Name Screening AI Helps Digital Economies Play by Real-World Rules

ROSETTE

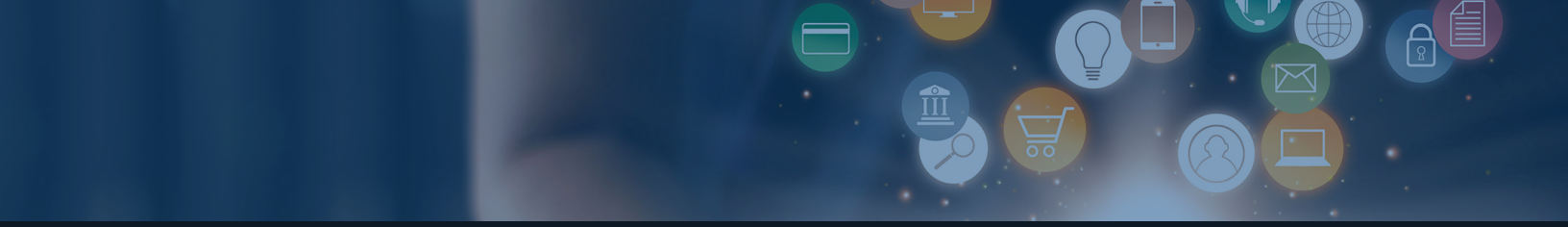
Payments played a critical role in transitioning physical goods from storefronts to e-commerce to online marketplaces. The same trends are happening for digital content. Digital content ecosystems such as gaming, social platforms, and next-gen marketplaces are becoming more prevalent and economically important. The success of these online platforms and user-generated content- (UGC) driven economies is creating a need for people and businesses to get paid directly for their digital content, which is often a regulated activity. The traditional financial system, however, is limited for digital content and doesn't offer the flexibility to unlock its full potential.

Enter Tilia, a trusted financial system that supports instantaneous, micro, high-volume payments for any digital ecosystem. Originally, Tilia was spawned in 2005 by Linden Labs, the creator of Second Life which had one of the earliest digital economies. Today as a licensed money transmitter in the U.S., Tilia provides payments infrastructure technology to any digital economy by facilitating secure, regulated transactions at scale. It handles everything from in-game payments to pay outs, tax management, chargebacks and fraud protection, customer support, and compliance with financial regulations.

The challenge

A key component of the Tilia financial compliance service is screening all payers and payees for sanctioned individuals and organizations. Fraud checks and sanctions screenings require accurate name matching across multiple languages, as user registrations might not be written in English or in Latin scripts.

The screening system needed name matching accurate enough to permit true matches to fall above a threshold match score (for example, 80%) and non-matches to fall below. To test various systems, Tilia created a test data set containing known entities to compare against the Office of Foreign Assets Control (OFAC) sanctions list. The list largely contained names that would not match and a smaller number of names that would.



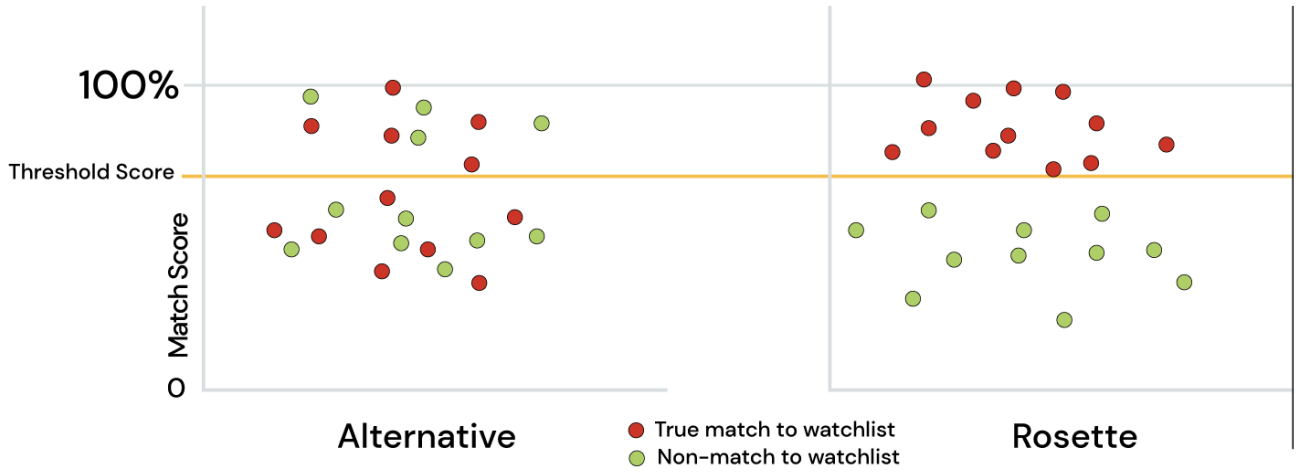
Tilia provides payments infrastructure technology to any digital economy by facilitating secure, regulated transactions at scale.

