



# Text Mining for Chinese Social Media

*Plus Alpha Consulting enables analysis of customer feedback from Chinese social media with Rosette®*

Established in 2006, Tokyo-based [Plus Alpha Consulting Co., Ltd.](#) specializes in software as a service (SaaS) for marketing, with visualization of information as its core offering. By applying text mining technology and data visualization, Plus Alpha develops, sells, and consults with customers to bring value to their businesses.

## The Challenge

Many companies have adopted Plus Alpha's SaaS text mining service, "Mieruka Engine," to analyze various data — such as questionnaires, call logs, blogs, and daily sales reports— to help understand their customers' needs. As Japanese companies started to enter the Chinese market — an environment considerably different from that of Japan — they required a tool to analyze local customer feedback in Chinese. In order to meet this demand, Plus Alpha planned a Chinese version of the "Mieruka Engine."

For automated analysis of Chinese text, Plus Alpha needed:

1. Tokenization of Chinese text (which is written without spaces between words) and part-of-speech analysis to enable keyword extraction
2. The ability to analyze both Chinese scripts: simplified Chinese (used on mainland China) and traditional Chinese (used on Taiwan).

## The Solution

Plus Alpha chose to integrate the Chinese module of Rosette Base Linguistics, which accurately divides Chinese sentences into words and parts of speech, from which keywords can be extracted with high accuracy and efficiency. Rosette provides pre-processing of Chinese text for “Mieruka Engine” to support simplified and traditional Chinese characters, and thus process data from any part of the Chinese-speaking world. It can instantly analyze various customer feedback collected in the Chinese market. Chinese information that appears on social media, such as Facebook and Twitter, can also be constantly monitored.

By strengthening such things as the dictionary function to support the language characteristics unique to Chinese, it became possible to quickly and accurately analyze answers to open-ended survey questions asked in the Chinese market, as well as customer feedback in Chinese collected by call centers.

“With Japanese companies increasingly expanding into the Chinese market, the demand for Chinese market analysis has also increased rapidly,” said Katsuya Mimuro, president and CEO of Plus Alpha Consulting. “With the introduction of the Base Linguistics Chinese module, our ‘Mieruka Engine’ could meet the needs of these companies in a timely manner, making it possible for them to appropriately analyze the needs of the Chinese market, and further grow sales of Japanese products in the Chinese market.”

## The Results:

Using Rosette Base Linguistics’ Chinese module, Plus Alpha Consulting’s “Mieruka Engine” has been supporting Chinese since May 2011 as the “Mieruka Engine Global Option,” enabling its users to meet the demands of the Chinese market.

---

Rosette® provides businesses and government agencies with text analytics in 55 languages. [www.rosette.com](http://www.rosette.com)

---