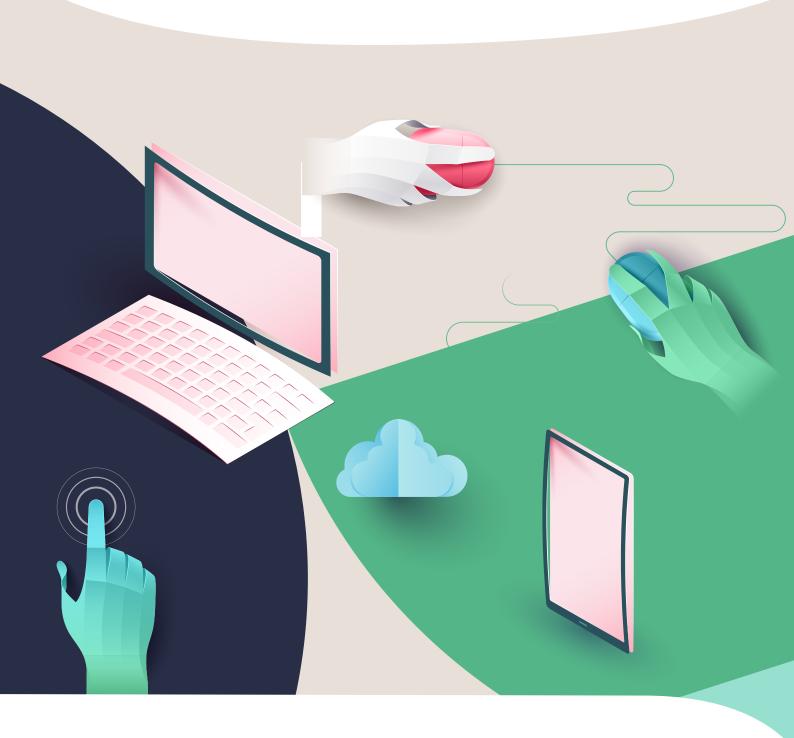
SEO UNCOVERED [REMASTERED]

An Advanced Guide To Organic Search



CHAPTER 2
Technical



CONTENTS

INTRODUCTION	
SITE OPTIMISATION	4
ON-PAGE OPTIMISATION	5
IMPORTANT HTML ELEMENTS FOR OPTIMISING YOUR PAGES	5
Meta title tag	5
Meta description	5
Header tags	5
<h1></h1>	6
<h2></h2>	6
<h3></h3>	6
Image source tag	6
Alt attribute	6
Text link	
Image link	7
Image link (with nofollow)	8
Semantic markup and schema.org	8
REDIRECTS	9
The 404	10
The 'soft' 404	10
URL best practice	11
What makes a good URL?	
CANONICALISATION	12
Common URL duplications	12
Best practice	12
301 and canonicalisation	12
SITE STRUCTURE	13
CONTENT OPTIMISATION	17
How content should be optimised	
THE 'DUPLICATE CONTENT' MYTH	18
Filtration of duplicate content occurs at three stages	19
Do you have an internal duplicate problem?	20
How to fix them	20
How to avoid them	20
Expertise	20

TECHNICAL

"Without a strong technical foundation, creative efforts can (and often do) fail, while successful efforts cannot be correctly measured."

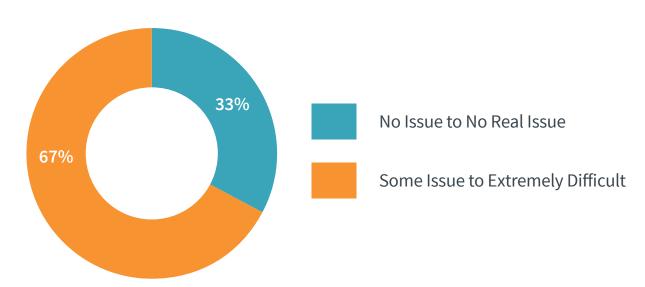
Alan Reeves, Director of Search, Click Consult

INTRODUCTION

With 67% of respondents to Click Consult's most recent survey expressed some level of difficulty with the implementation of technical SEO, it is necessary to approach this subject in some detail. As an agency which prides itself on leading rather than following best practice. We sympathise.

