

THE EXPERT'S VOICE® IN WEB DEVELOPMENT

Joomla! 3 SEO and Performance

*DOMINATE SEARCH ENGINE RANKINGS
AND ATTRACT MORE VISITORS
TO YOUR SITES*

Simon Kloostra

Apress®

www.allitebooks.com

For your convenience Apress has placed some of the front matter material after the index. Please use the Bookmarks and Contents at a Glance links to access them.



Contents at a Glance

About the Author	xiii
About the Technical Reviewer	xv
Introduction	xvii
■ Chapter 1: How Search Engines Work	1
■ Chapter 2: Preparing Your Joomla Environment	15
■ Chapter 3: Joomla Global Configuration.....	21
■ Chapter 4: Metadata	29
■ Chapter 5: URL Structure.....	37
■ Chapter 6: Header Tags	45
■ Chapter 7: Optimize Internal Links and Anchor Text	51
■ Chapter 8: Optimize Images	53
■ Chapter 9: Microdata or Rich Snippets	55
■ Chapter 10: Duplicate Content.....	61
■ Chapter 11: 404 Errors and Pages.....	69
■ Chapter 12: robots.txt	79
■ Chapter 13: Sitemap.....	83
■ Chapter 14: Simple SEO Extensions	87
■ Chapter 15: SH404SEF.....	93

■ CONTENTS AT A GLANCE

■ **Chapter 16: SEO Site Audits 109**

■ **Chapter 17: Off-site SEO 117**

■ **Chapter 18: Joomla Performance and Speed 127**

■ **Appendix A: Glossary..... 159**

■ **Appendix B: Checklist..... 163**

Index..... 165

Introduction

This book is aimed at Joomla enthusiasts and professionals who want to increase the number of visitors to their Joomla websites by improving their performance in search engines such as Google, Bing, and Yahoo!. This means not only ranking high but also improving the presence of a site in the search pages. This book is written for Joomla 3, but it can be used perfectly for Joomla 2.5 as well.

It will show you how to optimize your Joomla website for search engines: first, by using only the Joomla core options and, later in the book, with extensions, including a chapter dedicated to SH404SEF, which is probably the most popular full search engine optimization/search engine friendly (SEO/SEF) extension for Joomla. (Beginners can skip this chapter, as the book does not hinge on this subject.)

You will also gain general background information on SEO-related topics, as this is necessary for fully understanding what you are doing. The main focus of this book will be on on-site optimization, although some attention will also be paid to off-site optimization, such as linkbuilding and setting up Google Analytics and Webmaster Tools.

Second, the book is about performance, explaining how to create a lightning-fast website, first by optimizing the Joomla core options and, second, by using some specialized extensions.

To understand this book, you should have at least a basic background in managing Joomla, meaning that you understand the basics of creating articles and menus and are able to install and configure extensions. If you do not yet have this experience, it would be wise to read a starter book on Joomla first. An excellent primer would be *Beginning Joomla!: From Novice to Professional*, by Dan Rahmel.

No knowledge of SEO is required; this book will instruct you as you go along. Chapter 1, especially, will provide the principal introduction to the world of SEO. If you already have some familiarity with SEO, you can also skip this chapter and jump straight to Chapter 2 for information related specifically to Joomla.

CHAPTER 1



How Search Engines Work

Before we start optimizing Joomla for search engines, let's quickly go through some of the basics of how Google ranks a web site. Feel free to skip this chapter, if this is already familiar to you. I will only discuss the basics here; more advanced search engine optimization (SEO) topics will be discussed as we go, while we are configuring Joomla.

You can attract an audience to your web site in many ways. Roughly, the most important ones are the following:

- Organic search (Google)
- Paid search (the ads around the organic traffic)
- Social media
- Links from other web sites to yours
- Visitors directly typing in your URL

In this book, I will focus mainly on the first method, organic search; the others are less Joomla-specific. I will, however, spend some time on the other channels too, to get you on track.

Organic search does not just mean getting a high rank in Google, it is also about being noticed on the search pages. If the result for your site stands out in a positive way, people might be more inclined to click it, even when yours is not the first site on the page, but the third or fourth. This is often underestimated, and grasping its importance can seriously get you more visitors.

Also, SEO is about optimizing *all* pages of your web site, not just the home page. Many of my tips are about optimizing all individual pages of your site. Of course, the home page will often bring you the most organic visitors, but optimizing the other pages, too, helps them appear in the search results as well. Second, the overall optimization of the site will be appreciated by the search engines and possibly result in higher rankings.

Note that where I refer to *Google*, I also mean other search engines as well. Roughly, they all work the same way, as far as we are concerned in this book, so when I mention Google, I basically mean search engines in general.

The Scanning Process

Google is constantly scanning and crawling all web sites online with so-called robots, bots, or spiders. Nothing physical about these, but that's just what they are called. This is an automated software procedure, used to determine how web sites are built up. How frequently your site is visited by these robots (as I will refer to them from now on) depends on your site. A small site that hardly changes and has been sitting there for years may only receive a weekly visit from a robot, but very dynamic sites such as news blogs may be spidered multiple times a day, up to thousands for very large sites.

Starting from the home page, these robots basically read all code, follow links (mainly internal but also external), and try to interpret what the web site is about. It is very important to understand that Google basically reads code. It cannot (yet) understand resources such as images, Flash, sound, video, etc. The robot almost acts like a blind person trying to understand the world around him. A lot of guidelines that Google recommends to web masters are actually retrieved from accessibility guidelines that help disabled persons read/hear web sites. Google uses **context** to understand these non-text resources, such as images. As an example, if you want to display an image of a **bird**, you should make sure it is named **bird.jpg**.

Google also uses the structure of the HTML of a page to guess what the web site is about, using the order in which content is created (important stuff first, less relevant later), but also it appreciates any effort in structuring the code, such as using the correct tags in your HTML. As an example: **<section>** elements are believed to be more relevant than **<aside>** and **<footer>** elements. Also, other elements, such as **<h1>**, **<h2>**, **<p>**, **<quote>**, etc., can be interpreted and used to build structure for headings, paragraphs, quotes, etc.

So for now, the main thing to remember is: **Google basically reads text**, at least as far as optimizing the actual web site is concerned, while also taking into account structure and hierarchy. This is an extremely simplified way of looking at things, but it may help you to understand the process.

Google also uses **off-site** factors, such as links and popularity on social media, but for a new site, that is only relevant once the site is live. I will pay some attention to off-site SEO later, but our main focus will be **on-site SEO**.

Google Updates: Panda, Penguin, and Hummingbird

Google is making constant changes in its algorithms. Some of this is simply because people are trying to trick Google, and it wants to prevent that. Second, it is really trying to offer the most relevant answers to what people are looking for. Many of these changes are only small changes, sometimes on a daily basis, but a few major updates to the algorithms have been shaking the rankings more than others. You may have heard of the **Panda**, **Penguin**, and **Hummingbird** updates. These updates were mainly bad for sites that offered irrelevant content, tried to rank through irregularly acquired backlinks, etc., basically not providing what a user would really search for. Many of these updates favored content on smaller authoritative sites over larger generic sites, so your on-topic Joomla site may benefit from the ongoing Google changes.

I could devote a full book to these updates, but basically, you should remember that Google wants you to create sites with relevant content that are useful for visitors and structured in such a way that Google can easily understand what the sites are about. Try to avoid anything that sounds like a trick (hidden links, buying backlinks, stuffing keywords, etc.). Just ask yourself: is this really useful for my visitors, or am I just doing this because it should trick Google to believe that this is the case? Even if the trick still works, it probably will not in two years' time, and you might fall hard.

Unique and Quality Content

Before anything else, you have to stand out from other web sites competing for the same Google ranking, so make sure you have interesting and unique content that your visitors love to come back to. Make sure you write for your visitors, keeping Google in your mind all the time, but prioritize the former. If you switch these around, and only write for Google, your visitors will get bored with your content and leave the site.

What is also very important: Google loves **fresh content**, so, if possible, maintain a blog, or allow your visitors to leave comments. If your visitors can leave comments, they may assist in creating content, though you may have to moderate this, of course.

For the rest of this book, I will assume that your actual content is fine, as this is not a book on SEO writing, though you will receive plenty of tips as we go.

Keywords vs. Themes/Topics

Probably, you already heard a lot about keywords in relationship to SEO. Some still think that SEO is basically just cramming a bunch of keywords in your text, and Google will automatically rank you for those superfluously used keywords. This may have been how things worked for a while, but certainly not anymore. It is even counterproductive now and seen as bad practice.

Today, search engines want you to write high-quality content, and you will probably agree that text with just a few unnaturally used keywords is not of high quality. Instead, write text that reads pleasantly and is spot on your topic. This will appeal to your visitors and probably also to Google. If you are an expert on your topic, most probably, your keywords will occur in a logical frequency anyhow.

However, this does not mean that keywords are not important. On the contrary, it is probably one of the most important stages in the SEO process. For example, the use of synonyms, alternative terms, singular/plural usage, etc., is perfectly natural and increases the chances of ranking higher, either for one of these alternative terms or for the general theme/topic.

And, most important, do not totally forget about the most logical keywords. If you have a web site about your architecture business, make sure to use at least the word *architecture*. (I have actually seen sites on which this was done incorrectly!)

We will look at **keyword research** in a separate chapter.

SERPs, Title, Metadata

You will probably use Google on a daily basis, so you will be familiar with how a page with search results will look like. In SEO terms, such a page is a **SERP**, which stands for **search engine results page**. Memorize this; we will be using this abbreviation a lot.

What you also have to be aware of, is the following: how the record for a web site looks like in the SERP is directly related to how you configured your site!

Let's take a look at a random result that we find when we look for the keyword combination *joomla seo*. One of the results, as shown in Figure 1-1, is a page for the extension MijoSEF.

MijoSEF - Joomla SEO & SEF URLs - Miwisoft LLC

miwisoft.com/joomla-extensions/mijosef-joomla-seo-sef-urls ▼

MijoSEF can be used for Joomla SEF URLs, Metadata and Sitemap generator. It is a great way to improve your Joomla SEO.

Figure 1-1. SERP snippet for joomla seo

You will always at least see three items here:

- Title
- URL
- Meta description

The URL speaks for itself, but **title** and **meta description** are extremely important.

Title

When we speak about the title in web design, this can mean two things:

- The **title** as a user would understand it when he/she is looking at a web page, i.e., the biggest headline displayed. While this is also a very important SEO factor that I will discuss later, this is not what we are talking about here.
- **Page Title**. To be exact: This is the `<title>` tag, as it is found in the source of a web page.

When we check the source of the **MijoSEF** page, we can easily detect the title element (Figure 1-2).

```

5 <meta http-equiv="content-type" content="text/html; charset=utf-8" />
6 <meta name="keywords" content="joomla, extensions, pages, sef, seo, mijosef" />
7 <meta name="author" content="Gilbert Lici" />
8 <meta name="title" content="MijoSEF - Joomla SEO & SEF URLs - Miwisoft LLC" />
9 <meta name="description" content="MijoSEF can be used for Joomla SEF URLs,
Metadata and Sitemap generator. It is a great way to improve your Joomla SEO." />
10 <title>MijoSEF - Joomla SEO & SEF URLs - Miwisoft LLC</title>
11 <link href="http://miwisoft.com/joomla-extensions/mijosef-joomla-seo-sef-urls"
rel="canonical" />

```

Figure 1-2. Title element in source COE (common operating environment) view

When you compare the excerpt from the Google SERP page and the snippet above, it is very easy to see that the title, as defined by the administrator of the web site, is actually taken over by Google. It corresponds to the result of the page in Google. This means that you can manipulate how your search results look in the SERP! You can imagine that this can be very powerful: before visitors even reach your site, you can interact with them.

I will talk a lot about the title tag later on. For now, it is important to remember that search engines allow 65–70 characters for this in their SERPs, which is why you should make sure to keep within those boundaries. You can make longer titles, but these will be broken off, which looks ugly and should be avoided.

Also, remember that people do read this to decide whether or not to click the result. Imagine what will happen if you limit the title to just the word *Home*, which is something you will see a lot with Joomla web sites....Not informative at all!

Meta Description

In addition, the **meta description** (Figure 1-3) is taken from the source of the web page we were just looking at.

```

6 <meta name="keywords" content="joomla, extensions, pages, sef, seo, mijosef" />
7 <meta name="author" content="Gilbert Lici" />
8 <meta name="title" content="MijoSEF - Joomla SEO & SEF URLs - Miwisoft LLC" />
9 <meta name="description" content="MijoSEF can be used for Joomla SEF URLs,
Metadata and Sitemap generator. It is a great way to improve your Joomla SEO." />
10 <title>MijoSEF - Joomla SEO & SEF URLs - Miwisoft LLC</title>
11 <link href="http://miwisoft.com/joomla-extensions/mijosef-joomla-seo-sef-urls"
rel="canonical" />

```

Figure 1-3. Meta description in source code view

So, again, we have another means of influencing the way your web site looks in the SERPs. Only this time, we have some more space that we can use. If you limit the length of the meta description to something between 150 and 165 characters, you will usually be on the safe side, and no text will be cut off.

Note that it is not guaranteed that Google will take over your configured title and meta description: if it is totally unrelated to your content (which would be a bad habit), Google will generate its own snippet. Also, depending on what the user looks for, another snippet might be more appropriate. A user could search for a keyword that you did not configure in your meta description, while the body text contains the exact keyword that

the user is looking for. This is no problem. Of course, you cannot put every keyword imaginable in the little space you have. As an example: on www.joomlaseo.com, I have a page about image optimization. The meta description I configured is the following:

There are many tools that help you to reduce image size without reducing quality. Also you can use optimization techniques like sprites and base64-encoding.

If someone searches for *Joomla image optimization*, this is also the snippet shown in Google. However, if somebody specifically searches for *Joomla image optimization tinypng*, Google will take a snippet from the body text to include this word and show it in the search page, like this:

Also you can use optimization techniques like sprites and base64-encoding. ... siteground joomla hosting. joomla seo logo ... tinypng-reduce-image-size-joomla.

In this case, it is okay that Google ignores my configured meta description (partially) and adds a correct snippet from the body text.

In contrast to the <title> element, the meta description is **not** used by Google to determine how your web site is ranked in the search results (unless you use **duplicate** meta descriptions on several pages of your web site, in which case you might be penalized!). This does not mean that meta descriptions are not important: to use meta descriptions positively, you should make them attractive for visitors, so that they will click the result, even if it is lower on the SERP. Use them as a mini ad for your site. I will discuss this in more detail later.

Keyword Analysis

I already mentioned that Google is less aimed at **exact keywords** these days, but more toward **topics/themes**. You should, however, make sure that the words you use to describe your theme match the words that users employ to search for it in Google. Also, a variety of keywords is better than using a limited set multiple times. This is why it is extremely important to perform keyword research.

When writing your content, you should make sure that you know the keywords that your visitors type into Google. Otherwise, you may lose a lot of potential traffic from not knowing the keywords that potential visitors are using. For example, if you write about an illness, it could be that many users would search for the Latin name of this illness. Keep in mind, however, that you should write primarily for humans, and then optimize for Google (in that order). So make sure to include keywords in a natural way. If you feel more comfortable doing so, you might start with writing your content just as you normally would and check correct use of keywords in the editing stage.

There are a lot of approaches to this. Let's research our keywords with the example of a webs hop for books in mind, so the main keyword will obviously be *books*.

First, just brainstorm, and place yourself in the position of someone looking for books. Think of all the alternative terms that you could use to look for books, such as the following:

- Novel
- Biography
- Encyclopedia
- Etc.

Also look for combinations of words, such as

- Books online
- Books for women

Do not forget singular and plural forms of the same keyword. Especially if the list grows, you should organize this. Use a spreadsheet or another tool you like.

With the list of keywords you now have, go to google.com and search for the word *books*, or any other keyword you have come up with, and scroll to the bottom of the page with results. At the bottom of the page, Google will show you further suggestions for phrases to search for (Figure 1-4), and these will probably be much used ones.

Searches related to books

kindle books	best selling books
books to read online	book recommendations
good books	books a million
barnes and noble	ebooks



Figure 1-4. Suggestions from Google at the bottom of the SERP

Long-Tail Keywords and Local SEO

It is always tempting to aim for the top position for popular keywords. If you manage to get to the top, your visitor numbers could be huge. The problem is, however, that your competition will probably also be huge. Getting a number-one position for a keyword such as *cars* is almost impossible, considering all the competing large companies with massive SEO budgets. Instead, you should be looking for more specific terms, more related to the specific audience, such as *blue vintage cars*, and so forth. You could even optimize for exact keyword combinations that are very rarely searched but result in high conversions. Someone searching that specifically probably knows exactly what he/she is

after, and if your site returns as a result, chances are very high that the searcher will like your site. This is what is called optimizing for **long-tail keywords**. The nice thing about sites that have been optimized really well is that these long-tail keywords pretty much emerge by themselves. Sometimes, proper sites can be found in the SERPs for thousands of keyword combinations, some of which you might not have known people searched for. The tools discussed in the next paragraph will help you to search for those.

Another way to optimize for specific situations is **local SEO**. If you are selling cars in, say, Timbuktu, you do not need to show up in searches for cars worldwide. Make sure you optimize for your local city or area. You can do this by using your city here and there in your body text or headers (prevent over-optimization), but also by emphasizing your address in a footer module repeated on every page (especially if you mark up the address with **microdata**).

It is well worth the effort to do some further reading on the subject. The relevant chapter from Moz's SEO beginner's guide is a very helpful start for that: www.moz.com/beginners-guide-to-seo/keyword-research (the whole guide is very interesting as well).

Tools for Keyword Research

Apart from common sense and brainstorming, it can often pay off to use a tool for keyword research. A very basic but useful tool I access is the web site at www.seorch.eu. Under the menu Google Tools, it has the **Suggest** option, which will give you many additional options. You won't receive metrics such as search volumes or popularity, but it might help you to build a list of keywords. There are also more advanced tools, however. Let's start with some free ones.

Google AdWords Keyword Planner

Google AdWords is a tool aimed at advertisers whose ads you often see alongside the organic search results. Not that we want to do any advertising for now, but hidden in Google AdWords is a free tool that is very helpful in getting further details on your keywords: the **Keyword Planner**.

You can access the tool without a Google account (with endless CAPTCHAs), but without a Google account, you might as well stop reading this book, so go to www.adwords.google.com, sign up, and navigate to **Tools ► Keywordplanner**.

Click **Search for new keyword and ad group ideas**. You can also search for keyword suggestions for a URL (your competitors!) or a product category, but the most basic option is the list of keywords that you have. Also, you can set some filters, such as a country, if you are only targeting a specific country or language. You can fill in multiple keywords, but let's first just use a few. If we enter a few keywords related to *books* (use filters if needed), the tool might give us a list like the one shown in Figure 1-5.

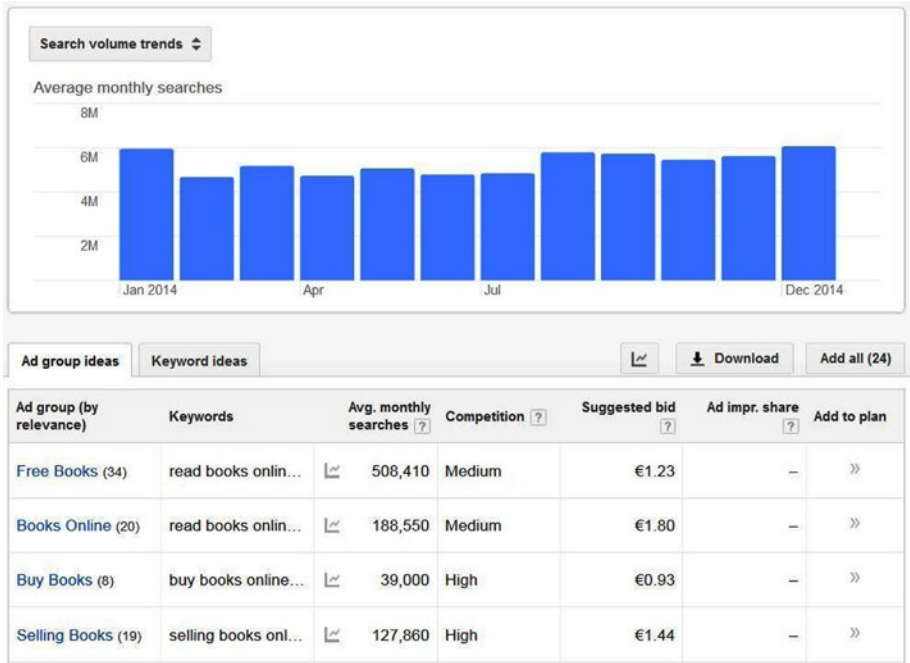


Figure 1-5. Keyword group in Google AdWords, including volumes

You first get an overview with search volumes for your combined set of keywords, which isn't that useful, but below that you will be provided with a list of **Ad group ideas**. Just choose one that best describes your audience (you can repeat this for more groups) and click it (Figure 1-6).

Ad group: **Books Online** 2 of 24 ad group ideas

← → Add all (0)

Keyword (by relevance)	Avg. monthly searches	Competition	Suggested bid	Ad impr. share	Add to plan
read books online	60,500	Low	€1.10	–	»
books to read online	18,100	Medium	€0.72	–	»
second hand books online	4,400	High	€0.42	–	»
reading books online	8,100	Medium	€1.42	–	»
books online to read	880	Medium	€1.18	–	»
kids books online	9,900	Medium	€1.02	–	»
discount books online	720	High	€0.81	–	»
order books online	1,900	Medium	€0.94	–	»

Figure 1-6. Keywords in an AdWords ad group

You will be provided with a list of keywords related to this ad group. Check for relevant keywords, preferably looking for **high search volume** and **low competition**. You can order the rows by clicking the column headers, and you can also apply filters (maybe only look for low competition). Look, too, for long-tail keywords—keywords of low volume but extreme relevance, which might possibly lead to high conversion.

Make sure you keep the process organized, and use a spreadsheet tool if necessary.

By the way, you might also want to check the web sites of your competitors. You can also use the AdWords tool to retrieve keywords from any web site on the Web.

This is just a brief introduction to how you can use the tool. If you spend some more time with it, you can obtain a wealth of information here.

WebCEO Keyword Research Tool

Another free tool that can be accessed if you register for an account is available at www.webceo.com. It requires that you sign up for an account, but you get a lot of useful information in return, so I advise you to take a shot.

Once you have registered, you must first create a project by submitting a URL for your web site. Please do so. In the process, you can also connect to your **Google Analytics** account and provide some additional information about keywords and competitors, but most of that is not required and not needed for the step we perform now (it can be very useful later, though, as I will explain in Chapter 16).

From the project navigation, click **Keyword Research**. This will give you a few options to choose from, as shown in Figure 1-7.

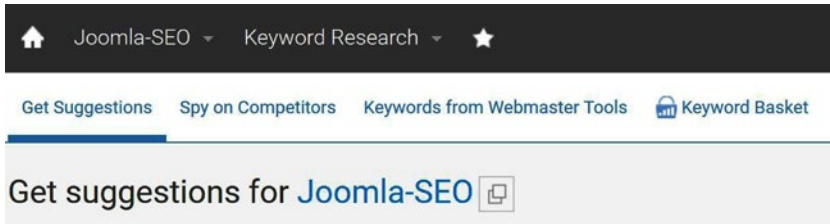


Figure 1-7. Keyword research options for WebCEO

The following options are available:

- **Get Suggestions:** This is the one I will explain further.
- **Spy on Competitors:** Spying is a very useful SEO practice! Unfortunately, it is only available to paid subscribers.
- **Keywords from Webmaster Tools:** If you connect your Google Webmaster Tools account, you can import keywords from it.
- **Keyword Basket:** These are some options for your top keywords.

The **Get Suggestions** feature is an especially handy tool, as it allows you to retrieve useful suggestions quite easily. First, enter one (or more) of your earlier collected keywords in the **Suggest** field. A list of related keywords will then be shown, as you can see in Figure 1-8, in which I searched for suggestions for the set of keywords *bookstore*, *books online*, and *bookshop*.

Keyword	Global searches	Bid competition	Search Trends	KEI
books online	60,500			4.56
bookshop	18,100			19.58
bookstore	74,000			29.32
read books online	60,500			23.05
book online	22,200			19.43
free books online	110,000			10.49
reading books online	8,100			4.42

Figure 1-8. Keyword Suggest tool results

You see a number of interesting columns listed. **Global searches** (Local searches can be added), **Bid competition**, and **Search Trends** pretty much speak for themselves, but the **KEI** is a crucial one here. KEI stands for **Keyword Effectiveness Index**, and the lower it is, the more interesting these keywords are to optimize for. A low KEI means that there are many global searches with low competition, while a high KEI indicates the opposite number of counts (maximum value is 100). So, in the example in Figure 1-8, *reading books online* seems to have a much more interesting value for us than the combination *read books online*. Of course, you should diversify your keywords, so this is not a perfect example, but it shows the opportunities. Just play around with your own keywords.

There are plenty of other tools available, but most of those are part of paid packages, so for a beginning SEO agent, they are probably too costly.

Use Your Keywords

Assuming your web site is in place with the right template and layout, we can now start filling the web site with data, i.e., articles, modules, and menus. At this stage, we will have to think about how and where we start filling the web site with text and keywords.

Generally speaking, try to use your keywords as much as possible, in all relevant areas of the site, without **stuffing** the keyword. Of course, the article text is what you will have to pay most attention to, but other relevant areas could include the following:

- Domain
- URL
- Meta title
- Meta description
- H1-heading and in further H2 and H3 tags
- Images (name and alt-text)
- Links

Most of these bulleted items will be discussed in detail in separate chapters later in the book.

As far as article text is concerned, which is where most of your content will go, focus on your readers, but without forgetting about how Google would like you to write.

Although not a complete guide in SEO writing, the following few rules of thumb will be helpful to keep in mind:

Use your main keyword **early in your sentences and paragraphs**, if readability does not suffer. For example, compare the following two sentences:

- After hours of testing, we found out that bicycle X was the best.
- Bicycle X was the best one, as we found after hours of testing.

If *bicycle* is our main keyword, the second example would work best, while readability is not affected.

Something else you need to pay attention to is the percentage of a single keyword used in your text, the **keyword density**. There is no exact ratio that you should stick to, but you should definitely not go far above 5% with this, meaning five occurrences of a single keyword within 100 words. Also, play with **synonyms** and **abbreviations**. Using alternatives for the same keyword allows you to play with the keyword density a bit. Using synonyms also makes you target related alternatives in Google.

There are several online tools available to check the ratio of the main keywords. A tool I often use is www.seorch.eu, which also gives you the keyword density for the most used words of your page.

Apart from the basic body text, also make sure to build your **html headings** correctly. Google has to understand the hierarchy of your text, and you must use your headlines to do so. In HTML, these are your `<h1>`, `<h2>`, `<h3>` tags, etc. I will go into more detail on the H1 issue later, but you will have to make sure that you always use an `<h1>` on every page, but never more than one! Google sees the `<h1>` as one of the most important SEO factors on a site, so use it correctly.

Google SEO Starter Guide

Everything discussed so far is basically what Google suggests to web masters building web sites. As long as you follow Google's guides, you should be fine. If for any aspect of SEO you wonder whether it is really worthwhile, just check what Google says. If it says do it, do it! You want to be in its rankings, so play according to its rules. Google also created a pdf called the **Google Search Engine Optimization Starter Guide**, which you can find online at the following URL: www.google.com/en/webmasters/docs/search-engine-optimization-starter-guide.pdf.

It is not very difficult to read, and if you follow it, miracles will be worked for your efforts.

CHAPTER 2



Preparing Your Joomla Environment

There are some considerations before actually jumping into the search engine optimization (SEO) settings of your Joomla installation. These are issues such as hosting, domain, URL, etc., so I will discuss these first.

Hosting

Hosting is extremely important for your site. At least make sure that you are on a reliable host. Preferably, look for hosts that specifically optimize for Joomla and for performance and security. In Chapter 18 on performance, I will pay closer attention to what you should look for in a good host.

Hosts that often have issues such as downtime, problems regarding file permissions, backdated PHP versions, or hacked sites can be bad for your site, as the IPs close to yours may get blacklisted, and you may suffer too, especially if you share the same server as a hacked site. Also, installation issues can sometimes occur.

Once you have chosen your host, also consider which package to choose. Sometimes, the upgraded packages offer advantages in performance, but SEO can also be a factor.

Having a dedicated IP address and being on a dedicated server can make you less vulnerable to other bad/blacklisted sites hosted near your environment. Having sites such as these sharing your IP number or server could also harm your SEO.

In addition, having an SSL certificate might be a good idea. While traditionally it is mainly used to secure payments from web shops, more and more sites are fully SSL-secured. This is also due to Google having announced that it prefers sites that fully run on SSL connections. Note that SSL can sometimes cost you some performance and is a bit more complex to install.

Domain/URL

Choosing a good domain is very important. The URL is a very important SEO parameter, and it is used by Google as a ranking factor. This is why **exact-match domains**—domains that contained the exact focus keywords for your business—used to be popular.

An example is www.joomlaseo.com, which contains both the terms *joomla* and *seo*. While this could still be a good idea for some sites, it is not really necessary anymore.

Also, Google may look just a bit closer to your site to assess if its content is really reliable and matches the domain. This is because this technique has been abused many times in the past.

Branding is becoming more important these days, so opting for a good brand as the domain is more the way to go, in many cases. You can compete perfectly by using other areas of your site to use your keywords. An example is the company Moz. A few years ago, it changed its domain from SeoMoz.com to Moz.com. This helped build the brand, while it threw away its main keyword from its URL!

Some other points to consider:

- **Keep it short** to allow people to simply remember your URL. Many sites have about 10% of users who simply type in the URL.
- Choose a reliable extension (TLD or top-level domain), such as **.com**, **.net**, or a country-specific extension. Avoid extensions such as **.guru**, **.info**, etc. Also, TLDs that recently have become available (such as **.ninja**, etc.) should be avoided for the time being, until they have proven themselves.
- It can sometimes pay off to buy an already existing domain/site. These sites may have a long-standing reputation, and you may easily benefit from this, from an SEO perspective. You can acquire domains though www.flippa.com.

Joomla Version

Make sure you use Joomla 3. Joomla 2.5 and earlier versions are no longer supported. If you still use them, make sure you upgrade. They may soon contain security risks that will no longer be patched, whereas Joomla 3 contains many features that are beneficial for SEO.

Also, make sure you always keep your current Joomla version up to date, as regards all your extensions. This avoids getting hacked. Hacked sites may get blacklisted by search engines, which could make all your SEO efforts useless. Some hosts offer auto-updating features that may help you to keep up to date.

Start on a Test Environment

If you are going to work on a new site, make sure you create and optimize the web site on a **localhost** environment on your own PC. Preferably, you should not be working on the live domain on which your site is going to be (not even on a subdomain). First of all, it looks amateurish to anyone passing by, but you will also create URLs while you are building the site, and Google may already start indexing these URLs. However, it is very likely that these URLs are going to change while you are developing, meaning that Google will encounter errors for changed URLs, and you may be penalized for this very early, maybe even before the site is finished.

Of course you can set the site to an offline status, but it is best to work locally—on your computer. Apart from working on the SEO aspects, you will also be working on your template, images, overrides, etc., and all this is much easier if you just work on your PC or Mac. (Advanced users can block access to all but their own IP in `.htaccess`.)

Creating an environment to run your web site locally is easy. There are several options, but XAMPP, which can be obtained from www.apachefriends.org (Figure 2-1), is used most frequently.



Figure 2-1. Download and installation of XAMPP

XAMPP installs a set of software including Apache, MySQL, PHP, and PERL, which is all you need to run your local web server. There are versions for Windows, Mac, and Linux.

Depending on your operating system, you can simply place your web site in the destined folder and access it with your browser.

Template

Optimizing your web site for search engines starts at a very early stage. Actually, you should be keeping this in mind when you are building your template. The reason is that Google likes structured content, and structure starts with the way the template is set up.

Custom or Template Provider?

Whether you create your own template or use one from a template provider depends on your own skills. If you are technically proficient, go ahead and build your own templates, as long as you know what you're doing. Write clean HTML and keep in mind new features such as microdata.

Commercial templates can be as good as your own. Designs are often very attractive, though some say that, due to the included template frameworks, these templates are heavy and load slowly. If you stick to well-known template providers, this should not be an issue. Well-trusted template providers include

- YOOtheme
- JoomlaArt
- RocketTheme
- Gavick
- JoomlaBamboo
- JoomlaShack
- JoomlaShine

The companies behind these providers are trustworthy, and their teams are technically skilled. Also check whether they are quick to update their templates when new features arrive in Joomla. Often, commercial templates contain many overrides for article display, and when something in the Joomla core article display is changed, these overrides may require updating too.

Performance is also something to look at. Slow templates give users a bad experience and may chase them away. More on this is included in Chapter 18.

Preferably, stay away from free templates. Quite often they are available on scammy download sites or bloated with hidden backlinks. They probably are more harmful than good to your SEO efforts. Only if you really trust the source of the free template should you probably go ahead and use it. This could be the case with the commercial template providers I mentioned previously. These providers usually offer two or three templates for free, and these should be fine to use.

By the way, the templates that come with Joomla 3 are pretty good, if the design suits you. The Protostar template is especially good. It will always include the latest Joomla updates with regard to microdata and other SEO improvements, and it is responsive to mobile usage.

Clean and Valid Code

As always, your code should be valid HTML, so make sure you run the code through the W3C validator check at www.validator.w3.org. By the way, don't panic if your code is not 100% valid, but try to get as close to 100% validity as possible. I use the Webdeveloper Toolbar in Firefox for this (Figure 2-2). You can download it from www.addons.mozilla.org/nl/firefox/addon/web-developer. This gives you a constant little view in your browser when testing. It tells you if your code is valid or not, and it also checks for CSS and JavaScript errors.



Figure 2-2. Webdeveloper Toolbar showing JavaScript errors

Also, keep your HTML clean and readable, by avoiding inline JavaScript and CSS as much as possible. This simply keeps the **text/code ratio** as high as possible. Remember, SEO is a matter of dozens of little steps, and even the smallest step might just give you the edge over your competition.

HTML5

Using HTML5 gives you even more opportunities to create hierarchy in your code. Using tags such as `<section>`, `<aside>`, `<article>`, `<footer>`, etc., makes it even easier to distinguish the crucial parts of a page just by looking at the code. HTML5 was finally completed in October 2014 and provides a much cleaner and clearer way for HTML to be constructed. The previous HTML standard was completed in 1997 and contains few of the features and tags that allow a page to be optimized for SEO.

Usually, an `<article>` will receive greater SEO value than an `<aside>` or `<footer>` that is shown on every page of your site.

Order of Code

If possible, build your template in such a way that the most important parts of the page show up first in your code. This will not always be easy, but it can be done. The most important information on any page will usually be in the **Component** position (usually the actual article), not in the surrounding modules. So, you should aim to put the component as early in your code as possible.

If you put your main component before the sidebars, this would be ideal. Some template providers use this technique in their templates. The templates from JoomlaShine are an example. In the raw code, they place the main component before the sidebars. So, where usually the order will be

1. Left sidebar module
2. Component
3. Right sidebar module

JoomlaShine manages to put the component first in the code, followed by the two sidebars. Then it cleverly uses CSS positioning, so that, visually, the left sidebar comes first, while in the HTML, the component comes first.

This is pretty advanced stuff, and you really do not have to follow it, but remember, it might just be this little step that you need to make a big difference.

Responsive Design

Make sure your template has a responsive design. The share of users watching your site on mobile devices is still growing. Some sites already have 50% of their audience on mobile devices. In addition, Google seems to move in a direction in which mobile design gets an advantage in the rankings. It actually announced that mobile design will be a ranking factor as of April 21, 2015.

Also, if you are starting with a new site, there are plenty of responsive templates on the market. There is not a single template provider that does not create responsive templates. If you have an older site, however, you might think of upgrading it soon.

So, let's assume we have a clean installation of Joomla with a template optimized for SEO and performance. We can now start to configure Joomla. In the next chapter, let's go over the settings one by one.

CHAPTER 3



Joomla Global Configuration

As soon as you have installed Joomla, your first action will usually be to go to the **Global Configuration** settings. These settings are extremely important, not only for search engine optimization (SEO), but also for other areas of your configuration, such as setting the editor, e-mail settings, performance settings, cache, etc. If you optimize all the parameters in this menu, you will have a great start to building a well-optimized Joomla site in every aspect. I will first go through the fields that are important for SEO. In Chapter 18 on performance, we will change some additional settings.

Global Configuration: Site Name

For SEO, the most important parameters are all located on the first tab that you see, the **Site** tab. The other tabs are also important, but less so for SEO. I will, however, review them in the chapters about performance.

Back to the Site tab. The very first parameter we see is **Site Name**, under **Site Settings** (Figure 3-1).

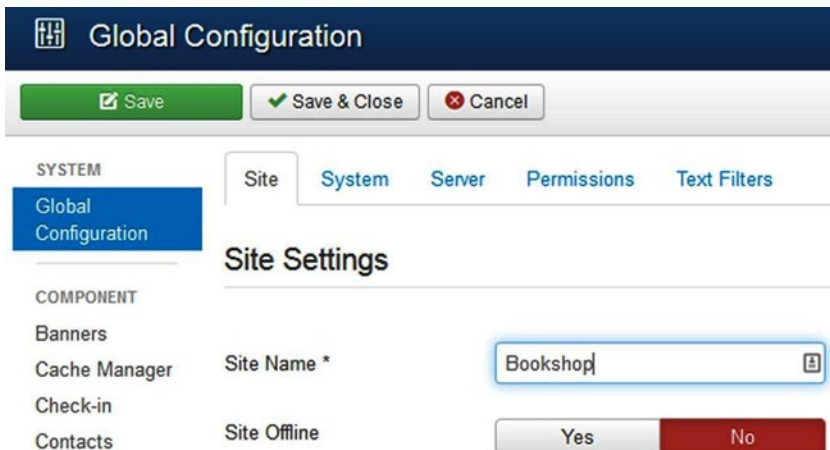


Figure 3-1. Site Name

Site Name is a mandatory setting that you probably entered during the installation of the site, and you can review it now. Keep it short, because it is used in many places, depending on your configuration (sometimes in the logo position or in your page titles). Later on, we will do further configuration and probably have to append things to the site name. The Site Name is often (but not always) used in two major locations:

- Occasionally, in the **logo position** (depending on your template)
- More frequently, as part of the **page title** (more on that later)

If your web site is for a **brand**, such as Coca-Cola or Joomla, the site name will probably be the actual brand name, but it can be anything. For example, for www.joomlaseo.com, the Site Name is simply Joomla SEO, which contains the two keywords it most wants to rank high for.

