SEO GUIDE

Version 1.2

it'seeze websites

brilliant designs – better results

What is SEO?

SEO – **Search Engine Optimisation** – is the process of improving a site's or page's ranking within search results through a combination of **on-page** changes (such as creating good headings) and **off-page** changes (such as gaining links from third-party sites).

How search engines rank sites

Each search engine has a vast database, called an **index**, storing information about all the pages they have found and how the pages are connected to each other by links. The index is created using software known as a **robot**, **crawler**, or **spider**. Robots find new pages by following links from existing pages. Some search engines allow you to manually submit new sites to them, but this isn't necessary as long as there are links pointing to the site.

The ranking of a page in the search results is determined by a combination of **relevance** and **importance**. Relevance depends on how much a page relates to the user's search terms. Importance is independent of any particular search query, and depends on factors such as the number and variety of incoming links.

Domain names

Years ago domain names had a big influence on ranking, and an **exact match domain name** – one that matches the search query – could significantly boost rankings. Because exact match domain names are more an indication of a site's budget (due to the high prices commanded by such domain names) than its quality, search engines have steadily reduced their influence on ranking. Instead, domain names should be chosen with a human audience in mind – a domain name should be short, memorable, and easy to type.

Titles and meta descriptions

The title and meta description are two examples of **metadata** – information that is not shown directly within a page but is available to search engines.

The **title** (often incorrectly called the title tag) is shown on the browser tab containing the page and in the browser title bar when viewing the page, and is the main link in the search result. The **meta description** (often incorrectly called the description tag) is often used as the longer **snippet** of text that appears in the search result.

A N Other Limited: Hand-made furniture in Totnes, Devon
www.example.com/
We offer a range of high-quality hand-made chairs, tables, and other furniture, built by expert craftsmen at our studio in Totnes, Devon.

The title is the single most important item to optimise in order to improve a site's perceived relevance. A search engine will regard a page that includes the user's search terms within its title as more relevant than a page that only includes the search term within the page text. Key elements to include within a title are:

- brand
- location (if applicable; it may be omitted for national companies)
- descriptive terms, including possible search terms

In the example above, the page title starts with the brand name, followed by some descriptive terms including possible search terms such as 'hand-made' and 'furniture', and ending with the location.

Away from the home page the brand can be left until the end of the title – for example, "New furniture designs from A N Other in Totnes, Devon". People are more likely to click on a search result when they see relevant words at the start of the title. As the home page will be the highest ranking page on the site for a search for the brand name, it is the only page whose title is worth starting with the brand name.

A title that is too short does not take full advantage of the ability to include search terms within it. A title that is too long is truncated by the search engines, and the additional text past the truncation point does not help ranking. The recommended title length is between 20 and 65 characters.

The meta description does not affect a page's ranking. However, it is often shown in search results and hence can act as a 'pitch' to encourage users to click through to the page. It should be a few sentences, around 80 to 160 characters in length (longer descriptions will be truncated), that give users a compelling reason to believe the page is worth reading. Note that for particular search terms search engines may decide not to use the meta description and instead use some relevant text from the page as the snippet.

In the it'seeze system, the SEO button on the toolbar allows you to edit the page title and description.

Keywords and keyword stuffing

Keywords are the search terms for which you believe your page is relevant. For example, keywords for the it'seeze home page include 'it'seeze', 'website', 'web design', and 'responsive'.

It is possible to tell search engines your intended keywords using **meta keywords** (often incorrectly called the keywords tag). While this was important in the early days of search, the major search engines no longer use this information as it is easy to abuse. Instead they

attempt to algorithmically determine the terms for which a page is relevant based on the page content.

For a search engine to know that a page is relevant for a term, that term (or closely related terms, such as different spellings or parts of speech) must be used on the page. Using a keyword more often on the page will generally increase the relevance of the page for that term. However, **keyword stuffing** – overuse of a keyword – will instead cause search engines to penalise a site by decreasing its ranking. Keyword stuffing can be recognised by unnatural writing, with certain words occurring unusually frequently. If text reads naturally it is unlikely to be guilty of keyword stuffing.

Another form of keyword stuffing is the inclusion of long lists of keywords, usually at the footer of a page. This is commonly used to list all of the locations a company serves. While a short list of half a dozen locations is fine, long lists will cause search engines to penalise the site.

Headings

Headings on websites are not just larger or more prominent text; they are explicitly indicated by invisible pieces of code called **heading elements** (often incorrectly called heading tags). Keywords that appear within heading elements have a greater effect on relevance than keywords that appear in normal text. However, this effect is relative to other text on the page, so a page where all the text is a heading will be treated the same as a page where none of the text is a heading; this prevents heading elements from being abused.

In the it'seeze system, the heading component allows you to insert heading elements. The most important heading on a page should be set as the **h1** heading. Further headings can be given the levels **h2** to **h6**, although it is rare to use levels past h3. There can be multiple headings of a particular level on a page, but in general there should only be one h1 heading.

Text equivalents (alt text)

Search engines are fundamentally text-based; while many search engines provide image and video search tools, these tools rely on finding text associated with the images or videos to be able to determine their subjects.

