The Past, Present, & Future of SEO

A guide to the evolution of search along with advice on how to get the most from your SEO campaigns today, improving ROI while still working within the latest Google Best Practice Guidelines.
TABLE OF CONTENTS

INTRODUCTION 3
   CHANGE IS GOOD 3
   CATCH ME IF YOU CAN 4

THE ORIGIN OF SEARCH 5
   THE BIRTH OF THE BOT 5
   SEO IS BORN AND SO IS SPAM 6
   SEARCH GETS SOPHISTICATED 7

GOOGLE CLEANING UP THE WEB 8
   PAYING YOUR WAY TO THE TOP 8
   WEBMASTERS GAIN CONTROL 9
   LINK EARNING 9
   HUMAN INTERVENTION & PANDA 10
   IMPROVING THE SEARCH EXPERIENCE 12
   PENGUIN 13
   A SIGN OF “BIG THINGS” TO COME 14

A TIMELINE: THE EVOLUTION OF SEARCH 15

TECHNICAL SEO TIPS 17
   CLEANING UP YOUR ACT 17
   PLANNING STRATEGICALLY 19
   BEST PRACTICE RECOMMENDATIONS 20

CONTENT & ONLINE PR 22
   LINK EARNING AND DIGITAL ENDORSEMENTS 22
   LINKS APPEAL TO GOOGLE’S SENTIMENTAL SIDE 23
   SEO NO LONGER AN OUTSIDER 23
   WHAT MAKES “GREAT CONTENT” GREAT? 26
   HOSTING CONTENT VS. GUEST POSTING? 26
   SUMMARY 27

CONCLUSION 28
   SEO: WHAT DOES THE FUTURE HOLD? 28
   SEO SHOULD NOT STAND ALONE 28

SEO GLOSSARY 30
INTRODUCTION

The SEO industry is evolving rapidly and the most successful online businesses are those that are keeping up-to-speed with the latest changes and updates from Google.

Natural search remains the largest driver of traffic and revenue for most sites and this paper is designed to help you better understand the ever-changing SEO landscape. By understanding how the SEO landscape has evolved, we hope to help you improve your natural search rankings today and attract more customers online sustainably, ethically, and in line with the search engines’ recommendations.

CHANGE IS GOOD

The face of SEO is vastly different now than how it was 10 years, 5 years, or even 1 year ago. In order to keep your website’s position in the search engines consistent and avoid the wrath of Google, adhering to best practices is essential.

Google’s aim is to “make the web a better place” for its users, and it is constantly rolling out algorithm updates that focus on achieving this aim. Their focus is on presenting the most relevant results to searchers, those that are highly authoritative, credible, and user-friendly. We feel this proactive shift has been, and continues to be, a hugely positive step forward both for site owners and users alike.

Google’s latest updates have been controversial, but for those truly putting the user at the heart of all their digital marketing activities, the updates have delivered tangible improvements in rankings and ultimately revenue.

While it can be hard to keep up-to-date with the frequent changes Google is making, investing time in understanding the developments is critical to your long-term success online. To get the most out of the money you invest in SEO, you must first have a clear understanding of:

a) The search engines’ overarching aim and their vision for the future of the web.

b) What your SEO partner is doing on your behalf (i.e. what they do for their fees).

c) The direct impact SEO has on other parts of your business (e.g. your PR, marketing, or sales teams) and the relationships that need to be in place to maximise its impact.

d) The wider reasoning and experience that goes into forming the recommendations your SEO agency give to you.

Hopefully, this paper will help answer any questions you may have and give you some additional insight into the past, present and future of SEO.
CATCH ME IF YOU CAN

In the past, many companies tried to trick Google into ranking their sites higher than they deserved – it was a never-ending game of cat and mouse in which, ultimately, the user was left neglected and frustrated. In response to the increasing amount of manipulation and as technology has evolved, Google has gotten better at identifying such trickery and self-enforcing its guidelines.

Google is now looking well beyond the number of links to your site alone; instead, it is also taking into account authority, how usable and easy to navigate your site is, how fast it loads, and penalising sites that put more emphasis on over-optimisation for the search engines than on quality customer-centric content and navigation.

Unlike in the past, Google’s algorithms can now spot unethical practices easily, and increasingly brands of all shapes and sizes are being penalised for their use of them. This includes companies that have dabbled in poor practices in the past, either knowingly or unknowingly. There is nowhere to hide anymore, and this means that if you’re not prepared to do SEO properly, it’s better not to do it at all.

Google’s guidelines are crystal-clear, and no excuse exists for companies that choose to ignore them. Penalised sites can recover, but for many, the decrease in traffic witnessed when a filter is triggered or a penalty imposed is a serious blow to their operations. By keeping track of Google’s policies and keeping your users in mind in both on-page technical SEO updates and your off-page strategy, you can maintain your rankings and ensure long-term profitability for your online business.

While Google is savvier today than ever before, it hasn’t always been so in control of the results it served to searchers. Let’s have a look at how the search engines came to need such robust and stringent measures to control web spam…
THE ORIGIN OF SEARCH

While the history of the internet would form a book in its own right, we can trace the origins of search all the way back to just after the creation of the first electronic computing systems in the 1950s. In the 1960s, Gerard Salton, a professor of Computer Science at Cornell University developed, with his group, the SMART Information Retrieval System (System for the Mechanical Analysis and Retrieval of Text). It was the first system created to retrieve data and was used to search the University’s vast electronic archives.

Skip forward almost three decades (during which time the internet was in its formative years) and the Archie search engine was born. Archie was a tool for indexing FTP (file transfer protocol) archives, allowing people to find specific files on the internet. Created at McGill University School of Computer Science in 1987, it is considered to be the first search engine, although it was only ever used within the confines of the University. Both search tools, SMART and Archie, were extremely primitive, only allowing people to find files if they knew the exact name of the file they were trying to retrieve.

Separate to the search protocols springing up around the world, in 1980, Tim Berners-Lee, an independent contractor at the European Organisation for Nuclear Research in Switzerland, had started to consider the problems of sharing data and information with his physics peers globally – particularly problematic given there were no common machines or presentation software.

In March 1989, Berners-Lee wrote a proposal for ‘a large hypertext database with typed links’, but it generated little interest. His boss however encouraged him to begin implementing his system on a newly acquired NeXT computer workstation. Berners-Lee considered several names for his invention, including ‘Information Mesh’ and ‘The Information Mine’ (which was turned down because it abbreviated to TIM, the founder’s name). Eventually they settled on ‘World Wide Web’ and the “Web” as we know it was born.

By Christmas 1990, Berners-Lee had built all the tools necessary for a working web, including the HyperText Transfer Protocol (HTTP) and the HyperText Markup Language (HTML). On August 6th, 1991, the web became a publicly available service on the internet – its aim was to help share data and create links to information that could be accessed from anywhere.

THE BIRTH OF THE BOT

By December of 1993, three fully fledged ‘bot-fed’ search engines had appeared on the web: JumpStation, the World Wide Web Worm, and the Repository-Based Software Engineering (RBSE) Spider. Each was a primitive version of what we consider to be the early search engines.
JumpStation was the first search engine that behaved and appeared to the user in a similar way to what we are used to; however, it gathered info about only the title and header from web pages and used a simple linear search. It shared the same basic concept as Google does in that it used an index which was searched using a user-entered keyword query, presenting its results in the form of a list of URLs that matched those keywords.

