A New Level of Corporate Data Management and Analytics

December 2022

This report is a collaborative effort by the members of the Data Management and Analytics Community (DMAC).

Editor: Marius Schmid
# Table of Contents

1 Introduction .......................................................... 2

2 Strategic Aspects ....................................................... 3
   2.1 Cloud as a Strategy .............................................. 3
   2.2 Becoming Strategically Agile .................................. 4
   2.3 Becoming a Data-driven Organization ....................... 4

3 Architectural Aspects ................................................. 6

4 Organizational Aspects .............................................. 7
   4.1 IT-Business Collaboration ...................................... 7
   4.2 Funding Logics .................................................. 8
   4.3 Data Governance ............................................... 9

5 Conclusion ............................................................. 10
1 Introduction

Financial organizations face a unique combination of challenges in corporate data management. External factors like regulatory requirements and evolving customer needs call for flexible and cost-efficient solutions. Simultaneously, financial organizations face an enormous business complexity that bears resulting IT complexity challenges, whilst attempting to adopt new technology stacks that enable additional business value. The large amounts of (newly) available data (e.g., geospatial data) could be leveraged in different ways, but legacy systems (e.g., transactional systems) that have historically evolved also need to be accounted for. Last but not least, increasingly visible decentral and agile approaches both in software development and on the business side need to be aligned with the way data and analytics are managed in such traditional, centralized organizations. Thus, financial organizations and their IT departments must deal with several threats and opportunities that are challenging to oversee and manage.

In an effort to help addressing these ongoing challenges, the Data Management and Analytics Community (DMAC) was established back in 2012. DMAC is a loose collaboration of large European banks organized and moderated by the Institute of Information Management at the University of St.Gallen (IWI-HSG). DMAC’s aim is enabling its members and guests to share insights on “eye level” and consolidating effective practices for guiding the development of their corporate data management organization and landscape. This entails business opportunities and threats in the financial industry, as well as relevant technological and other industries’ trends. Participants hold various senior management positions in data management-related fields (e.g., Head of Data Warehousing, Chief Information Officer, Chief Data Officer) across several large European banks. The community gathers multiple times per year in workshop formats, fostering understanding, reflection, and exchange between participants. Over the last ten years, DMAC was supported by Commerzbank, Credit Suisse, DNB, DZ Bank, Erste Bank and UBS.

This report summarizes selected key insights derived from the longstanding exchange in the DMAC community as well as additional interviews with participants, drawing up an outlook of an emerging new level of corporate data management and analytics. It should be noted that discussions within DMAC are quite detailed and traditionally open, but cannot be presented here in detail for confidentiality reasons. The sharable insights are divided into three key focus areas: Strategy, architecture, and organization (see Figure 1). Though portrayed separately, we observe a high interrelatedness across these areas.
2 Strategic Aspects

Strategy is a crucial driver for the successful transition of corporate data management and analytics. Key activities can be divided into a strategic focus on cloud technologies, efforts to become strategically agile and the transition towards data-driven organizations.

2.1 Cloud as a Strategy

Facilitating the adoption of cloud technology is expected to allow large banks to better prepare for changes that occur in the current, increasingly fast-paced competitive environment. Consequently, the shift from on-premise solutions towards cloud-based offerings goes beyond technological upgrades and is seen as an integral part of a future-proof strategy. Primarily, cloud technology promises to deliver enhanced performance and scalability by enabling the use and integration of additional hardware, software, and data that would be difficult and costly to provide through in-house development. By accessing a variety of easily integrable tools from external cloud vendors, banks can more quickly and with less organizational overhead leverage the newest technologies, e.g., in the field of artificial intelligence. Simultaneously, cloud technology scales easily and enables synergies by providing a unified infrastructure. In the current data management and analytics world, different infrastructures co-exist. Integrating software tools in such an environment is a lengthy process. By transitioning to the cloud, several administrative tasks and data engineering-related tasks could be delegated to the cloud vendor. This leaves the IT department of banks with fewer integration tasks and less lifecycle management. Instead, they can focus more on requirements engineering and operational aspects, increasing speed and flexibility of solution delivery.

