

ON PAGE TECHNICAL SEO AUDIT

3RD EDITION



A K O O Z A I G U I D E



Legal Notice:

By reading this e-book, you agree to the following terms and conditions.

Under no circumstances should this e-book be sold, copied, or reproduced in any way except when you have received written permission.

As with any business, your results may vary and will be based on your background, dedication, desire, and motivation. Any testimonials and examples used are exceptional results, which do not apply to the average purchaser and are not intended to represent or guarantee that anyone will achieve the same or similar results. You may also experience unknown or unforeseeable risks which can reduce results. We are not responsible for your actions.

The material contained in this report is strictly confidential.



CONTENTS

A LITTLE BIT ABOUT KOOZAI	4
ABOUT THE AUTHORS	5
ABOUT THIS WHITEPAPER	6
HOSTING SERVERS / PROVIDERS	7
GOOGLE ANALYTICS CODE	10
GOOGLE WEBMASTER TOOLS	11
GOOGLE WEBMASTER TOOLS RESEARCH	13
ROBOTS.TXT	18
SITEMAP.XML	20
HTML VALIDATION	21
SITE SPEED	22
.HTACCESS FILE	25
URL STRUCTURE	26
DUPLICATE CONTENT	27
AUTHORSHIP	28
MARK UP, RICH SNIPPETS, TAGS AND SCHEMA	29
SUMMARY	30



A LITTLE BIT ABOUT KOOZAI:

Koozai are a multi-award winning Digital Marketing Agency. We're one of the UK's fastest growing technology companies having been ranked 23rd in the Deloitte Technology Fast50.

We help businesses connect with their customers online – providing a range of industry leading services from Organic Search Engine Optimisation (SEO) and Pay Per Click (PPC) Management, Social Media and Brand Management.

We thrive on building long lasting client relationships and delivering true value for money. We're passionate about what we do - and that shows in our work.

At Koozai, there are no lengthy contracts, just world class Digital Marketing. Koozai will help you build your brand online and achieve a ROI that can be clearly measured against your bottom line.



koozai

INTELLIGENT DIGITAL MARKETING



WE THRIVE ON BUILDING LONG LASTING CLIENT RELATIONSHIPS AND DELIVERING TRUE VALUE FOR MONEY. WE'RE PASSIONATE ABOUT WHAT WE DO - AND THAT SHOWS IN OUR WORK."



ABOUT THE AUTHORS



Dean Marsden (@Koozai_Dean)

Dean Marsden is focused on the core things that keep a brand looking great online, from brand management to video marketing. He excels at delivering video marketing solutions for clients and specialises in converting website visitors into customers.



Emma North (@Koozai_Emma)

Emma has more than 5 years' digital marketing experience and has worked on dozens of websites in a wide range of industries. She has a passion for both SEO and PPC and is driven by the need to develop her digital skills and knowledge. She is always exploring innovative solutions for new problems encountered in the ever-changing digital world.



ABOUT THIS WHITEPAPER:

There are a number of critical and technical issues that any good webmaster should review in order to improve the on-page SEO of their website.

Our completely up-to-date third edition includes the very latest information when performing an on-page SEO audit.

Covering everything from hosting types to code performance issues, understand the core technical issues that can hold back an SEO project.

There are a number of critical and technical issues that any good webmaster should review in order to improve the on-page SEO of their website.

In this whitepaper we set out all of the technical aspects of on-page SEO and how to use them to your advantage.



“THERE ARE A NUMBER OF CRITICAL AND TECHNICAL ISSUES THAT ANY GOOD WEBMASTER SHOULD REVIEW IN ORDER TO IMPROVE THE ON-PAGE SEO OF THEIR WEBSITE.”



HOSTING SERVERS / PROVIDERS

Hosting your website on the right type of server is really the first step or choice you will need to make when setting up a website.

However, even if you already have a website up and running, it is usually still possible to migrate it over to another server. This paper will not cover migrating your website (for further information on this, please see: <http://www.koozai.com/resources/whitepapers/website-migration-new-site-guide/>) but we will explore some of the benefits of using the two main types of hosting servers.

If you would like to check your server type, you can use:

<http://www.whatismyip.com/tools/server-headers-check.asp>

You will then get a handy report like the one below, which also contains other pieces of technical information that may be of use.

```
HTTP/1.0 200 OK
Date: Wed, 30 Apr 2014 09:57:52 GMT
Server: Apache
Set-Cookie: bb2_screener_=1398651872+72.55.191.54; path=/
Set-Cookie: PHPSESSID=q9nne919jfe8uj50Vlc4ssd3; path=/
Expires: Thu, 19 Nov 1961 08:52:00 GMT
Cache-Control: no-store, no-cache, must-revalidate, post-check=0, pre-check=0
Pragma: no-cache
X-Pingback: http://www.koozai.com/blog/mirpc.php
Link: ; rel=shortlink
X-Powered-By: PleskLin
Vary: Accept-Encoding
Connection: close
Content-Type: text/html; charset=UTF-8
```

“ LINUX SERVERS ARE USUALLY RECOMMENDED BY HOSTING PROVIDERS.”

Linux Servers

Linux servers are very common and are usually recommended by most hosting providers, unless you have a specific need for a Windows server. Around 60% of publicly accessible websites are hosted on Linux servers.