Search marketing is a field that is in constant flux – with each generation of computing enabling higher levels of sophistication to the algorithms of leading search engine providers, the time and effort to keep up with the increasingly complex requirements of technical best practice are extensive.

Technical representing a challenge: 67 vs 33



SITE OPTIMISATION

Ideally this is where the process of technical optimisation will begin. If a brand's website is imagined as a physical container, then it can be assumed that, before adding content, it is necessary to ensure that container is not likely to leak.

Technical optimisation allows consumers that reach for a specific brand's 'container' to see that:

- It can be found: Which means that the site must be correctly indexed. It is during a 'site:search' that problems with site architecture can be discovered, revealing linking issues or Google penalties which cause a homepage to rank lower than expected comparative to other pages on the site.
- It has the right labels: When the brand name or branded goods or terms are searched for, the correct pages should always appear in an order a consumer would expect. If not, this can be another indication that there are penalties or indexing issues.

Though this process involves some relatively simple steps, it is fundamental to the SEO process. If the domain is not operating as it should, all other efforts may be ultimately rendered fruitless. Though this is how optimisation of a domain should begin, there are other steps which are necessary at this early stage to ensure the best results. You should ensure that:

- You are correctly managing content updates, redirects, and domain migrations
- You are using the correct redirects for the required purpose, and that 404 pages are relevant, unique and branded
- Your site is using the correct technologies, avoiding reliance on Flash and Javascript
- Your site's navigation is clear and hierarchical and that there is a well organised and exhaustive XML Sitemap, and inter-page linking.
- There are no or very few unnecessary links and duplicate pages.

The technical process of optimisation consists of multiple steps, repeated multiple times, in a variety of different areas of a website across multiple types of device. For the best results this should be an ongoing process, but the following pages will follow a typical route through the optimisation process, in sufficient detail for any brand to implement a good technical strategy.

ON-PAGE OPTIMISATION

In this context, 'on-page' refers to 'back-end' content, rather than the content visible to a consumer – that which is written with search crawlers, and Google's robot.txt in mind. If the level of technical SEO here is good, there will be greater freedom down the line to be creative with your content

IMPORTANT HTML ELEMENTS FOR OPTIMISING YOUR PAGES

Meta title tag

Though it is ultimately limited by pixel length, your meta title tag should be as accurate, concise and precise a description of your content as is possible in 56 characters. It is one of the most important sources of information for Google's algorithm when it comes to deciding relevance (and therefore affects SERPs position) as well as aiding in the ease of browsing for the consumer – it should always contain key terms relating to the page.

Meta description

The best explanation of the 'meta description' tag is as a short summary of the page – a clear and concise 155 character description of the on page content (with a minimum of 90 characters recommended). However, as this is the text visible in SERPs, it should also serve to draw consumers to the page – featuring, wherever possible, some form of enticement or call to action. A good 'meta description' can drastically improve your click through rate.

<meta name="description" content="This is a concise web-page description">

Though your brand may publish a novelty blog titled '10 uses for unwanted Christmas knitwear,' to keep the bounce rate low and your potential readers are given a full and attractive account of onpage copy (deterring the searchers that may be looking for recycling or upcycling advice), you may wish to describe it as: "A Humorous List Of Unexpected Uses For An Unpopular Christmas Gift."

Header tags

Header tags summarise the pages you create for both consumers and search engines – and should be easily read by both.

<H1></H1>

Though wit and humour are desirable for a page title, it is more important in the digital age for it to convey the information for which you wish to rank. While 'Book lack in Ongar' is a great headline, Google's search algorithm is looking for 'Funding cuts for Essex libraries result in severe book shortages' for search queries regarding library funding. Though there is no definitive limit, best practice dictates that you should attempt to keep your titles unique, between 60 and 80 characters, and feature keywords near the beginning.

<H2></H2>

This is a more appropriate place to flex your creative muscle and generally serves as a sub-title. The <h2> tags are also good for use as stand-alone subheadings when seeking to break up content for ease and speed of understanding. Again keywords should feature somewhere in the subheading – as each header is important to framing the relevance of page to search queries.

