

— Point of View Data-Centric Company



Table of Contents

1 Executive Summary	3	3.4 Data-centric Approach to Business Strategy Development	15	6 Technical prerequisites for Data-Centricity	31
2 Introduction	5	3.5 Recommendations for Strategy in a Data-centric Age	16	6.1 Autonomy and Data Democratization in a Data-centric Enterprise	31
2.1 What Are Data?	5	4 How Data-centricity Improves Market Position	17	6.2 The Infrastructure Prerequisites for a Data-centric Company	37
2.2 Categories of Data	6	4.1 Data-centricity: An Increasingly Critical Factor in Market Competition	17	6.3 Recommendations	39
2.3 Three Levels of Data Integration in Companies	6	4.2 A Widening Gap: How Data-centric Market Leaders Gain a Decisive Competitive Edge	17	7 Compliance	40
2.4 Challenges and Best Practices of Becoming Data-Centric	8	4.3 Recommendations for Marketing in the Data-centric Age	20	7.1 Introduction	40
2.5 Launching the Process of Building a Data-centric Organization	9	5 Data-centric Corporate Functions	24	7.2 Application of Regulatory Statutes Poses Major Challenges	40
2.6 Recommendations	9	5.1 Finance	24	7.3 Development of a Compliance Framework	42
3 Strategy in a Data-centric Age	10	5.2 Human Resources	26	7.4 Challenges	44
3.1 Strategic Relevance of Data	10	5.3 Research and Development	28	7.5 Result	44
3.2 Strategy in Data-centric Companies	11	5.4 Recommendations	29	7.6 Recommendations – User-centric Approach to the Legally Compliant Use of Technologies	44
3.3 What to Include in a Great Data Strategy	13				

_ 1 Executive Summary

Data serve as the foundation for all decision-making processes that, freed from the confines of silo views and with a 360-degree perspective, can improve products and services and even result in completely data-based business models. Past research has demonstrated that data centrality leads to better performance and productivity in tandem with higher market value. Moreover, the quality and speed of the decision-making process, enriched by higher development productivity and customer insights, can attain a higher level.

Today's companies find the wealth of the data available to them growing at an exponential rate as the collection of the information never ceases.¹ Yet the number of DAX companies (seven of the 40) that have designated a position of chief data officer ("CDO") for the management of their data remains low.² Data strategy begins with a data-centric mindset and incorporates data and analytics into all corporate functions; the company's ability to meet constantly changing business and market

conditions in the data-centric age is sharpened and value for the business as a whole is generated. Becoming a truly data-centric company demands an emphasis on data as the core element of all operations; they must be regarded the starting point for making decisions and the optimization of processes and essential components of new products and services.

As of this moment, comprehensive data-centricity can translate into a competitive advantage, although the scope can vary from one industry to the next. Still, it is now a critical factor in market competition and a key factor in future business models. Marketing activities such as price, promotion, place, and product are being adapted to assure data-based optimization of services in various industries. Data-centricity will enable companies to identify fields where they can obtain meaningful insights that will make a significant difference in their market placement and customer base.

Finance departments pursuing a data-centric approach establish a position of strength as they are able to assess data value and share the insights and recommendations they have derived from the data with senior management and other departments within the organization, staking out their role as custodians of enterprise data and the ultimate authority for analytics.³ The rise of big data and the advent of other data-related technologies in the past decade have brought about significant changes for human resources departments as well. The amount of information related to employee activities in an organization and stored by Google, Microsoft, Amazon, and Facebook has reached a staggering volume of 1,200 petabytes.⁴ R&D, with its focus on process innovation development, the updating of existing products,

1,200 petabytes
of information.

¹ The technology section in chapter 3 will touch briefly on the means for securing the highest data value.

² Detecon Research

³ <https://hbr.org/resources/pdfs/comm/workday/FinanceData.pdf>

⁴ How Much Data Is Created Every Day? [27 Powerful Stats] | SeedScientific

and new product research and development, is another department enjoying huge potential. Redefining the R&D process in a dynamic decision loop that directly analyzes the market/customer data for updates to the product/service is not possible without a data-centric approach in R&D.

The democratization of data access inevitably demands changes in the way people think about the handling of data. In short, data should be treated as a product that is ready for use and reliable. The data mesh platform is an internationally designed and disseminated data architecture with centralized governance and standardization to ensure interoperability and a shared and harmonized self-serve data infrastructure. Rather than applying top-down decisions regarding the formats in which data are stored for future users, the information is stored in its original form so that future users are completely free to decide what transformations will best serve their needs. The solution desired by most organizations is a one-stop platform for the performance of analytics and the development of meaningful insights

without that do not require the assistance of a central IT team. The central IT team still exists, but it does not own the data.

The general impression appears to be that strict laws would cause Europe to fall far behind China and the United States as such regulation would prevent the data that are or could be produced from being used for data-driven innovations. The assumption is that there would be few data-driven innovations coming out of Europe in future. Detecon's analysis has identified three problem areas for companies that have had to abandon data-driven innovation projects involving new technologies:

The solution most companies want is a **one-stop platform.**

1. The impact and requirements of the applicable data protection regulations were not included in the development from the outset.
2. Internal compliance processes are found to be static, complicated, and time-consuming and require extensive resources, sometimes from outside sources, for audits and consulting services.
3. Some companies prohibit as a general principle the use of certain technologies such as machine learning or the exploitation of their own data because of the complexity and ambiguity of interpretations.

Authorities have only recently begun to enforce fully the penalties established in the GDPR, actions that can have far-reaching consequences for both companies and individuals. Whatever approach is taken, we believe that the use of implementation and translation tools for laws and regulations will enable users to develop legally compliant, data-driven innovations that will create future security for entrepreneurial activities.

_ 2 Introduction

2.1 What Are Data?

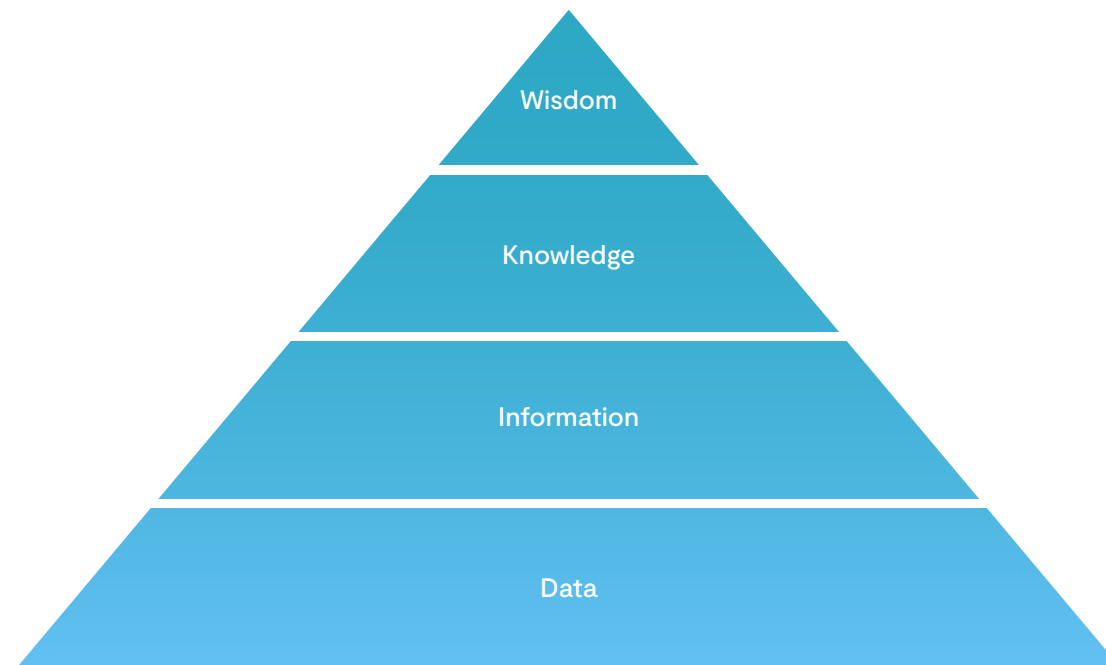
The Oxford Learner’s Dictionaries (n.d.) define data as “facts or information, especially when examined and used to find out things or to make decisions.”

The data-information-knowledge-wisdom hierarchy (DIKW) depicts data as the bottom level of the pyramid. The processing of the data produces information.⁵

In turn, the analysis or interpretation of information creates knowledge that, when applied, results in wisdom.⁶ Naturally, all levels can and should be generated to one degree or another in every company.

Companies have acquired enormous amounts of data that can prove to be useful and valuable, but they must learn how the conduct of analyses or modeling can identify the use and value generation from this store-house.

Figure 1: Levels of the DIKW Hierarchy⁷



⁵ (Ackoff, 1989, p. 3) – Ackoff, R. L. (1989). From data to wisdom. *Journal of applied systems analysis*, 16(1), 3–9.

⁶ (Ackoff, 1989, p. 3; Kitchin, 2014, p. 9–10) – Ackoff, R. L. (1989). From data to wisdom. *Journal of applied systems analysis*, 16(1), 3–9. AND Kitchin, R. (2014). *The data revolution: Big data, open data, data infrastructures and their consequences*. Los Angeles: SAGE Publications.

⁷ Adaption from Rowley (2007, p. 164) – Rowley, J. (2007). The wisdom hierarchy: representations of the DIKW hierarchy. *Journal of information science*, 33(2), 163–180.

2.2 Categories of Data

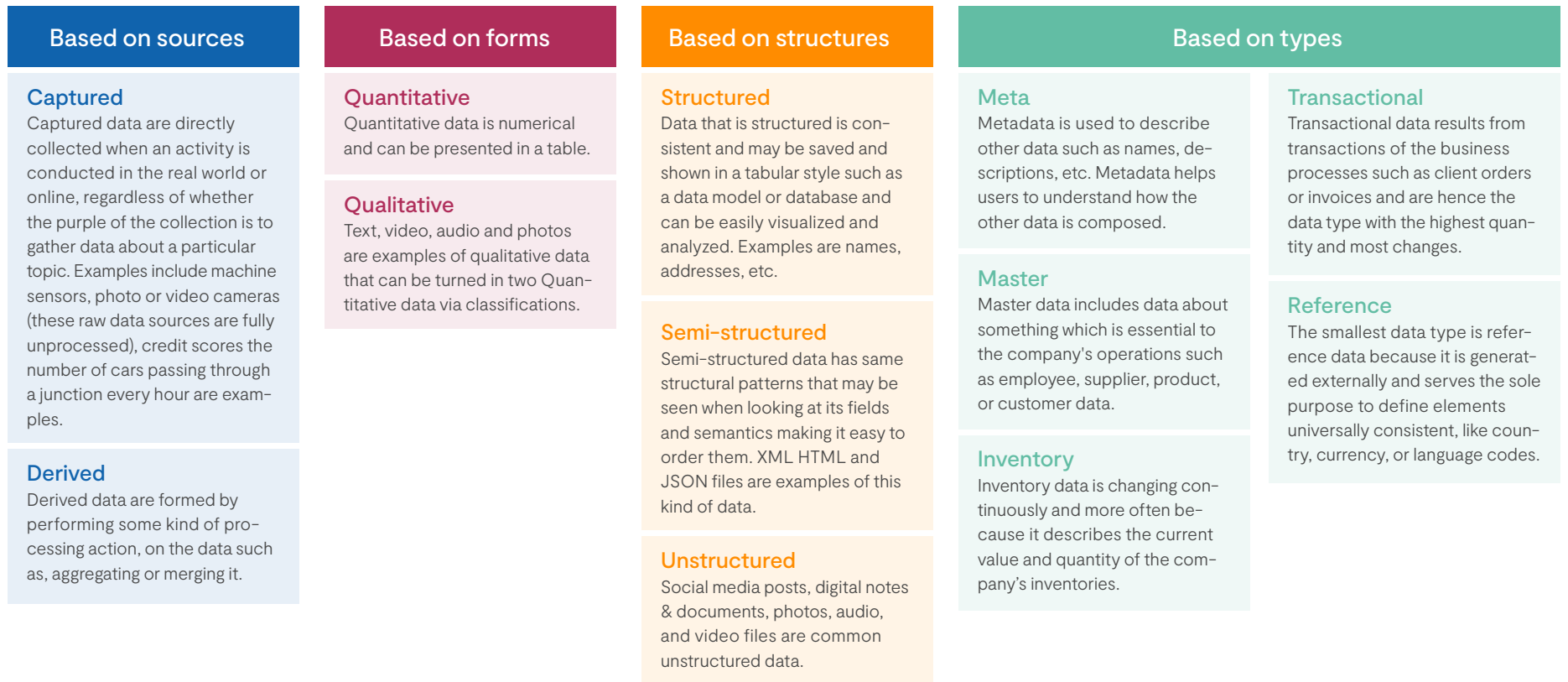
Data can be categorized from a company perspective according to the categories shown on the right.

2.3 Three Levels of Data Integration in Companies

Schloo⁸ identifies three levels of data integration i.e., data-informed, data-driven, data-centric.

The least complex level is data-informed as it uses solely descriptive analytics based on historical data⁹ for the creation of dashboards. There is no central data strategy or the use of data science to carry out advanced analytics such as predictive, prescriptive analytics.¹⁰

Figure 2: Categories of Data



⁸ (Schloo, 2021, p. 43) – Schloo, B. (2021). Data-centricity: A framework approach to assess the value of data for companies (Master Thesis, htw University of Applied Science).

