlink the relevant web pages to each hub. Like that, it's easier to rank well. Google determines context and relevancy not only by what's on your web page, but also by the web pages around it and linking to it.

Example :

hub	relevant web pages
Sport	baseball, cricket, biking, fitness ...
IT	web development, programming, IT security ...
Health	yoga, raw food, diet ...

Think of the hub as the master document that acts as an overview or gateway to all of your individual content web pages. The hub is the authority page. Often, the hub is a link-bait web page or a category-level web page. It's typically the web page with the most inbound links and it's often a landing page for your Website.

## Secure content against scrapers

The practice of fetching your content and re-publishing is called "scraping," and the scrapers perform remarkably well in search engine rankings, they often outrank the original Websites.

### Ping

Pinging is a process to notify the search engines that you have added new information in your Website or changed the configuration. After that, search engines will crawl and index your Website. Pinging is helpful when used wisely. You should use Ping services after writing any post or changing some content of your Website. I will recommend you to use it just once in a day. If you will use it more frequently, search engines will consider it as spam and ban your Website.

### Links

Use absolute links <http://phoenix-it-mos.net/web-design-and-development/> and NOT relative ones like [../webdesign](#). If you do it this way, the link remains pointing to your Website when a scraper picks the link up and copies the content.

## Spam

Source : [MOZ](#)

### Introduction

**Spamming means** creating pages and schemes designed to artificially inflate rankings or abuse the ranking algorithms.

**Manipulative techniques generally won't help you** and they often result in search engines penalties on your Website.

To **identify spam**, see Google's Webmaster [Guidelines](#) and Bing's Webmaster [FAQs](#).

### Web page spam

#### Keyword Stuffing

One of the most obvious and unfortunate spamming techniques, keyword stuffing, involves littering keyword terms or phrases repetitively on a page in order to make it appear more relevant to the search engines.

#### Manipulative Linking

One of the most popular forms of web spam is manipulative link acquisition. This means attempting to exploit the search engines use of link popularity in their ranking algorithms to artificially improve visibility.

#### Cloaking

A basic tenet of search engine guidelines is to show the same content to the engine's crawlers that you'd show to a human visitor. This means not to hide text in the HTML code of your Website. When this guideline is broken, the engines call it "cloaking" and take action to prevent these pages from ranking in their results.

## Low Value Pages

Low value pages are technically not considered as web spam. The search engines have methods to determine if a web page provides unique content and value to its searchers. The most commonly filtered types of web pages are thin affiliate content, duplicate content and dynamically-generated content pages that provide very little unique text or value.

## Domain level spam

### Links

Just as with individual web pages, the search engines can monitor the kind of links and quality of referral traffic sent to a Website. Websites that are clearly engaging in the manipulative activities described above on a consistent or seriously impacting way may see their search traffic suffer or even be banned from the index.

### Trustworthiness

Websites that have earned trusted status are often treated differently from those that have not. For the search engines, trust most likely has to do with the links your domain has earned. If you publish low-quality, duplicate content on your personal blog and additionally buy several links from spammy directories, you're likely to encounter considerable ranking problems. However, if you post that same content on Wikipedia, even with the same spammy links pointing to the URL, it would still rank tremendously well. Such is the power of domain trust and authority.

### Content value

As we've seen, an individual web page's value is computed in part based on its uniqueness and the visitor's experience; likewise is the entire domain's value assessed. The search engines simply don't want thousands of copies of Wikipedia filling up their indexes, so they use algorithmic and manual review methods to prevent this. When users click on a result, then quickly hit the back button on their browser and try another result, this indicates that the result didn't meet their expectations. It's not enough just to rank for a query. Once you've earned your ranking, you have to prove it over and over again.

## Clean URLs

Clean URLs are also called semantic URLs. These URLs are intended to improve the usability and accessibility of a Website.

Example for a clean URL : <http://domainname.com/subject/>

Example for non clean URL : <http://domainname.com/filename.ext>

Filenames and missing trailing slashes are deprecated in clean URLs!

If you change to clean URLs think about creating a new sitemap !

