



## Tackling data challenges for modernizing legacy technology platforms

October 2020



**MAKING AN  
IMPACT THAT  
MATTERS**  
*since 1845*

Financial services firms, more specifically broker-dealers and asset management firms, struggle with modernizing their legacy technology platforms. These platforms hold most, if not all, of the critical data for these institutions. Given these platforms have been built over the years, the data is fragmented across the enterprise with no coherent data strategy around curation, translation, integration, storage, quality governance, and compliance. Therefore, these institutions spend most of their transformation budget on solving data challenges and focusing very little on real business needs.

“A successful legacy technology modernization invariably depends on the ability to harness the power of data to achieve operational excellence, comply with regulatory mandates, enable valuable insights for your business, and provide a superior client experience.”

To keep up with evolving business needs, banks typically focus their technology transformation efforts across five key problem areas (“modernization barriers”).



However, the explosion of data in the digital age has compounded the complexity of technology transformation across each of these five areas.

*“IDC predicts that our global datasphere—the digital data we create, capture, replicate, and consume—will grow from approximately 40 zettabytes of data in 2019 to 175 zettabytes in 2025 (with one zettabyte equaling one trillion gigabytes).”<sup>1</sup>*

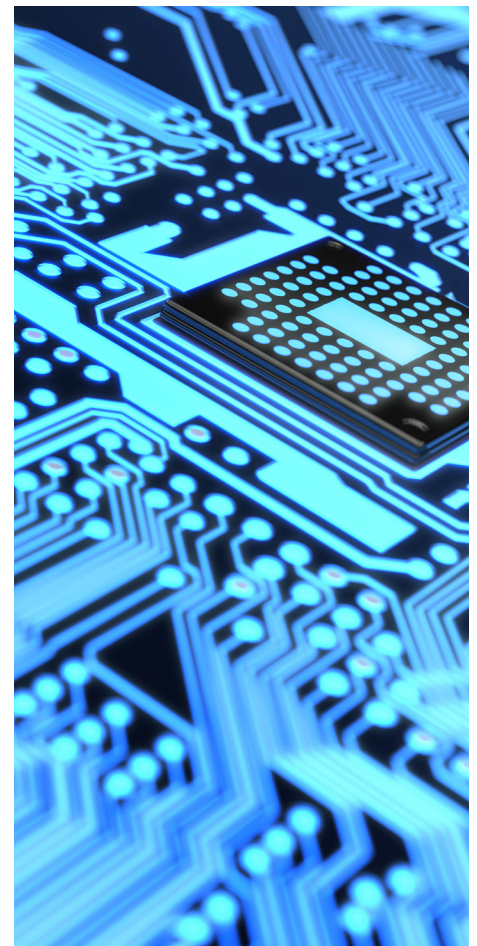
Broker-dealers and asset management firms continue to automate and digitize increasingly larger parts of their businesses, creating more dependencies on data and amplifying their existing data challenges.

These are expensive and frustrating problems for senior executives, making them wonder, “What If...”

	Legacy platforms could <b>seamlessly support</b> new business needs <b>without costly transformations?</b>
	Processes could <b>scale more efficiently</b> by <b>eliminating manual interventions</b> and workarounds?
	Legacy data could enable <b>customer analytics and AI</b> to help <b>run and grow the business faster?</b>
	Improvements in <b>data quality</b> could <b>reduce operational overhead</b> and make processes more efficient?
	Your firm could <b>achieve predictive digital controls</b> and regulatory compliance using data?

**To achieve all of this and tackle the five modernization barriers mentioned above, global firms need a well-formulated data strategy.** How firms choose to deal with these challenges can say a lot about their strategy and their continued viability in a market that is facing increasing disruption.

In this paper, we examine the barriers in detail and discuss pragmatic data initiatives to help address each of them. We also present two successful case studies involving large-scale legacy modernization at broker-dealers and asset management firms and illustrate how data challenges were successfully addressed in these transformations.



1 IDC, The Digitization of the World From Edge to Core, November 2018.

# Solving modernization through Data strategy – Illustrative case studies

Disparate data sources (e.g., multiple front-office systems) often stand in the way of transforming middle-office and back-office operations to more efficient functions. Case study 1 below illustrates how **data pipelines** and a **golden data source** (built harnessing the power of managed big data platforms on cloud, in this case) can help modernize your business without necessarily disrupting the existing on-premises or legacy source systems.

— | **Case study 1.** Disparate data sources lead to modernization challenges | —

*“How do I ensure accurate CAT regulatory reporting of all equities and options trades daily? How do I tackle the risk of underreporting and/or incorrect reporting of equities and options data?”*

 - RegOps lead, global investment bank

- Fragmented order management life cycle, resulting in certain parts of the order chain not being systematized
- Multiple data standards and lack of data lineage capabilities creating linkage issues for complex orders spanning across almost a dozen front-office systems
- Infrastructure scalability issues on-premises posing challenge to meet processing requirements (e.g., huge data volumes, strict SLA of T+1 8 a.m. ET)

 **500M+**  
**REPORTABLE RECORDS**

 **Multiple**  
**FRONT-OFFICE SYSTEMS**

 **Solution**





We leveraged our experience of cloud and big data to help the client solve their regulatory reporting issues:

- Set up a new big data pipeline on cloud with dynamic infrastructure provisioning algorithms to address market volatility, SLA requirements, and 5X higher regulatory reporting volume
- Created connectors to transfer data from on-premises front-office systems to the cloud data pipeline for daily reporting
- Created a big data repository on cloud (golden source), which validated, enriched, and linked 150+ attributes across 10+ front-office application feeds to a common staging standard
- Analyzed more than 25 critical order flows and established data lineage using advanced data tagging to help business users address regulatory requests faster