The WWW Worm indexed titles, meta descriptions and keywords and URLs; however, both search engines’ main failing was that they listed results in the order that they found them, with no discrimination or care for accuracy or relevance. Equally, because these early search algorithms did not do any real link analysis or cache full pages of content, if you didn’t know the exact name of what you were looking for it was extremely hard to find it. Thus usability at this stage was very poor.

**SEO IS BORN AND SO IS SPAM**

In 1994, WebCrawler was formed at the University of Washington. It was the first search engine to crawl the entire contents of a file as well as file names and metadata, and also offered advertising in a similar way to how Google does today. It was so popular though that few people could effectively use it during the daytime.

The introduction of WebCrawler may well have paved the way for the initial concept of SEO. The way it indexed sites opened the results up to corruption and manipulation, as people caught on to the fact that by keyword-stuffing their pages’ content, they could appear more frequently in its search results.

From here, more sophisticated search engines began appearing; Lycos, AltVista, and Infoseek to name a few. Infoseek was the first search engine to let people submit pages to the index in real time (this was in 1994), and this popular feature was a search spammer's paradise. Despite each having its own developments and adaptations, like AltVista's natural language searches, many were still only relatively small-scale in terms of the number of pages indexed.

Lycos started with a catalog of just 54,000 documents, but within two years it had extended its catalog to more than 60 million documents (more than any other search engine at the time). *(Lycos, by the way, is still in existence today, owned by Daum, Korea’s second largest search engine).*

It was around this time that search engines started to adapt the way they indexed sites, taking into account the number of links pointing to a site and the relevance of the anchor text surrounding those links. This, however, led to an increase in web spam, and as a result SERPs were largely unclear and unhelpful for searchers.

Interestingly, Yahoo! which was also founded in 1994 by David Filo and Jerry Yang, initially chose to outsource their search service to other providers, considering it secondary to their directory and
other features. (It wasn’t until 2002 that they realised the importance and value of search and started aggressively acquiring related companies).

SEARCH GETS SOPHISTICATED

1996 was a pivotal year for search. Larry Page and Sergey Brin, both Stanford computer science grad students at the time, began collaborating on a search engine they called BackRub, so named because of its unique ability to analyse the "back-links" pointing to websites included in its index.

BackRub operated on Stanford servers for more than a year, and at the start, their technology wasn’t taken seriously. Despite suggesting that ‘not all links on the web were equal’, and proposing a search engine results page (SERP) that used links in combination with their authority and reputation to rank results, they were unable to sell the idea to others. BackRub ranked pages using citation notation theory. If someone cites a source they usually think it is an important one; hence, in digital terms, links acted as citations. In the PageRank algorithm Google is famed for (named after Larry Page), links counted as votes, but some votes were more influential than others. The ability to rank and the authority to vote for others depended upon your own site’s authority: how many people link to you and how trustworthy those links are.

In 1997, the pair decided that BackRub needed a new name. After some brainstorming, they settled on Google, a play on the word “googol,” a mathematical term for the number represented by the numeral 1 followed by 100 zeroes. The name reflected their mission to organise a seemingly infinite amount of information on the web.

In 1998, shortly after Google’s formal incorporation, PC Magazine reported that Google ‘has an uncanny knack for returning extremely relevant results’ and recognised it as ‘the search engine of choice’ in its Top 100 Best Sites. Sergey tried to sell their PageRank technology, but nobody was interested in buying or licensing it.

MSN Search was launched around the same time, but Microsoft did not get serious about search until after Google proved the business model. Until Microsoft saw the light, they primarily relied on partners like Overture, Looksmart, and Inktomi to power their search service.

In 1999, AOL selected Google as a search partner, and Yahoo! followed suit a year later. Google has grown and gained market share year by year ever since. It only operated in English up until May 2000, when it rolled out its first foreign language versions: French, German, Italian, Swedish, Finnish, Spanish, Portuguese, Dutch, Norwegian, and Danish.

Amidst this vast array of innovation and development, the search industry saw numerous smaller services come and go – many struggling to balance the need for relevance and quality with the commercial aspects of running a search engine profitably.
GOOGLE: CLEANING UP THE WEB

Site owners and spammers were quick to cotton on to Larry Page’s PageRank algorithm, and it was this that caused the onslaught of paid linking practices that tarnish the SEO industry even now. SEO at this time was essentially synonymous with spam and black hat trickery.

The game of cat and mouse had begun. Searchers wanted relevant results, Google wanted credible, authoritative results, but spammers were just interested in being top of the SERPs and getting more traffic. They would do anything they could to use Google’s citation notation approach to their advantage and paid linking became rife. In addition, spammers used an array of so called black-hat techniques to boost their rankings, including keyword stuffing pages and putting white text on a white background to deceive the search engines.

Google’s growth continued and so did the uptake of spam tactics among the webmaster community. In 2003, Google launched a significant update – the ‘Florida’ update.

Florida caused a serious ripple in the SEO community – it was the update that put Google updates on the map. Many sites lost ranking, and business owners were furious. The update spelled the beginning of the end for 90s SEO tactics and started to turn the industry upside down.

It wasn’t as well publicised as updates today are, but it appeared to affect sites employing repetitive inbound anchor text on links, keyword stuffing, and sites whose links were not backed up with the relevant vocabulary on its pages. In short, a vast number of sites, many of which had ranked at or near the top of the results for a very long time, simply disappeared. This update was the first update that aimed to make SEO complicated in order to try to deter web spam. Before Florida was rolled out, all you needed to do was buy links with your target keyword in the anchor text, and with enough repetition your rankings were almost certain to improve.

Florida had shaken old-school SEO techniques, and spammers soon developed new ways to manipulate Google’s SERPs. People found new intricate and creative ways to pass on PageRank, and soon people were using less overt tactics like buying in links in bulk or using link networks, link exchanges and link triangles to obtain authoritative links in order to boost their rankings.

PAYING YOUR WAY TO THE TOP

In 2005, Google realised it needed to try to reduce unethical paid linking practices that had become the norm among spammers. While it was technically OK to buy links (e.g. for display advertising), the purchase of a link could not be with the intent of influencing the SERPs. To answer the problem, Google rolled out the “nofollow” link and encouraged webmasters to use it on their site where relevant.

Initially the rel="nofollow" attribute was tasked with cleaning up comment spam and paid links. The nofollow link would act as a message to Google’s bots to say that any such links shouldn’t influence
the rankings of the linked site. The nofollow link wasn’t a negative vote for the website as such; it was just a way to make sure that spammers received no benefit from link spam. Google encouraged people to use the attribute anywhere that users could add links by themselves, and also in any genuine paid advertising links. The implementation was in theory a good idea, although Google relied solely on trust to ensure webmasters used the nofollow links appropriately.

The introduction of Google’s nofollow link meant that buying links for the purpose of buying PageRank became trickier, and it was around this time the concept of “link building” was born. Generating links for the purposes of SEO alone became a laboured, manufactured process and people were forced to think of new ways to obtain much needed “followed” links.

Between 2005-2007 Google was combining different forms of media to enrich its results, moving away from a pure algorithmic approach to a more hybrid approach. It added maps, local listings, scholarly sources, email, calendars, and video (with the acquisition of YouTube in 2006) to name but a few. By 2007, Google was also actively mixing recent news into its organic search results. This meant link building was able to become ever increasingly diverse although many still opted to take shortcuts (like link buying) wherever possible as it was much less labour-intensive.