Global Configuration: SEO Settings

The **SEO Settings** section in **Global Configuration** (Figure 3-2) is extremely important. Setting these correctly before uploading your web site to the live server is crucial. Changes in these settings will usually change your URLs! Every time you perform an action in which URLs are changed for a live site, alarm bells should go off, as this could possibly lead to Google encountering removed URLs, which might very well lead to penalties (which results in a lowered ranking, either for this specific page or, with many errors, the whole site). Therefore, you should always plan these actions (more on this later). Preferably, set these settings correctly from the start, and never touch them again!

SEO Settings

Search Engine Friendly URLs	<input checked="" type="radio"/> Yes	<input type="radio"/> No
Use URL rewriting	<input checked="" type="radio"/> Yes	<input type="radio"/> No
Adds Suffix to URL	<input type="radio"/> Yes	<input checked="" type="radio"/> No
Unicode Aliases	<input type="radio"/> Yes	<input checked="" type="radio"/> No
Include Site Name in Page Titles	<input type="text" value="No"/>	

Figure 3-2. SEO Settings in Global Configuration

Let's walk through the parameters in these settings and discuss them one by one.

Search Engine Friendly URLs

SEF URLs, as the term *Search Engine Friendly URLs* is usually abbreviated, are switched on by default in Joomla versions 2.5 and higher. You should leave this untouched, but it is useful to recognize what this parameter does, as it will help you to understand more in demanding chapters later in the book.

So, take a minute to switch SEF URLs off (Figure 3-2), and check your site on the front end. Make sure your site has at least two articles and two menu items; otherwise, there is nothing to be demonstrated.

Check how the URLs in the navigation bar look now. For example, a single article can be displayed with a URL such as this:

```
index.php?option=com_content&view=article&id=4
```

The number 4 at the end of the URL corresponds to the **ID of the article** in the database. If you navigate through the menu, you will often see, too, an **Item id** in the URL referring to the menu item ID in the database, as in the following:

```
index.php?option=com_content&view=article&id=4&Itemid=110
```

Note that the same article is already displayed here through two URLs (remember this, as there will be much, much more about it later).

For the purpose of troubleshooting, it can sometimes be useful to view URLs such as this, as the component responsible for this URL is now visible—in this case, `com_content`.

Apart from showing you this unusual example, those strange URLs make no sense, of course, so we are going to set the value for SEF URLs back to Yes (Figure 3-2). Users can now much better understand what the URL is about, and so does Google.

Suppose we now have a menu item called Portfolio. With SEF URLs switched on again, it will now have a URL such as the following, which is much cleaner:

```
index.php/portfolio
```

What you should keep in mind, though, is that the strange non-SEF links are *not gone!* If you remember one of these URLs, and type it in your browser's address bar, you will still see the same page as previously. This means that we now have found three URLs for the same article.... It's nothing that can't be fixed, but we will need extensions for this, to be discussed later.

Use URL Rewriting

URL rewriting (Figure 3-2) sounds complicated, but it isn't. The purpose of this parameter, for now, is to get rid of the `index.php` bit that is still displayed in our URLs. Imagine that we have a portfolio page that we created by linking a menu item called Portfolio to a list of portfolio items, the URL would be like this:

```
/index.php/portfolio
```

`index.php` does not tell us or Google anything, so we should get rid of it. To do so, we must to perform two actions at the same time. Note that your host should support URL rewriting, but if it doesn't, move immediately to another host. This should really be a standard option.

- First, we need to set **Use URL rewriting** to **Yes** and save the page.
- Second, go to the root of the Joomla installation. You will find a bunch of directories and single files there. Now it depends a bit on which file we have to look at, but for the majority on Apache web servers or local XAMPP installations, this will be the `htaccess.txt` file. Now rename this file to `.htaccess`. (For IIS servers, rename `web.config.txt` to `web.config`.)

Windows users cannot perform the renaming in the Explorer window. To achieve this, go to a **Command Prompt** and browse the folder that holds the `htaccess.txt`, using the `cd` command, and issue the following command:

```
rename htaccess.txt .htaccess
```

You can also open up `htaccess.txt` in an editor such as **Notepad** and save it as `.htaccess`. I often hear that people find it scary to do this step, but I strongly advise you to do so now, in the early stages of your site development!

If we now navigate to the **Portfolio** page, the `index.php` part is gone. Next to that, we can now make full use of the Apache URL rewriting functionality, which gives us many advanced and useful options that we will use later.

Again, `/index.php/portfolio` is not gone. It is still accessible, giving us yet one more URL leading to the same content!

Adds Suffix to URL

Leave the default for parameter **Adds Suffix to URL** (Figure 3-2) set to **No**. It can add to the URL suffixes such as `.php` or `.html`. Adding non-SEF characters Adds Suffix to URL to URLs is really useless, and this is what this setting is for. Note that, historically, `.html` suffixes were considered good for SEO, but that is no longer the case.

Unicode Aliases

Leave the default for **Unicode Aliases** set to **No** (Figure 3-2). Most sites should not ever touch this. The setting allows for **utf8** glyphs in the URL, which may be necessary for non-Latin characters in the URL, in rare cases, or for non-Western languages.

Include Site Name in Page Titles

Okay, time for some considerations again. The most important consideration for this parameter is: **How do we want to display our page titles?** We are talking about the title as it is displayed in the Google results page (and also in your browser's tabs): the `<title>` element.

Together with the meta description, the page title defines how our page is shown to anyone searching for information on Google. This title is the main headline showing up in Google, so it has a great influence on whether someone clicks the result for your site or not. If you set it up correctly, the **CTR** (Click Through Rate) in the SERP will go up. Again, remember to use the keyword early in the title, and make sure to limit this to 65 characters, to prevent breaking it off.

Before we set this option, you have to know how page titles are generated in Joomla.

- They are taken from the **menu alias**, if an article is **directly linked to a menu**, or
- They are taken from the **actual article** if the article is **not directly linked to a menu**, such as in blog or list views.

There are more possibilities, but these are the main ones. Note that titles can usually be overridden at some level.

So, when creating an article, always keep in mind the length of the title. If you don't override it at some level, the title of the article will be prepended or appended to the site name (as defined in Global Configuration). The number of characters for both should not add up to more than 65.

Okay, so what are the options for this setting now? The **Include Site Name in Page Titles** option (Figure 3-2) allows you to use the site name that we just configured in the page titles. All three options could be fine for your site, so I'll discuss them one by one:

- **Before:** If your Site Name contains your main keywords already, it is best to use the Before setting. This means that every page on your web site will contain the main keywords, possibly giving you some more freedom to set up article or menu titles. For example, on `Joomlaseo.com`, the site name is Joomla SEO. This means that if I write an article about **Canonicalization** (also with this article title), the page title will be set as **Joomla SEO - Canonicalization** (the dash is placed automatically). If I did not have the site name prepended, I would have to put Joomla SEO in the title of every article, which is not very convenient for readers.
- **After:** This is more useful if your site name is a brand name such as, say, Coca-Cola. You want people to know that they will be accessing a Coca-Cola site when they see it in the SERP, but it may not actually contain the main keyword (say, soda).
- **No:** There could be many reasons not to use the site name. It gives you more freedom to use the 65 characters available for a more unique page title.

These are just guidelines, so feel free to act independently, as long as you realize the goal: to include your keywords in your page titles in a natural way, with or without your site name included.

Finally, page titles can be overridden on many levels, so you are not forced to stick to the site name-article title format.

Global Configuration: Metadata Settings

In the introductory chapter on Google, I explained how important metadata are. Now, in the fields in your Global Configuration, under **Metadata Settings** (Figure 3-3), you have the opportunity to fill in something appropriate. If you do, this will ensure that no page ever shows empty metadata in the source, even if you forget to add it later for a specific article. For now, leave it empty. In Chapter 4, I will explain the reasoning behind this choice.

Site Meta Description

First, we find **Site Meta Description** (Figure 3-3). I will discuss the purpose of meta descriptions and how to configure them in much more detail later. For now, leave the field empty.

Metadata Settings

The screenshot shows the Joomla! Global Configuration interface for Metadata Settings. It features three main sections: 'Site Meta Description' with a text area containing 'Meta description here'; 'Site Meta Keywords' with a text area containing 'leave empty|'; and 'Robots' with a dropdown menu set to 'Index, Follow'. The 'Site Meta Keywords' field is highlighted with a blue glow.

Figure 3-3. Metadata Settings in Global Configuration

Site Meta Keywords

Google pretty much ignores meta keywords, as they have been abused too much in the past. Some say they could even be harmful to your SEO, but this is only true if you abuse them. I would advise you to leave them out completely, but it won't hurt to use two, three, or four keywords in the Global Configuration, under **Site Meta Keywords**, and only there. Do not spend any effort in adding them to menu items or articles.

Robots

The default setting for this is **INDEX, FOLLOW**, meaning that all pages must be indexed by the search engine's robots, and the links have to be followed. Leave this on the default setting, which is exactly what you want.

One reason for deviating from this could be a situation in which, say, a copy of a production site is restored on another domain for testing purposes, and you want to avoid Google encountering two identical sites, which could lead to duplicate content penalties.

Another reason could regard a site that is under construction. As I mentioned earlier, it is best to work entirely in a local environment, but sometimes this is not feasible, such as in the case of a team having to work on the same project, and all its members must be able to access the project site from multiple locations.

In these two situations, you could consider setting the parameter to **NOINDEX, NOFOLLOW**, which would tell Google to completely ignore this site.

■ **Tip** If you change this now, **change it back when the site goes live!** It has happened to many projects before.... You spend a lot of time working on SEO and then forget this one little parameter, and all your efforts are wasted, at least until you correct this.

The Robots parameter is also available in article and menu items. For specific pages, you might consider setting it to something alternative, such as **NOINDEX, FOLLOW**. This could be useful for pages that do not contain relevant information, such as the login page for your site. I will discuss this later.

Further parameters under Metadata Settings are not relevant for SEO, so they will not be discussed here.

Joomla Search

This paragraph is a bit of an odd one. It is not directly related to setting up your SEO configuration; however, setting up the **Search component** will allow you to gather very useful information from your users. Usually, you need tools such as Google Analytics to see what people are doing on your Joomla site and where they go. Using the Search component, you will be able to collect some very useful information.

First of all, the component allows your users to search the content of your web site, but you can also set it to collect the information that your users are looking for. Knowing this might reveal some questions that your users expect to find answers to on your site, also allowing you to add this information, if it isn't there yet.

So, how does this work? First of all, you have to switch on the gathering of these statistics. The reason it is not switched on by default is that, sometimes, when many search queries are performed, this can weigh on the database. As long as you keep an eye on this, it shouldn't hurt, and you can also choose to monitor this behavior only temporarily, until you've learned enough. To switch it on, go to **Components ► Search** and click the Options button at the top-right corner (Figure 3-4).

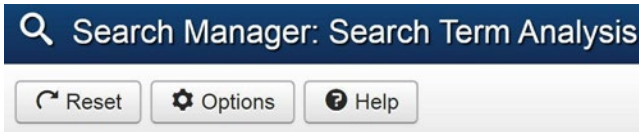


Figure 3-4. Search component Options button

Then set **Gather Search Statistics** to **Yes**. Of course, do not forget to publish a **Search module** as well; otherwise, your users cannot search at all. From that moment on, all search queries will be logged, and you will see a screen like the one in Figure 3-5.

 A screenshot of the Joomla! Search Manager interface showing a table of logged search queries. The table has three columns: "Search Phrase", "Hits", and "Results". The data is as follows:

Search Phrase	Hits [▲]	Results
content delivery network	1	Off
cdn	1	Off
cloud flare	1	Off
.ht	1	Off

Figure 3-5. Logged search queries from users in the Search component

And, of course, use this information. Do not just let it sit there! If you find that users are searching information on your site with synonyms that you did not think of, include them in your keyword list and optimize pages for this. You might even find that users search for information that is not yet on your web site. For example, if you find that users search for information about the nutritional value of the chocolate bars that you sell, create an article about this. It will satisfy not only the users you already have, but it may also result in other users coming from Google, because Google finds that you are supplying helpful answers.

CHAPTER 4



Metadata

I already mentioned that on the Global Configuration screen, you should not enter a meta description. The reason is that if we set it there and do not set any further meta fields for specific articles or menus, all pages of our Joomla site will have the exact same meta description, meaning duplicate meta descriptions, and Google really doesn't like these.

You can imagine why. Actually, you are telling Google that all these pages are about exactly the same subject, and this is not very helpful to Google. Also, it is not useful for anyone reading this meta description as the summary of your page in the SERPs (search engine results pages). Most probably, Google will ignore your meta description for many pages and simply make up its own. I cannot stress enough how important it is to optimize individual meta descriptions and titles for all your pages. This really helps all these individual pages to receive good rankings in the SERPs, not just your home page. Most sites that I see that are badly optimized lack these.

This is why I advise you **to leave the meta description empty in the Global Configuration!** This may be contrary to advice you heard elsewhere, but Google would rather have you leave it blank, as opposed to having the same description duplicated all over the place. From the following link, check this YouTube video by the Google Webmaster folks for some background on their view: www.youtube.com/watch?v=W4gr88oHb-k.

Remember that the global settings are for the entire site. If on a lower level any of these metadata settings is set (in an article or menu item), it will override the global setting, so if you plan on providing meta descriptions on **all** lower levels too, feel free to fill in a global meta description. It will be overridden, in any case.

Metadata Setup per Article or per Menu Item

This is a good moment to reflect on further setup of the site. Many configuration options depend on how you link to your articles. The main difference is between the following two options:

- Will you link every article to **its own menu item**? This will often be the case on small brochure-like sites. For example, an article entitled **Contact our business** will be attached to a menu item called **Contact**.
- On larger sites, many articles will not be directly linked to a menu item one by one but will be made available in a **blog or list view** or as **featured** articles.

Many choices we will make are dependent on this, so if the desired setup is already clear to you now, you will know which path to follow later in this book.

For now, also start to think about how your site has to be set up, as far as detailed metadata is concerned. You can set your metadata on one or two levels.

- On an article level
- On a menu item level
- On both of the preceding

There is a hierarchy in what level “wins.” **Menu item** settings override **Article** settings, which override the **Global Configuration** settings. As long as you ensure that every URL of your site has the correct meta description, it doesn’t matter how you achieve this from a search engine optimization (SEO) perspective, but it pays to take the most efficient route, which is usually to fill in metadata per article, and not in the menu items, except in some specific cases. I will talk about this in more detail later.

Metadata for Every Separate Page

So, the metadata settings that you configured in Global Configuration are just a start. For most of these settings (e-mail, cache, robots, etc.) one option globally is fine, but the meta description should be overwritten as much as possible on an article or menu level.

Meta Description

This is what you really have to override as much as possible. If you don’t set separate meta descriptions in your pages, they will all look the same in the SERP, or, more likely, Google will simply take a snippet from your page and show it, usually ending with three dots, because of broken-off sentences. Also, you might not end up very high in the rankings, while your click through rate (CTR) will be relatively low.

Also, make sure you set a meta description on every single page and make a concerted effort to make the description readable and attractive, as this really helps your individual pages to stand out. Compare the two SERP snippets in Figure 4-1. Both are found among the top ten for the keyword *bookstore*. Even if they are listed like this, a fair amount of users is likely to click the second item. (We want books; we don’t want to be included on a mailing list.) It probably works much better as a call to action, resulting in a higher CTR. You just have to take some time to write a nice meta description. Sometimes, SEO is also a marketing skill!

University of Calgary Bookstore: Welcome

calgarybookstore.ca/ ▼

Store Announcements. Sign up for our mailing list and know when all the sales are happening. Buyback is open Wednesdays. Check if we are buying your book ...

Stanford University Bookstore Apparel, Merchandise, & Gifts

www.stanford.bkstr.com/ ▼

The Stanford University **Bookstore** is your source for team apparel, gifts and accessories for women, men and kids. Buy online, free in-store pickup.

Figure 4-1. Bad and good examples of meta descriptions

So, the goal is to set up a unique meta description on every URL of your Joomla site, but how you should actually implement this depends on many factors. It can be achieved in multiple ways. One method is not better than any other, as long as you arrive at unique descriptions. This is true for simple Joomla content, i.e., articles for which you can set the meta description on three levels: globally, per menu, per article. For all metadata settings, menu settings always override article settings, and both override Global Configuration. So, for unique meta descriptions, you can choose to set them up as follows:

- **Per Menu:** In this case, leave the article meta description empty, to avoid doing things twice. This should work fine for simple sites, where every article is linked to its own menu item.
- **Per Article:** The option to work with menu settings could work fine, but any article not directly linked to a menu item will be missed and will have an empty meta description. This could be a separate article, but it also counts for articles that are linked to a menu item within a blog or list view. Any article that is missed will display the meta description shown in the Global Configuration, which should be avoided. Therefore, setting meta descriptions in the article is the best way to go. Leave the metadata settings in the menu items empty.

Metadata: Article

To actually set your meta description (and robots setting, see Chapter 12), go to the **Publishing tab** of your article (Figure 4-2).



Figure 4-2. Article metadata

I advise you to make a habit to enter the meta description immediately when you create the article, as it can be easy to forget it. (Remember: Make it unique and compelling, by using keywords and no more than 160 characters.) You can always add it later, if you prefer.

In case you prefer to set your keywords as well, this follows the exact same logic as the meta description, but best practice is simply to save yourself the effort and leave the Keywords field empty.

Metadata: Menu Item

For menu items, the meta description is simply available from a separate tab called Metadata (Figure 4-3).

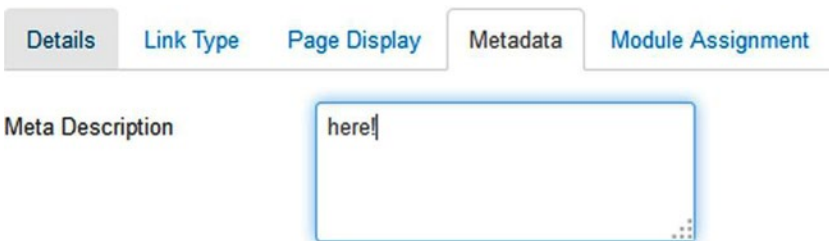


Figure 4-3. Menu item Metadata tab

I advised you to set metadata on an article level, so if this menu item is simply a link to an individual article that already has a meta description, do not enter it here again.

However, there are also views that cannot receive a meta description from individual articles. If your menu is a link to a **Featured articles** view, or a **weblinks** list, etc., the actual URL on which this overview is shown does not have a meta description yet. For pages such as this, set it in the menu item metadata settings.

If you are not sure if the URL on the front end actually has a meta description, simply check the source view of the page and look for the following string:

```
<meta name="description" content="Metadescription here....." />
```

Make sure that every page that is accessible on the front end has a unique meta description. Basically, the only exception is pages for which you have set a **Robots** value to **NOINDEX**. These pages should not appear in Google and thus do not require a summary.

Page Title

Remember: Page titles make up the headline of the result of your pages in Google, configured through the `<title>` element in your HTML source code. Note that this is definitely **not** the main headline in the text of the page (this is referred to as the Heading, discussed later). You will also see the page title displayed in the browser tab (Figure 4-4).

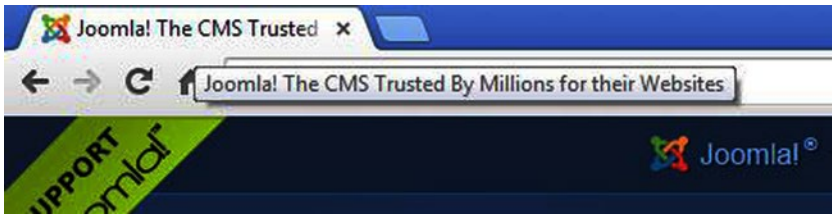


Figure 4-4. Page title shown in browser tab

Unfortunately, the way page titles are configured in Joomla is sometimes a bit hard to understand and configure. You must first understand how they are created.

Your page titles will be created for you automatically, depending on your setup. First, it depends on whether or not you have configured Global Configuration to show your **site name** before or after the title or not at all (read again, just in case). Next, the page title will be created differently, depending on the page you are at.

- For articles linked directly to a menu item, the **name of the menu item** will be taken.
- For articles shown in a blog or list view, it will be the **title of the actual article**.

Let's look at an example.

- Site name: **Kung Fu Shoes** (set up to be prepended)
- Article title: **Great for walkers and hikers in the outdoors!**

When this article would be shown in a blog view, not linked to a menu, the page title would be

Kung Fu Shoes - Great for walkers and hikers in the outdoors!

This is a perfect title. It is 61 characters and has a few nice keywords (shoes, walkers, hikers, outdoors). Now let's see what the page title will be when this article is linked directly to a menu item called **Great**. The page title will now be taken from the menu item, leading to a title like this, for the same article, as follows:

Kung Fu Shoes - Great

This is too short, and the keywords walkers, hikers, and outdoors are missed! Luckily, you can overcome this, and you really should.

- Go to the corresponding menu item, in this case, menu item Great.
- Go to the Page Display tab.
- There, you will find a field for Browser Page Title.

Now this field can be used to set a more useful page title (do not forget Site Name might be appended or prepended, if configured), as shown in Figure 4-5.

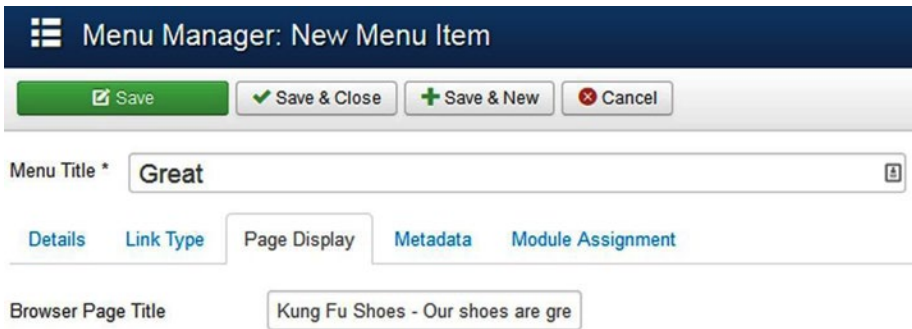


Figure 4-5. Page Display tab for alternative Browser Page Title

Use this! Remember: This action will arrange how the search result looks like in Google. Many Joomla site administrators often forget to do this. You still sometimes see web sites in the rankings of Google that still only show the word *Home* as the page title. Imagine how bad their CTR will be....

Managing Metadata

Managing metadata can be a daunting task if your site grows larger. The only way to verify correct metadata is to enter the separate articles or menus and check them. Especially if your Joomla back end is a bit slower, it can take ages to verify these.

In cases such as this, you might consider managing your metadata using an extension. There are many to be found in the **Joomla Extensions Directory**, or **JED**. I discuss some of these extensions later in the book. Particularly for beginning SEOs, two extensions are recommended: **OSMeta** and **Easy Frontend SEO (EFSEO)**.

OSMeta especially works very close to the Joomla core. It simply updates the core metadata fields and makes them available in a nice overview, allowing you to update them from both Joomla and the extension. Page titles can also be set, which cannot be done in Joomla (apart from changing the Joomla article). The basic version is free, while the PRO version allows for some advanced features (metadata for menu items and K2 support). Additional information about these extensions can be found later in this book.

Robots

Among the metadata information in individual articles and menu items, we also find the Robots parameter. In our Global Configuration, we have already set the parameter Robots to **INDEX, FOLLOW**, meaning that all pages of our web site should be indexed and followed by Google and other search engines. For the majority of the pages on your site, this will be the desired setting, and you should not have to worry about this field anymore. However, there might be exceptions.

Some pages of your site might as well not show up in the search engine indexes. Pages such as login pages or pages with your terms and conditions only serve a very specific purpose for visitors already on the site, but they will not be the pages you want to attract visitors with. Actually, these pages are not valuable for inclusion in the Google index. In fact, they are more or less diluting the value of the other pages on your site a bit. By only allowing pages with true value to be indexed, these pages might only get a small ranking advantage.

Some even say, too, that blog and list view pages should be altered. For these pages, it could be an option to set the Robots parameter to **NOINDEX, FOLLOW**. This means: Do not index this page, but just go ahead with crawling the links that you find on it. (There are also reasons not to do so; see Chapter 5 on URL structure). For pages with search results, this is definitely the best value.

If you are going to change this value for Robots for separate articles, you can set this in the actual article (**Publishing tab**). For the blog/list views, set it in the menu item (**Metadata tab**), as shown in Figure 4-6.



Figure 4-6. Robots parameter options

The parameter allows you to choose among four different options, as follows:

- **INDEX, FOLLOW:** This is the default value. It tells search engines to treat this page as a completely normal page, meaning it should include the site in the search engine's index and follow any link that it finds, in order to keep spidering the site.
- **NOINDEX, FOLLOW:** Do not index this page but keep following all links encountered on it. This could be a useful value for pages that do not offer additional value by themselves but provide access to articles that are. You could use this value for pages with search results or category lists, etc.
- **INDEX, NOFOLLOW:** Include this site in the index, but do not follow any links.
- **NOINDEX, NOFOLLOW:** Do not include this site in the index and follow no links. This is the value already discussed under the global Robots setting, allowing you to prevent any access to robots. This is useful for keeping this page (or complete site) out of Google.

Use these values with care. If you do not feel comfortable using them, do not touch them. You could do harm to your site if you set the values incorrectly. There are reports of web masters accidentally setting a site to NOINDEX, NOFOLLOW, after which Google removed the site from the index within 48 hours, and it took weeks to get back into Google again, after the NOINDEX was removed.

CHAPTER 5



URL Structure

Using a valid and logical URL structure is vital to your search engine optimization (SEO) efforts. It plays a major role in Google algorithms. These rules are not hard to guess; they are the same as those you would want URLs to adhere to yourself, as a user.

- They should be short, so that they are easy to remember.
- They should be readable to humans and search engines (search engine friendly [SEF]), not something like this:

```
/index.php?option=com_content&view=article&id=1.
```

We have already achieved this by setting **Use SEF URLs** in our Global Configuration.

- Avoid extensions such as .php, .html, etc. They are meaningless and can be easily avoided. Also, avoid including the term `index.php` in the URL (achieved by **URL rewriting**), or anything you can do to avoid meaningless characters.
- Use **hyphens** instead of spaces to connect words in the URL (no underscores or other diacriticals), as in `/url-with-multiple-words`. Happily, this is how Joomla already works, so there's no need for you to worry.
- Use your keywords in the URL. If someone posts a link to this page in a forum or such simply by pasting in the URL, this has the advantage of automatically including the keyword, as in the following: www.website.com/keyword.
- Preferably, URLs should not be more than two levels deep. So, avoid URLs such as the following (unless you're Amazon.com):
`/level1/level2/level3/level4/page.html`.
- Use a logical structure if you have multiple levels. A good example could be `/webshop/shoes/sneakers`.

Now that we know what we want to achieve in our URLs, we must know how Joomla builds its URLs.

URLs in Joomla

First, the URL structure of a Joomla site is defined by the **menu structure**. Only if the menu structure does not provide for a URL (an article not accessible through a menu item) is it defined by the category/article structure. An article that is linked to a menu item will simply receive a URL based on the **menu item alias** it is linked to. If the same article is linked to a second menu item, it will also receive a URL for the second menu item alias. I will give you a few examples of what this could result in, as understanding this mechanism is crucial for building great URLs.

- An article is linked to a menu item called Shoe repair. The URL will be /shoe-repair.
- An article is linked to a menu item called Shoe repair, and this is a submenu of the menu item called Maintenance. The URL will be /maintenance/shoe-repair.
- The article is not accessible through any menu. The URL is will be /123-article-alias (more on this unusual number later).
- The article is not accessible through any menu, and it is a member of a category. The URL will be 8-category-alias/123-article-alias.
- The article is available in a blog view under the menu item called Shoe repair. The URL will be /shoe-repair/123-article-alias.

Based on Alias

The way in which URLs are set up resembles slightly the logic of page titles.

Where page titles are derived from the title of an article or menu item, a URL is created by its alias. Whether this is the alias of the menu item or of the article again depends on how the article is linked (menu item vs. blog view), as in the preceding examples. Now, usually, this will give us a nice URL, but it does give us an extra option to further tweak it. You could use this so that your page titles are attuned to humans as much as possible, while the URL is optimized a bit more for Google.

Second, in the case of menu items, we have the same issue here as with page titles: the default menu item alias is simply derived from the menu item title, with lowercase letters and with hyphens instead of spaces. As menu item titles are usually very short, they are nondescriptive. However, you could manipulate the alias so that it contains your keywords more.

Now, whether or not this technique is actually useful is open to discussion. A more descriptive URL is usually longer, and harder to remember, but most people don't type in full URLs anyway. On the other hand, you could also do the reverse: create a shorter alias for a long article title. Time for some examples.

Example 1: Menu Item URL with Keyword

Say we have a contact page for a shoe brand, which is FootXXX. Now our domain is also footxxx.com, so the URL for the contact page is then footxxx.com/contact. Now, if you change the alias for the menu item for the contact page to contact-for-shoes, you managed to squeeze in the main keyword in your URL, making it footxxx.com/contact-for-shoes....This is illustrated in Figure 5-1.

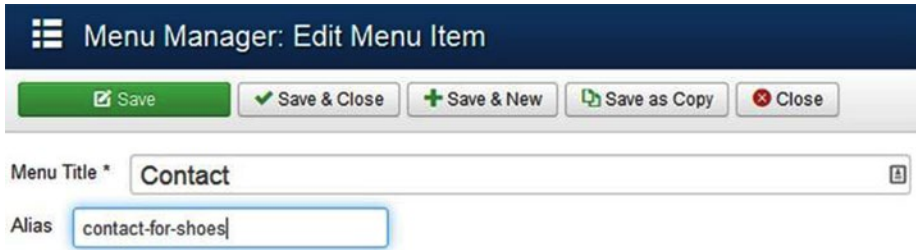


Figure 5-1. An improved alias for a menu item equals a better URL

This example may not seem very practical, but there may definitely be situations in which you can make good use of this technique.

Example 2: Shorter URL

Let's assume we have a site with a blog, simply linked through the menu item called Blog, and that we recently published an article entitled "The Future of SEO in the 20th Century." In this case, the URL will be based on the menu item + the article ID + the article alias. If you simply leave the article alias to what Joomla! generates for you, it will simply take the article title and replace spaces for hyphens. The URL for this article will be

/blog/123-the-future-of-seo-in-the-20th-century

This is a very long URL, and it is hard to remember and type. You can solve this by using a shorter alias, maybe seo-future, resulting in a much nicer URL, as follows:

/blog/123-seo-future

Just one thing to remember: **Only do this for new articles.** If you do this for an article that already exists and has been indexed by Google, it will no longer be found under the URL that Google stored, resulting in Page not Found errors and possible penalties for this URL. This is nothing you can't solve (see Chapter 11), but remember to take this into account.

Category Structure

It is crucial to any site with more than a dozen articles that the structure be established in a well thought out manner. How you do this is not very important. You might as well draw it out on a piece of paper first, but the structure of your articles, and especially categories, is very important. If you just start adding content without a plan, simply adding categories as you feel the need to, or even worse, without the use of categories, you will get lost in your content, especially with larger sites. If you get lost yourself, your visitors and Google will be lost as well.

It is very likely, then, that you may want to start structuring at some stage, probably a few months after the site is live, but now your URLs have already been indexed by Google. So, simply placing articles in other categories cannot be done anymore, because this will change your URLs (alarm bells going off, remember), and visitors and Google will encounter errors (**404 - Page not found**) when accessing the old URLs.

The way to avoid this is by carefully planning which categories you need, whether you require subcategories, and how you will place the articles in which category. If done correctly, you will have a nice structure to your content from the start, and this will be reflected in your URLs.

Menu Structure

About as important as your category structure is the **menu** structure. Make sure it is set up in a logical way. Also, remember that your URL structure is primarily based on the menu structure. It often makes sense to keep the menu structure similar to the category structure. For example, if I have a blog, the articles in the blog are often stored in the category Blog, and the Blog category is made available through a menu item also called Blog. This results in a perfectly well-organized structure that is transparent to anyone. Of course, you are not forced to do so, as long as you find other ways to create a logical structure.

Furthermore, do not squeeze too many items in a menu. A rule of thumb is to have no more than seven items per menu. If you have more, build a hierarchy, using either submenus on separate pages, or use drop-down menus. Again, this should not be too complicated. Three levels are the most you should use, and even that number can potentially confuse users. With deeper structures especially, a breadcrumb module can help organizing this.

Having a complicated and hard-to-understand menu structure will not immediately kill your ranking, but is definitely bad for usability and will make people leave your site earlier, unless you find alternative ways to create an overview and logic to assist your users.

Breadcrumb Module

The breadcrumb module is a very nice way to help visitors keep track of where they are on your site. It gives them a trail of how they got to the current page, allowing them to navigate through that route. (See Figure 5-2.)

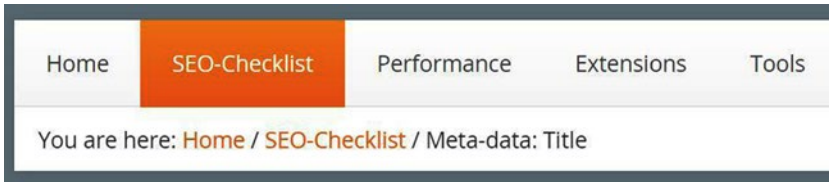


Figure 5-2. Breadcrumb module

You should be aware that the breadcrumb module is a reflection of the menu structure. So, if you have a blog that is located as a menu item called Blog under another menu item called Communication, the breadcrumb path will be

Home >> Communication >> Blog >> Article-title

■ **Note** All parts of the breadcrumb are clickable, except for the current entry, which is the last one in the trail.

However, if you have another link to the same article, say, in a footer menu item, in a menu item called Featured, the breadcrumb path will be

Home >> Featured >> Article-title

Preferably, avoid this (because you are creating a duplicate URL for the same content). If you have to link again, use a menu item of the type Menu Item Alias, as in Figure 5-3.

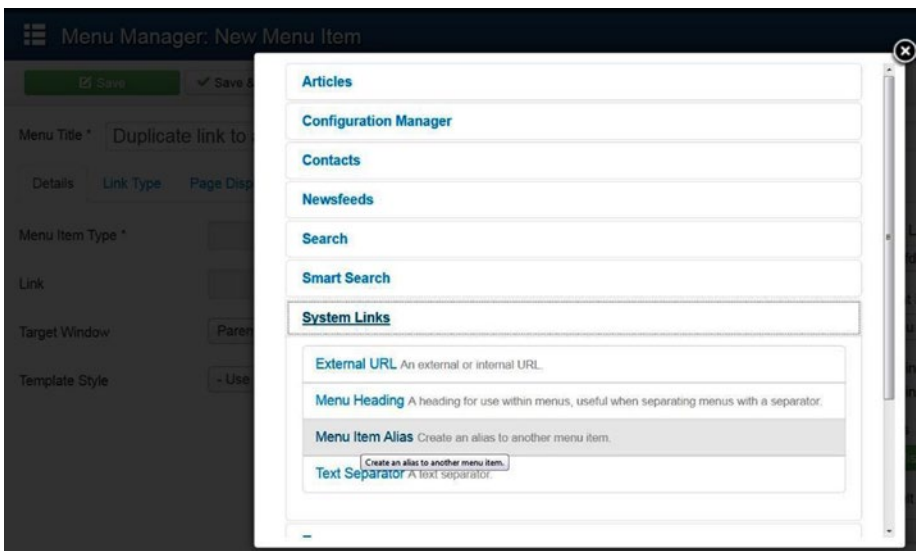


Figure 5-3. Use menu item aliases to link twice to the same article

Another very powerful feature of breadcrumbs is that they can be shown in the Google SERP. To achieve this, you have to mark them with microdata (more on that later). Doing so will make these items directly clickable from Google.

Another thing to note: show the breadcrumb module anywhere on your site, but exclude it from the home page. It makes no sense and adds no value here.

URL Hierarchy

If you've followed along so far, you should have a nice hierarchy already, both in your categories and in your menus, where both navigation and category structure follow the same pattern. Make sure especially not to create too many items in one level. If you need a dozen or more main categories, try to create a substructure with subcategories and submenus. Eventually, you will end up with a pyramid-like structure. Actually, it might be best to sketch this structure first, before you create your categories and menus in Joomla. (See Figure 5-4.)