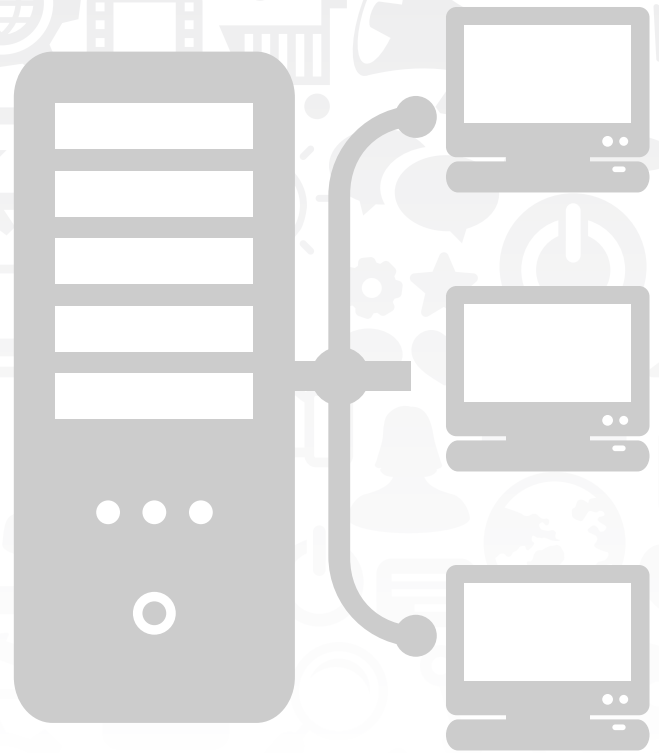
Linux offers some very attractive features and functionality that, in general, enable easier or simpler methods or processes for implementing changes to technical aspects of your website.

A great example of this is enabling “Gzip Compression”. This will be covered in more detail in the Website Performance section of this paper; but in short, to perform this task on a Linux server, a small piece of code is added to a file and uploaded to the root of the site.

Unless you specifically need to use a Windows server, we would always recommend using a Linux server. The primary reason for this is that Linux servers do not require you to have admin access to make basic changes. Linux is also a free / open source system and as such is nearly always cheaper to use than Windows.

Windows Servers

Windows servers require a greater



level of knowledge and experience to use and as such they can offer a great deal of flexibility. This flexibility does come at a price though, because server administrators are often highly skilled and qualified individuals. Therefore, it can cost a lot to have someone make changes to it if you are unable to make them yourself.

“WINDOWS SERVERS REQUIRE A GREATER LEVEL OF KNOWLEDGE AND EXPERIENCE TO USE AND AS SUCH THEY CAN OFFER A GREAT DEAL OF FLEXIBILITY.”

Windows is specifically good for integration with other Windows based tools and programs, such as:

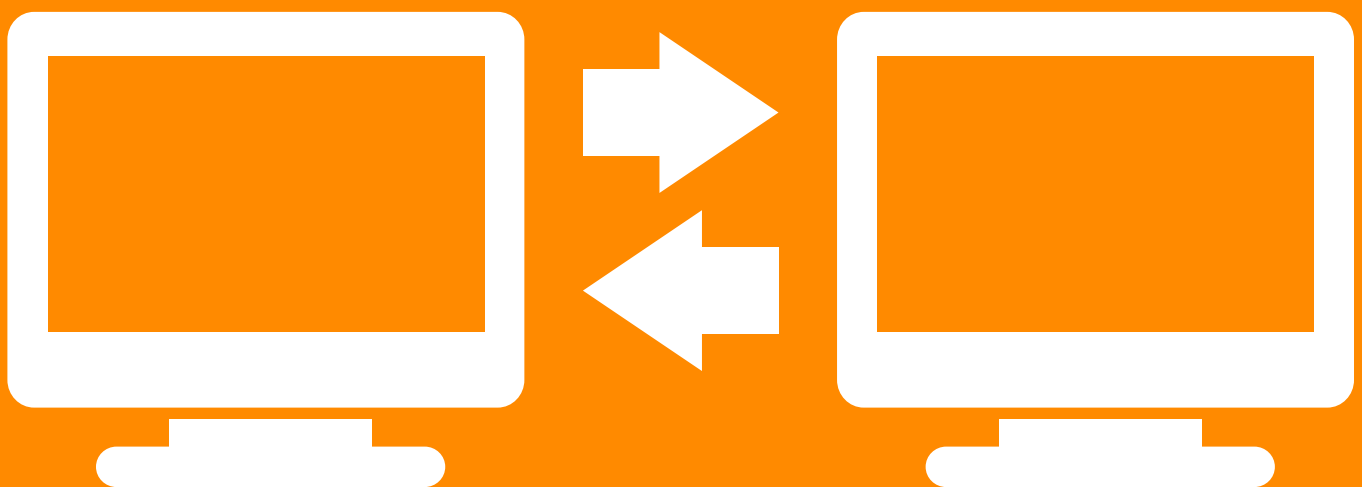
- ASPX
- .Net Framework
- Visual Basic

setting up 1,000 redirects, can be very labour intensive.

If you are using IIS and have SEO issues, the best way to resolve this is to use the official [IIS Search Engine Optimization Toolkit](#).

Windows servers are also able to seamlessly integrate Excel, Access, and some other Windows applications into your website.