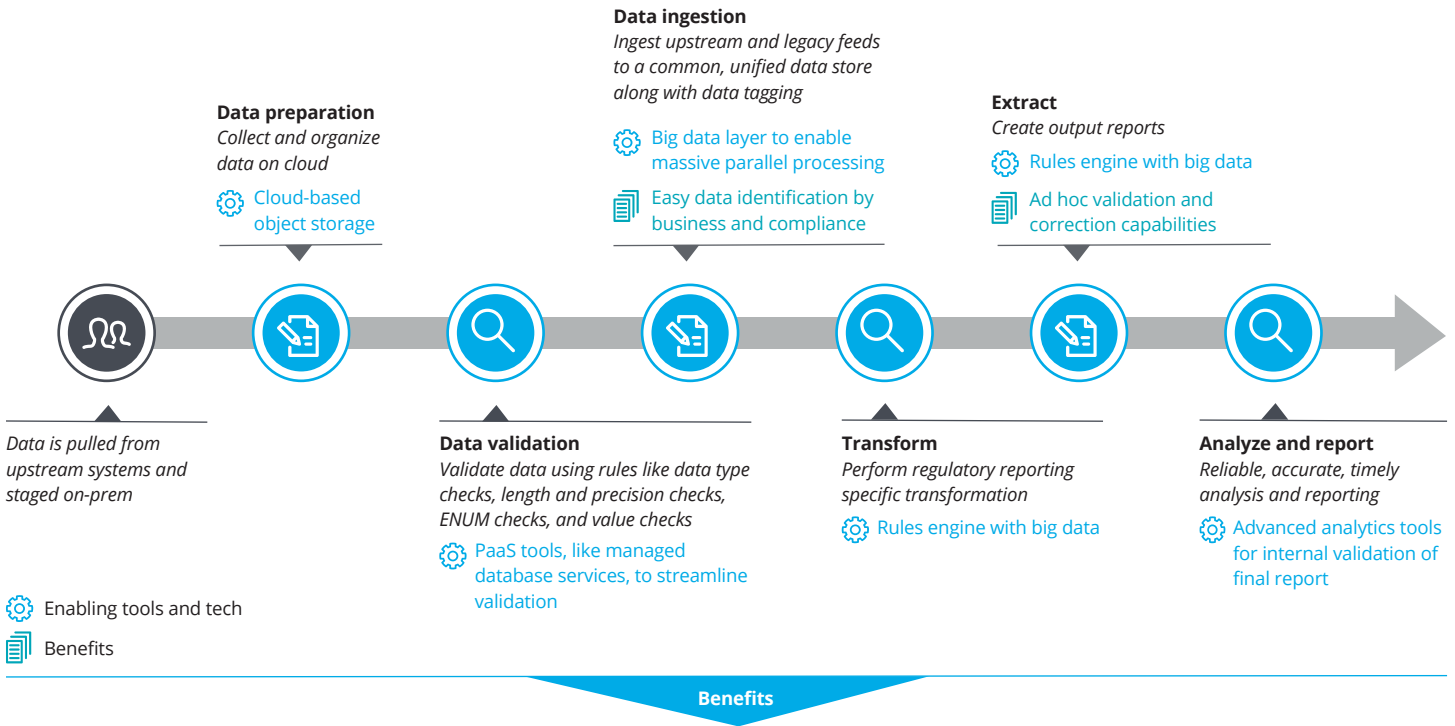
*Helped modernize reg reporting infrastructure using cloud and big data*

 **Impact**

- Achieved the regulatory SLA of processing ~300-800 of millions of data daily within T+1 8 a.m. ET, even during highly volatile trading in March 2020
- Saved *data preparation cost and time* for regulatory reporting teams across the bank through the creation of the golden source data repository on cloud
- Facilitated *heightened understanding of front-office data* through data lineage, helping the regulatory team identify edge cases and improve reporting quality

-  Golden source
-  Cloud-based PaaS platforms
-  Big data pipeline
-  Data lineage

Regulatory reporting data pipeline – Solution overview



- Structured data** flows enabling efficiency levers through reduced data preparation across the bank
- Business-driven workflows** to help enable parallel pipelines for independent flows
- Increased transparency** into transaction processing due to data lineage
- Processes organized** in a logical manner versus organized by departmental structures and green-screen designs
- Improved service quality** via reduction in fragmentation of work and with auto-scaling abilities of infrastructure

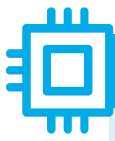
Another modernization issue that financial institutions often deal with is tackling **multiple infrastructure environments** in the complex interplay of hybrid cloud, SaaS offerings, and external vendors. Case study 2 below provides some perspectives into this.

—| **Case study 2.** Solving integration challenges for a dual IT stack |—————

“How do I **reduce manual interventions** in the brokerage business processes (e.g., money movement, fund access) to **meet rapid growth in volume and clients**?”

