

# McKinsey on Payments

## Special Edition on Advanced Analytics in Banking



4

The analytics-enabled collections model



13

How machine learning can improve pricing performance



20

Combating payments fraud and enhancing customer experience



28

Using data to unlock the potential of an SME and mid-corporate franchise



36

Hidden figures: The quiet discipline of managing people using data



44

Using analytics to increase satisfaction, efficiency, and revenue in customer service



52

Designing a data transformation that delivers value right from the start



60

Building an effective analytics organization



68

"All in the mind": Harnessing psychology and analytics to counter bias and reduce risk



80

Mapping AI techniques to problem types



86

Data sheet: Advanced analytics



# Designing a data transformation that delivers value right from the start

The CEOs of most financial institutions have had data on their agenda for at least a decade. However, the explosion in data availability over the past few years—coupled with the dramatic fall in storage and processing costs and an increasing regulatory focus on data quality, policy, governance, models, aggregation, metrics, reporting, and monitoring—has prompted a change in focus. Most financial institutions are now engaged in transformation programs designed to reshape their business models by harnessing the immense potential of data.

Chiara Brocchi  
Davide Grande  
Kayvaun Rowshankish  
Tamim Saleh  
Allen Weinberg

Leading financial institutions that once used descriptive analytics to inform decision-making are now embedding analytics in products, processes, services, and multiple front-line activities. And where they once built relational data warehouses to store structured data from specific sources, they are now operating data lakes with large-scale distributed file systems that capture, store, and instantly update structured and unstructured data from a vast range of sources to support faster and easier data access. At the same time, they are taking advantage of cloud technology to make their business more agile and innovative, and their operations leaner and more efficient. Many have set up a new unit under a chief data officer to run their data transformation and ensure disciplined data governance.

Successful data transformations can yield enormous benefits. One US bank expects to see more than \$400 million in savings from rationalizing its IT data assets and \$2 billion in gains from additional revenues, lower capital requirements, and operational efficiencies. Another institution expects to grow its bottom line by 25 percent in target segments and products thanks to data-driven business initiatives. Yet many other organizations are struggling to capture real value from their data programs,

with some seeing scant returns from investments totaling hundreds of millions of dollars.

A 2016 global McKinsey survey found that a number of common obstacles are holding financial institutions back: a lack of front-office controls that leads to poor data input and limited validation; inefficient data architecture with multiple legacy IT systems; a lack of business support for the value of a data transformation; and a lack of attention at executive level that prevents the organization committing itself fully (Exhibit 1). To tackle these obstacles, smart institutions follow a systematic five-step process to data transformation.

## 1. Define a clear data strategy

Obvious though this step may seem, only about 30 percent of the banks in our survey had a data strategy in place. Others had embarked on ambitious programs to develop a new enterprise data warehouse or data lake without an explicit data strategy, with predictably disappointing results. Any successful data transformation begins by setting a clear ambition for the value it expects to create.

In setting this ambition, institutions should take note of the scale of improvement other organizations have achieved. In our experience, most of the value of a data trans-

formation flows from improved regulatory compliance, lower costs, and higher revenues. Reducing the time it takes to respond to data requests from the supervisor can generate cost savings in the order of 30 to 40 percent, for instance. Organizations that simplify their data architecture, minimize data fragmentation, and decommission redundant systems can reduce their IT costs and investments by 20 to 30 percent. Banks that have captured benefits across risk, costs, and revenues have been able to boost their bottom line by 15 to 20 percent. However, the greatest value is unlocked when a bank uses its data transformation to transform its entire business model and become a data-driven digital bank.

*Actions: Define the guiding vision for your data transformation journey; design a strategy to transform the organization; establish clear and measurable milestones.*

## 2. Translate the data strategy into tangible use cases

Identifying use cases that create value for the business is key to getting everyone in the organization aligned behind and committed to the transformation journey. This process typically comprises four steps.

