

Why Vital4 Chose Rosette for People and Company Search in Its Due Diligence Platform

ROSETTE

Whether it's for financial compliance or pre-employment screening, due diligence is all about finding and matching names against watch lists, negative news reports, SEC filings, and more. Vital4 is a standout data aggregator, providing and updating thousands of data sources 24x7 for due diligence vendors and corporate departments.

The data Vital4 provides is not just from watch lists but also negative news gleaned from articles around the world in many languages, and information about politically exposed persons (PEPs). Their strength is in aggregating thousands of sources more than their competitors and maintaining full FCRA (Fair Credit Reporting Act) compliance for U.S. customers, meaning that as soon as a person is removed from a watch list their name must be removed from Vital4 databases within 24 hours. Ingesting the data in whatever format it arrives and then making it accessible is a huge job.

The challenge

For Vital4 clients, all searches center on people and businesses. However, traditional keyword-based search did not provide the accuracy that Vital4 needed to satisfy its customers.

Its existing search solution lacked true fuzzy name matching capable of matching names that were very likely the same but appeared different

due to typos, phonetic errors, nicknames, initials, truncations, misordered name components, or were simply written in two different languages.

Additionally, when searching news articles, a search on "John Smith" might return an article that contained, "Smith explained that John and his team were responsible for the damage..." The article would be returned only due to the proximity of "John" and "Smith," but, it clearly doesn't qualify as an article mentioning a "John Smith."

Manually having human editors tag every article with the people and companies mentioned would have solved the problem, but was impractical with thousands of sources being refreshed daily.

The solution

Vital4 started by looking for fuzzy name matching and found Rosette® by Babel Street, an AI-driven platform that enables customers to match identities, analyze languages, extract and link

information, and interpret content. Additionally, Rosette is extremely compatible with Solr and Elasticsearch — both of which offer flexible search indices, making it an easy decision for Vital4 to migrate to Elasticsearch from their previous, rigid search system.

Intelligent fuzzy name matching

Vital4 clients expect the effortless experience they get from major search engines that can find relevant documents even with an imprecise query. Rosette offers Vital4 the ability to deliver highly relevant name search results that match the expectations of their users.

The greater precision of Rosette made it stand out. While many fuzzy name matchers rely on generating thousands of name variations in hopes of making a match, Rosette offers a sophisticated hybrid approach, using both phonetic matching and a statistical engine with a wide array of algorithms that contain knowledge about how names are written and pronounced in multiple languages, and how they are transliterated across scripts.

For every name search of people or organizations, Rosette considers 15+ variations including nicknames, abbreviations, truncations, and initials. Also, by flexibly comparing all the tokens between two names, Rosette overcomes any issues of misordered or missing name components.

Most importantly, every match is delivered with a match score that empowers the user to decide what should be considered a match. That same

scoring and weighting system integrates with the search engine's relevancy scoring algorithm to produce a single list of relevancy-ranked results.

"For name matching, we didn't see any other offerings that fit our needs as well as Rosette," Patrick Deeb, former CTO of Vital4 said.

AI-Powered Tagging

"Once we moved to a more flexible search index solution, it opened up the ability for us to make better use of the Rosette tools available to us," Deeb said. "We were able to start running all of our articles through entity extraction, and create an array of those entities to be searched on, rather than searching on the entire text of all the articles. This immediately added quality to the results."

Vital4 uses Rosette entity extraction to automatically identify entities (people, places, locations, etc.,) in each document. The scope of due diligence is such that Vital4 has to consider news articles in every language for its database, and so it makes full use of Rosette's multilingual features.

Vital4 looked at other entity extraction solutions, but the choice of Rosette was based on its quality results.

"We compared [Rosette] to some other text analytics and it seemed like the best," Deeb said. "The search we left behind was just not the kind of experience our clients wanted. It came down to what our clients were asking for — fuzzy name matching."

Babel Street is the trusted technology partner for the world's most advanced identity intelligence and risk operations. The Babel Street Insights platform delivers advanced AI and data analytics solutions to close the Risk-Confidence Gap.

Babel Street provides unmatched, analysis-ready data regardless of language, proactive risk identification, 360-degree insights, high-speed automation, and seamless integration into existing systems. We empower government and commercial organizations to transform high-stakes identity and risk operations into a strategic advantage.

Learn more at babelstreet.com.