The disadvantages of using a Windows server, as mentioned, are the cost and knowledge required. As such, making a change like in the example above of enabling “Gzip Compression” will require you to log into the server and make changes through the IIS control panel. Making bulk changes, such as





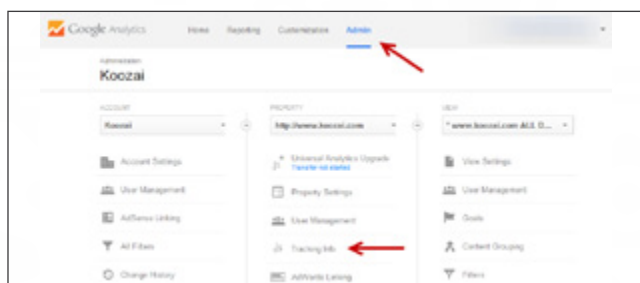
GOOGLE ANALYTICS CODE

Installing Google Analytics asynchronous tracking code on your website is a simple yet vital step in the process of implementing on-page SEO tools.

You will need to create a Google Analytics account via the link below if you do not currently have one:

<https://www.google.com/analytics/>

Once you have an account, simply go to the 'Admin' panel and click 'Tracking Info' as shown in the image below:



Once you have done this, you should scroll down the page to see some code and instructions, as illustrated in the example below. On this page you should choose the relevant settings for your tracking requirements, although normally the default settings are fine for most websites.

This is your tracking code. Copy and paste it into the code of every page you want to track.

```
<script type="text/javascript">
var _gaq = _gaq || [];
_gaq.push(['_setAccount', 'UA-12345678-1']);
_gaq.push(['_setDomainName', 'kooza.com']);
_gaq.push(['_trackPageview']);

(function() {
var ga = document.createElement('script'); ga.type = 'text/javascript'; ga.async = true;
ga.src = (https? => document.location.protocol + 'https://') + 'http://www.google-analytics.com/ga.js';
var s = document.getElementsByTagName('script')[0]; s.parentNode.insertBefore(ga, s);
})();
</script>
```

The latest code from Google should always be copied and pasted into the source code of your website immediately before the closing head tag which looks like this: `</head>`

This enables the code to be called very soon after the page starts to load, which allows you to capture more data. Installing the code further down the page can result in data loss or incomplete data within Google Analytics.



GOOGLE WEBMASTER TOOLS

Google Webmaster Tools (GWT) is a powerful and free tool provided to help webmasters understand how their sites are being indexed.

This allows them to optimise their sites for improved visibility within the Search Engine Result Pages.

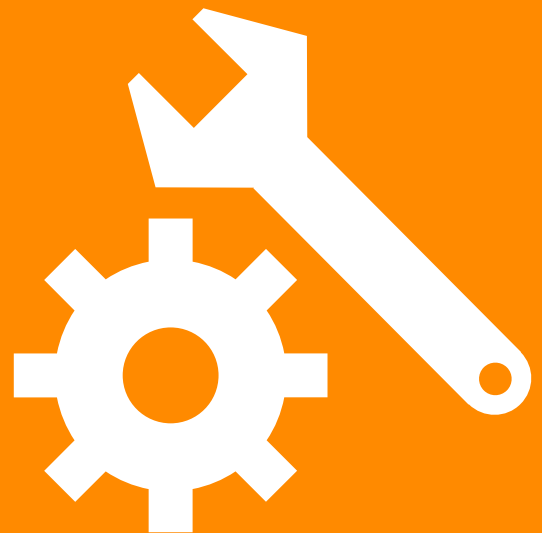
Verification of Webmaster Tools

This next stage is a simple but essential step in the SEO process. There are multiple ways to verify to Webmaster Tools that you are the administrator of your website. To get started, visit: <https://www.google.com/webmasters/tools/>

The recommended verification method is quite advanced but is the most secure option. A TXT Record must be created within the DNS management of your domain registration control panel. Google provides step-by-step instructions for

control panels of the most common domain registration companies.


If you do not have access to your domains control panel or are simply confused by DNS records then we recommend one of the following alternative verification methods:



THIS SHOULD BE THE FIRST PORT OF CALL FOR ALL WINDOWS IIS ERRORS.



- HTML file upload – add an HTML page that Google provides to your webhosting
- HTML tag – add in a line of code to the header section of your website's Home page
- Google Analytics account – if you are using the same login as your Google Analytics account then Webmaster Tools can use the analytics code to verify your site
- Google Tag Manager – Google Tag Manager is an advanced way to quickly add different types of Google code to your website's pages, however it does require a longer setup process than other methods.



in the top right-hand corner. Choose the 'Webmaster Tools Preferences' option in the drop-down navigation.

You are then presented with options to choose your language and set what type of emails to receive and to which address. We recommend setting the email type to 'all issues' to make sure you don't miss anything.

Setting up Email Alerts

To set up email alerts, navigate to the main Home page of your Webmaster Tools account (straight after you've logged in) and click on the cog symbol

“TO SET UP EMAIL ALERTS, NAVIGATE TO THE MAIN HOME PAGE OF YOUR WEBMASTER TOOLS ACCOUNT.”

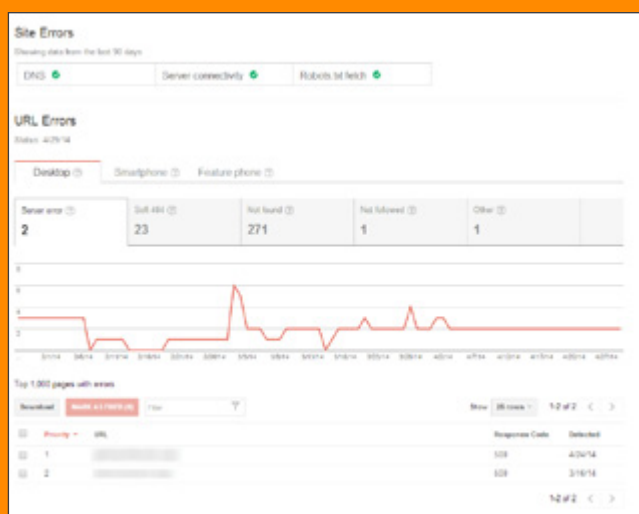


GOOGLE WEBMASTER TOOLS RESEARCH

Below we will discuss some of the more relevant aspects of Google Webmaster Tools (GWT) to perform a technical SEO audit.