### Example for a solution

The clean URLs can be implemented by a SILO structure.

Example :

[index.php](#) (the homepage)

[/about/index.php](#)

[/services/index.php](#)

[/consulting/index.php](#)

[/web-development/index.php](#)

[/system-management/index.php](#)

[/contact/index.php](#)

If you call for example <http://domainname.com/services/consulting/> the content of the index page in this subdirectory will be displayed.

With this approach, we have a duplicate content problem. There are two URLs serving the same content : <http://domainname.com/meaning/> and <http://domainname.com/meaning/index.php>

In order to have one unique URL, you can use the following .htaccess code. It redirects the <http://domainname.com/directory/index.php> to <http://domainname.com/directory/> based on the rewrite rules with mod\_rewrite.

Code :

```
RewriteCond %{THE_REQUEST} ^[A-Z]{3,}\ /meaning/index\.php
```

```
RewriteRule ^ /meaning/ [R=301,L]
```

Of course, you have to turn on the rewrite engine first !

## Domain URL

Use descriptive keywords for the domain name.

## Possibilities for blogs

- Own domain : yourblog.com
  - Google will list the blog as a separate domain (of course) and it will be shown in the country and international SERP
  - You can sell the blog (domain) later
- Subdomain : blog.yoursite.com
  - Google will list the blog as a part of the Website in the country SERP, but not outside the country SERP.
  - This version will enhance the ranking of your Website and preserve the Websites authority.
- Subfolder : yoursite.com/blog
- Eventually replace "blog" for subdomain and subfolder by a more significant name



## Duplicate content

A homepage can basically be accessed by :

<http://www.domainname.com/index.php>

<http://www.domainname.com/>

<http://domainname.com/index.php>

<http://domainname.com/>

This will produce a duplicate content problem for search engines ! So, ideally you choose only one URL to be served. If you decide to use the non www version of the URL (<http://domainname.com/>) you can follow the steps below.

### First step

- Redirect the Website URL from www to non www in the .htaccess

```
# rewrite engine basics
```

```
Options +FollowSymLinks
```

```
RewriteEngine on
```

```
RewriteBase /
```

```
# redirect to non www
```

```
RewriteCond %{HTTP_HOST} ^www\.(.+)$ [NC]
```

```
RewriteRule ^(.*)$ http://%1/$1 [R=301,L]
```

- Links : make sure that all links, especially the internal ones, pointing to the non www URL
- Webmaster Tools : for new Websites to add - add the site URL in GoogleWT and BingWT without the www
- Webmaster Tools : for Websites already added with the www - adjust the preference to the non www version

## Other problems

After the first step, your Website will still serve two URLs (duplicate content) :

<http://domainname.com/index.php>

<http://domainname.com/>

The other problem is the URL <http://domainname.com/index.php> is considered as a **non clean URL**, (you have to avoid this).

## The second step

You have to redirect the <http://domainname.com/index.php> to <http://domainname.com/> in the .htaccess

```
# redirect index.php to URL
```

```
RewriteCond %{THE_REQUEST} ^[A-Z]{3,9}\ /index\.php\ HTTP/  
RewriteRule ^index\.php$ http://domainname.com/$1 [R=301,L]
```

## Another problem ?

After the second step your Website will still serve two URLs. A domain with and another without a trailing slash (**duplicate content**):

<http://domainname.com/>

<http://domainname.com>

**The usual duplicate content rules don't apply to the homepage !** <http://domainname.com> and <http://domainname.com/> are regarded as the same thing. The browser removes the trailing slash by default.

If you are a bit paranoid and you want to get rid of a potential duplicate content problem, use the canonical meta tag to prioritize the URL with a trailing slash.