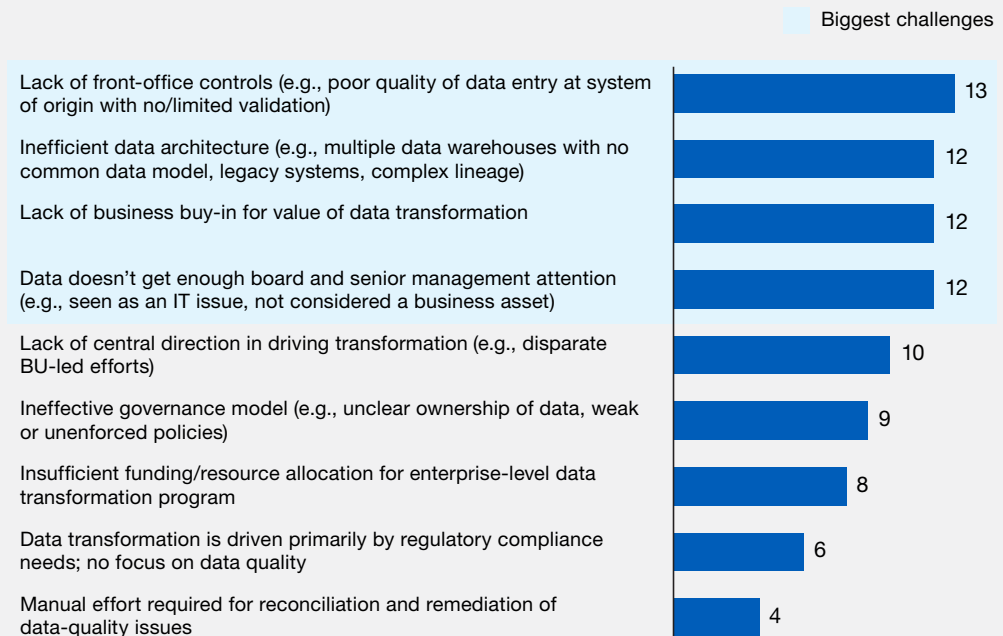
In the first step, the institution breaks down its data strategy into the main goals it wants to achieve, both as a whole and within individual functions and businesses.

Exhibit 1

### Typical challenges in data transformations at banks.

#### Challenges faced in improving data quality at the enterprise level, ranked by perceived importance

Number of participants ranking the challenge in 1st or 2nd place, from a total of 43 respondents



Source: McKinsey analysis

Next it draws up a shortlist of use cases with the greatest potential for impact, ensures they are aligned with broader corporate strategy, and assesses their feasibility in terms of commercial, risk, operational efficiency, and financial control. These use cases can range from innovations such as new reporting services to more basic data opportunities, like the successful effort by one European bank to fix quality issues with pricing data for customer campaigns, which boosted revenues by 5 percent.

Third, the institution prioritizes the use cases, taking into account the scale of impact they could achieve, the maturity of any technical solutions they rely on, the availability of the data needed, and the organization's capabilities. It then launches pilots of the top-pri-

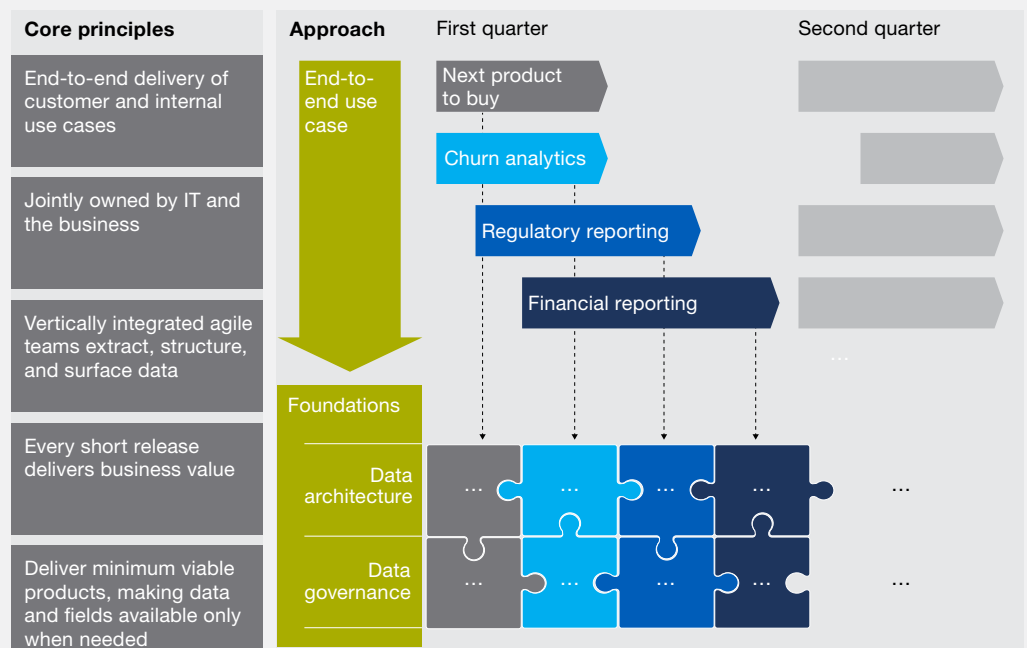
ority use cases to generate quick wins, drive change, and provide input into the creation of a comprehensive business case to support the overall data transformation. This business case includes the investments that will be needed for data technologies, infrastructure, and governance.