 - **Process principal, brokerage operations**




- Highly manual, fragmented data ingestion processes from internal and external recordkeeping systems required 40+ FTEs to manage data exceptions using 150+ end-user computing tools (EUCTs)
- Operational data and recordkeeping were done in separate systems, resulting in inflexible processes and creating data integration challenges
- High risk profile of critical systems and data stores resulted in resistance to change
- Underinvestment in core brokerage technology over the past 10+ years created data gaps

 **150+**  
**EUCTs**

 **10+**  
**YEARS OF LEGACY PROCESSES**

 **Solution**

- Reengineered processes using design thinking principles to make them lean and agile
- Developed a cloud-based operational data store (ODS) for addressing all data needs of reengineered processes
- Ringfenced the legacy internal and external recordkeeping systems and designed a state-of-the-art ingestion process to create cloud ODS and exposed data elements via microservices-based plug-and-play architecture
- Integrated a cloud-hosted workflow management SaaS solution with a personalized workspace to increase data transparency
- Established a digital factory to enable a sustainable model of product delivery

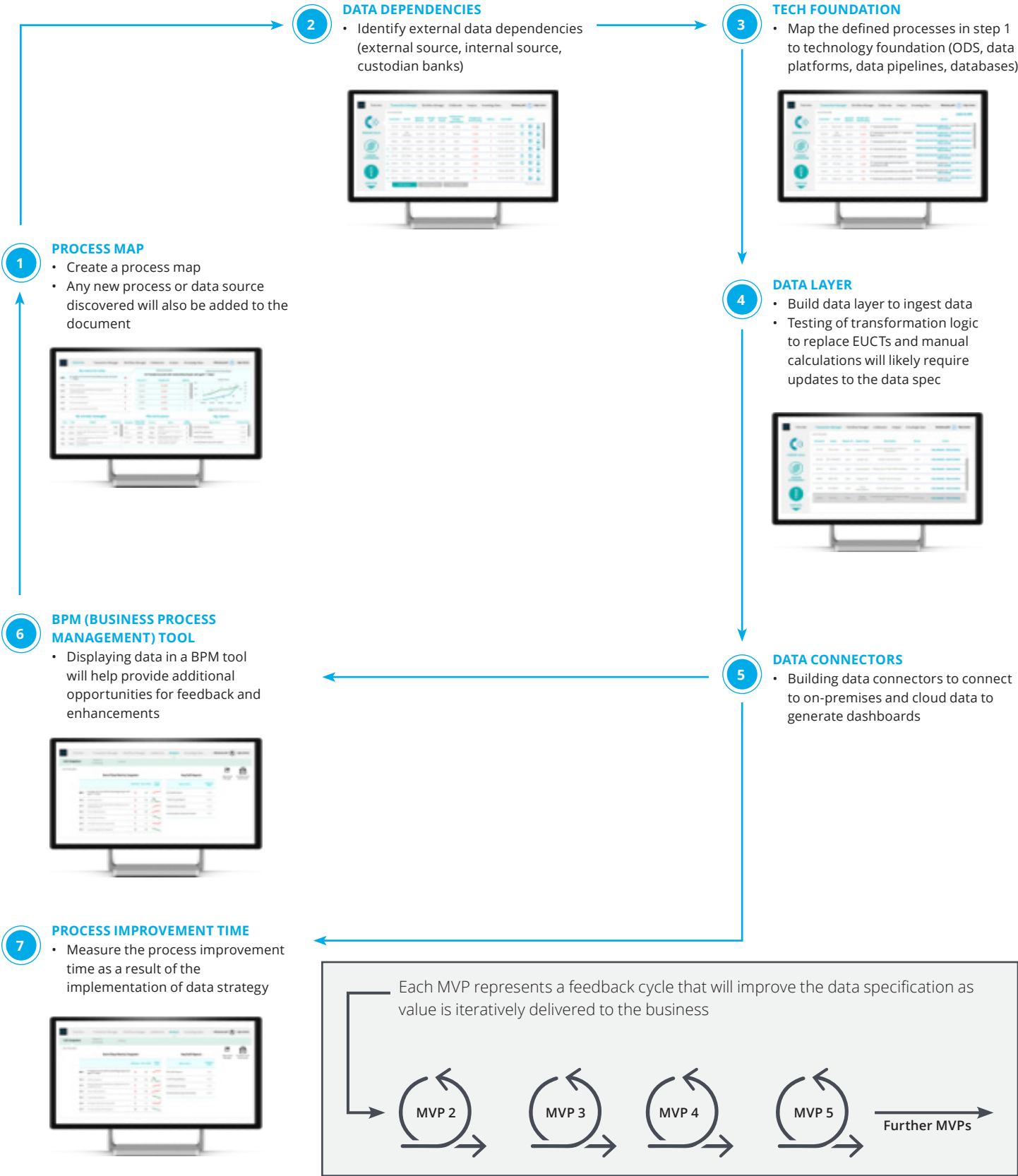
-  Workflow management
-  Scalable and unified operational data store
-  Business process Re-engineering and automation

**Helped automate process breaks and modernize exception management workflows in cloud using microservices**

 **Impact**

- Enabled **straight-through processing** due to better data quality and more aligned processes
- Improved **client experience** due to reduced EUCTs and lower levels of errors and exception-based processing
- **Released 30+ FTEs** from maintenance of legacy systems and processes and reallocated them to focus on higher-value work
- Embedded **automated controls** into process (e.g., segregation of duties, failed data transfer alerts)
- Created a **single source of truth** through the new cloud ODS for reporting KPIs, KRIs, and other analytics use cases
- Provided **visibility** into metrics and work volume to **empower staff** to prioritize and manage their work

Integrating dual IT stack – Solution overview



Case study 2 makes a strong case for accelerating modernization using well-managed data on new cloud environments while still carrying on run-the-bank activities at on-premises environments. As illustrated in the case study, this involves complex interplay of data across multiple environments, requiring ringfencing of legacy systems, integrating hybrid cloud environments seamlessly through modern data ingestion processes, and exposing data through a microservices-based API economy to the new cloud-based processes. All of these are essential elements of a well-formed data strategy.

To summarize, the two case studies are testimonies of how banks are increasingly using data strategy initiatives to address their perpetual issues related to modernization. In the next few sections, let us take a closer look at some of these issues.



