



# Babel Street Match Comparison to Common Alternatives

Babel Street Match was evaluated against three common alternatives using a dataset with 7,571 names, with at least 10 variants for each name.

These alternatives included:

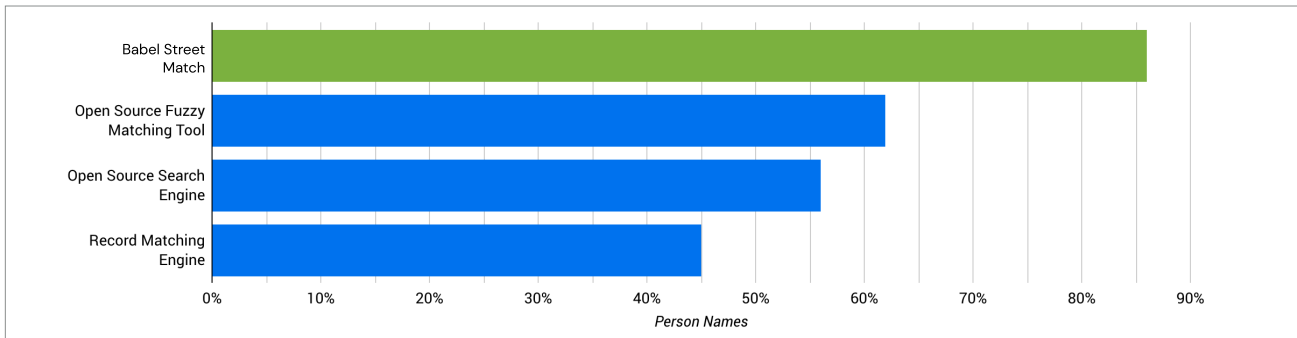
- An open source fuzzy matching tool
- An open source search engine
- A record matching engine

Testing and analysis show that these alternatives fall short of Match because they lack

script/language support, lack essential name phenomena support, and use rigid or overly simplified methods to calculate match scores

## Superior accuracy

Where a correct match is defined as matching a "gold standard" version of a name to one of its variants, Match outperforms the alternatives by 24% or more for person name matching.



## Coverage of match phenomena

Phonetic similarity

Kailey ↔ Caylee ↔ Kaylie

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 ↔ Мао Цзэдун ↔ 毛泽东 ↔ 毛澤東

Semantically similar names

PennyLuck Pharmaceuticals, Inc. ↔ PennyLuck Drugs, Co.

Semantically similar names across languages

San'in Telegraph and Telephone Corporation ↔ 山陰電信電話株式会社

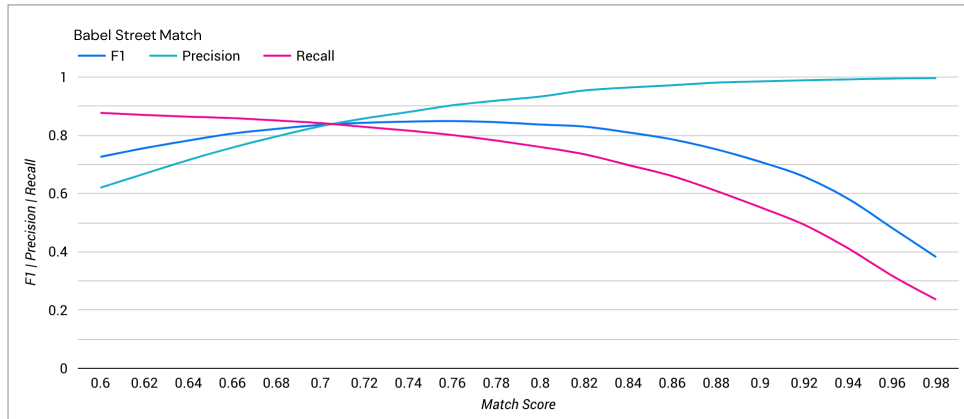
Organizational aliases

Boston Brewing Company ↔ BeantownBeer

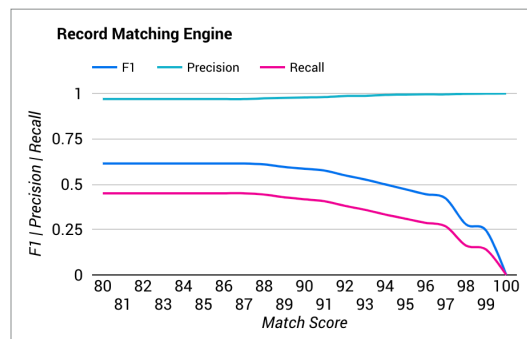
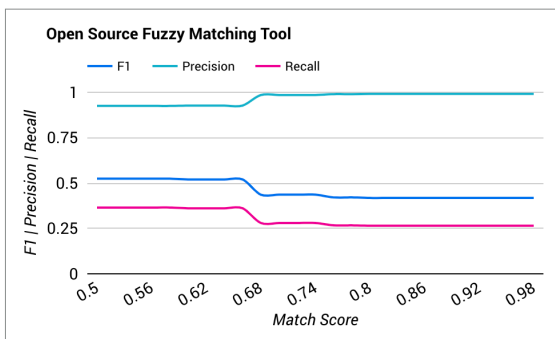
## Useful match scores

Match outputs a nuanced match score as a decimal ranging between 0 (no match) and 1 (perfect match); this match score can be used to balance precision and recall.

As the graph below shows, the precision and recall of Match meet at a point around .72; users of Match can look at lower scores to see more possible matches, and at higher scores to find only the most similar matches.



By contrast, the open source matching tool and the record matching engine operate in a binary “match” (score=1) or “no match” (score=0) paradigm without a range to indicate degrees of match. In this case, it is less clear what threshold will produce the desired balance of precision versus recall.



Furthermore, the open source search engine does not provide a comparable score. Thus it is not possible to compare match scores across multiple queries or configure business logic around the results.

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](https://babelstreet.com).