At this point in time, Google was saying mostly the same things as it does today; its guidelines were largely the same. However Google couldn’t enforce its algorithm, and its only real option at this point was to take manual action against individual cases that came to its attention in the hope of scaring webmasters out of their black hat habits.

WEBMASTERS GAIN CONTROL

Another significant change Google introduced was in 2009 when it brought in its rel=canonical tag. Literally meaning the ‘legitimate’ or ‘official’ version, ‘canon’-ical tags were introduced to help Google’s bots know which page was the original version of your content and eased the congestion duplicate content caused users on the SERPs. It was focused on helping search engines understand your site structure better and allowed webmasters more control over which of their pages were visible in the results.

This was one of many updates Google introduced that tried to mimic a human’s logical view of a website. Google’s bots were unable to make decisions on which of two almost identical pages were best suited to the search query typed so it used canonicals to help make these decisions. Often, through no fault of the webmaster, identical pages might be accessible via different URLs, and canonicals cleared up any confusion over which to display and improved the quality of the results. It had a big impact on reducing duplicate content in the SERPs, which ultimately resulted in a better experience for the searcher.

LINK EARNING

It wasn’t until the mid to late 2000s that the concept of link earning rather than link buying had really taken off in SEO circles. Link earning had happened incidentally up until this point but few SEOs
were fully engaged with its merits. Early adopters in the SEO world now started to realise, however, that there was little future for only optimising for the ‘now’ and they started to think about earning links through the provision of interesting content, engagement in key communities online, and generally good digital marketing and promotion that encouraged natural linking and social sharing.

Many SEOs, however, still clung on to the concept of using lower-value techniques, including paying for links, in order to increase their PageRank (regardless of it being explicitly against Google’s guidelines). Unfortunately, however, Google was still unable to differentiate algorithmically between earned and paid links, so the industry continued to apply black-hat techniques to trick Google into ranking sites higher than it should have. These techniques, unfortunately for Google, were still proving effective.

In June 2010 Google rolled out its new web indexing system called Caffeine which was designed to speed up crawling, expand the index, and integrate indexation and ranking in almost real-time for authoritative sites.

Caffeine provided ‘50% fresher results’ and was the largest collection of web content offered by Google to date. Google had overhauled its infrastructure to make the indexing process quicker and the search results more fluid. To keep up with the evolution of the web and to meet rising user expectations, Google had needed the much better indexing system Caffeine provided.

Google’s accompanying update at the time read:

“Content on the web is blossoming. It’s growing not just in size and numbers but with the advent of video, images, news and real-time updates, the average webpage is richer and more complex. In addition, people’s expectations for search are higher than they used to be. Searchers want to find the latest relevant content and publishers expect to be found the instant they publish.”

The old index, to put it in perspective, had several layers, some of which were refreshed at a faster rate than others. The main layer would update every couple of weeks. To refresh a layer of the old index, Google had to analyse the entire web, which meant there was a significant delay between when they found a page and when it was available to users. Caffeine meant Google analysed the web in small portions and updated its search index on a continuous basis, globally. As it found new pages, it was able to add these straight to the index. Caffeine impacted the speed and freshness of results but also allowed Google to incorporate other signals, such as social signals, into the indexing and ranking process.

HUMAN INTERVENTION & PANDA

Over time, Google had reached an impasse – web spam was still a serious issue that damaged the searcher’s experience. Despite offering clear advice in the form of its Guidelines, some sites were still actively engaged in falsely manipulating the SERPs and Google didn’t like it.
Google was forced to take more decisive action against site owners and SEOs who were spamming the web but in 2011 its algorithm still couldn’t easily identify spam on a mass scale. Human moderators however could. Manual penalties occurred, mainly when companies were openly criticised by experts who started to investigate seemingly unfair or unlikely SERPs.

**The Dirty Little Secrets of Search**

The most famous example of this was brought to the fore by *The New York Times* on February 11th, 2011 in their article titled “The Dirty Little Secrets of Search”.

Author David Segal questioned why, in the U.S. SERPs, was popular department store and online retailer J.C. Penney appearing at the top or near-top of the SERPs for almost every product they stocked. While with over 1,000 stores they were influential, it seemed unlikely they were the most relevant site for so many search queries, from ‘dresses’ to ‘home décor’, ‘comforters’ to the overly specific ‘grommet top curtains’. The retailer was even outperforming brands, e.g. beating Samsonite to position 1 rankings for ‘Samsonite carry-on luggage’. Suspicious, Segal employed an expert in online search, Doug Pierce of Blue Fountain Media in New York, to investigate.

What they found was a shock to everyone. Someone (at the retailer’s SEO agency) had paid to have thousands of links placed on hundreds of sites scattered around the web, all of which led directly to JCPenney.com. Worst of all, the links were often irrelevant, links for ‘dresses’ for example tacked to the bottom of nuclear engineering sites, casino sites, and property portals. J.C. Penney’s amazing results were derived from their agency’s use of deceptive black hat methods, which while not illegal, were not within Google’s stringent guidelines. Pierce described their optimisation as the most ‘ambitious attempt to game Google’s search results’ that he had ever seen. The findings, once publicised, shook people’s trust in Google’s SERPs and opened up the intriguing world of ‘black hat’ optimisation to the world. J.C. Penney’s top-ranking days were soon to come to an abrupt end. Google immediately issued a manual penalty (cited as corrective action against Penney) lasting 90 days. The penalty saw the site drop dramatically from the SERPs for all its key terms. ‘Samsonite carry on luggage’ went from #1 to #71, ‘living room furniture’ dropped from #1 to #68 and the impact on revenues was undoubtedly hugely significant.


Panda, arguably Google’s most famous update, rolled out on February 23, 2011. It was a massive game changer for the SEO industry and focused on stamping out spam link networks, content farms, link directories and thin content. A major update, it rolled out over many months, hitting Europe in April 2011.
Content farms allowed anyone to upload content, without moderation, and solely for the purpose of including a link. They were useless to a genuine searcher and caused an influx of low-quality content onto the web. Thin content, in a similar vein, was any content (often found on blogs) that was poorly written, suspiciously short, and included a lot of white space. The aim of thin content was to optimise for long-tail keywords and create doorway pages to other pages on a site funnelling traffic to a less relevant or keyword specific conversion page. Google wanted to emphasise the importance of quality content that was deep, engaging, and that encouraged social sharing, genuine comments, and interaction amongst readers - and Panda was a positive step toward this.

Panda also took issue with duplicate content and sought out sites that had lazily copied meta tags, titles, and descriptions on to multiple different pages.

Not only did Panda devalue the pages found to be offending in its SERPs, it spread the negative effect to the entire site and went as far as to also devalue any site they linked out to as well. Anyone previously engaged in unethical practices like buying links from content farms began to worry, although few at this stage took it upon themselves to proactively remove problem links.

**IMPROVING THE SEARCH EXPERIENCE**

Another significant change to affect site owners was the introduction of schema.org in June 2011 – the new initiative from Google, Bing, and Yahoo! that created a common vocabulary for structured data markup on web pages.

Schema.org allowed search engines to understand and parse certain types of specific information. It used this additional information to help add to the user experience. Site owners could improve how their sites appeared in major search engines by tagging people, addresses, telephone numbers, reviews, products and other important elements of their content in a way that the crawler bots could identify. Their aim was to make results richer and more informative.

In 2012, Google rolled out its page layout algorithm update which focused on correcting the amount of content presented on the page when a result was clicked. They did so after receiving complaints from users who said it was sometimes difficult to find the actual content thanks to the often unnecessary amount of advertising ‘above-the-fold’. Google wanted to encourage webmasters not to litter their site’s prime real estate with ads which would be detrimental to the user experience.

