# CAPCO

## AUGMENTED REVIEW:

## AN ESSENTIAL STEP TO GETTING STARTED WITH PKYC

# INTRODUCTION

Imagine a state of KYC, done right. Where the cost and time of processing each file is drastically reduced and changes in risk are flagged immediately. Ability to understand and manage client and reputation risk? Tick. More timely, accurate data on your client base? Tick. Welcome to the world of perpetual KYC (pKYC).

Our first paper on pKYC ([link here](#)) proposed a method on how to really get to know your clients using a perpetual approach to KYC. We detailed the need to change the current approach, how pKYC is a compelling alternative, and an overview of the tools and technology required to get there.
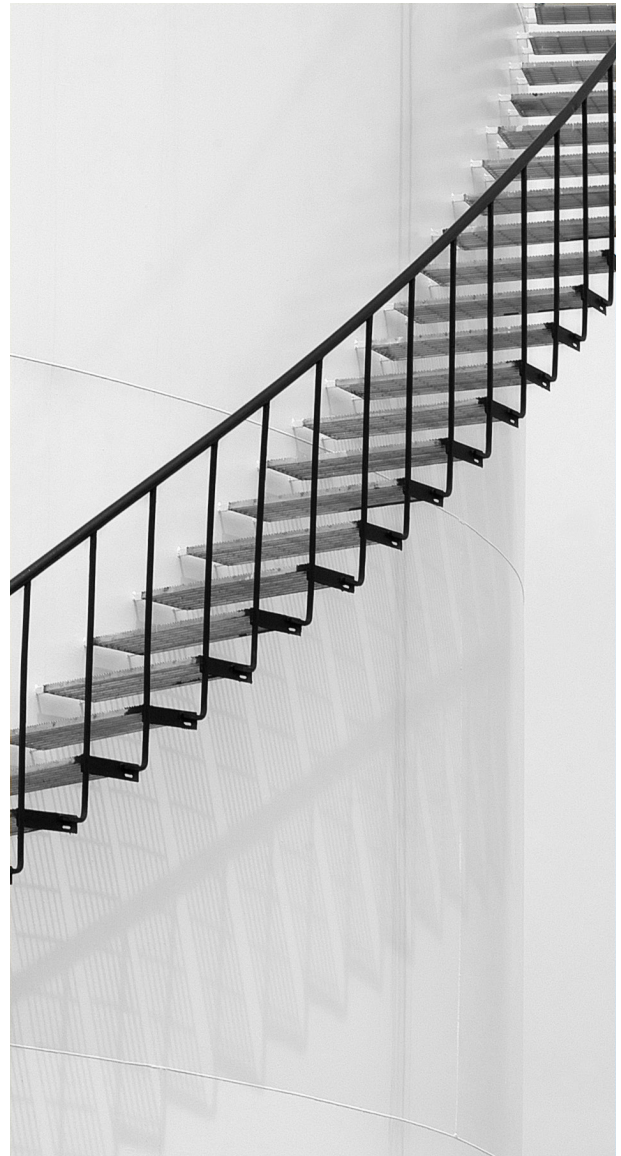
> ❝
>
> *The journey of a thousand miles begins with a single step.*
>
> **LAO TZU**
>
> ❞

**In this paper, we focus on those first steps to getting started. In particular, we focus on what we see as an essential interim step, which we have termed 'Augmented Review'.**

# SO, WHAT IS AUGMENTED REVIEW?

**Augmented Review adds API-enabled KYC sourcing and data aggregation within the periodic review process, augmenting human effort with automated research capabilities.**

| | Augmented Review | pKYC |
|---|---|---|
| **Optimised Processing & STP Autozmation** | ✓ | ✓ |
| **API-enabled KYC Research** | ✓ | ✓ |
| **Automated Multi-Source Data Aggregation** | ✓ | ✓ |
| **KYC Source Priority & Hierarchy Matrix** | ✓ | ✓ |
| **Manual (Batch Run) Material Event Review** | ✓ | |
| **Batch-Run Risk Decision Engine** | ✓ | |
| **Automated Material Event Triggered Review** | | ✓ |
| **Automated Risk Decision Engine** | | ✓ |
| **File Review Cycle Initiation** | Periodic Book of Work | Continuous Monitoring & Event Based Review |

**Figure 1:** *The main differences between augmented review and pKYC*

It is important to highlight that whilst augmented review can be seen as a steppingstone to pKYC for some, it is legitimately a target state for others. Most importantly, it is a significant diversion from current KYC processes as it enables banks better assess their client's risk and eliminates time consuming data amalgamation during identification, verification, and outreach processes.

