Improving KYC data acquisition: *A strategic data exchange model*

Capgemini invent

Know your customer (KYC) has traditionally been a highly document-driven process, requiring corporations to submit extensive paperwork to multiple financial institutions. However, with changing technology and regulatory expectations and the need for greater efficiency, the industry is leaning towards a data-driven model. In this new paradigm, real-time, structured, and validated data plays a critical role in enhancing compliance, reducing manual intervention, and improving operational efficiency.

To obtain KYC data, corporate investment banks (CIBs) have two approaches. First, they can manually connect with multiple registries, regulatory bodies, and exchange websites to collect data. This is a laborious and inconsistent process that demands significant resources. Or an alternate method is to utilize premium data aggregators like Moody's, D&B, Markit, Refinitiv, etc., which offer **70 to 80 percent of the KYC requirements** from verified sources in a digestible format. By leveraging this approach, the need for direct data extraction from numerous sources is reduced significantly. This streamlines compliance efforts and improves data accuracy.

At the same time, companies maintain **11 banking relationships on average** and must provide at least **10 key documents per bank** for compliance. This results in **100 to 200 KYC document requests annually**, causing inefficiencies, duplication of effort, and compliance fatigue.

This paper introduces a **strategic data exchange model** that is built on a proof-of-work mechanism, differentiating it from blockchain-based KYC models. In this new model, the client provides reliable data to the provider in case of any gaps, leveraging the CIB's infrastructure and in turn earns incentive from the CIB. The CIB earns credit from the data provider, and the data provider benefits from reduced long-term data acquisition costs. In contrast to blockchain-based KYC models, which frequently struggle due to fragmented regulatory acceptance, the inability to support multi-jurisdictional compliance, and lack of standardized verification, this model provides a centralized but scalable approach.

By implementing this model:

- **CIBs** lower operational costs, improve KYC compliance, and reduce manual verification efforts.
- **Companies** reduce KYC fatigue by minimizing redundant document submissions, leading to improved efficiency in banking relationships and they earn incentive for supplying information.
- **Data aggregators** enrich datasets with direct corporate input, enhancing their value proposition and market competitiveness and reducing the cost of data acquisition.

With the growing complexity of regulatory requirements and the rising cost of compliance, CIBs must adopt innovative, collaborative models to enhance efficiency. This data exchange framework presents a scalable, cost-effective solution that strengthens the financial ecosystem while positioning participating banks at the forefront of digital transformation.

The challenges: *High cost and fragmentation*

Global premium data providers, on average, invest an estimated $50 million to $100 million annually in acquiring and maintaining high-quality corporate data. Even with such investments, they still face data gaps, requiring their clients to conduct manual verification.

At the same time, cumbersome KYC experiences diminish overall client satisfaction. Research indicates that companies spend an average of **1,500 hours annually** responding to KYC requests, leading to inefficiencies and delays in business transactions.

Exploring alternatives: *KYC utility for blockchain solutions*

Blockchain is considered a potential solution for simplifying KYC processes. By enabling decentralized, tamper-proof records, blockchain seeks to establish a single, reusable KYC profile that financial institutions can securely access. This approach reduces duplication of effort and enhances transparency by ensuring that, once verified, KYC data remains unchangeable and traceable. Additionally, blockchain promises to provide greater security by using cryptographic verification, which reduces the possibility of identity fraud and unauthorized data access.

Nevertheless, despite its potential, blockchain-based KYC models present several drawbacks.

- **Regulatory fragmentation:**
  Different jurisdictions have varying regulations, and decentralized KYC models struggle with cross-border compliance due to the lack of standardized legal frameworks.
- **Absence of a unified trust mechanism:**
  Without a central verification authority, banks hesitate to rely on third-party validated data, reducing the credibility of blockchain-based KYC records.
- **High infrastructure costs:**
  Permissioned blockchain models require substantial investment in infrastructure, security, and maintenance, making large-scale adoption costly.
- **Data immutability vs. privacy regulations:**
  Blockchain's immutability conflicts with data protection laws like GDPR, which mandate the ability to modify or delete user data upon request.