The final step is to mobilize data capabilities and implement the operating model and data architecture to deploy the use cases through agile sprints, facilitate scaling up, and deliver tangible business value at each step (Exhibit 2). At one large European bank, this exercise identified almost \$1 billion in expected bottom-line impact.

*Actions: Select a range of use cases and prioritize them in line with your goals; use top-prior-*

Exhibit 2

## Banks can deliver end-to-end use cases at speed via agile sprints.



Source: McKinsey analysis

*ity use cases to boost internal capabilities and start laying solid data foundations.*

### **3. Design innovative data architecture to support the use cases**

Leading organizations radically remodel their data architecture to meet the needs of different functions and users and enable the business to pursue data-monetization opportunities. Many institutions are creating data lakes: large, inexpensive repositories that keep data in its raw and granular state to enable fast and easy storage and access by multiple users, with no need for pre-processing or formatting. One bank with data fragmented across more than 600 IT systems managed to consolidate more than half of this data into a new data lake, capturing enormous gains in the speed and efficiency of data access and storage. Similarly, Goldman Sachs has reportedly consolidated 13 petabytes of data into a single data lake that will enable it to develop entirely new data-science capabilities.

Choosing an appropriate approach to data ingestion is essential if institutions are to avoid creating a “data swamp”: dumping raw data into data lakes without appropriate ownership or a clear view of business needs, and then having to undertake costly data-cleaning processes. By contrast, successful banks build into their architecture a data-governance system with a data dictionary and a full list of metadata. They ingest into their lakes only the data needed for specific use cases, and clean it only if the business case proves positive, thereby ensuring that investments are always linked to value creation and deliver impact throughout the data transformation.

However, data lakes are not a replacement for traditional technologies such as data

warehouses, which will still be required to support tasks such as financial and regulatory reporting. And data-visualization tools, data marts, and other analytic methods and techniques will also be needed to support the business in extracting actionable insights from data. Legacy and new technologies will coexist side by side serving different purposes.

The benefits of new use-based data architecture include a 360-degree view of consumers; faster and more efficient data access; synchronous data exchange via APIs with suppliers, retailers, and customers; and dramatic cost savings as the price per unit of storage (down from \$10 per gigabyte in 2000 to just 3 cents by 2015) continues to fall.

In addition, the vast range of services offered by the hundreds of cloud and specialist providers—including IaaS (infrastructure as a service), GPU (graphics-processing unit) services for heavy-duty computation, and the extension of PaaS (platform as a service) computing into data management and analytics—has inspired many organizations to delegate their infrastructure management to third parties and use the resulting savings to reinvest in higher-value initiatives.

Consider ANZ’s recently announced partnership with Data Republic to create secure data-sharing environments to accelerate innovation. The bank’s CDO, Emma Grey, noted that “Through the cloud-based platform we will now be able to access trusted experts and other partners to develop useful insights for our customers in hours rather than months.”

*Actions: Define the technical support needed for your roadmap of use cases; design a modular, open data architecture that makes it easy*

to add new components later.