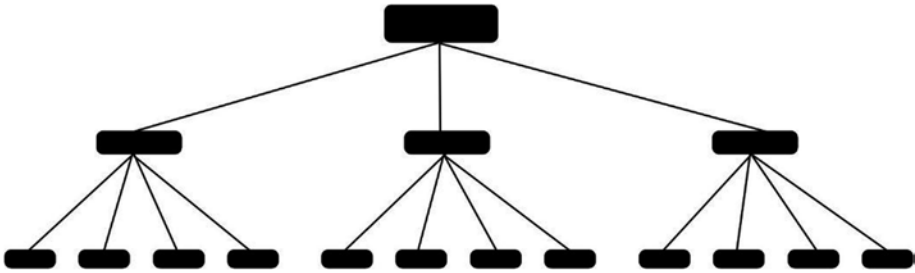


Figure 5-4. Pyramid structure for your URLs

In addition, if you regularly read blogs about SEO, you may have come across the term *link juice*. It does not sound very scientific, and it is not defined very strictly, but basically, it is a way to describe how the link value flows in a site. Pages receiving many internal links, especially if they have a logical structure, receive a higher value. In the preceding figure, this will usually be the pages higher up in the pyramid, with the home page dominating all and the main navigation, then the submenu items, and then the end points receiving less value or link juice.

This does not mean that the end points are not important, but they may be more appropriate for targeting the long-tail keywords, while the home page and the category pages are more attuned to your main keywords. Note that this flow can be emphasized by using breadcrumbs and showing the link to the category in an article.

However, what you often see on many sites is that the category pages do not contain any content by themselves! Sometimes, the category is just a bin with a label—and nothing else. In such cases, the pages with so much link juice are underrated, with so-called thin content: little or no text. Therefore, it is best to add a category description to your categories, and, of course, set this description to the value Show in your menu

items. An example would be the Performance section of my blog on www.joomlaseo.com/Performance. It contains many articles about performance, all linking back to this page through the breadcrumb trail. Of course, this is not always fitting for your situation, but be aware that you can use this to your advantage. If you decide to set your site up like this, make sure not to set category views to NOINDEX. (See the section “Robots” in Chapter 3.)

For the same reason, for simply showing submenus, avoid using menu item types, such as **External link** or **Page separator**, as intersections in the pyramid, which is something you often see.

Doing this correctly can greatly enhance your chances of receiving site links from Google in the SERP (search engine results page). (See Figure 5-5.)

Joomla! The CMS Trusted By Millions for their Websites

www.joomla.org/ ▼ Vertaal deze pagina

New York, NY, - January 20, 2015 - **Joomla**, one of the world's most popular open source content management systems (CMS), announced today the launch of ...

[Download](#) - [Extensions Directory](#) - [Demo](#) - [Joomla 3](#)

Figure 5-5. Site links for www.joomla.org

You will not only enhance your chances of getting these links but also influence which ones you get. Many sites do have site links, but often they do not receive the ones they would like. You can get rid of some site links you do not want by telling Google to remove them (in your Webmaster Tools account). However, but you cannot tell Google directly which ones you *do* want. You will require site structure to achieve this.

By the way, note that if you Google specifically enough, almost any site will have site links, but really valuable sites will also receive this in more general searches, and they really help you to stand out from the crowd.

Numbers in the URL

In one of the preceding examples, you may have seen that sometimes **numbers** are included in the URL. This does not occur for menu items, except if URLs are generated through the category/article structure, resulting in URLs such as the following:

/8-category-alias/123-article-alias

These numbers actually are the IDs in the database for these articles and categories. You will recognize them when you check the Article Manager in the back end of Joomla. In the overview, you will see this ID as the last column. The reason these numbers are there is for historical reasons, to prevent identical URLs in case two articles have the same alias. Identical aliases are not allowed, in any case, so these numbers are usually pretty useless.

Ideally, we would get rid of these numbers; they are non-SEF characters and add no value either for humans or search engines. Unfortunately, there are no ways to get rid of these numbers yet. This might change in the near future, as there is a lot of development regarding the Joomla Router, but for now, we are pretty much stuck with them. Currently, the only solution is to hack the Joomla core, but it is beyond the scope of this book to discuss that. If you really don't want numbers, you should either create menu links for every article or use an extension that builds cleaner URLs, such as SH404SEF.

CHAPTER 6



Header Tags

Header tags are the visual headlines and titles that you see on your site. They are vital as a visual sign of importance, but, in addition, Google uses them as such. Of course, you can create this effect by enlarging the paragraph font and changing its color, but that should be avoided. You should really use the correct HTML tags for this effect, which are the HTML tags `<h1>`, `<h2>`, etc., until `<h6>`. If your CSS is correct, the tags will be styled appropriately, usually with a visual hierarchy, in which the `<h1>` is the largest and the rest are progressively smaller. Besides being visually recognizable, they can be used also to create structure and hierarchy in your texts. In addition, I advise you not to use more than the first three levels, as it makes the structure harder to understand, though, of course, there can be situations for which you may have to disregard this.

The `<h1>` is a special case: it is the topmost level and should be sitting alone at the top of the pyramid. There should be no more than one occurrence of an `<h1>` in a specific page generally. (There may be exceptions, like a blog view with several articles, each having an `<h1>` within an `<article>` tag, but this is advanced territory.) Visually, having multiple `<h1>`s should not be too bad, but the `<h1>` is also used as an important indicator by Google—as the single headline associated with an article—and you do not want to confuse Google! For the most part, this is under your control. An exception could be a template that is configured to set the Site Name in the logo position as an `<h1>`, which would be really bad practice, but it happens sometimes. (You can check this in your source code, if necessary.)

Default Headings in Joomla

So, for the biggest part, creating this hierarchy is something you do in your editor window while creating the actual article. Most Joomla editors, including TinyMCE and JCE, allow you to do so very easily. The only issues are the top levels `<h1>` and `<h2>`. They are mostly controlled by Joomla for a number of specific situations.

In many cases, the most important heading tag on a Joomla site with default settings is `<h2>`, and not `<h1>`. This is owing to the following logic:

- For a Joomla article, the article title is taken as the `<h2>`.
- For the menu items, the heading is an `<h1>`, taken from the menu item title. The default setting is **not to show it**, however.

So, for most pages, an `<h1>` is not displayed; they only have an `<h2>`, as you can see in Figure 6-1, which shows a screenshot of a Joomla site with default sample data installed.

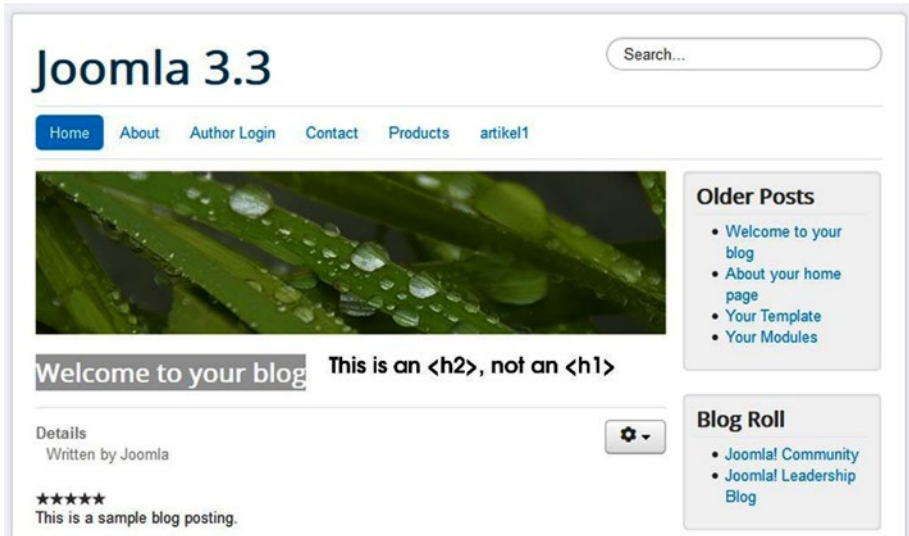


Figure 6-1. No `<h1>` heading shown

Actually, this may be the case with many of your pages too, so you need to fix this. How this is fixed depends a bit on your setup. Generally, there can be two situations.

Example 1: Articles Shown in Blog/List/Featured Articles

In this situation we have an `<h2>` as the title of the article, and we can only manipulate this from within Article Manager.

Click the **Options** button at the top right of the articles overview screen, as shown in Figure 6-2. This allows you to set global options for the article displayed.



Figure 6-2. Options button in Article Manager

Here, set the **Show Title** parameter to **Hide**. (See Figure 6-3.)



Figure 6-3. Set Joomla Show Title to Hide

Now, the actual title of the article will not be used, and you will have no title at all. This is even worse, of course, but can be fixed easily by typing the title in the actual editor window in the content, as shown in Figure 6-4.



Figure 6-4. Creating an `<h1>` in your content (code view)

Now you can set the title to whatever you want and make it an `<h1>`! Problem fixed. It's a little more work, but especially for smaller sites, this should not be an issue. Also, for inexperienced Joomla administrators, it is very straightforward to use.

Example 2: Menu Items Directly Linked to Articles

If you still use the default Joomla settings, this situation will give you an `<h1>` with the menu item (if the **Page Heading** is set to **Display: Yes**), plus an `<h2>` with the article title.

As the menu item will usually be short (something like *Home* or *Contact*), you will have an automatically generated `<h1>`, but it will be very short. Second, an `<h2>` with the article title will be present on the page as well, directly after the `<h1>`, resulting in two headlines right after each other.

There are a number of possible solutions to this situation.

- Use the procedure, as described previously, with the article title set to **Hide** and, additionally, **Show Page Heading** set to **No** (in the menu item). Again, type the desired `<h1>` title in the editor of the article.
- Override the current `<h1>` by going to the **Page Display** tab of the menu item, set **Show Page Heading** to **Yes**, and enter something more appropriate for Page Heading (Figure 6-5). This will now be displayed as an `<h1>`. Do not forget to set **Show Title** to **Hide** in order to prevent two consecutive headlines.

The screenshot shows the Joomla! article settings interface. At the top, the 'Menu Title' is set to 'Contact'. Below it, the 'Alias' is 'contact'. There are five tabs: 'Details', 'Link Type', 'Page Display', 'Metadata', and 'Module Assignment'. The 'Page Display' tab is selected. Under this tab, there are three settings: 'Browser Page Title' (empty), 'Show Page Heading' (set to 'Yes'), and 'Page Heading' (set to 'This will be the new H1 heading').

Figure 6-5. Using the Page Heading field to create an `<h1>`

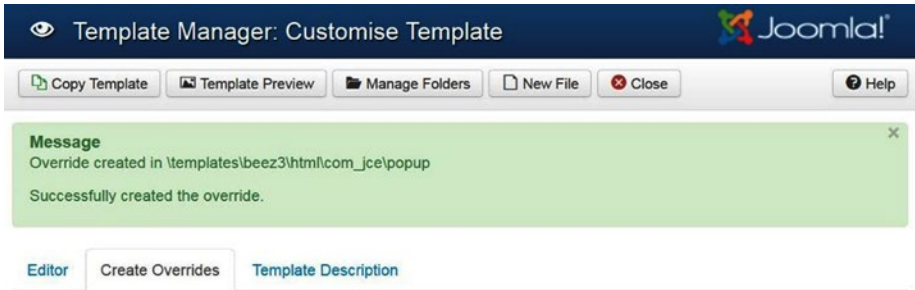
There are probably more ways to play with the article and menu settings to achieve something similar. It depends on how creative you are and how well you know Joomla.

The problem is that all these methods require manual intervention and extra configuration. Maybe you have a very large existing site and need a solution without going through all these options.

Template Overrides

Another often-used solution is to create a template override. While template overrides are usually for more experienced users, this is probably the easiest override that you can imagine. Also, creating template overrides in Joomla 3 has become extremely easy, so why not take a shot.

Go to **Extensions** ► **Template Manager** ► **Templates** and select the current template. Then go to the **Create Overrides** tab, click **com_content** (hidden options are now shown) and click **Article**. Note that if an override has already been created earlier (often the case for commercial templates), you will be notified. No problem, keep following along. You will either receive a message that the template override has been created or that it is already there. (See Figure 6-6.)



Modules	Components	Layouts
<ul style="list-style-type: none"> mod_articles_archive mod_articles_categories mod_articles_category mod_articles_latest mod_articles_news 	<ul style="list-style-type: none"> com_contact com_content archive article categories 	<ul style="list-style-type: none"> content edit editors form links

Figure 6-6. Creating a template override for `com_content/article`

Now switch to the **Editor** tab again, and then navigate to

```
html > com_content >> article >> default.php
```

A code-editor window opens up. (See Listing 6-1.) Now look for the following section, around line 45. (This may change in future Joomla versions, but the principle will remain.)

Listing 6-1. Template Override for `<h1>`

```
<div class="page-header">
<h2 itemprop="name">
<?php if ($params->get('show_title')) : ?>
<?php if ($params->get('link_titles') && !empty($this->item->
readmore_link)) : ?>
<a href="<?php echo $this->item->readmore_link; ?>" itemprop="url">
<?php echo $this->escape($this->item->title); ?></a>
<?php else : ?>
<?php echo $this->escape($this->item->title); ?>
<?php endif; ?>
<?php endif; ?>
</h2>
```

You will notice the `<h2>` tag. If you replace the opening and closing tag with an `<h1>`, your titles will appear correctly.

Note that you can achieve the same in Joomla 2.5 (which is also still possible in Joomla 3), by copying the file `default.php` located in

```
components\com_content\views\article\tmpl
```

to

```
templates\YOUR_TEMPLATE\html\com_content\article
```

(If needed, create the `html` and `com_content` folder in your template folder) and perform the same changes there.

If you make a mistake and need to go back, just delete the file `default.php` in the folder `html\com_content\article` (for both Joomla 2.5 and 3 versions).

Now check on the front end to ensure that all pages that formerly had an incorrectly created `<h2>` now have an `<h1>`.

Header Tags Extension

Another option could be to use an extension. There are a few that can assist you with this task, but a very good one could be the Header Tags extension. The Header Tags plug-in (developed by Conflate) is a plug-in that is capable of achieving this. You can configure it so that it checks for an `<h1>` and, if none is found, changes the first `<h2>` into an `<h1>`.

Just make sure that you check the outcome. A possible situation in which things might go wrong is if your template provider used an `<h1>` or `<h2>` to wrap the site name as a logo. This is a totally incorrect use of headers, and the plug-in will not solve this issue for you (and you can't expect it to). You will have to look up the code responsible for this in the template, replace the tags by a `span` or `div`, and use CSS to make them stand out.

CHAPTER 7



Optimize Internal Links and Anchor Text

Using internal links on your web site is a great way to provide structure and context to your visitors. It also helps to spread the desired link juice. Moreover, Google loves these internal links, provided they are built correctly, as they help it to easily discover all the pages of your site and understand their structure and internal relationship. This is one of the reasons you should never incorrectly set the Robots parameter to NOFOLLOW, as it would prevent the Google bots from following these links.

So, first, you should place and, second, create these internal links in the right way. There are several ways to build internal links.

- First of all, the links that make up your menu are also internal links. These are called the **sitewide** links (the same on all pages of your site) and are treated specially.
- **Breadcrumbs** are also great for creating internal links. Especially useful for large and complicated sites, they can be a great help to visitors for understanding the site. The same holds true for Google. If configured correctly (using microdata), Google can even display the breadcrumbs in the search engine results pages (SERPs). To use them, just publish a breadcrumbs module.
- Create text links manually. Whereas for menu links and breadcrumbs the anchor of the links may be limited because of the way Joomla works or owing to space limitations, links created manually within the text of an article or **Custom HTML** module are fully within your control.

Especially for links in your text, optimizing the anchor text is crucial. Anchor text is the words you click that make up a link. While there is not a lot you can do with the anchor text for a menu item called Services (because of space limitations in the menu), linking to the services page in the text of an article allows you to optimize the anchor text: the actual text that you click.

For example, do not create the link only by using the word *services*. Use something such as *shipping services*, ideally adding your main keywords. This is much more relevant. You should make a concerted effort to avoid anchor texts such as *click here*, *about us*, *download*, etc. For all of these, it is very easy to find more descriptive terms. See Figure 7-1 for an example.

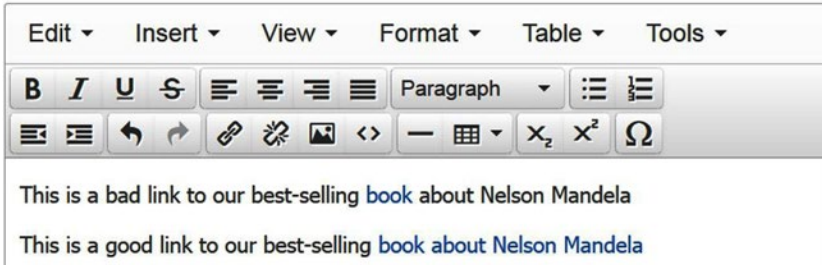


Figure 7-1. Creating correct anchor texts

You may think that for your visitors these links are very obvious, due to the way you set them up (underlines or different colors in your CSS or the context around it), but remember that search engines require the contextual reference of the anchor text to understand where the links are leading (or at least they still do).

You could emphasize important pages of your site by relatively linking a lot to them internally, though you should not overdo this (avoid over-optimization). If done right, they might show up as site links in the SERP.

Later on, in the link-building process, anchor texts become even more important, when building links from other sites to yours.

CHAPTER 8



Optimize Images

Earlier in this book, I explained that the spiders and robots that search engines use to crawl your web site cannot “see” images. They require textual context in order to understand images. For images, you can provide this context using either

- File names
- Alt text and title

So, when preparing your images, make sure that you use descriptive file names for your **.jpg**, **.png**, and **.gif** files. There is no problem in making them very descriptive. You can easily use more words to describe them, if you use hyphens to separate these.

For example, if you want to show a screenshot of how the JCE editor can be used to provide context, just use an image name such as **jce-for-setting-alt-text.jpg**. This perfectly describes what your screenshot is trying to tell. Make sure not to cheat, though. If the image is not actually related to the keyword, do not simply put your keywords in the file name of an image solely for SEO optimization purposes.

In addition to optimizing your file name, you should also provide an **alternative text** for the image. The original purpose of the alt tag is to be displayed when the image cannot be displayed (due to errors or slow connections) or to be read out loud by screen readers for disabled persons.

The alt text has to be typed using the editor you use in your Joomla articles. This could be the default **TinyMCE** editor, but I would advise you to use the JCE editor instead. The basic version of JCE can be downloaded for free, although some very useful plug-ins for JCE can be downloaded for a small fee (one license can be used across all your sites). For this specific purpose, however, the free version is fine.

When you add an image using JCE, by default, it will take the file name of the image, strip off the extension (jpg, png, etc.), and use that as the alt text. So, if you made an effort to use correct naming for the image, you can simply leave it up to JCE to set the alt text. If the suggested alt text is not sufficiently accurate, you can also set it manually, as shown in Figure 8-1.

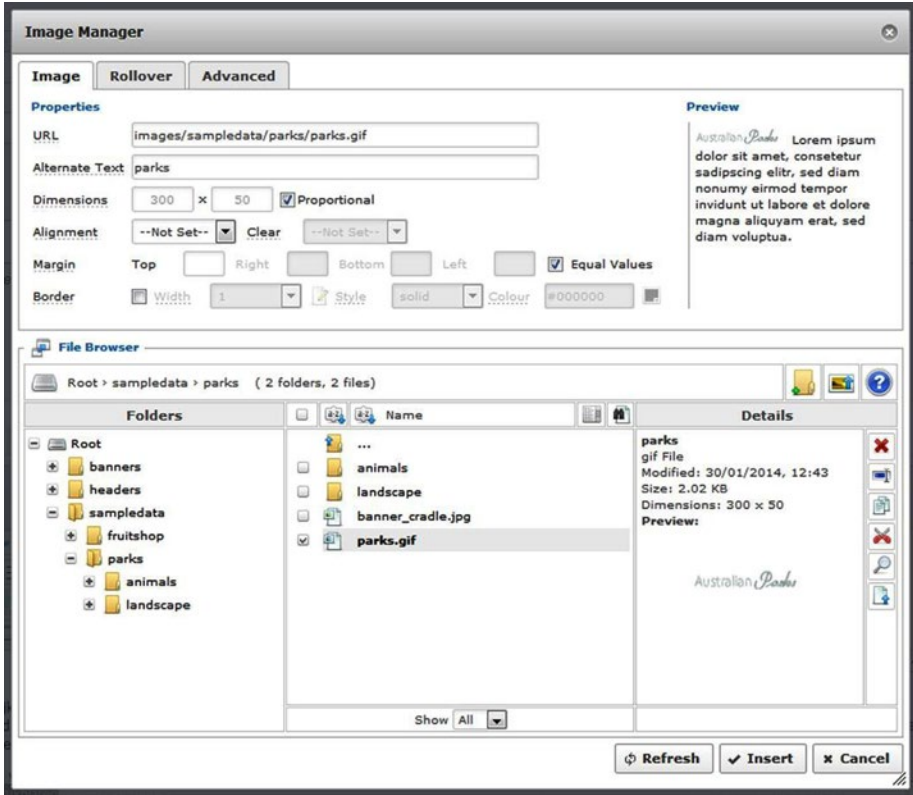


Figure 8-1. Use the JCE editor for alt text for images

If done correctly, your source code will include the alt text like this:

```

```

If you want to verify that all your images are set up correctly, you can use a tool on the Feed The Bot web site (www.feedthebot.com/tools/alt/) that checks the existence of an alt text on all your images.

This is a short chapter, but make sure you optimize your images accordingly. Make a habit of immediately renaming your files, and always add the alt text when uploading them to your article. Most sites contain many images, and these images can really be a boost to your search engine optimization, and not only for sites for artists/photographers, etc.

Note, however, that images can be heavy to load, so performance is almost as important as setting your alt text and file names. Read more about this in Chapter 18.

CHAPTER 9



Microdata or Rich Snippets

The use of **microdata** (sometimes called **structured data** or **rich snippets**) is not yet fully implemented in Joomla at the time of writing. Some, however, are already hard-coded in the most important parts of Joomla, such as the actual articles and the Contacts component. Joomla 2.5, unfortunately, misses out on this, though there are alternative solutions that I will discuss.

What is microdata? Microdata is the use of special metadata in your existing HTML tags that provides further meaning and classification. While a `<div>` or `` does not provide any relevance, by adding microdata, this is achieved. It's best to demonstrate this with an example. Let's assume we start off with the following bit of HTML:

```
<div> My address is 12 Joomla Street </div>
```

When we use microdata, this turns out to be like the code in Listing 9-1.

Listing 9-1. Address Marked Up with Microdata

```
<div itemprop="address" itemscope itemtype="http://schema.org/  
  PostalAddress">  
  My address is  
  <span itemprop="streetAddress">  
  12 Joomla Street  
  </span>  
</div>
```

This is much more code than the same information without microdata, but no one needs to guess that this is about an address. You can read a lot about the syntax to be used at www.schema.org.

What is especially nice about microdata is that Google loves them! It makes full use of them in its SERP pages, where you can often see nicely styled breadcrumbs, ratings, video thumbnails, prices for web shop products, etc. Check out the example in Figure 9-1, which has both a breadcrumb path and the review stars displayed.

Joomla Review - Joomla CMS - PHP Joomla - PHP / MySQL

php.about.com › Technology › PHP / MySQL › PHP/MySQL Basics ▾

★★★★★ Rating: 4 - Review by Angela Bradley

Joomla is a content management system based in PHP.

Figure 9-1. Microdata reflected in the Google SERP

Microdata in Joomla 3

Joomla 3.3 has hard-coded microdata in your content. In the near future, there will probably be possibilities for configuration. There may be a switch to turn them on or off on a global level, as well as on more specific levels, such as for all articles or contacts or specific articles/contacts. Also, it may be possible to mark up your article according to different schema.org types. Default markup will probably be Article, but alternatives could be TechArticle or BlogPosting. This is currently not implemented yet.

In addition, not all parts of Joomla are enriched yet. Currently, only the most used parts of Joomla have received them, such as Articles and Contacts. Modules such as the breadcrumbs module have currently been left out, so for these, you might still be looking for overrides or extensions (see Chapter 10 for Joomla 2.5).

Now let's show an example of how the code of a Joomla 3 article now looks (see Listing 9-2).

Listing 9-2. Joomla Microdata in Article

```
<div class="item-page" itemtype="http://schema.org/Article" itemscope="">
  <div class="page-header">
    Article heading here....
  </div>
  <dd itemtype="http://schema.org/Person" itemscope="" itemprop="author">
    Author here...
  </dd>
  <div class="content_rating" itemtype="http://schema.org/AggregateRating"
itemscope="" itemprop="aggregateRating">
    Rating here...</div>
  <span itemprop="articleBody">Article body text here....</span>
</div>
```

As you can see, article, article header, author data, rating, and article body now have been enriched with schema.org data. The same counts for your Contact form now (see Listing 9-3).

Listing 9-3. Joomla Microdata in Contacts

```

<dl itemprop="address" itemscope="" itemType="http://schema.org/
PostalAddress">
  <dd><span class="contact-street" itemprop="streetAddress">
    Street here...
  </span></dd>
  <dd><span class="contact-suburb" itemprop="addressLocality">
    City here...
  </span></dd>
  <dd><span class="contact-postcode" itemprop="postalCode">
    Postal code here...
  </span></dd>
</dl>

```

What is worth mentioning here is that it is now very important to use all fields in the contact form correctly, if you are configuring a contact in Joomla. The current display of the contact form is not always very pretty, and quite often, people just use the field for the address to type in some other non-address-related information, just because it looks prettier, like a welcome text. If you do this now, you'll have some random text that is marked up with address-related microdata. This is really not advisable. This is also true for other fields that are marked up with microdata. Make sure you use the field for its intended purpose.

■ **Note** If you cannot see the microdata showing up in your HTML code, it could be that you have a template override that is preventing the new code from being displayed. You can check if an override is present by accessing the site at `\templates\YOUR_TEMPLATE\html\com_content\article` and checking if there is a file called `default.php` that requires updating.

Microdata in Joomla 2.5

Unfortunately, Joomla 2.5 does not benefit from the microdata implementation. Luckily, there are alternative solutions. Some of these solutions can also be used in Joomla 3 for parts of Joomla that do not have microdata yet, such as the breadcrumbs module.

Now, for incidental use, you may be able to type the microdata information in the code view of your editor. For the JCE editor, there is even a microdata plug-in (beta) that facilitates incidental use of microdata (Figure 9-2). If you have a JCE subscription, you can download it from the JCE web site (www.joomlacontenteditor.net) as an additional plug-in.

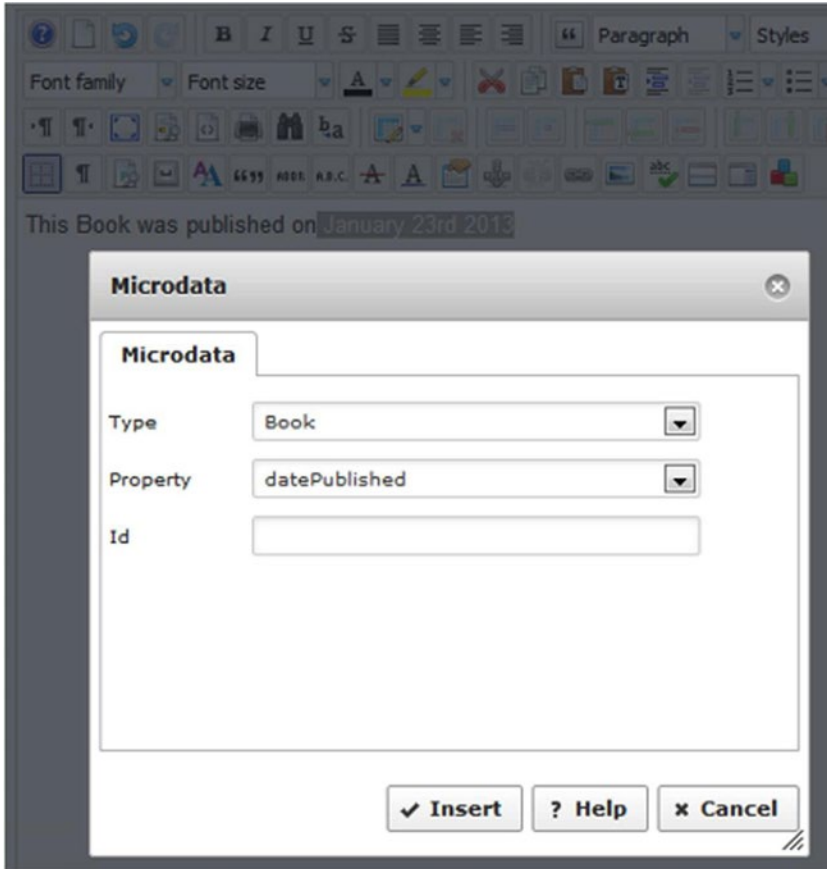


Figure 9-2. Microdata plug-in for JCE

For structural use, you probably have to extend Joomla.

One option is to create template overrides, so that the right tags are placed in the correct location. If you have the time and the skills, you can do this manually, but you can also buy sets of overrides from the folks at www.microdataforjoomla.com.

Of course, there are also plug-ins for this; see Chapter 10.

Your best option is probably to upgrade to Joomla 3, as it will keep getting better microdata implementation.

Microdata Extensions

Whether you are using Joomla 2.5 or 3, there probably still are parts of your site that could use microdata not yet currently available in Joomla. If you are using Joomla 3, keep in mind that there is a lot of planned development regarding microdata, so if you now need an extension, chances are that in a few months you may not require it anymore.

Currently, there are many plug-ins, usually targeting one specific piece of functionality in Joomla, such as the plug-in **Rich Snippets Vote** (www.deconf.com), which does this just for the voting system, or **Microdata Google Breadcrumbs** (www.expressive.nl), for breadcrumb functionality, obviously. Also, some larger extensions ship with optional microdata, such as **Hikashop** (www.hikashop.com), which is optionally capable (switch on the Hikashop microdata plug-in) of inserting microdata in their product pages, resulting in the price displayed in the search results (see Figure 9-3).

JA Amazon S3 - Cloudfront CDN Component for Joomla ...
www.joomlart.com/joomla/extensions/ja-amazon-s3 ▼ Joomla!Art ▼
 \$50.00 - In stock
 JA Amazon S3 Component brings the power of cloud computing to **Joomla!** Hybrid hosting is now possible, load site from your server and files (images/ ...

Figure 9-3. Price displayed in SERP, due to microdata usage

Further Microdata Implementation

At the time of writing, Joomla 3.4 was only just out, and a lot of issues surrounding microdata still need to be sorted. Some ideas as to where Joomla is heading are already known, however, and these may soon enter Joomla, so let me expand on this a little bit.

One option could be to choose whether microdata are marked up as Microdata or an RDFa format. This is a bit of a technical discussion and probably goes beyond most, but it could be interesting to have for advanced users.

What is more interesting for general users is that they may be able to choose a content type for their articles. Currently, it is hard-coded as Article, but the planned implementation is to allow users to choose between more, as shown in Figure 9-4.

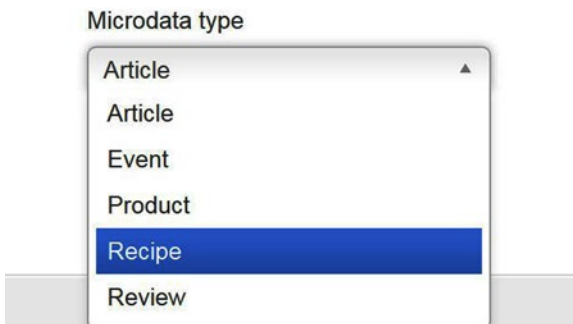


Figure 9-4. Possible options for microdata in future

Once this implantation is added to Joomla, you will probably be able to choose more than these options, but this is how things are planned. You will be able to set a default microdata type in the Article options, while you will be able to override this for specific articles in the **Publishing** tab of your articles.

Also, it would not be unusual to expect that standard modules such as the breadcrumb module will soon have microdata. Just make sure you keep reading about the latest developments. You can do so by accessing www.joomla.org but joomlaseo.com will also certainly publish blogposts related to this subject as soon as new features are added to Joomla.



Duplicate Content

Okay, this will be one of the longer chapters in the book. Duplicate content is an issue with many content management systems (CMSs), and Joomla is simply one of these. Before I dive deeper into this, why is duplicate content such an issue?

First of all, Google simply loves unique content, and any content that is duplicated is inferior to the original content. Now, between web sites (cross-domain duplicate content), this may not always be an issue, if you are the one who created the content. People sometimes “steal” content by copying interesting articles and using them on their own web sites or in data-feeds, but Google usually knows who the original owner is and ranks them higher. This is also why you should avoid doing the opposite and stealing content from others. Google does not like this, but many of your readers will also recognize that you are just repeating some popular opinion and find your site less trustworthy, or a lousy source of information. This should be common sense, and it is also not the issue I want to discuss here.

I am referring to **duplicate content within our own web site**. Now, of course, you should avoid reusing the same content in different articles, but this, too, should be obvious to you, and, again, is not the issue I am talking about here. I am actually addressing technical duplicate content, such as the same content being shown on an SEF URL and one or more non-SEF URLs. In the ideal situation, we should not have any technical duplicate URLs. These technical URLs may affect the overall ranking of your site (slightly).

However, what is also good to be aware of is that there is no need to panic right away if you spot duplicate content, because of all the technical duplicate links leading to the same articles. Google may usually notice that these links are probably there due to the CMS you are using and may not punish you all that severely. Especially for inexperienced users, fighting all these duplicate URLs may be complicated, and leaving those technical duplicates will not be all that bad. You can build a site that scores pretty well in search engine optimization (SEO) without getting rid of all of these. However, if competition for a ranking is really stiff, and you manage to get rid of them, this just might give you the edge over your competition.

Duplicate Content in Joomla

We already ran into a few examples, as you may recall. Let's go back to our test site and look at one of the articles. Suppose we have a look at a site with sample data installed. It could contain an article called **Sample sites**, and be linked to a menu, which is also called **Sample sites**. This means that the article is accessible through the following URL:

`/sample-sites`

(I will leave out the domain for clarity.) There are, however, many more possible routes to the same article, including the following:

- `/sample-sites` (the one we just discussed)
- `/index.php/sample-sites` (same with `index.php`)
- `/index.php?option=com_content&view=article&id=2` (article view)
- `/index.php?option=com_content&view=article&id=2&Itemid=10` (same, but linked to a menu through the `ItemId`)
- The same article with and without `www` in the URL
- With or without a trailing slash
- With parameters such as `/sample-sites?font=large` or others

And there are more possibilities for duplication. Often for blog sites, the **Home** menu item is a view of the **featured articles**, which results in duplication of the home page and the URL `/component/content/?view=featured`.

You see, there are a lot of possible routes to the same URL. Again, there is no need for immediate panic. Google often knows that Joomla sites function like this and will not punish you too severely for this by devaluing your rankings excessively. However, at times, you may find warnings about this in your **Google Webmaster Tools** account. Also, Google will be able to determine that the non-SEF links are technical duplicate URLs and not index them, but, ideally, you should help Google and only offer SEF URLs. Anything you do that helps Google get a clear picture of which parts of your site should or should not be indexed may help build Google's trust in your web site. If you are really serious about optimizing your web site for search engines, you should fix these duplicates.

Happily, there are quite a few ways you can get around this. Quite a few duplicates can be sorted with some tweaking of Joomla and small extensions. However, to really get the most out of Joomla, you will probably need one of the major SEO/SEF extensions, such as SH404SEF, MijoSEF, etc., but these will be discussed later, in Chapter 15. First, I'll cover some duplicate content issues and ways to address them without these.

Canonical URLs

What are **canonical URLs**? The best synonym for canonical is probably *preferred*.

Canonical URLs are a way of leaving the duplicate URLs intact, while adding code that tells bots and spiders that this URL is actually a duplicate of the provided preferred URL. Then there can be no confusion anymore as to what the preferred URL is for a set of duplicates. Say we have an article with the title **Meta data**, in a blog view of the category **Checklist**. It will probably have a URL like this:

```
/checklist/meta-data
```

Now the same article can be displayed under its **non-SEF URL**—maybe something like this:

```
/index.php?option=com_content&view=article&id=2&Itemid=10
```

Now these two URLs show the same content, so we need Google to state that it should only look at the first one. Canonical URLs are meant to arrange just that. The non-SEF URL should have a canonical URL pointing to the SEF URL. In the source code, this could look like this:

```
<link href="/checklist/meta-data" rel="canonical"/>
```

The same should work for pages such as e-mail and print views of your articles. These are accessible through another URL than that of the actual article you want to print or e-mail, so they should have a canonical URL pointing to the actual article.

Canonical URLs are only needed in the duplicate pages that point to another page. The actual preferred URL does not need one, as it would only point to itself. This is the way things work since Joomla 3.4. Before that version, all pages had a canonical URL, whereby the preferred pages had a canonical URL pointing to themselves.

Even if there are more URLs with the same content, as long as they are canonicalized, the non-preferred URLs will not be indexed, but only the preferred URL. Google advises canonicals as one of the preferred methods of treating duplicate content in your CMS.

In Joomla 2.5, canonical URLs were not part of the Joomla core, but in Joomla 3 they have been introduced. The canonical functionality is activated through the **System-SEF plug-in**. There is no way to switch on canonicals, unless you switch off the SEF plug-in, but this would leave you with non-SEF URLs. The only option that can be configured with the plug-in is **Site Domain** (Figure 10-1). This can be a useful option, however, if you use **parked domains** to point multiple domains to the same web site (which itself is not a good idea). In this exceptional case, fill in the preferred domain. If you only have one domain (99.9% of all sites), leave this empty.

