# Proposal for a New Keyword Research System

## By Kathia Teran

Based on my research, I found different tools and sites that could help speed-up the collection, categorization, storage, and presentation process of keyword research at Croud.

## I.  Data Collection

### Supermetrics

There are various tools that help find the best keyword data out there, and SEMRush is perhaps the most popular. Nevertheless, there are other sites that help with additional information that SEMRush doesn't necessarily provide, like Answer the Public or Also Asked.

To speed-up the collection process, it would be helpful to have all of these sources readily available in Google Sheets. Supermetrics does just that. It has its own Add-on in Google Sheets and has API access to over 30 data sources on their basic plan, including SEMRush, Google Analytics, and JSON capability to use other site APIs that are not present in the list. Additionally, it also offers weekly or monthly scheduled updates on the data collected.
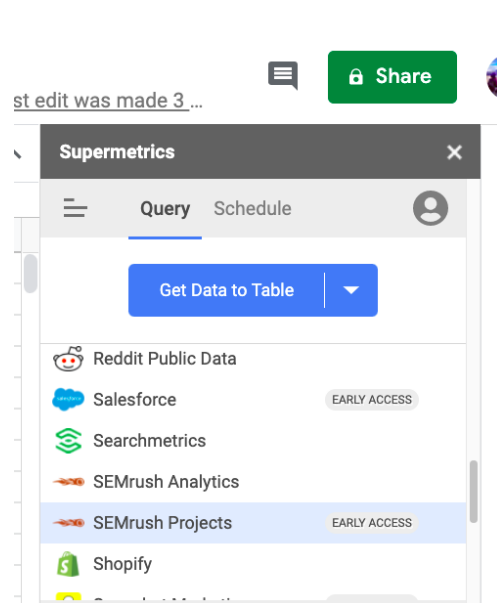


**Data Sources for Pro:**
Bing Webmaster Tools, Custom JSON/CSV/XML, Database, Facebook Ads, Facebook Insights, Facebook Public Data, Google Ads, Google AdSense, Google Analytics, Google BigQuery, Google My Business, Google Search Console, Instagram Insights, LinkedIn Pages, Mailchimp, Microsoft Advertising, Moz, Optimizely, Pinterest Public Data, Quora Ads, Reddit Public Data, SEMrush Analytics, Shopify, Snowflake,Stripe. Tumblr Public Data, Twitter Ads, Twitter Public Data, Vimeo Public Data, VKontakte Public Data, Yandex.Direct, Yandex.Metrica, YouTube

**Prices:**
Pro: one user $99/month
Super Pro: one user $189/month
Enterprise: multiple users, must ask for a quote

https://supermetrics.com/pricing/supermetrics-for-google-sheets

### MonkeyLearn

Gathering data for sentiment and semantic analysis is not necessarily direct and easy. In order to get better insights, the extraction and classification of keywords sometimes needs a more customized search. MonkeyLearn offers the ability to easily customize the extraction and classification of texts, as well as their own trained classifiers. In particular, they have an effective sentiment analysis tool. Additionally, they have their own API for Google Sheets:

Analyzed text:







**Pricing:**
API: $299/month

https://monkeylearn.com/

## II. Automated Keyword Categorization

There are many tools that provide automated capabilities for data analysis using machine learning. And many of them don't require machine learning knowledge. Nevertheless, there were 3 specific tools I found in my research that provide the best and most effective techniques for automated keyword categorization: BigML, Kraken AutoML by Big Squid, and of course Google's AutoML NLP.

### 1. BigML

Having done extensive research to find a true competitor to this tool, I can objectively say that BigML is likely the easiest and most effective for automated keyword/data categorization. It is meant to be used by people with little knowledge in machine learning, and it has a free API for Google Sheets.

I tested the tool to categorize serveware keywords I extracted from SEMRush as eco-friendly or not:
- First, I trained the data using two popular eco-friendly sites (The Little Market and Ten Thousand Villages), categorizing specific serverware terms as 'eco-friendly'.
- Then extracted the first 200 keywords for 'serveware' from SEMRush's Keyword Magic Tool, and ran the automation on Google Sheets using the BigML API:

As you can                                                                                                see, my model is not perfect, but I know that it had to do with issues with my trained data, and not with the application itself.