### Defining the problem for financial institutions

Many financial firms have **disparate data sources** across **multiple environments**. These are often caused by years of decision-making at the corporate level, such as acquisitions, siloed org structures, and more than a decade of underinvestment in technology. All of this can result in **fragmented platforms** across asset classes and geographies. These barriers often prevent even the best-intentioned firms from (a) attaining increased **automation levels**, (b) performing platform consolidation (due to **lack of a “golden source”** of data), and (c) transforming IT to operate in a **modern environment**.

As a result, some financial firms are struggling to **achieve greater delivery speed** (e.g., faster client onboarding, enhanced participant experiences), **increase efficiency** (e.g., access to real-time pretrade insights, such as compliance checks), and successfully **meet regulatory and compliance deadlines**. Hence, firms may need **pragmatic “data initiatives” (which include both data strategy and data integration initiatives)** to help them overcome these barriers and transform their operations through technology investments in order to increase profitability.

The next few sections closely examine the data challenges faced by financial institutions across the five modernization barriers and subsequently summarize potential options to tackle these issues.



### 1 Scale barrier

Most financial firms operate at a huge scale with **complex global operations** and a **high legacy IT footprint**. This often translates to the **unavailability of high-quality, easily accessible data for new technology solutions**. Data complexity, coupled with a stringent regulatory regime, buries firms under operational issues and leaves them with very little appetite for innovation. Thus, firms constantly play catch-up when it comes to adopting new technology.

#### Scale barrier illustrative example:

Geographical scale introduces multiple complexities in a bank's data operations due to country-specific regulatory requirements. In case study 1 above, the US subsidiary organization was part of a large European parent entity that had operations across the globe. The parent entity's trading desks from the United Kingdom, Hong Kong, and Brazil, routed orders to the US entity, and US regulation required applicable trades to be reported. These trades contained information identifying overseas clients, hence posing a risk under GDPR. To address such challenges, the firm had to identify each foreign trade by working closely with different desks and front-office systems and appropriately anonymize any sensitive information. This significantly increased reporting complexity, making it difficult to maintain and enhance systems in the future.



2



## Technology barrier

Technology barriers create **operational overheads**, most of which are related to data issues. Let us illustrate this using case study 1 above, where the client's regulatory reporting application was required to aggregate data across 10+ source systems. Each source system had its own data structure that required complex ETL operations, including validation, preparation, and aggregation of data to a common reporting standard. Performing such complex ETL transformations daily for a high volume (~300M to 1B records daily) created an operational overhead for the bank and posed a perpetual risk of missing the daily regulatory reporting SLA. This risk was further amplified as the existing on-premises data warehouse solution was not equipped to handle such high volumes, forcing the bank to transform their entire regulatory reporting stack.

This is not a problem for regulatory reporting alone. **Lack of data standards** often create operational overheads, and many firms feel a painful need for a **"golden source"** data platform that can process data from multiple sources into a unified standard, reduce operational overheads, and serve as a single source of truth. However, coming up with a unified data model that serves all stakeholders is not enough to address the problem. Firms should additionally address the following key technical considerations while designing such a "golden source" for data:



### Infrastructure

Have you considered building the new data platform on cloud to avoid potential scalability challenges?

.....



### Automation

What is your current level of automation for repetitive data preparation tasks, which are required to clean up structured and unstructured data?

.....



### Massive volumes

Have you considered redesigning your processes leveraging massive parallel processing of big data to avoid bottlenecks of serialized processing?

The advent of cloud, RPA, and big data technologies can offer a potential solution to address these technical challenges. However, these technologies often come with their own set of complexities. In our example, the bank spent significant effort in creating a common big data pipeline on public cloud to transform and ingest records from all 10+ on-premises source systems into a cloud data warehouse ("golden source"). However, once this was achieved, all downstream regulatory reporting applications started consuming clean data, significantly reducing the operational overhead.

**Integration challenges** also introduce operational overheads and prevent firms from efficiently scaling up processes in line with user and data growth. As an illustrative example, consider case study 2 above, where the bank reimagined their operational processes to reduce manual interventions. These new processes depended on operational data, which still needed to be collected from various internal and external record-keeping systems, moved to the new cloud operational data store, and exposed appropriately for further consumption by new processes. This involved integrating data across multiple technology stacks and environments. Integrations of such a complex nature need to keep certain critical considerations in mind for a successful implementation, as highlighted below.



### Data cataloging

Have you taken a step back to first create a data catalog to identify what data you have on cloud already, what data is higher-quality versus lower-quality, what data is sensitive versus not, what data is from a system of record versus other sources, and so on?

.....



### Data connectors

Do you have capabilities to integrate multiple systems across on-prem and cloud, such as building APIs to connect on-premises with cloud or building integration archetypes and patterns which can be scaled across the org?

.....



### Data volumes

How are you tackling the explosive rate of growth in data volumes? What is your archival and security strategy?

In our case study, the bank solved this by coming up with a **state-of-the-art data ingestion process** to push data into the cloud data store and exposed them by designing a comprehensive microservices framework.



### 3 Organizational silo barrier

Organizations operating in silos can be possessive of their data and not open to giving up ownership to the rightful owner. **Fragmented data ownership** across front, middle, and back offices makes it extremely difficult to have a **360-degree view of data domains**, resulting in suboptimal processes and workarounds built to address them. This could also cause data breaks and create the need for a number of reconciliations.

#### Organizational barrier illustrative example:

A bank may want to enable better customer journeys and self-service capabilities for their customers across multiple asset classes using a 360-degree view of their customers. As a result, the bank's customer experience team may need to do the following:

1. Ask other business units (like pretrade, risk, and client onboarding) for customer data
2. Collect, cleanse, and transform the data for analytics

However, the customer experience team may have limited context of how the data was generated by other business units, leading to gaps in interpretation. The other business units could also feel concerned—once the data is shared, how will the customer experience team use it? Will they share this data externally? What if they find errors or inconsistencies and start pointing fingers?