Central to the Augmented Review process is the efficient gathering and the processing of data. At minimum, data obtained using Augmented Review needs to be able to be automatically processed and pre-populated into a single location or file for the KYC Analyst's review.

In our experience, a data strategy is key to defining the data sources and attributes for Augmented review, i.e., what data is needed, where it is, the required standards, and how to get it. This involves the following:
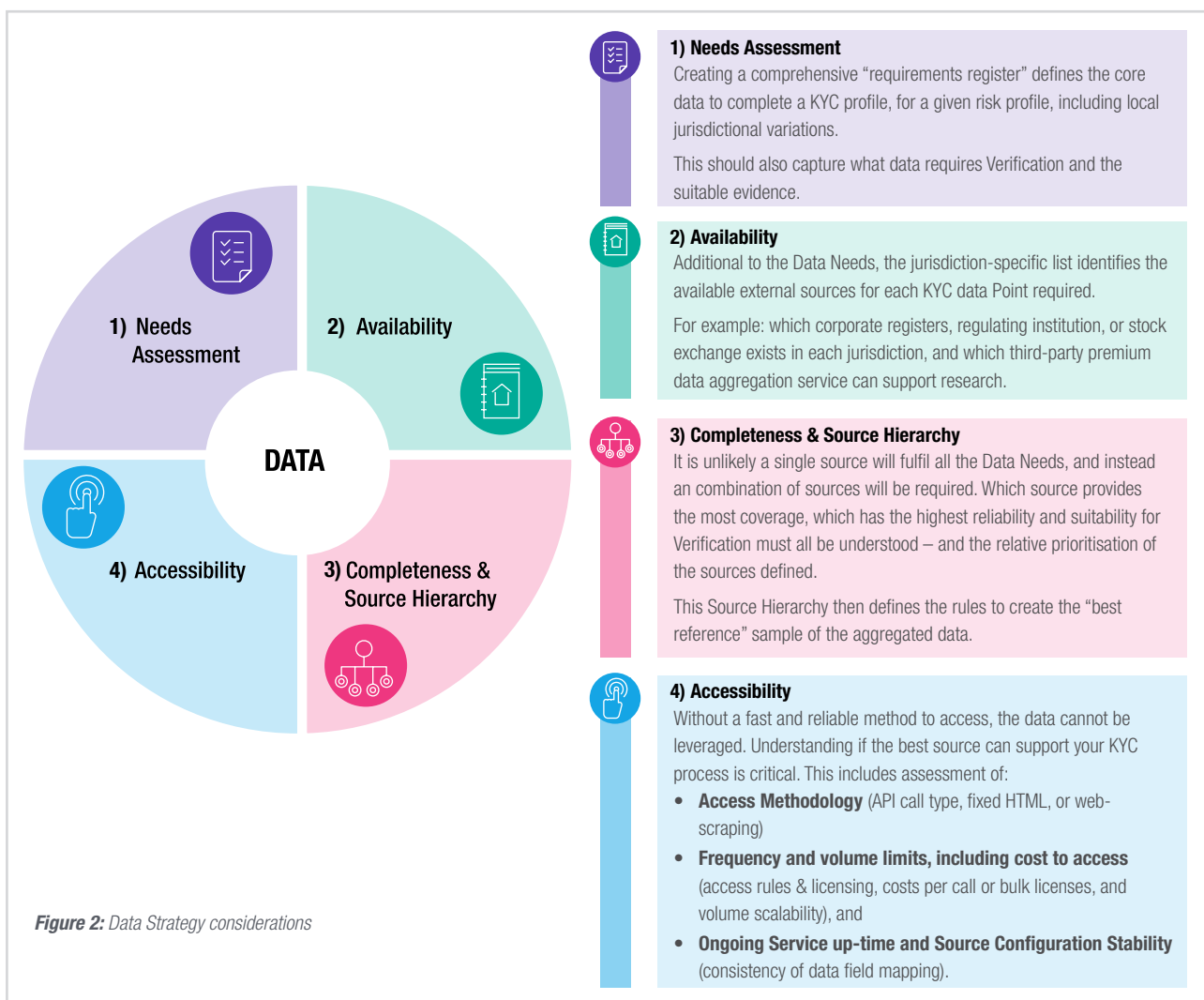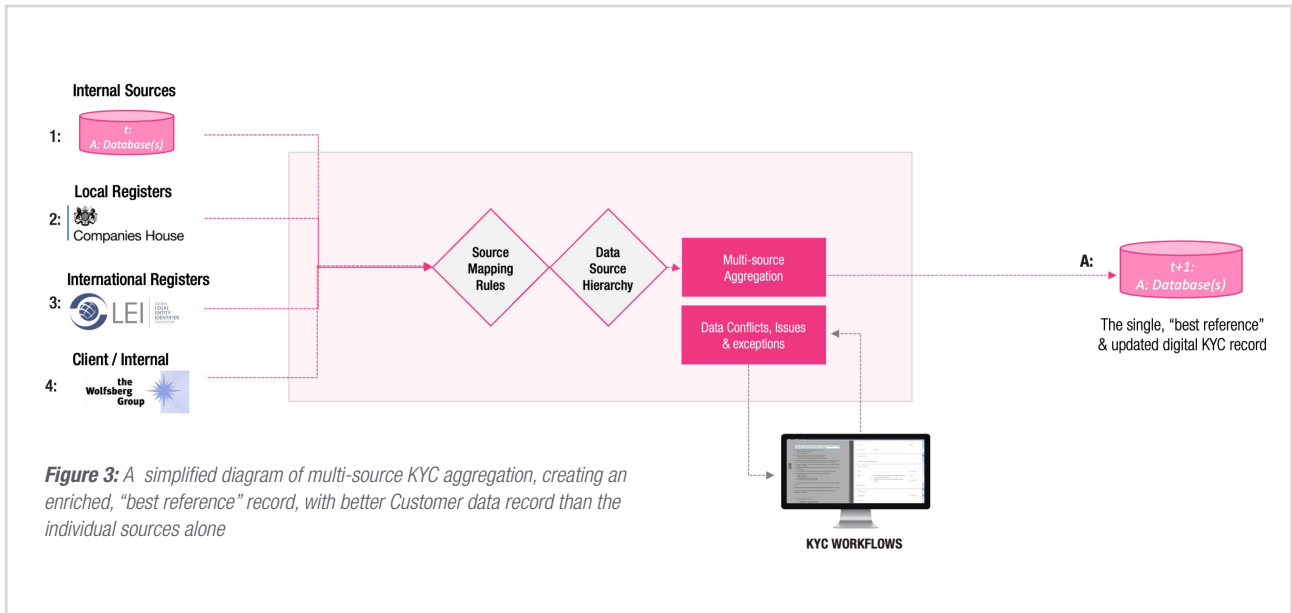