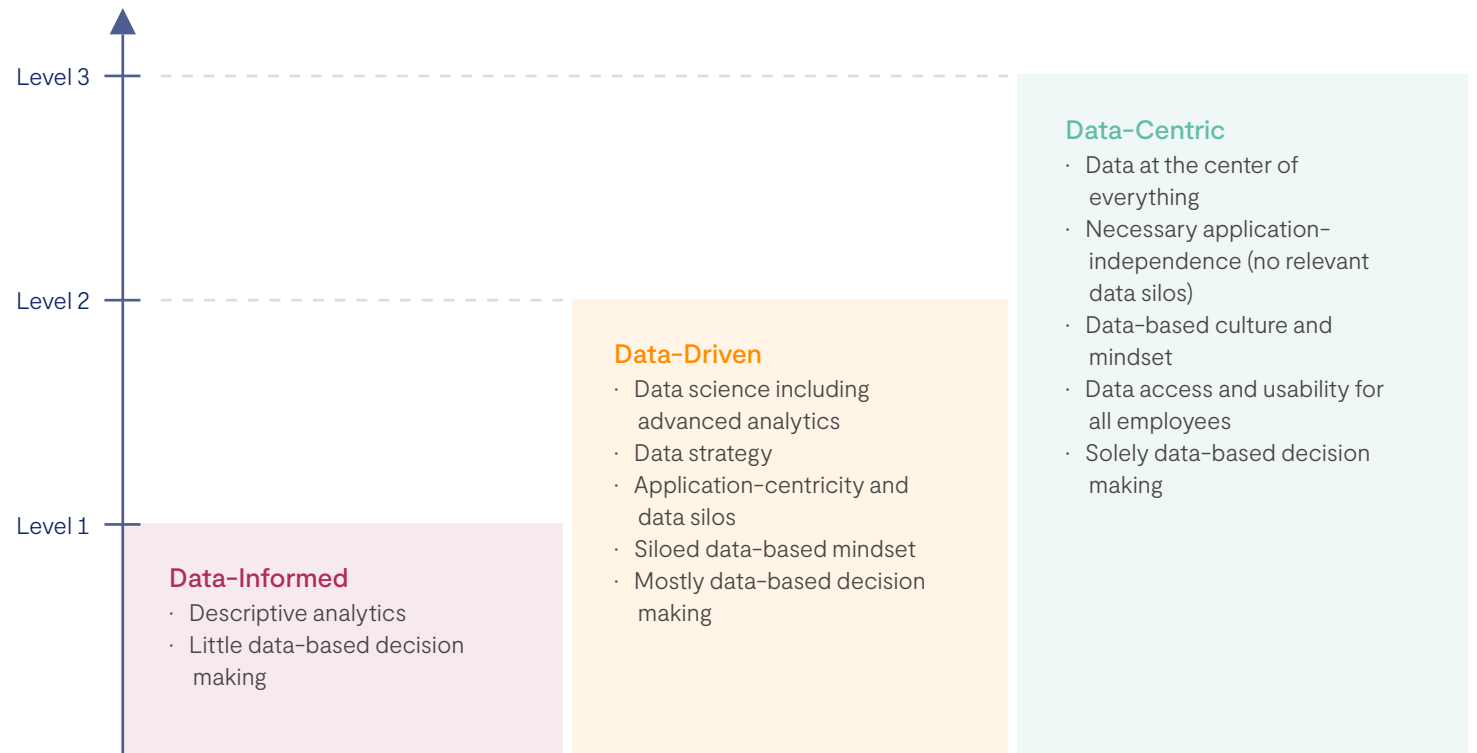
⁹ (Berndtsson et al., 2018, p. 1) – Berndtsson, M., Forsberg, D., Stein, D., & Svahn, T. (2018). Becoming a data-driven organisation. In 26th European Conference on Information Systems (ECIS2018), Portsmouth, United Kingdom, June 23–28, 2018.

¹⁰ (Berndtsson et al., 2018, p. 1). – Berndtsson, M., Forsberg, D., Stein, D., & Svahn, T. (2018). Becoming a data-driven organisation. In 26th European Conference on Information Systems (ECIS2018), Portsmouth, United Kingdom, June 23–28, 2018.

A data-driven approach utilizes advanced analytics (predictive and prescriptive) and has established or is in the process of establishing a data strategy. The principal idea behind data-driven integration is the use of data for making decisions with the aid of an overarching data strategy and data science.

Data-centricity is the final and highest level of data integration, the stage at which data are the heart and soul of everything. In a data-centric system, data serve as the foundation for all decision-making processes that, freed from the confines of silo views and with a 360-degree perspective, can improve products and services and even result in completely data-based business models. From a technical perspective, this implies the existence of an architecture that is independent of the application and that collates all the data into a single model and the same format so that they can be used and combined easily. From an organizational perspective, the transformation to data-centricity is a change process during which employees must be motivated, educated, and trained to support the establishment of a data-centric culture and mindset.

Figure 3: Level of Data Integration in the Company¹¹



¹¹ Schloo (2021, p. 43) – Schloo, B. (2021). Data-centricity: A framework approach to assess the value of data for companies (Master Thesis, htw University of Applied Science).

2.4 Challenges and Best Practices of Becoming Data-Centric

Specific examples the challenges that may arise are available to support companies during the process of their transformation into data-centric operations. Experts speak of grouping these challenges into two categories: “data and general technical challenges” and “organizational challenges.”

The lessons learned and best practices mentioned by experts as **success factors** during the process of transformation into a data-centric company can be grouped into **general, technical**, and **organizational** lessons learned.

The challenges, lessons learned, and best practices are summarized in the table on the right.

Table 1: Challenges and Lessons Learned and Best Practices of Becoming Data-centric¹²

Category	Challenges	Lessons learned and best practices
General	<ul style="list-style-type: none"> Shared data are not equivalent to equally accessible data – identification of easy-to-use mechanisms applied by engineers to make data visible, accessible, and usable while complying with data privacy regulations and avoiding bottlenecks 	<ul style="list-style-type: none"> Financial and personnel resources Transformation is a process, not a short-term project Establish an overall strategy/purpose and work, use case by use case; if an approach proves to be unsuccessful, drop it – “fail fast and fail often”
Technical	<ul style="list-style-type: none"> Lack of knowledge about existing data Insufficient data and/or poor data quality Identification of valuable data objects Efficient processes for storage, processing, and sharing of masses of data without any loss of speed in real-time data activities Diversity of data types and systems within a company in combination with varying requirements for data confidentiality 	<ul style="list-style-type: none"> Curated data teams and datasets to generate the same KPI values while avoiding duplication of activities and enabling self-service analytics Data catalog creating transparency about what exists where Save only relevant data as storage and maintenance cost money Invest in the harmonization of important data Data collection must be automated Automate compliant data access and deletion By default, data must be available to everyone. If not the case, explanation is required
Organizational	<ul style="list-style-type: none"> Organizational/mindset silos and change process Employees’ resentment or rejection of the change Employees fear loss of their jobs because of redundancy, especially when operations are automated Trusting the results of the data Power plays of individuals/departments unwilling to share the data Persuading all management levels to support the transformation Finding and training product owners and the links between the business and technical teams 	<ul style="list-style-type: none"> Centralized data governance and data ownership Need to involve all levels from the top in support of the overall strategy and from the bottom to secure full acceptance and obtain detailed insights about all units Identify pioneers (individuals from all units) so that their input and lessons learned can be utilized and their whole-hearted support is secured; they can be used to disseminate the full process through the entire organization Close and active collaboration among teams of business and technical employees Build-up of internal knowledge and simultaneously external (technical) expertise (if not already available) to accelerate the process Generate the willingness to make the general mindset change and to share data by demonstrating the specific benefits for the individuals and/or how they will be able to perform their duties and responsibilities more effectively Promote the use of data and the related tools by making their use indispensable

¹² Schloo (2021, p. 50) – Schloo, B. (2021). Data-centricity: A framework approach to assess the value of data for companies (Master Thesis, htw University of Applied Science).

2.5 Launching the Process of Building a Data-centric Organization

Past research has demonstrated that data centrality leads to better performance and productivity in tandem with higher market value. Moreover, the quality and speed of the decision-making process, enriched by higher development productivity and customer insights, can attain a higher level. What is more, data-centrality lowers hardware and software requirements by eliminating unnecessary applications, reduces the time and cost of integration data from new sources or a combination of various sources, and puts an end to duplication of the same work.

Essentially, any business that wants to become data-centric must consider various aspects that can be broadly categorized as business value or strategic, technical, and organizational concerns. Above all, it means generating business value by placing data at the center, creating a technical infrastructure that enables all employees to make data-based decisions from sharing and using the data, and promoting an organizational environment in which the way employees think is guided by a data-centric mindset.

Implementing data-centrality to some degree opens the door to the use and exploitation of company data. The factors that should be considered for the identification of potential value are economic, environmental, and social in nature while the ultimate goal concerns the economic benefits: increased profits and revenues, cost savings, and employee and customer satisfaction. In this respect, the company's vision and strategy should act as the underlying guiding principles and ensure that the identified use cases actually support the goals the company is striving to achieve.

2.6 Recommendations

- Shared data are not equivalent to accessible data; establish an overall strategy/purpose and work, use case by use case; if an approach proves to be unsuccessful, drop it – “fail fast and fail often”.
- Create a data catalog that creates transparency regarding what exists where.

- Curate data teams and datasets to generate the same KPI values, avoid duplication of activities, and to provide self-service analytics.
- Data collection must be automated. Automate compliant data access and deletion.
- Establish centralized data governance and decentralized data ownership.
- Identify pioneers (individuals from all units) so that their input and lessons learned can be utilized and their whole-hearted support is secured; they can be used to disseminate the full process through the entire organization.
- Build-up of internal knowledge and simultaneously external (technical) expertise (if not already available) to accelerate the process.
- Generate the willingness to make the general mindset change and to share data by demonstrating the specific benefits for the individuals and/or how they will be able to perform their duties and responsibilities more effectively.

Data collection must be automated.

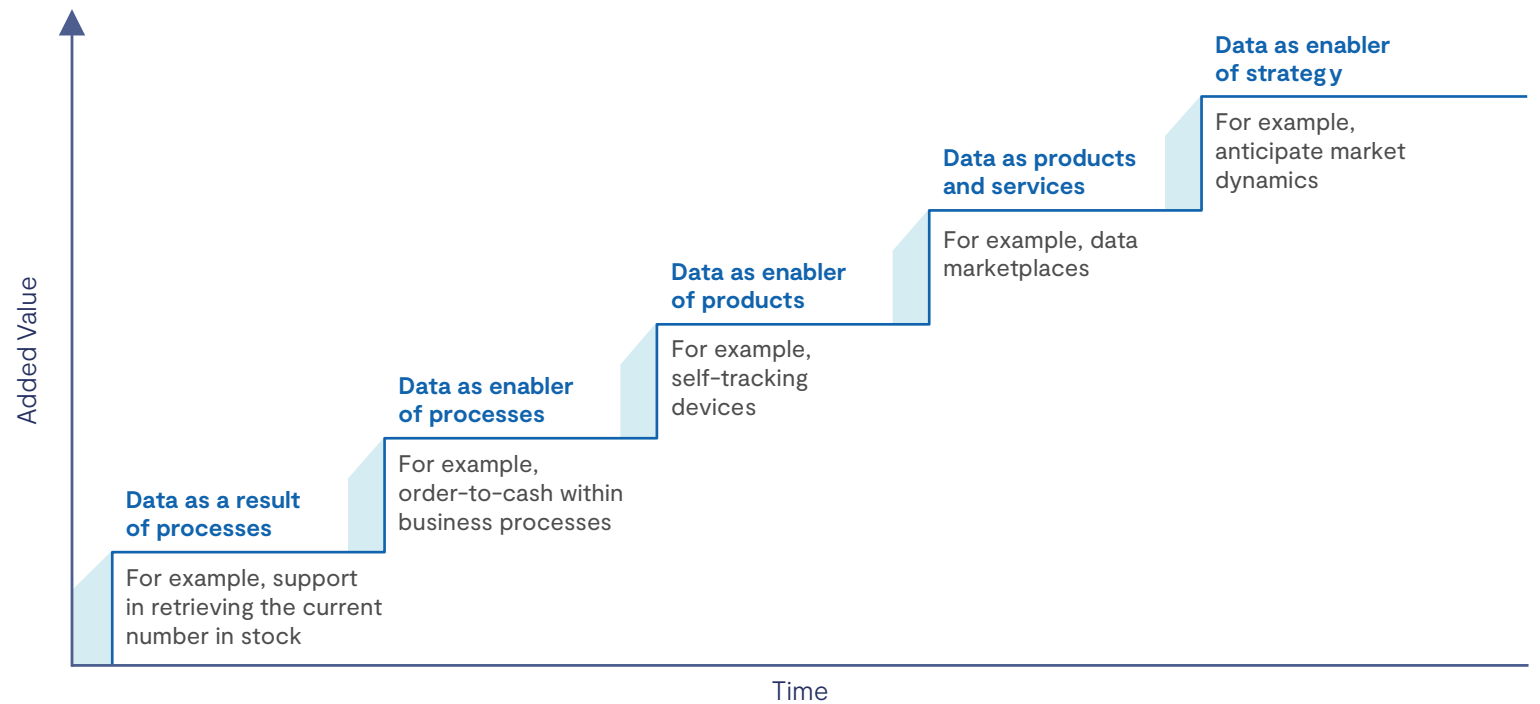
_ 3 Strategy in a Data-centric Age

3.1 Strategic Relevance of Data

Can it be a coincidence that no fewer than four of the five most valuable companies are established and recognized players in the digital world who exploit data as an important strategic asset and a key value driver? Hardly! No, we at Detecon, along with investors around the globe and many researchers,¹³ firmly believe that data-centricity, especially the use of data in strategic decision-making and the creation of value through data, lays the foundation for long-term and sustainable market success and a competitive advantage. A survey of more than 900 global business analysts found that 84 percent of the respondents believed it is essential to anchor key business decisions and strategies on the bedrock of data, yet half of them also indicated that their organizations do not at this time routinely use data to drive decisions and strategies.¹⁴

In Germany, just over half (22 out of 40) of the companies listed on the DAX have appointed chief digital (rarely: chief information) officers with the designated

Figure 4: The Evolution of Data (adapted from: Gür & Spiekermann, 2020)



¹³ (Gür & Spiekermann, 2020) – https://www.isst.fraunhofer.de/content/dam/isst-neu/documents/Publikationen/Datenwirtschaft/ISST-Report/Fraunhofer_ISST-Report_Data-Strategy-Praxis-Report.pdf

¹⁴ (Forrester, 2020) – Data-centric businesses are 58% more likely to exceed revenue goals | TechRepublic

role of helping to steer the organization through the digital transformation. The presence of such an officer is a good start, but the effectiveness of the position will be limited until data-centricity has been established throughout the entire company. Even fewer DAX companies (seven of the 40) have appointed a designated chief data officer (“CDO”) for the strategic management of data.¹⁵ SAP is one example of a company that argues strongly in favor of investing in such a position and for the CDO to be a “strategic decision for managing data effectively for business growth, realizing efficiencies, and minimizing risk.”¹⁶

Even though not all companies have reached the same level of maturity, technological progress has increased the potential of data to create value over the course of time as is illustrated in [Figure 4](#).