A cost-effective alternative: *The strategic data exchange model*

Unlike traditional blockchain-based KYC models, which often struggle with scalability and compliance, our strategic data exchange model offers a more practical and cost-effective approach. Instead of relying on distributed ledgers, this model enables a structured and incentive-driven data-sharing framework. (Similar models have been successfully implemented in various industries, demonstrating their effectiveness and practicality.)

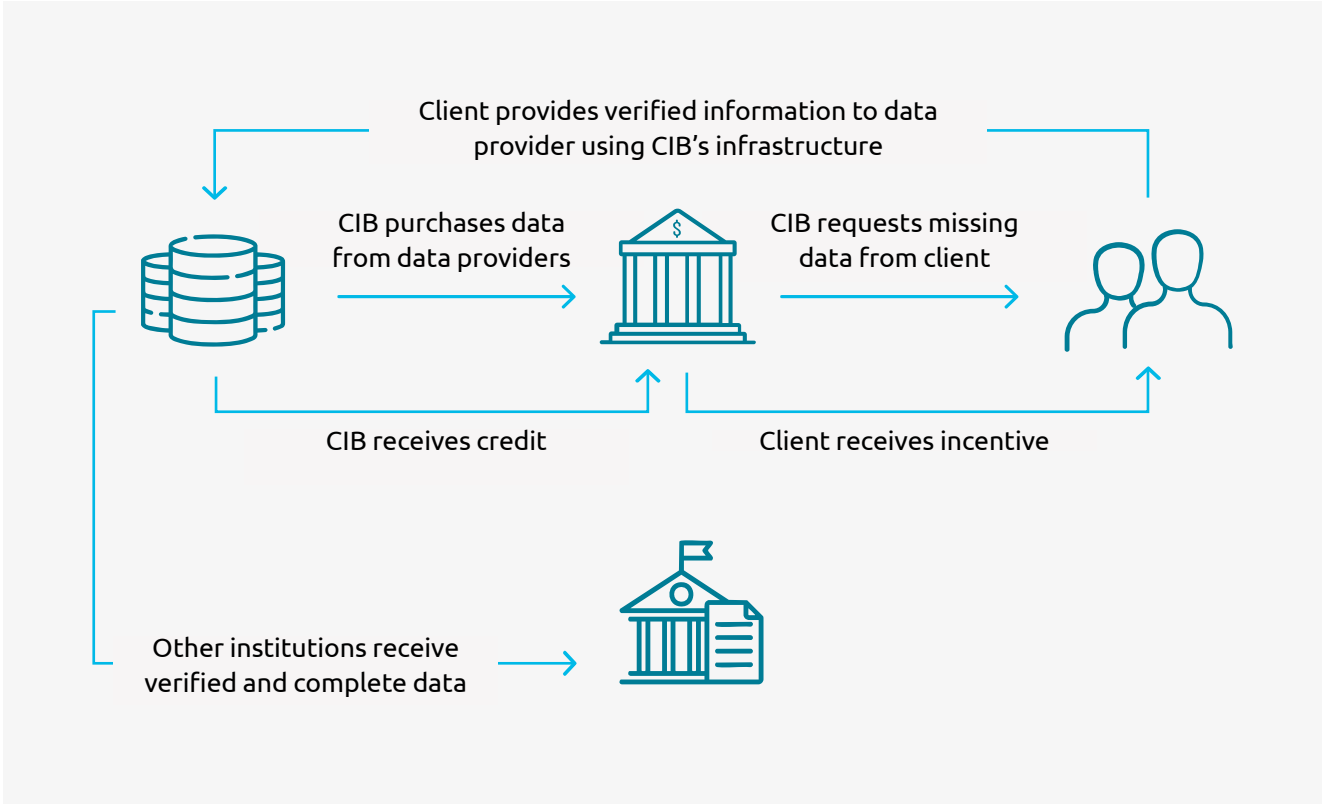
How it works

- **Direct data procurement:**
  The CIB establishes an exclusive data acquisition agreement with a premium vendor, optimizing costs through bulk purchasing.