**1) Needs Assessment**
Creating a comprehensive "requirements register" defines the core data to complete a KYC profile, for a given risk profile, including local jurisdictional variations.

This should also capture what data requires Verification and the suitable evidence.

**2) Availability**
Additional to the Data Needs, the jurisdiction-specific list identifies the available external sources for each KYC data Point required.

For example: which corporate registers, regulating institution, or stock exchange exists in each jurisdiction, and which third-party premium data aggregation service can support research.

**3) Completeness & Source Hierarchy**
It is unlikely a single source will fulfil all the Data Needs, and instead an combination of sources will be required. Which source provides the most coverage, which has the highest reliability and suitability for Verification must all be understood – and the relative prioritisation of the sources defined.

This Source Hierarchy then defines the rules to create the "best reference" sample of the aggregated data.

**4) Accessibility**
Without a fast and reliable method to access, the data cannot be leveraged. Understanding if the best source can support your KYC process is critical. This includes assessment of:

- **Access Methodology** (API call type, fixed HTML, or web-scraping)
- **Frequency and volume limits, including cost to access** (access rules & licensing, costs per call or bulk licenses, and volume scalability), and
- **Ongoing Service up-time and Source Configuration Stability** (consistency of data field mapping).

**Figure 2:** *Data Strategy considerations*

## Can we build it? Should we buy it? Assessing Internal Development

By its nature, multi-source KYC requires an aggregation step to create the single customer record. The process of aggregation at scale requires three broad capabilities:

1. **Cross-Source Data (Key Name) Mapping:** for example, recognizing that *Legal.Name, Entity_Name, Legal.Entity* refer to the same attribute

2. **Data Source Preference & Hierarchy:** for example, primary sources typically take precedence over secondary sources

3. **Aggregation Tooling:** the mechanism able to process the above rules automatically and at scale.

**Figure 3:** *A simplified diagram of multi-source KYC aggregation, creating an enriched, "best reference" record, with better Customer data record than the individual sources alone*

To weigh which data aggregation solution to pursue – firms will need to understand their existing technology architecture, the available market solutions, and their own in-house development capabilities. Depending on their current capabilities and needs, should the firm buy-in a ready-made but potentially more expensive solution? Or could the business develop the capability and have full control of the features?

Developing an in-house solution has benefits but also comes with challenges. Whilst this approach offers full control over functionality, due to the significant complexity of the data process changes required, in-house development may extend project timelines and functionality realisation later than using a ready-built solution.

Depending on your project constraints, it could be best to look for a short-term tactical solution initially and subsequently transition to a longer-term strategic solution to gain stakeholder buy-in quickly.

# AUGMENTED REVIEW – A MINIMUM VIABLE PRODUCT (MVP)

## Why use an MVP?

Our proposed approach to Augmented Review recognises the scale and risk-critical nature of the KYC process. Using a Minimum Viable Product (MVP), and iteratively expanding to different client groups, can better manage the delivery risk. A key benefit of this approach is that firms can fully demonstrate the functionality with a tightly parameterised control group (as a proof of concept) and have fewer risk parameters and variables to account for.

It is worth noting that many of the below steps can be taken in parallel versus sequentially.

## Picking the first client group – Initial Client Cohort

To start, a first group of clients will need to be selected for your deployment.

**For smaller deploying organizations**, it is possible that your client base is relatively homogeneous, with similar KYC data needs, and smaller overall volume of clients. In this case – no specific group will clearly be a "low hanging" option, and all could be considered for the first cohort.

**For larger deploying organizations**, it is likely that several different sizes and geographies of clients all exist, with significantly different data sources and requirements – therefore an initial client cohort will need to be selected, and their risk and overall complexity will strongly influence the risk and complexity of the program.

