

Breaking the  
cycle: improving  
and maintaining  
customer data



**EY**

Building a better  
working world

# INTRODUCTION

If there is one issue that keeps financial crime compliance executives up at night, it is the quality of their customer data.

Almost every financial institution struggles to keep customer data fresh, accurate, complete, and available for downstream anti-money laundering (AML) controls. The ugly but universal truth is that data maintenance is challenging and expensive, causing it to be deprioritized until it becomes unavoidable. Even in organizations that take a proactive approach to data management, the opportunities for breakdown are many: ambiguous data policy requirements, disparate and contradicting systems of record, poor data taxonomies, irreconcilable data sources, large-scale customer acquisitions, off-kilter Know Your Customer (KYC) refresh cycles and insufficient change management controls, to name a few. What's more, bad data begets bad business outcomes: poor customer insights, poor AML risk decisions and poor technological adaptability. The end result is a financial services industry hamstrung by a shared pain point with endless complexities and no clear path to resolution.

Fortunately, this doesn't have to be the case. By applying a strategic data management approach, financial institutions can improve both customer data and the maintenance regime that governs it. Designing a rightsized data remediation solution tailored to a firm's business model, customer base and compliance environment can help teams avoid common pitfalls when collecting customer data to enhance quality. Likewise, uplifting legacy data maintenance processes and controls can help firms sustain data integrity for years to come. Pairing these efforts, in turn, can break the historic cycle of customer data remediation and subsequent degradation, transforming data into a tool for business growth.





## Don't pass go: defining a sustainable data governance model

Before beginning a remediation journey, organizations must align on an internal data governance framework to manage the collection, ownership, lineage, quality, testing, security and administration of customer data.

Not only is remediating into a target model much easier than designing a model post-remediation, but it also means that principle guides the final framework rather than convenience. Moreover, applying target-state controls to a pre-remediation data set can point out gaps, disconnects and shortcomings, which can then inform the final shape of the remediation journey. For example, if forthcoming AML standards require the capture of customer postal codes to help

disposition negative news screening alerts based on location, organizations should develop corresponding controls to identify instances in which postal code data is unavailable or unretrievable from underlying data sources. Deploying these controls will not only demonstrate risks in the broader control environment but also highlight variances in data quality ripe for remediation.



# Choose your own adventure: common remediation journey types

The next step to improving customer data is to remediate existing data gaps. Traditionally, the remediation of legacy customer information – typically via KYC refresh – has been resource-intensive and manual in nature: large outreach campaigns, armies of seconded analysts to validate customer inputs, and significant potential for human error. As a consequence, costs run high and customer experience slips. While customer-facing data remediations will always remain a staple of the industry, the proliferation of personal data in recent years and the digitization of the workplace have enabled firms to pursue alternative remediation capabilities. These days, leading institutions typically employ some combination of three main remediation strategies, outlined below:

- ▶ **Internal remediation:** For many, poor data quality is not so much a symptom of incomplete customer information as a side effect of disparate or conflicting internal data sources. Data may be siloed within business lines or collected in different systems at different points in time for different purposes under different requirements or scenarios. In these cases, organizations may opt to perform internal reconciliations that seek to enhance data quality by matching data elements across

sources or “cleaning” data by resolving discrepancies or conflicts through manual review and investigation.

- ▶ **Third-party remediation:** If customer data cannot be sourced internally, organizations will need to look externally. Private third-party data providers can scrape hundreds of data sources around the globe to deliver reputable customer data insights via application programming interface (API) pulls. Similarly, as world governments step up efforts to deter financial crime, many are developing national business registries that collect data under penalty of law. Private and public data sources alike can be used to supplement internal data sets or enrich missing customer information in lieu of customer outreach.