## Duplicate content

### Basically you have the following situations which generate duplicate content :

1. **Home page** variants like www and non www as well as index.html and /
2. not having **Clean URLs** : having /dir/index.html and /dir/ for the web page
3. having duplicate **web pages**
4. having the same "bio" in all **social media**
5. republishing blog **articles**

### Basically you can do the following against duplicate content :

1. **Home page** : using the "rewrite " and "redirect" in the .htaccess
2. **Clean URLs** : using SILO and "rewrite" in the .htaccess
3. duplicate web pages :
  - use **redirect** in the .htaccess (301)
  - use the **canonical meat tag**
  - **noindex** of the duplicate pages (not recommended by Google)
4. **Social media** : personalize the "bio"
5. **Articles** : do not copy

## Redirect

When multiple web pages with the potential to rank well - are redirected to a single **web page (A)**, they not only stop competing with each other, but also create a stronger relevancy and popularity signal overall. This will positively impact **web page (A)** to rank well in search engines.

## Robots

### Introduction

Robots specify an access policy for spiders (web crawlers, robots, user-agents). While robots.txt may restrict crawling, only Meta Robots and X-Robots will remove URLs from search results.

	<b>works on</b>	<b>links</b>	<b>user agent choice</b>
<b>Robots.txt</b>	all files/dirs	-----	yes
<b>Robots meta</b>	web pages	all links	yes
<b>X-Robots</b>	all files/dirs	all links	yes

## Robots.txt

Robots.txt is also known as the robots exclusion protocol. The robots.txt file must reside in the root of the domain.

### 1)

If you want all spiders to index everything from your Website, you basically don't need a robots.txt file. If you want to **exclude** some files or directories for some (or all) spiders, then the robots.txt is a good method to do that.

Examples:

Exclude all spiders from the whole site	User-agent * Disallow: /
Exclude the Alta Vista spider	User-agent scooter Disallow: /
Exclude some directories for all spiders	User-agent * Disallow: /data Disallow: /private_homepage Disallow: /keys

### 2)

Aside from the include/exclude mechanism, you can indicate to the search engine the **location of the sitemap**. <sitemap:www.yourWebsite.com/sitemap.xml>

### 3)

The **Crawl-delay** indicates the speed in milliseconds at which a spider can crawl a server. The crawl-delay is not part of the [original robots.txt specification](#). But it's no problem to include it for those spiders that understand it. Example : [Crawl-delay: 13](#)

### 4)

Some **speed tests** "require" a robots.txt file.

## Robots meta tag

The robots meta tag is not only another exclusion method. It can be used to allow to index or not a specific web page and it can be used to allow to follow or not all the links of a web page. If the tag is not defined, the default is "index, follow".

Examples :

The web page will be indexed, the spider will follow the links

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

```
or <meta name="robots" content="all" />
```

The web page will not be indexed, the spider will not follow the links

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

```
or <meta name="robots" content="none" />
```

The web page will be indexed, the spider will not follow the links

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

The web page will not be indexed, the spider will follow the links

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

The **name** can be replaced by the user agent of a specific spider. If the **name** is **robots**, it means it is valuable for all spiders.

Some user agents : Google (googlebot), Yahoo (slurp), Bing (bingbot)...

## X-Robots

The X-Robots-Tag can be used as an element of the HTTP header response for a given URL. Any directive that can be used in an robots meta tag can also be specified as an X-Robots-Tag.

You can add the X-Robots-Tag to a site's HTTP responses using .htaccess and httpd.conf files that are available by default on Apache based web servers.

The benefit of using an X-Robots-Tag with HTTP responses is that you can specify crawling directives that are applied globally across a Website. The support of regular expressions allows a high level of flexibility. You can use the X-Robots-Tag for non-HTML files like image files where the usage of robots meta tags is not possible.

**Example** for image files :

```
<Files ~ "\.(png|jpe?g|gif)$">  
  Header set X-Robots-Tag "noindex"  
</Files>
```

## Technical on-page SEO

### Introduction

Technical on-page optimization refers to all measures that can be taken directly within the Website in order to improve its position in the search rankings. The term "technical" means all elements on a web page except the content.

The on-page factors are controlled by you in what you are doing on your web page.

Examples of on-page optimization includes the links, meta tags, structured data markup, mobile optimizing, HTML code, etc.

### Link structure

#### Overview

A crawlable link structure is vital for SEO. The search engines crawlers must be able to follow the links of a Website in order to get all web pages.