#### 4. Set up robust data governance to ensure data quality

The common belief that problems with data quality usually stem from technology issues is mistaken. When one bank diagnosed its data quality, it found that only about 20 to 30 percent of issues were attributable to systems faults. The rest stemmed from human error, such as creating multiple different versions of the same data.

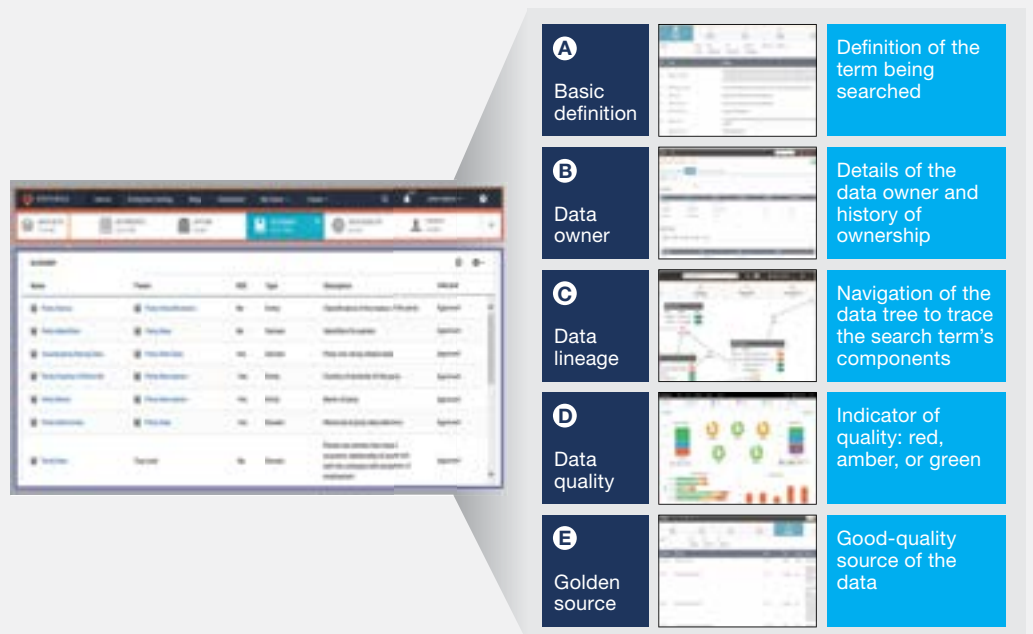
Robust data governance is essential in improving data quality. Some successful financial institutions have adopted a federal-style framework in which data is grouped into 40 to 50 “data domains,” such as demographic data or pricing data. The ownership of each

domain is assigned to a business unit or function that knows the data, possesses the levers to manage it, and is accountable for data quality, with metadata management (such as mapping data lineage) typically carried out by “data stewards.” A central unit, typically led by a chief data officer, is responsible for setting up common data-management policies, processes, and tools across domains. It also monitors data quality, ensures regulatory compliance (and in some cases data security), supports data remediation, and provides services for the business in areas such as data reporting, access, and analytics.

Best-in-class institutions develop their own tools to widen data access and support self-service data sourcing, like the search tool

Exhibit 3

A custom-designed search tool provides users with key information on data elements.



Source: McKinsey analysis



one bank created to provide users with key information about the definition, owner, lineage, quality, and golden source of any given piece of data (Exhibit 3). Organizations with readily accessible information and reliable data quality can deliver solutions much more quickly and with greater precision. They can also create enormous efficiencies along the whole data lifecycle from sourcing and extraction to aggregation, reconciliation, and controls, yielding cost savings that can run as

high as 30 to 40 percent.