The Detecon point of view: In the years ahead, the leaders of data-centricity will dominate their industries and build up a sustainable competitive advantage while companies neglecting data-centricity will be left behind. In view of this inevitable conclusion, developing an excellent data strategy and realizing it throughout your

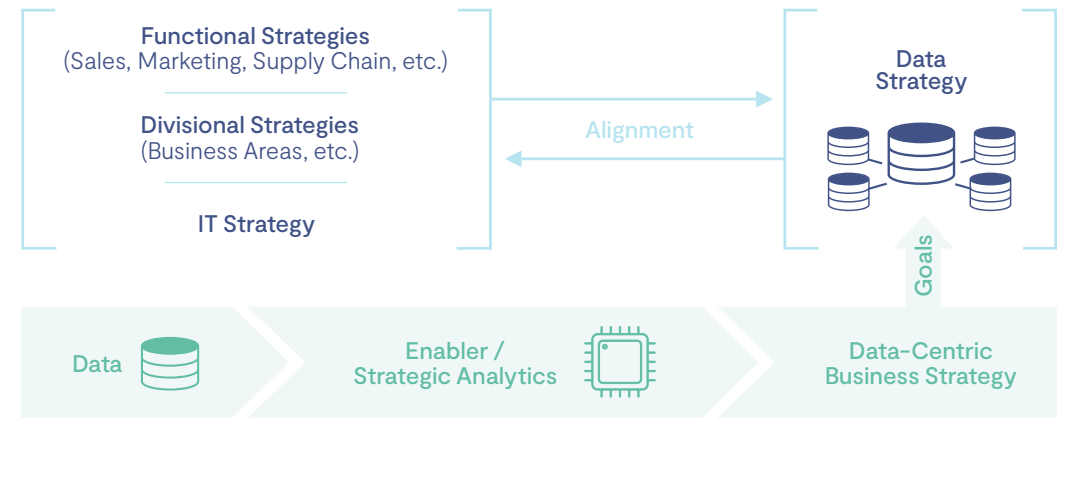
organization as the means of driving the generation of added value is essential if your company is to be steered to success! The transformation into a data-centric enterprise that also utilizes data as an enabler of the general business strategy is a key strategic decision demanding both top-level and company-wide commitment. Now is the time to make data a strategic issue!

3.2 Strategy in Data-centric Companies

At its very core, every strategy is about winning the game – so in our view, companies that neglect data cannot claim to be pursuing a winning strategy. The strategy in data-centric companies evolves in two principal ways, both of which push the generation of added value to higher levels (see Figure 5). The first essential step is to draw up a clear data strategy that is based on the outstanding goals of the business strategy. The

data strategy promotes the effective use of data in all business units and serves to drive the generation of added value, putting into companies’ hands the means to achieve data leadership and the related competitive edge. The second stage is the specific use of data as an enabler of the business strategy, a process also known as strategic analytics. A business strategy is decisive for

Figure 5: Strategy in Data-centric Companies



¹⁵ Detecon Research, 2022

¹⁶ SAP, 2020

growth, positioning, and focus, making it the primary driver of company-wide generation of added value. A data-centric business strategy is guided by strategic considerations based on decisions powered by data. Leveraging data when making decisions encourages more informed directional guidance and reduces the likelihood of bias causing errors in judgment, further facilitating generation of added value.

As the effective reallocation of resources during the strategy process is in most cases the single most power driver of revenue growth,¹⁷ a data strategy and the allocation of resources relating to data management must be firmly embedded in and explicitly described in conjunction with the business strategy. Very few companies, however, have reached the point of incorporating data into strategy. Below are the examples of two companies that have explicitly included data goals in their business strategy:

- **Allianz** “Digital for us means [...] enhanced data availability and analytics.”¹⁸ Allianz recognizes data and analytics to be important drivers for future success with the overall aim of implementing a data-first decision-making culture.¹⁹
- **BASF** “Digitalization is an integral part of our business. We want to significantly improve the availability and quality of our process data.”²⁰

These goals stated in the business strategy are the starting point for the development of a data strategy that must subsequently be aligned with other strategies (functional, divisional, or IT) with the ultimate aim being the establishment of a data-centric company.

Yet organizations often fail to define a profound, distinctive, and far-reaching data strategy.²¹ The GAS region is still in an early stage of development when it comes to data strategies as indicated by a survey of executives in

which 65 percent of the respondents noted that their companies do not have a data strategy, not even as one of the components of an IT strategy.²²

A study of more than 500 companies examining the use of data as an enabler of the business strategy found that organizations that treat data as a strategic asset and place them at the heart of the business strategy are more effective when deducing relevant strategic insights, enhancing business performance, and reshaping processes, products, and services.²³ The logical conclusion is that top-level executives should pursue a strategic analytics approach to their business strategy, one in which data analytics informs key strategic decisions. The output of the data-centric business strategy can be monitored by the use of interactive dashboards providing real-time information about key metrics. Such features also improve business agility and the ability to initiate successful and proactive strategic moves.

¹⁷ (Atsmon, 2016) – <https://www.mckinsey.com/business-functions/strategy-and-corporate-finance/our-insights/how-nimble-resource-allocation-can-double-your-companys-value>

¹⁸ (Allianz, 2022) – <https://www.allianz.com/en/about-us/strategy-values/strategy.html>

¹⁹ (Shahkarami, 2020) – <https://www.agcs.allianz.com/news-and-insights/expert-risk-articles/grd-data-management.html>

²⁰ (BASF, 2022) – <https://www.basf.com/global/en/investors/basf-at-a-glance/strategy.html>

²¹ (Gür & Spiekermann, 2020) – https://www.isst.fraunhofer.de/content/dam/isst-neu/documents/Publikationen/Datenwirtschaft/ISST-Report/Fraunhofer_ISST-Report_Data-Strategy-Praxis-Report.pdf

²² (Hewlett Packard Enterprise, 2021) – <https://www.hpe.com/de/de/newsroom/press-release/2021/10/umfrage.html>

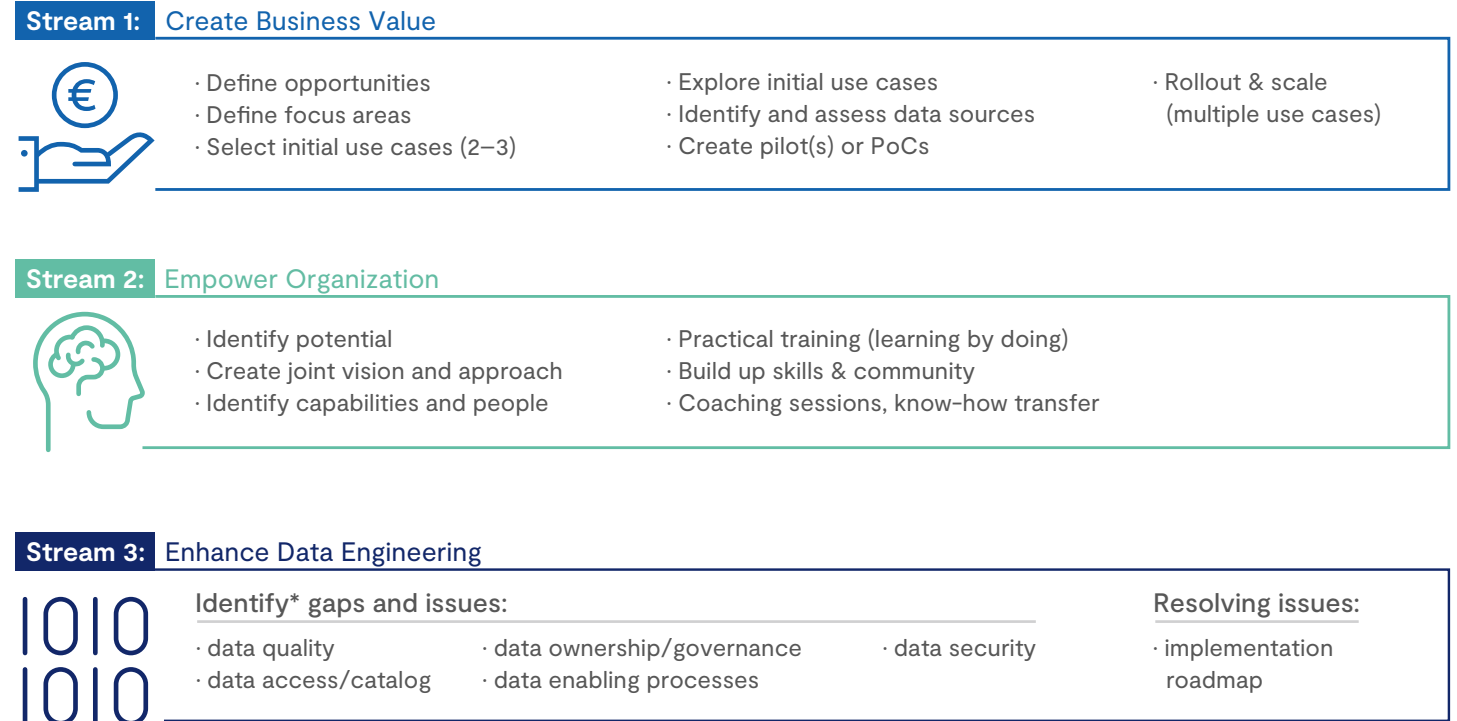
²³ (Stackpole, 2021) – Making the business case for a chief data officer | MIT Sloan

The Detecon point of view: A data strategy must define a timeline and the required resources in addition to the technical and support conditions for its realization and be successfully integrated into the company. The data strategies currently found in many companies and the related activities are often too low-level, too short-term, and too far removed from overall business strategies to create a data-centric enterprise.²⁴ If optimal interplay between business and data strategies is to become possible, data can no longer be considered separately. Moreover, in our VUCA world, it has become more important than ever before to make the right decisions quickly and readjust strategic directions whenever necessary. This leads us to expect strategic analytics and value-generating, data-centric business strategy to take their place as key success factors and competitive advantages for organizations in the coming years.

3.3 What to Include in a Great Data Strategy

A great data strategy begins with a data-centric mindset and integrates data and analytics into all corporate functions, laying the groundwork for the company's ability to meet constantly changing business and market condi-

Figure 6: The Three Streams of Data Strategy Development



* (from streams 1 and 2 and additional analysis)

²⁴ (Ladley & Redman, 2020) – <https://hbr.org/2020/03/use-data-to-accelerate-your-business-strategy>

tions in the digital age and to generate added value throughout the enterprise. While not everyone in the company – and beyond – may automatically become enthusiastic about data-centric initiatives, the executive leadership must instill a strong conviction of the necessity to launch the company on this transformative journey to data leadership. Success metrics for evaluation of the progress should be carefully considered with the aim of maintaining the motivation of all and of ensuring that the company is truly benefiting from the data-centric initiatives. Another important point is to specify the required resources and the timeline for reviewing the data strategy.

Starting from the goals derived from the data-enabled business strategy, the data strategy development process can focus on three key streams with the ultimate target of boosting value generation throughout the entire organization (see [Figure 6](#)).

The first stream is the foundation for the business logic of the data strategy as its focus is on the creation of business value. Quite frequently, a key strategic issue when implanting a data strategy is the lack of clarity

about the business value and the resultant lack of support from top management. The target in the “Create Business Value” stream is therefore to demonstrate clearly the business value of data. This is achieved by identifying specific opportunities and focus areas, exploring initial use cases, creating pilot projects of PoCs, and (in more advanced phases) the rollout and scaling of multiple value-creating use cases, just to mention some examples.

