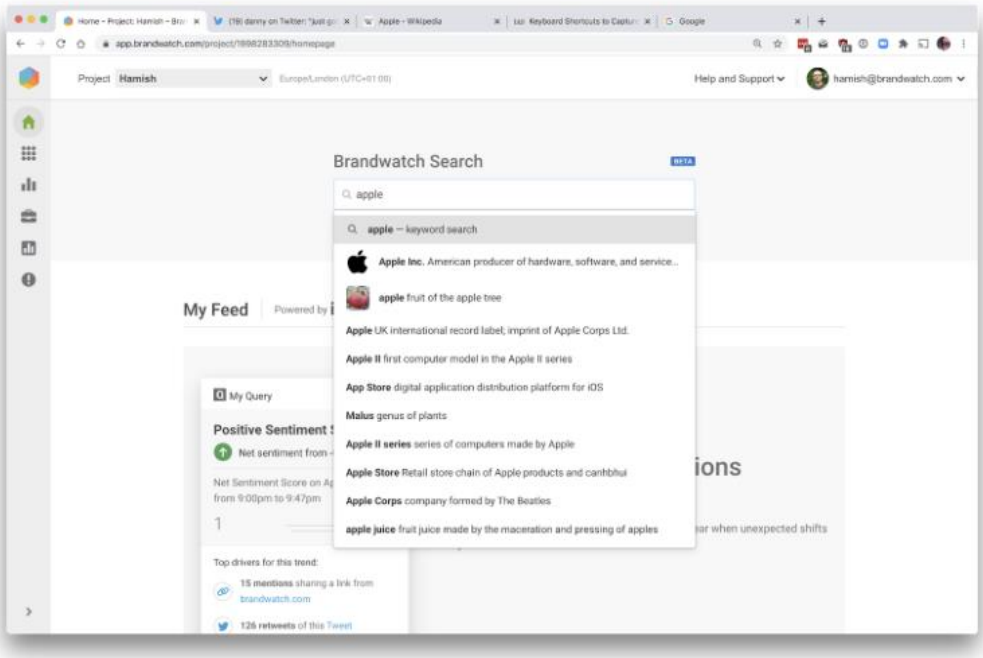
What is AI Entity Search?

Brandwatch offers two ways of searching for data across our data archives: keywords or AI entity search. AI search and AI topics were developed in-house by Brandwatch, trained using our proprietary AI, and are not direct integrations. We do not send any data outside of Brandwatch with these features.

The main difference between the two is that a manual keyword-based search will return mentions that match exact keywords, whereas an AI entity search takes into consideration the definition of the terms that you are searching for.

Say for example that you wanted to search for the brand Apple. A manual keyword search only containing the word "apple" would also return mentions of fruits and any other apple based product names or brand names, so you would need to add additional context to your search in order to get to the relevant content (e.g. Apple AND (iphone OR ipad... etc)). However, using AI-powered entity search, you pick the concept that you are searching for — in this case the brand Apple, not the fruit — and the system will return mentions of the entity that you have selected.



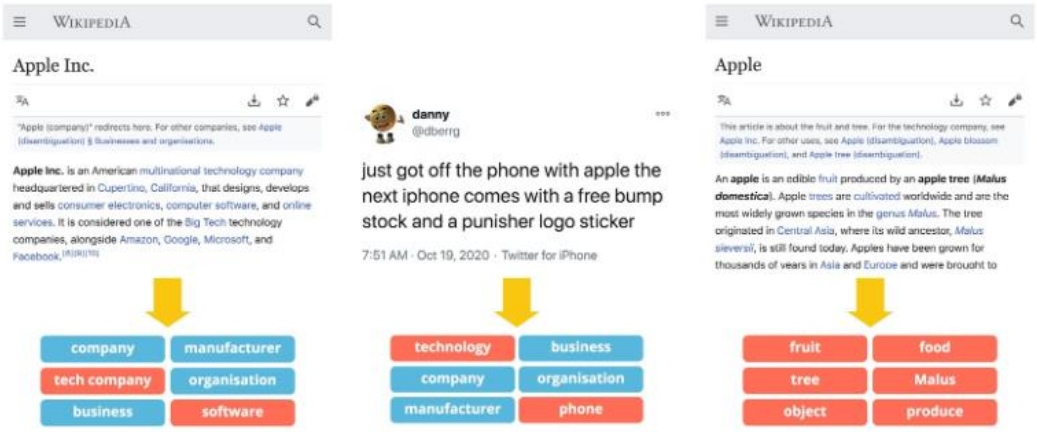
How does entity search work?

Our AI technology works completely differently to keyword matching and retrieves mentions via a three step process:

1. **Noun extraction.** First, we run all mentions in our data library through noun-phrase extraction, which pulls out all of the important noun phrases from a mention. In the example below, we find iphone, apple, logo, and a few others.
2. **Entity candidate generation.** Secondly, we generate a list of possible candidate definitions for each identified entity. For example, Apple the company and apple the fruit are both entity candidates for the word "apple."
3. **Entity disambiguation.** In the final and most comprehensive step, we work out which entity candidate is the one being used in the given mention.