Crawl Errors

The below image shows the section of the GWT dashboard that displays information about crawl errors.



Here you can click on the tabs to see a list of the pages / URLs on your site that Google has identified the specified error on. Errors are categorised by type of device to make it easier to spot desktop or mobile errors only.

You should regularly check GWT for these errors and if you identify any, seek to resolve them as soon as possible.

Click on each link in the list to learn more about the error and find out where it was linked from incorrectly. This can either be from another page on your website, or alternatively from a page on an external website.

Crawl errors are essentially problems that Google's spiders / robots encountered when attempting to index a site. Errors can be missing pages /

“ CRAWL ERRORS ARE PROBLEMS THAT GOOGLE'S SPIDERS ENCOUNTERED WHEN ATTEMPTING TO INDEX A SITE.”

broken links, server errors etc. Some of the most common errors are:

Soft 404 Error - These occur when the server does not return a proper 404 error but Google bot recognises the page as missing.

Not Found Error - This is a missing page and the majority of 404 errors will come under this heading.

400 Error - A web server error telling you that your web browser or whatever application you use to access the webpage being loaded has accessed that page incorrectly or the request was corrupted.

401 Error - Usually this is a failed login attempt, essentially someone trying to access a page they do not have permissions to access.

403 Error - This is a forbidden page, similar to a 401, although there is no login available - meaning that any attempt to access it through a web browser will return a 403 error.

500 Error - This is a generic error code and could apply to any internal server error. This could include, but is not limited to, the server being overloaded, load balancers failing, or being unable to complete a server side request.

Sitemap Errors - Errors that are found in your Sitemap.xml file.

Timed Out - Usually this is returned because the server remains busy, but can also indicate incredibly poor load times, such that the robot crawling your site has given up trying to load a page after a set amount of time.

It is also possible to check error and server response codes using tools like Screaming Frog, which you can download for free here: <http://www.screamingfrog.co.uk/seo-spider/>

It's best practice to edit the source of the link so that the original link is corrected. This is easier if it's on your website, however it's much harder if the link is from an external website. In both cases, it's also important to create a 301 redirect from the old incorrect link to the correct live page so that users and search engines can find the page every time.

When you have fixed an issue you can mark it as fixed to ensure you know which ones you have done in the future.

URL Parameters

The "URL Parameters" section of GWT allows you to specify how Google is to deal with certain dynamic URLs containing specific or generic parameters. For example if you have an

ecommerce website, your ecommerce system may generate URLs that sort products by price or size. Essentially these are the same pages and you should use the URL Parameters feature to indicate that these dynamic URLs do not need to be indexed.

There are too many parameters to list and these will vary depending on how your site is built. Google Webmaster Tools will usually identify the types of parameters identified on your website through crawling it, however you can manually add your own if it doesn't.

Using the URL Parameters feature alongside your robots.txt file means you can help Google concentrate on indexing your live and unique content.

HTML Improvements

This section of GWT will show you a number of issues related to HTML. However this will not show validation problems, but rather issues with Meta:



Click on any of the text highlighted in blue to go through to a page detailing the URLs where the specified errors were detected. Common issues include Meta that isn't a good length or that has been

duplicated elsewhere, because Meta data should be unique on each page of the website and fit within the character limits to display correctly in the search results.

Malware

This section of GWT normally shows "Google has not detected any malware on this site." However, if it shows anything else then you may have some big problems!

If your site has become infected it can usually be attributed to some vulnerability that a hacker has exploited. This could lead to anything from your site's content being changed, to malicious scripts being executed on the computer of someone else visiting your website.

There is no standard answer for what to do in this situation, other than to say that you may need to restore your



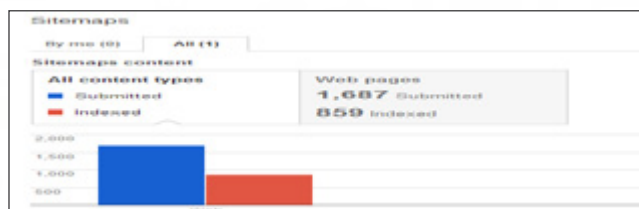
website from a backup and hire someone to look to address the vulnerability that facilitated the hack in the first place.

If you search for the specific Malware file in Google you can also usually find support from other affected users.

Sitemaps

As mentioned earlier in this whitepaper and in more detail a little later on, you should always use a Sitemap.xml file and submit it to Google via Webmaster Tools. This section of GWT shows you how Google is using it and if there are any related issues.

You can test your Sitemap in this section and also see how many pages from the Sitemap have been indexed as the image below shows:



In this case we can see that 859 of the 1,687 pages in the Sitemap are indexed, meaning 828 are currently not. This

“YOU SHOULD ALWAYS USE A SITEMAP.XML FILE AND SUBMIT IT TO GOOGLE VIA WEBMASTER TOOLS.”