There are several factors to consider – but generally, Customer Type and Size will most strongly influence the online KYC Data available, aka the Data Footprint:
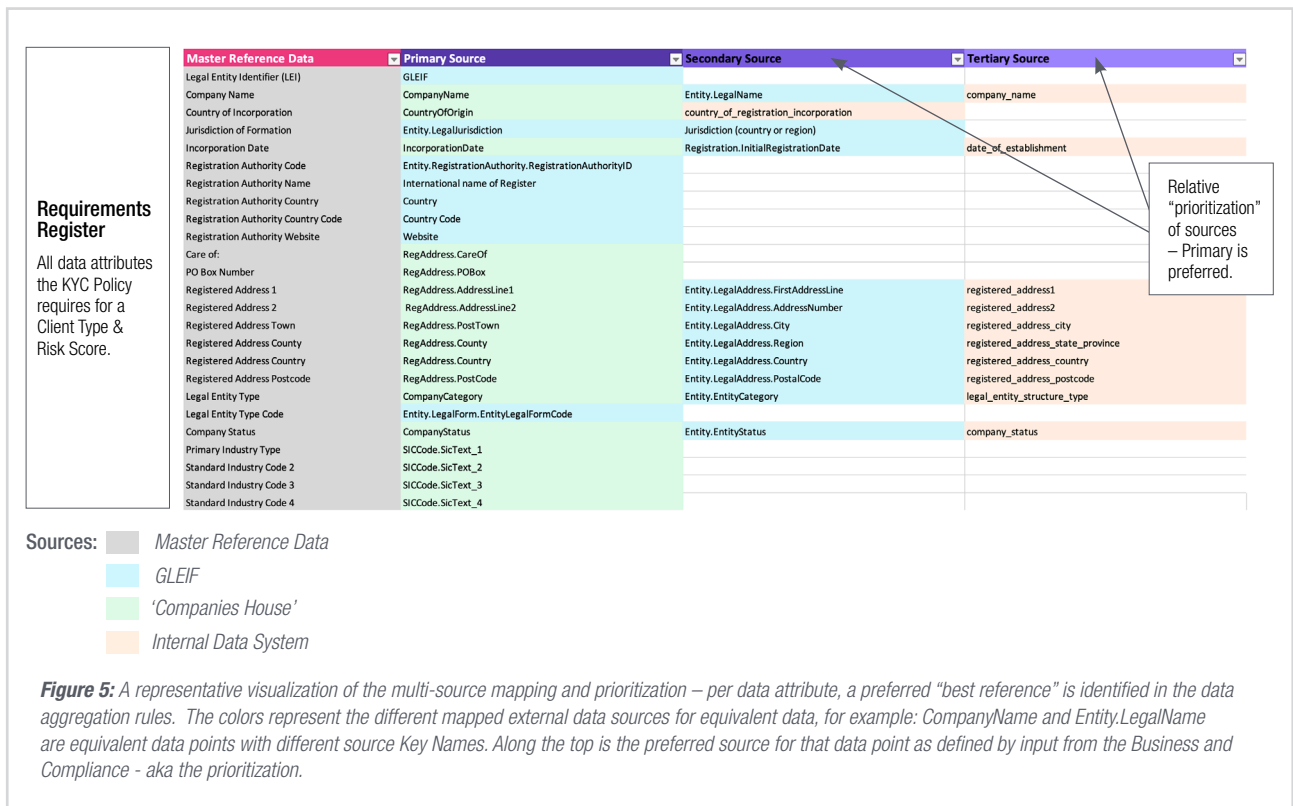