Below you can see how properly categorized keyword sheets would look like:

## Training Data > Create Model > Classify

| Keyword | Intent | confidence |
|---|---|---|
| pizza near me | local | 0.95053 |
| dentist chicago | local | 0.84605 |
| buy a laptop | transactional | 0.85871 |
| cheap deal | transactional | 0.86295 |
| what is the meaning of life | informational | 0.86586 |
| why is the sky blue | Informational | 0.71521 |
| how to say hi in italian | informational | 0.49929 |
| how to tie shoelaces | informational | 0.49929 |
| xbox 360 | transactional | 0.94075 |
| 32 gb ram | local | 0.51427 |
| bed for dogs | transactional | 0.59253 |
| hair salon | local | 0.97873 |

## Keyword Research ☆ 📁

File   Edit   View   Insert   Format   Data   Tools   Add-ons

↶  🖨  🖹 | 100% ▾ | £   %   .0   .00   123 ▾ | Calibri

designer denim jacket mens

| Keyword | Category 2 | confidence |
|---|---|---|
| black skinny cargo pants | Trousers and Ch | 0.95674 |
| men's winter puffer jacket | Jackets and Coa | 0.99779 |
| black chino outfit ideas | Trousers and Ch | 0.98385 |
| funnel coat mens | Jackets and Coa | 0.99412 |
| striped ringer tee mens | T Shirts | 0.94867 |
| red fleece gilet | Jackets and Coa | 0.96923 |
| charcoal trousers | Trousers and Ch | 0.99045 |
| extreme destroyed jeans | Jeans | 0.43849 |
| denim jacket male | Jackets and Coa | 0.99779 |

BigML is a powerful tool because it can speed-up the process of categorizing and understanding all sorts of other information, such as: search intent, consumer affluence, funnel stage, product types, etc.

It is a 'Freemium' tool, so there are other services they offer that come with a cost. Nonetheless, as I mentioned, their Google Sheets API is free of cost, and is all that is needed apart from signing-up to their site for free.
**Pricing:** https://bigml.com/pricing#subscriptions

**https://bigml.com/**

### 2.  Kraken AutoML
This powerful tool by Big Quid can offer essentially the same capabilities as BigML, but it needs more customization and it does come with a higher cost:

---

### ☺ Let's start with the basics.

The Kraken no-code **automated machine learning (AutoML)** platform **simplifies** and **automates** data science tasks:

- data prep + cleaning
- algorithm selection
- model training + tuning
- model deployment

### ☺ Kraken AutoML is for everyone.

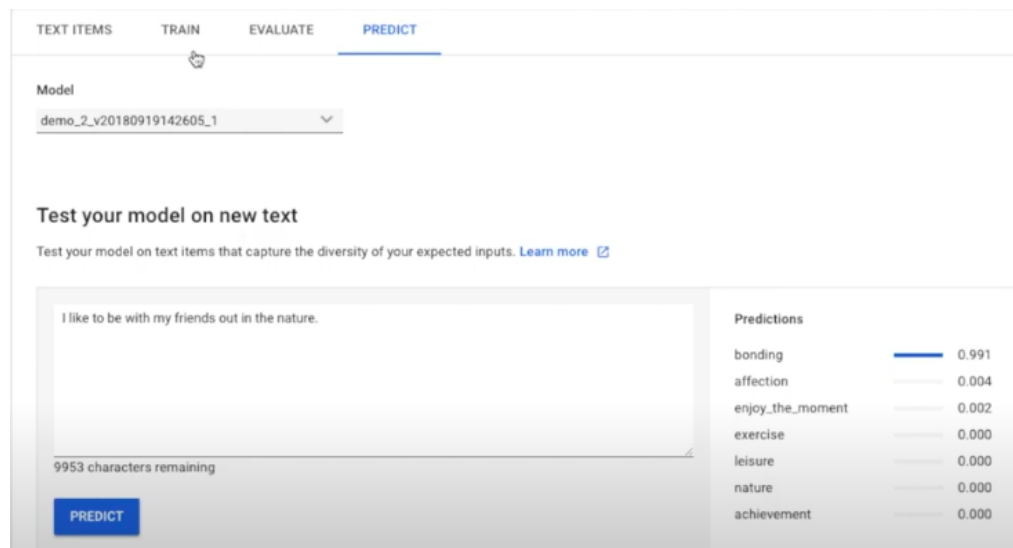Kraken was built with analysts and engineers in mind. If you've done data analysis before, you're ready!

Kraken's no-code, easy-to-use interface and integrated SONAR© training make it easy to become a citizen data scientist. Advanced features allow data scientists to work faster and more efficiently.

