



Why you need a digital data architecture to build a sustainable digital business

Sven Blumberg, Oliver Bossert, Hagen Grabenhorst, and Henning Soller

Companies that succeed at meeting their analytics objectives let business goals drive the technology. Here's how they structure a data architecture that works.

Data architecture has been consistently identified by CXOs as a top challenge to preparing for digitizing business. Leveraging our experience across industries, we have consistently found that the difference between companies that use data effectively and those that do not—that is, between leaders and laggards—translates to a 1 percent margin improvement for leaders. In the apparel sector, for instance, data-driven companies have doubled their EBIT margin as compared to their more traditional peers.

Using data effectively requires the right data architecture, built on a foundation of business requirements. However, most companies take a technology-first approach, building major platforms while focusing too little on killer use cases.

Many businesses, seeing digital opportunities (and digital competition) in their sectors, rush to invest without a considered, holistic data strategy. They either focus on the technologies alone or address immediate, distinct use cases without considering the mid- to long-term creation of sustainable capabilities. This goes some way toward explaining why a 2017 McKinsey Global Survey found that only half of responding executives report even moderate effectiveness at meeting their analytics objectives. The survey found the second-largest challenge companies face (after constructing a strategy to pursue data and analytics) is designing data architecture and technology infrastructure that effectively support data-and-analytics activities at scale. We found that eight out of ten companies embark on digital data enablement by making their IT departments responsible for the data transformation—with very grand implementation programs—and a small set of business use cases.

This strategy is quite different from that employed by next-generation digital leaders, who typically embark on transformation from a business perspective and implement supporting technologies as needed. Doing the technology first produces more problems than successes, including:

- **Redundant and inconsistent data storage.** Only two in ten banks we've worked with have established a common enterprise data warehouse, which is essential for creating a single source of truth for financial and customer data.
- **Overlapping functionality.** Every bank we've worked with has at least one business function supported by three different technological systems.
- **A lack of sustainability.** The solutions at which financial institutions typically arrive are often quick fixes that ignore the enterprises' larger aspirations for datafication. For example, one insurance company extracted and replicated data from its warehouse each time it was needed rather than building data architecture that would allow it to store each customer element only once, thereby reducing costs and eliminating inefficiencies.

These problems have real business consequences. Meeting leading-edge business requirements, such as real-time customer and decision support, and large-scale analytics requires the integration of traditional data warehousing with new technologies.

The two-speed data-architecture imperative

Today, enterprises must cope with increasingly large and complex data volumes (worldwide, data storage doubles every two years) coming from diverse sources in a wide variety of formats that traditional data infrastructures struggle, and most often fail, to operationalize. Developing new business capabilities—such as individual pricing for customers based on real-time profitability, as some insurance companies have done, automating credit decisions that lead to improved outcomes for banks and greater customer satisfaction, or running automated, more cost-effective strategic marketing campaigns as we've seen in the chemicals sector—demands new ways of managing data.

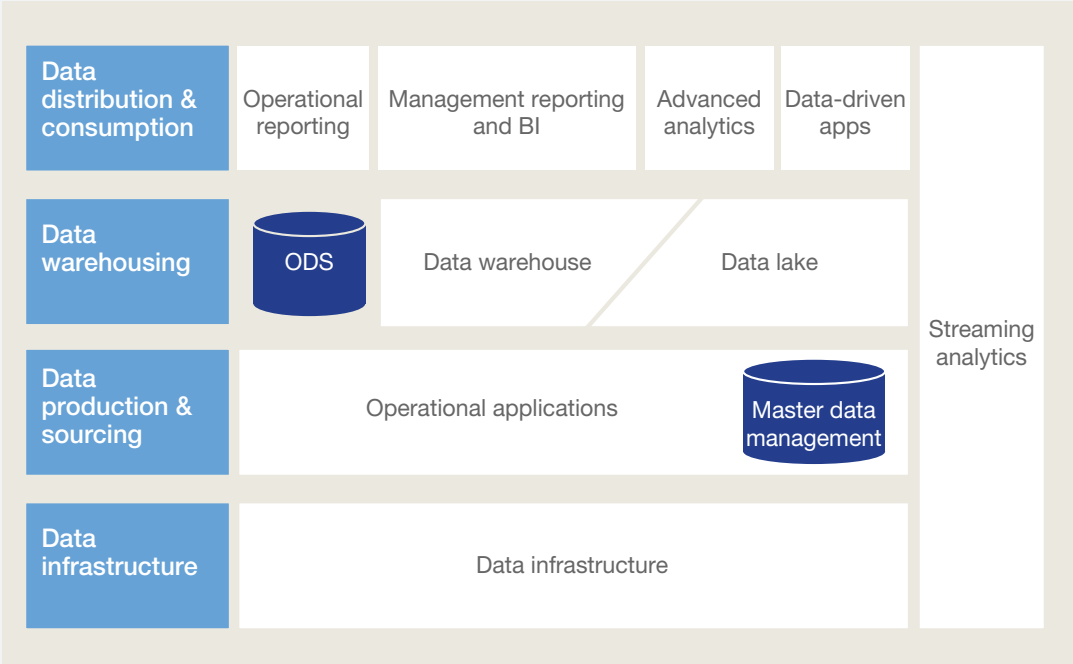
This does not mean, however, that legacy data and IT infrastructures must be trashed, or that new capabilities need to be bolted on. It does mean that the traditional data warehouse, through which the organization gains stability and financial transparency, must be scaled down and integrated with the high-speed transactional architecture that gives the organization the capability to support new products and services (as well as real-time reporting).