**Figure 4:** *A visual of how automated data collection complexity changes with client size*
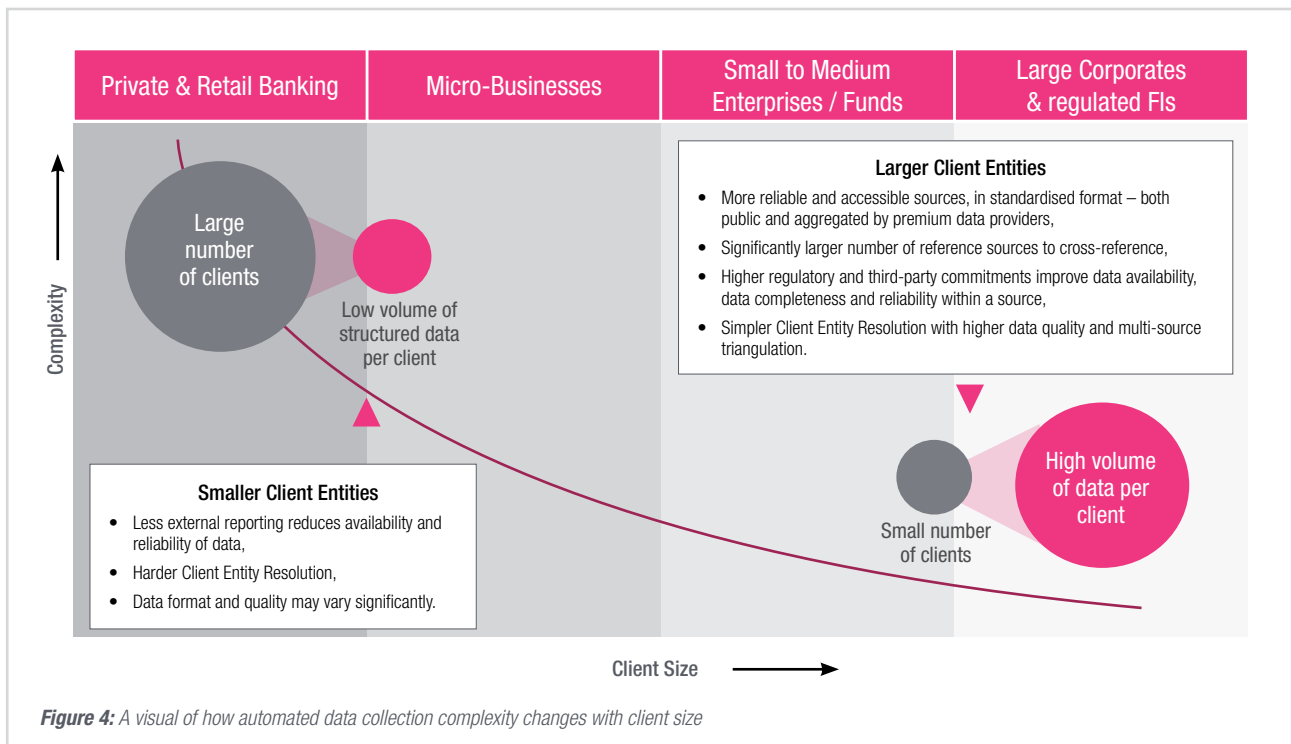
In Capco's experience, a good initial cohort should consider the following factors:

1. **Data Completeness and Availability** – better data, in fewer sources, simplifies the work and complexity to create a client profile and risk assessment. Often data completeness and availability are dependent on geographic locations of the client.

2. **Client Risk and Customer-Based Representation** – there is a trade-off when selecting a client cohort based on risk: lower risk entities can simplify delivery in having typically more straightforward data needs and lower compliance risk overall. Alternately, choosing a cohort with higher risk clients (with typically higher KYC costs) may demonstrate a better return on investment.

3. **Data Complexity** – greater complexity in the data requirements to collect, but also in the data transformation and processing steps needed for the organization to leverage, will also add complexity to the work and timelines of the MVP.

## Automated Source Match & File Pre-Population

After selecting an initial client group and completing the data strategy assessment, a Source Prioritization Matrix can be constructed. This is a KYC reference table outlining: all the data requirements, for each client type, with a prioritized list of sources for each data attribute.



**Requirements Register**

All data attributes the KYC Policy requires for a Client Type & Risk Score.

Relative "prioritization" of sources – Primary is preferred.

| Master Reference Data | Primary Source | Secondary Source | Tertiary Source |
|---|---|---|---|
| Legal Entity Identifier (LEI) | GLEIF | | |
| Company Name | CompanyName | Entity.LegalName | company_name |
| Country of Incorporation | CountryOfOrigin | country_of_registration_incorporation | |
| Jurisdiction of Formation | Entity.LegalJurisdiction | Jurisdiction (country or region) | |
| Incorporation Date | IncorporationDate | Registration.InitialRegistrationDate | date_of_establishment |
| Registration Authority Code | Entity.RegistrationAuthority.RegistrationAuthorityID | | |
| Registration Authority Name | International name of Register | | |
| Registration Authority Country | Country | | |
| Registration Authority Country Code | Country Code | | |
| Registration Authority Website | Website | | |
| Care of: | RegAddress.CareOf | | |
| PO Box Number | RegAddress.POBox | | |
| Registered Address 1 | RegAddress.AddressLine1 | Entity.LegalAddress.FirstAddressLine | registered_address1 |
| Registered Address 2 | RegAddress.AddressLine2 | Entity.LegalAddress.AddressNumber | registered_address2 |
| Registered Address Town | RegAddress.PostTown | Entity.LegalAddress.City | registered_address_city |
| Registered Address County | RegAddress.County | Entity.LegalAddress.Region | registered_address_state_province |
| Registered Address Country | RegAddress.Country | Entity.LegalAddress.Country | registered_address_country |
| Registered Address Postcode | RegAddress.PostCode | Entity.LegalAddress.PostalCode | registered_address_postcode |
| Legal Entity Type | CompanyCategory | Entity.EntityCategory | legal_entity_structure_type |
| Legal Entity Type Code | Entity.LegalForm.EntityLegalFormCode | | |
| Company Status | CompanyStatus | Entity.EntityStatus | company_status |
| Primary Industry Type | SICCode.SicText_1 | | |
| Standard Industry Code 2 | SICCode.SicText_2 | | |
| Standard Industry Code 3 | SICCode.SicText_3 | | |
| Standard Industry Code 4 | SICCode.SicText_4 | | |