*Actions: Assess data quality; establish robust data governance with clear accountability for data quality; provide self-service tools to facilitate data access across the whole organization.*

## 5. Mobilize the organization to deliver value

Successful data transformations happen when a company follows an approach driven by use cases, promotes new ways of working, and mobilizes its whole organization from

Exhibit 4

### Data governance is rolled out domain by domain.

Cluster data into domains		Plan the roll-out of data domains in waves			Implement data governance in each data domain	
Systems of records	Data domains	Roll-out plan by data domain			Area	Key activities
   Customer data		Year 1	Year 2	Year 3	Data management	Data domain definition Identification of common data elements Mapping of data lineage Population of data dictionary Choice of "golden sources" Definition of security requirements
		Pilots	2nd wave	4th wave		
  Collateral data		Q1 & Q2	Collateral data Accounts Product catalog General tables Employees Mortgage Consumer lending Consumer deposits	Business banking Securities Insurance Private banking	Data quality	Assessment of current data quality Definition of KQIs Design of data-quality controls
		Q3 & Q4	Risk management MIS data Finance data Regulatory reporting Operational risk	Commercial lending Commercial real estate Leasing Treasury management Capital markets		
Number of systems	Number of domains	1st wave	3rd wave	5th wave	Data technology and tools	Definition of tools required for implementation
~1,400	50–100					

Source: McKinsey analysis

the beginning. Adopting a use-case-driven approach means developing target data architecture and data governance only when it is needed for a specific use case. One European bank implemented this approach in three steps (Exhibit 4):

First, it identified the data it needed for key use cases and prioritized those data domains

that included it. Typically, 20 percent of data enables 80 percent of use cases. Second, the bank developed a rollout plan for implementing data architecture and governance in three to four data domains per quarter.

Third, the bank set up a cross-functional team for each data domain, comprising data stewards, metadata experts, data-quality

Exhibit 5

### Banks need new roles to compete effectively in a data-driven market.

Senior executive	Translator	Data scientist	Data owner	Head of data governance	Data quality manager	Data technology manager
Digital culture						
Design and agile thinking	Design and agile thinking	Design and agile thinking	Fundamentals of data management	Fundamentals of data governance	Fundamentals of data quality	Fundamentals of data technology tools
Use-case reflections	Source of value	Source of value	Data culture	Data management	Data management	Data management
	Best practices in data engineering	Best practices in testing and piloting	Data quality		Data quality tools	Data modeling
	Best practices in data management	Technical leadership program				Data design
	Best practices in data modeling	Data science				Data software ecosystem
	Technical leadership program	Advanced analytics				
	"Train the trainer" approach					

Source: McKinsey analysis



experts, data architects, data engineers, and platform engineers. Before data was ingested into the data lake, these teams worked to identify key data elements, select golden sources, assess data quality, carry out data cleansing, populate the data dictionary, and map data lineage. Each team worked in agile sprints in a startup-like environment for three to four months. A central team took care of value assurance and defined common standards, tools, and policies.

This approach delivered numerous benefits for the bank, including rapid implementation, capability building, and the creation of tangible business value at every stage in the journey. During any transformation, calling out and celebrating such achievements is critical. As the CDO of JPMorgan Chase, Rob Casper, observed, “The thing that achieves buy-in and builds momentum better than anything is success . . . trying to deliver in small chunks incrementally and giving people a taste of that success [is] a very powerful motivator.”

More broadly, senior executives need to champion their data transformation to encourage widespread buy-in, as well as role-modeling the cultural and mindset changes they wish to see. Formal governance and performance-management systems, mechanisms, and incentives will need to be rethought to support new ways of working. At

the same time, most organizations will need to develop new capabilities; only 20 percent of the banks we surveyed believe they already have adequate capabilities in place. Given the scarcity of external talent, in particular for key roles such as business translators, organizations will need to provide on-the-job training for employees involved in the transformation, and complement this effort with a data and analytics academy to build deep expertise in specialist roles (Exhibit 5, page 58).

*Actions: Adopt a use-case approach to the whole journey; establish central governance to ensure cross-functional working, the use of standard methods, and clear role definition; build new data capabilities through hiring and in-house training.*

\* \* \*

In the past few years data has been established as a fundamental source of business value. Every financial institution now competes in a world characterized by enormous data sets, stringent regulation, and frequent business disruptions as innovative ecosystems emerge to break down the barriers between and across industries. In this context, a data transformation is a means not only to achieve short-term results, but also to embed data in the organization for long-term success.

**Chiara Brocchi** is an expert in McKinsey’s Milan office, where **Davide Grande** is a partner. **Kayvaun Rowshankish** is a partner and **Allen Weinberg** is a senior partner, both in the New York office. **Tamim Saleh** is a senior partner in the London office.