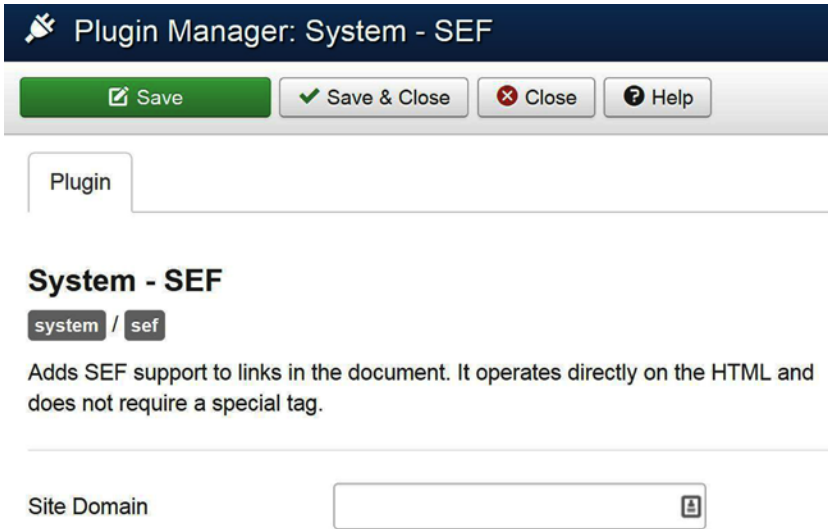


Figure 10-1. System-SEF plug-in

Currently (Joomla 3.4), canonicals in Joomla are sometimes correct but also sometimes not. Any page accessible as `index.php/article` is correctly canonicalized to the same without `index.php`, which is correct usage, but there are also situations in which they do not work as desired yet. By the time you are reading this, I hope that a correct implementation will have been created.

If you still have issues with your canonical URLs, a solution could be to install the **Canonical plug-in by StyleWare** (www.styleware.eu). No configuration is required, just install and activate the plug-in. Alternatively, advanced users could hack Joomla (be extremely careful!!!) to leave out the output from the Joomla core canonical plug-in (look up file `/plugins/system/sef/sef.php` and update lines 51–55. Also, you can unset the canonical tag (if you really want to) in your template's `index.php`, using the code in Listing 10-1.

Listing 10-1. Code to Unset Canonical URLs in Joomla (Solution: Robert Went)

```
$doc = JFactory::getDocument();
foreach ( $doc->_links as $k => $array ) {
    if ( $array['relation'] == 'canonical' ) {
        unset($doc->_links[$k]);
    }
}
```

301 Redirects Using .htaccess Tuning

When you activated the **URL-rewriting** feature in the Global Configuration, you also renamed the htaccess.txt file to .htaccess. The creation of SEF URLs is now maintained by some settings in this file, but it also allows you to do much more now. In Chapter 18 on performance, we will also activate a lot of useful features using this file, but for now, we can make great use of it by setting up so-called **301 redirects**.

A 301 redirect is actually a means of telling anyone accessing a specific URL (including search engines) to go to another URL, in the process also notifying the user that this is a permanent redirect. Any search engine encountering such a link will usually update its index in a few days to use the new URL. What is also nice is that the link juice or link value is also transferred to the new page!

There are many ways you can set this up. You can set up 301 redirects per URL, but you can also use advanced filtering techniques to pick up sets of URLs—maybe even a full domain, in the case of a migration to a new domain.

One great possibility is to either enforce or prevent the use of www in your URLs, getting rid of 50% of your duplicate URLs. Google sees www.site.com and site.com as two different URLs, and this means that for every URL of your site, a duplicate version is available. A few lines of code in your .htaccess file quickly removes 50% of your duplicate URLs. I strongly advise you to use this. You can either enforce www or non-www usage. In general, I would advise you to enforce non-www use, but in some situations, you may have to do the reverse (in the case of sites already indexed by Google as www or if you use CloudFlare, which specifically requires www). You can use the following examples in Listings 10-2 and 10-3, replacing www.site.com with your own site's URL.

Listing 10-2. Reroute www to Non-www

```
RewriteCond %{HTTP_HOST} www.site.com
RewriteRule (.*) http://site.com/$1 [R=301,L]
```

Listing 10-3. Reroute Non-www to www

```
RewriteCond %{HTTP_HOST} ^site.com [NC]
RewriteRule ^(.*)$ http://www.site.com/$1 [L,R=301]
```

This code does not simply reroute every URL to the home page but actually transfers it to the corresponding full URL, so that, for example,

`www.site.com/this-url`

will be redirected to

`site.com/this-url` (or vice versa)

Another (advanced) example (Listing 10-4) is this code to get rid of trailing slashes, in case your URLs are accessible with and without a trailing slash (reroute `/this-url/` to `/this-url`):

Listing 10-4. Remove Trailing Slashes from the URL

```
RewriteCond %{REQUEST_FILENAME} !-f
RewriteCond %{REQUEST_FILENAME} !-d
RewriteRule ^(.+)/$ http://%{HTTP_HOST}/ [R=301,L]
```

And there are many more possibilities. Just a slight warning: With code such as this, your `.htaccess` file can be extremely powerful, but any error, even a small typo, can make your site crash completely. Especially if you receive blank pages or server errors (code 500), there is probably something wrong in your code. Usually, returning one step may solve the problem; otherwise, simply restore a `.htaccess` copy from a fresh Joomla installation.

Robots.txt

A slightly tricky way of getting rid of all your non-SEF URLs could be to block them in your `robots.txt` file (more on this file in Chapter 12). It does not actually make these duplicates disappear, then, but you instruct Google to leave any specified URL out of the index. It just takes a very short line of code to do so, to be appended to the end of the `robots.txt` file, which is located in the root of your Joomla site.

```
Disallow: /*?*
```

This string looks up any URL with a query string, i.e., a question mark (all non-SEF URLs!). Make sure you do not have any URLs that could be blocked incorrectly then!

Whether this is a recommended method is questionable; setting a correct URL or rerouting to the correct URL is probably a better option. If you want to exclude articles or categories, it is better to use the Robots metatag, which you can set within the articles. It is preferable not to use `robots.txt` for that.

Extensions

There are many extensions that provide functionality to route all traffic to the SEF URLs. Most of these, unfortunately, are quite complicated and part of larger extensions that also provide more advanced features, such as **SH404SEF** (which I will discuss in detail). Second, they often change your existing URL structure, so they are less recommended for an existing site or for inexperienced users.

For users who prefer to stay away from these larger extensions, the plug-in **Non-SEF to SEF** from **Kubrik extensions** is a very nice alternative. You can download it for free for noncommercial use, though you can request a donation-key, if you want to support the developer.

Webmaster Tools

If the methods discussed so far do not solve your duplicate content issues, a final option might be to instruct Google to remove the duplicates from the index through your **Google Webmaster Tools** account.

Duplicate content: Webmaster tools If you do so, remember that you should also do this in your **Bing Webmaster** account.

Use this as a last resort. Sorting these issues from within your own site is preferable by far.

CHAPTER 11



404 Errors and Pages

A 404 page is one showing a **Page not found** error. This is something that should be avoided as much as possible. For visitors of your site, this looks ugly or unexpected, especially with Joomla's default error page (Figure 11-1).



Figure 11-1. Joomla's default 404 Article not found page

Additionally, search engines do not like them at all, so they should be prevented as much as possible. In case prevention is not possible, at least you have to offer a nice-looking page.

Note that 404 pages do not necessarily have to impact your site in a negative way. Of course, Google will not penalize you for typos that your visitors make. Also, if you have a page that is simply not there anymore for valid reasons, a 404 error is not necessarily bad. Say you have a web shop, and one of your products is discontinued and will no longer be sold, the page for the product will become useless, and if you keep it there, visitors will be annoyed, because they cannot order the product. In cases such as this, simply remove the page. For some time, Google will flag this page with a Page not found error, and after a while, it will remove the URL from the index, without affecting the rest of your site. This is valid use. What would be bad is if you had a menu link still pointing to the product that was no longer available, which would leave Google encountering this link every time the site is spidered. It is not hard to grasp that the latter is a case of bad design, and situations such as this hurt your search engine optimization (SEO).

Causes of 404 Errors

The reasons why a 404 error is generated can differ. The main issues are often the following:

- A user tries to type in the address of a link to your web site in the address bar of his/her browser, and he/she simply misspells it. This is something that cannot be prevented by you, unless you cater to some frequent typos!
- Your site is only accessible either with or without www, and other options lead to errors when users type them.
- You make other changes that lead to changed URLs on your web site, such as switching off URL rewriting after your site has been live for a while.
- A large Joomla upgrade, such as 2.5 to 3, might result in some changes (usually not, however).
- You completely redesign a site or replace the old one.
- Changes are made in the suffix settings in Global Configuration, adding or removing `.html` suffixes.
- You remove or rename an article or menu on your web site, and its URL simply does not exist anymore. This is probably the most frequent cause of 404 pages on many Joomla sites.

Preventing 404 Errors

Barring typos or other maladroitness from users, you should usually be able to do mitigate 404 errors. The majority of problems exist either as a result of changes we might make as the webmaster or through unexpected (to your users) setup of the site. For example, if your site has a URL `/products/chairs/chair-model-x`, make sure, too, that `/products/chairs/` is a valid link. Users expect this and sometimes navigate by the URL. With that in mind, we should at least make sure our site is set up correctly.

An issue such as errors occurring because a site is not accessible through www can be easily prevented. Of course, we do not want to keep both the www and non-www live on our site, but by creating a 301 redirect (explained in Chapter 10), users and robots do not notice irregularities. They are simply routed to the correct location, while bots also receive the Permanent Redirect notice and will continue to use the changed location.

On a live site, what you should **ALWAYS** keep in mind is what the consequences are of any action undertaken. Remember that your URLs are based on the menu and article **aliases** you created, so if you really have to rename an article or menu item, make sure to leave the alias in place. Happily, this is the default behavior of Joomla (see Figure 11-2).

Article Manager: Edit Article

Save Save & Close Save & New

Title *

Alias

Figure 11-2. Do not rename article-alias when renaming articles

As long as you keep in mind that your actions could result in changed URLs, there should be no reason not to perform the changes you want, as long as you **PLAN** them and take action (such as redirecting the changed URLs). You can even plan for expected user behavior, such as an often misspelled word in the URL.

Joomla's Redirect Manager

The standard solution for a changed URL is, again, the 301 redirect method. You can perform these via complicated `.htaccess` tricks or extensions, but, happily, Joomla ships with a nice feature in the core (2.5 and higher), which is the **Redirect Manager**. For inexperienced users, this is a really good way of managing changed URLs. You can access the tool via **Components** ► **Redirect**.

In case you encounter the message **The Redirect Plugin is disabled**, go to the **Plugin Manager** and enable the **Redirect plugin**. The reason that the plug-in is disabled by default on new installations is that for some sites, many errors are logged, and the database gets filled. As long as you keep an eye on this, there can be no harm in activating it. You can always switch it off later.

At first, nothing is shown in the **Redirect Component**, but you can easily test it. Just go to the front end of your Joomla site and type in a URL that does not exist, such as `/test-404`. You will see an error, and in the back end, a record will have been created in the database (see Figure 11-3).

Expired	URL	New URL	Referring Page	Created Date	404 Hits	ID
<input type="checkbox"/>	test-404			2014-06-25	1	1

Figure 11-3. 404 error logged in the Redirect Manager

If you click on the record, you can now update it if you want (see Figure 11-4).

Figure 11-4. Enter a Destination URL for a nonexistent URL

Here, you can enter a **Destination URL**. This can either be a full URL or a relative URL (make sure it is an existing one). Do not forget to **Enable** the status before saving.

If you now go to the URL that earlier resulted in an error, it will be routed to the new location. For smaller sites, this can be just the solution you need to fix changed URLs. After you make any changes or upgrades, either wait a few days and reroute the records generated by users hitting the old URLs or force the errors that you will expect to occur and correct them, before any user or bot encounters an error. The latter course of action is the preferred one. If you have the option, it is better to be proactive than reactive. This prevents anyone, either users or search engines, from encountering the error.

By the way, note that you can also apply mass updates. If multiple records have to be redirected to a single URL, you can easily do so. Note the Destination URL at the bottom of the screen in the overview. If you select multiple records and click it, you can enter a single destination for all the records selected. Note that they will be enabled directly.

■ **Note** By the way, you can also use the Redirect Manager to create short URLs for your site. Just create a new record, with the short URL as the source and the actual URL as the destination. You can now forward the short URLs to your audience, so they only have to remember the short URL.

The disadvantage of the **Redirect Manager** is that you must manually type or copy every separate destination URL, and any typo you make in a destination URL will lead to a new error.

Also, I already mentioned that for some sites, new errors are recorded all the time. These records may seem odd and completely unrecognizable to you, but usually, they appear to be created by hackers trying to locate a vulnerable plug-in on your site. This can really fill up your database a lot (extensions such as ExtraWatch, extensions.joomla.org/extension/extrawatch, can monitor database size for you). By periodic cleaning, you can sometimes prevent this, but if it keeps occurring, you have to prevent the logging of new 404 records. Luckily, since Joomla 3.4, this is now possible. If you check the Redirect Plugin in the Plugin Manager, you see the parameter Collect URLs. If you switch it off, no new records will be recorded, but the configured redirects will remain active.

Extensions

An alternative to the Redirect Manager could be an extension. A popular choice for this could be **RedJ** (www.selfget.com/products/redj.html). It works similarly as the Redirect component, but it allows for macros and regular expressions. Especially for large-scale changes of a structural nature, this could be very useful. Also, full SEO /SEF extensions such as **JoomSEF**, **SH404SEF**, and **MijoSEF** offer nice ways of rerouting 404's as part of their functionality.

Finally, a completely different extension is **Dynamic 404** by **Yireo**. It can reroute any 404 error to an existing page of your site, based on predefined settings, which are also customizable. Note that this is a commercial extension.

301 Redirects with .htaccess

Advanced users can also achieve **301 redirects** without any extension, by using the **URL rewriting** functionality in their `.htaccess` file. One advantage of this is also performance. Any URL request for your site will always be picked up by the `.htaccess` file before Joomla is activated. If you manage to create the redirect here, no performance hit occurs.

The possibilities with this file are endless, and you can find many resources and forums in which you can find extremely complicated bits of code, but just take a few general bits of code that might often be useful. Redirecting with `.htaccess` can be done on a per-URL basis, but also with wildcards using regular expressions.

The following line should already be present in your `.htaccess` for all these actions.

```
RewriteEngine On
```

As it is already there by default, do not repeat it.

EXAMPLE 1

So, let's start with a simple example in which we just want to take a source URL and a destination URL.

```
RewriteRule ^oldpage1.html$ "http://www.newsite.com/newpage1" [R=301,L]
```

Using this technique can be very useful in cases in which a site receives a completely new URL structure. Extract the old URLs from a sitemap or such and put them on a spreadsheet. Then define the corresponding new URLs and link them.

Then set them up in `.htaccess`. You will have to do some work manually, but it is all planned, and it can all be done before even publishing the new site.

EXAMPLE 2

If you still have `.html` suffixes that you want to get rid of, or your old site used them and the new one does not, you could use this bit of code. It will simply recognize the `.html` suffix and reroute you to the exact same URL without the `.html`.

```
RewriteCond %{REQUEST_URI} \.html$
RewriteRule (.*)\.html$ /$1 [R=301,L]
```

Remember that `.htaccess` is a very powerful tool that allows you to achieve many of your purposes and goals. The downside of this is also that it is extremely easy to mess up your site completely. Particularly if you receive server errors (error code 500) or looping redirects, your `.htaccess` code may be incorrect and require correcting. Just go back to the original configuration, then, or simply replace the file with an `.htaccess` from a clean Joomla install.

Offering a Custom 404 Page

You now know that if 404 errors cannot be prevented, a very ugly error page is displayed that is also completely different from the design of the rest of your site. It might even scare users away, because it almost seems as if they are on a different web site.

If you have bought your template from an established template provider, a custom 404 page will often be provided by them, but if they haven't, or if you build your own templates, you should solve this problem.

Fortunately, doing so is very easy. You can even do it without using an extension, just a little bit of coding.

First, simply create an article on your web site that displays an attractive text that you would like to offer your visitors. You could offer your apologies for the error (never offend your users, even if the error is their fault), in addition to some links to possible locations on your site—first, a link to the home page, or to a search page. Optional, although recommended, is to set the Robots value for the article to `NOINDEX`, `NOFOLLOW`, to prevent Google from indexing it.

Save and Close the article. Then, in the Article Manager, look up the article **ID** listed at the very end of the list, as shown in Figure 11-5, in which the ID is the number 9.

Status	Title	Access	Author	Language	Date	Hits	ID
<input checked="" type="checkbox"/>	Welcome to your blog (Alias: welcome-to-your-blog) Category: Blog	Public	Super User (Alias: Joomla!)	All	2011-01-05	5	3
<input checked="" type="checkbox"/>	404-article (Alias: 404-article) Category: Uncategorised	Special	Super User (Alias: Joomla!)	All	2011-01-01	8	9
<input checked="" type="checkbox"/>	About (Alias: about) Category: Uncategorised	Public	Super User (Alias: Joomla!)	All	2011-01-01	22	1

Figure 11-5. Article created to serve as 404 message, with Article ID 9 highlighted

So, we remember number 9.

Now we are going to create a new PHP file for the template that is active for your Joomla site. Online, you often find the advice to start with the file located at

`templates/system/error.php`

However, we will only need two lines of that file, so do not use it. Instead, create an empty PHP file at the following location:

`templates/YOUR_TEMPLATE/error.php`.

Open the file in your code editor. Now, update it, so that the complete code looks like that in Listing 11-1.

Listing 11-1. Create a 404 Error Page in Joomla

```
<?php
    defined('_JEXEC') or die;
    if (($this->error->getCode()) == '404') {
        header('Location: /index.php?option=com_content&view=article&id=9');
        exit;}
?>
```

Note that 9 corresponds to the article that we created for the 404 error. This is the simplest form, and good enough for most sites, so stop here if you are satisfied and the rest appears too complicated. However, you could make this a bit more attractive.

Advanced 404 Page: Link to a Menu

We are currently linking to the non-search engine friendly (SEF) URL of the error page. If you link the 404 article to a (hidden) menu item called **Error**, you can replace the ugly long URL for the 404 page simply with `error`, so that

```
header('Location: /index.php?option=com_content&view=article&id=9');
```

becomes

```
header('Location: /error');
```

After all these steps, check the front end of the web site, navigate to a nonexisting URL, and check if you indeed end up at the article we just defined, at the URL `/error`. In case of any issues, either correct your errors or just remove the `error.php` file you created, to go back to the original configuration.

More Advanced: Correct HTTP Header

An addition that might even be better is to replace the line starting with

```
header("Location....
```

with:

```
header("HTTP/1.0 404 Not Found");
echo file_get_contents(JURI::root().'/index.php?option=com_content&view=
article&id=999');
```

Note that on some hosts, the slash before `index.php?` has to be removed in order to prevent errors. To avoid errors, the full code will be the following:

```
<?php
defined('_JEXEC') or die;
if (($this->error->getCode()) == '404') {
    header("HTTP/1.0 404 Not Found");
    echo file_get_contents(JURI::root().'/index.php?option=
com_content&view=article&id=9);
    exit;}
?>
```

What this does is to send a **404 HTTP header** to the browser, **stay on the same URL**, and **retrieve** the content of the 404 page, instead of just going there, which is technically more correct.

Monitor Crawl Errors in Google Webmaster Tools

Note that next to monitoring 404 errors on your own site, either in the Redirect Manager in Joomla or whatever extension you use, you should also check the Crawl errors in Google Webmaster Tools (under **Crawl, Crawl errors**). Any errors you find there are often (but not necessarily) a sign of errors or of a bad setup of your site, especially if you see a sudden rise in errors. If a drastic rise occurs, you will even receive an e-mail from Google telling you this. If you do, make sure to review the errors and see if you can find out why they are occurring. Either prevent the error again then or use 301 redirects to prevent further errors.

CHAPTER 12



robots.txt

In your Global Configuration settings, and in the articles and menus, you have often seen, and possibly used, the Robots field. This is used to indicate whether certain pages of your web site should be indexed or not.

While not entirely the same, the reason for the existence of the `robots.txt` file (Listing 12-1) in the root of your web site is somewhat similar. You can use it to inform spiders and robots from search engines which part of your web site's folders should be opened for spidering. When you open it up in an editor, you will see the following, except for some commented out lines:

Listing 12-1. robots.txt File Content

```
User-agent: *
Disallow: /administrator/
Disallow: /bin/
Disallow: /cache/
Disallow: /cli/
Disallow: /components/
Disallow: /includes/
Disallow: /installation/
Disallow: /language/
Disallow: /layouts/
Disallow: /libraries/
Disallow: /logs/
Disallow: /modules/
Disallow: /plugins/
Disallow: /tmp/
```

The first line (`User-agent: *`) tells all robots that they should follow these rules (you can set rules for different bots, such as mobile and non-mobile, etc.), and the rest is basically a set of **disallow** rules for most of the folders in your Joomla installation. (Make sure to include the trailing slash). As a consequence of this, you should not create menu items with names corresponding to system folders, such as **includes**, or **templates**, as they will not be indexed. Your site will not work for such locations, in any event. Actually, when you try to create menu items with corresponding names, Joomla will not allow you to (see Figure 12-1).

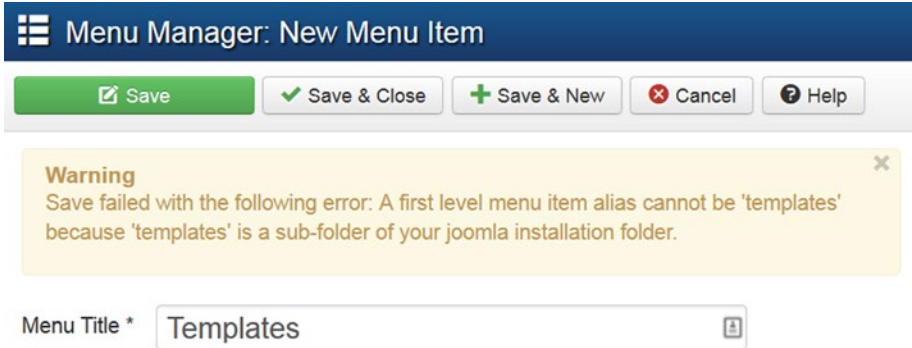


Figure 12-1. “Templates” menu title disallowed

Old robots.txt files—Check Them!

Depending on **when** you originally installed your Joomla site, there may be entries in the robots.txt file that should not be there. At the time the site was installed, the file was probably perfectly fine, but search engine optimization (SEO) is constantly evolving, and you may be stuck with the old configuration. The robots.txt file is not replaced when you upgrade Joomla (not even from Joomla 2.5 to Joomla 3). Especially, if you still find any of the three lines shown in Listing 12-2 in your file, remove them.

Listing 12-2. These Lines Should Be Removed from robots.txt

```
Disallow: /images/
Disallow: /media/
Disallow: /templates/
```

In Joomla 2.5 especially, the line for /images was often still found. This meant that Google was not allowed to access the images for your articles. Particularly if you have a site that depends on images, such as a photography site or a travel site, this could seriously penalize your SEO efforts! Earlier we spent a lot of effort optimizing our images, giving them a correct file name and an alt description, all of which becomes pretty useless if the images are not accessible to Google. You are basically saying: do not index these images. Of particular relevance is that they will probably not end up in the Google image search.

Something that was also changed in Joomla 3.3 is the removal of the lines for /media and /templates. For /media the issue is pretty much the same as for /images: some extensions store their images there (for example, the web shop extension HikaShop). Not having this folder accessible is bad for the indexing of the media contained in it. In addition, the /templates folder should be accessible.

So, just **remove** these line from the robots.txt file. You may also want to check periodically whether the robots.txt file has been overwritten by a Joomla update, although this seems to go smoothly at present.

Do Not Block Resources

The reason we also have to ensure that the `/templates` folder is **not blocked** in `robots.txt` is something you might not expect immediately. It is related to mobile design.

Google must be able to access the `/templates` folder, in order to use your CSS and JavaScript files. If it has no access, it cannot determine whether your site has a mobile design or not. To verify if Google has any issues checking your site, you can do the **mobile friendly test** that Google offers (www.google.com/webmasters/tools/mobile-friendly). In Figure 12-2, you can see how the test will view a web site that has a blocked `/templates` folder in `robots.txt`.

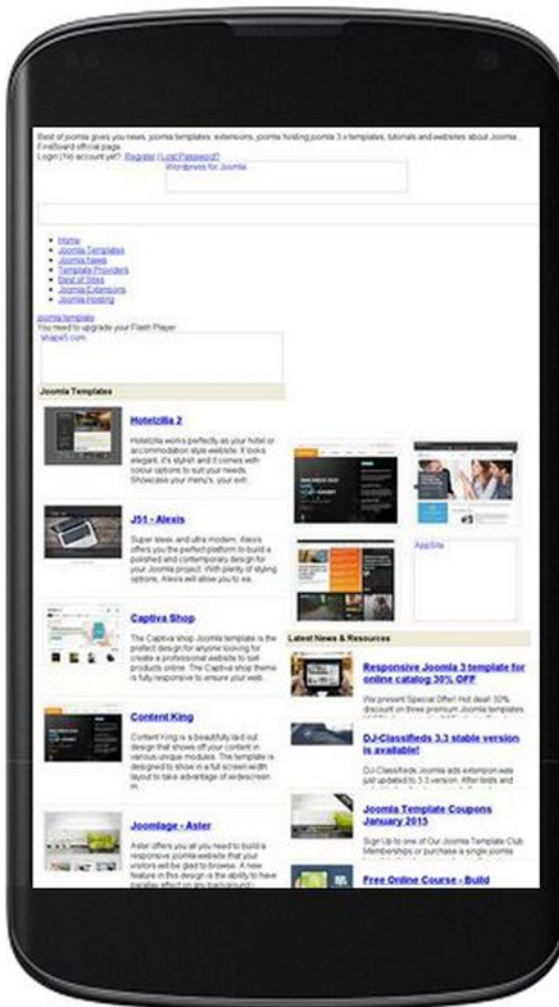


Figure 12-2. Joomla site with blocked resources

You can see that this is a really basic view, as if the CSS has not been loaded, which is exactly what occurs when the Google bots attempt to access this site. Now, Google cannot determine whether your site is mobile-friendly (responsive) or not.

This should be avoided at all costs. Mobile-friendly sites do not yet score significantly better than non-mobile sites, but all signs indicate that this may change in the near future. Via e-mail, Google already sends to owners of sites registered in Google Webmaster Tools warnings that their sites are not mobile-friendly and, in addition, mobile sites are marked as mobile-friendly in mobile search results, so you'd better make sure your mobile view is correct.

In the preceding example, this could perfectly well be a mobile-friendly site. Google is just not capable of verifying that, due to blocked access to the `/templates` folder, in which all the CSS and JavaScript files are residing. Luckily, the mobile-friendly test shows which locations are blocked. In this case, it is mainly the case for the `/templates` folder, so make sure it is not blocked in your `robots.txt` file.

If you see any other locations that are blocked, make sure you make further adjustments to `robots.txt`. Some plug-ins store their CSS in their own folder. So, maybe the mobile-friendly test finds a blocked CSS file that is from a plug-in that stores its CSS in its own folder, say something like the following:

```
/plugins/this_plugin/resources/css/css.css
```

Google cannot access this file because the folder `/plugins` is blocked as a whole by default. Best is not to unblock the whole folder, as it contains many files, and it is advisable not to let Google have access to anything in there, unless needed. So, instead of unblocking the whole folder, grant **specific access** to the blocked subfolder, by using an **allow** rule in `robots.txt`, like the following:

```
Allow: /plugins/this_plugin/resources/css/
```

This ensures that only this specific folder is opened, while the rest of the `/plugins` folder remains locked because of the line `disallow /plugins/` earlier in the `robots.txt` file.

In the end, make sure to check the validity of the file. You can use the **robots.txt Tester** feature for that in Google Webmaster Tools. Incorrect setup of the file could result in parts of it being forbidden accidentally, and you sure don't want that!

CHAPTER 13



Sitemap

Once you have built your web site, written all content, added metadata, and set up your URLs exactly as you want, your site is almost ready to be published. Preferably at this stage, but possibly later, you should create a file called `sitemap.xml` and store it somewhere in your web site, ideally in the root of your site.

The file is meant to help search engines understand the structure of the web site and which URLs are there. If you do not create a sitemap, most engines will be quite capable of sorting out your URLs for themselves, but, usually, they consider it a big plus if you supply one; it makes life easier for them.

You can supply multiple sitemap types—an **HTML** one for your users and one (or more) **XML** files for search engines.

Sitemap for Visitors

If you have a really small site, you might wonder whether an HTML sitemap really adds value. There are many small sites with only a few pages on which all content is linked through but a few menu items. Personally, I do not create a separate sitemap page for my visitors for sites like this. For larger sites, though, or sites with a complicated structure and multiple drop-down options, an HTML sitemap can be valuable. Disabled persons, especially, often like to navigate using a simple page from which they can oversee the whole site with simple text links. Also, you can use such a page for your own benefit, to see how well-balanced your site is, of course.

A page like this can easily be created by showing the menus of your site, with the full structure opened up. You can simply create it manually as an article, but you can also create an article with the **menu module** within. To do so, simply use a bit of code like the following in an article:

```
{loadmodule mainmenu}
```

This will result in the **main menu** being displayed within an article, output as a simple list. Any changes in the menu will also result in a change in the sitemap.

A second option is to use a sitemap extension. Popular choices are **OSMap** (free, previously known as XMap) and **JSitemap** (commercial). They allow you to create a sitemap with configurable options for your users. See Chapter 14 for a description of these extensions.

Sitemap for Search Engines

Google also likes to find a `sitemap.xml` file containing the URLs of your web site, so I advise you always to do so, regardless of how small or large your site is. Not even Google is always able to find every page on your site. Particularly if you have a page that cannot be reached by a menu link or internal link, Google may simply not encounter this page. A `sitemap.xml` file can reveal this page then, and even if your navigation is fine, a `sitemap` still makes life easier for Google. It is not for nothing that Google actually pushes you to submit a `sitemap`, as you will notice when you start using Google Webmaster Tools.

Especially if you put the `sitemap.xml` in the root of the site, Google will usually find it there and use it. You should also submit the `sitemap` in your Google Webmaster Tools, but I will get to that later. An example of a `sitemap.xml` could be as follows (Listing 13-1). Note that this is a simple example. You can also specify relative importance using a `priority` attribute, multi-language variants, etc.

Listing 13-1. Snippet from a `sitemap.xml` File

```
<?xml version="1.0" encoding="UTF-8"?>
<urlset .....>
  <url> <loc>http://website.com/</loc></url>
  <url> <loc>http://website.com/webshop</loc></url>
  <url> <loc>http://website.com/contact</loc></url>
  etc.
  etc.
</urlset>
```

So, you supply an XML file with specific formatting rules, plus the URLs of your site (omit ones that you do not want to have indexed!). There is no need to create one manually, as there are free `sitemap` generators online, such as www.xml-sitemaps.com and others. If you supply your URL, they will crawl your site and generate a `sitemap` for you. If you have a site that changes a lot, you may have to repeat this regularly, depending on how often content changes. Note that Google has some guidelines on how the `sitemap` should be formatted (see <https://support.google.com/webmasters/answer/183668?hl=en>). Note that for very large sites, `sitemaps` can be split into multiple files.

An alternative could be to use one of the previously mentioned `sitemap` extensions. They can also generate a `sitemap.xml` file. Usually, they do not generate it in the root of the site but somewhere deeper. If you submit your `sitemap` to your Google Webmaster Tools account, you should use this location then. Some extensions are even capable of pinging the file to Google and Bing, so that you do not have to do this.

If the `sitemap.xml` file is not located in the site root, it is also good practice to provide the location in `robots.txt`, even if you did not register it in a Webmaster Tools account. Then any search engine will be able to locate your `sitemap` file based on the provided location in `robots.txt`. Add a line such as the following to the bottom of `robots.txt`:

Sitemap: http://joomlaseo.com/index.php?option=com_jmap&view=sitemap&format=xml

Image Sitemap

If your site is heavy with images, and images could be useful for search engine optimization (SEO), an image sitemap should be created too (submit this one to Webmaster Tools as well). There are quite a few web sites that have a fair amount of traffic due to their images being listed in the **Google Image search**, sometimes up to 30% or more for sites such as photography and artist sites. It can really pay to optimize for this by adding an image sitemap. An image sitemap shows all URLs with images, including the exact path to the images, as in the code in Listing 13-2.

Listing 13-2. Snippet from an Image Sitemap File

```
<url>
<loc>
http://joomlaseo.com/e-book-joomla-3-seo-performance-nederlands
</loc>
<image:image>
<image:loc>
http://joomlaseo.com/images/joomla-seo-logo.png
</image:loc>
<image:title>joomla-seo-logo</image:title>
</image:image>
</url>
```

Creating an image sitemap is a bit harder than a standard XML sitemap. There are no free online image sitemap generators, and not every sitemap extension is capable of creating image sitemaps. If you need one for this, the commercial **JSitemap** can do this for you (see Chapter 14).

See Chapter 14, too, for some additional sitemap extensions.

CHAPTER 14



Simple SEO Extensions

Though you can optimize your site really well by just using the Joomla core features, it may often be easier to get some help from some simple extensions. They can also often solve issues for you that you would otherwise be stuck with. This chapter lists some simple ones. The extensions listed in this chapter are useful for easy management of some aspects of search engine optimization (SEO) and are usually quite suitable for newcomers to the world of Joomla SEO. Chapter 15 of the book is dedicated to **SH404SEF** and its competitors. These extensions are more suited to experienced users who need an all-in-one solution.

The extensions listed here are available for Joomla 2.5 and 3, and they are mostly free to use, at least for noncommercial use, unless stated otherwise. The following will be covered:

- **OSMeta**: For easy metadata management
- **Easy Front-end SEO**: For easy metadata management from the front end of the site
- **Non SEF to SEF**: Get rid of non-SEF URLs
- **Canonical**: Sets correct canonical URLs
- **OSmap**: Creates easy sitemaps
- **JSitemap**: Creates advanced sitemaps
- Some others

OSMeta

OSMeta (www.alledia.com/extensions/osmeta) is a relatively new extension, but, in my humble opinion, this is one of the best SEO extensions, as it simplifies metadata management significantly, without adding a lot of weight to the site.

The basis is originally taken from **SEOboss**, and much improved. The basic version already offers a very good option to maintain **article metadata** and page titles. The option to add and edit **page titles** is especially useful, as this is often difficult in Joomla. Without an extension, you have to go to the **Page Heading** tab of every menu item, while for individual articles, you cannot even set it.

The Pro version offers some further possibilities, such as editing a **menu item’s metadata** and **K2** support. I can highly recommended using this extension. It also keeps very close to the Joomla core: any metadata entered in the Joomla article is also available in OSMeta and vice versa. The extension uses the core fields in the Joomla database as much as possible.

So, how does it work? First, download the extension from www.alledia.com and install it. Then go to **Components** ► **OSMeta**, as shown in Figure 14-1, and you will see the central dashboard.

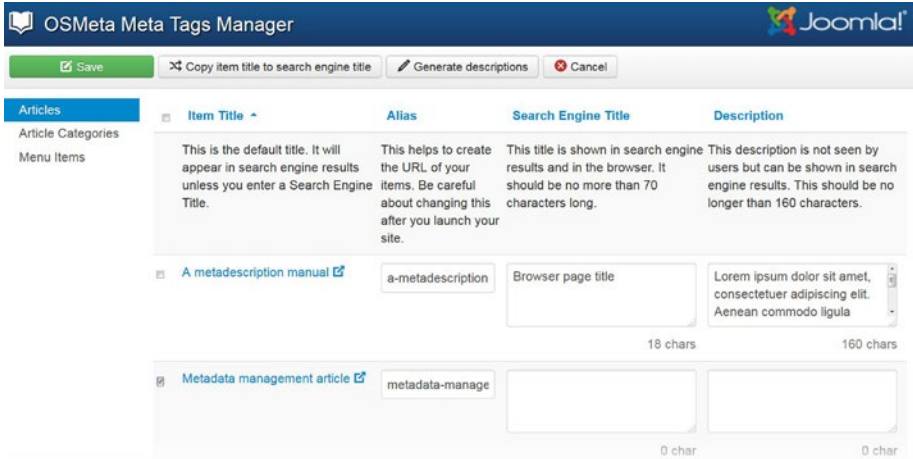


Figure 14-1. OSMeta dashboard

You probably recognize your articles here. If you had previously filled in any meta descriptions, you would see them here. Now, you can update them from here. You can edit/add multiple items from this dashboard, instead of having to go to the individual articles. Also, you see some buttons at the top of the screen that might come in handy.

- **Copy item title to search engine title:** Simply copies the article title to the browser page title
- **Generate descriptions:** Copies the first 160 characters of the article body text to the meta description field

Both options can be useful. After copying the items to their destination, it is best to verify and tune them. Especially for meta descriptions, the values might be broken off sentences, but at least you will have something to begin with. For optimal results, you really have to spend some time creating good nice meta descriptions!

You may also notice that there are small counters next to the fields, allowing you to keep track of whether you are approaching the maximum advised length for these fields (70 for titles and 160 for meta descriptions). Very useful!

You can perform the same actions for Categories and Menu-items (if you have a Pro subscription). Front-end editing is a feature that is being developed right now and may soon become available.

Easy Front-End SEO

Easy Front-end SEO (EFSEO) is meant for easy metadata editing in the front end. The plug-in is available from www.kubik-rubik.de as a free extension, with all functionality in place. For commercial use or if you want to support the developer, you can buy a donation code, which gets rid of the notification in the admin area.

As the name suggests, this plug-in is very easy to use for all users, from beginners to advanced. There is no need to worry about where you should add your metadata (article, menu, global configuration); you just browse the site from the front end, and if you see pages with missing or incorrect metadata, you can simply add or edit it from there.

Depending on how you configure the plug-in, you can either leave metadata already entered in Joomla in place or always overwrite it. If you want, you can even let EFSEO generate the meta description, keywords, and title automatically.

After you've installed the plug-in, you have to set it up from the **Plugin Manager** (see Figure 14-2).