Without clear incentives for cooperation across groups, initiatives like this typically present risks without benefit and call for data-sharing alignment.

Another key organizational barrier is related to **data governance**. More than 70%<sup>2</sup> of employees have access to data they should not, and 80% of analysts' time is spent simply discovering and preparing data. Even if data access is controlled by centralized policy within the organization, many applications require access to data stored in upstream applications. We have seen such cases where all users of the downstream application are thus unknowingly granted unrestricted read access to the entire upstream application. How can organizations segregate data access through "need-to-know" roles? Aligning each application's view of data and presentation with underlying data is also a key requirement.

Considerations to address these kinds of organizational barriers are:



#### Control

Have you established the right set of controls to ensure data is accessible to the correct set of people? For example, even if data access is controlled within the organization, an application accessing that data could be letting all its users view the data unknowingly.



#### Governance and quality

Is data governance and quality a small part of the job for many people across the company who are the closest to the data and understand it in the context of the business? This helps to clean up data on an iterative, constant basis whenever and wherever quality issues are identified.



#### Org structure

How do you ensure that analytical talent is spread across the organization? Do you have a high percentage of employees responsible for analytics in each team? Are executives acting as analytics champions?

4 

**Talent barrier**

Today's top technical talent does not wish to work on legacy systems and platforms because they correctly assume that it could limit their ability to **develop transferable skills** in a very competitive job market. Since traditional financial firms are still burdened with many legacy platforms, they may find it difficult to attract the right talent to help them thrive and compete with new and more tech-savvy competitors. This makes it important for firms to identify and **develop their future talent model**. Typical considerations that firms need to address while designing their technology talent model for the future are:



**Skills**

Have you determined the skills that will be critical in your future IT operating model?

.....



**Maintenance**

How do you ensure the maintenance of existing systems in the future?

.....



**Resource management**


How do you recruit people with new skills that will be critical for your future IT organization? How do you retain people whose skills are critical today?

.....



**Institutional knowledge**

How will you capture existing institutional knowledge that is a key source of competitive advantage?

5 

**Ageing capitalized infrastructure barrier**

Rigid legacy data infrastructure may be fully capitalized and low-cost to run, but it **limits data and analytics capabilities** and prevents firms from realizing the full potential of data. Even as firms make the transition from a customized set of legacy (e.g., mainframe) applications to a configurable set of cloud-native applications, it is essential to preserve the valuable intellectual property lying in existing, aging infrastructure. Firms should consider the following key aspects when tackling this barrier:



**ROI**

What is the ROI for moving to a new infrastructure versus maintenance of legacy capitalized infrastructure? What are the risks associated with retaining business on aging infrastructure?

.....



**Cost of migration**

What is the cost of developing new capabilities on existing infrastructure versus ringfencing legacy infrastructure and building new business on cloud?

.....



**Architecture options**

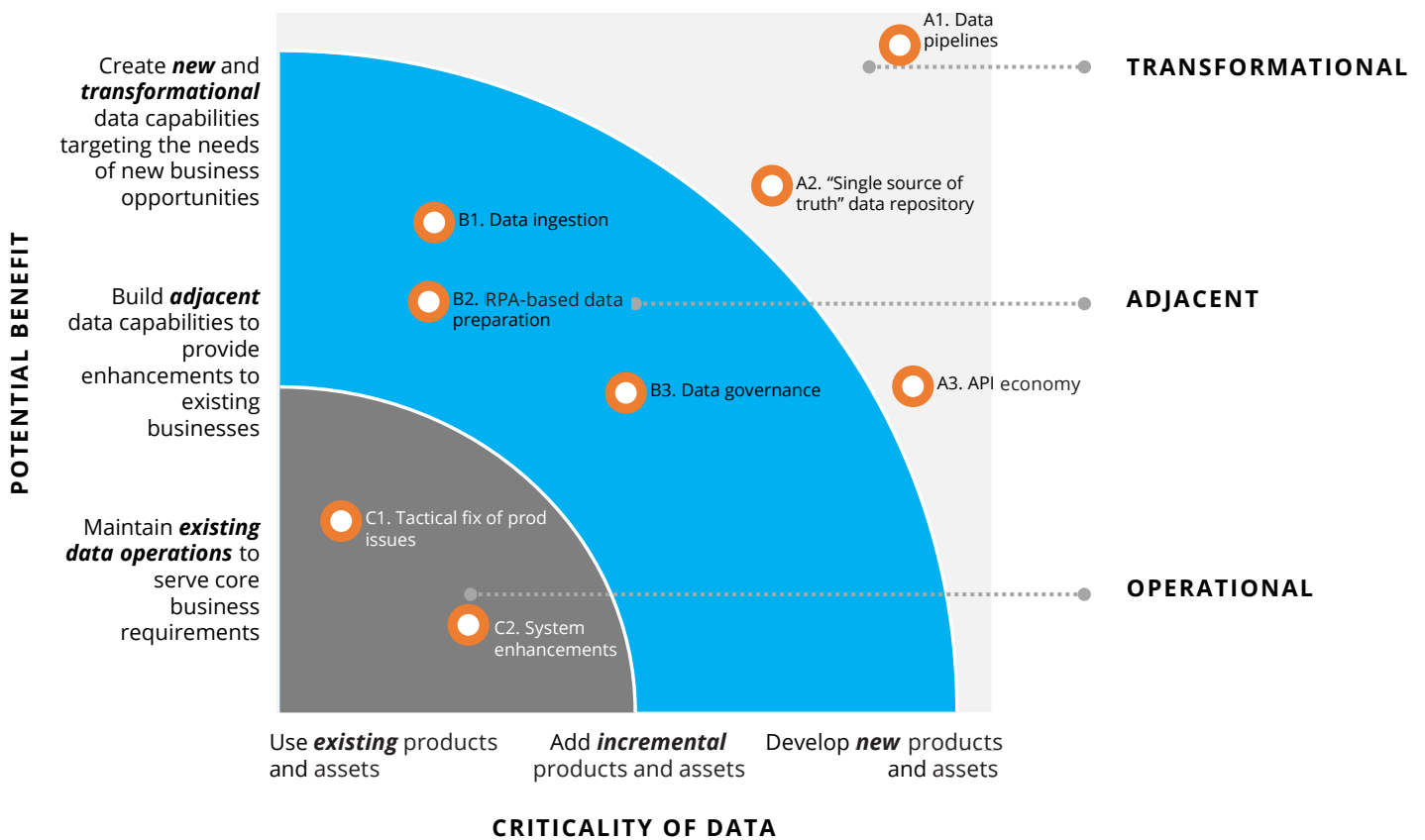
What are the architecture options for connecting data between aging infrastructure and cloud?