The second stream of the data strategy development seeks to empower the organization so that it can implement the use of data across all sectors of its business and achieve the ultimate goal of enterprise-wide value creation. The “Empower Organization” stream is important for the design of a data strategy because it addresses key strategic issues, including the lack of data skills in the workforce, the lack of clarity regarding employee benefits of a data culture, and the general resistance to cultural change. In the long term, cultivating a data-centric corporate culture presupposes the democratization of data intelligence by providing all employees – not just developers or data scientists – with the skills required to place data at the heart of decision-making and busi-

ness operations. This organizational enablement can be achieved by means of strategic change management initiatives arising from a company-wide vision and approach, initial capability assessments, and continuous skill development from the introduction of practical training programs, coaching sessions, and knowledge communities.

The third stream – “Enhance Data Engineering” – seeks to improve technical capabilities and resolve key strategic issues such as inadequate data quality, a lack of data access, and the unavailability of processing and analytics tool. The technology needed to attain the highest data value must be taken into account and the implementation road map for its realization must be prepared. The full length of the data value chain must be reviewed: the collection, management, transformation, visualization, and storage of data that will support data-centric value creation for the company.

Data thinking is a useful and innovative framework that can be utilized during the development of a data strategy. The framework combines data science with design thinking, so the focus is not solely on data-analytics technolo-

gies; it goes on to encompass the design of user-centric solutions with strategic business potential. Data thinking applies design thinking methods for the identification of customer needs and creative ways to find new solutions for the mastery of data-driven challenges.

3.4 Data-centric Approach to Business Strategy Development

By and large, strategists have rarely made any use of data and advanced analytics for the development of the business strategy,²⁵ and yet the right analyses offer potential for significant improvement of business strategies and strengthening of the executive leadership's position!²⁶ A data-centric business strategy based on strategic analytics offers tremendous opportunities for value creation and business agility to organizations.

One major benefit is the improvement of decision-making that is achieved by the use of a more objective and fact-based set of information. Data-centric decision-making utilizes data in support of strategic business deci-

sions that are in alignment with the organization's goals, objectives, and initiatives.²⁷ Two outstanding success stories from the digital business world could not have been realized without a data-centric decision-making approach and the data-centric business strategies developed in conjunction with it. If they had not moved in this direction, Amazon would still be a simple online bookstore and Netflix would still be mailing out DVDs that cannot even be played in many households today.²⁸ Both of these companies, however, made key strategic business model decisions based on data and realized their organizational success by pursuing data-centric business strategies. Fast forward to today: many people cannot even picture their everyday lives without these two companies.

Strategic analytics can generally prove to be a useful tool for identifying trends that are still in their early stages or for revealing a variety of opportunities for business growth and strategic value creation. Intelligent analyses – including sentiment and network analysis of data from publicly

accessible sources such as myriads of websites, company descriptions, patent filings, news sources, clinical trial reports, geospatial, demographic, and M&A data, or academic papers, just to mention a few – can uncover patterns pointing to emerging trends, go-to-market demands, or recognition in good time of changes in customer sentiment.²⁹ What is more, these analyses can reveal new and attractive customer segments and acquisition targets, new applications for existing offerings, or even ideas for products, services, and new business models. Starbucks, for instance, makes its decisions about the acquisition of locations in consideration of analyses of demographic and traffic data as a tool for ensuring high success rates for its investments.³⁰

What is more, mathematical modeling and simulation can help to anticipate complex market dynamics by approximating real-world behavior. Agent-based modeling and

Amazon would still be a simple online bookstore and Netflix would still be mailing out DVDs.

²⁵ (Ladley & Redman, 2020) – <https://hbr.org/2020/03/use-data-to-accelerate-your-business-strategy>

²⁶ (Mulligan et al., 2021) – The strategy-analytics revolution | McKinsey

²⁷ (Tableau, 2022) – <https://www.tableau.com/learn/articles/data-driven-decision-making>

²⁸ (Miller, 2019) – Data-Driven Decision Making: A Primer for Beginners (northeastern.edu)

²⁹ (Mulligan et al., 2021) – The strategy-analytics revolution | McKinsey

³⁰ (Stobierski, 2019) – The Advantages of Data-Driven Decision-Making | HBS Online

simulations, Monte Carlo analyses, and a variety of machine learning approaches can be employed to realize differing scenarios of future situations and to assess their risks or to highlight important trade-offs and assumptions associated with a series of strategic choices. Methods of this type can be especially useful whenever managerial intuition is inadequate to account fully for the implications of the actions of multiple independent actors such as competitors or customers.³¹

Additionally, the strategic analytics approach to business strategy frees companies to become more flexible resilient, proactive, and agile. By anticipating the strategic moves of their competition and becoming a first mover thanks to the timely identification of business opportunities, companies pursuing a data-centric business strategy can lay a solid foundation for success and strengthen their competitive position on the market.

Leveraging the capabilities for utilizing data in conjunction with advanced analytics in strategy development for the radical improvement of strategic decision-making and of the overall business strategy will help any enterprise to stay ahead of the game in this data-centric age.

3.5 Recommendations for Strategy in a Data-centric Age

Becoming a data-centric company means placing data at the center of everything as they are the enablers of strategy development, the basis of decision-making and process optimization, and a part of new products and services. Indeed, the data inventory must become a strategic asset and key value driver for the organization.

We have the following recommendations for companies wanting to establish this prerequisite for long-term and sustainable market success and competitive advantage.

- (1) Explicitly include data as a value driver when deliberating possible strategies; secure a clear demonstration, externally and internally, from the top level management that the organization is on a journey to becoming a data-centric company.
- (2) Develop a convincing data strategy (taking into account the three streams “Create Business Value”, “Empower Organization”, and “Enhance Data Engineering”) that promotes the effective use of data in all business units and positions to drive company-wide value creation.

(3) Employ strategic analytics in data-centric business strategy development while leveraging data-powered decision-making to obtain more informed directional guidance, improving value creation and business agility.

(4) Keep in mind that a data strategy must define a timeline, the required resources, and the technical and support demands for its realization and successful integration into the company’s operations. The data strategies and their activities currently found in many organizations are often too low-level, too short-term, and too disconnected from the general business strategies to create a data-centric enterprise. Data can no longer be considered in isolation if optimal interplay between business and data strategies is to be achieved.

(5) Data thinking is a useful and innovative framework that can be effectively utilized during the development of a data strategy. The framework combines data science with data thinking with the consequence that the focus is not solely on data analytics technologies; instead, it extends to the design of user-centric solutions with strategic business potential.

³¹ (Mulligan et al., 2021) – The strategy-analytics revolution | McKinsey

_ 4 How Data-centricity Improves Market Position

4.1 Data-centricity: An Increasingly Critical Factor in Market Competition

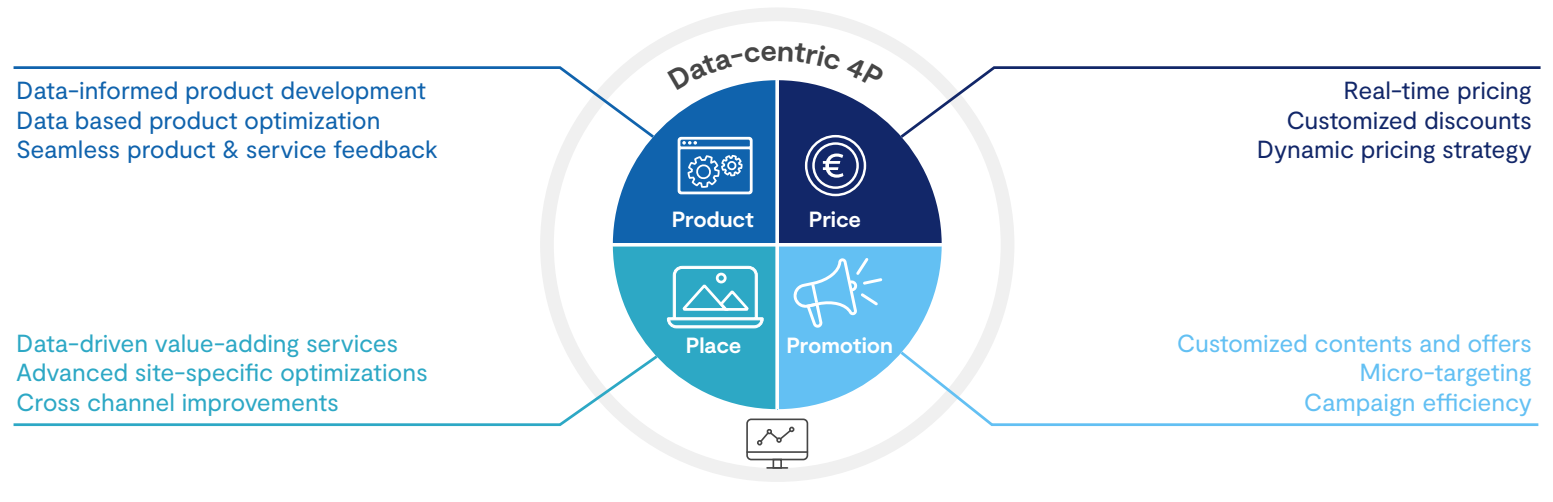
While, as briefly described above, the inclusion of data as a key starting point during strategic development is becoming increasingly inescapable, they have already achieved a position of crucial relevance and practical at the interface between companies, their customers, and their competitors in day-to-day market competition. Companies are desperately seeking to surpass their competitors by enhancing their skills in analyzing customer data and exploiting them for marketing purposes. By taking a data-centric approach, they can significantly expand the range and heighten the quality of their insights for the optimization of their own products and to expand their business model through the introduction of additional services and products complementing their current portfolio onto the market.

This chapter will explore the means by which marketing can capitalize even further on data-centricity and discuss the key factors that must be considered when moving in the direction of a data-centric business model.

4.2 A Widening Gap: How Data-centric Market Leaders Gain a Decisive Competitive Edge

In many industries, real-time analytics and optimization have long since become standard practice rather than a novelty. Data-centric optimizations are applied extensively across all aspects of marketing.

Figure 7: Data-centric Approach to 4P



€ Price

Pricing is a simple example of the use of data-based optimization that is especially prevalent in the B2C sector. Real-time dynamic pricing based on data for factors such as competition, consumption trends, and similar sources is widespread. Amazon alone makes several million price adjustments on its own platform every day. Surprising perhaps, technology companies were not the first adopters of dynamic pricing; that honor goes to the hospitality and travel industry (especially airlines as their core mission is to market each flight at the highest possible capacity while maximizing the profit margin per seat). In the meantime, virtually all large retail companies are moving to adopt dynamic pricing as a means of maintaining their competitiveness and optimizing margins. Trading and comparison platforms have already increased significantly the level of transparency on many markets. Today's customers are able to compare prices in only seconds, so suppliers are under even more pressure to set prices appropriately and adapt them to changing market conditions. Even the offline retail sector is adopting dynamic pricing optimization that has become possible with the spread of digital price tags.

Today's customers are able to compare prices **in only seconds.**

Suppliers will certainly continue to make use of dynamic pricing schemes so as to keep up with the pace of activities on their particular markets. Nevertheless, they should be aware of the challenges of dynamic pricing models that have become apparent in recent years, especially if they want to prevent significant harm to reputation and loss of trust in their brands. The ride services Uber and Lyft are warning examples; they were massively criticized publicly when, following an assassination attempt, the demand for rides in the impacted neighborhood skyrocketed and prices were dynamically raised to five times their previous level. Reaping the rewards of data-based pricing requires companies to give special consideration to the unique characteristics of the market on which they operate and to brand perception. Ultimately, they must be prepared to refine their data-based pricing policies and to be more sensitive and considerate in the future.

Promotion

Customer-related data analysis is essential to a modern company to much the same degree as competition-oriented data analysis. Companies draw on a broad range of disparate data sources so as to obtain an assessment of potential customers and their needs that is as accurate as possible and so as to tailor more specifically both their product line and their marketing approach. Personalized content is more effective at reaching customers and can significantly raise the conversion rate as well. More in-depth knowledge of customers translates into a more efficient use of marketing budgets.

An impressive example of the opportunities offered by data centricity is the Deutsche Bahn's "No Need to Fly" marketing campaign. The starting point for the campaign was the determination of the international travel destinations attracting the interest of individual Facebook users. Photos of German destinations similar to the preferred international sites were selected and the air and rail fares to both destinations from the specific user's location were calculated. The comparison of the two destinations were subsequently displayed to the user as an advertisement. This data-centric approach posted a click-through

rate that nearly quadruple that of other campaigns, and the cost per click of the campaign was reduced by 59 percent in comparison with previous campaigns. This example illustrates the potential to be found in making greater use of data-centric approaches in marketing for achieving both greater cost efficiency and heightened impact.

The relevance of customer data in marketing, already at a high level, became obvious when iPhone manufacturer Apple introduced a new opt-in model for the analysis of user data by app and website providers in spring 2021. Since the introduction of this feature known as “App Tracking Transparency,” every app must obtain the user’s consent to the analysis of his or her data for marketing purposes. Since a large proportion of users did not consent to an analysis of their personal customer data, the targeting options of many advertising campaigns were significantly restricted and the effectiveness of the campaigns declined. The result of the introduction of the new feature was an almost immediate loss of revenues totaling \$10 billion for the major platform operators Facebook, Twitter, Snapchat, and YouTube in 2021.



