airbnb

Use Case:
Identity Verification

Segment:
Sharing Economy

Products:
Rosette

Functions:
Name Analytics

Availability:
API or SDK

Rosette Name Matching Keeps Guests and Hosts Booking "Without Reservations"

Airbnb puts the "unique" into travel by connecting private owners renting out a room, apartment, castle, or villa to travelers staying a night, a week, or a month. This community marketplace of accommodations spans 34,000 cities and 190 countries. More than 17 million guests have over 800,000 listings to choose from. People can book accommodations or monetize extra space through the web or their mobile phone.

> "given the international nature of their business, Airbnb's Verified ID process helps users match names that originate in multiple languages, and in more than just the Roman A-to-Z alphabet"
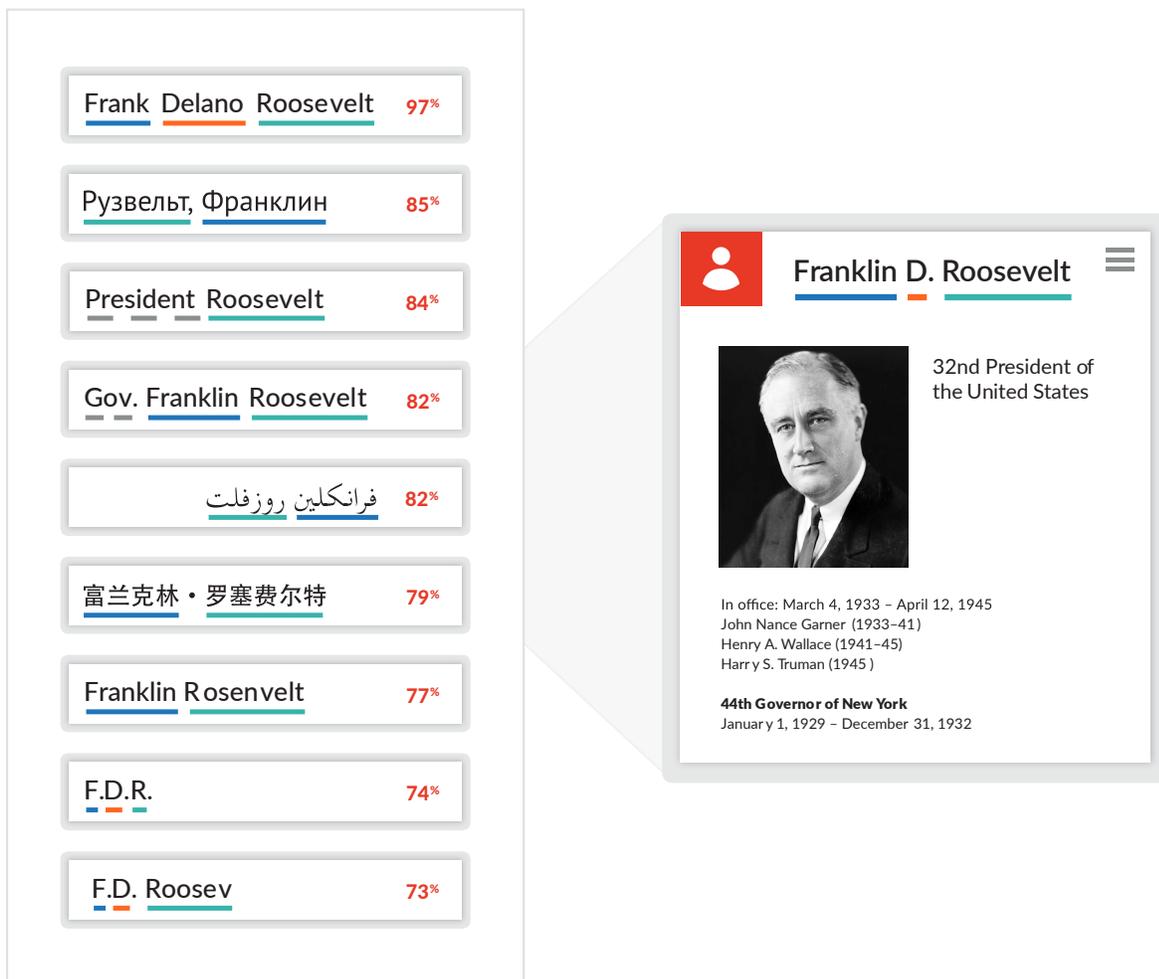
The Challenge

Trust is foundational to the sharing economy. There is no place for anonymity in a community based on trust. That's why Airbnb is dedicated to providing their users with the best decision-making tools possible. While profiles and reviews have been core to the Airbnb experience since the beginning, it became more important to find new ways to connect the online and offline worlds to provide users with even more tools to make informed decisions. That's why Airbnb introduced Verified ID.

