



Rosette enables fintech to reduce false positives by 75% for KYC/AML



In Europe, businesses from the largest financial institutions to individual bookkeepers, attorneys, and real estate brokers must identify their customers to comply with European anti-money laundering (AML) screening regulations. These requirements are often burdensome, involving paper forms, collecting passport information, and ultimate beneficial owner (UBO) data. Finland-based financial technology provider, DOKS® introduces cost and labor savings through cloud-based know your customer (KYC) and AML tools and processes.

DOKS electronically collects and stores data for the entire KYC/AML process, digitizing the paperwork journey. It references data from local trade/business registers to determine UBOs and potential politically exposed persons (PEPs) in the organization, verifies identities through local banking IDs of Nordic customers and other electronic ID services, and assesses risk by screening against local and country-specific watchlists.

The challenge

DOKS has to match names and identities when:

- Screening against sanctions lists
- Verifying electronic IDs

DOKS' in-house name matching tool wasn't able to handle the multitude of variations in personal and company names and the increasing volumes of sanctions list screening. Their existing tool was also producing too many false positives in electronic ID verification. The cost of a DOKS customer missing a match to a sanctions list could result in significant penalties and damage their reputation.

Although nearly all DOKS transactions are in languages that use the Latin alphabet, the data on corporations, entities, and owners might be in any language or alphabet.

Using Rosette, the number of false positives has decreased by up to 75% in the best cases, yet hasn't compromised the reliability of the results.

—Petteri Tenhunen, Co-founder and Partner of DOKS

The solution

DOKS evaluated commercial and open-source name matching options. Besides accuracy — including the capability to match names in different languages and scripts — the ability to process large volumes of data quickly was a core requirement. Rosette® by BasisTech met the need for accuracy and speed because it plugs into Elasticsearch, which DOKS engineers were already familiar with.

Rosette has two other features that attracted DOKS: match confidence score and explainability:

Match confidence score: Rosette produces an intuitive score from 0 to 100%, enabling the user to establish a threshold score above which two names are considered a match. Some businesses may use a higher threshold to reduce false positives. Others may set a lower threshold to minimize missing true matches.

“This relevancy score in Rosette is essential,” said Petteri Tenhunen, Co-founder and Partner of DOKS. “It makes it possible to configure the name matching according to the customer’s specific requirement and reduce false positives.”

Explainability: With each name comparison, the AI and technologies powering Rosette provide the reasons behind each match score calculation. Because DOKS is under continuous audit by third parties, it must be able to show the rationale for assigned scores.

“When reviewing our due diligence reports, customers are often very interested to learn why our system ended up with a certain conclusion and why the match percentage was 80% instead of 65%,” said Tenhunen.

The screenshot shows the Rosette Match Score interface. At the top, it displays a 'Match Score' of 63% with 'Clear' and 'Compare' buttons. Below this is a 'Match Score Computation' table with columns for 'deaz' (Weight: 48%), 'chuy' (Weight: 46%), and 'a' (Weight: 7%). The table lists three match components: 'jesus' (Weight: 27%) with an 'OVERRIDE' match score of 0.557; 'alfonso - lopez' (Weight: 46%) with an 'INITIAL_MATCH' score of 0.366; and 'diaz' (Weight: 27%) with an 'HMM_MATCH' score of 0.381. At the bottom, the 'Final Score' is shown as 0.639.

	deaz Weight: 48%	chuy Weight: 46%	a Weight: 7%
jesus Weight: 27%		OVERRIDE 0.557	
alfonso - lopez Weight: 46%			INITIAL_MATCH 0.366
diaz Weight: 27%	HMM_MATCH 0.381		

Final Score
0.639

The administrative interface in Rosette explains the 63% match score for “Jesus Alfonso-Lopez Diaz” versus “Deaz Chuy A.” Rosette found a statistical match (HMM) for diaz and deaz; a nickname match (override) for Jesus and Chuy; and initials match for A. and Alfonso-Lopez.

“Name matching is central to reducing false positives, and we found few commercial providers with that focus,” said Tenhunen. “Our goal is to reduce the time our customers spend managing regulatory obligations. Rosette is state-of-the-art and complements our KYC product, which aims to deliver the best solution to our customers.”

The impact : Fewer errors in KYC/AML screening

Since DOKS switched all of its customers to the system using Rosette, the feedback has been positive. For every match to a sanctions list entry, DOKS customers are shown the queried name and the initial information from the list, such as date of birth, title, and place of residence, so they can quickly eliminate false positives and decide on the next course of action.

“Using Rosette, the number of false positives has decreased by up to 75% in the best cases, yet hasn’t compromised the reliability of the results,” said Tenhunen. “That means significant savings for our customers running their AML processes in DOKS.”

This relevancy score in Rosette is essential. It makes it possible to configure the name matching according to the customer’s specific requirement and reduce false positives.

—Petteri Tenhunen, Co-founder and Partner of DOKS

[Learn more about Rosette Name Indexer and request a demo.](#)

About BasisTech

Data analytics and machine learning are critical to verifying identity, understanding customers, anticipating world events, and uncovering crime. BasisTech provides businesses and governments with advanced analytics and AI-powered solutions for deriving insights from multilingual text, connecting data silos, and discovering digital evidence. Our **Rosette**[®] text analytics platform employs classical machine learning and deep neural nets to extract meaningful information from unstructured data. **Autopsy**, our digital forensics platform, and **Cyber Triage**, our incident response tool, serve the needs of law enforcement, national security, and legal technologists. **KonaSearch** delivers deep search across Salesforce and other data sources.