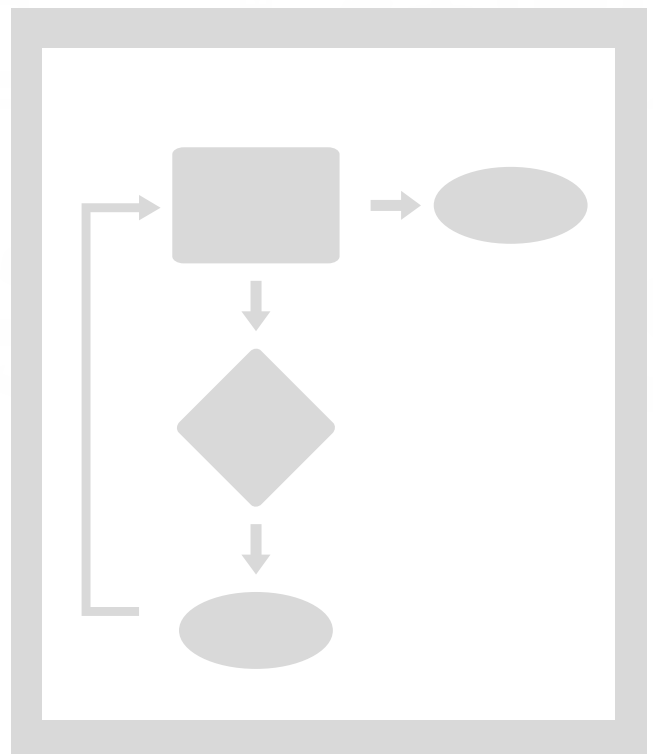
could indicate an issue stopping Google from adding the missing pages to its index, or poor quality pages that Google refuses to index.

Blocked URLs

This section tells you which URLs your robots.txt file is blocking and is a great place to see how Google perceives the parameters that you have specified within your robots.txt file. If you notice here that pages are blocked that should not be, you should amend the robots.txt file and retest straight away.

Manual Actions

Within the Search Traffic header, the Manual Actions section lists any manual webspam actions which have been applied to your site by Google. These can



include penalties for unnatural inbound or outbound links, thin content, keyword stuffing, or simply “pure spam”.

Although you should receive a message in Webmaster Tools if a manual penalty is applied to your site, it is recommended that you check this section frequently to ensure nothing has slipped through the net. Doing this will help you to spot any issues which could be having a negative effect on your site and rankings.

If you do receive notice of a manual action, you need to resolve the issue as quickly as possible and submit a reconsideration request.

Issues with backlinks (“unnatural inbound links”) tend to be the most complicated and lengthy to resolve. Check out our guide on [Backlink Analysis](#)



[and Removal](#) for more detailed advice on unnatural links penalties and how to remove them.

Links to Your Site

Google provides a list of your site’s backlinks, referring domains and referring anchor text through the “Links to Your Site” section of Webmaster Tools. Here you can carry out backlink health checks by ensuring that your link profile does not appear spammy with an unnatural number of links with optimised anchor text or from a single domain.

Although there are many backlink analysis tools available that you can and should use to stay on top of your backlink profile, this section is important because it gives you an insight into some of the links that Google is actually taking into account when reviewing your site.





ROBOTS.TXT

The Robots.txt file is simply a text file that exists on the root of your domain and is the first port of call for search engine spiders when they crawl your site.

The robots.txt file also carries useful functionality, such as indicating to search engines which files or directories of your website should be indexed and which should not.

Making a robots.txt file is very easy and is done simply by using Notepad or any text editor; open a new file and save it as “robots” and the default file type will be .txt (which means text). Below is an example of the Koozai Sitemap:



Specify Access

All robots.txt files should open with the following statement:

User-agent:

However, this statement on its own is incomplete as the command is specifying to the “user agent” or spider/crawler that the commands underneath apply to. So you can specify a different set of rules for Google and Bing if you wanted to do so.

For example, the following would apply only to Google:

User-agent: Googlebot

In order to define a set of commands for all search engines simply use the following:

User-agent: *

Blocking Files and Folders

It is sometimes relevant to “disallow” access to a directory, file or type of file from being indexed. This is usually because Google only spends a certain

amount of time indexing websites and in that time you want them to index the important stuff. It can also prevent pages that are in development from appearing in search results.

To disallow the indexing of a directory of images located at www.example.com/images/ you could use the following command:

Disallow: /images/

If your website has dynamically created URLs as the result of search functionality on your website, you will want to exclude all of those dynamically generated pages. It is possible to exclude specific file types or URLs containing specific characters. Using Jpeg files as an example, this is how you could exclude all JPEGs from being indexed like this:



Disallow: /*.jpg\$

Allow Access To Everything

To simply allow all of your site to be indexed by all crawlers just add the following code to your file:

User-agent: *

Allow: /

Sitemap Reference

It is important to link to your Sitemap from the robots.txt file as search engine bots will follow this link to your Sitemap when it first visits your website.

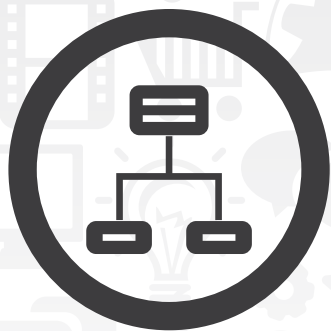
For example:

Sitemap: <http://www.koozai.com/sitemap.xml>

Checking Robots.txt File

If you want to make sure the right information is being blocked then use the robots.txt testing tool in Google Webmaster Tools. This can be found under the “Blocked URLs” heading in the Google Webmaster Tools menu.

“IT IS IMPORTANT TO LINK TO YOUR SITEMAP FROM THE ROBOTS.TXT FILE.”