- Be sure that all web pages are interlinked
- In order to build a crawlable link structure use the `<a href="destination"></a>` tag only !
- Avoid web pages with hundreds of links, this will be interpreted as spam
- Stay away from using frames and iframes for web pages to be indexed
- Do not block web pages with robots.txt or the meta robots tag

Internal links are not equally qualified by the search engines. When you link your internal web pages within your content hub together, keep in mind a few important points:

- Links from the content (contextual) are more valuable than navigation links (menu)
- Links higher up on the page are more valuable than links further down, links to your most important content pages higher up in the HTML text
- Links in HTML text are more valuable than image links

### Breadcrumbs

Breadcrumbs are equally beneficial to the search engines as they are to users. They provide logical access to all the pages of your Website.



## Contextual links

A contextual link is the clickable text (keyword) within the written content of your web page. Contextual linking can be a very strong SEO linking technique for your Website.

## The canonical meta tag

Canonicalization deals with duplicated content on Web pages.

The canonical tag tells search crawlers which page is the singular, authoritative version that should count in web results. Google supports the tag on both single domains and across multiple domains. The canonical meta tag indicates the search engines which unique URL to index.

**Example** (to put in all duplicate web pages)

This is the page to index :

```
<link rel="canonical" href="http://www.example.com/mypage.html/" />
```

## The hreflang meta tag

hreflang deals with alternate language/country definition - using hreflang makes a better ranking in SERP "allover the world". If you use the hreflang meta tag, Google will show the **search result** in each country specific Google search engine, independent in which language the search query was.

If hreflang is not used, Google will figure out the language and country :

- from the ccTLD (country code Top-Level Domain)
- the language you're declaring in <html> and/or <meta>
- the language you are using on your Website

The implementation of hreflang makes sense if you have a **multi language Website** with web pages that have their counterpart in one or more languages. Using hreflang allows you to define precisely the language country combination for a specific URL.

Example for the German language :

- de-ge for Germany <http://example.com/ger>
- de-ch for Switzerland <http://example.com/swi>
- de-at for Austria <http://example.com/aus>

Example :

- **using** the hreflang  
Search for "Site web et blog recommendations" in Google.com as well in Google.fr gives a high ranking for the search result in both search engines.
- **not using** the hreflang  
Search for "Site web et blog recommendations" in Google.com as well in Google.fr gives a high ranking in the search result in Google.fr. No search result in Google.com !

## Implementation overview

You have three possibilities :

1. in the <head> of Web pages
2. in the sitemap
3. in the HTTP header for non-HTML files

## Implementation example

If page A links to page B, page B must link back to page A, otherwise the "hreflang" may not be interpreted correctly.

- home page on the English Website <http://phoenix-it-mos.com/>

```
<head>
<link rel="alternate" hreflang="en" href="http://phoenix-it-mos.com/" />
<link rel="alternate" hreflang="fr" href="http://phoenixfr.com/it/" />
</head>
```

- home page on the French Website <http://phoenixfr.com/it/>

```
<head>
<link rel="alternate" hreflang="fr" href="http://phoenixfr.com/it/" />
<link rel="alternate" hreflang="en" href="http://phoenix-it-mos.com/" />
</head>
```

## The title tag

```
<title>here comes your title</title>
```

The title is very much responsible for helping your web page to not only climb up the search engine ranks, but also to convince your potential visitors of the relevance of your web page.

The optimal length for a title tag is usually 50-60 characters.

Put the most important keyword in the natural and descriptive title, write it for users, then optimize it for search engines. Each web page should have its own (unique) title !

## The meta description tag

```
<meta name="description" content="your content comes here" />
```

A meta description is a snippet of information that appears below the title in a SERP. Its major purpose is to provide useful and relevant information about the content of a web page. A good meta description convinces the visitor to click on the link for your web page. The meta description length should be less than 150-155 characters. Anything longer will be truncated.

Put the important keywords in the meta description, write it for users, then optimize it for search engines. Each web page should have its own (unique) meta description ! Call for action and telling the visitor what he/she can benefit from the web page are good strategies.

