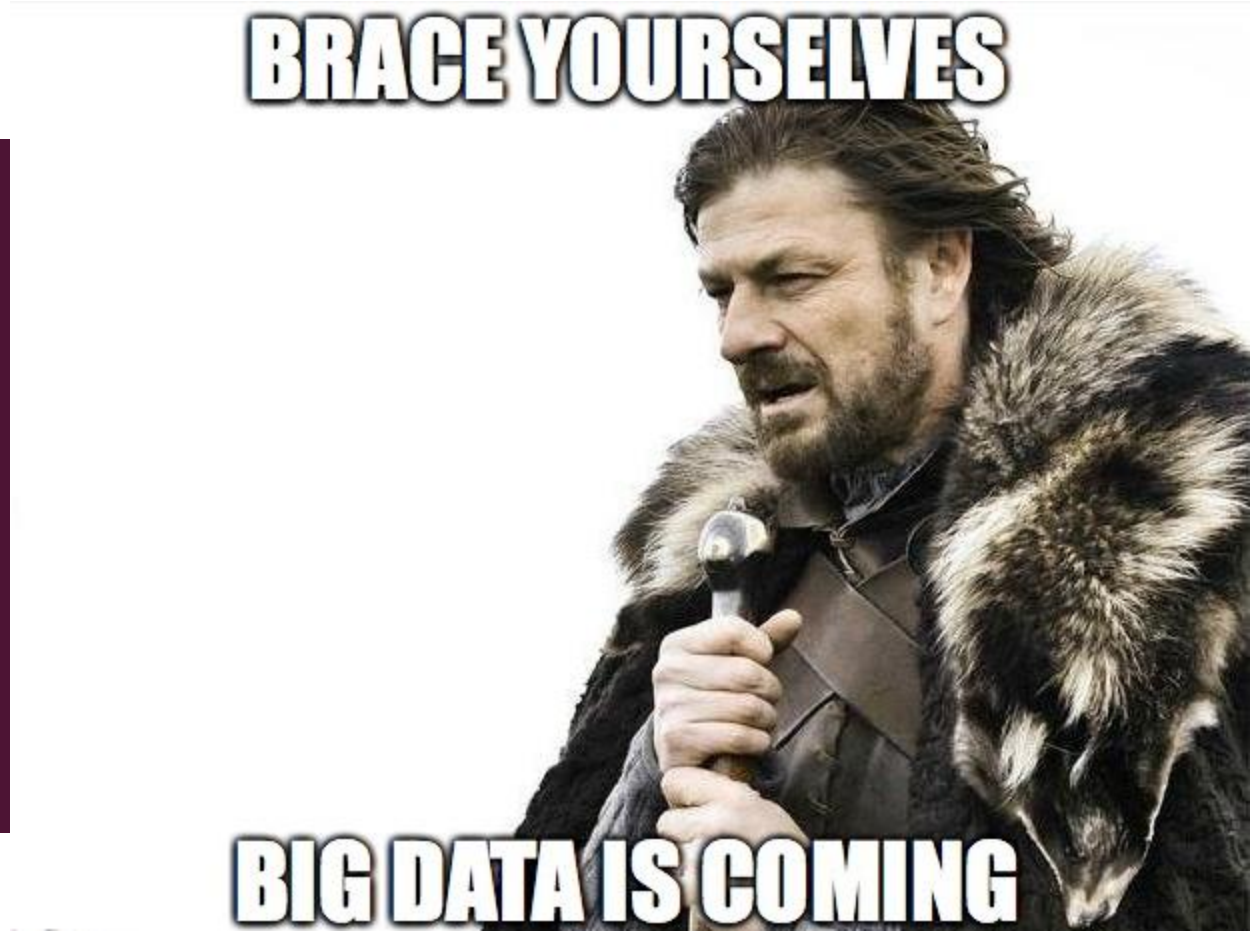
BIG DATA AND SEO



WHAT IS BIG DATA?

- To understand the idea of Big Data and how it relates to SEO, perhaps it's best to begin by looking a bit at the history of the web and human interaction with it. Currently, three periods of the web have been identified by computer historians:
- **Web 1.0** – the first period of the web, which lasted from 1991 to 2004. This period was saw most users of the web as content consumers as opposed to creators. Sites weren't interactive and were static. Webmasters did not use programming languages to create pages but rather Server Side Includes and Common Gateway Interface.
- **Web 2.0** – the second period of the web. There are debates about regarding when Web 2.0 actually ended. Some historians suggested it ended in the early 2010s whereas others suggest that we are only seeing the end now. This period is marked by the introduction of multiple self-publishing avenues making it easier for everyone to become a content creator on the web. For example, social networking sites, comment sections on news articles, blogs, etc. No longer are creators limited to those who know programming. Dynamic websites are the norm.