SITEMAP.XML

The Sitemap.xml file is an XML document containing a list of pages on your website that you have chosen to tell Google and other search engines to index.

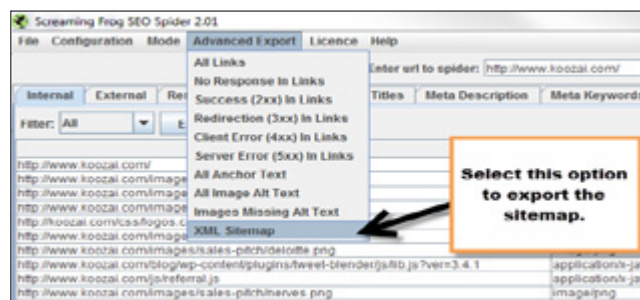
Google may decide not to index all of these but uses this file as a guide to the pages available on your website.

The Sitemap can also carry other information about each page, such as when it was created and last modified, and its importance relative to other pages on your site.

This is also the second part of call for search engine spiders looking to crawl and index your website, so it is very important that you have one present on your site. There are a number of free tools that can create a Sitemap for you, including the two below:

- Screaming Frog (<http://www.screamingfrog.co.uk/seo-spider/>)
- XML-Sitemaps.com online tool (<http://www.xml-sitemaps.com/>)

You will require the latest Java update for your operating system for Screaming Frog to work, but once you load it, simply follow the below instructions to produce a Sitemap:



Select a folder to save your Sitemap to. We recommend calling it "sitemap.xml".

It is also important that you submit your Sitemap to both Google and Bing Webmaster Tools in order to expedite the process of indexing your site.

“THE SITEMAP.XML FILE IS AN XML DOCUMENT CONTAINING A LIST OF PAGES ON YOUR WEBSITE.”



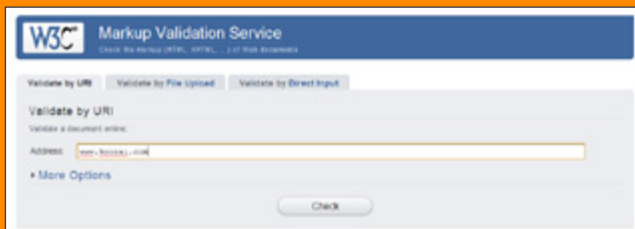
HTML VALIDATION

Having good and compliant HTML is important for a number of reasons; not least to ensure that your website and all of its features render or work properly.

HTML is really a vast and exhaustive topic and one that we cannot cover in a single paper, but you can check that the HTML on your website is compliant to standards by going to the following URL:

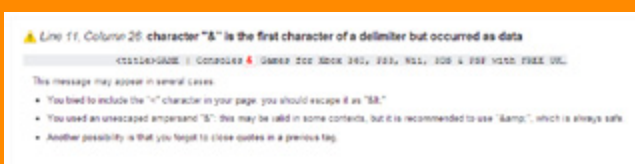
<http://validator.w3.org/>

Simply copy your website URL into the “address” field and click “check”:



If your site is fully compliant you will see a green verification tick. If not, you will see a list of any errors the validator finds.

An example error is shown below:



The tool runs a real time test so you can publish changes to your website and run a fresh test to immediately see if everything is set up correctly.

“ IF YOUR SITE IS FULLY COMPLIANT YOU WILL SEE A GREEN VERIFICATION TICK.”



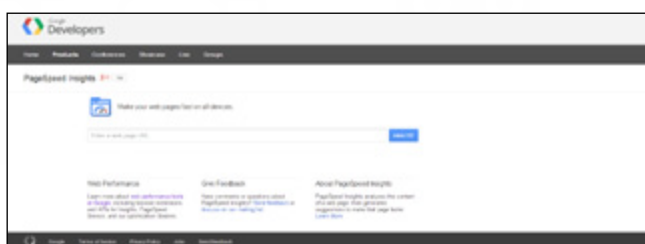
SITE SPEED

Web site performance is a major factor in the ranking of your website, for the obvious reason that users do not want or expect to have to wait for slow webpages to load.

As such, search engines take into account the speed and performance of your website when ranking your pages in the SERPs.

To check your website's speed and identify ways in which you can improve it, one of the best tools is Google's PageSpeed Insights:

<https://developers.google.com/speed/pagespeed/insights/>



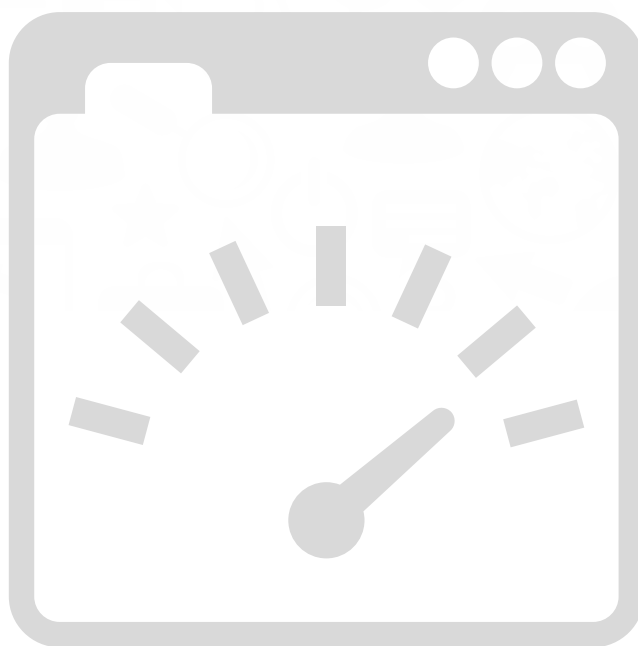
Combine External Scripts and Files

This is probably above average in terms of difficulty, unless you are well versed in creating and editing scripts. That said, it is still very much worth doing because it reduces the total amount of files that web browsers need to load.