- ▶ **Customer-sourced remediation:** Of course, the most direct (and common) way to source customer data is to contact the customer. Periodically requesting personal or business information from customers (via KYC reviews) is a routine regulatory practice and a primary avenue for customer data collection. To reduce the likelihood of a customer submitting false information, customer-provided data is typically validated, either via evidentiary documentation or third-

party verification. Given the potential impact to customer experience, many institutions seek to target remediation efforts by triaging the highest-priority data elements or customer subsets. For example, data points like full name, address and date of birth might be more important than annual income or professional industry from an operational, business or compliance perspective; as a result, customer outreach might only ask for these basic fields and rely on alternative methods to source secondary characteristics to limit customer friction.

Organizations must understand their business model, data geography and remediation objectives when determining their optimal remediation strategy. Performing data lineage analyses beforehand can confirm data is flowing from source to application properly and highlight gaps or discrepancies; authorizing pilot programs and fact-finding efforts can help inform feasibility and efficacy; and developing forecasting models can help with anticipating and managing costs. Selecting a fit-for-purpose remediation approach lays the groundwork for the project’s overall success and should be the starting point for every organization pursuing customer data improvements.

# Sticking the landing:

## leading remediation practices



After selecting the appropriate remediation strategy, the second step is execution. Remediation paths can and should be tailored to an organization's needs and objectives, but a few key considerations can increase the functionality of each:

### Internal remediations

- ▶ **Engage data owners:** Data owners are priority stakeholders in any remediation project. They understand the location, accessibility and quality of the data in question, not to mention the architecture that maintains it. Engaging these individuals early and often pays dividends in understanding the tactical pathways available to an internal remediation project. Can distinct systems of record be merged to enrich data sets in an automated way? Does the data lineage allow for connectivity between applications, or does reconciliation have to be done manually? What are the unique identifiers shared between disparate databases that can be used to uplift one source or another? Answering questions like these will inform the final shape of the remediation.
- ▶ **Clean and scrub:** Depending on an organization's data architecture, the most effective approach for an internal remediation might be cleaning, reconciliation or both. Data should first be located, merged and standardized with a common identifier. Entries should then be de-duplicated; missing, incomplete or inaccurate values should be added, modified or deleted using an exogenous data set. Contradictions should be resolved through investigation, potentially via external desktop research at approved sources, or reconciliation within a defined data hierarchy. While some components of an internal remediation can be automated, most will retain a large manual effort, given the nature of the work. Project leaders should plan accordingly.
- ▶ **Develop hierarchies:** Conflicting data sets can throw data accuracy into question. Database A says the customer resides in Miami, whereas Database B shows



“

Multiple factors come into play when choosing a data provider — cost, reputability, limitations, entity match rate, methodology — but institutions should be sure to consider the fit between the third-party platform and their own data architecture.

London: Which is it? In these instances, it helps to have defined data hierarchies that can automatically reconcile such discrepancies. Qualitative assessments of databases – methods of enrichment, frequency of updates and underlying data suppliers – can help differentiate between competing values and better enable internal remediations.

### **Third-party remediations**

- ▶ **Assess sources:** The success of a third-party remediation is primarily driven by the quality of the external data source. Multiple factors come into play when choosing a data provider – cost, reputability, limitations, entity match rate, methodology – but institutions should be sure to consider the fit between the third-party platform and their own data architecture. Does the external data source have the requisite data elements that are in scope for the remediation? Does the third party’s design allow for the deployment of process accelerators like APIs and robotic process automation? How well does the third-party source align with internal systems of record? Assessing external data sources for compatibility and quality facilitates an effective third-party remediation.
- ▶ **Take a risk-based approach:** Despite the rapid growth of business intelligence providers in recent years, there is no gold standard for external customer data sources. Each platform has incomplete or inaccurate values, so reliance