Now that we have covered each barrier in detail, let us see some of the new ways to tackle each issue.

**Using data strategy and data integration initiatives to tackle institutional barriers**

Since data issues are at the heart of each of these five barriers, firms need to undertake specific data strategy and data integration initiatives to resolve these issues. These initiatives can be broadly categorized into three areas:

- 1 **TRANSFORMATIONAL** - Developing new breakthroughs in data management and integration capabilities that don't yet exist within the company
- 2 **ADJACENT** - Expanding existing data management and integration capabilities to meet requirements that can be addressed within one to three months
- 3 **OPERATIONAL** - Supporting lights-on activities for ongoing data and integration capabilities



“Transformational” initiatives are led directly by CXOs and BU heads, while “adjacent” and “operational” initiatives can be led by portfolio managers, product owners, and project controllers. Each data strategy initiative can help firms address multiple modernization barriers. The data strategy initiatives mentioned in this article focus on the “transformational” and “adjacent” categories and are summarized in the below table along with illustrative complexity and impact across the five barriers.

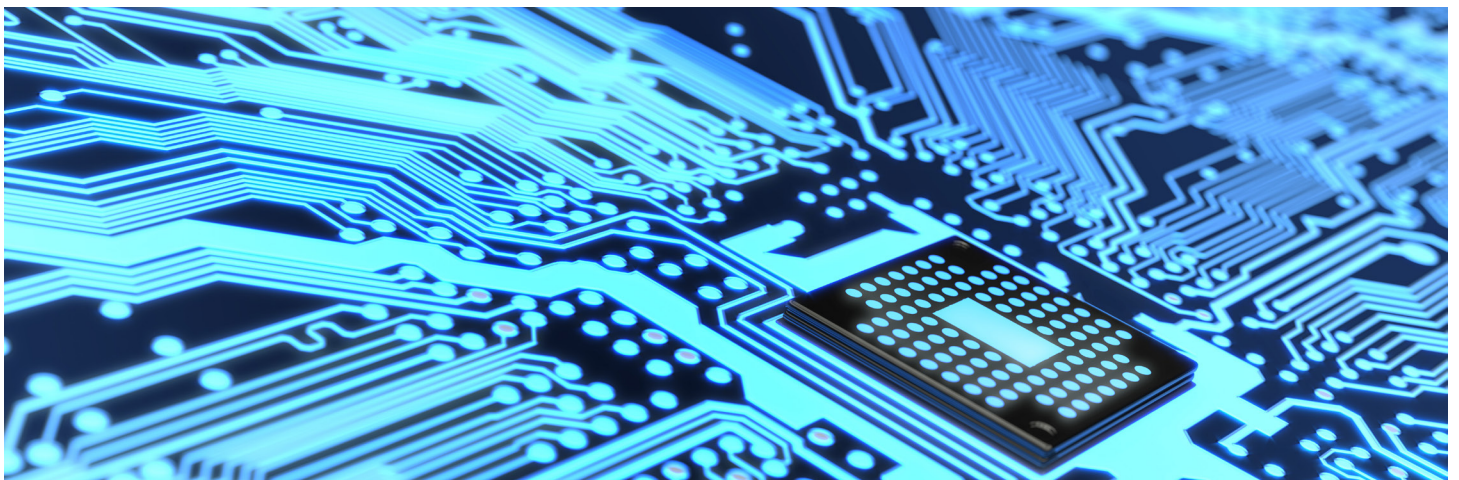
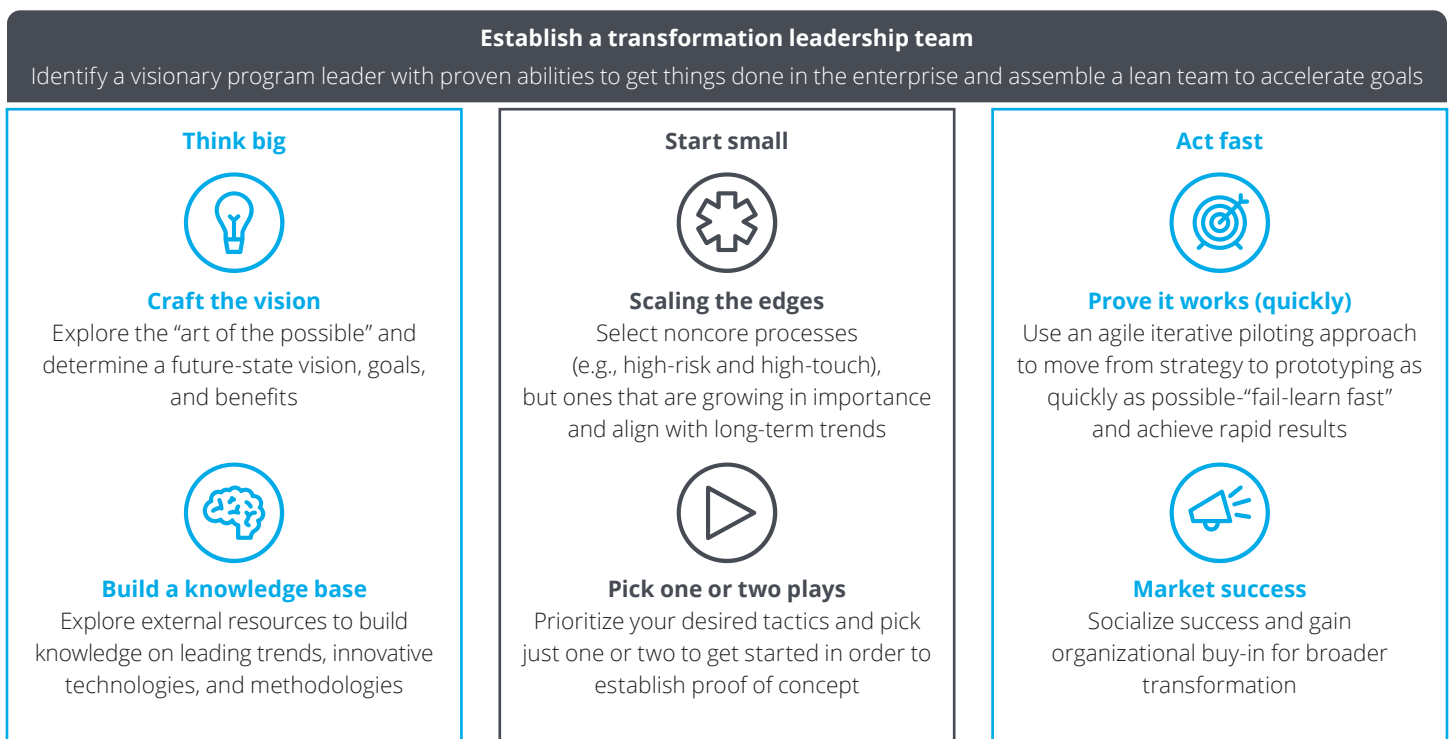
		← Impact on modernization barriers →					
Category	Data strategy initiative	Complexity	Scale	Technology	Org. silo	Talent	Infrastructure
Transformational	<b>Build data pipelines for strategic advantages</b> <ul style="list-style-type: none"> <li>Ingest and transform high volumes of data from multiple sources for the consumption of multiple internal or external products</li> </ul>	High	●	●	●	◐	◑
	<b>Set up a “golden source of truth” platform</b> <ul style="list-style-type: none"> <li>Create a central data hub from multiple disparate sources to reduce data duplication and improve data governance and operational efficiencies</li> </ul>	High	◑	◑	●	◐	◑
	<b>Think “API economy”</b> <ul style="list-style-type: none"> <li>Expose digital services and assets using APIs in a controlled, secure, managed, and orchestrated way to streamline development</li> </ul>	High to medium	◑	◑	◑	◐	◑