<H3></H3>

This is a tertiary emphasis tag – or your 'sub-subheading', formatting of which should adhere to its role as a paragraph leader, or stand-out line. As above, ensure this header is relevant to the text it precedes or is linked to – using keywords from the text to add relevance to the passage to promote ease of reader and search engine understanding.

Image source tag

Your 'image source tag' is responsible for providing the source for any inserted image, ensuring that all consumers see the same image.

Alt attribute

The alt attribute is vital when placing images anywhere on your site. Not only is it the text which will be voiced by screen readers, it is also how Google attributes subject relevance to any image. It is best to approach it with the screen reader in mind, however, as this will ensure the best result for both purposes – as Wikipedia states:

Keep in mind the purpose and context of an image and what would be useful to someone who cannot see it.



Text link

Your Anchor Text is the format in which a text link is placed (hypertext). The anchor text (the word or phrase to which a link is attached or 'anchored') and the text around it should provide easily read signals as to the content of the destination page.

If, therefore, someone wanted to place a link to their blog, the code and result would look something like this:

Code:

```
If you would like to know more about this fascinating subject,
you should <a href="https://genericbloggingsite.com/myblog.html">visit my blog</a>
where there are several more posts on or around this topic!
```

Result:

If you would like to know more about this fascinating subject, you should <u>visit my blog</u> where there are several more posts on or around this topic!

Image link

The above is the necessary format if, for any reason, you would like to use an image rather than a phrase to serve as the portal for an external link. While the picture serves as the clickable link, it is still vital to remember that relevance is still determined by the alt attribute.

A clickable picture link would look something like the following (where the
br /> is code for a line break):

Code:

Result:

The following image links directly to the Click Consult home page:



Image link (with nofollow)

The above is the format for a picture link which would be 'nofollow'ed. This type of link is predominately used when linking to sites which you may not completely trust or have control over – such as fora, comments sections etcetera. For example, the following is a 'nofollow' picture link:

Code:

Result:

The following image links directly to the Click Consult home page:



As you can see, there is no visible difference (unless you have a browser plug-in to highlight the nofollow aspect); the 'nofollow' is strictly a back-end addition to a link, used by search engines to determine whether to pass Page Authority from the linking page to the destination page.

Semantic markup and schema.org

When optimising a location specific landing page, it is vital you employ semantic markup, to allow Google to understand the local context of the information shown on the page, thereby allowing your landing page to show for location specific search queries.

Use of <u>Schema.org</u> (as below) allows the page to be understood as specific to a location and therefore enables your landing pages to be tailored to local audiences.

REDIRECTS

When you remove pages from a site, whether temporarily or permanently, put redirects in place from the old page to the most suitable current page so that visitors to the page do not encounter a 404 error from broken links and bookmarks.

These redirects also have the same benefit for site crawlers, directing them through to an appropriate page. While the use of 302 and other redirects add no concrete SEO value, the 301 redirect passes the majority of authority from the old to the current page, and is therefore recommended.

Redirects are generally added from the server, requiring access to the .htaccess file for apache servers and web.config for Windows IIS servers.

200 - OK/Success

Requires little explanation – this is a code read by the browser only; your page should be displayed to the consumer.

301 - Permanent Redirect

In cases where there is duplicate or similar content which are feeding of the same traffic, a 301 will redirect all traffic to the stronger content, new content or a new website/URL.

302 - Temporary Redirect

There are few reasons for a 302 but essentially it asks Google to retain indexes for both pages, it does not pass on authority however, so the authority of the first page may be lost.

410 - Gone (Permanently Removed)

Used when content has been deliberately removed and will not be available again in the future. It indicates to engines to purge this page from indices.

500 - Server Error

For use when there is no more appropriate code available – when the server encounters an unexpected condition that defies a more specific code.