Show Title field	Yes
Maximum number of characters	65
Show Title icon	Yes
Show Description field	Yes
Maximum number of characters	160
Show Description icon	Yes
Show Keywords field	Yes

Figure 14-2. Easy Front-end SEO plug-in parameters

You can set up user permissions (per ID or per group), and then you set which metadata parameters have to be configurable. Your options are

- Title
- Description
- Keywords
- Generator
- Robots

Usually, just Title, Description, and, possibly, Keywords should be fine.

Optionally, you can also set the plug-in to auto-generate the metadata from the **Automatic Mode** tab. While this may be nice to start off with, review the information. The plug-in takes just the first 165 characters from the body text, so this might not always be sufficiently relevant.

In the **Advanced** tab, you can even set how the front-end icons are displayed.

Now that all is configured, when you go to the front end and log in, you will see a nice set of icons sitting at the top right of your screen, which shows you for which parameters data is missing, in the order you configured (see Figure 14-3).



Figure 14-3. Easy Front-end SEO view from the front end

The image in Figure 14-3 indicates missing meta description and keywords, while title, generator, and robots are in place. Clicking the icon allows you to edit the record and correct it. As with OSMeta, the option to set your page title is really nice!

Note that this information is stored in the EFSEO plug-in tables, so if you de-install the plug-in, the metadata is gone.

Non-SEF to SEF

Some of the extensions I mentioned, such as SH404SEF, are very good for URL redirection, to prevent duplicate content, but they are also quite complicated, and they often change your existing URLs. When the only thing you want to achieve is to get rid of those duplicates, **Non-SEF to SEF**, also from Kubrik extensions, is a very useful plug-in (`joomla-extensions.kubik-rubik.de/nsts-non-sef-to-sef`).

Basically, it tries to find the best corresponding SEF URL for a non-SEF URL and reroutes it with a **301 redirect**. It does so using the internal Joomla logic. It also checks whether an article is directly linked to a menu item or not, so as to generate the best corresponding SEF URL.

The plug-in requires no configuration, except for switching it on. The only option is to exclude certain components. It is recommended, as it cleans up a lot of possible duplicate URLs and is really easy to use.

Canonical

The **Canonical** plug-in by **StyleWare** (`styleware.eu/store/item/26-styleware-content-canonical-plugin`) solves some issues regarding the setting of canonical URLs in Joomla 3. It is recommended in the event you have issues with the way Joomla 3 sets canonical URLs. There is no configuration required; you only have to install and publish it.

Keep in mind, however, that the way Joomla generates URLs is still under development, as it is an accepted fact that (at the time of writing) canonical URLs in Joomla are not perfect. If the problems are solved, you might choose to de-install the StyleWare plug-in.

OSMap (Previously XMap)

If you do not want to create your sitemap manually, you can use **OSMap** (www.alledia.com/extensions/osmap). It is actually a replacement for the previously very popular extension called XMap. XMap was probably the most popular extension for creating sitemaps in Joomla, but support has stopped. Luckily, support and development has been taken over by Alledia. OSMap can create both a sitemap for your users and a `sitemap.xml` for search engines. I already discussed OSMap briefly in Chapter 13.

Alledia aims to create extensions that make life easier for Joomla administrators, and this is an extension that fits this description. It is very simple to set up and gives you a sitemap in a few minutes, both HTML for visitors and XML for search engines. Also, it provides support for image sitemaps and multilingual sites.

JSitemap

A commercial alternative to OSMap that receives good reviews is **JSitemap** (storejoomla.org/extensions/jsitemap.html). I will discuss it here, as it is probably better than the (free) XMap extension.

JSitemap comes in two flavors, with the more expensive version offering more options. Although the difference between the two is small, I would opt for the expensive version, as it also offers support for image sitemaps.

After installation, you should first review the configuration. Usually it is fine, but there might be reasons to deviate. Check your data sources. By default, all content is shown in your sitemap, but if you have parts or pages that you do not want in the index, there are ways to exclude them. You can do this on a per-menu basis or by using wildcards.

Also, check your image sitemap. Often, there are images, such as icons for e-mail and print view, that offer no value and are repeated all over the site. By using filters, you can exclude them, so that only useful images are put in the image sitemap.

Once configured, all data is available from your central dashboard (Figure 14-4), with options to view your sitemaps, but also to send them to Google and Bing, even without a Webmaster Tools account (better to still register, though).

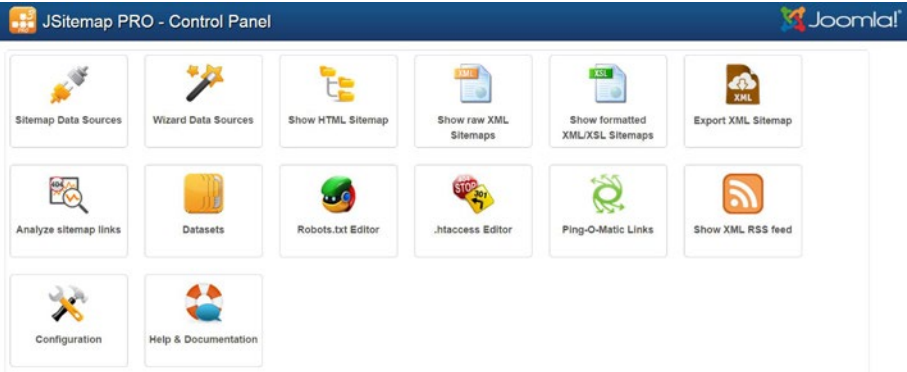


Figure 14-4. JSitemap dashboard

Other SEO Extensions

If you check the Joomla Extensions Directory, you will find many more extensions that could possibly help your SEO efforts, or that claim they could, at least. Use common sense to decide which ones could be useful in your personal situation, taking care only to install extensions from reliable sources, as always. Some popular extensions that might be worth looking at are

- **RS-SEO (www.rsjoomla.com/joomla-extensions/joomla-seo.html):** Used for managing SEO metadata, on-site and off-site. It ranks your pages for presence of metadata and titles, etc., but it can also track PageRank, keyword performance, and competition. This is a commercial extension.
- **iJoomla SEO (seo.ijoomla.com):** This is also a commercial extension that offers functionality similar to RS-SEO.
- **SEO Generator (www.suchmaschinen-optimierung-seo.org/seo-nachrichten/seo-generator):** Similar to Easy Front-end SEO, this plug-in can automatically generate metadata. Setting metadata manually will always be preferable to automated solutions, but there may be situations in which automating the process might be required.
- **Header Tags (www.joomla-specialist.net/joomla-extensions/header-tags-plugin.html):** I already discussed this extension. It is useful for setting up your <h1> to <h6> tags.
- **SEOSimple (github.com/nocean/SEOSimple):** This extension can be employed to automatically use the starting chunk of your article text as the meta description. If you need an automated solution, use it, although manual creation is always preferable.

CHAPTER 15



SH404SEF

A separate chapter has been dedicated to **SH404SEF**, because it is one of the most powerful and popular extensions for managing search engine optimization (SEO) for Joomla web sites. The extension used to be one of the Editor's Choice extensions on the Joomla Extension Directory, and it is very popular among Joomla SEO administrators. Later in the chapter, I will also discuss some of SH404SEF's competitors.

By the way, if you are new to the subject of SEO in Joomla, I advise you to skip this chapter for now and stick to one of the extensions mentioned in Chapter 14. There, I discuss some extensions that are far easier to manage and better suited to beginners.

SH404SEF is not just an SEF extension, it also offers more functionality. Actually, almost all the tools to manage SEO that were discussed so far can be managed with SH404SEF. Features included are

- SEF URLs (other, but better, than Joomla's)
- 301 redirect for non-SEF URLs
- Managing metadata from one dashboard
- Canonical URLs
- Custom 404 page
- Easy management of 404 errors
- Security features
- Managing social media
- Short URLs
- Google Analytics dashboard integration

You might not need some of these features, which is why some prefer to use more lightweight extensions with less complicated options and features, but that's also a matter of personal preference. For most users, it simply offers an all-in-one package for SEO and more.

Plan Ahead!

You should always plan ahead when building your web site, but with SH404SEF installed, planning is even more important. Changes in your category structure or renaming stuff can easily break your old URLs and mess things up. Also, you should never just install and switch on the extension on a live site, because your URLs will surely break. Also, if you find you don't like the extension, de-installing it does not always occur without problems.

The one thing to keep in mind is that SH404SEF **builds URLs differently than Joomla**. In Joomla, the way that URLs are built depends on your setup. For articles attached to a menu, URLs are based on the menu item they are attached to, whereas articles not linked to menus are built on a **category/article** basis as well.

SH404SEF **always** uses the category/article structure as a base, regardless of how features are linked or shown. The big advantage of this is that it does not matter how articles are assigned and what a site structure is like. Even if it is linked via two menus, the URL of an article will always be the same, while in the Joomla core, this would lead to two different URLs (duplicate content!).

How exactly the URLs are created is configurable. Usually, you will set the category not to show in the URL for the **Uncategorized** category, while for very deep category structures, you may choose to include only the last one or two categories. Based on these configured settings, your URLs are set.

Preparations

SH404SEF can be downloaded (subscription required) from www.weebly.com (previously the extension was owned by Anything Digital). Weebly has some more extensions that are worth checking out, but SH404SEF is probably the company's best-known one. You can get access to the extension, which allows you to install it on multiple domains, for \$39 a year.

We will install the extension on a web site about books. If you want, you can create the test setup for yourself, so that you can follow along.

- We create a clean Joomla 3 web site without any sample data installed.
- Next, under the **Uncategorized** category, we create two more: **Novels** and **Biographies**.
- In the Uncategorized category, we place three articles: **Home**, **About us**, and **Contact**.
- In Novels, we have articles about *Moby-Dick* and *The Catcher in the Rye*.
- In Biographies, we have **Al Capone** and **Nelson Mandela**.
- For Novels, we create a menu item called **Our Novels** as a category list, and for Biographies, another menu item called **Our Biographies**. The other three articles are linked to their own menu item as **Single Article**.

The only thing we configure in Joomla before going to SH404SEF is to **switch on URL Rewriting** in the **Global Configuration** settings, plus **rename** `htaccess.txt` to `.htaccess`, of course. By the way, you can use SH404SEF without URL Rewriting, but that leaves you stuck with the `index.php` part in the URLs.

If we just use the default Protostar template, the main menu now looks like that in Figure 15-1.

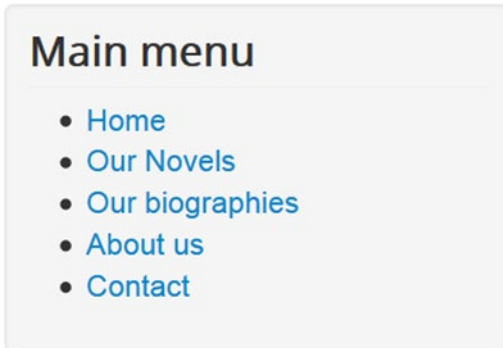


Figure 15-1. Menu, as per our example

For comparison later, before installing SH404SEF, let's take the URL for the Nelson Mandela book as an example and memorize it for later.

`/our-biographies/7-nelson-mandela`

Installation and Configuration

Now, go the **Extension Manager** and install the extension you downloaded from the Weeblr web site. You should receive a message that **SH404SEF has been installed successfully**. Now go to **Components** ► **SH404SEF**, where you will see a dashboard like that shown in Figure 15-2.

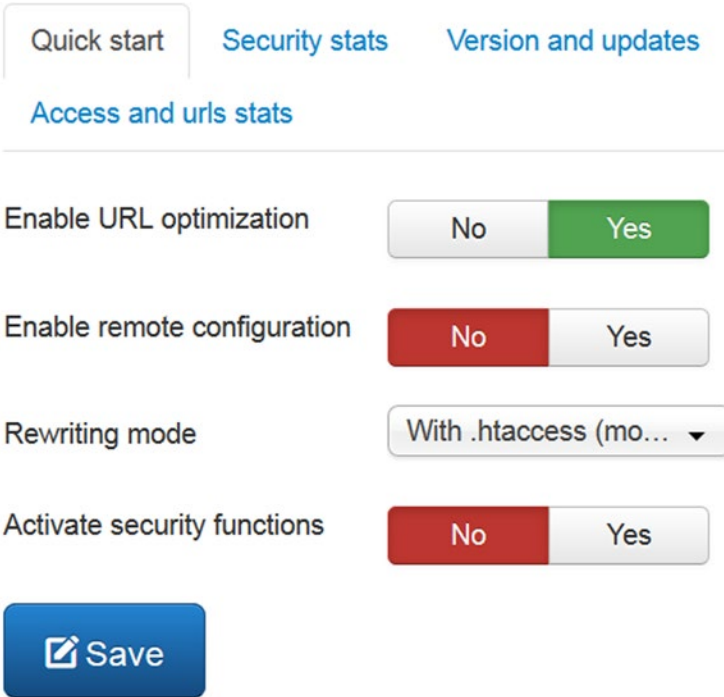


Figure 15-2. SH404SEF dashboard

Do not touch these settings yet! They will automatically optimize your URLs with default settings, and this is not what you want. Especially, do not touch the **Enable URL optimization** button, as that actually triggers your URLs to be changed.

First, click the green **Configuration** button at the top-left part of the screen. A big modal window opens (Figure 15-3), with hundreds of options that can be configured.

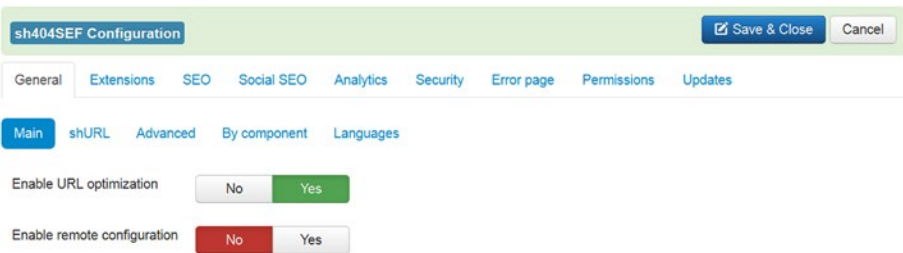


Figure 15-3. Configuration options

This may seem a bit intimidating at first, but, happily, we only need to change a few of these now; most of the parameters are fine. Note that I will only give you a sample of how I usually set up SH404SEF, so feel free to play with any parameter you find interesting. For now, I will only mention the basic settings that may have to be reviewed.

- **General—Main:** Note that the Enable URL optimization option is here once again: do not touch it for now. The other fields, with defaults, are fine, except for the following two:
 - **File suffix** is set to `.html`. Just clear the field, as there is no reason to add suffixes.
 - **All lowercase:** Set this to **Yes**, to prevent caps in the URL (capitalization can also be a cause of duplicate content).
- **General—shURL:** By default, **shURL** is enabled, which means that for every existing URL, a shortened URL is created. This could be a useful feature, if you must often inform people of your URLs, in case they have to type them in. The URLs are simply rerouted to the correct one (301), so you don't have to worry that they cause duplicate URL issues, but on most sites, they are not needed, so I usually switch off the feature.
- **General—Advanced:** Defaults work well, except these few:
 - **Rewriting mode** should be set to with `.htaccess` `mod_rewrite`.
 - **301 redirect www/non-www:** Set it—preferably to non-www (shorter), but www is fine too; just choose one.
 - **Log 404 errors:** It is currently set to **Yes**, which is good, especially for new sites. However, there is a potential risk of the database filling up, if a lot of errors are logged, so I usually switch it off after the site has been stable for a few weeks.
- **General—By Component:** No need to touch this one. A reason to use this could be a component not playing well with SH404SEF. In that case, you can set it to use the default Joomla SEF logic. Note, however, that if you do so, 404-logging will not work anymore, as SH404SEF will no longer be allowed to capture all traffic. Examples of such Joomla SEF default logic include HikaShop and JSitemap.
- **General—Languages:** This is only for multilingual sites. If I am sure that my site uses only one language, I usually set all values to **No**, as in rare cases, language strings such as `/en` are inserted in the URL, even if English is the only language installed.

- **Extensions—Joomla:** This allows configuration of your URLs, based on the Category structure. Defaults usually work fine, but you may want to tune them to your wishes.
 - Usually, I change all settings here to use an alias for title, category, and menu, as this allows me to differentiate, if required.
 - You can fine-tune how your URLs are built. Defaults work fine; especially, leave **Uncategorized items** to use only the title, to prevent the word *uncategorized* in the URL.
 - One more thing: Set **Insert content table name** to **No**. It adds the word */table* to the URL for list views (the menu item is the category list).
- **Extensions—Other options:** No need to touch these, unless specific configuration is required.
- **SEO—SEO Management:** Default settings are fine. I only change **Change multiple H1's** from **H2** to **Yes**. You should prevent multiple `<h1>` tags anyhow, but this can be used as a precaution.
- **SEO—Page title:** This uses the same logic as the **Site Name** settings in **Global Configuration**. I advise you to think carefully about what is best for your site. This results in the content of the `<title>` tag.
- **SEO—Joomla content page title:** Further options to tune `<title>`.
- **SEO—Mobile:** This was useful before responsive design evolved, as you could activate a separate template for mobile devices. It is no longer needed for responsive sites.
- **Social SEO:** Allows you to add social buttons to your articles. Be aware that for further configuration, you must navigate to the **Joomla Plugin Manager** and look up **sh404SEF core—Social buttons**. If you don't need it, disable the plug-in, as it adds a bit of JavaScript to your pages, which should not be loaded then.
- **Analytics:** Allows you to check your Google Analytics data from within your Joomla web site. Enter your web property from your **Google Analytics** account here (UA-xxx-xxx).
- **Security:** Especially if you are not already using a security extension such as **Akeeba's Admin Tools**, this could be very useful for securing your site. This is outside the scope of this book, however.
- **Error Page—Basic configuration:** This feature creates a default 404-error page within your template, moving away from the ugly Joomla 404-page. Settings on this tab work fine.
- **Error Page—Default 404 Page:** This creates a nice text for your visitors, in case they encounter a **Page not Found** error.

When done, save your settings. If necessary, review your settings. Now you can set **Enable URL optimization** to **Yes** in the main screen (already shown in Figure 15-2). In addition, make sure that **Rewriting Mode** is set to use `.htaccess` and save your settings.

Now, as soon as someone hits the front end of the site, all URLs will be generated automatically. Do this, and browse to the Nelson Mandela article again. Note the difference in the URL for the Nelson Mandela book.

- Before SH404SEF:
`/our-biographies/7-nelson-mandela`
- After SH404SEF:
`/biographies/nelson-mandela`

You will see that the Category/Article structure is nicely applied to the URL now. Note, too, that the original URL no longer exists. If you hit it, you receive a 404-error.

At this stage, we have our initial configuration ready, but we will have to review our results. We performed a lot of configuration in one go, and it's easy to miss a step or make a mistake. We will review our configuration in the next chapter.

URL manager

When we click **URL manager** in the menu to the left, we will get the following screen (Figure 15-4):

shURL	SEF url	Non-SEF url	Meta	Duplicates	Aliases	Custom
1	all-yu-need-to-know-about-joomla-and-seo	<code>index.php?option=com_content&Itemid=101&id=1%27&lang=en&view=article</code>		2		
2	Blog/	<code>index.php?option=com_content&Itemid=124&id=13&lang=en&layout=blog&view=category</code>		3	2	
3	Blog/adding-breadcrumbs-microdata-to-joomla	<code>index.php?option=com_content&Itemid=124&catid=13&id=71&lang=en&view=article</code>				
4	Blog/changing-domain-url-using-301-redirect	<code>index.php?option=com_content&Itemid=124&catid=13&id=70&lang=en&view=article</code>				
5	Blog/developments-and-news	<code>index.php?option=com_content&Itemid=124&catid=13&id=85&lang=en&view=article</code>				

Figure 15-4. URL manager

This is a list of the URLs created by SH404SEF. For most SH404SEF users, this is the focal screen for tuning and configuration. You can clearly see the relationship between the SEF and non-SEF URLs. You might want to copy the non-SEF URL in your browser, to note that, indeed, it takes you to the correct SEF URL.

What may also strike you is to see URLs for pages you did not configure, such as **lost-password** and **lost-username**, and even pages for the **print view** of your articles. What you may not have realized is that the corresponding non-SEF URLs can always be accessed on any Joomla site, even though you have not configured any login page. SH404SEF does nothing more than create a nice SEF URL for URLs that have always existed (if needed, you can set a **NOINDEX** value for **Robots** now, for pages that are shown here but that you don't want indexed by the search engine spiders).

You can also access any other feature that has been created for these URLs, such as

- **shURLs:** We did not activate them, but if you did, you can see them here.
- **Meta:** If metadata has been set, you will see a flag (which you can also click), allowing you to quickly check whether all pages have metadata entered. Note that only metadata configured in SH404SEF are checked. If you have already set your metadata in the Joomla article, this is not indicated.
- **Duplicates:** Sometimes, two different non-SEF URLs lead to the same SEF URL, which is actually how duplicate content issues are prevented by SH404SEF. If so, this is indicated. Usually, you will not touch these.
- **Aliases:** This indicates whether any aliases have been set. If you want to reroute another URL to the one we are looking at, you can do so by setting an alias, effectively creating a 301-redirect. If you have to set an alias, use the SEF URL for this. Do not enter a non-SEF URL.
- **Custom URL:** When the URL generated by SH404SEF is not the one you would actually like, you can simply set it to another one. The existing URL will be added as an alias automatically by SH404SEF, so that no errors will be generated if users still try to access it.

Each of these parameters also has its own screen. I will get to that.

Purge URLs

Above the list of URLs, we also see a red **Purge** button (Figure 15-5).



Figure 15-5. Purge button in the URL manager

This button will clear all your URLs, so, preferably, you should **never** touch this. The only time I usually do is when I am reconfiguring or building a site, if I've discovered that I have made a mistake in the configuration. As long as the site is **not live**, this is not a problem.

In such cases, correct the configuration, Purge your URLs, and regenerate them, by simply visiting the front end of your site.

It is preferable to do this before building metadata, titles, and custom URLs although, usually, these should not be deleted. Only the actual SEF-URLs get cleared. After having set your configuration correctly, you will never have to touch this button again.

Set Up the Home Page

You may have noticed that a separate URL has been created for the article that is linked to the **home page** of our site. The article called **Home** has the same content as the **actual home page**. This is a consequence of the rule that every article receives a URL based on the combination category/article. For Uncategorized items, we have elected not to include the category in the URL, so for the article called Home, a URL called /home/ is created.

The duplication of content can (must) be easily corrected, together with some other settings we still have to set: click the **home page meta** button (Figure 15-6) at the top of the list of URLs.



Figure 15-6. Home page meta button in the URL manager

Click it and a modal window will pop up (see Figure 15-7).

 A screenshot of a web form for configuring home page meta. The form has three tabs: "Title and meta tags" (selected), "Aliases", and "Social networking". Below the tabs, there are three input fields:

- "Page title" with a text input field containing "Hom page title here".
- "Description tag" with a text area containing "Home page meta here". At the bottom left of the text area is a green circular button with the number "141".
- "Keywords tag" with a text input field containing "Home page keywords here".

Figure 15-7. Home page meta configuration options

First, make sure to fill in your metadata here. This is one of the most important settings, as this controls how the home page is displayed in Google's SERP pages!

Second, click the **Aliases** tab, and enter the URL for the article assigned to the home page—in this case, simply home. Not doing so definitely leaves duplicate content in your site.

Aliases Manager

The **Aliases manager** gives you an overview of all aliases configured. Usually, you will work from the URL manager screen, but sometimes this overview can be useful for troubleshooting.

An example could be a situation in which you have set many aliases and accidentally created a loop.

shURL Manager

This screen is potentially useful for your **shURL** management, but you can also use it to view the information from the URL manager.

404 requests manager

This is an **extremely useful** screen. It helps enormously in situations in which you have to fight the **404 errors** that users encounter. For new sites, this may not be very relevant, but in situations in which you reconfigure an existing site, or changed the URL structure, it certainly is.

If you have followed along, you may have seen that before activating SH404SEF, the URL for the Mandela book was different than after, and accessing the old URL led to a 404 error.

This error was logged and is available for investigation in the **404 requests manager** screen (see Figure 15-8).



Figure 15-8. 404 requests manager

Were you still working with the core Joomla **Redirect Manager**, you would have to look up the correct corresponding destination URL, copy it, and paste it in as the URL to redirect to, which is quite cumbersome, if you have many errors. Now, with SH404SEF, that is extremely easy. Just click **Redirect to a SEF URL** and a new window will open (Figure 15-9).

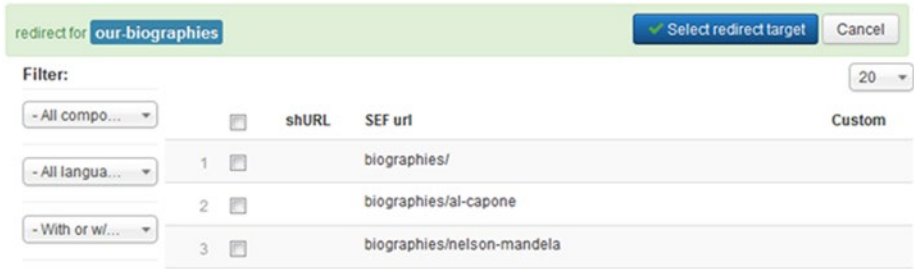


Figure 15-9. Choose a URL to direct to

Based on the erroneous URL, a list of **matching URLs** is suggested, which is correct in almost 90% of cases. If not, you can still set it to **not filter** and look up the correct one.

Sometimes, users or robots keep hitting odd pages in your site, generating many errors a day that are impossible to track and are useless. In cases such as this, it might be a good idea to stop logging 404 errors, by switching off the logging feature in the SH404SEF configuration. Otherwise, your database might fill up. The good thing is that if you stop logging, the original redirects remain active.

Titles and Meta Manager

This overview supplies a list of all your non-SEF URLs and the metadata entered for them (see Figure 15-10).

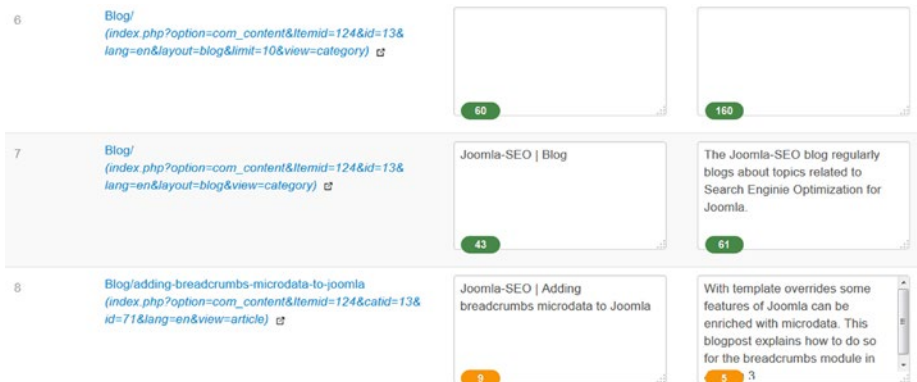


Figure 15-10. Title and meta manager overview

If two non-SEF URLs are set to the same SEF URL, you will see at least two records, which is why I prefer to work from the URL manager, as it is based on the SEF URL overview. Any data you included in the URL manager will be visible here, however, so to quickly oversee your metadata, this screen is still very useful.

One additional nice feature of SH404SEF is that all fields for entering title and meta description show a counter with the remaining number of characters left before reaching the advised 65 or 165 characters (Figure 15-11). This feature is also added to the Joomla **Article Manager**, which is especially useful for your page titles (though you should remember that **Site Name**, which is not included in the calculation, may be added!).



Figure 15-11. Character count for metadata

This may also be a good point to reflect on whether you should always use SH404SEF's **metadata management**. The point is that for new Joomla articles, a separate action is needed to add the metadata in SH404SEF. This might not always fit into your workflow, such as for sites offering many new articles on a regular basis. You might prefer to add the metadata at the same time, while still creating and editing the article in Article Manager. As long as you do this consistently, this result is quite possible, as the Joomla metadata will be used then.

Analytics Manager

If you have entered your Google Analytics credentials, you can check your data from the SH404SEF **Analytics Manager** (Figure 15-12). Especially for users who only monitor and access one site, this can be useful. It supplies the same data as that within Google Analytics.



Figure 15-12. Analytics Manager in SH404SEF

Disadvantages

Of course, there are also disadvantages to the extension. I already mentioned that many users find the extension complicated to use, but this depends on skill and experience. Among other disadvantages sometimes mentioned are the following:

- **Performance:** Some users encounter performance issues resulting from many database queries. With correct server setup, this should not be a problem, however.
- **Issues with third-party extensions:** There definitely are extensions that do not play well with SH404SEF, so before installing, check if this is the case for your installed extensions. Some extensions, excluding the extension from within the SH404SEF configuration screen, will work fine; with others, you may have a hard time.
- **Issues with upgrades:** SH404SEF adds complexity to your web site, so upgrades may become more difficult.
- **De-installing issues:** SH404SEF can be de-installed, but there are sometimes issues in which traces remain (files), causing unusual behavior on your web site.

Extensions Similar to SH404SEF

SH404SEF has a few competitors that offer similar, full SEO/SEF functionality. Best known are **MijoSEF** and **JoomSEF**. They work pretty well and may even beat SH404SEF in some respects. The best way to find out which one works best for you is simply to test it. What is especially nice about these alternatives is that they have a basic version that you can use for free, though you cannot access some of the more advanced features. Usually, however, free versions should provide you sufficient insight to determine whether you want to buy the full version.

MijoSEF

MijoSEF offers a similar dashboard as SH404SEF, with large icons instead of a basic UI (Figure 15-13):

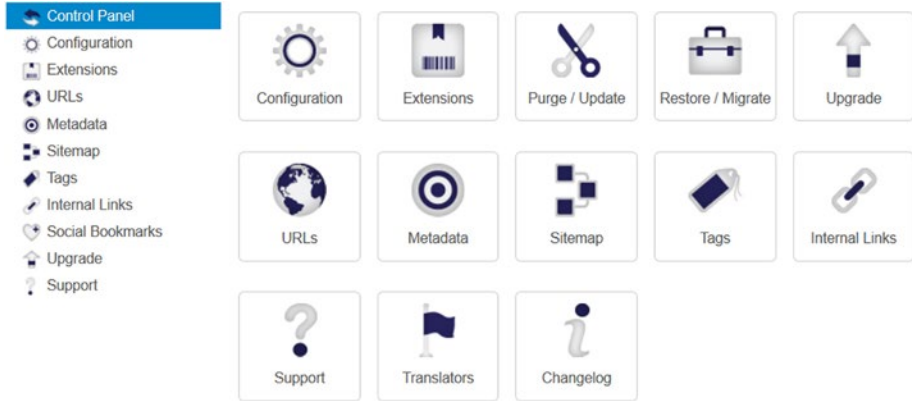


Figure 15-13. MijoSEF dashboard

MijoSEF offers some functionality that SH404SEF does not, such as

- Automated creation of internal links
- A sitemap
- Automated generation of metadata

URLs are also based on a similar configuration set (see Figure 15-14).



Figure 15-14. MijoSEF options

JoomSEF

Like SH404SEF and MijoSEF, **JoomSEF** offers a nice control panel (Figure 15-15).

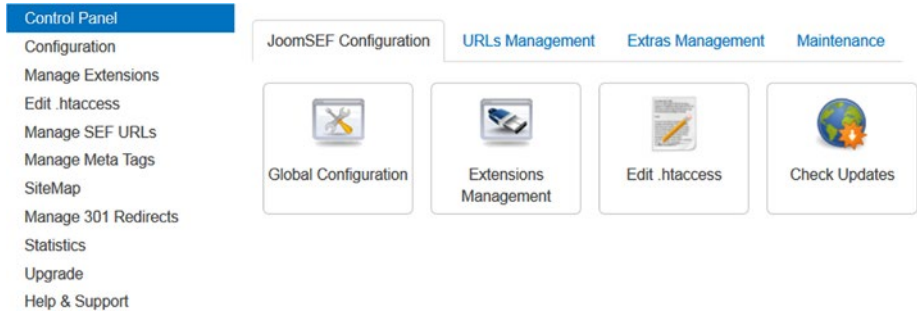


Figure 15-15. *JoomSEF Control Panel*

Especially in the Pro version, some features are available that can offer extensive possibilities, such as

- Editing `.htaccess` from the component
- Checking keyword density
- Auto-generation of metadata
- Setting up Cron tasks

Which one of these three large SEF extensions is best is up to you to decide. You might not even like them and will prefer to use Joomla core options or smaller extensions instead. As long as you know what you are doing, any choice could still result in a perfectly optimized Joomla site. In addition, check how these tools behave between Joomla upgrades. SH404SEF is fairly stable. In any case, make sure whichever tool you use does not have a habit of breaking between Joomla upgrades.

CHAPTER 16



SEO Site Audits

Okay, so our site is ready, we fully optimized it with the options available to us in the Joomla core and maybe a little extension here and there that you chose to install. Time to publish our site!

Of course, you should now monitor the site closely, to see whether it is running well and whether we have any 404 error messages, but that should be a continuous process.

What you might also want to do is run the site through a few handy search engine optimization (SEO) checking tools and audit it. There are quite a few tools for this, depending on what exactly you need. Some of these are very basic and free to use, for one specific purpose, while there are also massive tools that offer extensive SEO possibilities but are also extremely expensive. I will discuss a few handy ones here, plus one of the commercial SEO tools that offers a free limited version.

SEO Doctor Firefox Plug-in

You can easily look this one up in the add-ons directory for Mozilla. Install it, and you will see a little window (Figure 16-1) sitting at the bottom left-hand side of your screen.



Figure 16-1. *SEO Doctor button*

It is ranking every site you browse for SEO. It is not a check on minor details, just the basics, so aiming for 100% should not be too hard. If the score of your own site is below 100%, just click the score, and it will open to tell you what your issues are (Figure 16-2).

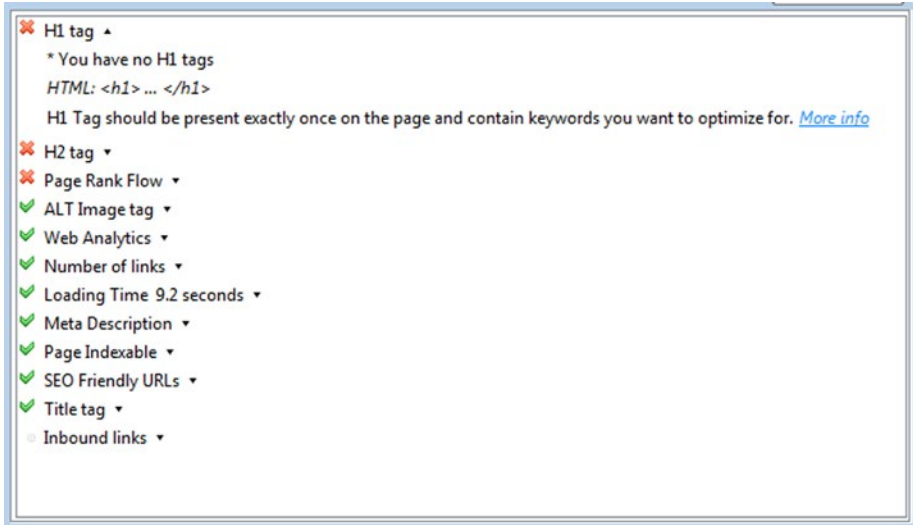


Figure 16-2. SEO Doctor details

In the example in Figure 16-2, it tells us that we have issues with SEF URLs, multiple H1 tags, and with Page Rank flow. You can even further expand on any ongoing issue. You might be surprised at what you forget to optimize. This is very easy and not too obtrusive.

MozBar

Moz is probably the leading provider for SEO software (though probably too expensive for most users). It has developed a nice free toolbar for Firefox/Chrome, however, that provides some good information—the **MozBar**. It provides a bar with some basic information sitting at the top of the browser, but clicking the details button opens up a lot of information about the page, such as metadata, links and images found, alt text, markup, microdata, and incoming nr. For links, MozTrust (the Moz version of Google’s Pagerank), etc., see Figure 16-3.

Tag/Location	Content	# of Characters
URL	http://joomla-seo.com/	21
Page Title	Joomla-SEO Search Engine Optimization tips for Joomla	56
Meta Description	Joomla-seo.com helps Joomla administrators to optimize websites for Search Engine Optimization. Rank high in Google with tips from our SEO Checklist.	148
Meta Keywords	Joomla, SEO	11
H1	Search Engine Optimization & Performance tips for Joomla! Share on Facebook Share Share on Twitter Share Share on Google+ Share	126

Figure 16-3. MozBar for Chrome and Firefox

It's very nice to gain quick access to a lot of information without diving into the source code. You also have quick links to more detailed information on the moz.com site, but most of that is only accessible with a Moz subscription.

By the way, there are plenty of other toolbars that aim to provide similar information. Which one you prefer depends a bit on the information you need, as some have a slightly different focus on specific subjects (backlinks, Pagerank, social media, etc.). Some potentially useful toolbars are

- SEOquake toolbar by SEMrush
- SEO toolbar by SEO Book.com
- And more

SEORCH.eu Web Site

The web site at www.seorch.eu has a lot of good tools, especially for beginning SEOs. I've already talked about the keyword suggestion tool it includes, but on the home page, it even allows you to perform a full on-page SEO review for your web site (Figure 16-4), with the option, too, of scoring it with your keywords.

SEO analysis

Date: 14 July 2014
 Website: <http://joomla.org>
 Keyword: joomla
 Permalink: <http://seorch.eu/seo-check/joomla-org-2014-07-14-10-14-38.html>

Hey! Maybe you .. want to monitor Google Rankings and support SEORCH?