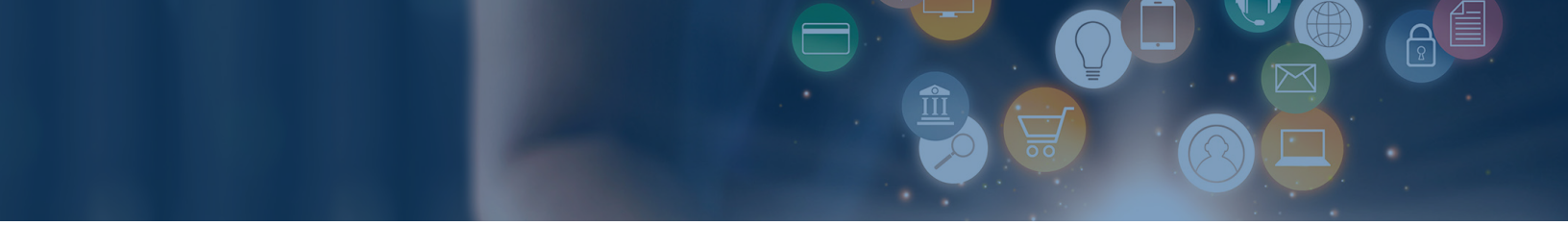
Michael Jacobi, Senior Engineering Manager at Tilia, and his team tested an open-source name matching library and the name matching technology of Rosette® by Babel Street.

The open-source library fuzzy matched names using a phonetic match (Soundex) algorithm and edit distance (how many character differences there are between two names). It generated thousands of false positives, and missed matches obvious to a human, such as a name with and without an umlaut (Stüber/Stuber). Given a match score threshold, the plan was for the compliance team to review all matches above the threshold.

"The open-source library's distributions of false and true positives were right on top of each other. The significant overlap meant we would have to choose between a high threshold that missed true positives or a low threshold that overloaded our compliance team with false positives," Jacobi said.

The test results from Rosette turned out to be exactly what Tilia needed.





The solution

When Jacobi ran the test set through Rosette, the true matches and non-matches separated into two distinct groups: true positives above the threshold score and non-matches (false positives) below.

Rosette easily met Tilia's requirements with its ability to evaluate names for more than a dozen variation types, such as nicknames, transliteration spelling variations, initials, out-of-order name components, phonetics, and the same name written in different languages.

Furthermore, for organizational name matching, Rosette also leverages advanced natural language processing (NLP) that checks name components by meaning (semantic similarity). This is particularly useful when matching names across languages, since organizational names are often translations of the meaning of each word.

The support for non-English names of people and companies was important because of the international nature of video games. Tilia receives user credentials in many languages other than English. Furthermore, that Rosette is self-hosted was another bonus for the fast-paced Tilia environment where every millisecond counts.

Phonetic similarity
Jesus ↔ Heyzeus ↔ Haezoos

Transliteration spelling differences
Abdul Rasheed ↔ Abd al-Rashid

Nicknames
William ↔ Will ↔ Bill ↔ Billy

Missing spaces or hyphens
MaryEllen ↔ Mary Ellen ↔ Mary-Ellen

Titles and honorifics
Dr. ↔ Mr. ↔ Ph.D.

Truncated name components
Blankenship ↔ Blankensh

Gender
Jon Smith ↔ John Smith (but not Joan Smith)

Missing name components
Phillip Charles Carr ↔ Phillip Carr

Out-of-order name components
Diaz, Carlos Alfonzo ↔ Carlos Alfonzo Diaz

Initials
J. E. Smith ↔ James Earl Smith

Name split inconsistently across database fields
Rip · Van Winkle ↔ Rip Van · Winkle

Same name in multiple languages
Mao Zedong ↔ Mao Цзэдун ↔ 毛泽东 ↔ 毛澤東

Semantically similar names
PennyLuck Pharmaceuticals, Inc. ↔ PennyLuck Drugs, Co.

Semantically similar names across languages
Nippon Telegraph and Telephone Corporation ↔ 日本電信電話株式会社

Organizational aliases
Very Fine Groceries, LLC ↔ VFG

Some of the many types of name variations that Rosette considers in every name comparison.

The impact

"Rosette is core to our services because we are required to do sanctions screening, and state regulators audit us periodically. They provide a list of names to run through our system and check the results to ensure our sanctions screening is sufficient for the money transmitter license obligations," Jacobi said. "Rosette has been passing well."

With the success of the sanctions screening, Jacobi says Tilia is now looking into potentially using Rosette in other areas, such as detecting users with multiple accounts trying to game the system by redeeming a one-per-person marketing promotion multiple times under different usernames.

The accuracy of Rosette did produce an unexpected problem. True and false positives were almost too cleanly falling on either side of the match threshold.

"We were concerned that if the compliance team didn't see a few false positives once in a while, they wouldn't know what to do if they ran into one," Jacobi said. "So, we've set the match threshold just slightly lower than we need to so at least 10-15 false positives a week get sent to the compliance team to keep them sharp."

Rosette is core to our services because we are required to do sanctions screening, and state regulators audit us periodically.

— Michael Jacobi, Senior Engineering Manager

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.**

BABELSTREET