## The heading tag

`<h1>this is the most important heading</h1>`

`<h6>` is the less important one. Headings are the simplest way to divide the content of your web page into different smaller sections to make it more comprehensible for the visitors. For SEO, these tags have a lot more importance than other HTML tags. Use descriptive (long tail) keywords in heading tags.

Search engines primarily take keywords from `<title>`, **content** and **heading tags** to define the context of a web page.

### The rel="nofollow" attribute

```
<a href="http://www.example.com" title="example" rel="nofollow">anchor text</a>
```

The "nofollow" attribute in the example above tells the search engines **not** to follow this specific link.

Here are some cases in which you might want to consider using nofollow :

- **Untrusted content**
- **Paid links**

A site's ranking in Google search results is partly based on analysis of those sites that link to it. In order to prevent paid links from influencing search results and negatively impacting users, webmasters should use "nofollow" on such links.

- **Crawl priority**

Search engine robots can't sign in or register as a member on your forum, so there's no reason to invite Googlebot to follow "register here" or "sign in" links.

- **Recursive links**

Link from the menu, which point to the page it resides

**Additional information :** [support Google](#)

## The structured data markup

### Introduction

"Structured data markup" (or structured data) is a standard way to annotate your content in web pages so that search engines can understand it. "Structured data markup" tells the search engines what your data means, not just what it says.

Example : If you add "author" to Smith, the search engine will now "understand", that Smith is the name of the author of the web page.

When your web page includes structured data markup, Google and other search engines can use that markup to index your content better and present it more prominently in search results.

Structured data markup makes your content eligible for two kinds of Google features :

- **Enhanced presentation in search results:** By including structured data markup appropriate to your content.
- **Answers from the Knowledge Graph:** If you're the authority for certain content, Google can treat the structured data markup on your Website as factual and import it into the [Knowledge Graph](#), where it can power prominent answers in search and across Google properties.

## Schema markup

Schema.org is an initiative launched on June 2011 by the operators of the world's largest search engines Google, Yahoo and Bing to create and support a common set of **schemas** and **schema formats** for structured data markup on web pages.

Data in the schema.org (semantic) vocabulary is using three schema formats: **microdata**, **RDFa**, and **JSON-LD**.

### - Microdata

Microdata is probably the most popular markup for structured data. It's an extension to **HTML5**. Microdata has the three basic elements **itemscope**, **itemtype** and **itemprop**, which are added to the HTML or meta tags.

The **itemscope** attribute should be placed in a HTML element that contains all of the data relating to an item.

The **itemtype** attribute is used to markup what an item is. The value of this attribute should contain an URL to a schema that defines all of the properties that the item can have.

The **itemprop** attribute is used to markup an item property. The value of this attribute should contain the name of the property. The contents of the element will be used as the value for the property.

### - RDFa

RDFa (**R**esource **D**escription **F**ramework in **a**tttributes) is an extension to **HTML5**. RDFa has the four basic attributes: **vocab**, **typeof**, **property** and **resource**. RDFa uses a number of properties to identify entities such as a person or an event. It applies its markup in HTML tags, such as `<div>` and `<span>`.

### - JSON-LD

JSON-LD (**J**avaScript **O**bject **N**otation for **L**inked **D**ata) is a JSON based data format. This format is valid XHTML. **The mark-up is invisible for users, this is against search engine guidelines. Google still favors Microdata and RDFa !**

## Microformat

Microformat allows to **implement** structured data markup which can not only be read, but also understood by search engines. Google supports a subset of the Microformat !

Microformat extends conventional HTML tags with semantic information in order to add structured data to a web page. With the Microformat, you'll mostly use the class attribute. This format is valid XHTML.

### - The rel="author" tag

The "rel" tag is part of the Microformat. Put a hyperlink on a web page with the "rel=author" tag. This tells the search engines (basically) that you are the author of that web page.

### - The rel="publisher" tag

The "rel" tag is part of the Microformat. Put a hyperlink on a web page with the "rel=publisher" tag. This tells the search engines (basically) that you are the publisher of that web page. You have to add the Websites URL in the G+ page in order to let Google incorporate the Website (pages) in the knowledge graph.