Sites that didn’t have much content above-the-fold were affected by this change, although it only noticeably affected less than 1% of searches globally, mainly affiliate sites.

Google reminded web developers, webmasters, and SEO’s that their overall aim was still to focus on delivering the best possible user experience. Increasingly it tried to pre-empt the searcher’s
intent and continued to roll out updates, like Venice (also occurring in 2012) which aggressively localised organic results and tightly integrated local search data into its SERPs, in order to improve the user experience and the relevance of returned results.

*It just so happened that the Venice update coincided with the long awaited rollout of ‘Search + Your World’ or Google+, Google’s own social media site that was set to rival Facebook.*

**PENGUIN**

Next up was the Penguin update, which came into play in April 2012. It was a watershed moment for SEO. This key update was dedicated to trying to identify and quash black-hat webspam. It targeted sites violating Google’s Webmaster Guidelines and undertaking any form of over-optimisation, including keyword stuffing, cloaking, link scheme participation, deliberate use of duplicate content (including article spinning) and unnatural links. It focused on expelling spammy sites from the results, like those inserting vast quantities of useless footer links, those overusing keywords in menus, and those with unnatural backlink profiles that relied on the overuse of exact-match anchor text. The scope of Penguin varied by language; languages with more webspam saw more impact.

Instead of killing authority generated by unnatural links or over-optimisation, the Penguin update meant the site actually suffered as a result. So-called “toxic sites” also passed on their toxicity to others, meaning sites suffered by association. Because many of the black-hat techniques targeted in this update often went hand-in-hand with buying links, the concept of paid linking was again under scrutiny.

This presented a new problem for spammers, who were being forced to proactively clean up their act. No longer could they hide from their wrongdoings, and ignorance was not a defence. SEOs and site owners frantically went about trying to identify and undo their past bad work. Effectively, Google had done a great job of making site owners themselves take responsibility for cleaning up the web!

The emphasis on earning links had never been greater. Increasingly webmasters needed to accept they could not manipulate the SERPs, and their ranking was reliant on two key things (as is still the case today):

1. Good quality, unique and customer-centric content.
2. A user-friendly, logical site structure and content that is easily navigated by both users and search engine crawlers alike.

Reputation, in terms of links, was once again earned, not bought, through general good marketing, in line with Google’s original citation theory. There were significantly fewer, if any, loopholes for SEOs to exploit, and previously resistant SEOs began coming around to the idea that they would have to work with Google, embracing the clean-up operation it endorsed, rather than working against it.
While we’ve focused on the negative impact these updates had on the black-hat SEOs and spammers, it should be noted that many sites that had historically always put users at the heart of their SEO activities prospered from the updates. By clearing away the spam, their sites rose in the search engines, and they were rewarded for their quality content and authoritative links.

A SIGN OF “BIG THINGS” TO COME

Today, Google is still working to improve the relevance and quality of its search results. In 2013, it continued to roll out extensions to its Penguin update and we are on for ‘a big summer’ of changes according to Google’s Matt Cutts.

Penguin 2.0 (technically the fourth update to Penguin), which started in May, is said to be a major upgrade to the algorithm. Still only focused on targeting black-hat SEO and those who try to buy their way into search results, Cutts suggested in his latest bulletin that these updates will be valuable to small and medium businesses who work hard to build compelling content and user-friendly websites.

Among a host of key changes Cutts has hinted at, it is understood that paid content and linked advertorials that violate quality guidelines are going to suffer. Sites should fully disclose when content is paid and not editorial, and such content should not affect rankings.

In addition, SERPs rife with spam (for example SERPs served for ‘payday loans’) are likely to witness even tighter filtering than ever before in a bid to clean up more of its results over the coming months. The on-going Penguin update will become increasingly aggressive, and there are murmurings that Google is working on ways to enhance the sophistication of its link analysis too. Keep your ear to the ground for what comes next…
A TIMELINE: THE EVOLUTION OF SEARCH

1960

SMART information retrieval – late 1960’s

1970

1st Search Engine - Archie

93 – 1st ‘all text’ crawler

1990

Predominant SEO hat colour

1. Early SEO was relevant - filenames & meta
   EASILY MANIPULATED

2. ‘all text’ crawling -> KW spamming born

3. 94-96 – inc. link #’s as quality signal

4. Link spam becomes prevalent

5. Link spam evolves -
   - Paid linking from ‘good sites’
   - PageRank manipulation - networks

1995

Search quality poor,
littered with spam

1998 Revolution - Google cleans up search

1st ‘all text’ crawler

91 - Invention of www -
primitive web search soon followed

Lycos

Use of statistical word relationships

Infoseek

Google!

Larry & Sergey invent
“Backrub”

Based on ‘citation notation’ -
not all links equal
TECHNICAL SEO TIPS

With Google’s algorithm actually changing 500+ times a year, keeping up to speed with what you should be focusing on is tricky. Many updates are minor adjustments and go unnoticed, but every few months a major change to the algorithm will be released that can have a massive impact on the search engine results pages and your rankings.

To maximise the results you get from your wider SEO campaign (and in some instances to avoid incurring penalties), you need to address the crawlability of your website and optimise it for users at the same time. Below are our key technical SEO considerations...

CLEANING UP YOUR ACT

Following Google’s more recent updates - now is the time to clean up your act. Often even the smallest on-site changes can have a significant impact on your rankings, so tackle these issues as a matter of priority if you haven’t done so already:

Quality content

Make sure that your website is rich in high-quality, unique, customer-centric content. If areas of your site are quite thin or were written only to help you rank for long-tail keywords, then think seriously about improving or removing them. Always write for the user and consider your wider content and online PR strategy in your plans to create new content for online or offline use. In order to effectively optimise a piece of content, just write naturally!

Over-optimisation

If you know you are guilty of over-optimising your pages, by stuffing keywords into your page titles, meta descriptions, menus, internal link titles, or in the body of your content, then this needs to be rectified. Issues arise when specific terms are utilised too frequently, and websites that are thought to be purposely over-optimising may have a keyword rankings filter applied which prevents the website from ranking to its true potential.

Page titles at the individual page level are probably the most significant on-page factor. Don’t repeat variations of your head terms in your title tags, and think about how it might appear to a user if you are in doubt. On most occasions, the page title appears in SERPs as the hyperlink you click to visit the site, so it should be conversion-focused, logical, and interesting for a human to read.

Duplicate content and canonical tagging

Duplicate content can be present on a site for genuine reasons, you may have localised language pages (e.g. US English and UK English pages), or simply have pages which are accessible by users and search engines via multiple complex product paths with differing URLs. Tracking parameters and session IDs appended to URLs are also common causes of duplicate content issues which present a problem for search engine bots trying to index your site.
Identify and remove all duplicate content on your site, and if you genuinely require duplicate pages, use canonical tags correctly to identify them.

Canonical tagging provides you with complete control over the URL that you wish to be returned in search results. Essentially they inform Google where the original or preferred version of your content can be found, and notify Google to ignore duplicate versions of the page. In addition, canonicals ensure that link popularity is consolidated to your preferred version of any given page.

**Links from bad neighbourhoods**

There is no escaping any poor practices you might have dabbled in in the past. You may not even have been aware of the links being built to your site, but either way, now is the time to face up to them. Use Moz’s Open Site Explorer, Majestic SEO, and Webmaster Tools to identify all the sites that are linking back to yours. If you discover a large volume of links coming from link or article directories or other low quality sites it is recommended that you try to get these links removed.