It is important to note that scripts and files should be consolidated based on type, so do not combine a JavaScript and a CSS file together. You should combine all JavaScripts together and combine all CSS files together where possible.

CSS

As mentioned above, combine your CSS files together into as few files as possible and ensure that any links to the original files are amended to point to the



new single file. Other aspects to look to optimise are:

- Reduce white space by removing gaps in code
- Do not duplicate parameters
- Check for errors
- Be as efficient as possible with your coding

You can use the W3C validator to check your site for CSS errors.

Enable Gzip Compression

Gzip Compression helps to compress files and data when being parsed from the hosting server to the web browser, although this does not include images. As previously touched upon, if you use a Linux server this is easily achieved through adding the below code to your .htaccess file. For Windows servers this will be a lot more complicated.

The code to use is below:

```
-----  
# compress the files
```

```
AddOutputFilterByType DEFLATE text/  
html text/plain text/xml text/css text/  
javascript application/x-javascript  
-----
```

```
# removes browser specific bugs
```

```
BrowserMatch ^Mozilla/4 gzip-only-text/  
html
```

```
BrowserMatch ^Mozilla/4\.0[678] no-gzip
```

```
BrowserMatch \bMSIE !no-gzip !gzip-  
only-text/html
```

```
Header append Vary User-Agent  
  
-----
```

Format for Links

When a link is clicked the server will try and determine the type of file that you are opening, by putting "/" after the link, e.g. www.example.co.uk/blog/. This means that the server will know that it is a directory page; which reduces processor time and increases the website's speed.



Hosting Package

Hosting your website within the country that your target audience is located in can help to improve load times.

Always look for a Tier 1 hosting provider who can guarantee 99.999% uptime, as well as a high level of performance.

Reduce DNS Lookups

This is not a simple task if you are currently using the less efficient method as it will require a site re-structure; but if you are in development stages or are looking to add a sub domain, think about reducing DNS Lookups.

Rather than using `www.example.website.com` you should use `www.website.com/example`; this is because the former requires two DNS (Domain Name Server) Lookups, rather than the latter which requires one.

Reduce Image Size

There are a number of different ways to reduce the size of the image files on your site:

“HOSTING YOUR WEBSITE WITHIN THE COUNTRY THAT YOUR TARGET AUDIENCE IS LOCATED IN CAN HELP LOAD TIMES.”

- Compressing images
- Reducing the quality
- Saving them in a JPEG format
- Calling images from CSS
- Using a program like SnagIt or Paint to strip the excess data from files created in Photoshop
- Reducing the resolution of the image
- Using less colours

Reduce the Number of Redirects

This is not always possible, but reducing the number of redirects present on a site will ultimately reduce the overall load times.

This should be executed properly by removing redirected pages from the Sitemaps, removing any links to those pages from the site and anywhere you can on the Internet. Then ensure that Webmaster Tools reflects this change, by checking to see that these pages are not showing as broken links.

Further tips on how to redirect files can be found in our “Website Redirects & Canonical Issues” guide at <http://www.koozai.com/resources/whitepapers/website-redirects-and-canonical-issues/>



.HTACCESS FILE

This is only relevant to websites hosted on a Linux server, as Windows servers cannot use this type of file to control server-side actions.

These files are not to be used lightly and any aspect of the website that could be affected by these changes should be thoroughly checked after implementation.

It is possible to take down your website through incorrect coding within the .htaccess file, so we cannot stress how important it is that all work is checked properly once complete.

Despite the risks involved in managing your own .htaccess files, they are a powerful way to help you resolve a number of basic technical issues with a site.

One common type of issue resolved with .htaccess code is canonical webpages. This is when a page of a site, or even every page of the site, can be accessed from multiple URLs, running the risk of the same pages being indexed multiple times and being considered duplicate content by search engines.

NoFollow Tags

Nofollow tags have been used incorrectly by many webmasters over the years. If you are linking to another site, don't automatically use the "no follow" attribute. If the link is relevant, hasn't been purchased and is being used in a positive context, then it shouldn't be applied. This will help Google to find the page and determine that it is relevant.

However, if you have functionality on your site that can cause an infinite loop issue where robots / crawlers get stuck going from page A to page B then back to page A, use a no follow to prevent this from happening.

You can see in GWT that if your Sitemap has 1,000 pages but 50,000 have been indexed, there may be an issue with Googlebot being stuck in some kind of loop.



URL STRUCTURE

URL structure is often overlooked by webmasters who do not consider user journey or have little SEO knowledge. URLs that use syntax unfriendly character combinations like the example below are hard to read by humans:

`mysite.com/p=1?id+19374363%20%aab.html`

Furthermore, using this kind of meaningless format means that you cannot use either the keyword or a relevant description of the page within the URL. Consider not only what this looks like on-page when someone is visiting the site, but also what it looks like in the search results.

Using poor URL structures can also lead to other issues concerning parameter handling; for example if you have search functionality within the site then you may have a “?” appear within dynamically

created search result pages on your site. If you have a good URL structure this is easy to deal with. Simply exclude any page URL that contains a “?” from being indexed. If you use poor URLs, excluding a “?” may mean that all pages cannot be indexed.