## Rich Snippets

Rich Snippets is a Google term for **implementing** structured data markup. It includes schema format (microdata, RDFa, JSON-LD) and a subset of Microformat.

Rich Snippets can be embedded **visibly or invisibly for users, in HTML tags or meta tags.**

Including Rich Snippets in web pages will :

- help Google algorithms better **index** and understand the content of your web pages
- display Rich Snippets within the **search results**
- display Rich Snippets within **Google+**



## Social media markup

### - Open Graph

Open Graph is a type of structured data markup used by Facebook to parse information of a web page.

With Open Graph you can tell Facebook what to display. You have to add the Open Graph markup into the `<head>` of your web page.

### - Twitter Cards

Twitter Cards is the structured data markup used by Twitter. With Twitter Cards you can attach photos, videos and media experience to tweets that drive more traffic to your Website. Users who tweet links to your content will have a “Card” added to the tweet that’s visible to all of their followers.

## Mobile optimizing

### Introduction

The W3C's recommendation says : "One Web means making, as far as is reasonable, the same information and services available to users irrespective of the device they are using".

I will discuss here only one of tree approaches. This approach uses **CSS media queries** to modify the presentation of a Website based on the size of the display. A key advantage of this approach is that designers can use a single template for all devices, and just use CSS to determine how content is rendered on different screen sizes.

There is a growing number of mobile/responsive-friendly, open-source frameworks like [Twitter-Bootstrap](#) or [Foundation](#) which help accelerate the process of building responsive Websites.

### The HTML viewport

Even for a basic page, you **must** always include a viewport meta tag. The viewport is the most critical component you need for building multi-device experiences. Without it, your Website will not work well on a mobile device.

Using the meta viewport value `width=device-width` instructs the page to match the screen's width in device independent pixels. This allows the page to reflow content to match different screen sizes, whether rendered on a small mobile phone or a large desktop monitor.

Adding the attribute `initial-scale=1` instructs browsers to establish a 1:1 relationship between CSS pixels and device independent pixels regardless of device orientation, and allows the page to take advantage of the full landscape width.

```
<meta name="viewport" content="width=device-width, initial-scale=1" />
```

### The CSS viewport

The meta viewport tag, while broadly supported, is not part of a formal standard. This behavior is being included in CSS as part of the [CSS Device Adaptation](#) specification. Until this specification is finalized and widely implemented, authors should continue to **use the meta viewport tag for compatibility with the corresponding @viewport styles.**

```
@-webkit-viewport { width:width=device-width; zoom: 1.0;}  
@-moz-viewport { width:width=device-width; zoom: 1.0;}  
@-ms-viewport { width:width=device-width; zoom: 1.0;}  
@-o-viewport { width:width=device-width; zoom: 1.0;}  
@viewport { width:width=device-width; zoom: 1.0;}
```

## The CSS @media rule

The CSS @media rule includes a block of CSS properties for a specific **media type** (device) and a specific **condition**.

### - Example :

```
@media screen and (min-width: 501px) /* for 501 px or more */
```

```
{
```

```
    CSS code for desktop comes here
```

```
}
```

```
@media screen and (max-width: 500px) /* for 500 px or less */
```

```
{
```

```
    CSS code for mobile comes here
```

```
        content width = 100%
```

```
        overall horizontal unit = %
```

```
        font : eventually use Google fonts
```

```
        replace background images by color codes
```

```
        look for a solution for sub menus
```

```
        do not use a global height for content boxes add padding top and bottom instead
```

```
}
```

## Dimensions

### - Font

The minimum font size should be 16px (1em / 100%), the line height should be at least 1.2em (19px)

### - Touch elements

Touch elements should have at least a margin of 32px (2em / 5mm) and should have a minimum height of 48px (3em / 7mm)

## **The sitemap**

Sitemaps help search engines find and classify content on your Website that they may not have found on their own.

XML is the most widely accepted format for sitemaps. RSS is another format used for sitemaps, RSS means "Really Simple Syndication" or "Rich Site Summary". You can also use a text file as sitemap. The text sitemap format is one URL per line up to 50,000 lines.