- **Identifying data gaps:**
  When gaps in KYC data arise, the bank collects the information from the client and directs their corporate clients to submit the missing information to the data provider. Banks provide the infrastructure for a centralized yet collaborative data exchange. Companies submit validated data through an API-driven mechanism, reducing redundancy and error.

- **Proof of work mechanism:**
  Corporate clients contribute validated data to enhance the aggregator's repository and earn incentive from the bank. In return, the bank earns credits from the vendor, which can be used to acquire enhanced datasets.

- **Data quality enhancement:**
  As more companies participate, the data provider's repository becomes increasingly comprehensive, reducing future gaps and streamlining compliance across the ecosystem.



Comparative analysis of a blockchain-based decentralized KYC solution and the proof-of-work inspired strategic data exchange model data-sharing mechanism

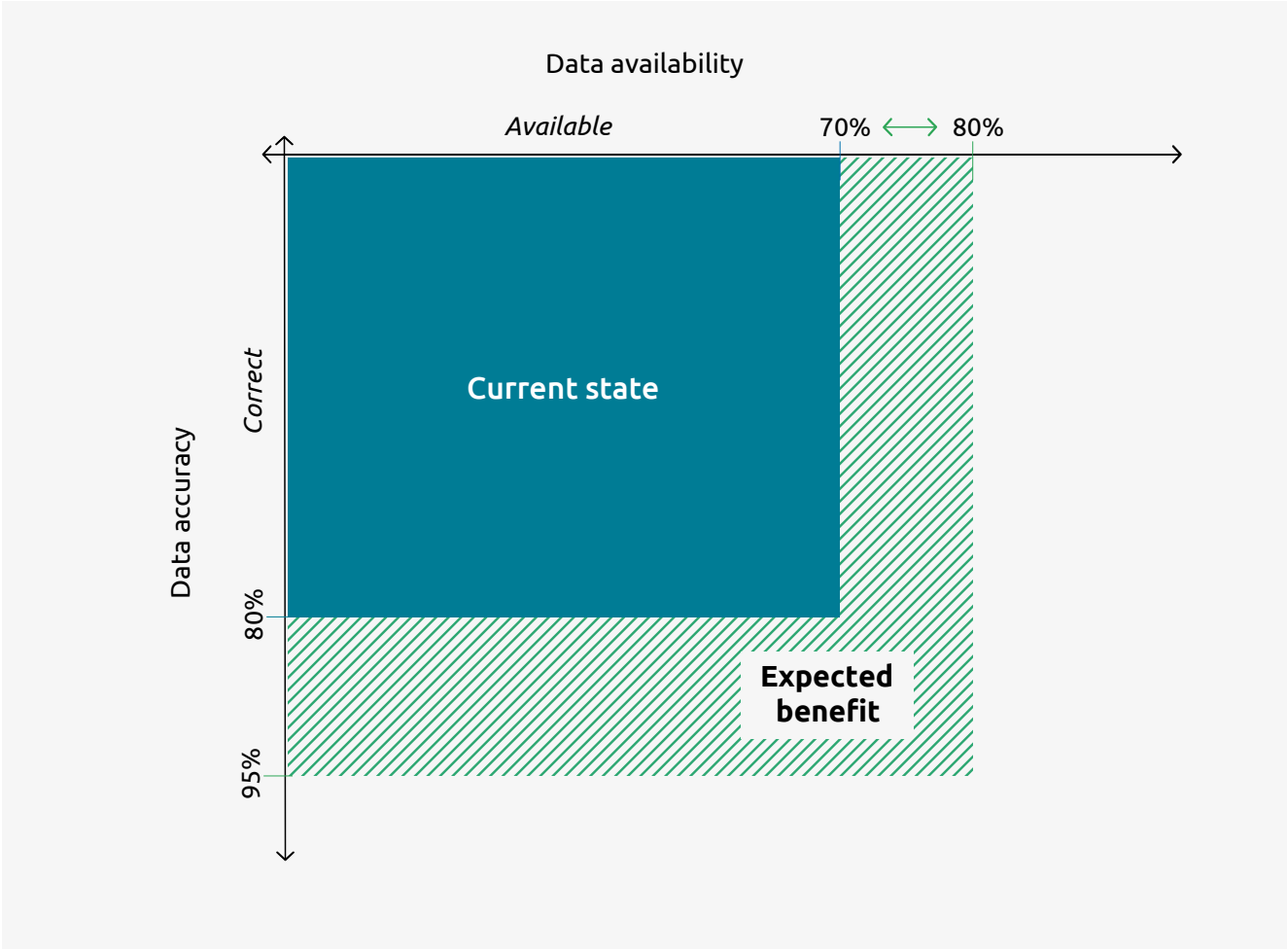
Feature	*Blockchain model*	*Strategic data exchange model*
Approach	Decentralized KYC ledger on blockchain	Centralized infrastructure with verified data aggregation. Not just utility but clients actively contribute validated data
Data control	Customer-owned, stored on a blockchain ledger	Bank-controlled, with corporate consent-driven submissions
Compliance	Struggles with regulatory adoption across jurisdictions	Designed for jurisdictional compliance and regulatory alignment
Data verification	Peer-to-peer sharing, no single validation authority	Verified by premium data aggregators and banks
Cost	High due to blockchain infrastructure and redundancy	Lower costs via incentive-based data enrichment
Adoption challenges	Banks hesitant due to reliance on third-party blockchain	Easier adoption as it leverages existing regulatory frameworks
Interbank data sharing	Enables reusability of KYC data across banks	No direct interbank sharing, but updates flow through providers
Operational complexity	Requires decentralized identity verification across multiple banks	Streamlined process where banks manage infrastructure and vendors verify data