The 4 most important OnPage SEO factors:

Title - <title> Tag	HELP? ?
The <title> Tag has the right length: 54 characters.	SHOW ME ✓
The <title> Tag has the correct number of words: 9.	✓
1 keyword(s) found in the <title> tag.	✓

Figure 16-4. www.seorch.eu page analysis

SEORCH is a recommended tool to check your SEO efforts. It even allows you to scan your whole site if you want—all for free!

Screaming Frog SEO Spider

The Screaming Frog tool (www.screamingfrog.co.uk/seo-spider) is a very popular tool among both beginning and professional SEO consultants. It is free to use for sites up to 500 URLs, which should probably be fine for most of us. This tool is a desktop tool that simply gives extremely useful information about your site in a rather basic, but extremely useful, way, with the option to export it all to a spreadsheet.

This gives you the overview of all pages of your site, with a lot of relevant information (Figure 16-5). This especially concerns metadata information, page titles, response codes and times, heading information, etc. If you are serious about SEO, you will probably need this tool at some stage!

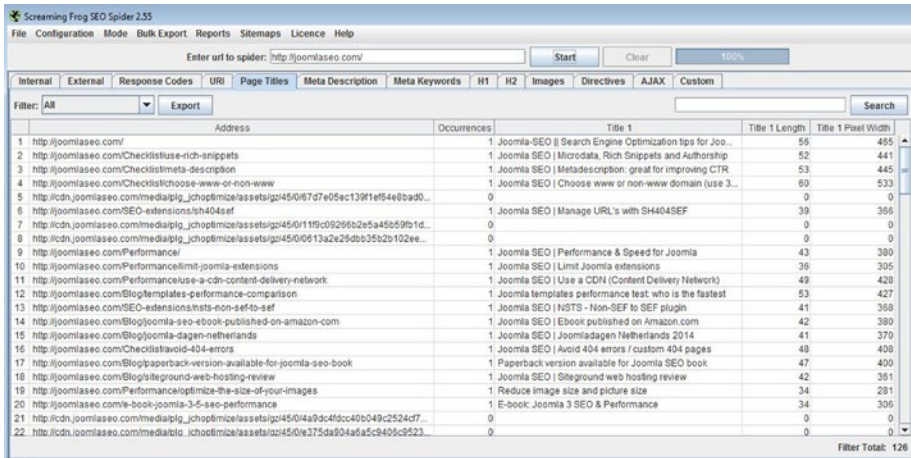


Figure 16-5. Screaming Frog SEO Spider site analysis

Site Audit with WebCEO

The tools mentioned so far are perfectly fine for a beginning SEO, and, especially if you only have one or two sites, they can give you most of the information that you need; they just take some more time. If you are going to get serious about SEO, however, you will probably need tools that will save you time by automating things. Reports that you would otherwise repeat manually every few weeks would be especially nice to schedule. There are many tools that can help you with this. In this section, I will show you one method, using **WebCEO**. This is a commercial service, but WebCEO also offers a free version that allows you to audit two sites, with some limited settings. So, to begin with, this tool could be a nice addition. I will mention some alternatives at the end of this chapter.

WebCEO aims to be an all-in-one SEO package. I already mentioned WebCEO as a tool with which to perform **keyword research**. Maybe you already signed up for it then, but if not, go to www.webceo.com and click the sign-up button. First, you have to submit at least one URL to create your first project (out of a maximum of two, for the free version).

Some of the features this service offers are

- Keyword research
- Optimization of internal links
- Optimization of landing pages
- Technical SEO auditing
- Tracking broken links
- Tracking backlinks to your site
- Finding opportunities for other backlinks
- Tracking social engagement
- Ranking tracking for your main keywords
- Integration with Google Analytics and Google Webmaster Tools
- And more

I have already discussed the keyword research options, but let's see what the other options are.

Technical Audit

The **Technical Audit** runs a rough check for technical issues, such as broken links and images, slow response time, server issues, 404 errors, etc., as shown in Figure 16-6.







FOUND ISSUES	
	Page Not Found Issues ▼
	Server Issues ▼
	Slow Response Time Pages ▼
QUALITY ASSURED	
	No Broken Images
	No Broken Anchors
	No Page Access Issues

Figure 16-6. Technical SEO audit

Obviously, all these errors should be fixed. All errors found are actual indications of pages or parts of pages that are really not working, and users will suffer from this. Clicking a red item will open a list with details about the issue(s) found. Make sure you solve these errors first. Some of this information can also be pulled from sources such as Google Webmaster Tools and other free sources, but this overview opens all issues in a single view, which is very convenient for troubleshooting, especially if you are less experienced.

SEO Analysis

SEO Analysis is a pretty massive check of many SEO-related features. The first check is the major one, checking for on-site issues such as duplicate metadata, correct length of metadata, URL structure, correct use of H1 headings, etc. (see Figure 16-7).

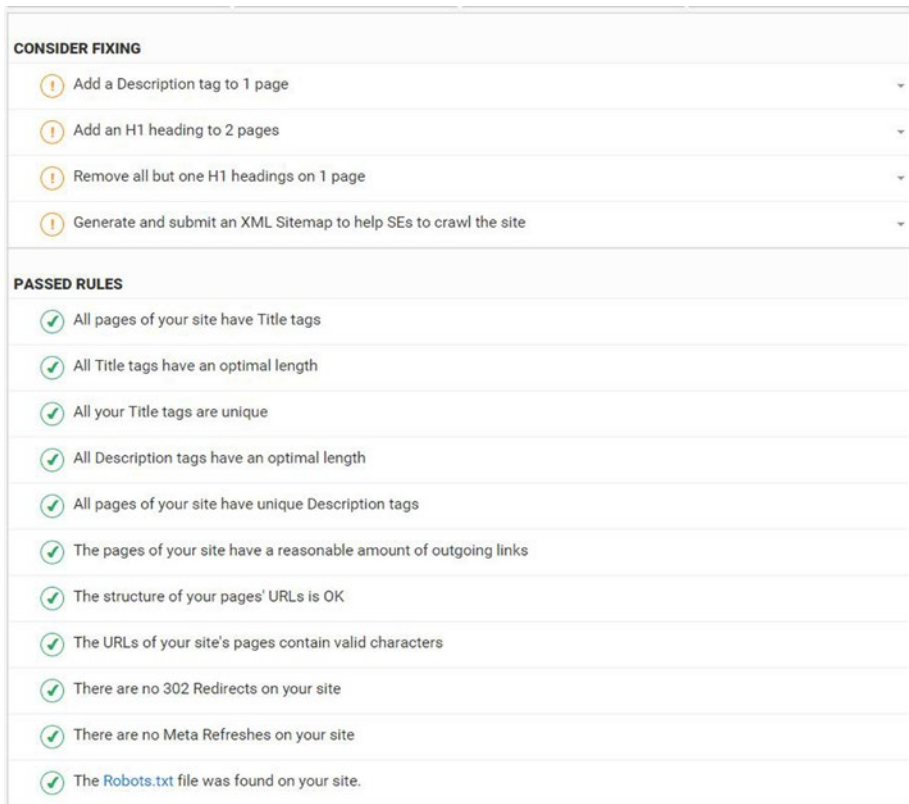


Figure 16-7. SEO Analysis

Again, all these issues should be fixed, if possible. In the example in Figure 16-7, it seems that, for some pages, a meta description is missing. Click the item to reveal the item in error (see Figure 16-8).

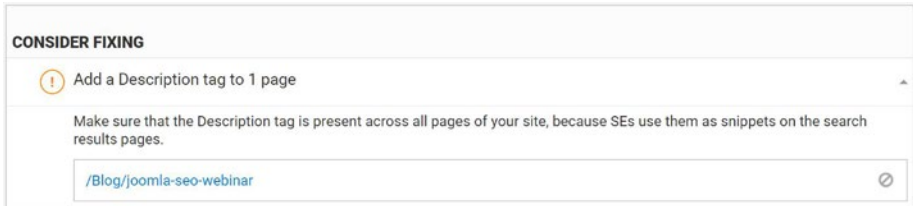


Figure 16-8. Missing meta description

There are a few other on-site checks that can be done from this section of the tool.

- **Google Pagespeed check:** Checks for performance issues using the Google Pagespeed tool (see Chapter 18 for more details on this)
- **Mobile SEO check:** Verifies that your mobile display has no issues

Landing Page Optimization

The **Landing Page optimization** is also a nice feature for SEO beginners. You can configure landing pages that you want to optimize for a specific keyword. The tool then checks if the keyword is used on all strategic places it should occur. The tool will not take into account any keyword variations, but apart from that, it performs a thorough check, as you can see from Figure 16-9.

Keyword Placement Map	joomla seo
Keywords in Title tags	!
Keywords in the H1 presence	!
Keywords in the H2-H4 headings	!
Keywords in the Description tag	!
Keyword presence in Body	1
Keywords in the URL	!
Keywords in the image ALT tags	!
Keywords in the image names	!
Pages with keywords as internal link text	!
Keyword density in Body text	<1 %

Figure 16-9. Landing Page optimization

Note that the check is on the exact keyword you have set to check for. In the example in Figure 16-9, I checked for **joomla seo**, but if I use **seo joomla** in a header, it will not recognize the term. So, as with any tool, you should know how to use it.

These checks can really help you to understand your site. It goes beyond the scope of this book to discuss all options WebCEO has to offer, but just try for yourself and play around with the tool. **Rank Tracking** can be an especially useful monitor, particularly if you worked for a client on an SEO project and you wanted to show him/her that rankings kept rising as result of your optimizations.

For the advanced features, you need a Pro subscription. This opens some of the currently inaccessible options and more thorough checks (such as more backlinks checked and a deeper search engine results page [SERP] depth for rank tracking, white label reports, etc.).

Other Tools for SEO Audits

Of course, there are other tools available with similar features. I will discuss the main candidates.

- **SemRush:** This is a very extensive toolkit for SEOs that also includes a Site Audit feature. One free project can be monitored with a free account.
- **Lipperhey:** This is a specialized tool mainly for detailed on-site SEO checks. A free account allows you to monitor one web site.
- **Majestic SEO:** One of the leading services for backlink checking. You can check your own sites for free by linking the tool with your Webmaster Tools account. For other sites, you need a subscription.
- **Moz:** Probably the leading company in the SEO world. Its tools are very nice (free for 30 days) but also pretty expensive.
- **Raven Tools:** Similar tools as Moz, but more affordable. This is probably a better choice for smaller SEO agencies, compared to Moz.

CHAPTER 17



Off-site SEO

At some stage, your site will be published and fully optimized. It is to be hoped that search engines will quickly add your site, with all its pages, to their indices. If you did a good job, this could go smoothly, but if you really want to aim high, you should also optimize your site outside the actual web site.

This process is called **off-site SEO**, as compared to **on-site SEO**, which we mostly worked on till now. Off-site search engine optimization (SEO) can have a tremendous effect, but doing it right can also be very time-consuming. Just check a few SEO agencies and look up their rates. Usually, they charge you per month, on a subscription basis, at rates of up to a few hundred dollars or euros a month and up, which is an indication of how time-consuming SEO efforts can become. However, some basic work from your side can get you going fairly well. This is what this chapter is about. For smaller sites in not too competitive areas, this will usually work really well.

The subject of off-site SEO is not Joomla-specific, of course; it is applicable for any web site out there. The nice thing is that a few initial steps can already help your efforts quite nicely. I will discuss a few, just to get you going.

Google Webmaster Tools

Whether **Google Webmaster Tools** should be treated under the topic of off-site SEO is questionable, but it is crucial that you do register, as this will help your efforts greatly. If you perform one step outside on-site SEO, it should definitely be setting up Webmaster Tools. This is a must for any webmaster.

So, the first stop after publishing your web site should really be Google Webmaster Tools. If you do not already have a Gmail account, get one now, and then sign up for Google Webmaster Tools at www.google.com/webmasters/tools. If this is the first time you use it, your list of sites being managed will be empty; otherwise, you will see them listed here. Assuming this is new to you, first click the **Add a site** button (Figure 17-1).



ADD A SITE

Figure 17-1. Add a site to Webmaster Tools

You will be requested to fill in your web site’s URL. Doing so will bring you to the verification screen shown in Figure 17-2.

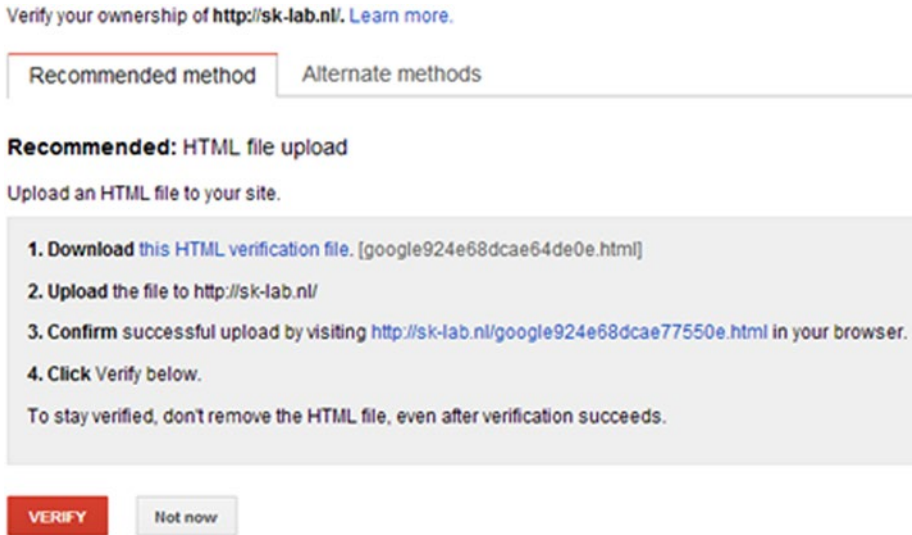


Figure 17-2. Verify your site

Just follow the instructions by grabbing the .html file supplied, uploading it to the root of your web site, and clicking the Verify button. You should receive a message such as that shown in Figure 17-3.

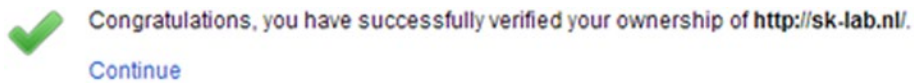


Figure 17-3. Verification is successful

There are also alternative methods, such as using your Google Analytics account, but the method described is probably the easiest to start with.

Setting a Preferred Domain

Now, the following may seem a bit strange to you, but my advice is to simply go along. If you just submitted your site as website.com, repeat the process for the same, but with www added—www.website.com—even if you configured your web site to be available only through one of these. The second verification takes a matter of seconds, because Google will ask you to use the same verification file (which is already accessible), so you do not have to retrieve and upload another file. Just two clicks and you are verified again,

resulting in two web sites in your Webmaster Tools. If you also have **SSL enabled**, you also have to verify the `https://` version, and any **subdomains** or **CDN-subdomains** must also be verified.

You will now receive an e-mail from Google confirming the site you verified, with steps to set up your configuration, which we will do right now.

Before continuing, make sure you work from the version that you will actually use (www or non-www), we will not use the other option anymore. We just leave it sitting there in our list.

Click the **Settings** icon at the top right of the screen and select **Site Settings** (see Figure 17-4).

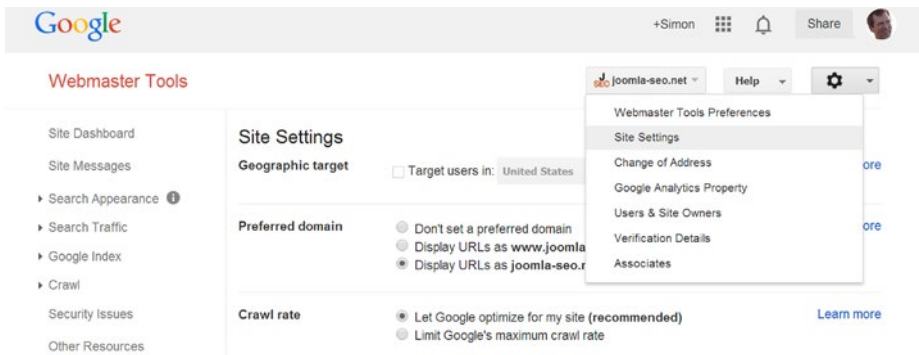


Figure 17-4. Site settings

You can now set the **Preferred domain**. Set it to the one you prefer to use and click Save. This step is especially useful if you have not technically enforced your site to use either one of the www or non-www options, but if you have, just perform this step. Every step you do that makes Google's life easier may be appreciated.

In addition, notice the **Geographic target** here. If your site is aimed at a specific country and this is not clear from your TLD (Top Level Domain, for domains such as .com, .net, .info, etc.), you can indicate it here.

Sitemap

There is one more thing that we are going to do now. If you followed along so far, you have created a `sitemap.xml` file in the root of your web site, or maybe somewhere else. You may have noticed that in the Webmaster Tools dashboard there is a question:

Why submit a sitemap? So, obviously, Google wants you to supply one.

Either click directly on the **Sitemaps** tab in your Dashboard or use the navigation on the left to navigate to **Crawl ► Sitemaps**. Currently, there is none, so click the button Add/Test Sitemap, shown in Figure 17-5.

ADD/TEST SITEMAP

Figure 17-5. Add a sitemap

You are requested to indicate the location of the sitemap now. You have the option to test it, to see if Google can reach and understand the file; otherwise, submit it right away.

Webmaster Tools will inform you of the number of URLs in your sitemap that you submitted (in blue). Initially, you only see the number of URLs submitted, but later, the indexed ones will appear in red (see Figure 17-6).



Figure 17-6. Sitemap submission successful

It seems that if you do this process very conscientiously with **correctly optimized** web sites, Google is quicker to index your sites. I usually have new sites with dozens of pages fully indexed in two days' time, and the sooner they are indexed, the better.

Remember to **resubmit** sitemaps every time your site changes, or make sure that your sitemap extension does it for you.

After initial setup, it may take one or two weeks before Google has checked your site in all details, but make sure to come back then, as you may receive very valuable information about your web site. Just investigate all options in the navigation menu for your site. I'll discuss a few.

HTML Improvements

Particularly interesting for new web sites is **Site appearance** ► **HTML improvements** (see Figure 17-7).

The screenshot shows the Joomla! Webmaster Tools interface. On the left is a navigation menu with options like Site Dashboard, Site Messages, Search Appearance (expanded), Structured Data, Data Highlighter, HTML Improvements (highlighted), Sitelinks, Search Traffic, Google Index, Crawl, Security Issues, and Other Resources. The main content area is titled 'HTML Improvements' and includes a sub-header 'Last updated Jul 14, 2014' and a note: 'Addressing the following may help your site's user experience and performance.' Below this are two tables. The first table, 'Meta description', shows 'We didn't detect any issues with the meta descriptions on your site.' The second table, 'Title tag', lists several issues with their corresponding page counts.

Meta description	Pages
We didn't detect any issues with the meta descriptions on your site.	

Title tag	Pages
Missing title tags	1
Duplicate title tags	0
Long title tags	0
Short title tags	0
Non-informative title tags	0

Figure 17-7. HTML Improvements

It could give you very useful information, especially about **duplicate content** detected, or at least duplicated metadata and titles. Note that, sometimes, the information is a bit slow to update, so don't panic right away. If the duplications are correctly spotted, make sure you fix them.

■ **Note** There may sometimes be a lot of warnings. Try to figure out the important issues, and don't worry too much about the small stuff. If you can solve issues, that's great, but not always really necessary.

Search Queries

The second very useful feature is the **Search Traffic** ► **Search Queries** option (see Figure 17-8).

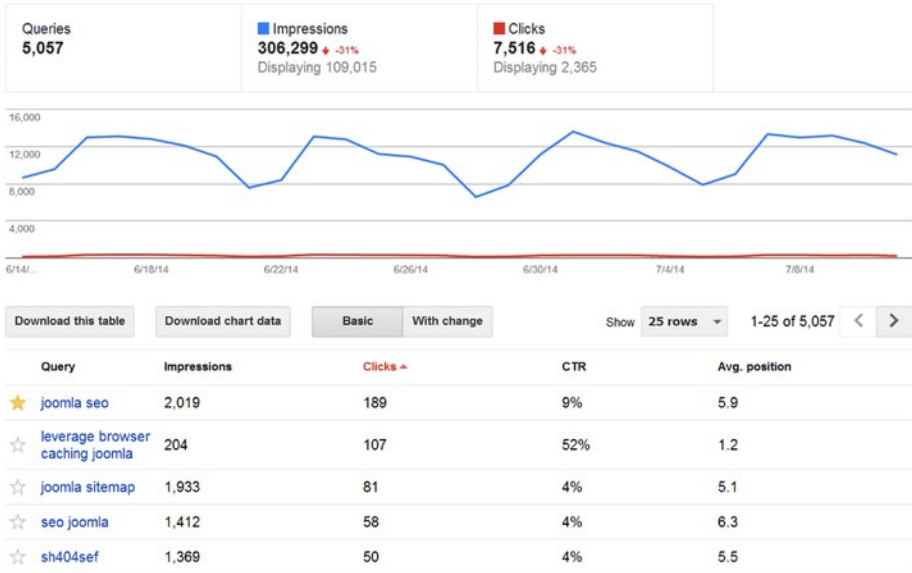


Figure 17-8. Search queries in the SERP

Especially since **Google Analytics** is not giving us all keyword metrics anymore (90% of traffic often shows as **Not provided**), this has become a very useful screen, with listings of keywords used to find your site, **Click Through Rate** (CTR), and position in the SERP. I find these options even more interesting than the data collected in Google Analytics. Particularly, it provides information about keywords that are **potentially** effective. If you order the list on the **Impressions** column, you may find some keywords for which you apparently are already showing up in the SERP but that are not yet receiving that many clicks. For example, you may find that for a very popular keyword, you are already ranking 13th. This may not get you much traffic yet, but now that you know this, you could optimize this keyword some more, possibly making it show up in the top 10.

In addition to that, the navigation menu shows valuable information about other things. Especially important could be

- **Links** to your site
- **Status** of pages in the index
- **Sitelinks**
- Presence of **structured data** (microdata)
- **Crawl errors**, such as 404 errors
- **Security issues** found (Google can even tell if Joomla is outdated)
- **Validating robots.txt** through the Robots.txt tester
- **Fetch as Google**, which allows you to check how Google sees your site and whether all resources that are needed for the site are actually accessible. Issues are usually related to blocked resources in robots.txt

Also, should Google at some stage find anything serious, such as malware, or bad practices that you performed, you may receive a notice about this, allowing you to act upon it.

Finally, even though Google has an immense market share, don't forget the competition: Bing and Yahoo!. They use the Bing Webmaster Tools at www.bing.com/toolbox/webmaster. Make sure to follow the same procedure with them. Even though they have a relatively small market share compared to Google, they might still deliver visitors to your site.

Google Analytics

Google Analytics is also something that you should definitely activate. While **Webmaster Tools** provides meaningful data about appearance and CTR in the SERP pages (user behavior **before** accessing the site), Analytics provides insight on how people behave **on** your site.

Besides telling you just how many visitors you have, it also tells you which are the most used pages of your web site, the **bounce rate** (visitors clicking the back button without browsing the rest of the site), but also where your visitors are coming from geographically, what OS they use, whether they use mobile devices, etc. Advanced users can also set goals and funnels to track **conversion rates**, and there is much more. Before becoming an expert in this matter, at least install the tracking code and check the data generated, or maybe you will use it in a few months, but at least your history data already starts to build.

To sign up, go to www.google.com/analytics, then fill in some details about your first site.

- Name for the web site: whatever you prefer, usually just the URL without the http://
- The actual URL
- Which industry best describes your site
- Reporting time zone

Having filled in this information, a code snippet will be generated, usually something like Listing 17-1.

Listing 17-1. Google Analytics Tracking Code

```
<script>
(function(i,s,o,g,r,a,m){i['AnalyticsObject']=r;i[r]=i[r]||function(){
  (i[r].q=i[r].q||[]).push(arguments)},i[r].l=1*new Date();a=s.createElement(o),
m=s.getElementsByTagName(o)[0];a.async=1;a.src=g;m.parentNode.insertBefore(a,m)
})(window,document,'script','//www.google-analytics.com/analytics.js','ga');
ga('create', 'UA-XXXXXXXX-XX', 'auto');
ga('send', 'pageview');
</script>
```

The UA-XXXXXXXX-XX part is the code that is specific to your site.

Now, you only have to make sure this code is embedded in your page. You can do this by editing your template `index.php`, but if you use a commercial template. They often cater to this. Some Joomla templates offer a field in which you just have to fill in the UA code; the rest of the code is then generated by the template.

As soon as you have installed this code, your statistics will start running, giving you, after a few days, something similar to what is illustrated in Figure 17-9.



Figure 17-9. Google Analytics dashboard

Some of these are nice indicators of how people experience your site.

- **Pages/Session:** The more pages people browse, the better, but it does not always mean much. For news sites, or blogs with particular interests, just one page view could still be fine. Often, viewers will spend a few minutes here, which is also an indicator.
- **Avg. Session Duration:** Higher is usually good.
- **Bounce Rate:** The number of people who just read this page and then use the back button in the browser. If people are just on the page to read your article and take their time, a high bounce rate is still not worrying, but if this is for a landing page, which should persuade people to fill in a form on the next page, this is an indication that you are not doing a good job.

As you see, none of these indicators by itself says much. You will have to combine the figures to make sense of them.

Google+

Google more and more tries to push its own social platform, Google+, these days. That is why, more and more, you see references to the platform in the search results pages. The most apparent one used to be the portrait pictures next to the descriptions in the SERP, though these have already been removed. Still, it might be a good idea to implement authorship, as it is a way to tell Google that you are the (original) creator of the content, which could help build Google's trust in you as a genuine content creator.

For this reason, I would advise you set this up as well. Just go to the Google+ signup page at www.plus.google.com/authorship and follow the instructions. You can verify your site by e-mail (with an e-mail attached to the domain), or by linking back from your web site to your Google+ page with a link configured with a `rel=author` attribute, like this:

```
<a href="https://plus.google.com/+MYPAGE?rel=author">My Google+ profile</a>
```

Apart from this, promote your page from within your Google+ page. Google+'s page could help your ranking. It may be a minor signal, but it helps.

Link-Building and Social Media

Getting links from sites other than yours is one of the best ways to show that you are being appreciated on the Web. Google definitely loves those links, but it is also quite picky as to whether those links are really relevant. Avoid links from spammy directories. The best links are links from blogs and sites that are about the same subject as yours. If you have a wonderful site with really unique and compelling content, those links may come naturally, because everyone wants to suggest this great site to their friends or colleagues. However, not all sites are like this, and you may have to help the process a bit, and for some sites, this is easier than for others.

In the instance that you have a web site about Joomla web design, you may already be a forum member of www.Joomla.org or of other Joomla forums, such as your template provider or host. Just make sure to set your signature on your blogposts with a link to your site. Also, you could register on resources.joomla.org as a Joomla-related business. It actually helps you to get noticed, while also creating a valuable backlink.

What are also great opportunities for easily creating links are YouTube, Slideshare, Vimeo, Instagram, Pinterest, etc. Usually, these services will allow you to post some comments next to the video, presentation, or whatever that you upload, and you can then insert a backlink to your site. Note that some of these links have a **NOFOLLOW** attribute, which makes them count less. If the actual video or presentation is relevant, it might also attract more visitors, so this works in two ways.

Other easy ways to create links from high-value sites are videos with your link in the description; your Twitter, LinkedIn, Pinterest, and Facebook profile; a presentation on Slideshare; etc. In case you have a choice between a Google platform and a non-Google platform, use the one from Google, as in the choice between Vimeo and YouTube, or even better, use both. Google should be impartial, but whether it really is 100%...As far as social media is concerned, use them all, but definitely use Google+. More and more, you see Google+ snippets in the search results. Social media is becoming more and more important as an SEO signal, but it goes beyond the scope of this book to dig much deeper into this.

Local SEO

For local SEO purposes, you should try to get links on reliable business directories. There are a lot of business directories that are pretty scammy, and you should avoid these, but correctly curated directories could actually help. Examples of legitimate directories are

- Google Places
- Yellow Pages
- Yelp
- Hotfrog
- Foursquare
- And more

Maintaining these business listings can become quite time-consuming, so you should consider using a service to do this for you. Some reputable (commercial) services can be found at

- www.yext.com
- www.moz.com/local

From Here

Additionally, posting regular updates on your web sites works miracles, as Google loves sites that keep spitting out new and interesting content. If possible, add a blog section to your site for this purpose or rewrite content if it needs updating. Also, staying up-to-date helps. You can often see that Google adds a last updated date in front of site descriptions in the SERP.

For further link-building, a great way of detecting possible link prospects is to look at your competition. Try to find out which sites link to theirs. You can find out by checking the backlinks to a site by entering the URL in the Open Site Explorer page at www.opensiteexplorer.org.

The SEO section of the book ends here, as we are getting more and more away from Joomla territory and more into the SEO world. SEO is a science by itself, so keep informed and read and blogposts, as the topic is constantly evolving.

Some SEO companies maintain blogs with useful information. The following are especially worth checking regularly:

- Moz.com Blog
- Majestic Seo
- Savvy Panda Blog
- Google Webmaster Central Blog



Joomla Performance and Speed

It may seem a bit odd to find a large section of a book about search engine optimization (SEO) dedicated to performance and speed, but to me, it makes sense. SEO is only partially about attracting visitors. The next step is to keep them on your site. Providing users with a good experience, by allowing them to browse fast through your site, is a crucial part of keeping them there. In addition, some research has shown that the time it takes to load a web site actually affects conversion rate, meaning that faster web sites sell more. You may have noticed yourself, if it takes too long to load a web site, you simply leave it and check the next site in the search engine results page (SERP). Even more important, Google has mentioned in one of its blog posts on the Webmaster Central site that site speed is considered one of the factors weighted for Google site ranking!

Generally speaking, you should aim to load your site in under two seconds. Two seconds is an interval that users expect to have to wait, so it should not stop them from browsing. As soon as your site is taking longer to load, your visitors might lose interest. Several studies have shown that conversion rates start dropping if pages take more than those two seconds.

I would advise you to take one of your web sites and really make an effort to accelerate its speed as much as possible. Even if it already loads in two seconds, you might try to make it load in under one second. Your site will feel super snappy and lively, which is something users will definitely appreciate.

Some of the measures I will discuss may seem trivial, but if you follow along, you will learn a lot. Once you have gained experience, it will be easier and easier for you to speed up your sites, and you will know which measures work for you and which ones don't. Small, simple sites, such as a portfolio or brochure site, may only require five minutes of configuration to make them twice as fast.

Be aware that some of the actions I discuss may sometimes break functionality, so, as with any major change you undertake, test and be prepared to go back to the original configuration, either by retracing your steps or by restoring your backup. The more complicated your site is, the more time-consuming it will be to speed it up.

Check Before You Begin

Before we dive straight into optimizing Joomla for speed, let's first see where we currently stand. If you understand what each optimization step does, you will learn which ones have a large effect and which have a minimal one. Your web site must be published and online to take advantage of these tools, but afterward there are several options for checking the speed of your web site. What is especially nice is that these tools not only rank your web site speed but also provide you with very useful tips. Just by following their advice, you may be able to find out what your issues are.

It can pay to check your web sites with more than one tool, as the different checks do not always agree 100% about the actions to be undertaken.

YSlow Browser Extension

YSlow can be installed in your Firefox or Chrome browser, using the Firefox and Chrome Play stores, allowing any site to be graded from there. Here is a screenshot of the extension at work (Figure 18-1).

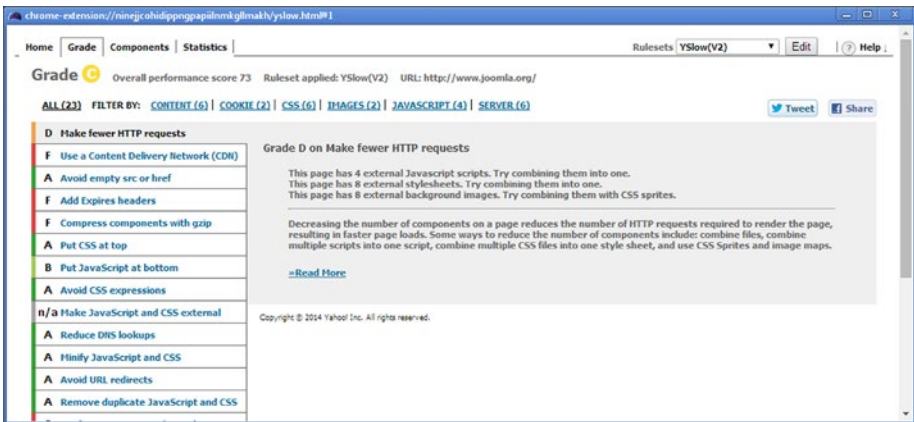


Figure 18-1. YSlow browser extension

If you do a lot of optimization on multiple web sites, this tool is very handy to have available from within your browser.

PageSpeed Insights

PageSpeed Insights (developers.google.com/speed/pagespeed/insights; see Figure 18-2) supplies both a mobile and desktop check, with different priorities. As this is a Google tool, you might make it your preferred check. Lately, PageSpeed has changed the way its checks work, adding more weight to the **Above the fold content**, something that is not that easy to comply with.

Google Developers

Producten > PageSpeed Insights

PageSpeed Insights 8+1 100+

joomla.org ANALYSEREN

Mobiel Desktop

68 / 100 Overzicht van suggesties

Dit moet worden gecorrigeerd:

- Gebruikmaken van browsercaching
 - › Informatie over de oplossing voor dit probleem
- JavaScript en CSS in inhoud boven de vouw die het weergeven blokkeren, verwijderen
 - › Informatie over de oplossing voor dit probleem
- Comprimeren inschakelen
 - › Informatie over de oplossing voor dit probleem

Figure 18-2. PageSpeed Insights

GTmetrix

GTmetrix (Figure 18-3, gtmetrix.com) collects data from both PageSpeed and YSlow in one nice overview. Its tips are usually very helpful.

Page Speed Grade: **(74%)** ↓ **C** YSlow Grade: **(70%)** ↓ **C** Page load time: 7.80s
 Total page size: 732KB
 Total number of requests: 43

Breakdown

Page Speed YSlow Timeline History

RECOMMENDATION	GRADE	TYPE	PRIORITY
Leverage browser caching	F (19)	Server	High
Specify image dimensions	F (29)	Images	High
Combine images using CSS sprites	E (53)	Images	Medium
Optimize images	D (65)	Images	High
Avoid CSS @import	C (70)	CSS	Medium
Enable gzip compression	C (75)	Server	High
Specify a Vary: Accept-Encoding header	C (75)	Server	High

Figure 18-3. GTmetrix performance test

It retrieves data from both PageSpeed and YSlow in one handy overview.

WebPagetest

The WebPagetest site at www.webpagetest.org also supplies useful data (see Figure 18-4).

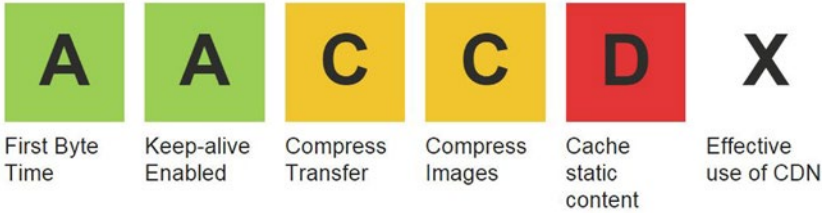


Figure 18-4. WebPagetest speed check

Pingdom

Pingdom tools at tools.pingdom.com (Figure 18-5) checks your site from random locations. You can use it on an ad hoc basis, but what is also nice is that you can sign up for permanent monitoring, up to once per minute. For one web site, you can do so for free, which gives you a useful overview over time.



Figure 18-5. Pingdom tools

Pingdom checks your web site from multiple locations around the globe, so it might come especially handy when you want to know how your site loads elsewhere, or when you want to verify the effect of a content delivery network (CDN).

So, now that we know where to begin, let's check what we can do to make improvements.

Web Hosting Provider

This may not always be the easiest parameter to change, but choosing the right host can have massive consequences. If you are serious about the speed of your web site, avoid cheap (or free) hosts, and select one carefully. With two different hosts, the same web site can differ in speed greatly. Differences of over 100% are not uncommon. Before choosing a hosting company, at least check the following:

- Online reputation
- Support
- Joomla specialization
- Up-to-date software and servers (especially PHP and MySQL)
- Fast servers
- Location of server (Asian web sites should be hosted in Asia, etc.)
- Whether it provides advanced techniques, such as Varnish caching, memcache, mod_pagespeed, etc.

Most hosts will claim that they are great at providing all of the above, but double-check. A perfect host does not exist, but, happily, there are quite a few good ones, usually depending on how much you are willing to pay. Some smaller, locally operating hosts have very good service, but to help you choose, here are some of the more reputable larger hosts operating worldwide:

- SiteGround.com (current host of the Joomla demo at www.demo.joomla.org and www.joomla.com)
- Rochen.com (former host of www.joomla.org)
- Bluehost.com
- CloudAccess.net
- DreamHost.com
- Rackspace.com
- Arvixe.com

Mean and Lean

To make Joomla load fast, your site should be mean and lean and only load what is required to satisfy the needs of your visitors. You should not only pay attention to the size of the items needed to load the site but also to the number of separate files needed. Every separate file requires a separate **HTTP-request** from the server. Generally speaking, making a web site load fast requires the following:

- Reduction in size (KBs)
- Reduction in number of HTTP requests
- Reuse (cache) of material that has already been downloaded or generated
- Allowance for resources to load simultaneously instead of sequentially

One thing to constantly keep in mind when building Joomla web sites is that every addition to Joomla may result in more time needed to load your web site. This does not only count for extensions but also for

- Images
- Fonts, both web fonts and @fontface embedded fonts
- Social media scripts (!)
- Templates
- Captchas
- Embedded videos
- JavaScript libraries
- Extensions
- Plug-ins
- And more

Building your template already requires that you start to think about what is really needed, such as whether you really require images for your backgrounds and buttons or whether this could also be achieved using CSS3 techniques.