While you may not be able to remove each and every toxic link, document your attempts, and you can always, as a last resort, submit your request via the Google Disavow Tool for consideration.

It is best not to have any kind of association with these types of low-quality sites (also known as bad neighbourhoods) as Google may consider you guilty by association. However don’t assume if you have had a drop in rankings that your backlinks are solely to blame – there may be other contributing factors that got you flagged. A thorough investigation is required. Additionally, while we advise you try to remove toxic links, don’t throw the baby out with the bathwater, so to speak – not all your links are going to be toxic. Time and time again, we see companies sending blanket emails to webmasters asking for perfectly good quality, useful, and contextually-relevant links to be removed. This is more detrimental than beneficial, so if you are unsure which link to remove, get an expert to help you.

**Over-optimisation of backlink anchor text and internal links**

By using links (internally or externally) you can confirm to Google that the page being linked to is about a topic similar to the page the link originates from. Google, however, is extremely sophisticated in that it can read the anchor text of links (and the words around it) to determine their real relevance, so don’t try to place too much emphasis on anchor text. Google targets sites that have purposely tried to manipulate the SERPs with over-use of exact match anchor text, so let links come with natural and sometimes noisy anchor text (“click here” or “this article” would be considered to be noisy anchor text, for example). By reading the text surrounding a link, Google can give the link context and meaning, so a link from a non-descriptive word with relevant words surrounding it, can be just as effective for SEO.

**Hidden text**

Hiding text does not necessarily relate to having white text on a white background. It is possible for websites to hide text using CSS templates which reveal text behind dropdowns. This is not
recommended if done only for SEO gain. Hiding huge paragraphs of text behind a dropdown at the bottom of a website should certainly be avoided at all costs, as it is unnatural and of no use to a user. Such practices could also damage your site’s chances of ranking well.

Redirects

Ensure that all redirects are implemented correctly. Wherever new pages are being created to replace old ones, ensure that the old page redirects with a 301 permanent redirect.

Footer links

The footer located at the very bottom of a web page is for secondary navigation only (usually reserved mainly for privacy statements, legal information and the like). It is best to avoid placing too many links in the footer because they are likely to be devalued by search engines automatically. If it’s a user-essential link it should be higher up in your page – this way it will give maximum value in terms of SEO.

Site Speed

Google uses site speed as a ranking factor. Site speed reflects how quickly a website responds to web requests, and faster sites don't just improve user experience; recent data shows that improving site speed also reduces operating costs. Assess the file size of any images on your site and reduce them as much as possible. Reduce any irrelevant ‘junk code’ present in your pages and use display advertising sparingly. Too many of these elements will slow page load times and could negatively impact your rankings.

PLANNING STRATEGICALLY

Perform detailed keyword research

Extensive keyword research must be carried out for all areas of your site. Ensure that each page of the site is optimised for the most relevant and highly-converting keywords. Think strategically about what users might be searching for and the intent that is associated with different combinations of keywords. Use Google’s Keyword Planner Tool to identify new keyword opportunities and help you decide which terms are best. You should also try to bring any PPC data you may have into your discussions about which keywords to target on your website.

While your website’s pages should reflect your chosen keywords, don’t write thin pages of content just for the sake of mopping up long-tail keyword searches. Unless the page adds value to a prospective customer or visitor, don’t upload it.

Make your site structure logical

Ensure your site’s navigation follows a logical structure; in terms of categorisation, navigation, and folder paths. Think about which pages are most important to your users and ensure they are easily accessible, in high/early positions in your menus. What the user is likely to be most interested in will then also be the first page Google’s bot crawls.
BEST PRACTICE RECOMMENDATIONS

Sitemap

Keep both your XML and HTML sitemap up-to-date with any changes that are made to the website. Their purpose is to help search engine spiders and your users find all of the site pages easily, so don’t neglect it. An XML sitemap file can contain no more than 50,000 URLs and must be no larger than 50MB when uncompressed. If your sitemap is larger than this, break it into several smaller parts. These limits help ensure that your web server is not overloaded by the serving of large files to Google. Specify all URLs using the same syntax and never include session IDs in submitted URLs.

Image optimisation

By optimising images correctly with alt tags, you will increase your likelihood of ranking for long-tail keywords. Alt text is the ‘alternative text’ that displays in the browser if the image fails to load or the user has images turned off. All content should be created with human consumption in mind, so label your images accordingly. Linked images should always use alt text because this is the only indication to the search engine of the context of the link.

Optimise PDFs

PDF files do not rank as well as HTML files in their own right. While you can write PDF files optimally for search engines, to get the best results, it is better to recreate the page in HTML.

If PDF files are essential, ensure they are created in a PDF generator that creates readable text (for example the full version of Adobe Acrobat). Some PDF generators create a file that is similar to a large image file and is not readable by the search engines. Make sure to use keywords carefully, as you would with any other piece of online content. When the PDF is generated, there will be an opportunity to input properties about the file. The title field is the most important. The title field will be used by Google as the title in the search results so make it conversion-focused. The description will be a snippet body copy itself. The keyword property and description field are ignored by Google in the same way as for HTML files, but since algorithms are prone to change, and because Google is not the only search engine in use, we recommend populating both.

Authorship

Correct use of Google Authorship (connecting your content to your author’s Google+ profile by correctly implementing rel=Author tags) has lots of value from an SEO perspective. First, it tells Google you are a real human being and adds trust and credibility to your content. Second, it helps you claim the credit for content or expert commentary you may provide to sites other than your own. (It can also help beat plagiarism as the original author always gets credit). Third, it is proven to boost click-
through rates, thus increasing the traffic to your site. Search results with rich snippets (like that of a Google Author headshot) are more prominent in the SERPs and do attract more clicks than results without.

**Structured data**

Structured data (also known as rich snippets) is an on-page mark-up that helps the search engines understand specific bits of content on your site. They are the extra bits of information that you see under search results. Their main purpose is to enrich SERPs and to help make your site stand out from the crowd. There are lots of rich snippets available for you to use, including breadcrumbs, events, music, organisations, people, products, recipes, reviews, software applications, and videos.

Look at the results on the right. The first result is a ‘Rich Snippet’ - it offers users a summary of what’s on the page and why it’s relevant to their query. The second result is a standard snippet without any structured data.

Google suggests using microdata, as outlined at schema.org, to mark up your content, and once marked up, test it using Google’s structured data testing tool.
CONTENT & ONLINE PR

Also known under the umbrella term, ‘off-page SEO’: website promotion, content marketing, and securing relevant links from other sites in order to improve a website’s rankings in the SERPs are all integral parts of your SEO strategy.

LINK EARNING AND DIGITAL ENDORSEMENTS

It’s no secret that links are the major currency of off-page SEO – Google and other search engines treat these as endorsements for your site, so the more highly authoritative endorsements your website receives, the more value it is awarded by search engines. Not every link is equal though. For example, a link from the BBC to your website is worth more than one from a blog that boasts only a minimal readership. This is because the BBC is more trustworthy, reputable and has more endorsements itself, factors which are reflected in its PageRank.

The most valuable links from an SEO perspective are those obtained from highly trustworthy sources, such as links from university sites (.ac/.edu), governmental sites (.gov), and charities (.org). Links from these sites are most beneficial, and as with all quality links, they can only be secured by offering useful, unique, interesting, and newsworthy content.