The cloud transition is also associated with several challenges, however. For example, adopting cloud technology bears a fundamental shift in the banks’ cost logic. Rather than the established fixed cost models with flat capacities, the cost structure on the cloud depends on usage. Currently, usage is not a relevant metric for calculating costs, particularly for users on the business side, making it difficult to predict the implications of a cloud transition. A usage-based cost logic will therefore require a significant cultural shift across the organization. Additionally, governance on
the cloud (e.g., data protection, secure integration with on-premise systems) is a critical aspect that is currently not well covered by providers and will likely require substantial effort and expenses by each bank. In general, it is therefore not clear whether the cloud transition will actually result in overall cost advantages compared to the current, on-premise environment or not. Across DMAC, companies are currently defining which aspects of their infrastructures can run on the cloud, with first migration projects underway. In this regard, the tension between what has been done historically and what is desired in the future is observed throughout. For Swiss banks in particular, the choice of a cloud vendor hinges on the fact that they have data centers in Switzerland. Aside from the technical and organizational implications, the regulatory boundaries therefore also complicate the transition to the cloud. Although cloud is evidently a central element of strategy, it is therefore not intended to shift all existing components to the cloud. Creating hybrid models (with certain aspects remaining on-premise, e.g. data warehouses) appears to be a more realistic strategy.

2.2 Becoming Strategically Agile

Next to the cloud transition, becoming strategically agile presents a central goal for all organizations. We observe that the accelerated digitization has become even more central during the Covid-19 pandemic. Thus, rather than continuing to think of agility in terms of setting up dispersed scrum teams, it is about becoming strategically agile as an organization. The idea is also to foster and leverage co-creation amongst business and IT. Strategic agility is associated with decreased inefficiencies and leaner controls, particularly regarding administrative tasks. In addition, instead of having steep hierarchies across several levels, the objective is to reduce management layers. Spotify’s approach of organizing according to value streams appears to act as a role model in this regard. In this model, flat capacities are defined in terms of funding – and the business side prioritizes what is most important to them. This enables organizations to avoid tedious funding discussions and organize in a more decentral way. In general, there is thus a shift away from the rigid cost focus. Such a fundamental shift in the structure of the organization and existing processes also requires a significant mind shift in the organization. At the same time, the regulatory aspects faced by banks impede the flexibility in trying to become more strategically agile.

"The high-level goal of many organizations, including our bank, is to become strategically agile, effectively reaching the next level in terms of digital maturity."

(Christoph Müller)

2.3 Becoming a Data-driven Organization

Regarding the overarching strategy of the organization, the phrase “data is the new oil” appears frequently. The bank’s business units generally see the value of data and want to leverage it – but how to actually achieve it often remains unclear. There appears to be a lack of understanding across the entire organization of what analytics actually could afford and does afford. One central element of realizing the value of data is putting it at the forefront of the organizational logic. Concretely, this means transitioning from an organization that thinks in and is driven by applications to a data-driven organization.
Achieving this transition requires both the technological foundation, as well as a significant cultural shift. In the current situation, data occupies a support role for business. In a data-driven organization, data becomes a key strategic driver and thus a crucial decision-making element in the organization. From an analytics standpoint, this promises to facilitate communication and scaling in the organization. Experts from the business-side collaborate with data experts in the early stages of projects already, preventing lengthy discussion at a later stage. Similarly, becoming a data-driven organization enables leveraging the open-source principle. Rather than having all technical know-how and expertise on programming tools remain on the IT side, the goal is to provide more of this to users (i.e., the business side), and allow them to develop additional tools and services themselves. This way, IT becomes less of a bottleneck. On the contrary, providing more access to users increases governance requirements to avoid uncontrolled development in the organization.

Despite the value of data, a critical challenge that remains is getting large amounts of consistent data in the organization. This is due to the diverse systems in place, with some leveraging real-time data with streaming whilst others are based on batch processing. In addition, in the context of banks, accessing and using customer data is a critical issue and customers need to agree to host their data in the cloud, for example. Establishing a data-driven organization is thus associated with several challenges and trade-offs.

“I think many people don’t understand what data centricity really means. [...] Many organizations have not yet managed to define a proper data and analytics strategy. [...] Nobody is denying that effectively using available data will become increasingly essential. The bigger issue lies in finding the right approach and the correct sequence of steps to actually implement this in our organization.”

(Florian Bleier)

Figure 2 summarizes the key strategic aspects determined across the community.
3 Architectural Aspects

Implementing the various strategy elements requires a fitting corporate data management architecture. We collected several aspects of a desired state, as well as the described challenges for moving towards this state.