## **SSL**

The latest moment to Implement SSL is for your e-commerce (order forms, transactions...) !

## Speed

In this section you will learn some important steps to make your Website faster. Why speed up your Website ? Because Google likes it and your visitors as well ! Choose a fast (and reliable) host, this will be your foundation for running your Website.

## Compression

Compression on the Web server is also called HTTP compression. The compression increases the page load speed. This concerns all transferred files. File examples : HTML, CSS, JavaScript, PHP, images, videos, etc. The most common compression schemes are **Gzip** and **Deflate**.

Many shared web servers do not support compression !

You can **check** the Gzip compression status in using the Firefox plugin "Web developer".

If Gzip is **enabled**, you will get a response header like :

Date: Thu, 23 Oct 2014 13:01:31 GMT

Server: Apache

X-Powered-By: PHP/5.4.34

Content-Encoding: gzip

Vary: Accept-Encoding

Keep-Alive: timeout=2, max=200

Connection: Keep-Alive

Transfer-Encoding: chunked

Content-Type: text/html

200 OK

If your shared host did not activate the Apaches Gzip module, you have to use the PHP's Gzip functionality. You have the possibility to activate compression via php.ini or .htaccess

## The CSS file

Avoid the `@import` rule. If you use it, your web page load time will slow down.

## The images

Image optimizing is the easiest and (most) efficient task to speed up your Website. Make sure that every image you use on your Website is optimized and compressed in order to consume as little bandwidth as possible. Start with the **header** and the **logo** and don't forget the **favicon**.

### The dimensions

1. Scale the images with Gimp (don't resize them via CSS)
2. Specify the width and height with CSS (for faster rendering)

### The weight

- Use an **appropriate** image **format** (jpg, png, gif...)
- **Remove** the **EXIF** data from the image with the Linux "trimage image compressor"
- **Compress** a maximum- example : `$ convert -quality 50% old.jpg new.jpg`
- Make **progressive JPEGs** : `$ jpegtran -progressive old.jpg > new.jpg`
- Compress **animated GIFs** : <http://ezgif.com/>
- Compress **PNGs** : <http://compresspng.com/>

## The browser cache

The static content will be cached in the browser . This gives a much faster response time for your Web pages ! The static content concerns CSS, JavaScript, images (including favicon), video files, etc. In order to optimize the browser cache, you have to define the resources and the time they reside in the cache.

1) Put the following code into the .htaccess

```
<ifmodule mod_expires.c>
ExpiresActive On
ExpiresByType image/gif "access plus 1 month"
ExpiresByType image/jpeg "access plus 1 month"
ExpiresByType image/jpg "access plus 1 month"
ExpiresByType image/png "access plus 1 month"
ExpiresByType image/ico "access plus 1 year"
ExpiresByType image/x-icon "access plus 1 year"
ExpiresByType text/css "access plus 2 weeks"
ExpiresByType text/javascript "access plus 2 weeks"
ExpiresByType application/javascript "access plus 2 weeks"
ExpiresByType application/x-javascript "access plus 2 weeks"
ExpiresDefault "access plus 1 week"
</ifmodule>
```

You have to add resources if the above code doesn't cover all resources of your Website !

2) No cache for the Homepage. Put the following code into the index.php after the `<html>` tag :

```
<?php>
header("Cache-Control: no-cache, must-revalidate");
header("Expires: Sat, 26 Jul 1997 05:00:00 GMT");
?>
```

## Minify

Minify removes all unnecessary elements like comments, white spaces, etc. in text files. This increases the page load speed. Minify concerns files like HTML, CSS, JavaScript, etc.

A good practice is also to combine files of the same type.

Minification is similar to compression. But code that has been minified is still plain text and can be used without decompression!

## HTTP headers

### Character set

Specifying a character set in HTTP headers can speed up browser rendering.

Add to the index.php file : `<?php header("Content-type: text/html; charset=utf-8"); ?>`

Check the response header, it must be **like** : Content-Type: text/html; **charset**=utf-8

There is also another method to implement this, but I selected the easier one. The other method uses the `auto_prepend` directive in the `php.ini` to add the necessary header.