The role of data coalitions in KYC optimization

Industry-wide collaboration through data coalitions plays a pivotal role in enhancing KYC data accessibility and standardization. By pooling verified corporate KYC data, banks and financial institutions can create a shared repository that minimizes duplication and improves data accuracy across the sector. Under a well-governed framework with proper consent structures, once a corporation submits its verified data, it becomes accessible to multiple banks – streamlining compliance efforts and reducing administrative burdens for all stakeholders. A single bank provides the necessary infrastructure, enabling other institutions to access updated data without direct interbank exchanges, thereby ensuring neutrality, preventing conflicts of interest, and reinforcing regulatory compliance. Ultimately, financial institutions, regulators, and industry players are working towards the same goal – efficient, standardized, and secure KYC processes.

Key benefits

This model will improve the availability and accuracy of data for better compliance and efficiency.



- **Cost optimization for banks:**
  Direct purchasing agreements lower data acquisition costs, and the credit-based system offsets future expenditures through vendor discounts, driving cost optimization for banks.
- **Enhanced KYC data coverage:**
  Increases data reliability through corporate-verified data, and leverages premium aggregators to avoid the complexity and cost of building individual API connections to registries, regulators, and exchanges.
- **Reduced compliance burden:**
  Companies benefit from a reduced compliance burden by avoiding repetitive KYC requests from multiple banks, and a single data submission ensures consistency while reducing reporting errors.
- **Competitive advantage and operational efficiency:**
  The model offers a competitive advantage by enhancing the bank's ability to meet regulatory requirements efficiently, while improved data quality accelerates customer onboarding and reduces compliance risks.
- **Scalability and industry-wide adoption:**
  By involving multiple banks, the model creates an industry-wide standard for KYC data collection, fostering greater adoption and regulatory acceptance.

