Why is holistic matching a game changer ?

A radically new approach to watchlist screening is boosting quality and reducing waste.

Over the last twenty years the financial industry has been continuously tweaking and tuning screening solutions to deal with increasingly complex regulations and ever-growing volumes. During all these years, however, the approach and algorithms at the heart of screening have hardly evolved, and the consequences for users in terms of costs and performance are becoming unsustainable. Fortunately, a new matching approach -**holistic matching**- has the ability to dramatically improve the situation, both in terms of efficiency and effectiveness.

What is holistic matching?

Most watchlist screening algorithms on the market today share the same two-phase approach to screening (be it for transactions or client records): first screen the name element, and if -and only if- there is a match on the name, then try to discard false positives using other available data elements. This second phase usually uses rules combining e.g. date of birth, address, ID number... to eliminate unneeded hits.

This two-phase approach is a direct consequence of the limited computing power that was available in the 1990's, and it represents a typical approach in computer engineering: if your system can't tackle a large problem, then split it in smaller chunks and solve each part sequentially.

While ultimately this approach gets the job done, operational and organizational impacts are significant: financial institutions have had to implement thousands of post-screening rules to discard false positives - each rule having to be audited by the relevant internal and external teams. As complexity and dependencies grow over time, after some time nobody dares to change these rules by fear of unintended consequences.

At the opposite, holistic matching takes a very different approach by ensuring ALL data elements available are matched at the same time and using the appropriate semantics to match each data point.

To illustrate this, let's consider a simple client record:

```
{"Name": "John Smith",
"DateOfBirth": "11/1/1968",
"CityOfResidence": "Jersey City"}
```

A screening engine using holistic matching will match all these data elements at once, while applying a different logic for each data point:

- The "Name" element will be matched as an entity's name, hence using fuzzy logic algorithms to detect all possible typographic, cultural and phonetic variations. Fuzzy techniques also include transliteration to compare different alphabets and Natural Language Processing (NLP) to isolate names in plain text. At the end, this will allow to match "John Smith" with watchlist records such J. Smith, John B. Smith, Smith John, etc.

- The "DateOfBirth" element will be matched as a date, with parameters allowing to set a confidence interval in terms of months or years of risk tolerance. For instance, if the risk tolerance is set at 3 years, listed individuals with a date of birth ranging from 1965 till 1971 would be considered valid matches for the 1968 date in John Smith's record.

- The "CityOfResidence" -and any location information- will be matched as spatial coordinates (like those provided by a GPS), with parameters allowing to set a "distance radius" confidence interval. This means any location data point is not compared to locations in the watchlist like two city names, but rather as a distance comparison between spatial coordinates. Indeed, from a semantics point of view, it makes more sense to compare two locations using GPS coordinates than to compare the letters of their names. To illustrate this with our example record, if the distance radius is set at 50 kilometers, watchlist entries located in New York (US) would be considered close matches, while a traditional name comparison between "Jersey City" and "New York" would yield no match.

The same logic would apply to ALL data elements available, such as place of birth, Passport (and any ID) number, etc. Any data element available will be matched at the same time and according to the appropriate semantics and matching logic.

As all of these individual elements feed a "meta score". A screening engine using Holistic matching would only raise alerts for which this meta score would be above a risk threshold set by the user. **Not only is this a much cleaner approach, but it also optimizes both effectiveness (catching "Jersey City" where traditional matching would have missed it) and efficiency (only returning list entries where the combined score of elements makes sense).**

What are the benefits of holistic matching?

Holistic matching offers three key benefits:

1. **Simplicity**. Matching all data points at once and applying the appropriate semantic and matching logic for each element removes the need for having thousands of post-screening rules, and in most cases removes the need for post-screening rules entirely. Instead, rules are replaced by much simpler (and intuitive) matching parameters such as "distance radius", "date of birth precision", etc. This is much easier to control by compliance operational teams and much easier to review by internal and external auditors. Moreover, these parameters provide extreme adaptability as they can be different for each single screening request.

2. **Performance and customer experience**. Performance of the screening process in terms of efficiency, latency and throughput is dramatically improved. This is essential to address the increasingly growing volumes of digital payments but also the SLAs of some types of transactions: real-time payments must be processed end-to-end in milliseconds. Through its higher efficiency (i.e., fewer false positives), holistic matching also allows to significantly reduce customer friction (and related operational costs).

3. **Effectiveness.** Holistic matching can also greatly improve screening effectiveness. By looking simultaneously at all data elements at once, one can detect a possible match on a record even if the "Name" component of the record is not a strong match but is reinforced by strong matches on other data points. This is something legacy name-based watchlist screening solutions can't do.

What's next?

These new and better screening technologies are available. Yet adoption by financial institutions is hampered by the inertia linked to legacy systems and technical debt. This is even more true for financial crime compliance, where regulatory scrutiny makes it more complex to adopt new technologies. Yet, immediate benefits can still be captured through leveraging holistic matching for:

1. **New use cases.** Such as real-time payments which saw their usage grow exponentially and for which financial institutions and fintechs often do not have appropriate watchlist screening solutions yet.

2. **Secondary screening**. While it can be difficult for a financial institution to justify the replacement of their legacy screening technology, they can still implement new solutions for secondary screening, using them as a tool to reduce the volume of alerts raised by their legacy screening solution to a manageable size. Secondary screening allows to capture at least part of the benefits and to grow confidence in the new technology over time.

Contact Us

✉	info@neterium.io
in	linkedin.com/company/Neterium
🐦	@neterium