Watchlist screening : Build or Buy?

Why fintechs need to think carefully about what technology to keep inhouse.

"Anything that gives you differentiation with customers, you should build [...] but for most back-office operations, and for things that won't give you differentiation with customers, you should buy."

(**Jeff Lawson**, founder and CEO of Twilio, in his book "Ask your developer")

In the past, to start a technology business such as a fintech, you had to build -at great cost- every single component that was required to deliver the service: user interface, back-office, operations… but also infrastructure, network or security. Today, all this changed radically, and fintech founders can rely on cloud services, open-source components and APIs to build their product. This allows them to concentrate their efforts on their core offering and hit the market years earlier than their older peers. For each required component, the "build or buy" decision is made quickly, often by development and product teams estimating the effort required to build the feature vs. the differentiating value it brings.

Watchlist screening – the regulatory requirement to screen client names and financial transactions against sanction, politically exposed persons (PEP) or adverse media watchlists – is a problem that looks deceptively simple.

Watchlist screening is a problem that looks deceptively simple, until the reality catches up, sometimes with an unpleasant reminder from the regulator

To a developer, it may look like all you need is to find whether names in customer records or financial transactions are present in these lists, basically nothing more than a glorified version of the CTRL-F search as it exists in Excel.

It looks so deceitfully simple at first sight that many fintechs, platforms -and even some adventurous banks- quickly task a team of in-house developers with crafting a solution to cope with the regulatory requirements. A quick Google search seems to indicate that "Levenstein edit distance" is the algorithm to use, Wikipedia even provides code samples, great! Problem solved… or at least so it seems. Until the reality catches up and the real complexity of screening comes to light, sometimes with an unpleasant reminder from the regulator, and board members suddenly aware of their personal liability.

It is all about SCALE

While it looks straightforward to screen a single client record against a sanction list containing a limited number of names of terrorists or backlisted entities, the problem is actually much more complex:

- *Millions of input records*. As fintechs grow, they quickly have more than a few thousand clients and transactions. What is really required is screening hundreds of thousands of transactions, or millions of names, every single day.
- *Millions of list entries*. Names or transactions have to be screened against ever growing lists: as an example, a typical PEP list contains over 4 million entries, frequently having many different names for the same person (unsurprisingly, listed individuals tend to have many aliases)
- *Unstructured data*. Input data itself is not always well structured, this is especially the case for payments transactions. Identifying individuals, companies, locations, ID numbers… in these transactions require advanced entity resolution techniques.
- *Fuzzy logic*. One has to cater for misspellings, different ways of writing names (e.g. "Smiths, Richard", "R. F. Smiths", "Richard Fitzgerald Smiths", "Dick Smitts"), translations of cities names (e.g. "London","Londres", "Londen", "伦敦"), and many other variations. Advanced fuzzy logic algorithms need to be implemented to cope with all various possible variations.
- *Different alphabets*. Many languages and alphabets are being used in the world (e.g. Cyrillic, Chinese, Japanese, Arabic…) and should be handled both in their original form and in a transliterated format, making sure "Игорь Ива́нович Се́чин" matches with "Igor Sechin", and "محمد سمير حداد" with "Mohammad Haddad".
- *Cultural diversity.* Korean, Chinese, Arabic, Russian, Hispanic… names have specific typologies and matching them like western names generates more false positives, and can even lead to missed detections.
- *It all adds up*. Of course, all the above cases actually combine, and you still have to find a name, even wrongly spelled, in a different alphabet and in the middle of a sentence in an unstructured data field of a transaction.

As we can see, scale transforms what appears to be a simple problem into a significant challenge. Unfortunately, it also appears that the basic approach does not scale at all, as edit distance algorithms become exponentially slower as data sets and lists grow, and they obviously can't handle complex fuzzy matching in multiple languages, cultures and alphabets. Moreover, scaling is not only about handling large data sets: performance is also a factor. When fintechs onboard new clients through an app, they expect sub-second response times or have the risk of losing their prospect. And when screening financial transactions, payment networks typically expect a screening system to reliably respond within a few milliseconds. Achieving all of the above within these constraints is very difficult.

Scale transforms what appears to be a simple problem into a significant challenge.

It is very hard to master the "magic spot" between effectiveness, efficiency and performance

The challenge of scaling is compounded by the need to meet key additional requirements around effectiveness, efficiency and performance.

Effectiveness is the license to play. *Effectiveness* is the most important requirement for any watchlist screening solution. This is what regulators expect: that no single name or transaction falls through the cracks. Regulated entities have zero risk appetite for 'false negatives' (i.e. missed hits) as the legal and financial consequences can literally end a business. The regulator's expectation goes way beyond 'exact matches': any watchlist screening system must implement at least fuzzy logic and be able to handle unstructured input.