Place

Aside from the optimization of advertisements and “customers also bought” placements, some data-based services present clear added value for customers and opportunities for differentiation from competitors. One example of this aspect is sizing for online clothing retailers. More and more retailers have begun to customize their automated recommendations for individual customers on the basis of the measurements previously provided by the latter. These recommendation models are further refined by the inclusion of extensive datasets comprising previous orders and returns from all customers to identify more precisely what sizes will best fit specific customers. Armed with this information, customers can make more informed purchasing decisions. The greater confidence in product selection both improves the conversion rate and reduces the high volume of returns prevalent in online fashion retailing.

Differentiation through data-based services of this type is of high strategic relevance for e-commerce. The challenge for providers is found in the acquisition of the large

volumes of data required to refine the recommendation models. Large retail brands and platform operators like Zalando or Otto have the reach essential for the development of these types of services that draws on their own customer data; they are in a position to develop proprietary solutions for their own platforms. Smaller stores and fashion brands rely instead on the integration of platform-independent technology providers such as Presize.³² These platform providers give smaller retailers the chance to compete with larger platforms because they maintain powerful recommendation models trained with the data from all participating online stores.

This example from the fashion retail trade demonstrates how companies can significantly improve their customer experience by analyzing the customer data they already possess. This opportunity is not restricted to retailers with high transaction volumes. Depending on the use case, companies may find it in their interest to share data with partners and competitors for the development of new services for their own customers and the strengthening of their own data ecosystem.

³² <https://textile-network.de/de/Fashion/CAD-CAM/Mega-deal-Presize-start-up-owned-by-Meta-now>



Product

Product development can also be optimized significantly through data centricity based on collected user data. Netflix is a good example. Thanks to its enormous user base, the streaming service can analyze its users' preferences with extreme precision. This analysis of user behavior is not limited simply to the programs that are watched; it reveals what scenes were repeated or skipped or the point at which users interrupted the streaming. Netflix uses the obtained data to integrate the more

Companies should attempt to collect data along the full length of the value chain.

attractive elements into its new films and series, matching customers' interests with a high degree of probability. The evaluation of user behavior at this level of detail represents a significant competitive edge over traditional television broadcasters. This is a prime example illustrating that data use is not limited to only one link in the value chain. On the contrary, the data from one stage of the value chain can generate extensive potential for added value at other stages. The lesson is that companies should attempt to collect data along the full length of the value chain, including, if possible, the upstream and downstream links – from suppliers to customers.

The examples above are representative of the broad range of opportunities arising from data centricity at all levels of marketing, from the optimization of prices to personalized customer targeting, data-based add-on services, and product development and optimization. They encourage and remind companies to make the most of the data available to them. As of the moment, comprehensive data centricity can result in a competitive advantage; as its adaptation expands across a growing

number of markets, data optimization will become a critical element for survival in a competitive environment.

4.3 Recommendations for Marketing in the Data-centric Age

Although the applications of data-driven insights into marketing activities are impressive, this facet of data utilization will ultimately be inadequate for many traditional business models. Indeed, traditional corporations will be forced to take even more dedicated and courageous steps to data centricity going well beyond the superficial flagship initiatives still being launched by many corporations (and that have their roots in a pre-digital era) because these efforts often stall after the development of a digitalized minimal viable product (MVP). They inevitably benefit no more than a small fraction of the company's current products and value creations.

This next section proposes three major steps that will reap the full benefits from the use of data in sustaining, stabilizing, and growing business on highly competitive markets.

1. Start your transition to a data-centric business model now

The initial actions for digitalization of a business and its current portfolio are (in the best case) based on the primary business data obtained from previous operations. Data-informed activities of this nature might aid in handling the most crucial short-term challenges encountered during stabilization of the business. They are only the first steps on the journey to data centrality, however; this move to data-informed decision-making based on previous business data will not secure a front-runner position for companies still operating with traditional business models – whether for product advancement or customer knowledge or on the market. This shortcoming will become especially evident when a comparison is made with new digital competitors who are building robust platforms, forging strong partnerships, and extending their value chains.

Successfully competing with new digital players demands of the corporate leaders that they leave the organization's comfort zone, broaden their view of the supply chain in its entirety, and explore new methods for engaging with customers and gaining deeper insights from available

data. By taking these steps, companies will be able to identify the fields from which meaningful conclusions that will make a significant difference in their market placement and clientele can be drawn.

2. Stay in motion on your path to a fully data-centric business model

Merely adapting the methods used by traditional companies to create value, however, is too narrowly focused as an approach to achieve the goal of becoming a fully data-centric company on the market. **Data-driven platform** business models (Amazon, Airbnb, Alphabet on the international stage; Check24 in the GAS region) create value from the data gathered and analyzed within their platforms, but not rely solely on the data acquired during their own interaction with customers. They also integrate the data from the partners into the development of insights that enable them to tailor offerings and establish additional revenue streams. Although this oft-quoted data-centric business model is not by any means the ideal business model approach for all companies, it should be kept in mind as an image of the enormous potential of data collected by businesses and their partners along the value chain.

3. Find an approach that is yours only, based on your unique identity and data

Fortunately, the data-driven platform model is not the final word; there are a minimum of two other types of data-centric business models of proven effectiveness in safeguarding or even growing revenue streams substantially across most traditional industries.

- **Data enhanced products and services** building on existing products and services
- **Data monetization** obtaining substantial and quantifiable benefit from own and third-party data

For instance, T-Mobile US has successfully completed a large-scale project for the integration of data from its multiple local sources into one central source involving the installation of an IT architecture compatible with data integration, providing to T-Mobile the instruments for the optimization of their marketing activities addressing their selected target groups. T-Mobile has been able to generate additional revenues from tailored cross- and up-selling offerings that in turn aided in safeguarding existing revenues through the implementation of a dedi-

cated churn management system used to identify key patterns of customer dissatisfaction such as bad coverage or contract terms unfavorable to customers. Thanks to all these actions, T-Mobile has been able to cut its churn rate by more than half within the last decade.

In contrast, American Express analyzed huge amounts of its own data and applied machine learning technology so that it could predict customer churn and initiate effective countermeasures. The company did not stop with this type of data-centric value creation, however.

Exploiting its position as a service provider to millions of companies, it launched a dedicated program called Amex Advance that could provide to its clients deeper insights concerning the behavior of the latter's customers. American Express itself benefits by reducing its

Figure 8: Three Generic Types of Data-centric Business Models

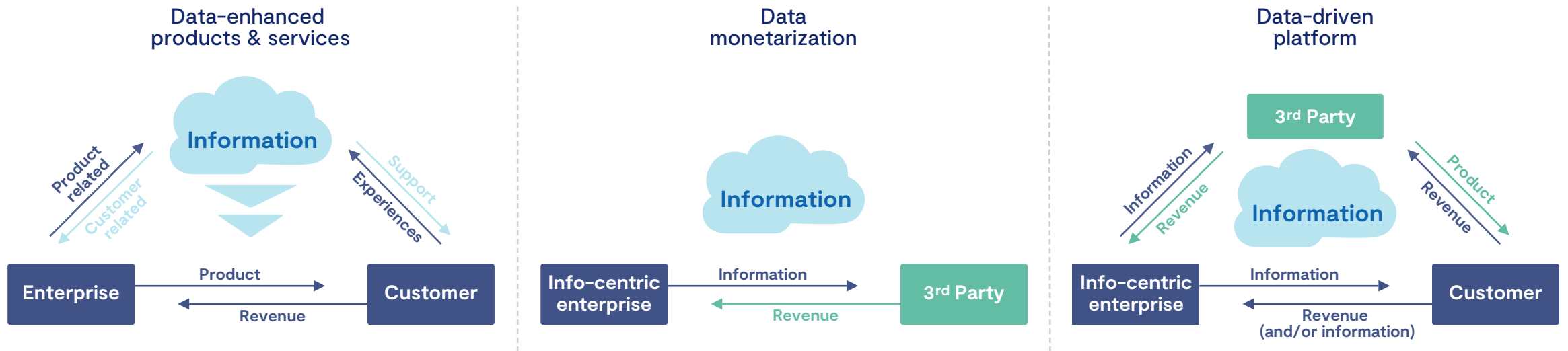
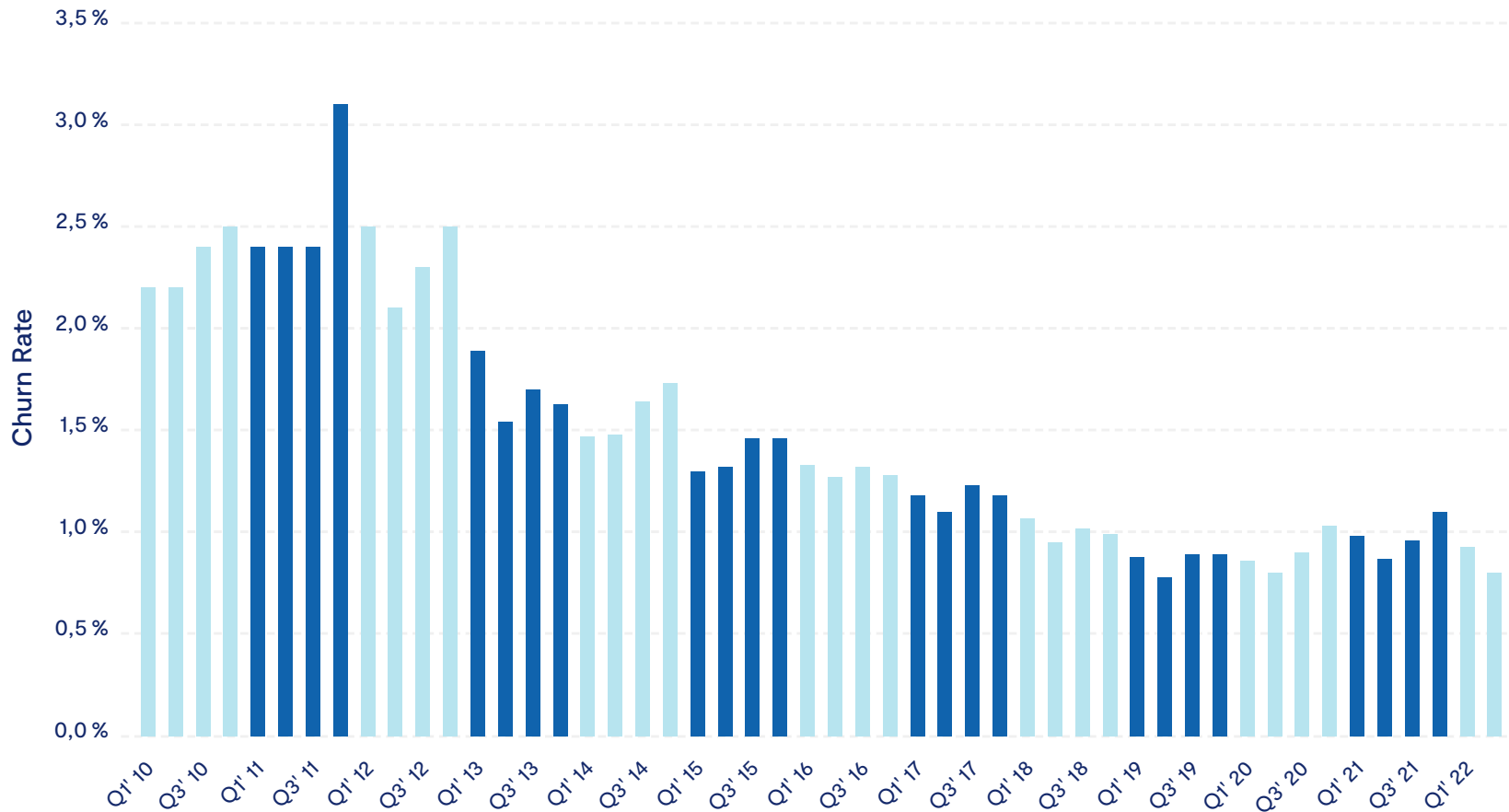


Figure 9: Churn Rate of Postpaid Contracts at T-Mobile US from 2010-2021³³



own exposure to the constantly growing market of digital payment providers, preserving its above-market margins and premium price level while simultaneously gaining new customers for the additional data-driven services.

In summation, the steps and the brief examples described above indicate that a data-driven mindset is the best basis for establishing a durable competitive advantage. While larger amounts of data lead to deeper insights, any move towards data centricity should build on the strengths of your present products and services on the market as this will enhance the chances of long-term success. In addition, the key role of the specific organizational capability to adapt and learn should never be underestimated as this is what determines the pace and the accompanying measures of the holistic transformation of organization. This will be described in greater depth in the following chapter.

³³ Statista, 2022

_ 5 Data-centric Corporate Functions

Corporate functions and R&D are fundamental to most organizations, and Detecon views data centricity as an approach that can achieve a unique placement of corporate functions and R&D, going beyond the mere compilation of data to bring together the decision-makers in the business and to provide to them the tools required to identify, optimize, and drive value.

5.1 Finance

The rising demand for data-related services concentrates above all on the supporting role of finance departments, who are facing steadily more exacting requests for data related to pricing and trade and for risk and compliance analyses.³⁴ Many finance teams, however (especially in large, sluggish conglomerates), are falling short in their attempts to satisfy these requests because of the low level of maturity of their digital analytics.³⁵

Detecon regards the following fundamental principles to be the most important pillars for companies looking to shift to a data-centric approach:

- Finance must be in the driver's seat for data value assessment.
- The use of advanced analytics in finance must be welcomed and insights that actually add value must be derived.
- The benefits of automation must be reaped and the finance department's position must refocus on value creation.

Finance executives are best qualified to answer questions about the value of data.