In addition, creating content also requires that you think in terms of efficiency. For example, do you really need to use three images in your article, or is one sufficient, or perhaps an image half the size?

Of course, there will always be a need for compromise. The fastest web site may also be a boring web site, showing only text, without additional resources apart from a CSS file. But when you are building a web site for a photographer, he/she may insist on a large, fancy slider, and you will simply have to install it to please your client. You still have the possibility of selecting a well-designed and efficient extension, of course.

Some even go as far as de-installing Joomla core components. If you do not use web links, get rid of it. Personally, I would rather leave core stuff untouched, but simply unpublishing the component also works well.

A special mention should also go to resources loading from third-party sources, such as **Twitter/Facebook/Pinterest** feeds, **Google web fonts**, etc. If you cannot avoid them, maybe you can at least move them away from the home page and only show them on the relevant pages. As for **web fonts**, sometimes it is more efficient to serve the fonts from your own web site, using **@fontface** embedding, than to serve them from Google web fonts (on occasion, this can really slow sites down).

Like SEO, performance should always be in the back of your mind. Use and install what you really need but, if you can, leave stuff out.

Joomla Caching

Caching is the mechanism by which to store certain views for users, so that these don't have to be retrieved from the database every time. These views are actually temporary files stored in one of the folders that Joomla uses: the cache folder. You can check the file system, to see how these files look like and what they contain.

- **com_*** files store views for components
- **mod_*** files store views of modules
- **page** files store views of full pages

The main advantage that caching offers is to make sites load faster, because they take a shortcut to already generated information. Indeed, the speed difference can sometimes be huge on some sites, though on others, the difference may be far less. It all depends.... You will only know for sure what the effect on your site will be by testing before and after caching.

The problem with stored views, however, is that information in Joomla may have changed since the last stored view was created. Particularly for sites with frequently changing content, this can be a disadvantage, for example, on news web sites or sites on which users create content and want to see how it looks on the front end right away. This may be a reason not to use cache, though there are ways to overcome this issue, as I will discuss in the "Extensions" section, later in this chapter.

Another potential problem is that some functionality simply breaks with cache switched on. The **System - Page Cache** plug-in is especially vulnerable to this effect. Other examples of affected functionality could be feeds or Captchas not working or checkout carts of web shop extensions.

What you should also know is that the Joomla cache options are only active for guest users. Users signed in to Joomla (front end) do not receive the cached views.

Joomla has quite a few possibilities built in to support different levels of caching. They all work differently, so let's go through them one at a time.

Global Cache Settings

The first parameter I will discuss is again in the **Global Configuration** screen. Go to the **System** tab, where you will find **Cache Settings** (see Figure 18-6).

Cache Settings

Cache *	OFF - Caching disabled ▼
Cache Handler	File ▼
Cache Time *	15

Figure 18-6. Global cache settings

By default, cache is set to **Off - Caching disabled**. The alternatives are **On - Conservative caching** and **On - Progressive caching**. You should be fine by selecting the **Conservative** option. Progressive caching does not give you a lot of additional performance, while sometimes reporting problems are that do not occur with Conservative caching.

The **Cache Handler** is set to **File**. Leave this on. Changing this is for advanced users. The only configured alternative is currently **Cache_lite**, but there are also other possibilities. This requires a specific server setup, plus you will have to manually update the `configuration.php` file in your Joomla root folder. An example is **memcache**, which speeds up database lookups (not supported on every host except the major ones). Memcache is a bit too technical for this book to expand on, as implementation also depends on the host, but if you're up for it, implement it, as it is really fast.

Finally, you can set the **Cache Time**. Feel free to customize it to your needs. Small static sites could use very long cache times, while others might have to set it shorter. The default of 15 minutes will suffice for most sites.

What this cache option actually does is create stored views of **Components**, the main content area for Joomla. These views are stored in `/cache/com_*` files.

Module Cache

Module cache is used for views of your modules. The default setting is to use the same value as the one set in Global Configuration (Global cache = off: modules cache = off and vice versa), but module cache can be **overridden** at a specific level, per module. Go to **Extensions ► Module Manager** and enter one of your modules. You will find the cache settings for the module in the **Advanced** tab (see Figure 18-7).

The screenshot shows the Joomla! Module Manager interface for configuring a module menu. The title is 'Author Menu'. There are four tabs: 'Module', 'Menu Assignment', 'Module Permissions', and 'Advanced'. The 'Advanced' tab is selected. The settings shown are:

- Menu Tag ID: (empty text input)
- Module Class Suffix: (empty text input)
- Caching: A dropdown menu with 'Use Global' selected (highlighted in blue), and 'No caching' as an option.
- Cache Time: (empty text input)

Figure 18-7. Module cache settings

If Caching is switched to On in Global Configuration, you can override it here and set it to No caching. You cannot do it the other way around. If Global Configuration uses no caching, you cannot switch it on in modules.

Next to that, you can also tune Cache Time (seconds are used here, while Global Configuration uses minutes).

Module cache can sometimes be useful if only a specific module encounters issues with the global cache setting, leaving the rest of the site nicely cached. An example could be a **breadcrumbs** module that fails to reflect the correct path when cached. If you experience this, switch off caching for this module.

Cached module views are stored in `/cache/mod_*` files.

System - Page Cache plug-in

Some Joomla! administrators are not aware that there is also a separate **cache plug-in**. You cannot configure it through any of the Joomla! administrator menus, but you will have to go to **Extensions ► Plugin Manager** and look up the **System - Page Cache** plug-in (before Joomla! 3.4, it was simply called System - Cache).

By simply enabling it, you switch on the plug-in. This plug-in stores views for entire pages, on top of the Component and Module view cache, so they work on top of each other. Switching on this plug-in can make quite a difference, even though you might already have switched on the global cache. Note that it is also known sometimes to cause issues and conflicts, so make sure to check your site carefully after switching it on.

There is one configurable option in the plug-in: the **Use Browser Caching** option (see Figure 18-8).



Figure 18-8. Use Browser Caching setting

Leave it switched off. The description is a bit misleading, and it actually only works for users who come back to the same page within a browsing session. It can also cause some issues.

Always check if your site still works correctly after switching on this plug-in, as it could quite often be the cause of a problem.

Cache Cleaning

At times, you may want to force your new content to be displayed on the site earlier than your predefined settings. To do so, you must clear the views that are stored in the **cache folder** in the Joomla file system. With cache switched on, this folder could have content such as that shown in Figure 18-9.

















 _system	7/14/2014 10:30 AM	File folder	
 com_content	7/14/2014 10:30 AM	File folder	
 com_languages	7/14/2014 10:30 AM	File folder	
 com_plugins	7/14/2014 10:30 AM	File folder	
 com_templates	7/14/2014 10:30 AM	File folder	
 mod_articles_archive	7/14/2014 10:30 AM	File folder	
 mod_articles_category	7/14/2014 10:30 AM	File folder	
 mod_articles_popular	7/14/2014 10:30 AM	File folder	
 mod_breadcrumbs	7/14/2014 10:30 AM	File folder	
 mod_custom	7/14/2014 10:30 AM	File folder	
 mod_menu	7/14/2014 10:30 AM	File folder	
 mod_search	7/14/2014 10:30 AM	File folder	
 mod_tags_popular	7/14/2014 10:30 AM	File folder	
 mod_weblinks	7/14/2014 10:30 AM	File folder	
 template	6/23/2014 11:39 AM	File folder	
 index.html	6/11/2014 7:46 AM	Firefox HTML Doc...	1 KB

Figure 18-9. Cached files and folder in the file system

You can easily recognize the views for components, modules, and pages here, and when you check them in a code editor, you can see that the HTML inside corresponds to the expected content (the line `JEXEC or die` prevents them from being directly accessible).

To clean up your cache and have a fresh site, go to **System ► Clear Cache**. You will receive a list comparable to the list of files in the cache folder. Check all, and click **Delete** (see Figure 18-10).

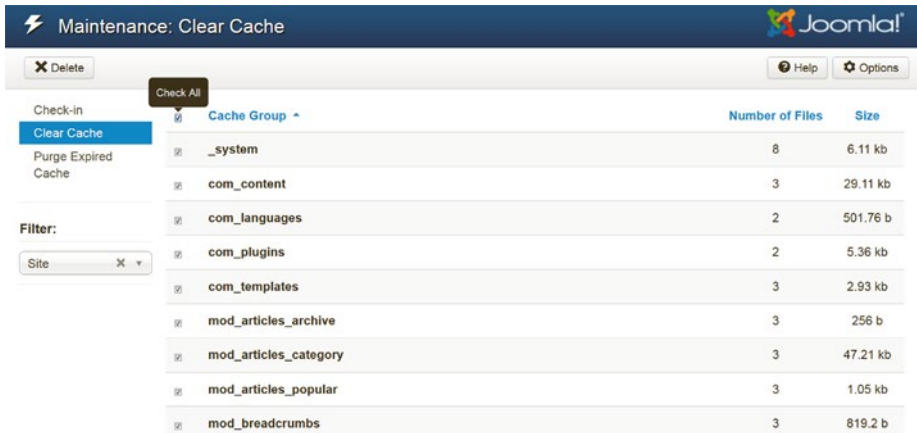


Figure 18-10. Cache cleaning

Also, go to **Purge Expired Cache** and clear this. This should not affect display of items on your site, as it removes cached files that have expired.

NoNumber Cache Cleaner Plug-in

Cleaning cache can become quite tedious sometimes. When you want to view the changes for an updated article directly, you have to go through quite a loop.

- Change in the article.
- Close the article.
- Go to **System ► Clear cache**.
- **Purge Expired Cache**, just to be sure.
- View your changes.
- New correction, new loop...

Particularly if you have to do this a few times, it can become quite tedious. In such cases, I recommend the **Cache Cleaner plug-in by NoNumber** (www.nonumber.nl/extensions/cachecleaner). It gives you a handy little button at the bottom of your administrator screen, available at all times, even when you are editing an article (see Figure 18-11).

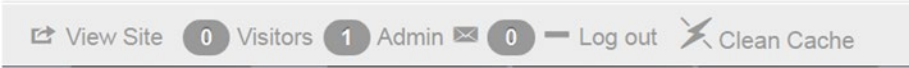


Figure 18-11. NoNumber Cache Cleaner plug-in

Now there is no need to leave the article you are currently editing. It clears all your cache with one click.

In the Pro version, you can even set it to periodically clear the cache, or after every save action. This can be extremely useful, as now you really do not have to worry about cached content. Especially when multiple non-tech users are allowed to edit and create articles, this can become even more useful, as the phenomenon “cache” may be completely unfamiliar to them. What is also very nice is that Cache Cleaner can clear third-party cache as well, such as cache from the JotCache extension, SiteGround Cache (if you use the SuperCacher options), and even cache on the CDN side (for some popular ones). Note that on large sites, there can be a performance issue if cache is frequently cleared automatically. Just test this.

JotCache

Apart from the NoNumber Cache Cleaner, there are more extensions that deal with cache. **Cache Control** (by **CrossTec**, crosstec.de) is a small plug-in that allows users to switch off caching for specific components, but this may be a bit limited.

For advanced purposes, **JotCache** (www.jotcomponents.net/web-programming/jotcache) is an extremely useful extension. With JotCache, you can exclude components, modules, menu items, module positions, and URLs, and you can even exclude specific browsers, leaving your site fast and with functionality intact. You can tune this on a detailed level. Particularly for components, you do not have to exclude a whole component. You can limit this to specific views of the component. In the event you have an issue with a category view, you can set JotCache to exclude **com_content** ► **category view**, leaving the other views cached.

When you install JotCache, you will receive both a plug-in and a module. First, you will have to enable the plug-in and set some basic options, but most of your configuration will have to be performed from the JotCache component (Figure 18-12).

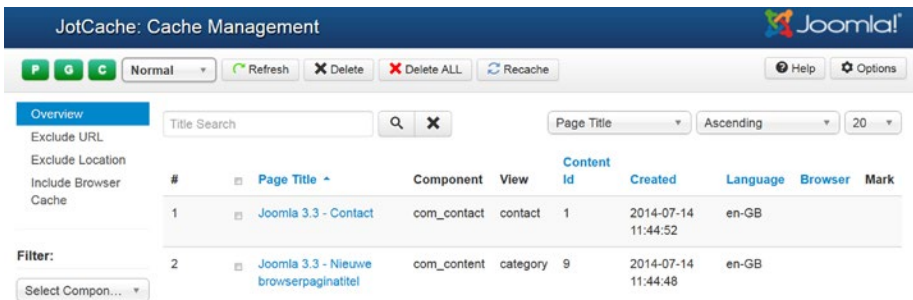


Figure 18-12. JotCache component

The following options are especially useful:

- **Exclude URL:** This allows you to exclude various components, or even specific views of these components. So if you find a component that does not like cache, check the check box, and it will no longer be cached.
- **Exclude Location:** Exclude location allows you to exclude different module positions in your template. Also, it is very cleverly developed.

If you have issues with the way Joomla handles cache, JotCache is highly recommended.

Leverage Browser Caching

The Joomla caching options discussed so far create stored views on the server. Leveraging browser cache allows you to reuse the resources of the web site that have already been downloaded on your PC in earlier visits, even another page from the same site. This can even be the HTML, but also CSS, JavaScript, fonts, and other static resources. The advantage of this mechanism is that these resources are already on your computer, so there is no need to use further bandwidth while displaying the web site.

What is also nice about this is that it is usually quite safe to use, and it will usually not break any functionality of your web site. The only downside is that some resources are not refreshed automatically, such as a CSS file that has been changed on the server. Most visitors will usually be able to solve this by clearing up the cache in their browser, but how serious this is again depends on your situation. (You can bypass this by using version numbers in your CSS, such as `/template/css/screen.css?v=2`. If the file changes, just change the version number.)

So how do we activate this? There are two prerequisites, which most probably are already met:

- **URL rewriting** through `.htaccess` should be enabled (you probably already did so).
- Your server should support Apache's `mod_expires` module (almost all hosts do). To prevent issues in case it is not supported (or on localhost), you should wrap the code in an `if` statement, as shown.

For each different file type, you then have to list an expiry interval, indicating how fast each should be refreshed: short for files that change often, such as HTML, long for files that hardly change, such as JavaScript and CSS.

This could result in a snippet like the one in Listing18-1, to be included in your `.htaccess` file.

Listing 18-1. Leverage Browser Caching in .htaccess

```

<IfModule mod_expires.c>
FileETag MTime Size
AddOutputFilterByType DEFLATE text/plain text/html text/xml text/css
application/xml application/xhtml+xml application/rss+xml application/
javascript application/x-javascript
ExpiresActive On
ExpiresDefault "access plus 1 seconds"
ExpiresByType text/html "access plus 600 seconds"
ExpiresByType text/css "access plus 604800 seconds"
ExpiresByType text/javascript "access plus 216000 seconds"
ExpiresByType application/xhtml+xml "access plus 600 seconds"
ExpiresByType application/javascript "access plus 216000 seconds"
ExpiresByType application/x-javascript "access plus 216000 seconds"
ExpiresByType image/x-icon "access plus 2592000 seconds"
ExpiresByType image/jpeg "access plus 2592000 seconds"
ExpiresByType image/png "access plus 2592000 seconds"
ExpiresByType image/gif "access plus 2592000 seconds"
</IfModule>

```

A default is available for file types not mentioned (access plus 1 second). You can expand this even further, if you have special file types, such as audio, video, @fontface embedded fonts, etc.

You also see some additional lines.

- **ETag:** Informs the browser that already downloaded images should not be reloaded
- **Expires headers:** Similar to ETag but more specific with regard to file types
- **AddOutputFilterByType DEFLATE:** Minifies HTML by removing blanks

Browser caching is very efficient, and your grades in Google PageSpeed will go up quite a bit as a result of using it. What is especially nice is that activating this feature in .htaccess hardly has any downsides, as long as you include the wrapping `<ifModule>` statement.

If you have **jbetolo** installed (discussed later), you can even automate generation of this code in .htaccess, by using the **Patch** button in the plug-in. Advanced users will probably want to set the code exactly to their own specifications, however.

Gzip

Joomla has **Gzip** support built in. This is the mechanism whereby content from your web site is zipped. The zipped file is sent to your browser and is unpacked and displayed. Currently, all browsers, including older IE versions from at least version 6, are supporting this, so there

should be nothing to hold you back from using it (as long as your server supports the `mod_gzip` module in Apache). How does it work? Simply switch it on (see Figure 18-13).

Server Settings

Path to Temp Folder

Gzip Page Compression Yes No

Figure 18-13. Gzip compression parameter

Go to **System > Global Configuration > Server Settings** tab.

Set **Gzip Page Compression** to **Yes** and save. The setting is immediately active after it is saved.

You might want to check whether the compression for your site actually works and what the compression rate is. You can do so by going to www.feedthebot.com/tools/gzip and entering your URL (see Figure 18-14).

gzip test

Compressed?	Yes
Compression type	gzip
Page size (Bytes)	30,488
Compressed size (Bytes)	7,583
Saving (Bytes)	22,905
Compression %	75.1%

Figure 18-14. Test Gzip compression

You see that the size in this example is reduced 75%, which can make quite a difference, especially for larger sites.

Reducing Image Size

If your site has a lot of images, **reducing the size of your images** is crucial. First of all, you should make sure that your images are saved in the lowest quality possible, without visually being affected. For `.jpg` images, a quality of 60% can save you more than 50% in file size, compared to the original one, while the image still looks good. You can set the

image quality from within the **Save for Web and Devices** feature in Photoshop. Also, make sure you take the correct format for the image (.jpg for photographs; .png for transparent images, logos, and uniform color; and .gif for small low-quality images).

For creating .jpg images, you also have the choice of **progressive** or **baseline**, which is something you can also set while using the **Save for Web and Devices** feature in Photoshop. Particularly with slow connections, you may notice the difference: baseline images load from top to bottom, while progressive images start with a slightly blurry view, getting sharper as the image loads (these .jpgs are built of multiple layers). For user-experience purposes, Google favors progressive images, but these may be slightly larger in size. Basically, it is a bit academic which type of image is better.

Lossless Compression

Something you often see when you are running your page speed tests is the term *lossless compression*. Images often carry all kinds of unnecessary metadata, such as the date of a picture, camera model, etc. Even if you save an image in Photoshop with the Save for Web and Devices option, not all this data is stripped off.

The **GTmetrix** web site has a nice built-in feature for you. When it finds images on your web site that could be further optimized, it not only notifies you, but a link to an optimized version of the image is supplied as well (see Figure 18-15).



Figure 18-15. Download optimized images from GTmetrix

Download the optimized version and use it to replace the original image, and you are one optimization step further.

Online tools that are there especially for this purpose are **Smush.it** (from Yahoo!), **TinyPNG**, **TinyJPG**, and **Kraken.io**. **TinyPNG** (Figure 18-16) especially does an excellent job of compressing files, but, unfortunately, only for .png images.

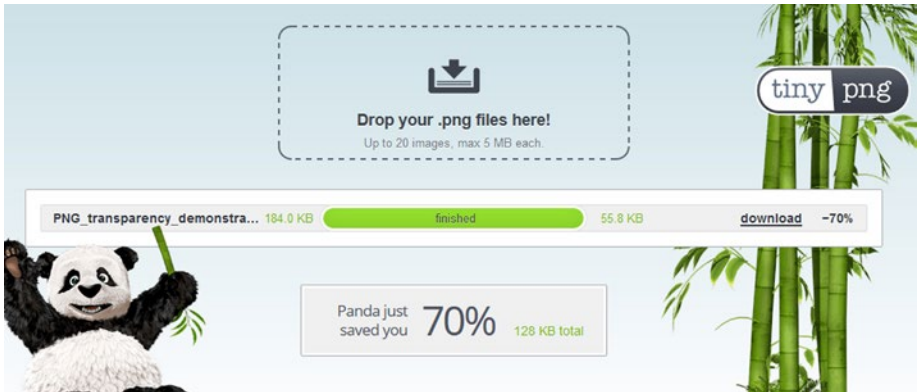


Figure 18-16. PNG compression using TinyPNG

Luckily, it has also provided TinyJPG for .jpg images. The easiest to use, however, is the Kraken.io web interface. You can simply offer .png and .jpg files together and download them all back in one zip file.

If you want to automate this process for your Joomla web site, you might take a look at the extensions **jbetolo** and **JCH Optimize**. These are extensions that have many advanced features, especially for CSS and Javascript optimization, but they also carry a little tool to scan your Joomla sites and optimize the images with Smush.it (see Figure 18-17).

Image optimization with Smush.it(TM)

-
- Replace the original files with the smushed version. Use with care. If returned file is empty due to failed call to Smush.it, we will leave the file intact.
- Search and process image files recursively in given directory structure. Use with care. If the files are many your request might time out. Also make sure not to breach your smush.it terms of use.
- Post fix of the generated files, applies only if replace above is not selected.

Figure 18-17. Compress images with Smush.it from jbetolo plug-in

Use this with care. Back up the images first. I will discuss these extensions in more detail later in the chapter.

Correctly Sized Images

What is also important regarding images is to use correctly sized ones. This is always an effort, especially in a content management system (CMS) such as Joomla, in which inexperienced users may have the possibility of uploading images, sometimes directly from their camera. The best solution is to use exactly sized images and not the editor to resize images. An alternative solution, however, would be to make it impossible for users

to upload large images altogether. Generally, it would be okay to limit the size of images to a maximum width of 600px. Even if users employ the editor to further resize an image, at least the size will have been restricted quite considerably already.

One of the best tools to achieve this is the **JCE editor** (www.joomlacontenteditor.net) combined with the **Image Manager Extended** plug-in (see Figure 18-18). It allows the setting of a default maximum width and height. The image is resized before uploading to the server, so this would be perfect.

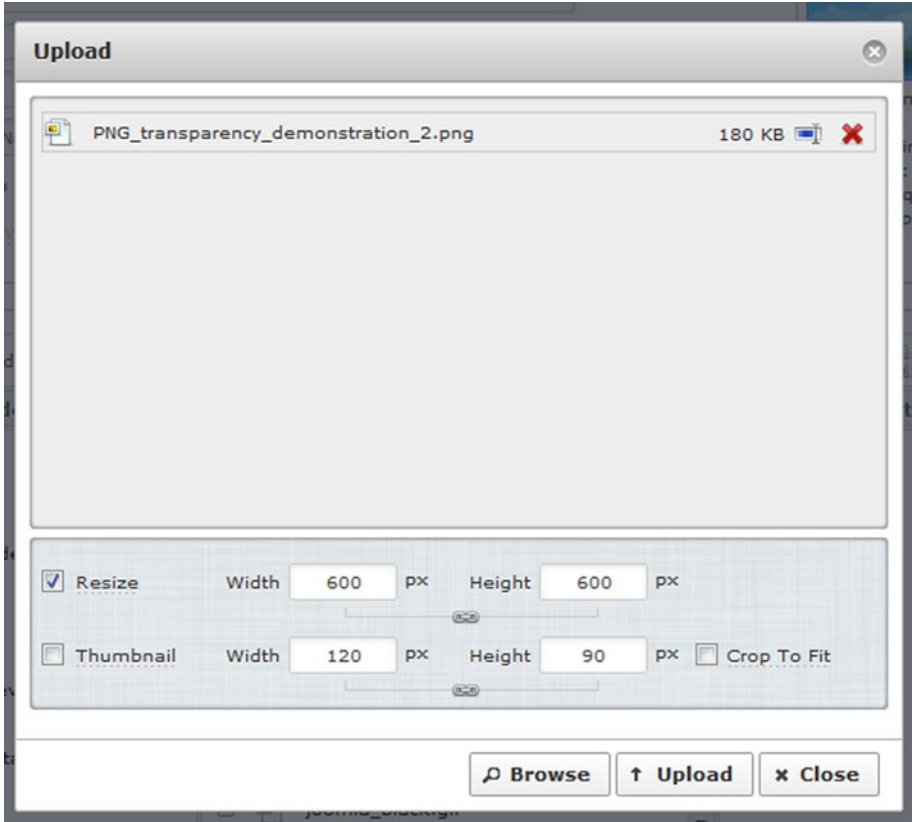


Figure 18-18. *Resize images with JCE*

There are also several plug-ins in the Joomla Extensions Directory that offer all kinds of image resizing options. Some are for very specific purposes. Some also simply force all images within a certain size. Others simply use an image that is resized and cached on the fly, so that it does not matter how large the image the user has uploaded is; visitors only see the optimized cached version.

A very useful plug-in that specifically aims to set different maximum image sizes for different devices is **XT Adaptive Images**. It allows desktop users to receive the normal large image, while users on smartphones receive of a configurable max-pixel-width, such as 480px wide.

Sprites

With **sprites**, you can combine a set of background images into one file, usually .png, for transparency purposes. Using advanced CSS positioning techniques, you can prevent loading quite a few different images, but load one instead.

This technique is especially useful for background images, so it would be best to optimize this when building your template. For normal images, this technique is less useful.

Experienced CSS users can hand-code this, but the extension JCH-Optimize (www.jch-optimize.net) can also do this automatically. I will discuss this plug-in in more detail later in this chapter.

Base64 Image Encoding

Especially for very small icons and images, you could consider to transform them using **base64 image encoding**. With this technique, you can use a large string representing your icon in your HTML and CSS. A little icon could be represented as follows:

```
<img src='data:image/png;base64,iVBORw0EUgA.....AAABJRU5ErkJggg=' />
```

This is an advanced technique that you should not just use for all your images. Actually, the code used to represent an image can become huge. As it takes memory and processing power from your browser to create the image, it can take a lot of time to process and display the image. This can really be full seconds for larger images! For really small images, though, the bonus of saving an HTTP request could be such that it is worthwhile to use this technique.

Using the technique is very easy, actually. There are many online encoders available, such as www.base64-image.de. Also, if you use the optimization plug-in **ScriptMerge** by **Yireo**, it can process these images automatically. You set a maximum size for any image to be selected, such as 2KB, and all images smaller than that will be processed as encoded images. Again, use with care!

Lazy Loading

Lazy loading of images is a technique that aims to upload images only when the user actually scrolls to a lower section of the site where the image is located. Images above the fold will be immediately loaded, but images under the fold will be initially replaced by a placeholder (this placeholder will create a whitespace for the image of the size of the actual image, to prevent collapsing text around the image).

When you actually scroll to this section of the site, the image will be downloaded from the server. This means that the initial size of the resources for the site will be less. Only a small bit of JavaScript has to be executed for this mechanism to work. Coding this yourself is difficult, but an extension comes to help: Lazy Load for Joomla (joomla-extensions.kubik-rubik.de/11fj-lazy-load-for-joomla).

Optimize CSS and JavaScript

CSS and **JavaScript** files more and more take up a large chunk of your web site's KBs. Especially with advanced templates and extensions, many separate CSS and JavaScript files are loaded, but Joomla core alone already loads quite a few files. Joomla by default loads the JavaScript library **jQuery**. While the separate files may not be that large, they add up quite a bit to the number of HTTP requests. You can clearly see this if you take a large site, run it through www.tools.pingdom.com, and check the **Waterfall** tab. You can see many separate files, all taking just a few milliseconds to load, but if these are loaded sequentially, this affects the speed of the site. So, what can we do about this? There are a number of options.

Reduce Your Template Files

Before we start optimizing our CSS and JS files, we might just check whether we actually need all those files. By default, Joomla loads quite a number of files. While we really need some of these, there are also a few that we might not really require. Especially if you build your own templates, there are some techniques to help you get rid of unwanted files. Note that if you use commercial templates, it might not be so easy to follow this step. Okay, let's first look at the files Joomla often loads. You might even see these files when you check the source code of your web site.

- `jquery.min.js`
- `jquery-noconflict.js`
- `jquery-migrate.min.js`
- `caption.js`
- `bootstrap.min.js`
- `mootools-core.js`
- `core.js`
- `system.css`
- `general.css`
- And more

You may have extensions installed that load their own CSS and JS files, sometimes even on pages that do not show any of the extensions' output. I will not go into much detail about each of these separate files, but if you find that one of them is not needed,

you can remove it. One of the files that you mostly really do not need is Mootools. For most functionality in Joomla, this has been replaced by JQuery. It is also a relatively large file, so getting rid of it usually saves some KBs (and one HTTP request). You can use a command like the one in Listing 18-2 for that (insert it in the <head> section of your template's `index.php` file).

Listing 18-2. Remove JavaScript Files in `index.php`

```
// Call the JS libraries (needed to unset them)
JHtml::_('behavior.framework');
JHtml::_('bootstrap.framework');
JHtml::_('jquery.framework');
JHtml::_('bootstrap.tooltip');

// Unset the JavaScript file(s) you want to get rid of
unset($this->scripts[$this->baseurl . '/media/system/js/mootools-core.js']);
unset other file(s).....
```

Note that if you want to unset only Mootools, you don't even require those first four lines, but for most other files, you do need those lines (Bootstrap files, JQuery, etc.). Again, be careful, and it is only for skilled users.

So, this is for JavaScript files, but you can do something similar for CSS files—again, by inserting a command in the <head> section of your template's `index.php` file.

```
unset($this->styleSheets[$this->baseurl . '/components/com_rsform/assets/
css/front.css']);
```

With some more advanced techniques, you can even do this per extension, as in the preceding example for RSform and that in Listing 18-3.

Listing 18-3. Remove CSS Files per Component

```
$unset_css = array('com_rsform' , 'com_othercomponent' );
foreach($this->styleSheets as $name=>$style)
{
    foreach($unset_css as $css)
    {
        if (strpos($name,$css) !== false)
        {
            unset($this->styleSheets[$name]);
        }
    }
}
}
```

So, using a few relatively simple techniques, you may get rid of some unnecessary files, which already gives you a head start with a relatively light template. Some more advanced tips can be found in a presentation by the Dutch company Perfect Web Team (www.perfectwebteam.nl/expert-sessie/joomla-templating). It's in Dutch, but still very useful, owing to the code examples, and with Google Translate, it becomes fairly readable.

If you do use a commercial template, look at your template provider first. There are a lot of differences between them. Some build very clean templates, while others literally require dozens of separate resources. In addition, some providers allow for optimization techniques in their template parameters. You can check my blog at www.joomlaseo.com/Blog/templates-performance-comparison for a comparison of different template providers.

A final option (not surprisingly) is to use an extension. Mootools can easily be removed by using the **Mootools Enabler/Disabler** (digitaldisseny.com); if necessary, on a page-by-page basis. Another extension is **JCC - JS CSS Control** (joomla-extensions.kubik-rubik.de/jcc-js-css-control) by Kubrik. This one simply allows you to remove any file, by simply entering the file you want to get rid of in the plug-in parameters.

Compress Your Files (Minification)

By **compressing** your files you basically remove all whitespace, leaving just a very compressed file. Reading your code will be near impossible, but all functionality will still work. You might leave your files uncompressed while developing and compressing them when publishing your site.

There are tools online to compress your code (just Google **compress css js online**) or use a code editor that does this for you, but you can also leave them uncompressed in your site and use an extension such as JCH Optimize, jbetolo, or ScriptMerge for that (see the following paragraphs).

Combine Your Files

Most Joomla sites use many CSS and JS files, leading to multiple HTTP requests. Combining all these would be a great way of reducing these requests. While for organizing your functions during development it might be nice to have separate files, your web site doesn't care. As long as you combine them in the correct order, there should be nothing holding you back from combining these files into one. Now, you only need one CSS file and one JavaScript file.

Again, you can sometimes do this manually (for example, at yui.github.io/yuicompressor), but this is also something that extensions such as JCH Optimize, jbetolo, or ScriptMerge can do for you, which I will cover in depth later in this chapter.

Defer JavaScript

You have probably seen this suggestion already in tools such as GTmetrix and Pingdom. If you **defer** your JavaScript, you tell your browser to load your JavaScript and then to **delay executing** the JavaScript until the entire site has been loaded. For many functions that require JavaScript, this is not a problem; they function just as well when they are deferred. You can implement the functionality by simply adding the word *defer* in your scripts, like this:

```
<script src="javascript.js" defer></script>
```

Be aware, however, that there are functions that require these scripts to be executed immediately. An example might be a collapsible Bootstrap module, or the e-mail cloaking plug-in.

This should not be too hard to implement, but you can also leave it up to JCH Optimize or jbetolo.

Run Scripts Asynchronously

Scripts can also be run **asynchronously** as soon as they are available, while the page continues to load. This feature is not fully supported: IE8 and earlier versions do not use it, but they won't crash or such, so the feature should be safe to use. It is implemented similarly to the defer tag.

```
<script src="javascript.js" async></script>
```

async plays together with the defer tag. Here is some further explanation from the www.w3schools.com web site:

- **If async is present:** The script is executed asynchronously with the rest of the page (the script will be executed while the page continues the parsing).
- **If async is not present and defer is present:** The script is executed when the page has finished parsing.
- **If neither async nor defer is present:** The script is fetched and executed immediately, before the browser continues parsing the page.

This feature is also included in ScriptMerge and JCH Optimize.

JCH Optimize

JCH Optimize has been mentioned a few times already. It can be a great help to speed up your web site. It offers many options to optimize various aspects of performance of your web site. It is especially good for combining and compressing CSS and JavaScript files, but it is also capable of setting defer/async parameters, setting a CDN domain, lazy loading images, optimizing images with Smush.it, and creating sprite images.

Let's see how we can configure it. The extension can be downloaded from www.jch-optimize.net (it is free to use, with a Pro version for only \$10). Download and install it from your **Extension Manager**. In the **Plugin Manager**, look for **JCH Optimize** and select it.

Enable it and go through the options you require (see Figure 18-19). Having read the previous paragraphs, you should have no trouble understanding what the different options do. On smaller sites, or when testing on your PC, you can probably start by activating all options. Then check if everything on your web site is still functioning as it should. Quite often, you may see that there are no issues at all, while your site is much faster.

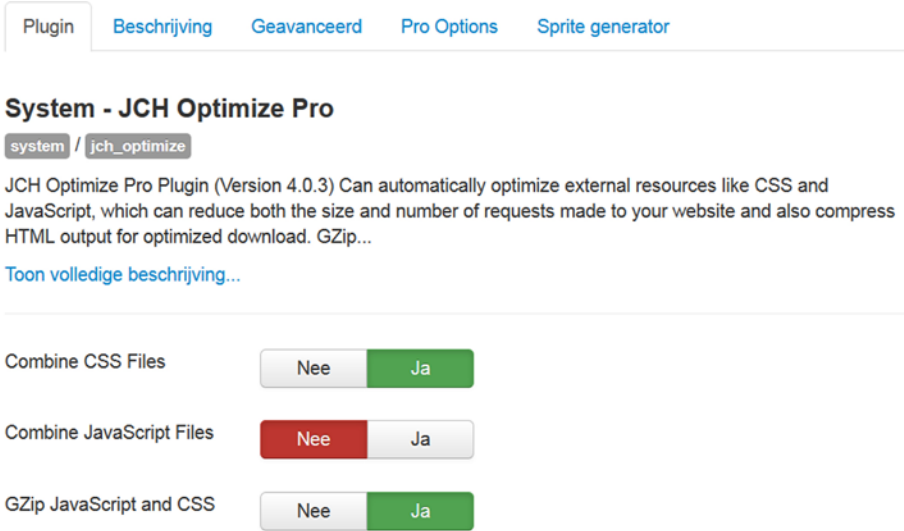


Figure 18-19. JCH Optimize plug-in options

In the **Advanced** tab, you can further fine-tune your actions. If there is a problem with one of the JavaScript files, you can **exclude** it from being combined, leaving the rest of the files combined. Also, you can choose to either **include** or **exclude** extensions from compression.

The **Pro version** of the plug-in allows you some even more advanced parameters, such as **asynchronous** loading of JavaScript, inline CSS and JavaScript, **lazy-loading** of images, etc. (Lazy loading of images means that they are only downloaded if you scroll down to where these images are located in the web page.)

In addition, support for CDN and Smush.it image resizing is included.

Don't be surprised, if things do break and stop working. Don't panic and, one-by-one, work through the options until you determine the problem. The JavaScript options especially can sometimes be problematic, so if you want to switch things off, start there. For example, form extensions can fail to display properly with some of the options. After setting **Defer javascript** to **No** and disallowing the inclusion of **inline javascript**, the form may function correctly again.

Jbetolo

Jbetolo is a free extension provided by www.jproven.com. jbetolo stands for “quickly” in Ethiopian, which pretty much sums up what this extension is meant to do. Roughly, it provides similar features as JCH Optimize. In part, they overlap, as you immediately see when you install it and check the plug-in's parameters, as shown in Figure 18-20.



Figure 18-20. *jbetolo* plug-in options

The main features are minification, combination, and compression of CSS and JavaScript files, while you can also include or exclude specific files and components. This is also what JCH Optimize has to offer.

Some features of JCH Optimize are not found in jbetolo, such as setting the `defer` attribute and asynchronous loading of JavaScript.

There are also some useful features, however, that are only available in jbetolo.

- In the **Tools** tab, there are options to set up the **leverage browser caching** code in `.htaccess`, without manually having to edit this.
- What is particularly interesting is the option to send the images from your Joomla site to Smush.it and receive the optimized versions, all simply with the click of a button!
- A big bonus is also the built-in **CDN support**, especially for MaxCDN and Amazon CloudFront, which can be implemented without an additional plug-in.
- Finally, **lazy loading** for images is available from the **Miscellaneous** tab.

Whether you would prefer to use JCH Optimize, ScriptMerge, or jbetolo depends on personal preference or the specific demands of your web site. If this is the first time you are using a plug-in such as this, JCH Optimize seems a bit easier to handle.