**Sources:**
- Master Reference Data
- GLEIF
- 'Companies House'
- Internal Data System

**Figure 5:** *A representative visualization of the multi-source mapping and prioritization – per data attribute, a preferred "best reference" is identified in the data aggregation rules. The colors represent the different mapped external data sources for equivalent data, for example: CompanyName and Entity.LegalName are equivalent data points with different source Key Names. Along the top is the preferred source for that data point as defined by input from the Business and Compliance - aka the prioritization.*

## How does it work?

The initial Source Matrix defines the data rules the data aggregator tool uses to sort and create the Best Reference available customer record:

1. APIs feeds pulls in the available data from the configured external sources
2. The Data Aggregator tool aggregates the internal and external sources into a single customer record by applying the prioritization rules, e.g., take the primary reference, or an acceptable secondary source if the primary is unavailable
3. Change Comparison: each data attribute is either updated with new data, or automatically marked as "No change" where the existing data is correct and in date
4. This "best reference" record is pre-populated into the KYC file, reducing research work effort through Augmented Review.

This process in effect automates the otherwise manual research and evaluation steps currently completed by human workers. The rest of the KYC process continues as per current state.

It is important to highlight that this matrix should be proactively updated over time to maintain the ability to obtain the "Best Reference" using observed performance accuracy.

## Business Benefits

Business benefits have varied depending on our clients' portfolios. However, it has been substantial including:

- 25% - 40% FTE time saves from implementing Augmented Review
- Coverage of 75% of their client footprint using external data feeds
- STP for KYC achieved for 'no change' clients
- Backing conversations to drive forward the agenda towards pKYC

## Ongoing Quality Assurance and Data Confidence – Continuous Improvement

After an Augmented Review, the KYC Files will pass through the typical Quality Assurance review workflows. However, by assessing the accuracy and validity of the key facts collected through Augmented Review, the findings of the QA review can also be used to:

- Assess the accuracy and reliability of the Augmented Review overall, and
- Also provide QA accuracy scores for individual sources of the data.

Using this feedback mechanism, evaluating the data source on a per data attribute basis, the Source Matrix, and each source's prioritization, can be continuously improved over time.

This quality assurance data is also critical for stakeholder reporting and for enabling pKYC should that be the strategic direction.

## Expanding from the initial client cohort

After conclusion of testing and sign off process for the initial MVP – having used the ongoing QA assessment of the processed files, and comparing the enhancement effectiveness, the second cohort can be considered.

As with the first Cohort assessment, the same Data Footprint considerations apply, and reference should be made to your business' specific needs and drivers of change to help identify the next best benefits realization for the effort.

# WHAT NEXT?

Augmented Review is not just an enabler for pKYC but can bring massive benefits and efficiencies to the current KYC process, lowering costs and manual touchpoints.