503 - Temporarily Unavailable

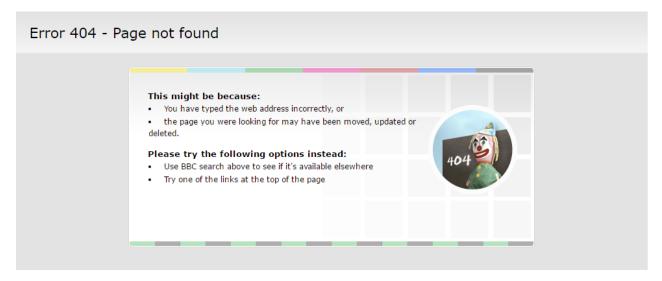
In cases where the server becomes overloaded, or requires maintenance, the 503 code will indicate the temporary nature of the outage.

The 404

The 404 is the most public of the above status codes, being the 'page not found' code. It is, therefore, often a nice touch for brands to personalise their 404 pages – either with links to pertinent content or to the homepage.

This gives the page a familiar branded feel, and can soften the blow of a broken link or page – as with the below offerings from language app 'Duolingo' and the BBC:





In the above example from the BBC the use of navigation options on the 404 also offers visitors an easy path to other areas of the site, improving the chances that the visit will not simply bounce.

The 'soft' 404

This, though showing a site visitor a 404 error page actually, presents site crawlers with a 200 status code, allowing the crawlers to view it as a live page. In other words, though the visitor may see your 404 page, a search engine such as Google will perceive the page as live.

URL best practice

Your URL has always been vital to your brand's online presence and, though domain name keywords are no longer as helpful as they once were, an exact match domain can influence click-through-rate. However, the important thing is that the URL is memorable, easy to spell and, preferably, short. This is the same, wherever possible, for all URL elements.

htttp://store.your_site.com/item_folder/item_page



- 1. Protocol
- 2. Subdomain
- 3. Root Domain
- 4. Top Level Domain
- Subfolder
- 6. **Page**

URLs should be kept simple and descriptive. If your site sells generic television remote-controls, then *http://www.channelchanger.com* is always going to serve you better than *http://www.televisions.com*, for example.

The 'televisions' URL would, because of the vaguely misleading nature of the URL, result in a much higher bounce rate due to irrelevant traffic. Yet this extends to subfolders too – if a subfolder contains blouses, for example, call it "blouses". Also, wherever possible, attempt to keep relevant content within the same subdomain – this will help with Domain Authority

What makes a good URL?

A good URL will always be as simple as possible. Though tempting to subfolder each product for the purposes of accuracy, the ideal target is to have your product as close to your domain name as possible and for it to be it clearly described. Not only does this make it easier for the consumer to remember, but if someone chooses to type out a link (for whatever vaguely implausible reason), it makes it less likely they will misspell the URL and cost you the link and referral traffic.

http://www.northwestdairies.com/half-pint



http://www.northwestdairies.com/p2&?=1234MZ32/pasteurised/semi-skimmed/organic/half-pint



In addition to impact on user experience, the use of variables (the section featuring the question mark) makes the URL more difficult to parse by search engines and consequently for returning in SERPs.

CANONICALISATION

Canonicalisation is the name for the process of redirecting search engines from one or more of your brand's URLs to a single, accepted (or canonical) URL. These multiple pages are common on, but can potentially cause problems with ranking – as the authority conferred upon content may be divided between multiple addresses, weakening the authority of each version and, therefore, decreasing the overall visibility of the page.

Common URL duplications

http://www.your_site.com http://your_site.com

http://www.your_site.co.uk http://www.your_site.com/index.html

Best practice

When optimising your site, you should always use the canonical tag (with few exceptions - such as for brands whose pages are repeated in multiple languages), ensuring all pages achieve the maximum Page Authority possible. All duplicate versions of a page should contain the same URL within the canonical tag – which should always be complete and identical, including the HTTP(S)://, or it may be ignored.

You can use link href=http://www.your_site.com/ ref="canonical" /> as a stand-alone tag placed between the <head></head> tags. This should always refer to your preferred version of the destination URL. This serves the double purpose of allowing you to control which of the URL's a search engine sees and ensures that your traffic is not split between duplicate content, improving the reliability of your reporting.