Efficiency reduces costs and transforms customer experience. When your fintech is starting and screens thousand customers and 100 transactions per day with a rate of false positives of 20%, this may seem manageable. But once you have hundreds of thousands of clients and thousands of transactions to screen every day, every single percent of improvement in the false positive rate will make a huge difference in operational costs (think how the resources tied up by alert review could be better made available for analysis or investigation). Even more importantly, customer experience is directly impacted by efficiency: any payment or client on-boarding unnecessary held by the screening system adds friction to customer conversion and retention. Increasing the efficiency of a watchlist screening solution is an art: it requires leveraging the latest technologies (AI, parallel computing...), understanding the subtilties of various languages and offering extreme configurability. Combining all this requires years or decades of field experience.

Customer experience is directly impacted by efficiency: any payment or client on-boarding unnecessary held by the screening system adds friction to customer conversion and retention.

Performance (latency, throughput) is the ticket to the premier league. Beyond the necessary effectiveness and efficiency, a number of use cases will require truly distinctive performance. A good example is the screening of real-time payments, which demands the screening process to be completed in less than a few milliseconds. Another one is the continuous monitoring of large client databases: as best practice calls for daily re-screening of client databases, sophisticated watchlist screening solutions need to process millions of records per minute to deliver this capability.

Things continuously evolve

Developing a watchlist screening solution is not a one-off effort, as things evolve fast and often:

- Watchlists change every day. Every day, new names are added (and removed) from sanction or PEP lists. Every day, new press articles are added to adverse media databases. Watchlist screening systems are expected to always screen against the most up-to-date lists. Handling daily updates of all different kinds of watchlists is an essential part of any screening solution, and considering the liabilities involved, this process needs to be flawless, monitored and audited.
- Regulatory Requirements continuously grow. Evolving regulatory requirements may create the need for new watchlists (like when the Ukraine-related sanctions imposed to screen against entities indirectly held by sanctioned individuals or entities) or new sources of information (e.g., Environmental, Social and Governance data). Regulators' evolving expectations may also require looking for new information in financial transaction fields (e.g., dual-use goods, specific asset classes).
- Transaction formats also evolve, and new payment methods emerge. The ISO20022 format is expected to be generally adopted in the coming years (while initially co-existing with legacy domestic and cross-border formats). Watchlist screening solutions will also need to increasingly deal with new payments methods such as instant payments, credit transfers running on credit card rails, cryptocurrency transfers...

All these changes require to continuously invest in evolving watchlist screening solutions, and the advantage of using a vendor solution is that the vendor does all this for you. Fintechs deciding to build their own solution need to keep significant resources constantly allocated to maintaining their screening tool, not only developers but also legal and compliance staff to translate new regulations into technical requirements. Moreover, new regulations are often driven by geopolitical crises and foreign policies, meaning they can pop up at any moment and be enforced very quickly, making resource and budget planning of updates close to impossible for a company also having other priorities.

Fintechs deciding to build their own solution need to keep significant resources constantly allocated to maintaining their screening tool.

Focus on your core business

The "Build or Buy" debate is common to all industries. For tech savvy firms such as fintechs, the temptation can be even more important to develop an in-house solution for what can appear to be a simple problem.

If you do, do it with your eyes wide-open: developing a state-of-the-art watchlist screening solution is a multi-million investment requiring deep field expertise and calling for significant on-going maintenance work. And experience shows that there is a high chance that – even with large investments - you may not get the results you expect, fall short on the regulator's expectations, and potentially hamper sales opportunities.

Eventually, it is important for fintechs to remember they should only build what truly differentiates their value proposition. For every component that is not part of the core proposition, it is often better to rely on a knowledgeable partner, and by doing so accelerate time-to-market, especially since fintechs operate in an extremely competitive environment. Using specialized components through APIs is the best choice to remain totally in control of user experience, and it strongly reduces project risk because technical integration is simple, standard and easy to test.

It is important for fintechs to remember that, in a highly competitive environment where time-to-market is key, they should only build what truly differentiates them.

While watchlist screening certainly is a required component for most fintechs (as part of the KYC process, customer monitoring or transaction screening), it is not core to the product they build. So, they should rather fulfil these crucial compliance obligations quickly and simply by embracing specialised Software-as-a-Service watchlist solutions, benefiting from the vendor deep experience and continuous investment in R&D. By doing so, they can save time and resources that can be dedicated to developing their truly distinctive value proposition.

Contact Us
info@neterium.io
linkedin.com/company/Neterium
@neterium