AlchemyAPI Drives Accurate Concept and Feature Extraction for Pocket

Cleaner data, vast language support sharpens Pocket's text, video and web page natural language processing

In the expanding universe of content on the Web, it's virtually impossible to read or watch everything you want at the moment you first discover it. This was the problem that the founders of Pocket hoped to solve for their users.

"Pocket gives you one place to put everything, and it goes with you," explains Pocket founder and CEO Nate Weiner. When you find an article you want to read or a video to watch, but don't have the time to read or watch it right away, you can save it in Pocket until you have free time to enjoy it. He suggests that you think of Pocket as one component of your content ecosystem: "If Flipboard is your news-stand, Pocket is your nightstand."

The Challenge: Gleaning high volumes of clean data

Pocket's challenge was to pinpoint concepts and entities within saved content that enabled accurate classification of text, video and webpages. The results needed to be robust and clean enough to feed in-house developed APIs with data that facil-itated grouping and ranking based on format, relevance and the amount of time it would take to read or watch what users save.

The early version of Pocket could ascertain data points such as the source domain and how content was consumed, but was unable to expose entities and concepts well enough to create the consumer experience Weiner envisioned.

The Solution: AlchemyAPI's entity and concept extraction APIs

AlchemyAPI's concept and entity extraction features begin each day with natural language processing (NLP) of more than 500,000 unique items for Pocket. The

pocket

Company Profile:
- Smart Consumer vertical
- Save for later platform

NLP Problems:
- Obtaining clean data
- High volumes
- Multiple language processing

AlchemyAPIs Used:
- Entity extraction
- Concept extraction

Volume of Content Analyzed:
- Over 500,000 unique items/day
- 1.5 million items/day total

Type of Data Analyzed:
- HTML, text

Result:
- Classification accuracy for winning user experience
- Higher performance of in-house developed classification and recommendation APIs

cloud-hosted APIs decipher the unstructured data and HTML; then actionable data, ready for the next operation, is returned to Pocket.

"We use the raw data from AlchemyAPI and bubble up from there," says Weiner. "The processed text, video and web pages are called into Pocket's applications to further categorize it."

Of course, all this happens behind the scenes for users. They simply open Pocket and select what interests them at that moment via the UI.

A new feature, also enabled by AlchemyAPI's NLP intelligence, called "Highlights" recommends that users rediscover what they've already saved based on the types of content most saved and shared across the Pocket platform.

"AlchemyAPI gives us a huge addition to our data set that we don't have other-wise," Weiner acknowledges. In comparative testing against two other NLP provid-ers, Weiner found that "the data is a lot cleaner and there's language support that competitors just don't have."

> *"AlchemyAPI proved to be the most focused on solving the problem we are trying to solve. And everything about the company indicates that they are going to be here in the long-run."*

Weiner was convinced that he found the best solution after connecting with AlchemyAPI's sales team and running a one-day, 500,000-item test. The results echoed what his colleagues had told him of their experience with AlchemyAPI: "AlchemyAPI proved to be the most focused on solving the problem we are trying to solve. And everything about the company indicates that they are going to be here in the long-run."

The broad language support is critical for Pocket in order to create a consistent experience for users almost everywhere around the globe. Approximately 40% of all Pocket users store English-based content and given that 1.5 million items are processed each day and 900,000 items are in languages other than English, it's important to have a program that can detect articles in all languages. It's that kind of feature functionality that makes AlchemyAPI's solution valuable for global enter-prises.

Conclusion

Pocket has the grand challenge of discovering articles all around the world each and every day for their users. This challenge would not have been met with such grace had it not been for the NLP features that AlchemyAPI provides for their cus-tomers.

"AlchemyAPI gives us a huge addition to our data set that we don't have otherwise. The data is a lot cleaner and there's language support that competitors just don't have."

Nate Weiner
Pocket Founder & CEO

About AlchemyAPI

The product of over 75 person years of engineering effort, AlchemyAPI is a text mining platform providing the most comprehensive set of semantic analysis capabilities in the natural language processing field. Used over 3 billion times every month, AlchemyAPI enables customers to perform large-scale social media monitoring, target advertisements more effectively, track influencers and sentiment within the media, automate content aggregation and recommendation, make more accurate stock trading decisions, enhance business and government intelligence systems, and create smarter applications and services. To learn more about our company and services, please call us at 1-877-253-0308 or email info@alchemyapi.com.