This is the two-speed principle.

This new, complex technical environment requires companies to closely examine business use cases before making costly technology decisions, such as needlessly ripping out and replacing legacy architectures. Instead, it is preferable to use a capability-oriented reconceptualization of data management as an enabler of digital applications and processes (Exhibit).

Exhibit

Best-practice reference-data architecture



To implement an end-to-end digital data architecture, an enterprise needs first to develop a point of view on its current and, if possible, future business requirements, sketch its desired, flexible data-management architecture, and create a roadmap for implementation. To begin, one must identify the key business use cases.

To do this, we recommend a thorough review of best-practice use cases across industries that address common value drivers (financial transparency, customer satisfaction, rapid product development, real-time operational reporting, and so on). Then, the company should compare those use cases with its market position and strategic direction, prioritizing those that best reflect the company's situation and aspirations. Once those reference use cases are identified, the company can begin to define target data-architecture capabilities. In this process, the business leads and technology follows.

The high-level structure in the exhibit above represents a layered data architecture that has been applied successfully by many organizations, across many industries, especially in finance. It extends to accommodate new digital capabilities such as collecting and analyzing unstructured data, enabling real-time data processing, and streaming analytics.

The exhibit shows a reference architecture that combines both the traditional requirements of financial transparency via a data warehouse and the capability to support advanced analytics and big data. In a phrase, it's a two-speed approach.

The two-speed architecture adheres to three core principles:

1. A limited number of components with a clear demarcation of capabilities to manage complexity while providing the required functionalities, such as advanced analytics and operational reporting
2. Layers that enable the transparent management of data flows and provide a single source of truth to protect against siloes and data inconsistencies (through the data warehouse, which models, integrates, and consolidates data from various sources)
3. Integration of state-of-the-art solutions with traditional components, such as the data warehouse, to satisfy such new requirements as real-time processing, and an operational data store (ODS) based on new database technologies

We have used this model to:

- Help clients think through and evaluate their options on an architectural level before discussing concrete technical solutions.
- Map technology components against capabilities to manage and avoid redundancies while identifying gaps.

- Create plans for stepwise transformations driven by business value while limiting business disruption.

Getting physical with digital

For example, one of the largest banks in Scandinavia, understanding the business potential of advanced analytics, big data, and better data management to improve fraud detection and prevention, ATM location, and other initiatives, was eager to begin its digital data journey. It was facing intense competition, and was considering making a massive, multimillion-dollar investment in its IT and data architecture.

A lot was riding on what the bank decided to invest in, where it decided to invest it, and how.

It began by identifying key use cases that reflected the organization's most compelling strategic requirements: improved fraud detection, optimized location and allocation of branches, and more granular customer segmentation.

Based on this determination, we helped the bank outline a target architecture, founded on the best-practice reference model, that would enable the capabilities the bank desired and assess available solutions. Instead of ripping out its entire IT infrastructure, the bank decided to add a single Hadoop solution that allowed for storage and distributed processing of the bank's extremely large and frequently unstructured data sets across thousands of individual machines. This was especially useful in scaling the bank's high-frequency requirements for its online fraud-detection processes.

For branch location, allocation, and optimization, a Hadoop data lake (a management platform that processes flat, nonrelational data) used the bank's geospatial and population-growth data to determine where best to locate new branches and ATM machines. To improve its customer segmentation, the bank tested a new customer algorithm on the Hadoop database before rolling it out on its legacy data warehouse. This eliminated the typically costly and time-consuming back-and-forth process of develop, pilot, assess, validate, tweak, and pilot again that characterizes traditional data developments.

In this way, the bank achieved its primary business goals. It added new, differentiating capabilities, such as real-time analytics, and created real enterprise value with a relatively small technology investment, not the massive one originally contemplated. This was achieved by deciding what to invest in, where to invest it, and how—*before* buying systems and software that might not have served it nearly as well. Crucially, instead of first buying the technology, the bank built an in-house analytics team, skimming off the cream of the local talent in the process.

Today, the bank is considered the leader in financial analytics in its market and sells analytics services to other financial institutions.

The bank knew that the time was ripe to get serious about digital transformation, made it a priority, and in doing so achieved what may well be an enduring competitive advantage, all without disrupting its business with a big-bang technological transformation. It started with a clear view of its business goals, kept them front and center, and created a two-speed data architecture that worked.

The lesson here is that for many companies, it is both doable and cost-effective to add analytics capabilities to an existing IT environment. But that requires a sound data architecture, and a well-grounded approach to data management.

Sven Blumberg is a partner in McKinsey's Düsseldorf office and Oliver Bossert is a senior expert in the Frankfurt office, where Hagen Grabenhorst is a consultant, and Henning Soller is an associate partner.