Important!

It is absolutely vital that no mistakes are made with the URL within the canonical tag. The smallest of mistakes can prevent the page ranking entirely.

301 and canonicalisation

Another useful tool is the 301, or 'permanent redirect' – if one considers the canonical tag as a 'suggestion' to site crawlers to accept your stated preference, the 301 redirect is a non-negotiable command to the same end.

SITE STRUCTURE

With your site and pages optimised, you can begin to map it for optimal user experience (UX) and crawler accessibility. Navigability is an extremely important part of both the customer journey and crawler indexing, but it is also vital to domain authority and, as a result, to conversion.

A sitemap is, therefore, integral to a well optimised site. As the sitemap impacts the flow of authority, the user experience and the performance of the site in SERPs, it is essential to ensure the map is logical, hierarchical, and flows from one end to the other. You should also make provision for regular maintenance as and when new content is added – as this ensures that nothing is masked by redirects, or in any other way hidden from site crawling and indexation.

You cannot accompany every consumer on their journey through your brand's site, but you can make sure it is easily navigated and correctly signposted – and the same goes for the journeys of site crawlers and robot.txt. If you want your pages to be correctly indexed and your Domain Authority to extend the full length of your site, then you need to ensure your site is correctly ordered.

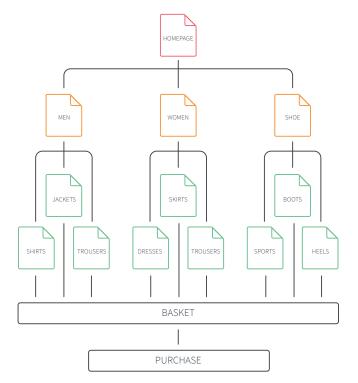
In the simplest possible terms, were a website a body, the content (blogs, products etc.) would be the flesh and skin, while the sitemap would be the skeleton – the frame around which everything else is built. For you to make progress, therefore, it is necessary for your foot bone to be connected to the ankle bone, ankle bone to leg bone and on and on.

If we think about a simple customer journey (eCommerce here, because eCommerce goals are easier to define):



As we can see, the customer arrives, identifies a product group, narrows it down, adds their product to the basket and converts. Simple.

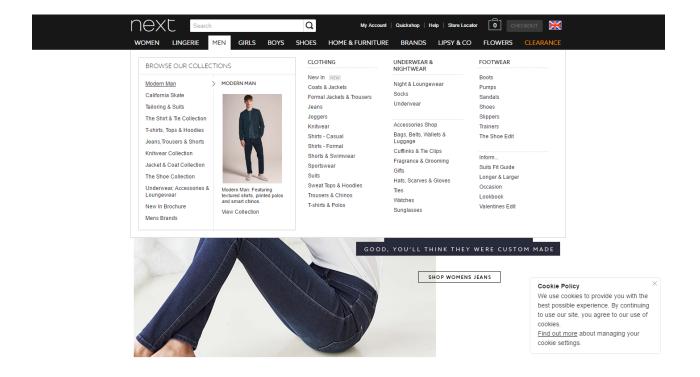
This simple pathway, however, has to be replicated for every journey on your site. Any consumer, site crawler or anything else with the intention to do so must be able to make a similarly simple journey through each layer of your site. It is for this reason that XML sitemaps are of paramount important.



To the left is a simplified visual approach:

Yet, even with this most basic approach
– with the homepage linking to minimal subcategories, which link to very few in turn
– we can begin to see how a journey can be impacted by limited in-site navigation. If a consumer enters the site above and wants to look at shirts following trousers, they would have to return to the 'Men' subcategory and retrace their steps. It's a small inconvenience, but an inconvenience nevertheless, and all inconveniences impact conversion to greater or lesser degrees.

It is possible to resolve this, as most sites do, with an ever-present, branded navigation bar featuring drop-downs to make virtually all pages reachable from anywhere on the site ensuring the user journey is simplified.