This customer-centric approach to SEO is more labour-intensive than historical black hat techniques were, but it is proven to deliver long-term results and will ensure harmony is maintained in the search engine ‘ecosystem’. It is much better for a site’s SEO to take longer creating one highly impactful and useful piece of content in order to maximise its ability to attract links, than to take a quick-fix scattershot approach churning out low-quality content that no one will want to read.

As previously mentioned, before Google improved its ability to automatically devalue a site that did not adhere to its guidelines, buying or spamming links (for the purpose of influencing the rankings) was commonplace among some in the SEO community. It did produce high rankings on the SERPs although the improved rankings were often short-lived. Today, however, these practices are not only ineffective, but will also likely result in a penalty which would be very detrimental to your website’s organic rankings.

The long-term health of your search rankings should always be the priority, and strategic, creative content that is designed to be able to attract links from these high PageRank sites naturally is the most effective way of not only improving rankings, but also safeguarding them. Websites with a healthy, natural and relevant backlink profile are harder to dislodge from their high SERP positions.
LINKS APPEAL TO GOOGLE’S SENTIMENTAL SIDE

Off-page SEO is done to prove to Google that your website is worth more to the searcher than your competitor’s site is. This proof is supplied by backlinks. Users may love your content, but that emotion alone means nothing to a search engine. The only way Google can determine your site is trustworthy and authoritative is to see people are linking to it digitally.

Google uses more than 200 factors to rank websites; site and page quality are the most important to off-page SEO. In the past it was all about the links, with reputation, trust, and authority being mere afterthoughts. Now they’re all imperative.

Today, the commonly used term ‘link-building’ is somewhat redundant. The building of links can be done artificially by machines and is devoid of genuine human endorsement. To improve your SEO you need to earn your links and interact with real human influencers. ‘Link building’ also implies a certain degree of manufacturing and is commonly misinterpreted as being about quantity of links rather than quality – a principle that adds no benefit to the web or its users.

After all, Google’s algorithm aims at providing an answer to a question. If you search for something online you’re seeking an answer; the price of cinema tickets, the weather tomorrow, your nearest petrol station, etc. Google’s entire business premise hinges on returning valuable information to the user, and that’s why human endorsements are key to a successful off-page SEO campaign. Humans know best which sites are valuable resources, are interesting, and answer their original question best.

SEO NO LONGER AN OUTSIDER

Digital or online marketing techniques are increasingly coming into line with traditional marketing practices and as such SEO as a term is beginning to be phased out of many marketer’s vocabulary. Thanks to its sordid history, it’s no surprise everyone is keen to replace it with a new buzz word. There are several replacement terms vying for its berth, with ‘content marketing’ and ‘online PR’ (and to some extent ‘inbound marketing’) currently being the most popular.

**Content marketing:** “The creation, promotion, and sharing of content in order to communicate with a particular audience and provide the audience with consistent, on-going valuable information.”

**Public relations:** “The practice of managing the spread of information between an individual or organisation and the public.”

By definition these two practices encompass what Google wants from webmasters. Add links into the equation and you have what was previously known ubiquitously as off-page SEO. This change of phrase is a recent phenomenon (as you can see on the next page) and reflects the switch from focusing on links alone to the creation of high quality, relevant content that naturally attracts links.
This means that for your entire marketing strategy to be successful, the days when SEO was siphoned off as a separate campaign are long gone. Instead, SEO needs to be incorporated into your overall strategy. Off-page SEO, PR, and other online promotions should be run in tandem with one another as a holistic approach to marketing. For example, if your PR agency is running a campaign that is generating a lot of coverage, then the SEO agency should be informed of this in order to get in touch with the publications where you’re mentioned and turn these mentions into links. If you’ve got a great story or great content, your SEO partner needs to know. The same is true of an advertising campaign that is likely to lead to coverage online. Teams should work together to make the most of one another’s ideas. Your SEO should know what your PR calendar looks like, and your ideas for content should be shared so that your SEO team can work with you to develop ideas in line with your wider business goals and ultimately secure links from your existing work.

The results of a holistic approach between all your agencies are far greater than if they are fractured and separated, as the two examples below show.
Case Study Example:
RAM Tracking Rush Hour Traffic Map

The Idea
To create an interactive map of traffic hotspots that could be hosted on RAM’s website. Using speed and location data obtained by the GPS vehicle tracking company, we analysed 7 million start and end points comparing the distance travelled during the rush hour commute of 8am to 9am with the later time of 11am to 12pm. The findings were formulated into an interactive Google map that could be regularly updated by adding new data.

The Strategy
Working in close conjunction with RAM’s PR agency, Search Laboratory compiled a strategy to outreach this content to relevant media publications.

- Particular emphasis was placed on tailoring the content to individual cities and regions in order to target local newspapers.
- The team focused on national media outreach to encourage nationwide coverage.
- Early communication with relevant media outlets, blogs, niche publications as well as experts and influencers on social media identified interest.

The Results
- The content was picked up by major national newspapers in the UK, including The Sun, Daily Mail, Metro and Daily Star as well as the BBC and ITV.
- The localised content was covered in regional publications such as the Manchester Evening News, The Birmingham Mail and The Yorkshire Post.
- Ram Tracking received over 20 high authority backlinks from a range of influential sites.
Many people overthink off-page SEO, which ultimately boils down to a simple equation:

**Great Content + Exposure = Natural Links.**

**WHAT MAKES ‘GREAT CONTENT’ GREAT?**

Thorough research is the first step on the road to creating great content. You must know what your website’s audience is interested in, and importantly, what they are not. On the surface this can appear to be a simple task. For example, if you are a women’s fashion retailer, then content on dresses is going to be in your readers’ sweet spot.

But do you know what type of content your audience reacts to? Are they turned on by list articles (e.g. 7 things you don’t know about X) and turned off by longer, more in-depth prose? Will they share content incorporating videos more readily than infographics?

Thorough research, and an element of trial and error, will help you answer these questions. Google Analytics should also be your friend when it comes to researching content, as this will take away some of the mystery of your audience’s preferences.

Garnering the opinion of influencers in your industry is one of the most effective ways of conducting this research. So before you put the proverbial pen to paper, you need to talk to influential bloggers and magazine or newspaper journalists to ensure whatever topic you are covering for content will succeed in attracting links. By canvassing these opinions, you can establish relationships, and it is these relationships that lead to natural links.

Rather than providing a sum of money, you are supplying great content that is being promoted within the industry by influencers, and if done right, this should lead to a genuine interest in the audience linking back to the post.

**HOSTING CONTENT VS. GUEST POSTING?**

To maximise a piece of content's potential for attracting natural links, it should be hosted on your own site. This is because if the content is hosted externally, then the audience naturally links to this third-party website to reference the article, meaning your site does not gain as much authority as it would if you had hosted the piece originally.
If the content is hosted on your site, then all the links will point to your website, thus ensuring all the authority flows to the page and on through other pages on your website. Links do not just improve the ranking and authority of the specific page they are pointed to, and that is why it’s critical to attract links to both deep pages and your homepage.

Guest blogging was once very popular within the SEO community as it was a relatively simple way to earn links on third party websites. This resulted in the web being populated with countless irrelevant, thin and poor quality guest posts that had no other purpose except acting as a vessel for a link. Google’s recent updates, Penguin particularly, looked to nip this problem in the bud, so be cautious about the level of guest blogging you choose to engage in. Certainly refrain from providing any content as a guest blog that you wouldn’t be proud to host on your own website.