Regulatory and legal considerations

- **Data privacy and compliance:**
  The model must comply with global data protection laws such as GDPR and CCPA. Banks and data vendors must ensure that any corporate-submitted data is used strictly for KYC purposes and is handled with explicit consent.
- **Validation and verification standards:**
  Regulators require that all KYC data be reliable and verifiable. To meet this standard, data vendors must implement robust validation mechanisms to ensure that corporate-submitted data meets regulatory requirements before being integrated into datasets.
- **Liability and accountability:**
  Clear contractual agreements must define responsibilities among banks, data vendors, and corporate participants. Banks should retain the right to dispute incorrect data while ensuring companies are not unfairly penalized for minor discrepancies.

Limitations

While the strategic data exchange model offers numerous benefits, its implementation may face limitations in certain countries due to regulatory, technological, and infrastructural challenges. For example, Singapore has PDPA, a stringent data protection law which restricts data sharing and storage. Additionally, laws may require data to be stored within national borders, complicating cross-border data integration. The model's reliance on advanced technological infrastructure can be challenging in regions with less developed systems, and achieving widespread adoption may be difficult in markets resistant to change. Cultural and operational differences also impact the willingness of companies to participate and share data, further limiting the model's effectiveness.

Overall, while the model presents a promising solution for KYC data acquisition, its implementation must be carefully tailored to address the specific regulatory, technological, and cultural contexts of each country.

Framework for implementation

Strategy and planning

- Identify premium data aggregator partners
- Define scope and regulatory requirements
- Assess internal data readiness

Data integration

- Establish API connectivity
- Implement data validation protocols
- Initiate pilot testing

Corporate onboarding

- Educate corporate clients on the benefits
- Create incentive structures
- Streamline submissions

Monitoring & optimization

- Track data quality improvements
- Measure cost savings
- Explore additional partnerships

Closing Remarks

The strategic data exchange model provides a scalable, cost-effective, and compliance-driven solution for KYC data acquisition. By fostering a collaborative ecosystem where banks, data vendors, and corporate clients actively contribute verified data, the model ensures a continuously improving and resilient financial compliance framework. Inspired by blockchain's proof-of-work principle, this approach incentivizes participants to enhance data quality while reducing redundancies and operational inefficiencies.

Regulators, with a dual incentive to maintain data integrity and enhance its completeness, play a crucial role in upholding standards and enriching the shared repository. By implementing this framework, corporate investment banks not only gain a competitive edge but also contribute to a more transparent, efficient, and self-sustaining financial ecosystem, where every participant plays a positive role in driving compliance and trust.

Authors

Hiloni Shah
Managing Consultant
hiloni-hitesh.shah@capgemini.com

Florent Palayret
Global Financial Crime Compliance Lead
florent.palayret@capgemini.com

Charles Dally
Bank Operations & Industrialization Global lead
charles.dally@capgemini.com

About Capgemini Invent

As the digital innovation, design and transformation brand of the Capgemini Group, Capgemini Invent enables CxOs to envision and shape the future of their businesses. Located in over 30 studios and more than 60 offices around the world, it comprises a 12,500+ strong team of strategists, data scientists, product and experience designers, brand experts and technologists who develop new digital services, products, experiences and business models for sustainable growth.

Capgemini Invent is an integral part of Capgemini, a global business and technology transformation partner, helping organizations to accelerate their dual transition to a digital and sustainable world, while creating tangible impact for enterprises and society. It is a responsible and diverse group of 340,000 team members in more than 50 countries. With its strong over 55-year heritage, Capgemini is trusted by its clients to unlock the value of technology to address the entire breadth of their business needs. It delivers end-to-end services and solutions leveraging strengths from strategy and design to engineering, all fueled by its market leading capabilities in AI, generative AI, cloud and data, combined with its deep industry expertise and partner ecosystem. The Group reported 2024 global revenues of €22.1 billion.

Get the future you want | www.capgemini.com/invent

Capgemini invent