WHAT IS BIG DATA?

- **Web 3.0** – the third period of the web, and the stage we are currently in at the moment. This stage is marked by the fact that the internet has become a data reading machine. The internet is now using the data you create as a way to profile you. A good example of this can be seen in how your web searches impact the type of adverts that Google chooses to display to you if you visit a site using Google Ads. This stage is also called the Semantic Web period as search engines are beginning to provide semantic meaning to your data and build relationship with ideas.

SO, HOW DOES THIS RELATE TO BIG DATA?

- More traditional definitions for the term big data referred to methods which allowed people to extract information or data sets that were larger than what traditional data processing software was able to deal with.
- However, currently, most use the term to user behavior analytics or predictive analytics. Basically, the information that can be inferred from the information gathered. (Think back to the Google Ads example.)

HOW DOES GOOGLE USE BIG DATA?

- Google uses several tools to gather predicative analytics in order to better serve its users. Some of examples how they gather big data to improve their services include:
  - **Knowledge panels**; these are built for high volume search terms in mind. The panels try to include all the information that most users generally want to know about a topic.
  - **Semantic relations**; they try to get the algorithm to make relations between the semantic meaning of words. For example, if you type in "cell phone", it will still bring up pages with the headline "mobile phone".
  - **Real-time data feeds**; think about when you are searching for the score of a sports game. If it is live in-play, Google will give you live updates of the minutes past, last play, current score, etc.
  - **Tracking cookies**; Google uses tracking cookies to make relations between two perhaps unrelated concepts. For example, people who like "A" generally tend to like "B" as well.

HOW DOES THIS AFFECT SEO?

- As previously mentioned, Google's algorithm tries to build profiles for its users and thus tries to determine correlations between certain search terms and the information you actually want. This is where Natural Language Processing becomes particularly important.
- If the search algorithm relates a search term to a whole load of concepts, you need to have those concepts on your page to help Google understand what your page is about. The algorithm uses big data to identify what sort of information people who search for a particular topic would also be interested in. (Remember the idea of keyword salience from the NLP presentation.)

HOW DOES THIS AFFECT SEO?

- The result of this is that it has become harder to rank on Google. So, previously, it was fine just to put your keywords on a page to get it to rank but now the algorithm requires you to do a whole lot more. Unfortunately, you don't always know exactly what is required and this is where research becomes more important.

SO, HOW DO I OBTAIN BIG DATA ABOUT MY TARGET AUDIENCE?

1. Keyword research – finding keywords that matter (Keyword Planner and SEMrush)
2. Improving your content with information gathered from your site (Google Search Console and Google Analytics)
3. Social media is more important than you think
4. Structured data is here to stay

1. KEYWORD RESEARCH – FINDING KEYWORDS THAT MATTER

- When you plan content, you need to know how people are looking for the content you are trying to create. Often this means not focusing on a handful of keywords but larger subsets to ensure that you are able to cover as much as possible.
- However, we should only focus on keywords that are actually relevant to the page we want to write. So, even if a keyword has a high search volume, if it isn't relevant to your page, don't include it. You need to keep user intent in mind. What does the user want to find when searching for this search term? Will my content provide that information?

1. KEYWORD RESEARCH – FINDING KEYWORDS THAT MATTER

- You can find keywords people are using to search for the content you would like to create through using tools such as SEMrush and Google Keyword Planner. Keyword Planner (which is accessible through Google Ad Words) is particularly important. This is Google telling you how users are searching for concepts you may want to write about and the terms that are related to it.

2. IMPROVING YOUR CONTENT WITH INFORMATION GATHERED FROM YOUR SITE (GOOGLE SEARCH CONSOLE AND ANALYTICS)

- If your page/site is getting traffic, you can optimize it by having a look at how exactly people are finding it. Sometimes, you'll find that this information conflicts with the keywords you obtained.
- Other times, it'll suggest that you add onto already existing content. For example, if you see people accessing your content through a search query/question that was not included on your page, you can add it.
- It's all about using the data that your site gives you to better answer the questions your audience is asking.