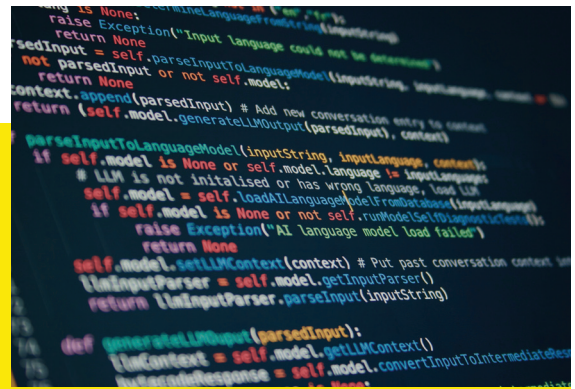
on a third-party data provider comes with inherent risk. To combat that risk, organizations should remain adaptive and pragmatic when faced with shortcomings in third-party data sets. For example, if the date of birth field is only partially available (e.g., MM/YYYY format) in the third-party system, consider applying a differentiated risk tolerance for the date of birth field to accept partial inputs to enable the enrichment of higher-priority data elements, such as name and address. Staying nimble will help the organization achieve remediation objectives in a third-party model.

### **Customer-sourced remediations**

- ▶ **Take a service-oriented approach:** It is hard to make customer outreach entirely painless, but it is easy to make it painful. The best way to cushion customer experience in a remediation is to anticipate customer needs. A comprehensive outreach strategy across multiple mediums – email, direct mail, text messaging, outbound call, in-branch, door knocking – gives customers autonomy over their remediation journey and better accommodates unhappy clients. It also improves the customer response rate by offering multiple remediation on-ramps to reach those with a strong preference for one notification medium over another.

- ▶ **Be deliberate:** Managing customer outreach should focus on minimizing disruptions to the customer experience. Where possible, institutions should seek to group accounts by parent customers to reduce duplicative outreach attempts and make the most of the customer's time. Monitoring outreach metrics, such as the number of contact attempts and their cadence, can help reduce customer friction, while targeting the right individual (i.e., authorized signer or primary account owner) increases the conversion rate.
- ▶ **Leverage technology:** The more an organization relies on technology to perform a data remediation, the less it relies on the customer. Digitizing the customer remediation experience through the creation of an online portal can help facilitate complete responses and standardize inputs; enriching or validating customer information via public or private data sources can reduce the need for re-outreach and bolster data integrity. Institutions should employ technology wherever possible in the remediation solution – from initial customer contact, to workflow management, to data verification – as a way to cushion customer experience.

# THE IMPACT OF AI



As artificial intelligence (AI) continues to enter the workplace in the coming years, it will present myriad opportunities to enhance the customer data management lifecycle. While the broad contours of a data remediation will stay the same for the foreseeable future, AI offers the chance to accelerate data capture, reduce customer friction, improve data quality, and better manage data validation. Potential use cases include:

- 1. Data extraction:** When customer data is stored in multiple documents across several onboarding systems, AI can be trained to navigate appropriate data sources and used to scrape relevant data points from disparate documents and collate for validation to reduce analyst effort. This not only facilitates internal data enrichment but also lowers customer outreaches where data is available.
- 2. Anomaly detection:** AI models can be trained to detect and flag anomalous data in real time, which reduces the need for future data remediation and enhances overall data quality at the source.
- 3. Real-time verification:** After-data capture, AI can be used to automatically validate customer inputs via third-party data sources in real time, expediting the data validation process during remediation.
- 4. Document authentication:** Similarly, AI can be taught to recognize and authenticate common customer document types (or even data attributes), identifying fake documents and allowing for faster document verification.

# Changing trajectories: maintaining data post-remediation

The final step to sustainably overhauling customer data is to implement processes and controls that support the retention of high-quality data.