5.1.1 Finance in the Driver's Seat for Data Value Assessment

As data become an ever more significant factor for innovation, competition, and the improvement of profitability in all industries, they have turned into a major asset, representing a value that organizations constantly seek to maximize. Finance executives are best qualified to answer questions about the value of data.

A data-centric approach allocates this responsibility to finance, positioning it to assess data value and share data-derived insights and recommendations with senior management and other departments within the organization and to stake out its role as the custodians of enterprise data and the go-to authority for analytics.³⁶ Ultimately, the best way to garner support from senior management is to demonstrate conclusively the value of data.

³⁴ <https://www.fm-magazine.com/news/2021/jan/how-finance-can-promote-data-led-decision-making.html>

³⁵ <https://hbr.org/resources/pdfs/comm/workday/FinanceData.pdf>

³⁶ <https://hbr.org/resources/pdfs/comm/workday/FinanceData.pdf>

The results of a recent study reveal that intangible assets (R&D, reputation, data, and the like) comprise approximately 90 percent of S&P 500 market value.³⁷ This figure alone raises the question of why organizations do not give adequate weight to these assets in the valuations. Due diligence in the handling of data should help to make more informed decisions.

Obviously, the integration of due diligence when processing data in any commercial or financial due diligence procedure is sorely lacking despite the crucial part it has to play, regardless of the investment or financing project that is under discussion.

5.1.2 Embrace Advanced Analytics in Finance and Create Insights That Add Value

Staying up to speed with the current pace of technological change is a daunting task, and there is certainly a temptation to cling to familiar working methods. As the masses of available data continue to grow, however, business leaders are seeking in-depth insights that will

tie business activities to long-term value, facilitate the modeling of scenarios in real time, and indicate how resources can be allocated efficiently.

There must be a shift in focus from the provision of reports to the delivery of added-value finance services – self-service analytics, spend analytics, cash flow forecasting, and customer or product profitability projections – that can offer proactive business guidance.³⁸

One such example is the British multinational airline group easyJet, which uses big data and advanced analytics to support its finance department in steering the pricing lifecycle, which involves the management individually of 500,000 flights and their ancillary services and 30,000 daily adjustments. easyJet has developed an algorithmic ecosystem incorporating historical flight performance, live analysis that can forecast and track demand progress, and events analysis to handle the variations in demand during the year.³⁹

5.1.3 Reap the Benefits of Automation and Refocus the Finance Function on Value Creation

Advanced analytics opens the door to the benefits of automation.⁴⁰ As machine learning and AI take over more and more of their routine tasks, finance professionals gain additional time to refocus their tasks. In this sense, Detecon expects an increasing shift in the activities of future data-centric finance functions from reactive responses, transactional activities, and retroactive analysis to more sophisticated, collaborative, and forward-looking procedures driven by the desire to gain insights.

One example of this change can be seen in Xero, a provider of cloud-based finance and accounting applications for small businesses, which offers accounting software based on computer capability for the simple posting of receipts and invoices. Moreover, it has introduced a new AI-based feature for bank account reconciliation that learns from millions of historical transac-

³⁷ <https://www.visualcapitalist.com/the-soaring-value-of-intangible-assets-in-the-sp-500/>

³⁸ <https://www.aberdeen.com/cfo-essentials/the-journey-to-data-driven-finance/>

³⁹ “How easyJet used big data to achieve many millions of revenue improvements. – StrataConf” Marketing Bluematrix on June 1, 2018. <https://www.bluematrix.com/post/how-easyjet-used-big-data-to-achieve-many-millions-of-revenue-improvements-strataconf>

⁴⁰ <https://hbr.org/resources/pdfs/comm/workday/FinanceData.pdf>

tions and further reduces manual data errors for its business customers. Citi, a multinational investment bank and financial services corporation headquartered in New York City is another example; it uses AI to streamline the time-consuming, highly manual processes of reviewing huge volumes of global trade transactions while ensuring regulatory compliance.⁴¹

5.2 Human Resources

In the past, data have generally been regarded as irrelevant for HR departments. Even if this were not the case, data did not play a strategic role, being used rather to calculate KPIs measuring person-day counts or absenteeism while insights critical for the business within complex data were ignored. While many companies still use head-day counts to manage their employees' sick leave, Previa, by contrast, has introduced a graph database that delves into the figures to uncover the hidden reasons causing absenteeism. Based on the results, companies can improve working conditions and foster the well-being of their workforce.

The rise of big data and the advent of other data-related technologies have led to significant changes in the position of human resources in the last decade. Google, Microsoft, Amazon, and Facebook maintain storehouses containing a mind-boggling 1,200 petabytes of information⁴² that is exclusively related to employee activities in the organizations.

Since HR is a people-oriented corporate function, we will examine how a data-driven approach could support the core activities of the department related to people – their recruitment, management, retention, and development.

5.2.1 Recruitment: Using Data to Find the Best Applicants

Without data, HR can do no more than make assumptions about the background education, experience, or personality when assessing prospective employees, inserting an element of uncertainty into recruiting decisions.

For instance, Stafory, an award-winning start-up, has developed Robot Vera, an application that aids major companies in their recruiting of personnel. Available in various languages, it searches for suitable candidates, sends job descriptions to them, and conducts interviews, reducing overall recruitment costs for the companies by 50 percent while working around the clock.⁴³

5.2.2 Management: Using Data to Improve Organizational Efficiency and Effectiveness

Total operational costs, whether for recruitment or administrative tasks, can be substantially reduced by the automation of processes using chatbots.⁴⁴ CNBC calculates that chatbots will reduce costs by as much as \$0.70 per interaction and to cut business costs by a total of \$8 billion as of 2022.⁴⁵

Telekom Romania has launched the chatbot ANA that automates the flows of internal HR activities and provides self-service support for employees. It aids in the processing of requests for leave and medical certifi-

⁴¹ "Citi Global Trade Uses AI to Digitize Compliance in Next Generational Project". Citigroup on April 29, 2019. <https://www.citigroup.com/citi/news/2019/190429b.htm>

⁴² How Much Data Is Created Every Day? [27 Powerful Stats] | SeedScientific

⁴³ „Meet Robot Vera, the Latest AI Development to Hit Human Resources“. J.B.F. News. By Maria on May 2, 2018. <https://www.jobboardfinder.com/news/robot-vera/>

⁴⁴ How Chatbots Help HR Managers Improve Employee Experience & Reduce Costs (botcore.ai)

⁴⁵ Chatbots expected to cut business costs by \$8 billion by 2022 (cnbc.com)

cates, displays remaining days of leave, and updates employee data using optical character recognition (OCR) technology. ANA also answers general HR questions, all the while perfecting and expanding the range of its performance by learning from direct interactions with its users.⁴⁶ Yet another example is Tata Steel, a multinational steel company from India with a workforce of more than 30,000. It has developed the HR bot Cara, which answers routine questions on HR policy, saving valuable productive time and effort for the HR team.⁴⁷

5.2.3 Retention: Using Data to Raise Employee Satisfaction and to Encourage Them to Stay with the Company Longer

Measurements of factors related to working conditions such as temperature, humidity, movements, and pulse rates can be used to determine whether the working conditions for personnel are sustainable.

Seeking to guarantee the **well-being** of its workforce, North Star BlueScope Steel utilized the IBM Employee Wellness and Safety Solution for the protection of their

IBM Employee Wellness and Safety Solution for the protection of the employees.

employees in extreme working environments.⁴⁸ This application is a research project incorporating IBM Watson’s cognitive computing power and sensors for the analysis of data collected from workers’ wearables and for the provision of data to North Star management in real time whenever the technology detects potentially problematic conditions.

Google’s PiLab is a pioneer in innovative people management. It conducts practical experiments within Google to determine the most effective approaches for people management and the maintenance of a productive environment by raising **employee satisfaction** (including the type of reward that makes employees the happiest). The lab has even improved employee health by reducing the calorie intake of the personnel at eating facilities on the basis of scientific data and experiments (simply by reducing the size of the plates).⁴⁹

5.2.4 Development: Help Employees to Grow Together with the Company – Using Data to Close the Gap Between the Current Competence Base and the Future Workforce Map

HR departments carry out strategic workforce planning by aiding employees in the design of a cross-functional career path, in the discovery of opportunities for personal promotion, or in the preparation for digital work situations, shifting from an operational to a strategic role for the long-term development of a company. The appeal of strategic workforce planning lies in its promise to adopt a data-driven approach to the structuring of the workforce plan – from the Markov model to dynamic programming, from stochastic modeling to (meta)heuristics.

NASA, for example, uses a graph database to map all information about employee skills, capacities, projects, and so on for the dynamic organization of the workforce and in preparation for future missions (e.g., Mars or moon projects). In creating the body of information concerning the prerequisite skills for various occupations, the association’s team used a database called

⁴⁶ “Telekom deploys DRUID AI chatbots to automate HR processes for 4000 employees.” Druid Enterprise Chatbots on May 28, 2021. <https://www.druidai.com/chatbots-in-the-press/druid-ai-chatbot-uipath-rpa-hr-automation-telekom>

⁴⁷ “Artificial Intelligence Is for Real”. Tata Group in September 2019. <https://www.tata.com/newsroom/business/artificial-intelligence-tata-group-companies>

⁴⁸ North Star BlueScope Steel Taps IBM Watson and Wearable Devices to Monitor Activity of Workers in Extreme Environments (prnewswire.com)

⁴⁹ How Google Is Using People Analytics to Completely Reinvent HR – TLNT

O*NET from the United States Department of Labor in conjunction with ESCO, the European Skills, Competences, Qualifications, and Occupations database.⁵⁰

5.3 Research and Development

The research and development (R&D) department focuses on process innovation development (e.g., thermal management solutions in the manufacturing sector), the updating of current products, and research into and development of new products.

Growing reliance on software and the availability of simulation and automation technologies have accelerated the innovation cycle. Corporations with a well-rounded R&D strategy are rapidly scaling up innovations that often threaten to upset established business models or steer industry growth into new areas. Nonetheless, a traditional, theory-centric R&D approach adhering to the traditional research process cannot maintain pace with the rapidity of changing market requirements. Only a data-centric approach in R&D can redefine the R&D

process as a dynamic decision loop that directly analyzes market/customer data for updates of the product/service.

5.3.1 Ecosystem Rebuilding

By pursuing a data-centric approach, a company or organization can proactively create an innovation ecosystem, effectively leverage external resources globally, and readily translate research into innovative, market-ready products and services.

Any new telecommunications network should be built in fulfillment of the current technical requirements of 5G. China Mobile, for instance, which boasted close to a billion mobile subscribers in 2020, has developed an AI R&D platform that provides comprehensive, high-quality, and tagged AI training databases that can be shared throughout the entire company and that supports centralized management of enormous quantities of multidomain data. China Mobile's goal is to build an AI ecosystem that will serve the network, market, service security, management, and other areas.⁵¹

5.3.2 Enhance Agility and Speed in Product Innovation with Reduced Time-to-Market

A data-centric R&D department will actively aid in the management of a project portfolio to maximize ROI and systematically manage KPIs to steer the organization.

Enthought,⁵² a digital solution provider focusing on science industries such as material science and chemistry, life science, semiconductors, and energy, has designed a suite of data analytical tools aimed at promoting digitally powered innovation for industries operating in scientific fields. The Enthought approach merges aspects of skill development, software, technology, and digital consulting that accelerates consumer product revisions involving less trial and error through the use of machine learning.⁵³

Only a data-centric approach in R&D can **redefine the R&D process.**

⁵⁰ NASA reaches for graph DB to find people, skills for Moon and Mars missions · The Register

⁵¹ GSMA (2019). AI in Network. Use Cases in China. <https://www.gsma.com/futurenetworks/wp-content/uploads/2019/10/AI-in-Networks-Use-Case-V.03-231019-Document.pdf>

⁵² Who We Are | Enthought, Inc.

⁵³ Enthought Tool Suite – Enthought Tool Suite documentation

5.3.3 Enable the Exchange of Data and Knowledge Among Various Parties

Scientific analyses, including experimentation or simulation, are among the most important elements in R&D tasks.

Both companies and research institutes can be expected to benefit from an interoperable and transferrable exchange of data via a central collaborative platform. An industrial data space gives rise to a “network of trusted data” that creates a fully connected data exchange network among data owners, data users, brokers, and certification authorities along the full length of the industrial value chain.⁵⁴ A data platform at a central hub can facilitate the exchange of information among various departments and allow companies to leverage their assets while concentrating on strategic activities.

Merck KGaA, a German health care, life science, and performance materials giant, has developed a data integration platform with the goal of advancing cancer re-

search. The platform supports the secure, transparent sharing of data among participating researchers and institutions, enables researchers to arrive at insights more quickly, and safeguards ownership of the data.⁵⁵

5.3.4 Effectiveness and Efficiency

An interoperable data-sharing platform available in real time and with possibly lower latency will surely speed up the research process, especially public health care management.

Google’s predictive analytics in health care illustrates a use case involving the application of data-centric approaches to the support of population health management. Prior to the COVID-19 pandemic, Google developed a predictive model based on machine learning to analyze the search data people had entered about their symptoms and to predict the spread of seasonal flu; the objective at the time was to enable local medical centers to prepare the infrastructure and provide necessary services in good time. Today, Google provides a dataset of

search trends for researchers studying the link between symptom-related searches and the spread of COVID-19 so as to obtain an earlier and more accurate indication of the reemergence of the virus in various regions.⁵⁶

5.4 Recommendations

Finance:

- (1) Integrate a data due diligence procedure into every commercial or financial due diligence.
- (2) Shift your focus from the provision of reports to the delivery of added-value finance services such as self-service analytics, spend analytics, cash flow forecasting, and insights into customer product profitability that will enable proactive business guidance.