3. SOCIAL MEDIA IS MORE IMPORTANT THAN YOU THINK

- Most people aren't aware of how much social media impacts rankings. For many, it is the primary method used for interacting with other people on the web. It is also an area where people make others (including search engines and other web technologies) of your likes, dislikes, what you would like to know, etc.
- It has been suggested that how much attention is paid to a particular idea/concept on social media directly impact search engine rankings as well. The actual reason for how it impacts rankings has been debated but many agree that it has got a lot to with engagement (CTR, bounce rate, etc.) as opposed to how many Facebook likes or retweets it gets.

4. STRUCTURED DATA

- We have already covered this before, so I won't go into too much detail about structured data. Basically, structured data provides Google with information about the page, which allows it to understand what it is about and how it relates to what content already exists. Why wouldn't you want to tell Google directly what your content is about?

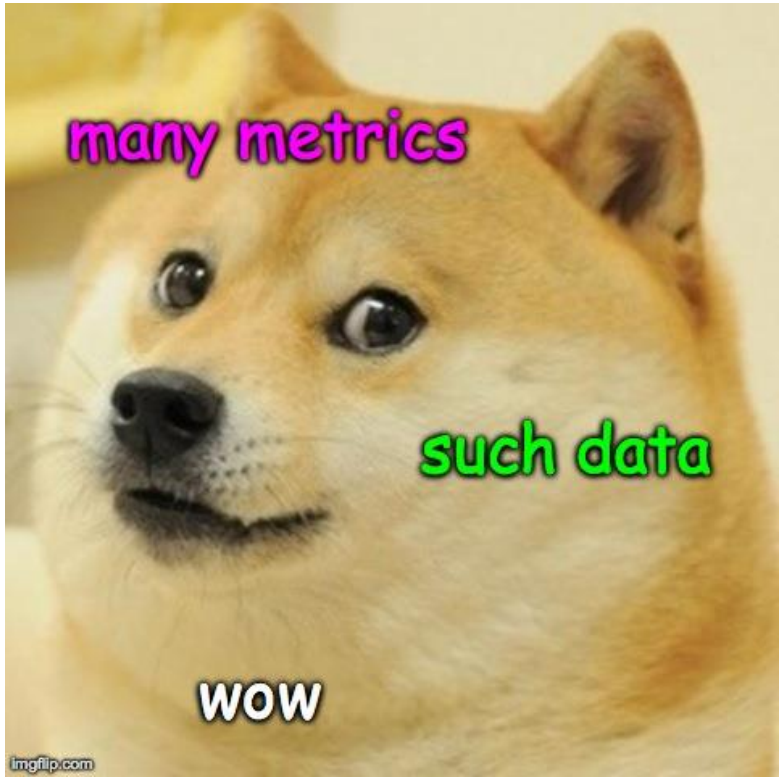
3. SOCIAL MEDIA IS MORE IMPORTANT THAN YOU THINK

- This is true of even YouTube – one of Google's own social media platforms. A good example of this was that PewDiePie once asked viewers to purposely dislike a video and it still appeared on YouTube's trending list. The reason for this was watch time. In fact, if you were to look at the full list of the most disliked videos on YouTube, you would notice that they were all trending at one time – and were at one stage being actively recommended to viewers by YouTube.

NOTES ABOUT BIG DATA

- One very important thing to note about big data is that the profiles created by Google aren't always astutely accurate.
- Thus, you could have all the relevant keywords and concepts you think to be associated with an idea on a page and it could still not rank. This is sometimes due to the fact that Google has created an erroneous profile for users searching a particular term, which then makes ranking for that term that much harder. This is because you do not know what to include on the page.
- However, this does not mean that we shouldn't have a look at big data – it could still prove very beneficial to your content.

QUESTIONS



SOURCES

- https://www.inc.com/aj-agrawal/4-ways-big-data-is-changing-seo.html
- https://reflectivedata.com/how-is-big-data-impacting-search-engine-optimization/
- https://www.smartdatacollective.com/use-big-data-seo/
- https://www.newgenapps.com/blog/5-reasons-why-your-seo-strategy-needs-big-data
- https://www.tmcnet.com/topics/articles/2019/03/13/441564-how-big-data-impacting-seo.htm
- https://www.vertical-leap.uk/blog/how-to-apply-big-data-to-seo/
- https://tweakyourbiz.com/technology/google/big-data
- https://www.smartdatacollective.com/how-web-3-0-going-change-data-access-as-we-know-it/
- https://www.investopedia.com/terms/b/big-data.asp
- https://moz.com/blog/relationship-between-organic-rankings-and-social-shares