---

These strategies require organizational alignment and cross-departmental coordination. Institutions should make a point to engage stakeholders (data owners, process owners, business line heads, etc.) during the data uplift, if not sooner, to agree on the post-remediation data governance regime. In particular, they should consider the following:

- ▶ **Continuous validation:** At the end of a remediation project, organizations should perform a series of controls to validate the quality, integrity, utility and completeness of the new data. While it might be tempting to stop there, one of the best ways to maintain data post-remediation is to continue reapplying those same validation checks on an ongoing basis. Periodically reassessing data entry points, mapping the flow of information between systems, retesting for missing or null values and reconfirming alignment of data to business needs are all effective methods to preempt data degradation. Deploying a governance regime in which process owners routinely review and attest to the health of their data provides an opportunity to monitor quality after a remediation journey.
- ▶ **Define a hierarchy for internal systems of record:** To reduce future data conflicts and simplify the data reconciliation process, organizations should interrogate the design, integrity and coverage of their systems of record to create a hierarchy. Establishing a waterfall model (e.g., primary reliance is placed on System A; if System A does not have the requisite data, turn to System B) clarifies data quality and reduces the need for periodic cleaning.
- ▶ **Standardize requirements:** Too often, poor data quality is the result of conflicting data requirements issued at separate times for disparate purposes. Before a remediation, future-state operating requirements should be issued by business and compliance teams and agreed upon by all downstream data users to inform the remediation scope and objectives. Post-remediation, business and compliance leads should enforce requirements with a view to maintaining quality and consistency.
- ▶ **Uplift and stress-test policies and procedures:** Once requirements have been aligned across data consumers and documented in operational procedures, compliance leaders should periodically stress-test their efficacy. Do business needs lead to a large number of exceptions to data requirements, resulting in data deterioration? Are requirements consistent with applicable laws and regulations? Is the data governance framework still rightsized? Reassessing the fit and purpose of the requirements regime is essential to preserving data integrity.



# Conclusion

When customer data is reliable and up to date, it opens up endless business opportunities. Getting there can be challenging, but a fit-for-purpose remediation solution and tailored data governance transformation go a long way toward making it a reality. Executives no longer need to shudder at the thought of their customer data quality: The road to improvement has been paved.



## CONTACT US



**Don Johnson**  
Partner  
Ernst & Young LLP  
[donald.johnson@ey.com](mailto:donald.johnson@ey.com)



**Vidhya Sekhar**  
Partner  
Ernst & Young LLP  
[vidhya.sekhar@ey.com](mailto:vidhya.sekhar@ey.com)



**Alex Treuber**  
Manager  
Ernst & Young LLP  
[alex.r.treuber@ey.com](mailto:alex.r.treuber@ey.com)



**Sandeep Dasari**  
Manager  
Ernst & Young LLP  
[sandeep.dasari@ey.com](mailto:sandeep.dasari@ey.com)

## **EY | Building a better working world**

**EY exists to build a better working world, helping to create long-term value for clients, people and society and build trust in the capital markets.**

**Enabled by data and technology, diverse EY teams in over 150 countries provide trust through assurance and help clients grow, transform and operate.**

**Working across assurance, consulting, law, strategy, tax and transactions, EY teams ask better questions to find new answers for the complex issues facing our world today.**

EY refers to the global organization, and may refer to one or more, of the member firms of Ernst & Young Global Limited, each of which is a separate legal entity. Ernst & Young Global Limited, a UK company limited by guarantee, does not provide services to clients. Information about how EY collects and uses personal data and a description of the rights individuals have under data protection legislation are available via [ey.com/privacy](https://ey.com/privacy). EY member firms do not practice law where prohibited by local laws. For more information about our organization, please visit [ey.com](https://ey.com).

Ernst & Young LLP is a client-serving member firm of Ernst & Young Global Limited operating in the US.

### **What makes EY distinctive in financial services**

Over 84,000 EY professionals are dedicated to financial services, serving the banking and capital markets, insurance, and wealth and asset management sectors. We share a single focus – to build a better financial services industry, one that is stronger, fairer and more sustainable.

© 2024 Ernst & Young LLP.  
All Rights Reserved.

SCORE no. 22964-241US\_2  
2403-4489779 BDFS0  
ED None

This material has been prepared for general informational purposes only and is not intended to be relied upon as accounting, tax, legal or other professional advice. Please refer to your advisors for specific advice.

**[ey.com](https://ey.com)**