

The Guide to External Data for **Better User Experiences** in Financial Services



Table of contents

Introduction	3
Chapter 1: What do you need to know about your customers?	6
Chapter 2: What data do you need?	10
Chapter 3: Integrating data into your KYC processes	12
Chapter 4: Making it swift and scalable	14
Final thoughts	17

Introduction

How well do you really know your customers? If the stats are anything to go by, the answer is probably: not well enough.

Fraud and identity theft are soaring, with losses in the US alone increasing by nearly 17% in 2019. Credit card and debit card fraud was up 35% in April 2020 compared to the previous year. Meanwhile, the technology used to keep bad actors out of financial services is falling short. In anti-money laundering (AML) detection 95% of system alerts turn out to be false positives.

All of which means millions of dollars wasted on investigating customers who have done nothing wrong, while the real criminals continue to slip through the net. On the one hand, financial service providers lose money on missed transactions by legitimate users. On the other, they also lose money from the collateral costs of fraud, including chargebacks and fraudulent claims. It's a lose-lose situation.

Clearly, the old Know-Your-Customer (KYC) methods aren't helping lenders to predict risk effectively. At the same time, the systems and processes they're using slow down the onboarding process for the customers they really want.

Done poorly, KYC creates substantial, frustrating bottlenecks. It also fails to detect real fraudsters by relying on the wrong indicators, putting you at even greater risk. The status quo is very bad business indeed.

Thankfully, there's a much better way. By combining smarter access to relevant data with extensive automation of the KYC process, you'll be able to make better predictions, more accurately. You'll know which parts of your protocol deliver results. Even better, you'll be able to achieve this in a smooth, streamlined, scalable fashion, speeding up the decision-making process, and delivering excellent user experience.

In this white paper, you'll discover:

- How the right data helps you build a system that detects actual cases of fraud without blocking legitimate transactions
- How to integrate this data into your KYC processes in ways that are streamlined and scalable
- How to create a frictionless experience for the customers you really want



Chapter 1:

What do you need to know about your customers?

Every type of financial business or service encounters its own, specific types of risk. That means they need different things from their data, including different information about their customers, before they can make an informed decision about working with someone. This makes it tricky to design a single, simple, customer-friendly process that every financial institution can use. The underlying goal here is to build a system that prevents fraud while letting legitimate transactions through, which is easier said than done.

If you're dealing with a loan application, for example, your primary concern might be fraud. Are there clues in the data that the applicant may have committed identity theft and is using someone else's details to gain access to credit?

If you're processing a new bank account application, you might be more worried about your AML obligations. Can this person prove the money coming into their account is from legal sources? Are there irregularities

in their transaction patterns that align with fraudulent behavior?

Exactly what you need to know depends on your business type. For example:

- **Banks and lenders**

For this sector it's vital to make sure that the person applying for an account or credit a) is who they say they are, b) can genuinely cover repayments, and c) acquires their money through legitimate means. At a basic level, this means establishing a person's identity, clarifying the source of their funds (which must be sufficient and legitimate), and accurately predicting the money laundering risk associated with that customer.

- **Investment brokers**

Brokers have to get detailed information about a prospective client's financial circumstances, level of investment knowledge, and risk tolerance. This dictates what they should and should not include in this person's investment portfolio. Data includes: annual income, net worth, employment status, identification numbers, investment objectives, financial and liquidity needs, tax status, existing

investments, experience, risk tolerance, and investment time horizon.

- **Insurance companies**

Detecting fraudulent claims is a major priority for insurance companies. It is especially important for them to get a picture of the person's trustworthiness and past behavior. This includes other claims they have made and any previous convictions - especially for fraud, and including those from outside the state or country. Data that shows signs of risky or erratic behavior, from cycling through changes of address or employment through to questionable activity on social media, may also affect their decision to insure. Additionally, insurance companies and brokers have to be especially vigilant against fronting (impersonating another person). This means they may also need data or AI programs that help them to predict whether a document / ID has been tampered with or forged.

Reducing the impact on the customer

With KYC fastly becoming a regulatory mandate, compliance with it typically centers around monitoring transactions, asking customers to sign clear, exhaustive acceptance policies, and comprehensive

identification procedures. There are three problems with this, however:

1. Subjecting above-board customers to this long and arduous process is hardly ideal. You want to avoid getting off to a bad start by creating barriers and frustrations.
2. If the process takes too long, the customer might jump ship and open an account with a competitor who has figured out a way to meet their KYC obligations faster.
3. Your investigations will only be effective if you have all the available information you need to assess the risk. Not just what the customer provides.

To get a complete picture, you'll need to look at external data sources, too. Plus, the more insight you can collate accurately and quickly about this customer from external sources, the less paperwork and manual checks you'll need to put them through, and the better the user experience will be.



Chapter 2:

What data do you need?

Once you've clarified what you need to know about your customer, you can pinpoint what you still don't know... and decide what external data will help fill in these gaps.