Always try to use a simple and logical URL structure that makes sense, both to humans and search engine crawlers. For example:

`www.example.com`
`www.example.com/page-name`
`www.example.com/category/page-name`
`www.example.com/blog`
`www.example.com/blog/post-name`

Often the simplest solution is the best; this is certainly true with URL structuring.

“ALWAYS TRY TO USE A SIMPLE AND LOGICAL URL STRUCTURE THAT MAKES SENSE.”



DUPLICATE CONTENT

Duplicate content is a huge problem for websites, as search engines do not want to rank pages highly in the SERPs if they use the same content as another page or site and are not adding any unique value to users.

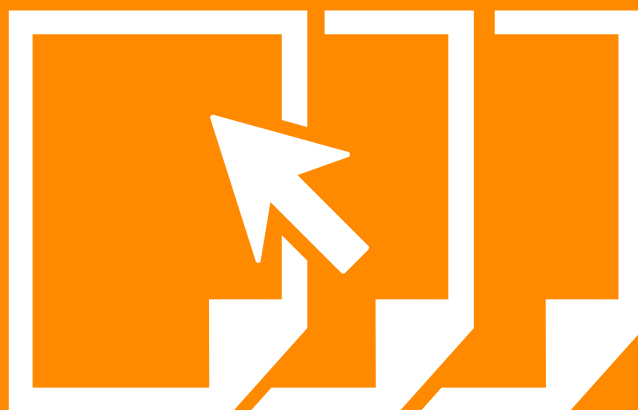
This is even more apparent since the Google Panda algorithm update was rolled out, which particularly targets poor quality and duplicate content. As a result, it has never been more important to identify and resolve any duplicate content issues that your site may have.

One of the best tools for this is Copyscape which has both a limited free version, as well as a very low cost advanced version, which is highly recommended. This tool is great for finding other websites that may have republished your content without permission. Use the following link to create a Copyscape account:

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

Keeping track of where your content is appearing on the net is vital to ensure that you remain the only website using the content you created. Vigilance here can also help protect against negative SEO attacks, where a competitor may duplicate your content in numerous places over the Internet to try and incur a penalty for your site.

“VIGILANCE HERE CAN ALSO HELP PROTECT AGAINST NEGATIVE SEO ATTACKS.”





AUTHORSHIP

Setting up authorship is now an essential part of building your trust and authority with Google by helping the search engine tie your content together.

This means that if you write a lot of good pieces, even across different sites, Google will understand that you wrote the content, giving you authority in that field.

In addition, Google Authorship can help your SERPs listings to stand out by including your Google+ circles information and profile picture with your listing:

The Real Metric For Measuring SEO Success



www.koozai.com/blog/search-marketing/measuring-seo-success/
by Emma North - in 811 Google+ circles

11 Apr 2014 - It's even easier to deem a decline in organic traffic as failure. But is this really the best metric to consider when measuring SEO success?

This involves setting up author pages on your website and then linking this, the

authored content and a Google+ profile together in the correct way:

- Link your Google+ profile to the author page on your site
- Link your content to the Google+ profile
- Link your author page to your Google+ profile

It is also possible to add an email address from the domain of the website (me@example.com) and then verify ownership through your Google+ account. This allows Google to determine authorship and thus attribute content to an author.

“GOOGLE AUTHORSHIP CAN HELP YOUR SERPS LISTINGS TO STAND OUT BY INCLUDING YOUR GOOGLE+ CIRCLES INFORMATION.”



MARK UP, RICH SNIPPETS, TAGS AND SCHEMA

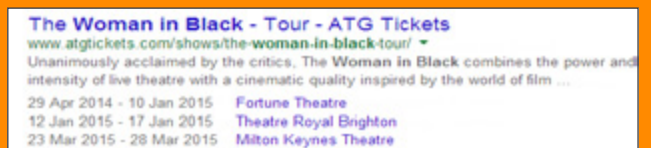
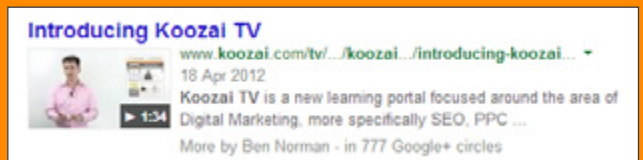
This is a powerful new way of adding code to your site that helps Google determine features of the page and your business.

For example, you can add review star ratings to your listings in the SERPs by adding code to the site that connects a review site to your website.

You can also mark-up data on-page to say “this is a price”, “this is a service” or “we offer this service in this specific place”. This can have a significant impact on your Click Through Rate (CTR), which has a direct impact on your rankings.

You can find the designated Schema markup at <http://www.schema.org> and below are some examples of recipes, videos and tickets.

When you have the code on your page, you can test it at: <http://www.google.com/webmasters/tools/richsnippets>



YOU CAN FIND THE DESIGNATED SCHEMA MARKUP AT [HTTP://WWW.SCHEMA.ORG](http://WWW.SCHEMA.ORG).





SUMMARY

We hope you found this whitepaper helpful in determining the technical flaws with your website and that it presented new opportunities for on-page SEO.

You might also enjoy our whitepapers on [Website Redirects](#) and [Getting Started With SEO](#) if you need additional resources on the topic.

For further information or assistance, please get in touch with us using any of the options below.



koozai

INTELLIGENT DIGITAL MARKETING



“ IF YOU DO NEED ANY FURTHER INFORMATION OR ASSISTANCE, YOU CAN GET IN TOUCH WITH US EASILY USING ANY OF THE OPTIONS BELOW.