Images in web pages can specify a **text equivalent**, often referred to as **alt text**. It is important to remember that although the text equivalent is read by search engines, it is intended for human consumption. For example, blind people browse the web using software that reads out the text on pages, and reads out the text equivalent for images. For a purely decorative image (for example, a decorative divider), the text equivalent should be left blank. For an image that conveys information, the text equivalent should convey equivalent

information. For example, a graph showing sales over the past year might have a text equivalent stating "Sales rose 12% over the past year, from 3167 units to 3547 units". The text equivalent was once commonly used for keyword stuffing, but search engines can now detect this and penalise pages that use this technique.

In the it'seeze system, the text equivalent can be specified by clicking on the edit button for an image component.

PageRank

In January 1998 Larry Page, one of the co-founders of Google, filed a patent titled *Method for node ranking in a linked database*. Building on earlier work with Google's other co-founder Sergey Brin, this patent describes an algorithm known as **PageRank**, which attempts to determine the importance of web pages based solely on the way they link to each other. Later that year Page and Brin founded Google, and results generated using PageRank were so much better than contemporary search engines that Google quickly became the most popular search engine, remaining so ever since.

The mathematics of the PageRank algorithm are described in detail in the patent, but there are two common non-technical explanations: the **voting model** and the **random surfer model**.

The voting model is more metaphorical than literal, but gives a good intuitive understanding of the algorithm. It says that when a page links to another page, it's like a vote for that page. The PageRank algorithm proceeds over multiple rounds; in the first round every page's voting power is equal, but in each successive round pages that received more votes in the last round have more voting power to give out in the next round. Eventually the results change little from round to round, and the vote a page receives is its PageRank.

The random surfer model is more precise, as it's a direct interpretation of the underlying mathematics. Imagine if a person surfing the web starts at a random page. When they've read the page, they do one of two things: there's a 15% chance they move to another random page, and an 85% chance they click on a random link on the current page. Suppose they continue this process for a large number of steps. The probability that they end up on a particular page is its PageRank.

The end result is that pages with more incoming links tend to rank more highly, especially if the pages that link to them also have many incoming links. As a result, SEO experts sometimes refer to links from higher ranking sites as passing more 'link juice' than links from lower ranking sites.

Navigation

One consequence of the PageRank algorithm is that a page can only rank if it has incoming links. Some older sites use Flash files for animated navigation, which causes problems for ranking as search engines cannot see links within Flash files; even if there are links elsewhere to the target pages, their potential PageRank is reduced by the 'missing' links. Flash files can by identified by right-clicking on them: right-clicking on a Flash file shows a context menu including an 'About Adobe Flash Player' item.

Splash pages are home pages with lots of graphical content, little text content, and a link to a **secondary home page** – a page that would otherwise be the home page. Splash pages harm ranking by 'absorbing' some of the PageRank that would otherwise belong to the secondary home page.

Link building

As a result of the PageRank algorithm, and similar algorithms used by other search engines, a large part of SEO concerns **link building**: the process of gaining incoming links in order to improve ranking. Some techniques to gain extra incoming links are discouraged by search engines, and links gained in such a way will be ignored if detected.

Natural links are the links that your site gains as a result of people just deciding to link to it, and are the best kind of link. If your content is interesting, entertaining, or controversial, it is likely to gain larger numbers of natural links. Gaining natural links is thus more a process of content creation and copy-writing than one of SEO. It is also a slow process, as to begin with few people will be looking at your site and hence it will take time to gain the first few links. When someone writes an article specifically in order to gain a large number of links, the article is called **link bait**. This term is generally used negatively and reserved for deliberately controversial content; it is less common to use the term for useful or entertaining content. A good way of accelerating the process of gaining natural links is **outreach**: contacting people you think will find your site useful. Outreach may or may not involve actually asking for a link, depending on how you think the recipient will respond; many bloggers, for example, don't like being asked for a link and will be more likely to link if not explicitly asked.

Self-created links are links to your site that you create on other sites. For example, you could link to your site in a comment posted to a blog, or in your profile page on a social network. While it's perfectly normal to have some links to your site of this form, if you have too many links of this form search engines will regard them as suspicious (**comment spam**)and ignore them. In 2005 Google introduced the **nofollow** attribute for links, which is a way for a site to tell search engines to ignore particular links – for example, links from Wikipedia and many blogs use nofollow and so have no effect on your ranking. This was intended to reduce the incentive for comment spam, although the effect has been limited.

Link exchanges and bad neighbourhoods

One common link building technique is the **link exchange**: linking to a site in return for a link back. As with self-created links, it's normal to have some links for this form, but having too many will look suspicious to search engines, especially if the links have been exchanged with irrelevant sites (sites on completely unrelated topics) or sites in **bad neighbourhoods** (groups of sites which search engines already regard as suspicious). Search engines will penalise sites that they detect are engaging in link exchanges. This is a harsher punishment than that for comment spam as search engines realise that comment spam could actually be a malicious rival attempting to frame you, whereas link exchanges involve changes to your own site and hence require your participation.

Google's "panda" update

Google launched the "panda" update (formerly referred to as the "farmer" update) on 23rd February 2011. This update attempted to assess the quality of sites. It was developed by training a learning algorithm using the results of real human assessment, which means it penalises sites for using techniques that real people don't like, such as excessive advertising or template content.

Google's "penguin" update

Google launched the "penguin" update on 24th April 2012. This update targeted common spam techniques, including:

- exact match domains (e.g. a-long-list-of-keywords.com)
- overuse of exact match links (a high percentage of incoming links using the same link text)
- low quality articles
- keyword stuffing, especially in internal links

It has been described an "over-optimisation" penalty, but it's really targeting outright spam.