**Pricing:** Starter: $1,500/month
https://www.bigsquid.com/kraken-platform

### 3. Google AutoML NLP

This powerful tool is easier to use than it appears. It takes a few more steps than BigML or Kraken AutoML, but it can help categorize and distinguish data very effectively.



**Pricing:** Free, limited to 150 requests and 15 analyses per month
https://cloud.google.com/natural-language/automl/docs

## III. Storage

MongoDB is a great tool for storing large unstructured data. Additionally, it has a direct integration with BigML and connection to Tableau for additional presentation options. But I specifically wondered if there is a tool where I can gather my data from different sources (SQL or NoSQL) and create presentations in the same space. Indeed, there is a tool called *Holistics*.

Holistics has a built-in EL(T) functionality to pull data from MongoDB (NoSQL) to a SQL database. After that, users can easily query and create dashboards/reports easily using SQL.

It also lets you gather data from Google Sheets, MongoDB collections, etc, and then offers great tools to best present this data very easily:
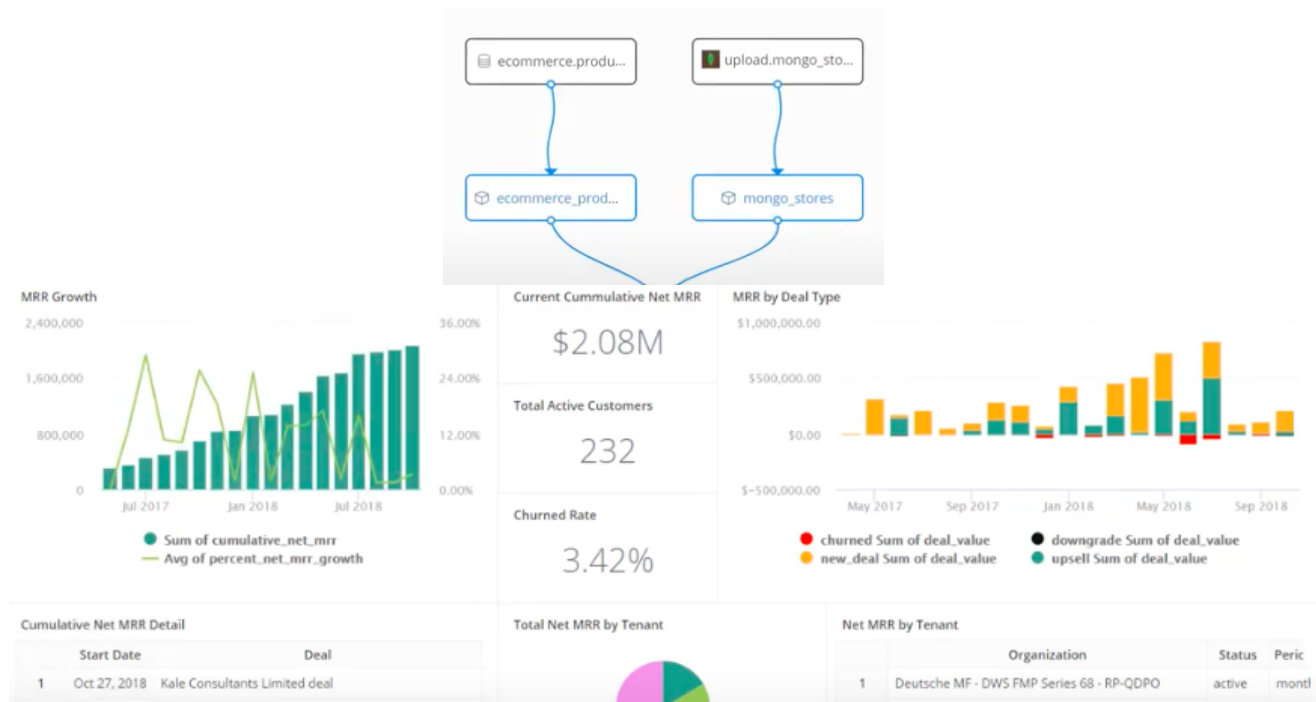


**Pricing:** Entry: $100/month for 10 users
https://www.holistics.io/pricing/


## IV. Special Mentions:

1. **Semantria:** Semantics analysis tool, only available for Windows and Linux:

   https://www.lexalytics.com/semantria

2. **Wordstream:** Free Keyword Research tool for additional international markets:
   https://app.wordstream.com/free-keyword-tool