### Vary: Accept-Encoding header

Specifying the Vary: Accept-Encoding header instructs the proxy to store both a compressed and uncompressed version of the resource.

If the Vary: Accept-Encoding header is missing, the proxy with a bug serves a compressed resource to a client that doesn't support compression.

Add the following code to the `.htaccess` – you have to mention the needed **file extension(s)** !

```
<IfModule mod_headers.c>
<FilesMatch "\.(js|css)$">
Header append Vary: Accept-Encoding
</FilesMatch>
</IfModule>
```



## Off-page SEO

Off-page SEO refers to activities outside the boundaries of the Website. Search engines analyze off-page SEO signals in addition to on-page factors to determine the Websites ranking.

The most important off- page factors are:

- Link building (quality links)
- Social Media (engagement)
- Forum (answer questions)
- see also : Web marketing (strategies/techniques)

## Link building

As from the perspective of SEO, **90%** of your effort should go into creating **great content**, and **10%** into **link building**, because good and relevant content does link bait.

Link building is the activity of acquiring links from other Websites to yours. The more high-quality Websites link to you - the more likely you are to **rank** well in search results. **Keep in mind that all good link building is relationship building.**

Link building should be a constant stream of activity **coupled with other initiatives** such as content creation, great customer service and social media. Combined, **all of these activities can lead to an increase in traffic and an increase in revenue for your business.** It is the **combination** and a **consistent** drive towards the goal that will generate success.

## Web marketing

### Introduction

**Web marketing** (Internet marketing) is the broadest approach and the umbrella for **SEO** and **SEM**. The goal is to attract as much customers as possible. The marketing connects the customers and the products / services in order to **realize sales**.

**Search Engine Optimization** (SEO) is about tactics to **improve** the **organic search results** for a Website.

**Search Engine Marketing** (SEM) is about the multitude of other type of search results that appear on search like : paid, local, video, social, etc and the things that are necessary to gain and manage **visibility** for those results - and finally to convert visitors to clients (**sales**).

### Keys

From the perspective of **marketing** **80 %** should be used for **publicity** and **20 %** for **content**.

Keep in mind that neither your customers' experience nor your brand starts with Twitter, Facebook, or your blog. **Social media** should take your existing brand and solidify it.

A realistic expectation to set is that you will see an **increase in traffic** if you are constantly working on improving your website, business, and online activity.

**Engagement** **matters more than any other factor** (in SEO) not because Google needs it to determine "good" sites. It matters because everything everyone does online can be tracked and analyzed.

Since **conversion rates** for e-shops **are usually very low**, conversion marketing can be a useful way to boost this number, online revenue and overall Website traffic.

You should always be focused on building a positive user experience and create quality content that attracts the right users and leads to conversions. SEO is the foundation for successful conversion.

## Monitoring and analyzing

Monitoring and analyzing is **one of the success keys**. You can detect mistakes, things to optimize as well as user behavior, etc. Monitoring and analyzing gives you the information to improve/develop your Website.

A good **SEO/marketing** campaign relies on not only implementing changes but also measuring the impact of those changes !

Use basically GWT and Google Analytics. In using **analytics** software, you will have insights of the changes you made and the performance of your Website.

**Observe** the important **metrics** in order to see how your Website performs.

**Stay up to date** on search engine news (algorithms...) and techniques (SEO, marketing) in order to optimize your web pages.

## Improving and developing

Improving and developing is an ongoing process after you realized your Website. Monitoring and analyzing will give you the data to improve/develop your Website. You can't improve your Website straight forward for specific search engines, because their algorithms aren't public.

Basically if you use "**best techniques**" you will rank better. This concerns : on-page SEO, link building and marketing.

With the combination of "best techniques", the news and the monitoring/analytics data you are able to design an **action plan** for improving and developing your Website.

Be active on **social medias** and create excellent **content** on a regular basis.

If you are improving and developing **on a regular basis**, you will get better ranking.