With this said however, don’t over-react. Content contribution (in the form of guest blogging or otherwise) on third party sites about relevant topics can still be beneficial if the post is genuinely written to add value to their readers. If you are an industry expert you may well be asked to provide expert opinion, but question every post you write. Does it provide the searcher with a unique viewpoint, an answer to their question, and the good quality content they seek?

**SUMMARY**

The industry is changing, as is the terminology, but links remain a constant – it’s just the method for earning them that has shifted. Off-page SEO is categorically a case of quality, not quantity. One link from a high PageRank website is more valuable than hundreds of directory or article submission site links, and if you haven’t already adjusted your expectations of what to expect from your content and online PR teams, now is the time.

A personable approach that cultivates the building of relationships with bloggers, journalists, and influencers coupled with well-researched, creative, quality content hosted on the client’s site is the winning formula for attracting natural links, but it is time-consuming. However, more natural links from authoritative sites = improved rankings on the SERPs = more traffic for your websites = more revenue.

“**The objective is not to make your links appear natural, the objective is that your links are natural.**”

Matt Cutts, Head of Web Spam, Google.
CONCLUSION

SEO: WHAT DOES THE FUTURE HOLD?

Search is becoming increasingly personalised, and the more information Google has about you, the better the results it will be able to present to you. It is already using your search history, your preferences and interests and those of your friends, your location, and your demographic to determine personalised results.

In time, Google will serve results predictively, based upon what it thinks you are likely to want at that time. For example, it’s your sister’s birthday is next week - Google will know your sister is your sibling because of the presence of certain social indicators, Google will also know that your sister likes crafts and jewellery, therefore Google might suggest that you buy some beads and charms for your sister in advance of her birthday.

Search, it is predicted, will therefore become less about targeting certain keywords and more about building a brand that people like and endorse (endorsements may come in many forms – links, Facebook likes, tweets, or shares). The more you receive, the more chance there is that your site will be shown in relevant searches. Build a brand that people like by providing a good user experience. Offer quality content that engages them, making them link (validly), share, and interact, and your site will be considered reputable and relevant by the search engines.

SEO SHOULD NOT STAND ALONE

The best results from SEO will be witnessed by businesses that take it seriously as a measurable, tangible, and transparent way to generate traffic for their site and leads/sales for their business. A holistic approach needs to be taken. SEO is not a stand-alone marketing element. It must overlap all other aspects of your marketing strategy in order to produce its full effect. SEO delivers the most ROI when it is able to use and build upon your wider marketing and PR activities.

Many companies, however, do not emphasise the importance of this inter-departmental collaboration to their teams, and cross-organisational collaboration is even harder to coordinate. By getting your SEO strategy to dovetail with other key marketing and PR strategies, each activity can be maximised in terms of potential return. Marketing, PR, IT, product, and sales teams will all be affected by SEO and should all have a good understanding of what your SEO teams (be it internal or external) are trying to achieve. SEO is a group process that might start with digital marketing and PR, but which should end with your sales department feeding back information about the quality of leads and volumes of sales that SEO generates. Without this group thinking and valuable feedback, you will be unable to refine and improve your SEO strategy moving forward, which will stifle the revenue you can derive from it.

By mixing the experience and skills found in your own marketing and PR teams with that of a specialist digital agency partner like Search Laboratory, you can create hugely impactful content that can significantly boost your rankings. If you think those internal relationships can be improved, speak to us about what we can offer to your teams in the way of training.
HAVE QUESTIONS?

If you would like to discuss the topics covered in this whitepaper in more detail, please get in touch. We would be happy to help answer any questions you may have:

Tel: (0113) 212 1211
Email: info@searchlaboratory.com
Web: www.searchlaboratory.com
SEO GLOSSARY

Above the fold
The area of your website which is visible to the visitor without them having to scroll up or down. In web marketing terms, it is considered prime real estate.

Algorithm
The technology used by a search engines to determine what pages to serve for a given search query. Search engines utilise many different and ever-changing algorithms in tandem in order to deliver what they deem to be the most relevant and high-quality search results.

Alt text
A written description of an image which is only displayed to the end user if the image is not deliverable. Alt text is an important part of image optimisation because search engine spiders cannot read or interpret images. The alt text tells a search engine what the image relates to and may also be used in website accessibility for the visually impaired.

Analytics (see also Google Analytics)
Analytics programmes assist in gathering and analysing key data about a website, its users, and their behaviour. Many paid analytics tools exist, but Google Analytics is the most commonly used free tool.

Anchor text
The visible text on which a hyperlink is placed within a web page. In the past, much SEO emphasis was placed on anchor text links (search engines do to some extent look for relevancy and similar themes running between linked sites), but today, anchor text has less relevance in the SEO mix.

Authority
Authority is the measure of trust and credibility a site is credited with by a search engine, which is used as a factor when determining the best results to serve to a searcher. Authority and trust are partly derived from the number of relevant inbound links from other trusted sites, which are deemed to be votes of confidence for that site and its content. Many factors are associated with authority, including the number of relevant inbound links, domain age, traffic, and site history, along with the rate at which unique, quality content is added to the site.

Backlink
An incoming link to a web page from a third party website. Backlinks can be “followed” and carry authority from one site to another or “nofollow” which does not.

Black hat SEO
Search engine optimisation techniques that are counter to the best practices outlined in Google’s Webmaster Guidelines are considered black hat. These tactics are aimed at falsely manipulating the search results. Google actively tries to penalise sites using such unethical techniques and regularly updates its algorithms to uncover new and emerging black hat techniques.

Bounce rate
The rate (shown as a percentage in Google Analytics) of visitors who enter a site and then leave it without navigating to or viewing any other pages. A high bounce rate may (among other things) indicate low-quality content, technical and usability issues, slow load time, or inappropriate advertising that promises something that the site’s landing page fails to deliver.

Canonical tag
A piece of HTML code used by search engine spiders to tell search engines which URL is the original version of your webpage. Literally ‘canon’ meaning the ‘legitimate’ or ‘official’ version – a canonical tag is used to help website owners avoid duplicate content (a factor which Google ranks negatively).

Cloaking
A highly black hat, unethical SEO technique that delivers different content to the search engine spider than is seen by a human user in order to trick the search engine and positively influence rankings. Sites undertaking cloaking
are likely to be penalised heavily by Google if caught.

**CMS**
A content management system (CMS) is a programme that allows the publishing, editing, and modification of content from a central interface. It makes webmaster tasks such as creating content or inserting media easier for those who may not be familiar with coding languages such as HTML. Popular CMSs include WordPress and Joomla.

**Code (see also HTML)**
Anything written in a language intended for computers to read, e.g. HTML, JavaScript etc.

**Comment spam**
The posting of fake or low-quality blog comments purely to generate an inbound link to another site - a black hat SEO technique that’s pointless and ineffective.

**Contextual link**
A link within site content that is related and relevant to the topic being covered.

**Conversion rate**
The rate at which a site or page achieves a quantifiable goal as determined by the website owner. Common goals to measure include newsletter sign-ups, forms completed, clicks, sales, or social media shares.

**Crawling**
The process by which a search engine spider (see also Googlebot) discovers new and updated pages to be added into Google’s index.

**Duplicate content**
Content on a website that is identical or similar to that found either on an alternate page on the same domain (or a third party domain if plagiarised). Google does not like duplicate content and it may affect a site’s overall rankings.

**Google AdSense**
AdSense is Google’s contextual advertising network, whereby publishers of all sizes are able to publish relevant advertisements near their content in order to share the profits from Google’s PPC advertising model.