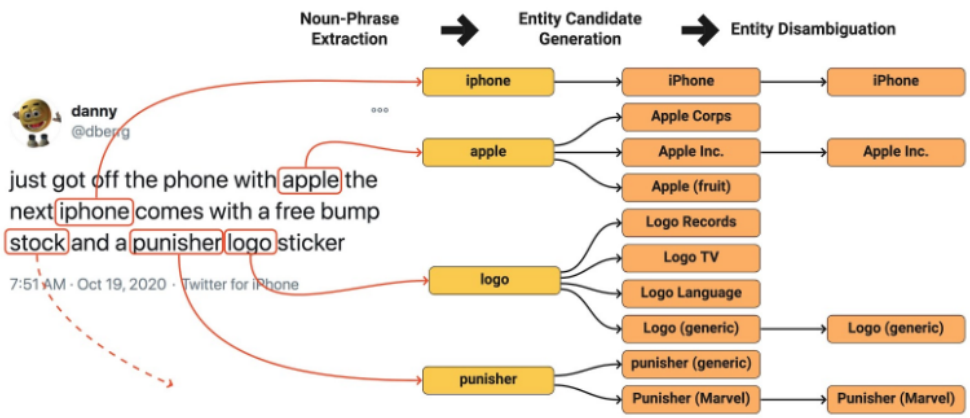
In order to do this, we use a language model, the text surrounding the entity in the mention (e.g. all the words that surround “apple,” in our example), and the Wikipedia descriptions for the entity candidates.

We use the language model to extract meaning from our mention and from the Wikipedia pages belonging to the identified entity candidates, and then we compare these to see which candidate is the best fit with our mention.



We train the language model as if we were going to ask it to generate the next word in a sentence, in a similar way to some more familiar software solutions. However, instead of actually generating the next word, we use the information that the model uses in order to generate that word. We look at the model’s understanding of the text we have given it and ask it for example: “What kind of concept is Apple in this context?” In doing so, the model tells us which entity candidate it would use in order to suggest the next

word in a sentence, which therefore provides us with our best fitting entity candidate — in our case, Apple the company.



Once mentions have been retrieved, you are free to analyze them exactly as you wish.

FAQs

When should I use AI-powered entity search and when should I use manual keywords?

AI-powered search terms are available for all entities that have their own Wikipedia page and enough data to go through the matching process explained above. You can use keywords or AI search interchangeably but for more ambiguous concepts, you will find that an AI-powered entity search provides you with a cleaner set of data in much less time than manual keywords.

Can I combine manual keywords and AI entities?

Yes! You can also combine multiple entities and/or manual keywords within a search using Boolean.

How much historical data is available with AI powered keywords?

When using AI-powered entity search in a query, we can only guarantee historical coverage back to July 2020.

Can I still use Boolean in my query?

You can use boolean with AI-terms. For information on how to use AI powered keywords when writing a query, see the **Creating a Query in the Query Editor** article.

What is included in the data?

The AI technology searches across the Brandwatch data archive. We are always working on improving the technology further to make it even more inclusive. For now, it is recommended to include adjacent entities, hashtags, and author handles separately to get the full view of your data when needed. For example, you may want to search only for the entity Apple or you may want to include other entities such as "iPhone" or "iPad," and specific hashtags or author handles such as "#AppleReleaseDay" to your search string.

Can I see the search string that is behind AI entity search?

AI-powered entity search uses a completely different way of retrieving mentions than the keyword matching you may have been used to. It uses the context of the full mention to decide whether or not it matches the chosen entity, rather than just keywords. Therefore, there is no specific search string associated with AI entity search.

I am seeing a difference in mention volumes between my manual keywords search and my AI entity search. Why is that?

If you are comparing a manual keyword search to an AI-powered entity search using an ambiguous brand or concept, then you would expect to see fewer mentions in the AI-powered search and the relevancy should be higher. If this is not the case, there may be restrictions or inclusions in your keywords search that are segmenting the data in a way that you haven't yet done with your AI-search.

You can use additional keywords and boolean alongside your chosen entities to get to the data you want. You may need to include specific hashtags, links, or author handles to your search that wouldn't be directly linked to the entity. For example, you may want to add "#appleproductlaunch" to your apple AI entity search.

Does entity search work in other languages?

Brandwatch Search works in all languages that have at least the "basic" level of support as specified in the **Language Support** article. However, AI-powered entity search is currently only supported for English language searches. Training our AI to analyze mentions to understand the specific entities they are referring to is a huge challenge. Brandwatch are the first in the industry to achieve this. We've started with English and plan to add other languages to the AI engine in the future.