Human Resources:

- (3) Using data to improve organizational efficiency and effectiveness: Total operational costs, whether for recruitment or administrative tasks, can be substantially reduced by the use of process automation, e.g. chatbots.

⁵⁴ Jürjens, J. (2016) The Industrial Data Space: Digital Industrial Platform Across Value Chains in All Sectors Of The Economy. Web: https://ec.europa.eu/futurium/en/system/files/ged/industrial_data_space.pdf

⁵⁵ “Merck KGaA, Palantir Form Joint Venture to Develop Cancer Data Platform”. Inside Precision Medicine on November 20, 2018. <https://www.insideprecisionmedicine.com/topics/informatics-topic/bioinformatics/merck-kga-palantir-form-joint-venture-to-develop-cancer-data-platform/>

⁵⁶ “Using symptoms search trends to inform COVID-19 research” Google Health. By Evgeniy Gabrilovich on September 2, 2020. <https://blog.google/technology/health/using-symptoms-search-trends-inform-covid-19-research/>

(4) Measurements of factors related to working conditions such as temperature, humidity, movements, and pulse rates can be used to determine whether the working conditions for personnel are sustainable.

(5) Using data to close the gap between the current competence base and the future workforce map: HR departments carry out strategic workforce planning by aiding employees in the design of a cross-functional career path, in the discovery of opportunities for personal promotion, or in the preparation for digital work situations, shifting from an operational to a strategic role for the long-term development of a company.

Research and Development:

(6) Only a data-centric approach in R&D can redefine the R&D process as a dynamic decision loop that directly analyzes market/customer data for updates of the product/service.

(7) By pursuing a data-centric approach, a company or organization can proactively create an innovation ecosystem, effectively leverage external resources globally, and readily translate research into innovative, market-ready products and services.

By pursuing a data-centric approach, a company or organization can proactively create an innovation ecosystem.

_ 6 Technical prerequisites for Data-Centricity

At the enterprise level, data centrality unifies the data infrastructure and data management system and migrates or integrates data from legacy systems into new systems allowing the use of data as required. The single units within an enterprise may have previously managed their data themselves, each one using a system incompatible with all others. If the strategic use of data and the assurance of data centrality are to be achieved, data integrability, data autonomy, and data democratization are indispensable. The training of employees in the use of available technologies may well be vital in this quest.

that, when wielded by the right people, can support organizations in making better, more informed, and more data-driven decisions. The adoption of data democratization leads to universal data accessibility throughout the entire company and to the acceleration of the process of gaining insights from data (speed to insight).⁵⁷

This democratization process cannot take place unless there is a parallel shift in the technological landscape: entirely new analytics solutions that offer a lower entry

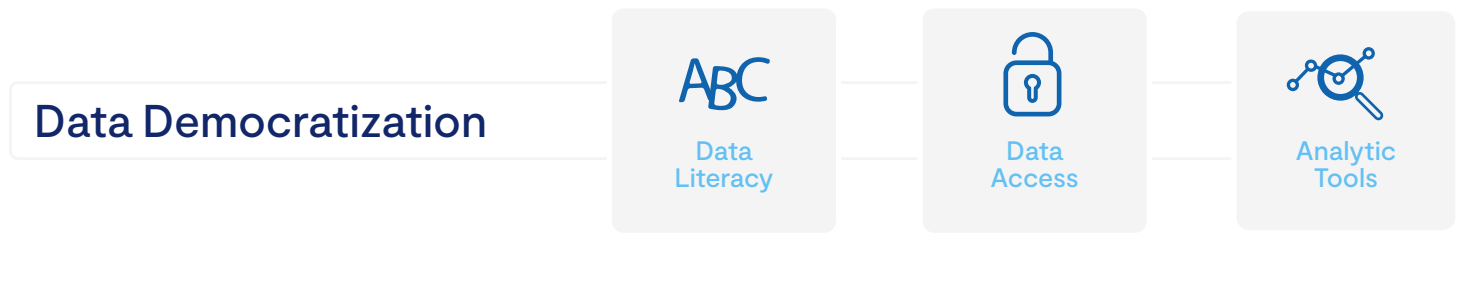
threshold than ever before must emerge. Its realization allows end users (even non-technical users and analysts) to utilize self-service analytic tools, to access historical data quickly, and to integrate, probe, and visualize them with the most suitable tools available. Just one example: graphic drag-and-drop user interfaces for cloud computing solutions have become available that permit the simple transformation of data or even the training of simple machine learning modes – without having to write a single line of code!

6.1 Autonomy and Data Democratization in a Data-centric Enterprise

6.1.1 Democratization of the Data and the Data Stack

For decades, the handling, analysis, and sharing of data in an organization was the responsibility of the IT departments and theirs alone. The insistence that all analytics pass through the bottleneck of the IT department delays data-driven business decisions, which cannot be made until the data become accessible. Data democratization is the recognition that data are the source of power and

Figure 10: Simple Visualization of the Elements of Data Democratization⁵⁸



⁵⁷ (<https://panoply.io/analytics-stack-guide/data-democratization-getting-started/>)

⁵⁸ (Reference – What Is Data Democratization? Why Is It Important for Institutions? [datateam.com.tr])

Still, if employees will be expected to access, manipulate, and interpret huge amounts of data, it is important that they understand what they are dealing with. Misinterpretation of data must be prevented. A huge inflow of data may become overwhelming and unmanageable if adequate consideration is not given to the end user's understanding of the data. Important considerations when putting data democratization into actual practice include the ways and means by which employees will access data and the quality of the insights that are being leveraged. Despite simplifications, new data users will still require a certain degree of data literacy if they are to make good use of the data. Even drag-and-drop software demands skill in its operation, and, above all, skills in the handling and interpretation of the data are vital if valid insights are to be drawn from the analysis. Data literacy helps employees to make data-driven decisions, to interact critically with data, to establish effective data governance, and to make ethical data decisions.⁵⁹

A Forrester report has found that 60 to 73 percent of all data within an enterprise remained unused for analytics as of 2019. One of the reasons stopping companies

short of becoming fully data-driven might be the lack of data skills among their employees, limiting data analytic capacity and preventing full workplace productivity. Three-fourths of the employees felt overwhelmed or out of their depth when working with data. In fact, 59 percent of employees reported a sense of burn-out when using business intelligence and data analytics tools. This level of unhappiness can be averted if employees are trained in the skills required to deal with data – i.e., data literacy.

Since data democratization will continue to be an empowerment process, the data literacy that is crucial can be achieved through training.

There are any number of excellent methods for the development of data skills among the workforce. First of all, software providers themselves have a keen interest in assuring access to low-cost (often even no-cost) learning tools and materials in a broad range of formats, whether texts, videos, or interactive, hands-on guided tasks. Training materials covering multiple levels of skill and tailored to reach a mass audience of end users are available as well.

In addition, the learning programs competing to offer affordable access to hundreds of thousands of high-quality, interactive courses in many languages are multitudes. Businesses committed to data centricity and to the significant improvement of data-driven decision-making should give thought to providing the space their employees require to educate themselves; they might, for instance purchase access to these platforms for their employees or offer time off from their work responsibilities so that they can concentrate on acquiring their new skills without distractions.

Investments in employees' data skills can also take the form of the organization of elementary and advanced courses on company premises. An internal certification course may heighten users' awareness of the dangers of misinterpretation and misuse of data. Another valuable solution might be hybrid mentoring programs during which employees progress at their own pace in online training programs while enjoying the support of mentors (colleagues who are experts in their fields) who supervise and pilot the learning process and provide assistance whenever difficulties or questions arise.

⁵⁹ (<https://www.datacamp.com/blog/why-data-literacy-is-important-for-your-team>)

Finally, the creation of communication channels or forums within the company for the users of a specific software can benefit employees in that they both provide and receive peer support whenever there are issues or concerns.⁶⁰

6.1.2 A Paradigm Shift: from ETL to ELT

The process of democratizing data access inevitably entails changes in thinking about the ways data should be handled. In the recent past, a new philosophy of data processing has been emerging, one that reverses the previous order of operations from extract-transform-load to extract-load-transform as this allows business users to utilize and integrate the data stored in the warehouse more directly than in the past. ETL stands for extract, transform, and load. Before we elaborate on this concept, let us first explain the idea of data pipelines.

A data pipeline is a series of data processing steps. The pipeline begins with the collection of the data from their source. This is followed by a stage of data processing involving a quantity of data received as input, and the result of each stage is output that becomes the input for the next stage. This procedure continues until the

processed data reach their final form and destination. Data pipelines allow data to flow from their source (e.g., a register of incoming calls or a smart device) to a data warehouse or lake and from a data lake to an analytics platform, whether they are transformed (filtered, enriched, aggregated, merged with other datasets) or where machine learning algorithms are running, and from there to a dashboard application or back to a smart device for the improvement of its future performance. Data originating from one source can feed multiple data pipelines.

For instance, a product review could generate data to feed a real-time report counting product mentions or a sentiment analysis application. Each of these applications relies on unique data pipelines leading to the desired result.

As numerous concerns about ETL pipelines have been raised recently, a new paradigm as well as new technological developments and possibilities have emerged. ETL pipelines have been criticized for several reasons. It is impossible to determine upfront what uses for the analyses, transformations, and raw data in the pipelines might appear in the future. The original idea behind the

Figure 11: Business Scenarios for ETL and ELT⁶¹

ETL	ELT
<p>✔ Source and target databases are different (e.g., Oracle source and SAP target databases)</p>	<p>✔ Source and target databases are same (e.g., Oracle source and target databases)</p>
<p>✔ Data volume is small or moderate</p>	<p>✔ Data volume is large</p>
<p>✔ Data transformations are compute-intensive</p>	<p>✔ Data transformations are less complex</p>
<p>✔ Data is structured</p>	<p>✔ Data is unstructured</p>

⁶⁰ (<https://www.astera.com/type/blog/5-best-practices-to-achieve-data-democratization/>)

⁶¹ (Reference - ETL vs ELT for Data Warehouse: What's The Best Approach? | Software Advice)

use of the data may evolve, and it may become necessary to restructure the data in deviation from the original concepts. Furthermore, the choice of transformation method for the source data in the ETL process is one of specific decision-makers and may well be final and irreversible. Perhaps only aggregated data in a daily, weekly, or monthly resolution are available while the analysis requires unstructured or fine-grained data in a second-by-second resolutions. There may also be a lack of clarity to later analysts or data consumers as to who transformed the raw data and why they were transformed and how the currently available data differ from the original data.

Advancements in the field of cloud-based computation and storage have made a new paradigm possible: the EL(T) approach. EL(T) means that the data are extracted (E) from their source in a raw format and stored (L – loaded) as-is so that they are available for any type of analyses in the future that might be in the form of a so-called reverse ETL process. Reverse ETL means that the data stored in the original format can be repeatedly re-extracted and transformed in a different way for every extraction, depending on what is required at the specific time. Since different data users may have differ-

ent requirements, ELT supports agile decision-making for analysts and promotes data literacy throughout the entire company.

In contrast to ELT, EL(T) decouples the extract-load steps from any optional transformation that may take place. The operational use cases, being unique, might or might not require a transformation process. The analytical use cases might require the transformation of the data at some point. The separation of EL from T allows analysts to choose whatever type of transformation they wish.

6.1.3 The Concept of Data Mesh

In 2020, Zhamak Dehghani proposed the concept of data mesh, a new enterprise data architecture framework that is expected to bring about revolutionary changes in data-driven decision-making.⁶²

The digital transformation of recent decades has led to an explosion in the quantity of generated data. The phenomenon has forced the development of increasingly powerful and scalable platforms for the storage and management of data and for their rapid extraction and processing, including real-time analytics. Parallel to this development, advanced analytical tools capable of

handling huge volumes of data have emerged. While it is noteworthy that providers have been creating advanced technologies to accommodate and process data, it has nevertheless been determined that even investments in state-of-the-art technologies do not automatically translate into faster data-driven decisions. In the author's opinion, this is a consequence of a maladaptive approach to data management and data architecture that assumes the objective of a centralized, monolithic, domain-agnostic data storage and management framework administrated by an elite group of highly specialized experts (whose accessibility may be highly limited) with no knowledge of the specific domains where the data originate.

The author stresses that the management and processing of data by teams isolated from the people generating or consuming the data cannot produce optimal results.