Operational	<b>Enable data ingestion</b> <ul style="list-style-type: none"> <li>Import data for immediate use or for storage from multiple environments for faster data access</li> </ul>	Medium	◑	◑	◑	◐	◑
	<b>Leverage RPA</b> <ul style="list-style-type: none"> <li>Automate repetitive tasks and manual processes related to resolution of data issues to improve operational efficiencies and data quality</li> </ul>	Medium	◑	◑	◑	●	○
	<b>Establish data governance</b> <ul style="list-style-type: none"> <li>Establish clear roles, responsibilities, and processes to ensure formal management of data throughout the enterprise</li> </ul>	Medium	●	◑	◑	◐	◑

Harvey balls indicate the impact of initiatives across each of the five barriers. Initiatives with “high” impact are more suitable to solve a barrier, and vice versa for initiatives with “low” impact.



# How can you get started?

Since the transformation budgets of firms are limited, it is important to design a strategic road map of data initiatives and only prioritize selected transformation initiatives that can deliver the most value. One rule of thumb that can be used to validate the strategic road map is Pareto's principle (commonly known as the 80-20 rule)-firms may often discover that only about 20% (or a smaller subset) of all existing data issues need a "transformational" approach. This can help firms design a practical road map that is achievable under a realistic IT budget. To get started on such transformational initiatives, a stable leadership team should also be in place to champion the cause and provide guidance throughout the transformation journey. Below is an illustrative framework on how firms can approach this.



# How can Deloitte help?

Through our rich experience, Deloitte has created solution accelerators to help financial institutions kick-start their strategic data and integration initiatives



## Canonical data model

- Helps firms quickly identify and prioritize attributes that should be considered for centralization in a unified view and understand what can sit outside of the unified view.