ScriptMerge

Another alternative with functionality similar to jbetolo is **ScriptMerge** by **Yireo**. You will recognize the same basic parameters to combine CSS and JavaScript files in the basic options, which also allow you to remove Mootools (see Figure 18-21).

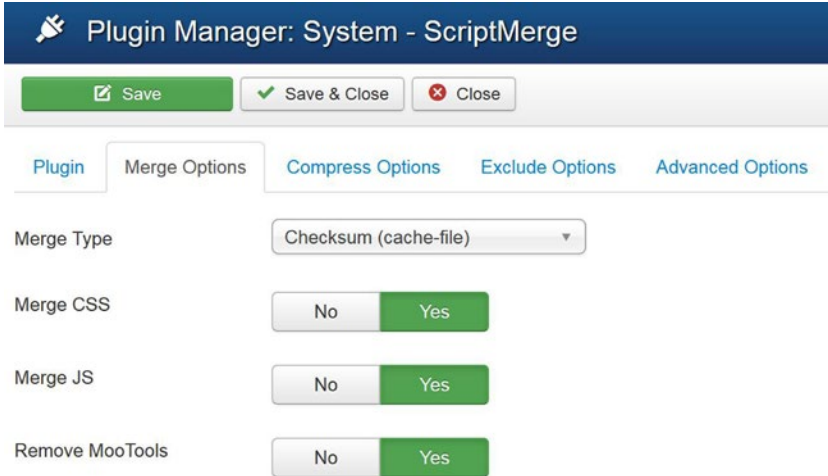


Figure 18-21. ScriptMerge's Merge Options

The merge options are the basic options that will hold the biggest value for you. The other tabs allow you to specify and fine-tune further options.

- **Compress Options:** This sets whether and, if so, how you want to compress your files (simple rules or more advanced).
- **Exclude Options:** As every now and then performance plug-ins such as these don't match with some other extensions, being able to set exclude options is vital. You can set exclude options on various levels (specific files, components, or menu items).
- **Advanced Options:** Various options. Unique options are the support for data URIs (Base64 encoding based on a maximum file size) and WebP (converts normal JPGs and PNGs into WebP images, reducing size, if supported by your host).

There are a few more extensions that aim to perform the same as JCH Optimize, jbetolo, and ScriptMerge, but these three are probably your best bet. If you run a performance test before and after installing one of the latter, you will be amazed at the results.

CDN (Content Delivery Network)

When you have a web site that is not only visited by local users but by people from all over the world, or within large countries such as the United States or Russia (see Figure 18-22), you may find that users far from where your web site is hosted encounter long page loads.

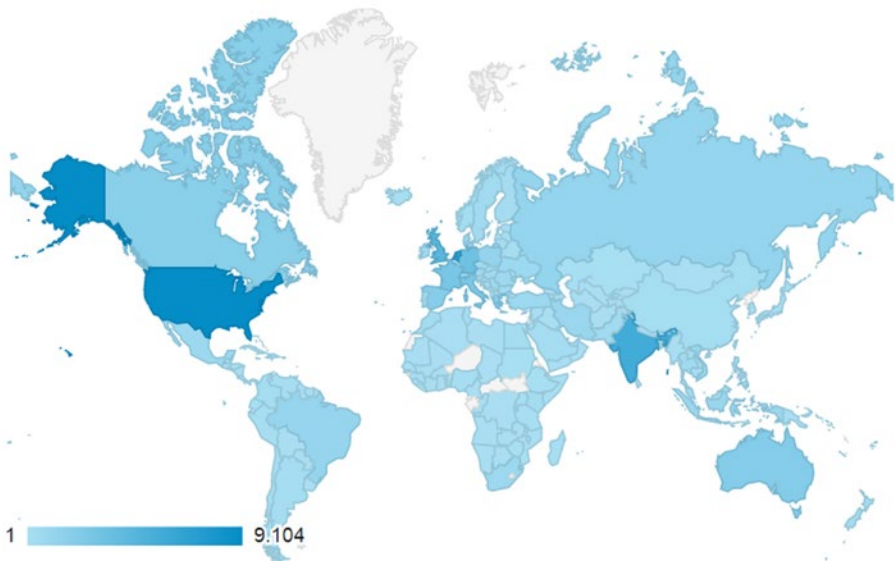


Figure 18-22. Sites with visitors across the globe or located in large countries need a CDN

This might even be more the case in developing countries with poor infrastructure and low-bandwidth connections. The only solution, then, is to get as close to your visitors as possible. This is where CDNs come in. CDN companies usually own networks with servers all over the world. The more servers they have and the more they are distributed, the better. If your content is distributed from a CDN, the static resources from the site are pulled from the CDN, while the changing content, i.e., the HTML files, is retrieved from the actual server where the site is hosted.

Static resources are files such as

- Images
- Video
- CSS
- JavaScript
- Web font files
- And more

Together, these files usually make up 80–90% of the size of your web site!

All this sounds quite complicated, but, happily, with the help of some extensions, solutions are not that hard to implement in a Joomla web site.

The first step is to sign up with a CDN provider. This may depend on the extension you want to use, so check your extension's web site first. Some popular CD's that often work well with Joomla are **MaxCDN**, **CloudFlare**, and **Amazon CloudFront**. After you have signed up with a CDN provider, you will probably have to enter the URL of your web

site. After that, it usually takes some minutes, up to hours, for the CDN to pull your files from the server and distribute them across its network.

Once the CDN is completed synchronizing, it will inform you of the subdomain on the CDNs servers where the files are hosted. This is often a long URL such as the following:

xxxxxx.cloudfront.net

With this information, install and configure your CDN extension. **CDN for Joomla** by **NoNumber** is a great extension that is extremely easy to use, but there are others, such as **JomCDN** by **CorePHP** (www.corephp.com), as well as still others that are specific to a certain CDN provider. jbetolo, ScriptMerge, and JCH optimize also support the use of a CDN.

In the case of the NoNumber extension (Figure 18-23), simply enter the domain supplied by the CDN in the plug-in parameter, and you're done.

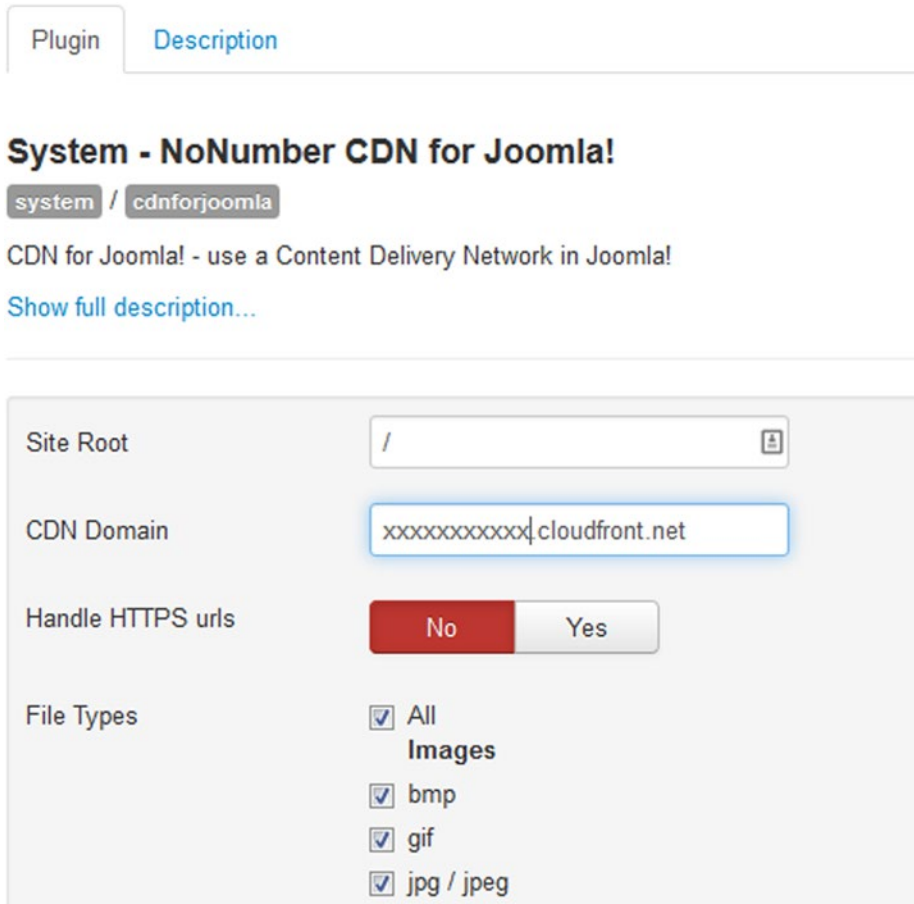


Figure 18-23. NoNumber CDN for Joomla plug-in settings

After activation, the **static files** will be loaded from the CDN. You can check this if you look into the source code of your files. A section for an image would then read as follows:

```

```

Without a CDN, this would simply be

```

```

It is also possible to replace the long, ugly URL for the CDN (<http://xxxxxx.cloudfront.net>) with something cleaner. You can make it look like the CDN is loaded as a subdomain from your own site, so it would appear as something like

```
http://cdn.yoursite.com
```

This takes some further configuration, and this probably goes a bit too far for this book, but there is an extensive tutorial for this on the www.joomlaseo.com web site.

Finally, some hosts, such as **SiteGround**, offer free CDN functionality right from their Cpanel for CloudFlare, without any configuration in Joomla, making it even easier to implement.

Optimize for Mobile Devices

Over the last years, users have started to use devices such as phones and tablets more and more to visit our Joomla web sites. This has encouraged us to embrace responsive design and offer them a nice user experience. We should also keep in mind, however, that many of these devices are still connected through slower networks, such as 3G and slower. Even though the situation improves quickly, we should take special care in optimizing performance for our web sites specifically for these devices, especially since the Google mobile-friendly update published on April 21, 2015.

Now for display purposes, we can sometimes use classes such as **hidden-phone** or **hidden-tablet**, so that they are not shown to the user, but this does not stop the code and images from actually loading. They are only hidden. This means useless data is being sent to the mobile browser, slowing down the connection.

Responsive design is usually based on detecting the size of screen or the viewport. Detecting the size of the viewport does not really detect whether we are on a mobile device, however. It just checks the width of the screen, which might as well be a resized browser window on a desktop.

A good technique that overcomes this issue is the use of **user-agent detection**. With user-agent detection you can actually define on what kind of device the web site is requested. This means that you have control over what is loaded or not.

A very useful extension that uses this technique is the **Advanced Module Manager** by **NoNumber** (www.nonumber.nl/extensions/advancedmodulemanager). This extension replaces the Joomla Module Manager with a much more advanced one. Now, you can not only restrict modules to certain menu items, but you can also use geographical information, detect the presence of other modules, time of day, etc. Among these options, you will also find one to **include or exclude mobile browsers** (see Figure 18-24).

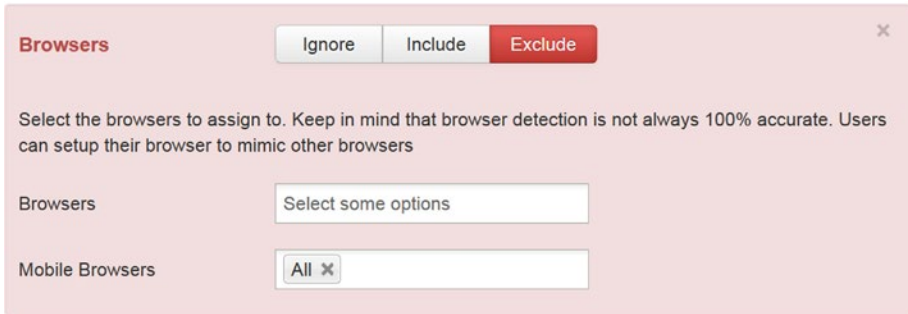


Figure 18-24. Exclude modules for mobile browsers with Advanced Module Manager

With the preceding settings, the module will not be loaded at all when the site is shown on a mobile browser, saving you a lot of bandwidth usage.

There is also a very small plug-in available that specifically does the user-agent detection for you, allowing you to use `if` statements in your template, depending on the user-agent encountered. This plug-in is not on the Joomla Extensions Directory but can be downloaded from the GitHub page of Rene Kreijveld (github.com/renekreijveld/UserAgentDetector). Once downloaded and installed, you can use code like the one in Listing 18-4 to load modules.

Listing 18-4. Exclude Module for Mobile with User-Agent Detection

```
<?php if($ualayout != "mobile") {?
<jdoc:include type="modules" name="slider-module" style="xhtml" />
<?php } ?>
```

This sample code will now only load the slider if the user-agent is not mobile. Particularly if you are making your own templates, this should not be too hard to implement.

Also, the **Simple Mobile Detection** plug-in by Conflate (www.joomla-specialist.net/joomla-extensions/joomla-simple-mobile-detection.html) employs this technique, allowing use of this anywhere in your content, including your articles.

Extensions

You may have noticed that I am a big fan of JCH Optimize, jbetolo, ScriptMerge, JotCache and NoNumber extensions. I have no relationship with the developers other than that I really appreciate the extensions, but, of course, there are many others.

Especially with performance, you will often find that a solution that works on one site does not work that well on another one, which could make you investigate other extensions. There are many more extensions on the Joomla Extensions Directory that might work for your specific situation, so feel free to check them out.

Final Check

After all your efforts, it is finally time to check your web site again. Run your site through the same YSlow, PageSpeed, or Pingdom test that you ran earlier.

In most cases, I expect that your web site will load in almost 70–80% of the time that it took before you began the optimization.

I wonder how close you can come to a 100% score!

APPENDIX A



Glossary

Certain terms in this book are frequently referred to. This glossary explains these.

Term	Explanation
301 redirect	HTTP-status code used to indicate that a URL has been permanently moved to another URL
404 error	HTTP-status code used to indicate that the URL does not exist anymore
404 page	The page that is offered to the user to indicate that the page is not found
Alt text	An alternative text attached to an image to describe its purpose to disabled readers (also used by search engines)
Anchor text	The actual words that make up a link
Bot	Also referred to as <i>Google Bot</i> , <i>robot</i> , or <i>spider</i> . Software used by Google to crawl web sites and determine their value
Bounce rate	The percentage of users who arrive at a web page and leave the site again
Breadcrumbs	A trail of links used to indicate the current location of the page in the web site hierarchy
Browser caching	Mechanism used to indicate that resources already downloaded on the computer of a user browsing the Web do not have to be reloaded on consecutive visits to the same web page
Caching	Mechanism used to indicate that resources already created on the server do not have to be re-created once they are there. Users browsing the site will be shown the cached version of this page.
Canonical URL	If the same content is offered on the same web site on multiple URLs, only one of these should be the canonical URL, i.e., the one that should be indexed by Google. The other ones should point to the canonical URL as the original source of the content.

(continued)

Term	Explanation
CDN	Content Delivery Network. A system of globally distributed servers used to offer resources such as images, CSS, and JavaScript files from the nearest location to the user
CSS	Cascading Style Sheet. Information used to apply styling to HTML pages
CTR	Click Through Rate. The percentage of users that sees a SERP and actually clicks these
Duplicate content	Identical content that is offered on multiple URLs on the same web site
Google Analytics	Statistics software offered for free by Google. Provides data related to user behavior on web sites
Google Webmaster Tools	Web service provided by Google to webmasters for free, allowing them to provide information to Google about their web site's infrastructure
G-zip	Mechanism to compress resources of a web page on the server before being sent to the user's browser
Heading	The main visual headline on a web page, configured as a <H*> HTML tag
.htaccess	System file delivered with Joomla. Used to switch on SEF URLs
Index	The complete list of web sites that are stored with a search engine and offered on its search page
Keyword	Word or phrase that web sites would like to be found by users searching in search engines
Keyword density	Percentage of a specific keyword that is found in a block of text. This should not be over 5%.
Keyword stuffing	Having a keyword density that is way above the advised upper limit
Link	Reference to a web site on another web site or another page within the same web site. Created with the <a> HTML tag
Link building	The process of generating incoming links from other web sites, usually in order to receive a higher ranking in Google
Link juice	The "value" that a link provides to its destination URL
Long-tail keywords	Keywords that are not very much searched for but that deliver customers with very specific wishes, often leading to high conversions
Meta description	The bit of text created on a web page created by the <meta name="description"> HTML tag. Used by Google as a snippet in the search results

(continued)

Term	Explanation
Meta keywords	The bit of text created on a web page created by the <code><meta name="keywords"></code> HTML tag. Not used anymore, so it should be avoided
Microdata	Also referred to as <i>rich Snippets</i> or <i>structured data</i> . A way of offering information marked up with additional information, so that it can be recognized by automated software (addresses, recipes, etc.)
Nofollowed link	A link with the attribute <code>rel="nofollow"</code> . Used to indicate that search engines should not attach any link value to the destination URL. Should be used for affiliate links
Robots.txt	System file delivered with Joomla. Used to inform search engines which part of the web site (URL or folder) is opened for crawling. Also used to indicate the location of the sitemap optionally
SEF URL	Search Engine Friendly URL. URL that can be easily read by humans and search engines, avoiding system characters
SEO	Search Engine Optimization
SERP	Search Engine Results Page. Page with (usually) ten sites shown for a search result
Sitelink	Small links shown as an addition to a specific search result in a SERP
Sitemap	Overview of all links on a web site, as far as they should be visible. This can either be an HTML version for humans or an XML version for search engines
Title	The page title, as shown in the browser's tab and as the first line for a search result for a web site in a SERP
URL Rewriting	Mechanism on Apache servers used to rewrite URLs in a user-friendly way. Also referred to as <code>mod_rewrite</code> in Apache.

APPENDIX B



Checklist

You can use this checklist as a means to optimize your own sites. You may want to adjust it according to your own settings. Note that this checklist only contains items related to SEO and performance. The order of the checklist is different from that followed in the book. This checklist is more oriented to an efficient workflow. Often, options for both performance and SEO can be found on the same screen; therefore, and it is more efficient to set these simultaneously.

Step	Action
1.	Choose a reliable web host and a domain.
2.	Build or buy a fast and optimized Joomla template.
3.	Install Joomla and keep it off-line until the site is ready. Install on a local or development server and upload once ready.
4.	Choose a site name that can be used in page titles and configure if and how it should be prepended/appended.
5.	Make sure Search Friendly URLs are switched on.
6.	Rename .htaccess.txt to .htaccess and switch on URL-Rewriting.
7.	Review the Metadata settings in Global Configuration.
8.	In Global Configuration, System-tab;, activate caching.
9.	In Global Configuration, System-tab;, activate G-zip.
10.	Activate the System-Page Cache plug-in in the Plugin Manager.
11.	Leverage browser caching by adding the necessary code in .htaccess.
12.	Force only www or non-www access of the site in .htaccess.
13.	Set up a logical category structure.
14.	Set up a logical menu structure, preferably following the logic in your categories.
15.	Make sure to apply metadata to every article you create. For items not related to articles, such as blog-views and lists, etc., apply these to your menu items. Use the keywords from the list of keywords that you made earlier in the process.

(continued)

Step	Action
16.	Check that your page titles are correct, especially for menu items.
17.	Avoid linking the same article to multiple menu items (use menu item aliases, if necessary). Avoid any kind of duplicate content.
18.	Make use of microdata.
19.	Create internal links and optimize the anchor text.
20.	Optimize images: use small images with stripped metadata; use logical file names and apply alt text.
21.	Configure a breadcrumbs module.
22.	Use 301 redirects to reroute old URLs to the new ones or for renamed articles/menu items. Use either the Redirect component, extensions such as SH404SEF, or .htaccess.
23.	Avoid 404 errors and configure a nice 404 page.
24.	Make sure no resources are blocked in robots.txt. Make sure the lines for /images, /media, and /templates are removed.
25.	Create a <code>sitemap.xml</code> file.
26.	Optimize CSS and JavaScript. Use the options that your template providers offer in the template parameters or use an optimization plug-in such as JCH-Optimize, ScriptMerge, or Jbetolo.
27.	Optimize for mobile devices. Use a responsive design and avoid loading heavy resources on mobile devices.
28.	For sites targeting large countries or the whole world, use a CDN.
29.	Set up Google Webmaster Tools and Google Analytics.
30.	Acquire backlinks from trusted sources.
31.	Acquire social media shares and set up social pages, such as a Facebook fan page, Twitter page, Instagram page, etc., that point to your web site.
32.	SEO audit your site with tools such as WebCEO, Screaming Frog, etc.
33.	Audit your site for performance, using tools such as GT-Metrix, Pingdom Tools, etc.

Index

■ A

Anchor text, [51–52](#)

■ B

Base64 image encoding, [145](#)

Bots, [2](#)

■ C

Cache cleaning, [136–137](#)

Cache plug-in, [135–136](#)

Canonical URLs, [63–64](#)

CDN. *See* Content delivery network (CDN)

Contact, [29](#)

Content delivery network (CDN)

globe/located in large

countries, [152](#)

hosts, [155](#)

NoNumber

extension, [154](#)

static files, [155](#)

static resources, [153](#)

■ D

Duplicate content

canonical URLs, [63–64](#)

description, [61](#)

extensions, [66](#)

in Joomla, [62](#)

301 redirects, [65–66](#)

robots.txt, [66](#)

Webmaster tools, [67](#)

■ E

Easy front-end SEO (EFSEO), [34](#)

automatic mode tab, [90](#)

metadata parameters, [89](#)

Plugin Manager, [89](#)

404 Errors

301 redirects with htaccess, [73](#)

extensions, [73](#)

Google Webmaster Tools, [77](#)

Joomla's redirect manager, [71–72](#)

prevention, [70–71](#)

■ F

Facebook, [133](#)

@fontface embedding, [133](#)

■ G

Google image search, [85](#)

Google web fonts, [133](#)

Google Webmaster Tools, [62, 77](#)

GTmetrix, [129](#)

Gzip, [140–141](#)

■ H

Header tags

default headings, Joomla, [45–48](#)

extension, [50](#)

template overrides, [48–50](#)

Hikashop, [59](#)

htaccess.txt file, [65–66](#)

Hummingbird, [2](#)

■ I

Internal links

- creation, anchor texts, 52
- description, 51

■ J

JavaScript files, 146, 148

Jbetolo, 150–151

JCH optimize, 149–150

JED. *See* Joomla Extensions Directory (JED)

Joomla caching, 133

Joomla environment

- clean and valid code, 18–19
- custom provider, 17–18
- domain/URL, 15–16
- download and installation, XAMPP, 17
- hosting, 15
- HTML5, 19
- localhost, 16
- order of code, 19
- responsive design, 20
- template, 17
- version, 16

Joomla Extensions Directory (JED), 34, 92

Joomla global configuration

- adds Suffix to URL, 24
- logged search query, 28
- metadata settings, 26
- meta keywords, 26
- robots, 27
- search component options button, 27
- SEF URLs, 23
- SEO settings (*see* Search engine optimization (SEO))
- site meta description, 26
- site name, 21–22
- site name, page titles, 24–25
- unicode aliases, 24
- URL rewriting, 23–24

Joomla performance and speed

- advantages, 128
- base64 image encoding, 145
- cache cleaning, 136–137
- caching, 133
- CDN, 153, 155
- combination, 148
- compression, 148
- CSS and JavaScript optimize, 146

extensions, 156

global cache settings, 133–134

GTmetrix, 129

Gzip, 140–141

image size

- correction, 143, 145
- reducion, 141–142

JavaScript, 148

Jbetolo, 150–151

JCH optimize, 149–150

JotCache, 138–139

lazy loading, 145

leverage browser

- caching, 139–140

lossless compression, 142–143

mean and lean, 132–133

mobile devices, 155–156

module cache, 134–135

NoNumber cache cleaner

- plug-in, 137–138

PageSpeed Insights, 128–129

Pingdom tools, 130–131

run scripts asynchronously, 149

ScriptMerge, 151–152

SEO, 127

sprites, 145

system-page cache plug-in, 135–136

template files reduction, 146–148

web hosting provider, 131

WebPagetest site, 130

YSlow browser extension, 128

JotCache, 138–139

JSitemap, 83, 85, 91–92

■ K

Kubrik extensions, 66

■ L

Lazy loading, 145

Leverage browser caching, 139–140

■ M

Metadata

- article, 29, 31
- description, 30–31
- JED, 34
- menu item, 29, 32
- OSMeta, 35

- page titles, 33–34
- Robots parameter, 35–36
- Microdata
 - address marked up, 55
 - extensions, 58
 - Google SERP, 55
 - implementation, 59
 - Joomla 2.5, 57–58
 - Joomla 3.3, 56
 - rich snippets, 55
 - structured data, 55
- Microdata Google
 - Breadcrumbs, 59
- Module cache settings, 134–135

■ N

- NoNumber cache cleaner
 - plug-in, 137–138

■ O

- Optimize images, 53–54
- Optimizing for long-tail keywords, 8
- OSMap, 83, 91
- OSMeta, 35, 87–88

■ P, Q

- 404 Page
 - article creation, 74–75
 - HTTP header, 76
 - Joomla, 75
 - menu, 76
 - templates, 74
- PageSpeed Insights, 128–129
- Panda, 2
- Penguin, 2
- Pingdom tools, 130–131
- Pinterest feeds, 133

■ R

- Rich snippets. *See* Microdata
- 301 redirects, 65–66
- Robots, 2
- Robots.txt
 - CSS, 82
 - file content, 79
 - folder is not blocked, 81
 - Joomla 3.3, 80

- mobile friendly test, 81
- “templates” menu title
 - disallowed, 79–80
- Run scripts asynchronously, 149

■ S

- Sample sites, 62
- ScriptMerge, 151–152
- Search engine friendly (SEF) URLs, 23
- Search engine optimization (SEO)
 - extension
 - canonical, 90
 - EFSEO, 89–90
 - Joomla Extensions Directory, 92
 - JSitemap, 91–92
 - non-SEF to SEF, 90
 - OSMap, 91
 - OSMeta, 87–88
 - SH404SEF, 87
 - StyleWare, 90
 - global configuration, 22
 - Google AdWords keyword
 - planner, 8–10
 - Google starter guide, 13
 - Google updation, 2
 - keywords uses, 12–13
 - keywords *vs.* themes/topics, 3
 - long-tail keywords and local, 7
 - meta description, 5–6
 - scanning process, 2
 - SERPs, 4
 - title, 4–5
 - tools, keyword research, 8
 - unique and quality content, 3
 - WebCEO keyword research
 - tool, 10, 12
- SEF URLs. *See* Search engine friendly (SEF) URLs
- SEO. *See* Search engine optimization (SEO)
- SH404SEF, 66
- Simple mobile detection, 156
- Sitemap
 - HTML, 83
 - Image Sitemap File, 85
 - search engines, 84
 - URLs, 83
 - visitors, 83
- Spiders, 2
- Sprites, 145
- Structured data. *See* Microdata

■ INDEX

■ **T**

Twitter, [133](#)

■ **U, V**

URL structure

alias, [38–39](#)

breadcrumb module, [40, 42](#)

category, [40](#)

hierarchy, [42–43](#)

Joomla site, [38](#)

menu, [40](#)

numbers, [43–44](#)

rules, [37](#)

User-agent detection, [155](#)

■ **W**

Web hosting, [131](#)

Webmaster

tools, [67](#)

WebPagetest

site, [130](#)

■ **X**

XAMPP, [17](#)

XMap, [83](#)

■ **Y, Z**

YSlow browser extension, [128](#)

Joomla! 3 SEO and Performance



Simon Kloostra

Apress®

Joomla! 3 SEO and Performance

Copyright © 2015 by Simon Kloostra

This work is subject to copyright. All rights are reserved by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed. Exempted from this legal reservation are brief excerpts in connection with reviews or scholarly analysis or material supplied specifically for the purpose of being entered and executed on a computer system, for exclusive use by the purchaser of the work. Duplication of this publication or parts thereof is permitted only under the provisions of the Copyright Law of the Publisher's location, in its current version, and permission for use must always be obtained from Springer. Permissions for use may be obtained through RightsLink at the Copyright Clearance Center. Violations are liable to prosecution under the respective Copyright Law.

ISBN-13 (pbk): 978-1-4842-1125-0

ISBN-13 (electronic): 978-1-4842-1124-3

Trademarked names, logos, and images may appear in this book. Rather than use a trademark symbol with every occurrence of a trademarked name, logo, or image, we use the names, logos, and images only in an editorial fashion and to the benefit of the trademark owner, with no intention of infringement of the trademark.

The use in this publication of trade names, trademarks, service marks, and similar terms, even if they are not identified as such, is not to be taken as an expression of opinion as to whether or not they are subject to proprietary rights.

While the advice and information in this book are believed to be true and accurate at the date of publication, neither the author nor the editors nor the publisher can accept any legal responsibility for any errors or omissions that may be made. The publisher makes no warranty, express or implied, with respect to the material contained herein.

Managing Director: Welmoed Spahr

Lead Editor: Ben Renow-Clarke

Technical Reviewer: Dan Rahmel

Editorial Board: Steve Anglin, Mark Beckner, Ewan Buckingham, Gary Cornell, Louise Corrigan,

Jim DeWolf, Jonathan Gennick, Robert Hutchinson, Michelle Lowman, James Markham,

Matthew Moodie, Jeff Olson, Jeffrey Pepper, Douglas Pundick, Ben Renow-Clarke,

Dominic Shakeshaft, Gwenan Spearing, Matt Wade, Steve Weiss

Coordinating Editor: Christine Ricketts

Copy Editor: Michael G. Laraque

Composer: SPi Global

Indexer: SPi Global

Artist: SPi Global

Distributed to the book trade worldwide by Springer Science+Business Media New York, 233 Spring Street, 6th Floor, New York, NY 10013. Phone 1-800-SPRINGER, fax (201) 348-4505, e-mail orders-ny@springer-sbm.com, or visit www.springeronline.com. Apress Media, LLC is a California LLC and the sole member (owner) is Springer Science+Business Media Finance Inc (SSBM Finance Inc). SSBM Finance Inc is a Delaware corporation.

For information on translations, please e-mail rights@apress.com, or visit www.apress.com.

Apress and friends of ED books may be purchased in bulk for academic, corporate, or promotional use. eBook versions and licenses are also available for most titles. For more information, reference our Special Bulk Sales-eBook Licensing web page at www.apress.com/bulk-sales.

Any source code or other supplementary material referenced by the author in this text is available to readers at www.apress.com. For detailed information about how to locate your book's source code, go to www.apress.com/source-code/.

Contents

About the Author	xiii
About the Technical Reviewer	xv
Introduction	xvii
■ Chapter 1: How Search Engines Work	1
The Scanning Process.....	2
Google Updates: Panda, Penguin, and Hummingbird	2
Unique and Quality Content.....	3
Keywords vs. Themes/Topics	3
SERPs, Title, Metadata.....	4
Title.....	4
Meta Description	5
Keyword Analysis	6
Long-Tail Keywords and Local SEO	7
Tools for Keyword Research	8
Google AdWords Keyword Planner	8
WebCEO Keyword Research Tool.....	10
Use Your Keywords	12
Google SEO Starter Guide.....	13

■ Chapter 2: Preparing Your Joomla Environment	15
Hosting	15
Domain/URL	15
Joomla Version	16
Start on a Test Environment	16
Template.....	17
Custom or Template Provider?.....	17
Clean and Valid Code	18
HTML5	19
Order of Code.....	19
Responsive Design	20
■ Chapter 3: Joomla Global Configuration.....	21
Global Configuration: Site Name.....	21
Global Configuration: SEO Settings	22
Search Engine Friendly URLs.....	23
Use URL Rewriting	23
Adds Suffix to URL	24
Unicode Aliases	24
Include Site Name in Page Titles	24
Global Configuration: Metadata Settings.....	26
Site Meta Description	26
Site Meta Keywords.....	26
Robots.....	27
Joomla Search	27

Chapter 4: Metadata	29
Metadata Setup per Article or per Menu Item.....	29
Metadata for Every Separate Page.....	30
Meta Description.....	30
Metadata: Article.....	31
Metadata: Menu Item.....	32
Page Title.....	33
Managing Metadata.....	34
Robots.....	35
Chapter 5: URL Structure	37
URLs in Joomla.....	38
Based on Alias.....	38
Example 1: Menu Item URL with Keyword.....	39
Example 2: Shorter URL.....	39
Category Structure.....	40
Menu Structure.....	40
Breadcrumb Module.....	40
URL Hierarchy.....	42
Numbers in the URL.....	43
Chapter 6: Header Tags	45
Default Headings in Joomla.....	45
Example 1: Articles Shown in Blog/List/Featured Articles.....	46
Example 2: Menu Items Directly Linked to Articles.....	47
Template Overrides.....	48
Header Tags Extension.....	50

- **Chapter 7: Optimize Internal Links and Anchor Text 51**
- **Chapter 8: Optimize Images 53**
- **Chapter 9: Microdata or Rich Snippets 55**
 - Microdata in Joomla 3..... 56
 - Microdata in Joomla 2.5..... 57
 - Microdata Extensions 58
 - Further Microdata Implementation..... 59
- **Chapter 10: Duplicate Content..... 61**
 - Duplicate Content in Joomla 62
 - Canonical URLs 63
 - 301 Redirects Using .htaccess Tuning..... 65
 - Robots.txt 66
 - Extensions 66
 - Webmaster Tools 67
- **Chapter 11: 404 Errors and Pages..... 69**
 - Causes of 404 Errors 70
 - Preventing 404 Errors 70
 - Joomla’s Redirect Manager 71
 - Extensions 73
 - 301 Redirects with .htaccess 73
 - Offering a Custom 404 Page 74
 - Advanced 404 Page: Link to a Menu 76
 - More Advanced: Correct HTTP Header 76
 - Monitor Crawl Errors in Google Webmaster Tools 77

■ Chapter 12: robots.txt	79
Old robots.txt files—Check Them!	80
Do Not Block Resources	81
■ Chapter 13: Sitemap	83
Sitemap for Visitors	83
Sitemap for Search Engines	84
Image Sitemap	85
■ Chapter 14: Simple SEO Extensions	87
OSMeta	87
Easy Front-End SEO	89
Non-SEF to SEF	90
Canonical	90
OSMap (Previously XMap)	91
JSitemap	91
Other SEO Extensions	92
■ Chapter 15: SH404SEF	93
Plan Ahead!	94
Preparations	94
Installation and Configuration	95
URL manager	99
Purge URLs	100
Set Up the Home Page	101
Aliases Manager	102
shURL Manager	102
404 requests manager	102

Titles and Meta Manager.....	103
Analytics Manager.....	104
Disadvantages.....	105
Extensions Similar to SH404SEF	105
MijoSEF.....	106
JoomSEF.....	107
■ Chapter 16: SEO Site Audits	109
SEO Doctor Firefox Plug-in.....	109
MozBar	110
SEORCH.eu Web Site	111
Screaming Frog SEO Spider	112
Site Audit with WebCEO.....	112
Technical Audit	113
SEO Analysis	114
Landing Page Optimization.....	115
Other Tools for SEO Audits.....	116
■ Chapter 17: Off-site SEO	117
Google Webmaster Tools	117
Setting a Preferred Domain	118
Sitemap	119
HTML Improvements.....	121
Search Queries	121
Google Analytics	123
Google+	125
Link-Building and Social Media.....	125
Local SEO	126
From Here.....	126

- Chapter 18: Joomla Performance and Speed 127**
- Check Before You Begin 128
 - YSlow Browser Extension 128
 - PageSpeed Insights 128
 - GTmetrix 129
 - WebPagetest..... 130
 - Pingdom..... 130
- Web Hosting Provider 131
- Mean and Lean 132
- Joomla Caching 133
 - Global Cache Settings..... 133
 - Module Cache 134
 - System - Page Cache plug-in 135
 - Cache Cleaning 136
 - NoNumber Cache Cleaner Plug-in 137
 - JotCache..... 138
- Leverage Browser Caching 139
- Gzip 140
- Reducing Image Size..... 141
 - Lossless Compression 142
 - Correctly Sized Images..... 143
 - Sprites 145
 - Base64 Image Encoding 145
 - Lazy Loading..... 145
- Optimize CSS and JavaScript..... 146
 - Reduce Your Template Files 146
 - Compress Your Files (Minification) 148
 - Combine Your Files 148

■ CONTENTS

Defer JavaScript	148
Run Scripts Asynchronously	149
JCH Optimize	149
Jbetolo	150
ScriptMerge	151
CDN (Content Delivery Network)	152
Optimize for Mobile Devices.....	155
Extensions	156
Final Check.....	157
■ Appendix A: Glossary	159
■ Appendix B: Checklist	163
Index	165

About the Author

Simon Klootra is known mostly for his blog, JoomlaSEO(joomlaseo.com), which is essentially devoted to the same subjects as this book: search engine optimization (SEO) and performance for the Joomla! content management system. In addition, he has written articles for the online Joomla! Community Magazine and occasional guest posts for several (Dutch) Joomla portals. He has conducted workshops and training sessions on numerous occasions, including the Dutch Joomla days.

Klootra is a professional website designer and works as an SEO consultant for SK-Web and Perfect Web Team, both based in the Netherlands. Naturally, SEO is a major focus of these websites. He also has a great passion for (speed) windsurfing.

About the Technical Reviewer

Dan Rahmel is a Windows programmer with more than 14 years of experience designing and implementing information systems and deploying mid-sized client/server solutions, using Microsoft SQL Server, MySQL, Microsoft Access, and Visual FoxPro. He has written over a dozen books, including *NutsandBolts Filmmaking*, *Visual Basic: Programmer's Reference* (first, second, and third editions), *Building Web Database Applications with Visual Studio 6*, *Sams Teach Yourself Database Programming with Visual Basic 6 in 24 Hours*, *Developing Client/Server Applications with Visual Basic 4*, among many others. His books have been translated into various languages, including Chinese, Japanese, Spanish, French, and Portuguese. He is a contributor to *DBMS*, *Internet Advisor*, and *American Programmer* magazines.