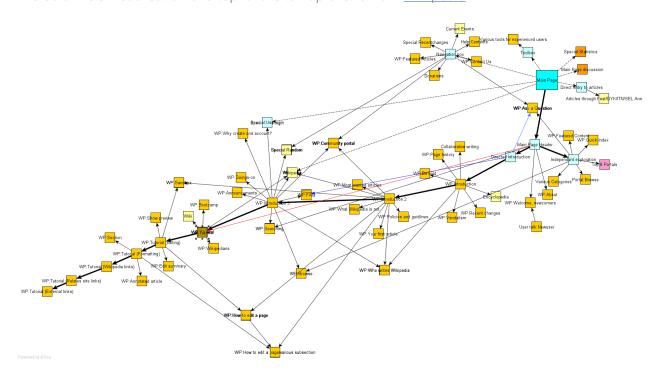


This is where the term 'link juice' comes into planning – Domain Authority has the equivalent of a trickledown effect, whereby homepage authority (as well as subpage) is passed to pages lower in the sitemap, but is occasionally blocked by broken links or poorly implemented redirects, meaning that the flow of link juice is halted. The easiest way to think about it is as a wine glass pyramid:



The link juice, as the metaphor always goes, is represented by the wine – it goes in at the top and flows down through the layers to the bottom. A broken link at any level is the equivalent of placing a board between layers, stopping the wine from reaching lower levels, wasting the link juice (diminishing the Page Authority) which can stop pages lower in the hierarchy from indexing entirely.

The below is a visualisation of a top level sitemap of and from Wikipedia.



As we can see, even this top level diagram is much more complicated than our original map, though your actual sitemap will, in reality, look more like the below:

```
<?xml version="1.0" encoding="UTF-8"?>
<urlset xmlns="http://www.sitemaps.org/schemas/sitemap/0.9" xmlns:xhtml="http://www.w3.org/1999/xhtml">
    <ur1>
        <loc>http://www.example.com/en/</loc>
             <xhtml:link</pre>
                rel="alternate"
                hreflang="x-default"
                href="http://www.example.com/"
             <xhtml:link</pre>
                rel="alternate"
                hreflang="en"
                href="http://www.example.com/en/"
            <xhtml:link</pre>
                rel="alternate"
                hreflang="fr
                href="http://www.example.com/fr/"
            <xhtml:link</pre>
                rel="alternate"
                hreflang="es"
                href="http://www.example.com/es/"
             <xhtml:link</pre>
                rel="alternate"
                hreflang="es-mx"
                href="http://www.example.com/mx/"
    </url>
    <url>
        <loc>http://www.example.com/mx/</loc>
            <xhtml:link</pre>
                rel="alternate"
                hreflang="x-default"
                href="http://www.example.com/"
            <xhtml:link
                rel="alternate"
                hreflang="en'
                href="http://www.example.com/en/"
```

Early efforts should be focused on creating logical flowcharts for primary consumer activities – if only to simplify the hierarchical structure of the site and make planning the user journey and preferred journey of site crawlers easier.

The next step is to create the sitemap for uploading to Google's Search Console (a guide to uploading can be found here). This can be done manually (as in the example above) line by line, or can be completed using an auto-generator such as xml-sitemaps.com.

The main things to remember are:

- That it follows your logical, hierarchical structure
- That all internal links work
- That all redirects in place are necessary and serving the purpose for which they're intended.

CONTENT OPTIMISATION

As stated earlier, the expectations of both search algorithms and consumers pertaining to site content has without doubt been a driving force behind the improving standard of branded content hastened by the changing nature of ranking signals used to decide SERP's positions.

Previous to the Penguin and Panda algorithm updates, black-hat operators within the SEO profession were much more able to game search engine algorithms by 'over-optimising' content (by producing content which was written for the algorithm rather than an audience). This is content many readers will recognise from the early days of the internet. Copy could often be keyword stuffed, virtually illegible and entirely unilluminating – written simply to increase traffic or rankings rather than to answer a query. So that a section of copy on a car insurance website (prior to the aforementioned algorithmic updates) may have read something like:

If you need car insurance, then the car insurance we have for your car insurance needs is the best car insurance for the lowest car insurance prices that you can purchase car insurance for when looking for car insurance.