Capco has implemented Augmented Review solutions at tier 1 clients. Connect with us to see how we can support you on your journey.
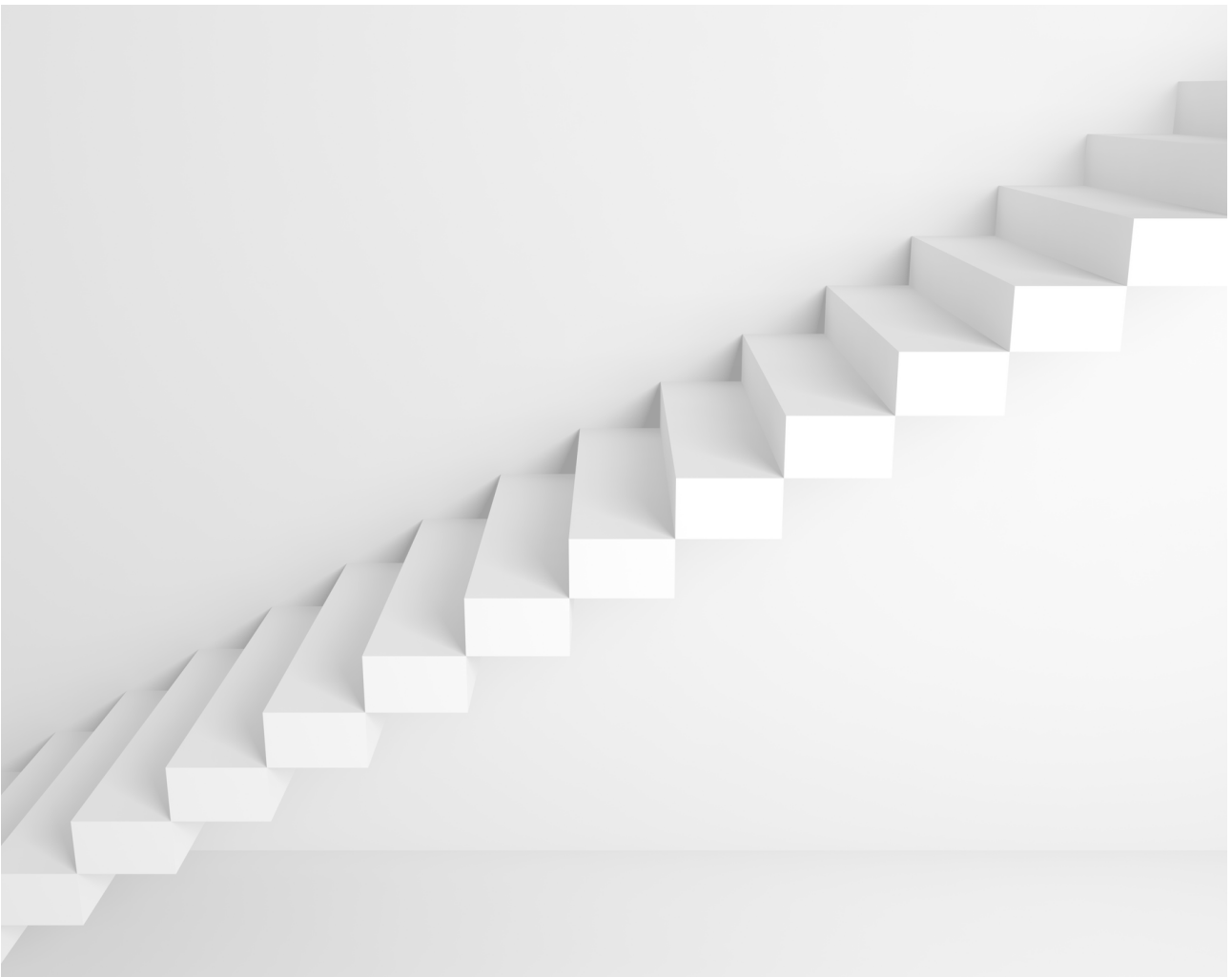
## AUTHORS

**Alex Saunders,** Financial Crime Risk SME
**Edward Pease,** Data Science Consultant
**Riddhi Sen,** Head of Data Science & Machine Learning

## CONTACTS

**Chris Probert**, Partner & Head of Data, chris.probert@capco.com
**Riddhi Sen**, Head of Data Science & Machine Learning, riddhi.sen@capco.com

---

## ABOUT CAPCO

Capco is a global technology and management consultancy dedicated to the financial services industry. Our professionals combine innovative thinking with unrivalled industry knowledge to offer our clients consulting expertise, complex technology and package integration, transformation delivery, and managed services, to move their organizations forward.

Through our collaborative and efficient approach, we help our clients successfully innovate, increase revenue, manage risk and regulatory change, reduce costs, and enhance controls. We specialize primarily in banking, capital markets, wealth and asset management and insurance. We also have an energy consulting practice in the US. We serve our clients from offices in leading financial centers across the Americas, Europe, and Asia Pacific.

To learn more, visit our web site at www.capco.com, or follow us on Twitter, Facebook, YouTube, LinkedIn and Instagram.

## WORLDWIDE OFFICES

| APAC | EUROPE | NORTH AMERICA |
|---|---|---|
| Bangalore | Berlin | Charlotte |
| Bangkok | Bratislava | Chicago |
| Gurgaon | Brussels | Dallas |
| Hong Kong | Dusseldorf | Hartford |
| Kuala Lumpur | Edinburgh | Houston |
| Mumbai | Frankfurt | New York |
| Pune | Geneva | Orlando |
| Singapore | London | Toronto |
| | Munich | Tysons Corner |
| | Paris | Washington, DC |
| | Vienna | |
| | Warsaw | **SOUTH AMERICA** |
| | Zurich | São Paulo |

## WWW.CAPCO.COM

**CAPCO**