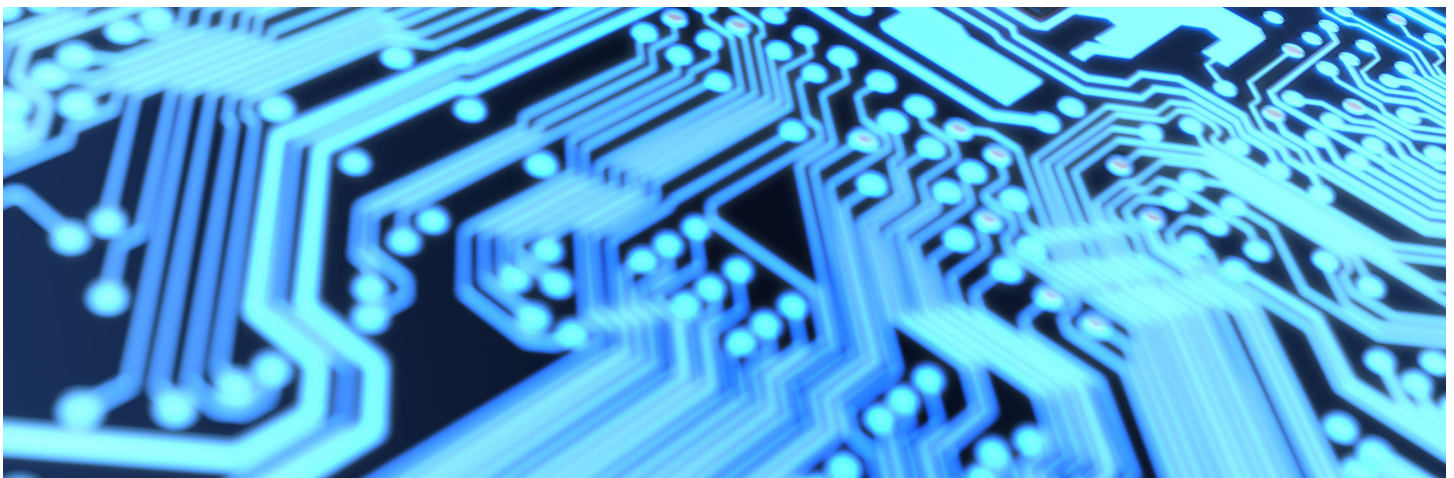
## Cloud-based data pipeline and data modernization

- This configurable framework addresses common integration challenges for firms operating their IT on a dual stack.
- The solution can ingest streaming data, as well as flat files, from on-premises data sources to cloud, perform reference data enrichments, and also create a scalable data pipeline on cloud for running big data transformation workloads. Each part of the pipeline is modular and can be utilized as a stand-alone tool for specific scenarios.
- The solution also comes with a plug-and-play analytics module that can enable tracking of data performance metrics and KPIs across the pipeline.



## Data validation engine

- The data validation engine provides detailed feedback to business and IT users on data quality issues using configurable rules through an out-of-the-box UI and also exposes REST API interfaces for integration with third-party apps and services.



# Authors

For more details on how these solutions can address your specific data challenges, contact the authors below.



## Sriram Gopalakrishnan

Data Strategy and Transformation  
FSS, Deloitte Consulting LLP  
[sgopalakrishnan@deloitte.com](mailto:sgopalakrishnan@deloitte.com)

Sri has more than 20 years of experience helping broker-dealers with large-scale initiatives related to technology transformation (cloud adoption, application rationalization, and process optimization), regulatory reform (living wills, DF, EMIR, and CAT), and building enterprise capabilities (onboarding, settlement, clearing, and collateral).



## Phillip Matricardi

Data specialist  
FSS, Deloitte Consulting LLP  
[pmatricardi@deloitte.com](mailto:pmatricardi@deloitte.com)

Phillip has more than eight years of experience as a specialist in over-the-counter (OTC) derivatives market reform triggered by Dodd-Frank and EMIR legislations, specifically collateral, central clearing, margin settlement, and related infrastructure (global standard establishment, collateral settlement system design, client onboarding). He also has seven years of experience in data analysis and operations for online marketing.



## Jayanta Ghosh

Cloud architect  
FSS, Deloitte Consulting LLP  
[jaghosh@deloitte.com](mailto:jaghosh@deloitte.com)

Jayanta has more than 13 years of experience leading technology advisory work within capital markets, having driven transformation initiatives focusing on system and infrastructure upgrades, application architecture changes, vendor and platform assessment, business and functional architecture blueprints, and managed business operation and projects.



## Abhimanyu Basu

Data Strategy and Products  
FSS, Deloitte Consulting LLP  
[abbasu@deloitte.com](mailto:abbasu@deloitte.com)

Abhimanyu has seven years of experience helping clients define and implement data strategy initiatives. He has worked extensively with banking and broker-dealer clients on regulatory reform (CAT), data modernization (cloud migration and big data pipelines), and M&A advisory (commercial due diligence, divestiture planning and execution). Abhimanyu has rich experience in product management, solution architecture, process improvement, and project management.



## W Snigdha Rao

Data Strategy and Analytics  
FSS, Deloitte Consulting LLP  
[wrao@deloitte.com](mailto:wrao@deloitte.com)

Snigdha has nearly four years of experience catering to financial services clients, working extensively on customer analytics (improving retention and dropout rates), digital transformation (XaaS, data analytics, and data management), and IT portfolio management (capability playbooks, process mapping, and technology portfolio playbooks). She has donned multiple hats across roles such as business analyst, product manager, and project manager.





As used in this document, "Deloitte" means Deloitte Consulting LLP, a subsidiary of Deloitte LLP. Please see [www.deloitte.com/us/about](http://www.deloitte.com/us/about) for a detailed description of our legal structure. Certain services may not be available to attest clients under the rules and regulations of public accounting.

This publication contains general information only and Deloitte is not, by means of this publication, rendering accounting, business, financial, investment, legal, tax, or other professional advice or services. This publication is not a substitute for such professional advice or services, nor should it be used as a basis for any decision or action that may affect your business. Before making any decision or taking any action that may affect your business, you should consult a qualified professional adviser. Deloitte shall not be responsible for any loss sustained by any person who relies on this publication.

Copyright © 2020 Deloitte Development LLC. All rights reserved.

Designed by CoRe Creative Services. RITM0564997