Though it may seem odd to the modern search engine user, this keyword stuffed copy would almost certainly have been enough to fool early search algorithms into believing that this site should rank well for car insurance queries – with its high percentage of keywords far more important than its potential ability to offer a quality service.

Now, however, after successive updates and refinements of its core algorithm, Google will rightly penalise unintelligible prose – choosing to view the tactic as indicative of a potentially reduced-quality consumer experience.

How content should be optimised

The focus for content optimisation is slowly, but with increasing pace, moving away from specific keywords and long/short-tail phrases to more natural language as Google's algorithm develops an increasingly complex vocabulary (a facility greatly increased by the incorporation of RankBrain). It is, therefore, becoming increasingly counterproductive to think in terms of keyword density when looking producing content for your brand's website.

Though keywords are still a factor, it is now possible for Google to identify synonyms, so that it is possible for content producers to focus on work which – to paraphrase Google's Matt Cutts – doesn't just appear natural, but is natural, reducing the need for copywriters to focus on keywords and phrases and to write as well as they are able.

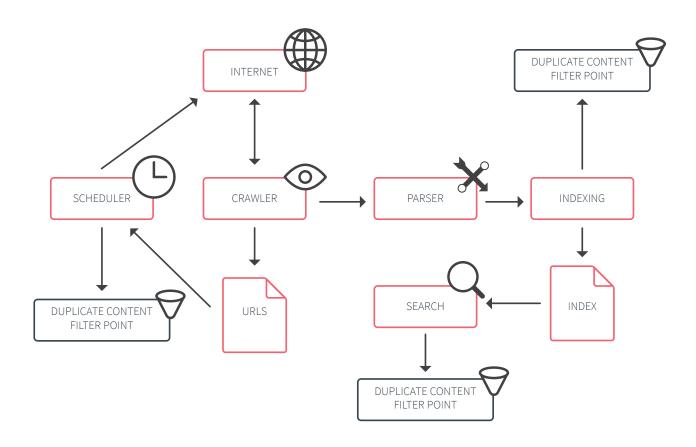
The most important aspect of brand content at this stage is ensuring that the html/semantic mark-up is performed correctly and in accordance with best practice. All images must have appropriate alt attributes, all text should be broken up with the correct header tags, while all links – whether internal or external, should be checked, tested and working to direct users through to the intended pages.

THE 'DUPLICATE CONTENT' MYTH

Discussion of content from a technical viewpoint invariably leads to questions regarding syndicalisation, product descriptions and repurposing; it is, therefore, as good a time as any to tackle the duplicate content myth.

As per a talk by Google webmaster John Mueller, the main concern for Google regarding duplication is that it wastes resources, crawl budget and time, which delays the pick-up of new content, making metrics more difficult to track.

Its major concern, for example, is not the first appearance of content (in response to a question during his hangout, Mueller stated that, if material is syndicated, Google will endeavour to return the best performing content according to an interpretation of intent/locality, not the content's original location). Google simply wishes to provide the best user experience, and this means returning a wide variety of search results – something with which a significant density of duplicate content can interfere.