The kinds of external datasets you should access might include:

- Standard **person data** relating to that individual's prior records.
- **Company directorships and beneficial ownership information.** This is key for spotting worrying trends, such as a single person opening or sitting on the board of directors for thousands of different companies — a common sign that this person is involved in creating shell companies for tax evasion purposes.
- **Location and other geospatial data.** This can indicate whether this customer's details are being used in multiple locations simultaneously, as well as other unusual patterns that could be a predictor of fraud.
- **Financial data.** This can include savings and investments, but also alternative data such as bill and rent payments, which provide

insights into their reliability and suitability for credit.

- **Home price data**, including details like mortgages and loans. This helps you to gauge levels of debt, fiscal responsibility and true asset worth. It can also flag up disparities between declared income and asset wealth, which may be a predictor of money laundering activities.
- **Employment history and associated records**. This also helps to account for income and wealth levels as well as assessing reliability and trustworthiness.
- **Sanctions screening lists**, such as OFAC SDN. This is a basic but crucial measure to ensure that you are legally permitted to do business with this person.
- **Politically exposed persons lists**. This checks for any concerning links to political figures.
- **Text scraping of media mentions**, to look out for negative coverage that could be a red flag.
- **Other online interactions**, mentions on social media, etc. Again, this can be a predictor for fraudulent or erratic repayment behavior.



Chapter 3:

Integrating data into your KYC processes

Now that you've tracked down the data you need, the next step is to integrate this into your KYC processes. This will help you to figure out whether the person requesting your services is who they say they are. Whether they have skeletons in their closet you're best off avoiding — or if they're a real person at all.

That means you need to combine the new data you've acquired with the basic information the customer has provided to generate a complete picture. This data enrichment helps you to build an increasingly accurate customer profile, ensuring you aren't caught out.

Let's say your prospective customer has provided you with a name, address and email address. Data enrichment tools let you sift through those external datasets, lifting out all other addresses, linked domains, businesses, or alternate emails/spellings that are associated with the initial email address, name, or residential address. Even better, data enrichment tracks down and adds new data based on the data you

already have. For example, based on a customer's registered address, it might add relevant information from census data, rental statistics, or average home prices. It might even be able to figure out if their IP address has ever been blacklisted.

This helps you to quickly identify everything that's known to be connected to this customer's details. It means you can broaden your search and spot any red flags in their online presence or published business dealings. It also means that the system can detect anything in this applicant's online or financial behavior that doesn't fit with the other information they have provided.



Chapter 4:

Making it swift and scalable

KYC is essential, but it can undermine the user experience if you don't handle it with care. You need to access the right data to make accurate decisions, wherever that data is stored. But you also need to get the results you want fast enough to avoid friction and delays.

Finding the data you need is typically a time and resource-intensive challenge. To make the process streamlined and scalable, you need to ditch manual processes wherever you can and get AI on your side. AI helps you by sifting through thousands of data sources in a matter of minutes, cutting out a lot of those headaches while actually leading you to more types of valuable data at the same time.

To start with, augmented data discovery is a powerful tool for searching and combining data from thousands of pre-vetted sources. This allows

you to enrich your data set and generate new, relevant features quickly, making it easier and more accurate to establish predictors for red flags in their history.

Why AI might be the answer

For example, let's say that you're a bank trying to decide on whether to provide a loan. Simply knowing the applicant's current income doesn't tell you enough about the likelihood that they will be able to make repayments. A person renting the cheapest apartment they can find in New York, for example, will probably find their salary stretched further than if they rent a typical house in Detroit. Localized house price and rental data can give you a lot more context. At the same time, demographic data on average incomes may provide more nuance. For example, let's say a person applies for a mortgage with a base salary of \$70,000. Their earning potential and thus their ability to repay will be very different depending on the actual job role. A radiation therapist at the peak of their career is unlikely to earn more than \$70,000 in the future. A first-year analyst at an investment bank on Wall Street will probably double this figure in bonuses. Augmented data discovery can help tease out the connections between datasets like these.

Secondly, you will need machine learning-backed predictive models to assess, rapidly, whether an applicant poses a risk. These models use complex algorithms to establish hidden points of connection between vast datasets, drawing out trends and patterns that you could never have detected or predicted if doing this by hand.

With a powerful model, driven by the right external data, you'll be able to decide on an applicant in a fraction of the time. This not only reduces risk and speeds up onboarding time from your perspective, but it also delivers a far better experience for the customer.

Final thoughts: Getting to the data you can trust

When designing the perfect KYC process for your financial services business, you need to balance ease and speed with accuracy. You can't afford to make mistakes and let fraudsters and criminals slip through the net. Equally, you can't afford to put valuable customers through such an exhausting process, that they abandon ship or are put off from using your services in the future.

This means that you need two things:

1. Access to accurate, up-to-date, relevant external data that you can absolutely depend on - to help you make the right decision
2. A way to build predictive models and feed them with your high-quality data, efficiently - to make your customer happy.

With the right platform, you can cover both these requirements in one place. Even better, you'll be able to automate your connections to external data, as well as data cleaning and harmonizing tasks, ensuring a smooth and frictionless process from start to finish.



About Explorium

Explorium offers a first of its kind data science platform powered by augmented data discovery and feature engineering. By automatically connecting to thousands of external data sources and leveraging machine learning to distill the most impactful signals, the Explorium platform empowers data scientists and business leaders to drive decision-making by eliminating the barrier to acquire the right data and enabling superior predictive power.

**For more information,
visit www.explorium.ai**