**Google Analytics (also known as GA)**
A popular service offered by Google that generates detailed statistics about a website’s traffic, sources of visits, visitor behaviour and that measures conversions. There is both a free and enterprise (paid) level of service.

**Google AdWords (see also Pay-Per-Click)**
AdWords is Google’s advertising platform and main source of revenue. Offering a pay-per-click model Google AdWords offers text, banner, and rich-media advertising options.

**Googlebot**
Also known as the Google robot, spider or crawler, the Googlebot is a web crawling robot used to help create the search results. Using a huge set of computers to crawl billions of pages on the web Googlebot uses an algorithmic process to determine which sites to crawl, how often, and how many pages to fetch from each. Googlebot’s crawl process begins with a list of websites generated from previous crawl processes. As Googlebot visits each of these websites, it detects links within them and adds them to its list of pages to crawl. New sites, changes to existing sites, and dead links are noted and used to update the Google index.

**Google Keyword Planner Tool**
A useful keyword research tool which estimates the competition for a keyword, recommends related keywords, and tells you what keywords Google thinks are relevant to your web page or website.

**Hit**
Once the standard by which website traffic was judged, Hits are now a largely meaningless metric replaced by ‘Page views’ or ‘Impressions’.

**HTML**
Hyper Text Markup Language or HTML is the mother tongue of the search engines. It is the most common authoring language used to create websites, and defines the structure and layout of a web page by using a variety of tags and attributes.
Impression
The exact number of times a specific web page has been accessed or viewed by a user (a page impression) or the number of times an advertisement has been served (with or without an additional action such as a click).

Inbound link (see backlink)

Index
A database of web pages and their content used by the search engines (noun). May also be a verb, meaning to add a web page to a search engine index.

Indexed pages
The pages on a website which have been indexed by the search engines.

Internal link
A link from one page on a website to another page on the same domain. Often used to add value to the reader, internal links also help search engines to understand what a website is about. By placing links to relevant related pages within the site will help demonstrate the relationship between pages and improve the usability of a website.

Keyword
The word or phrase that a user enters into a search engine to initiate a search.

Keyword density
The amount of times (usually as a percentage of the total word count) that a particular keyword appears on a web page. If this value is unnaturally high (such as with keyword spamming techniques) the site is likely to be penalised by Google (see keyword spam).

Keyword research
The method of determining which keywords are most suited to a specific site in order to help drive relevant traffic. Often used as the starting point for any targeted SEO or PPC campaign.

Keyword spam
Also known as keyword stuffing, this black hat technique involves stuffing content with an unnaturally high number of targeted keywords in order to unfairly influence search engine rankings.

Landing page
The web page that a user lands on when they click on a link in a SERP or in an ad.

Link bait
A web page or piece of content which has the main purpose of attracting incoming links and social media shares.

Link building
Content marketing and online PR activity (also known as off-page SEO) that is tasked with creating natural, high quality and relevant backlinks to a website or page in order to build authority and improve search engine rankings.

Link exchange
The act of building a reciprocal link in return for a backlink to your site. Usually of low-quality, and adding no value, reciprocal links are often considered an unethical, black hat SEO technique.

Link juice
An SEO term for the authority that is passed down via backlinks between authoritative sites.

Meta tags
Code in a web page that gives important information about that page to the search engines. Some meta information may be visible in the SERPs but is not visible on the page e.g. the meta description. It is very important to have unique and accurate meta titles and descriptions for each and every page on your site, because they provide the information that the search engines rely upon to determine what that page is about. Also, they act as the first impression that users get about your page within the SERPs.

Nofollow link
A HTML command found in either the head section of a web page’s code or within individual link code, which instructs search engine spiders to not follow either any links on the page, or a specific link. This attribute is used to prevent a link from passing link authority.
Open source
Software which is distributed freely with its original source code so that developers can modify and provide additional plugins for it as they see fit. Open-source software is often developed in a public, collaborative manner (e.g. WordPress).

Organic search results
Also known as natural search results, organic search engine results are those that are not sponsored or paid for and which are served based on a search engines calculation of relevance to the search query entered by the user.

Page view
The exact number of times a specific web page has been viewed by a user. Note there may be multiple page views per unique visitor.

PageRank (PR)
The value between 0 and 10 assigned by Google’s algorithm, which quantifies the importance it places on individual web pages and websites in their entirety. Hundreds of factors are considered when determining Google PageRank, and it is a logarithmic scale. PageRank is one of many factors used to determine search engine results.

Pay Per Click (see also Google AdWords)
Pay-Per-Click (also known more commonly as PPC) is an internet advertising model used to increase traffic to websites whereby an advertiser pays the search engine (or other publisher) when their ad is clicked.

Redirect
Any of several methods used to change the address of a landing page such as when a page or website is moved to a new domain.

Robots.txt file
A file in the root directory of a website used to restrict and control the behaviour of search engine spiders. Essentially, this file tells search engines which files and pages not to crawl or index. This may include development site's pages, closed offer pages, or pages not to be made publicly available/directly accessible from the SERPs.

Search engine optimisation (SEO)
The process of optimising and improving your site to help achieve a higher ranking in search engine results pages. The overall aim of SEO is to increase the site’s visibility among potential readers and its target audience, attracting more organic search traffic. The techniques used focus on making a website more accessible and crawl-friendly for a search engine and producing high-quality content that encourages natural backlinks from high quality third party sites to aid relevancy and authority. Also described as ‘the science behind publishing information and marketing it in a manner that helps search engines understand your information is relevant to relevant search queries’.

SEM
SEM stands for search engine marketing, and covers all aspects of digital and web marketing, including SEO and PPC. In the USA it is more commonly associated with paid search (PPC).

SERP
The search engine results page.

Site map
A page tasked with linking to every user-accessible page on a website. Its aim is to provide a more usable site by clarifying the structure of the site’s pages for users to navigate.

SMM
Social media marketing (SMM) is the act of marketing and promoting a site through social media channels.

SMO
Short for social media optimisation, SMO is the act of optimising one’s site for social media (e.g. by adding social sharing functionality and direct integration) and promoting its content using social channels (SMM). Similar to search engine optimisation SMO’s goal is to generate traffic and brand awareness for a website.

Spider
The search engine spider (also known as the Googlebot or crawler) is a robot used by search engines to find and add web pages to
their indexes. The name spider refers to its ability to ‘crawl’ web pages (see GoogleBot for more information).

**Splog**
The name for a spam blog which usually contains little, if any, content of value to human readers. Splogs are often machine generated and made up of scraped or spun content.

**Spun content (also known as article spinning)**
A black hat SEO technique whereby an author rewrites all or parts of existing articles and republishes them. Often article spinners use a machine to insert common synonyms in place of the original words, and this produces very low-quality content which is not at all reader friendly.

**Time on page**
The amount of time that a user has spent on one page before clicking to another page. Time on page (along with time on site) can be a good indicator of page quality and relevance.

**White hat SEO**
Ethical, sustainable SEO techniques which conform to Google’s best practice guidelines, and do not attempt to unscrupulously trick search engines and manipulate SERPs. Based on customer-centricity and the importance of quality content, white hat SEO techniques bring increased rankings and long term results.

**WordPress**
A very popular, easy-to-use and open source blogging and website management software platform. Visit WordPress.org for more information.

**XML sitemap**
An XML sitemap (often kept in the root directory of a website) is there to help search engine spiders find all of the site pages easily.