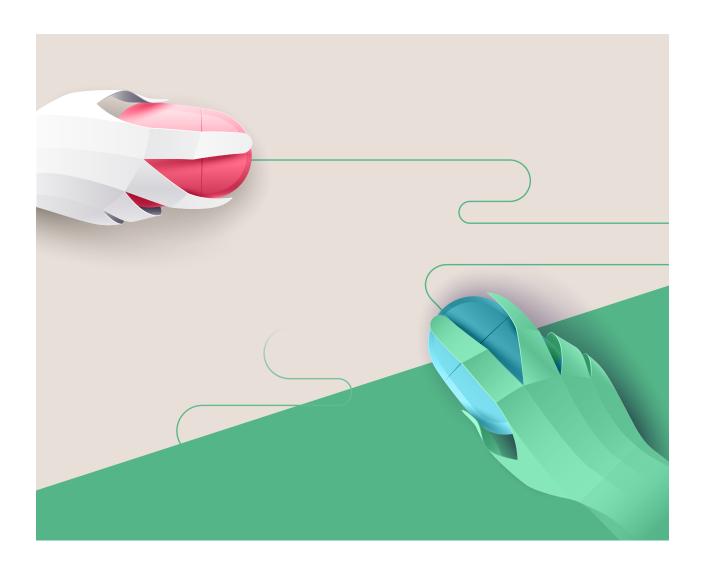


Filtration of duplicate content occurs at three stages

Scheduling: Google's bots cannot crawl the entire web at all times; some duplicate content is, therefore, detected during the scheduling process – which is the point at which the batches of URLs to be crawled are decided, increasing the efficiency of each crawl.

Indexing: Duplicate indexed items wastes storage space. Generally speaking, therefore, Google will only index one version of the content – unless, of course, it meets certain criteria such as those mentioned earlier, including for localisation, in which case both versions are indexed.

Search: Duplicated search results, Mueller stated, are potentially confusing and reducing the number of unique results on a page. This is why, even if the duplications are legitimate, you may see the phrase 'we have omitted some entries'.



Do you have an internal duplicate problem?

According to Mueller, clear indications that you have issues with duplicate content are if non-preferred URL's are appearing in searches, Google's Search Console is showing multiple instances of title and description duplication, or if site crawlers are returning a greater number of page-crawls than you have pages. In addition to this, there are a number of free web resources which can trawl your site for duplicate content.

How to fix them

Do **not** use robot.txt to resolve duplication issues – this simply disguises the duplication, hiding it from the site crawlers without resolving the problem, as both pages will be indexed and the resultant problems stemming from duplication will persist.

Google understands that some duplication is unavoidable (such as with press releases, or third party product descriptions), the key is to be consistent with your signals to Google and other search engines – ensure that you are using the desired URL everywhere (sitemap, canonical, href etc.), "rel canonical" is a good method of overcoming problems caused by multiple product pages (for size/colour etc.). Avoid URL variations in CMS, make appropriate use of 301 redirects wherever possible, and ensure your site is easy to read with a well-structured hierarchy. We would recommend extensive auditing for duplicate content both internally and externally to your brand's site.

How to avoid them

Be SURE with your content. Make it substantial, unique, relevant and engaging. If you repurpose content, give it a new spin or narrow the focus, don't artificially rewrite (removing spaces, change spellings etc.), and adhere wherever possible to webmaster best practice.

Expertise

These sections, perhaps, will go some way to explaining why SEO companies have become increasingly important as Google SERP's continue to gain significance in brand revenue generation. Though it is certainly not impossible for an individual to develop expertise in all of the key areas, or indeed to make a passable attempt and to gain some traction with their SEO endeavours, SEO companies tend to have spent the duration of their existence building a talent pool of experts in their particular field.

GET IN TOUCH WITH US TODAY

ABOUT US

Click Consult is a multi award-winning digital and search marketing agency with a focus on organic (SEO) and paid search (PPC). Part of global consumer brand business Ceuta Group, we have a team of 70-plus specialists and a portfolio of more than 60 clients worldwide.

Our complementary services include content marketing, outreach, social media, conversion rate optimisation (CRO) and international/multilingual search marketing. We can also offer training and consultation to support your teams or existing strategy.

Click was named SEO Agency of the Year 2019, adding to our long list of other awards and accolades, and also ranks within Econsultancy's 'Top 100 Digital Agencies', and Prolific North's 'Top 50 Digital Agencies'. We're also a Google Premier Partner, a Bing Select Partner and feature in The Drum Recommends.

YOUR BESPOKE ORGANIC SEARCH (SEO) ANALYSIS AND REVIEW

REQUEST NOW



0845 205 0292

OUR ACCREDITATIONS









YOUR BESPOKE ORGANIC SEARCH (SEO) ANALYSIS AND REVIEW

OUR AWARDS









OUR TECHNOLOGIES

monitor <u>TRAX</u>* rank <u>TRAX</u>* feed <u>TRAX</u>* link <u>TRAX</u>* pro <u>TRAX</u>* page <u>TRAX</u>