On the one hand, the sources of data that need to be processed, will remain diverse and thus the future corporate data management architecture will need to account for such diversity. More specifically, the foundation layer of a future architecture must enable to feed data in various forms and from different sources. In this regard, it is also desirable to shift infrastructure elements to the cloud provider and leverage the more flexible pay-per-use pricing models. On the end of the users, it is important to leverage more tools and functionalities. Thus, the architecture must be open to the outside and facilitate the integration of such external tools. In the current on-premise world, such a desired openness is often times simply not possible.

Aside from being more open, there is also an increased demand for real-time data, driven by customer needs. Thus, the architecture must account for streaming functionalities that enable (near-) real-time processing, for example in the trading sector. Multiple organizations also run Data Science Labs, offering “micro-services” to users, for example. Leveraging additional tools also requires more APIs to define clear interfaces between the user side, i.e., applications, and the underlying data. In between the foundational infrastructure and the user-side thus lies an additional layer of data consolidation. Multiple banks leverage data virtualization to prepare the source data for various forms of data usage. Here, different levels of data curation are becoming increasingly relevant. Data can be more or less curated not only technically, but also functionally – depending on the given use case.

Despite these different layers with accordingly diverse functionalities, the goal is to present a unified view to users, i.e., one ‘seamless’ analytics ecosystem. In this way, users have a single point of contact for their requirements and suitable resources can be determined by the banks’ data management and analytics team. The vision is to implement use cases according to their technological fit, rather than making purely cost-driven decisions. Figure 3 illustrates such an ecosystem and its constituent components.

![Figure 3. An Ecosystem of Systems](image)
Establishing such an architecture is accompanied by several key decisions and challenges, however. For the existing systems, it must be defined which components need to exist in different technologies, e.g., on-premise and cloud, and what can be unified into a single source. For the migration to the cloud, banks face different scenarios dependent on the existence and characteristics of the legacy systems, namely building solutions directly in the cloud, i.e., cloud-native, building solutions on-premise and then moving them to the cloud, or applying a hybrid approach. Thus far, no clear favorite or best practice has been established.

Another crucial challenge is guaranteeing data consistency in an architecture that enables real-time data and continuously loading data. Establishing a single source of truth for a specific point in time is non-trivial. In a scenario where data is curated at different levels, the question is also what level of curation is needed for what kinds of requirements, who may access what kind of data, and who is responsible for the actual curation and maintenance. In general, the architectural design cannot be detached from governance-related aspects. This becomes increasingly challenging as more functionalities are integrated. Aside from the systems that exist within the premise of the banks, legacy systems (e.g., ERP systems) may also present a critical issue. As legacy systems become more intelligent, how can generated insights be reused and customized, and how open will these legacy systems be to open-source software? Thus, architecturally speaking, the organizations are also dependent on the developments and possibilities of connected, external systems.

“The goal of the new digital core is to provide a central interface that makes data available to be used in operative processes and in analytical reports. This is supported by the creation of a new center that focusses on data management and provision in our organization.”

(Markus Ecker)

4 Organizational Aspects

The discussed future architectural foundation of corporate data management and analytics also heavily impacts organizational aspects. Key aspects include the future collaboration of IT and business, funding logics, and data governance.

4.1 IT-Business Collaboration

In terms of the collaboration between business and IT, different approaches can be observed. Some banks have data scientists situated strictly on the business side, whilst others have data scientists both on the IT side and on the business side, for example. Having data scientists positioned on the business side means they have a thorough understanding of the business, but may have more difficulties in finding and accessing the right data. IT then takes on a data engineering role to assist the business side in finding the needed data. Thus, IT offers consulting services to the business side. Ideally, as discussed with regards to architecture, the IT side is unified into one (analytics) ecosystem which the business side approaches with their given requirements. The IT side then functions as a gatekeeper to ensure quality, security, and consistency and elaborates how best to address the given requirement. Across organizational units, tensions are observed between the demands by data scientists, who want to access data immediately and freely, and the formal control processes implemented to ensure the mentioned quality, security, and consistency.
Determining access rights and defining efficient control processes is thus a crucial aspect for IT-business collaboration. In future setups, it is conceivable that developed solutions by business users may be made available for usage to other users and business units. However, this would entail a significant additional responsibility of maintaining such solutions beyond one’s own scope and usage. The business side generally refrains from such commitments and prefers providing services for themselves. Establishing a technological and cultural foundation where business side users are willing to co-create among each other is thus a central challenge of adapting the way IT and business collaborate.