Verified ID helps provide a connection between the online and offline identities. Airbnb users can earn a "Verified ID" badge on their profile by providing their online identity (via existing Airbnb reviews, LinkedIn, or Facebook) and matching it to offline ID documentation, such as confirming personal information or scanning a photo ID. The name provided by both channels must match for verification to succeed.

ROSETTE
TEXT ANALYTICS

What Is Name Matching?

Name matching is a key component of the Verified ID process. And, given the international nature of their business, Airbnb's Verified ID process helps users match names that originate in multiple languages, and in more than just the Roman A-to-Z alphabet. Users were already inputting their names using the Cyrillic alphabet (used by Russian, Bulgarian, Serbian, among other languages).

Frank Delano Roosevelt **97%**

Рузвельт, Франклин **85%**

President Roosevelt **84%**

Gov. Franklin Roosevelt **82%**

فرانكلين روزفلت **82%**

富兰克林・罗塞费尔特 **79%**

Franklin Rosenvelt **77%**

F.D.R. **74%**

F.D. Roosev **73%**

Franklin D. Roosevelt

32nd President of the United States

In office: March 4, 1933 – April 12, 1945
John Nance Garner (1933–41)
Henry A. Wallace (1941–45)
Harry S. Truman (1945)

44th Governor of New York
January 1, 1929 – December 31, 1932

Case Study: Airbnb
Rosette Name Matching Keeps Guests and Hosts Booking "Without Reservations"

What's So Hard About Name Matching?

Airbnb is loaded with talented engineers, so their first thought was to ask their developers to investigate building a name matcher exclusively in-house. Once they got into scoping out the work though, "We started to see the challenges across languages to be resource intensive and complex," said McCabe. "And there just aren't many solutions out there."

Most commonly, name matching is used in other industries for different use cases. Name matching functionality is often a component of a much larger solution that is typically expensive, unwieldy, and complex.

So building their own name matcher wasn't an attractive solution, but neither was buying an existing system.

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

Same name in multiple languages
Jesus ↔ Heyzeus ↔ Haezoos

Phonetic similarity
Abdul Rasheed ↔ Abdulrashid

Nicknames
William ↔ Will ↔ Bill ↔ Billy

Initials
J. E. Smith ↔ James Earl Smith

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

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

Missing name components
Phillip Charles Carr ↔ Phillip Carr

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

Truncated name components
McDonalds ↔ McD ↔ McDonald

Name split inconsistently across database fields
Dick • Van Dyke ↔ Dick Van • Dyke

Rosette name matching recognizes spelling variations when matching names of people, places, and organizations in 10 languages in their native scripts. Its fuzzy matching recognizes differences such as phonetic variation ("Cairns", "Kearns", "Kerns"); nicknames ("William", "Will", "Bill", "Billy"); swapped word order ("Ichiro Suzuki" and "Suzuki Ichiro"); cross-lingual matching and transliteration variations ("Sergei Rachmaninoff," "Sergey Rachmaninov," "Сергéй Васи́льевич Рахма́нинов"), and more.

"Other solutions that we could have adopted were built for different use cases, such as government and finance. We had a simpler use case," McCabe said. "So for us to find a company like Basis Technology willing to work with us to use a name matcher in other ways was very helpful."

Why Choose Basis Technology?

In the end, the choice came down to cost and flexibility.

Basis Technology's name matching is a functionality of Rosette®, highly flexible and accurate text analytics software. It intelligently matches names based on algorithms that take into account the origins and structure of names from different languages and cultures. The name matching functionality can be plugged into a system through a software development kit, or called as a web API by applications of any type.

Problem Solved

Once Airbnb chose Basis Technology's Rosette, their engineers did some manual tests and concluded the matches provided a good level of confidence.

"Basis is the glue in the solution for us. It's successful by how little we need to deal with it [the name matcher]. It's a fundamental piece, and also the piece we have to interact with the least once we integrated the solution," McCabe said. "In the end, we were much happier to go with something already built out in a few languages. It saved us a lot of time to focus on the actual verification piece," McCabe said.

In January of 2014, Airbnb is at 17 million guests all time and is adding about one million guests per month. In 2012, Airbnb estimated that a night was being booked every 2 seconds!

Rosette® provides businesses and government agencies text analytics in 55 languages. www.rosette.com