

# Plans for a new keyword research system

## Keyword Research System brief

*A key need for the US SEO team is a tool to automate our keyword research and categorization process. Currently, we gather data from 3rd party tools (SEMRush), analyze and categorize in Google Sheets, and then upload into our tracking tool (GetSTAT) manually.*

Example report downloads:

[SEMrush organic research report](#) (this is a very popular SEO tool used to get snapshots of organic ranking data from Google)

[STAT keyword tracker upload template](#) (this is our keyword rank tracking tool that we use to automatically track keyword rankings daily to report for clients)

You will already have a [login for STAT](#) with the password treatfork456 so you can get an idea of how keywords are currently tracked for our reporting needs.

---

It's important to note that there are essentially two elements to keyword research and tracking that we currently do manually:

1. Keyword research for pitching new clients which is a question finding as many **relevant** keywords as possible **quickly** (as automatically as possible) to show the biggest **opportunity** - "by ranking higher for these xxx keywords you could make \$x,xxx,xxx"
2. Tracking for campaign performance monitoring. We need to track the most important terms, ensure they are categorized expertly, and be able to report on [Share of Voice](#) for important keyword groups.

---

## The Research and Categorization

- We need the ability to ingest relevant keywords for a digital brand in any given space, e.g. eco kitchenware, business insurance, crypto trading, etc. We can get keywords from

anywhere really but SEMrush is likely to be best (further reading - [How Accurate are Website Traffic Estimators?](#))

- We also need to understand more than just keyword data that is from Google (via their AdWords keyword planner) - Google don't provide access to a lot of this data for SEOs so we would use tools like <https://answerthepublic.com/> to get Google Autocomplete phrases that show immediately popular search terms around a given topic as well as <https://alsoasked.com/> to get the "People Also Ask" questions that you often see in Google SERPs for a given keyword
- We may need to comprehend data in multiple international markets and across both mobile and desktop devices
- We need to be able to perform automated categorization based on various information for each term, including funnel stage, consumer affluence, [search intent](#), search volume trend data, etc.
- We need to automate additional categorization based on industry and site types: product type, color, material, mens / womens / kids, sub-location (e.g. city, state). This needs more than simple fuzzy string matching (e.g. forex, forex trading, fx, fx trading) but **keyword intent / semantic matching** (forex trading / currency trading) which could be derived from a semantic scoring public database
- The ability to use machine learning to extract notable insights from keywords using tools such as [Google's NLP](#)

## The Storage

- Keyword research performed should be stored in a MongoDB database with additional time-based dimensions; month, year, etc. as well as the metrics derived from keyword research API
- Ideally we will then enrich this database on a selective basis (e.g. monthly - cadence might need to be weekly in some instances) via APIs for tools we're using
- Considerations need to be taken to limit unnecessary scaling and bloating of the data stored (thresholds could be set for minimum monthly search volumes or limited number of keywords per domain)
- The database will need to be explorable by account managers via an easy to use query interface, e.g. [MongoDB Compass](#) - the ability to create custom reports / charts for export and future analysis / pitches is key

## What Needs to be Done?

- Please hypothesise a systems that takes into account all of the above - you don't need to physically build anything - simply detail
  - How do we get clean and structured data
  - How do we export it and what is it likely to cost us
  - How do we store it
  - How do we query it
  - Any hurdles to consider
- The process needs to be as automated as possible so scripts will likely need to be written to enable fast keyword research for the purposes of a pitch / new business proposal
  - Subsequent categorization and updates will need to be automated via APIs into the MongoDB database
- Please complete this work in any format you so wish - [A Google Doc](#) will suffice

## Extra Notes

- All of the above needs to be as automated as possible for the account manager
  - Ideally, a solution for retroactive analysis - if new keywords are added via APIs or manually, we'd like their past trend data to be accessible
  - Any questions please let us know
-