4.2 Funding Logics

Another organizational aspect that is fundamentally impacted by the discussed strategic initiatives is the funding logic. This is both with regards to budgeting and internal cost management or pricing. Historically, the IT side has commonly been funded through projects. Typically, business cases must be developed to justify budgets and realize projects. In the transition phase, organizations also offer a central budget for which departments can apply with use cases. Adopting new organizational forms like value streams with flat capacities completely changes the underlying funding logic, however. Here, instead of having profit- and cost-centers, an owner on the business side can prioritize what is most important to them and allocate budget from the flat capacity accordingly, making the budget allocation more flexible overall. On the other hand, the cloud transition promises a faster time-to-market and thereby decreased development costs, i.e., budget needs to be allocated in smaller portions.

Part of the funding logic is also how cost management and pricing models will evolve in new settings. Although cloud vendors offer more flexible usage, this flexibility also yields dynamic pricing, as discussed earlier. On the business side, users are accustomed to predictable costs, which is not the case in a pay-per-use environment. This could enable cost savings, but also significantly ramp up costs. With the current cost control, the risk resides with the IT side as such costs would directly affect their budget. Thus, control mechanisms will need to be introduced to reduce the risks of exorbitantly high costs. The transition to the cloud will therefore require generating an understanding across the organization for dynamic pricing as opposed to fixed costs and adapting cost control accordingly. Although these discussions are at an early stage, it is conceivable to create new internal pricing models, for example licensing fees, for the provision of internal IT-related services. Here, banks face a special situation, as they are divided in naturally more data-provisioning (e.g., capital market) and then more data-consuming divisions (e.g., wholesale). The added value of such models is not inherent and must be clear, however.

“A large challenge is related to the evolution of the IT budgeting model. Currently, we are not accustomed to “per-use” budgeting, so this is a completely new model that needs to be considered carefully.”

(Dr. Birgit Reischl-Lenz)
4.3 Data Governance

Aside from the funding logic, data governance is a critical organizational aspect. In terms of team setups, it is important to leverage diverse skill sets. Teams are commonly structured to include IT-related knowledge (e.g., IT engineers), business experts, and data scientists. When these teams deploy new solutions, i.e., micro-services, the responsibilities for maintenance need to be determined, however.

In the application-driven world, organizations commonly manage their security concepts with a “need to know” principle. Determining data access for the business side is crucial. Particularly data scientists are keen to access as much data as they deem necessary, but in the regulatory context of banks, data access must be strictly managed. Passing data from an operational application to an analytical application thus requires a formal process. For Swiss banks, the geographic location of developers is decisive for their access rights, for example. Particularly in the cloud, governing customer data is a key challenge, as elaborated earlier. Next to the regulatory aspect, it is also a technical question of whether users are even able to understand and use different data.

“We are in the process of providing access to the data in a way such that domain specialists and data scientists can more easily determine which data can be sourced in which quality and from where.”

(Martin Bilgeri)

Another crucial challenge related to data governance is defining responsibilities. In a world where solutions can be flexibly deployed, it must also be determined who is responsible for maintaining such co-created solutions. From an architectural perspective, it is about defining what aspects should be managed globally (i.e., enterprise-wide) and which aspects can be managed locally. Historically, the silo thinking brings the advantage of not being dependent on other departments in the usage and maintenance of solutions. Here, responsibilities are clearly defined. However, consistency suffers and applications are supplied with individual data. Thus, the organizational efficiency could be increased by managing certain aspects on an enterprise level. Creating incentives to take on responsibility for governing developed solutions will thus be a critical challenge for the future of data management in analytics. It may also entail defining entirely new roles strictly dedicated to data governance.

To summarize, Figure 4 depicts the key organizational aspects.

![Figure 4. Key Organizational Aspects](image-url)
5 Conclusion

In summary, we observe three main areas that are critical in the new level of corporate data management and analytics, namely strategy, architecture, and organization. All three areas are evidently interrelated and must be viewed as such. The transition to the cloud is a strategic goal that requires both a technological foundation in the data architecture as well as a cultural shift in the organization, e.g., with regards to the underlying funding logic and governance mechanisms. Similarly, becoming a data-driven organization hinges on a shift in the understanding of the role of data in the organization, next to enabling technical capabilities like near real-time processing. The transition process is accompanied by several challenges, like defining which parts of the architecture to offer as enterprise-wide vs. locally management components or developing internal pricing mechanisms for the new variable cost logic that comes with the shift to cloud technology. We observe these challenges across all DMAC banks as well as many guest participants. Overall, these issues are becoming notably more relevant for the overarching corporate strategy. The next level of corporate data management and analytics is thus at an early stage, yet clearly underway.