The difficulties begin with the generation of source data by teams with advanced knowledge of specific data, but who are frequently not data experts. Whenever data are collected for operational purposes, teams do not necessarily feel motivated to care about the data's usefulness to downstream consumers as long as the envisioned purposed has been achieved. The resulting datasets may

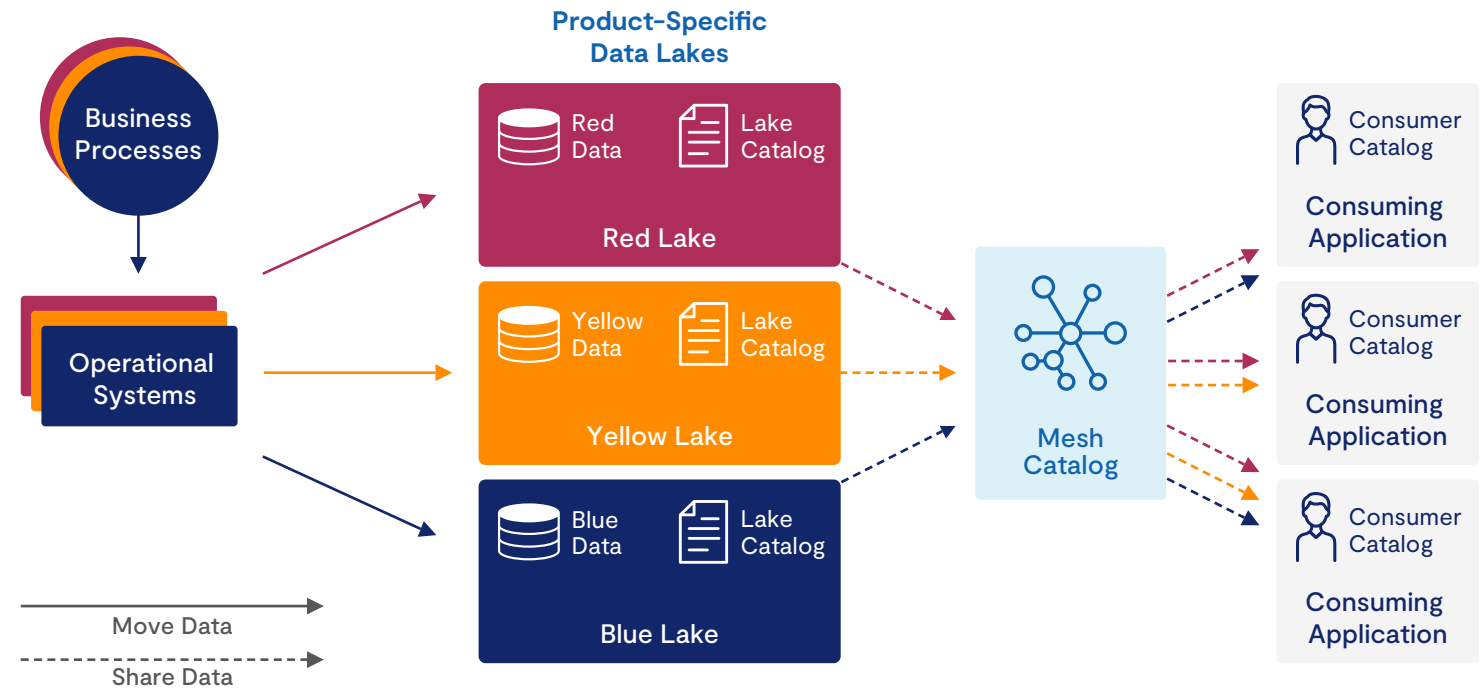
⁶² (<https://martinfowler.com/articles/data-monolith-to-mesh.html>)

contain errors, lack transparency, and be described so inaccurately that downstream users are unable to understand them; moreover, the documentation data engineers and analysts require to comprehend the material when they begin working with it may be missing. When end-user teams need reports on which they can base their data-driven decisions, they lack direct access to the data and the requisite data skills. They are forced to line up a team of data experts and assign to them the task of mining, processing, and adapting the data as appropriate to satisfy the consumers' needs. Yet data experts do not have the knowledge in either the domain of the team that generated the source data or of the target user team. Before they can even begin their actual tasks, they must become familiar with the subject matter and communicate extensively with the data creators and consumers before they can understand what types of reports or products they have been asked to create, and inquiries about the details of the datasets that only the latter's creators can answer are inevitable. The consequence: as ideas for new use cases proliferate and the backlog of requests to the data specialist continues to grow, processes are delayed and fast data-driven responses become impossible.

These delays can be avoided by moving away from the centralized data management framework commonly in use today, namely, the collection of all of the company's

data by a powerful data platform and managed by a small group of highly specialized data experts.

Figure 12: Structure of the Data Mesh Architecture at JP Morgan⁶³



⁶³ (Reference – How JPMorgan Chase built a data mesh architecture to drive significant value to enhance their enterprise data platform - Dustin Ward)

The author lists the principles that should be followed to make the new framework more data-centric:

1. Domain-oriented, decentralized data ownership and architecture (data are locally owned by the team responsible for their collection and/or consumption)
2. Data as a product
3. Self-service data infrastructure as a platform
4. Federated management of computing resources

The type of storage for both raw data and aggregated data products must be tailored to the needs of the specific consumers. Any data made available to consumers should be registered according to a standard scheme so that anyone requiring the data can also find them (self-service data infrastructure as a platform). The experts on the data-consuming teams then have what they need to perform the required analyses and create the requested data products such as reports, presentations, and dashboards.

The advantage of this proposal is that data experts have direct access to the team members with domain expertise and vice-versa, a situation in which mutual learning takes place. On the one hand, the data experts acquire domain knowledge and a better understanding of the data they are processing, making it possible for them to be utilized more meaningfully. On the other hand, the non-technical domain experts become familiar with the data and gain valuable skills by working daily with the data engineers.

Decentralization of this type significantly shortens the process from data collection to data-driven decision-making because teams needing data from other teams for their own purposes and the teams with the data communicate directly with one another; as each team has data experts with knowledge of the domain, mutual communication is further simplified.

This must not be understood to mean that decentralization should involve individual domains using incompatible data storage platforms. Such a situation would lead to the creation or persistence of data silos, and the efficient flow of essential data to their intended users would be difficult, if not impossible.

Modern and scalable platforms for the storage and processing of enormous quantities of data (cloud solutions that enable the creation of data warehouses and data lakes, e.g., data lakehouses) should be used, albeit it is necessary to rethink their architecture so that each team has full freedom to manage its own data while simultaneously assuring ease of accessibility to anyone else who may require the information. The governance of the architecture should be regulated by role-based access control. Organizations can assign different administrators to different sections of the catalog to decentralize control and management of data assets (federated management of computing resources). This hybrid model of a centralized catalog under federated control preserves the independence and agility of the local teams of specific domains while ensuring the reuse of the data assets among all teams and enforcing a common security and governance model globally.⁶⁴

The author writes that the data products created by each domain team should be discoverable, addressable, trustworthy, possess self-describing semantics and syntax, be interoperable, secure, and governed by global standards and access controls. In other words, the data should be treated as a product that is ready to use and reliable.⁶⁵

⁶⁴ (<https://databricks.com/de/blog/2021/08/30/frequently-asked-questions-about-the-data-lakehouse.html#ten>)

⁶⁵ (https://en.wikipedia.org/wiki/Data_mesh)

The author notes in his summary that the data mesh platform is an intentionally designed distributed data architecture, under centralized governance and standardization for interoperability, enabled by a shared and harmonized self-serve data infrastructure.⁶⁶

6.2 The Infrastructure Prerequisites for a Data-centric Company

Data infrastructure serves two purposes at a high level: the storage of the operational data (i.e., the data that are created and collected in the course of ongoing business such as records of financial transactions, sales operations, customer data, conducted projects) and the utilization, transformation, and analysis of these and other collected data in support of business leaders to secure better, data-driven decisions.

We have already mentioned that the trend to democratization of data access has prompted a shift in the approach to data pipelines away from extract-transform-load to extract-load(-transform). In place of top-down decisions mandating the format in which data are stored for future users, data are stored in their original format so that future users are completely free in their choice of the transformations that best suit their needs.

While the data warehouses commonly used permit the storage of structured data (usually in tabular form), they have serious limitations and are not suitable for data pipelines based on the EL(T) model. When data warehouses are in use, some of the raw data may be irretrievably lost. This happens after the extraction (E) – at the stage when the decision about the method for transformation of the raw data is made (e.g., their aggregation at specific time intervals (T)) and their subsequent loading into the target database (L).

Current technology trends – both the paradigm shift from ETL to EL(T) and the decentralization of data ownership in line with the data mesh philosophy – are better served by so-called data lakehouses, a combination of data warehouses and data lakes. Data lakes are platforms for the storage and subsequent extraction of data in their raw form while data warehouses store structured data. Data lakehouses combine the functions of the two platforms. They make possible, on the one hand, the saving, storage, and extraction of data in their original form and, on the other hand, the creation of data warehouses (structured, tabular databases) that can result (for example) from the processing of raw data stored in the data lakehouse.⁶⁷

In summary, an ideal data infrastructure for modern data-centric solutions should possess the following characteristics:

The tangible components of a data-centric organization embracing data democratization are data virtualization software, data federation software, cloud storage, and self-service applications for non-technical users. This type of infrastructure – when combined with a paradigm shift toward democratization of access to data and investments in the development of data skills among employees – will not only ensure the transition from the monolithic, centralized data platform paradigm to the decentralized data mesh framework, but will also create optimal conditions for significant improvement and acceleration of data-driven decision-making in all areas and at all levels that may be necessary, thereby satisfying the prerequisites for becoming a fully data-centric enterprise. Data mesh connects the various data silos, a feature that aids organizations in the integration of scalable and automated data analytics.

The shift from a data-driven organization to a data-centric one positions data as a key element of the organization's processes, and companies making this transition

⁶⁶ (<https://martinfowler.com/articles/data-monolith-to-mesh.html>)

⁶⁷ (<https://databricks.com/de/blog/2021/08/30/frequently-asked-questions-about-the-data-lakehouse.html#ten>)

Figure 13: Characteristics of a Modern Data-centric Solution

- 1** Be highly efficient and flexible
- 2** Be scalable i.e., capable of storing and matching storage space to the massive amounts of data generated; cloud management platforms are well suited for this purpose
- 3** Be capable of storing raw data in its original formats
- 4** Allow data to be stored in a structured and relational form
- 5** Be capable of integrating and transforming data from numerous and very diverse sources
- 6** Provide the ability to quickly extract huge volumes of data for analytical purposes and, to perform real-time data analytics
- 7** Easily integrable with external systems
- 8** Be equipped with or capable of being integrated with intuitive tools for data processing and analysis that are as accessible as possible for non-technical users
- 9** Enable decentralization of data ownership across the enterprise in accordance with the data mesh concept
- 10** Be reliable

are even now reaping the fruits of their labors. There are numerous benefits of data centricity for businesses: management and assessment of risks, improved accuracy of internal and external reporting, and management of costs, to name just a few. When it comes to choosing the optimal solution, however, we must recognize that there is no simple answer along the lines of “one size fits all.” Most organizations wish to have a solution providing a one-stop data platform that can be used to perform analytics and obtain meaningful insights without requiring the assistance of a central IT team. Although the companies will continue to maintain a central IT team, it does not own the data.

6.3 Recommendations

- Adopt democratization: data accessibility becomes universal throughout the company, accelerating the process of gaining insights from data (speed to insight). End users can utilize self-service analytic tools, access historical data quickly, and integrate, probe, and visualize them. For instance, graphic drag-and-drop user interfaces of cloud-computing solutions enable sim-

ple data transformations or even the training of simple machine learning models without requiring users to write a single line of code!

- Develop the data skills of employees: begin by facilitating access to low-cost (or even no-cost) learning tools and materials in a variety of formats. Moreover, various learning platforms compete in offering affordable access to hundreds of thousands of high-quality, interactive courses in many different languages. Invest in employees’ data skills by organizing elementary and advanced courses (such as internal certification courses) on the company premises. Finally, the creation of communication channels or forums within the company designed for users of a specific software program can support employees.
- Move away from the centralized data management framework now commonly in place based on the collection of all the company’s data by a powerful data platform and their management by a small group of highly specialized data experts. Install instead a new,

more data-centric framework featuring domain-oriented, decentralized data ownership and architecture (data are locally owned by the team in charge of collecting and/or consuming the data) that includes a self-service data infrastructure as a platform.

- Adapt to the demands of current technology trends: both the paradigm shift from ETL to EL(T) and the decentralization of data ownership in line with the data mesh philosophy benefit from so-called data lakehouses, a combination of data warehouses and data lakes. Data lakes are platforms for the storage and subsequent extraction of data in their raw form while data warehouses store structured data. Data lakehouses combine the functions of these two platforms. They make possible, on the one hand, the saving, storage, and extraction of data in their original form and, on the other hand, the creation of data warehouses (structured, tabular databases) that can result (for example) from the processing of raw data stored in the data lakehouse.

⁶⁴ (<https://databricks.com/de/blog/2021/08/30/frequently-asked-questions-about-the-data-lakehouse.html#ten>)

_ 7 Compliance

7.1 Introduction

Technology-savvy companies in Europe are often of the opinion that the laws, standards, and rules safeguarding data protection and security prevent innovations on a major scale. The general belief is that strict laws would cause Europe to fall far behind China and America. The data that are, or could be, obtained cannot be used for data-driven innovations, supposedly with the consequence that the number of data-driven innovations coming from Europe will be severely limited.

Other technological developments and innovations are now regulated by legal statutes as well. A bill proposing regulations that would ensure trustworthy artificial intelligence has been open for discussion since 2021; its provisions would impose sanctions for infringements that could amount to as much as 6 percent of a company's worldwide annual revenues ([Figure 14](#)).

7.2 Application of Regulatory Statutes Poses Major Challenges

A Detecon analysis has identified three problem areas for companies that have been forced to abandon data-driven projects involving new technologies:

1. The impact and requirements of the pertinent data protection regulations were not given consideration in the development from the outset. If this type of analysis of possible consequences is postponed, complex iteration loops usually result.
2. The organization's own compliance processes are regarded as static, complicated, and time-consuming, consuming major resources (sometimes from external sources) for audits and consulting activities. Moreover, many processes have not been adapted to include new methods such as machine learning or data self-services (employees' independent access to data lakes for the simulation of processes). If these compliance measures are not taken into account until the

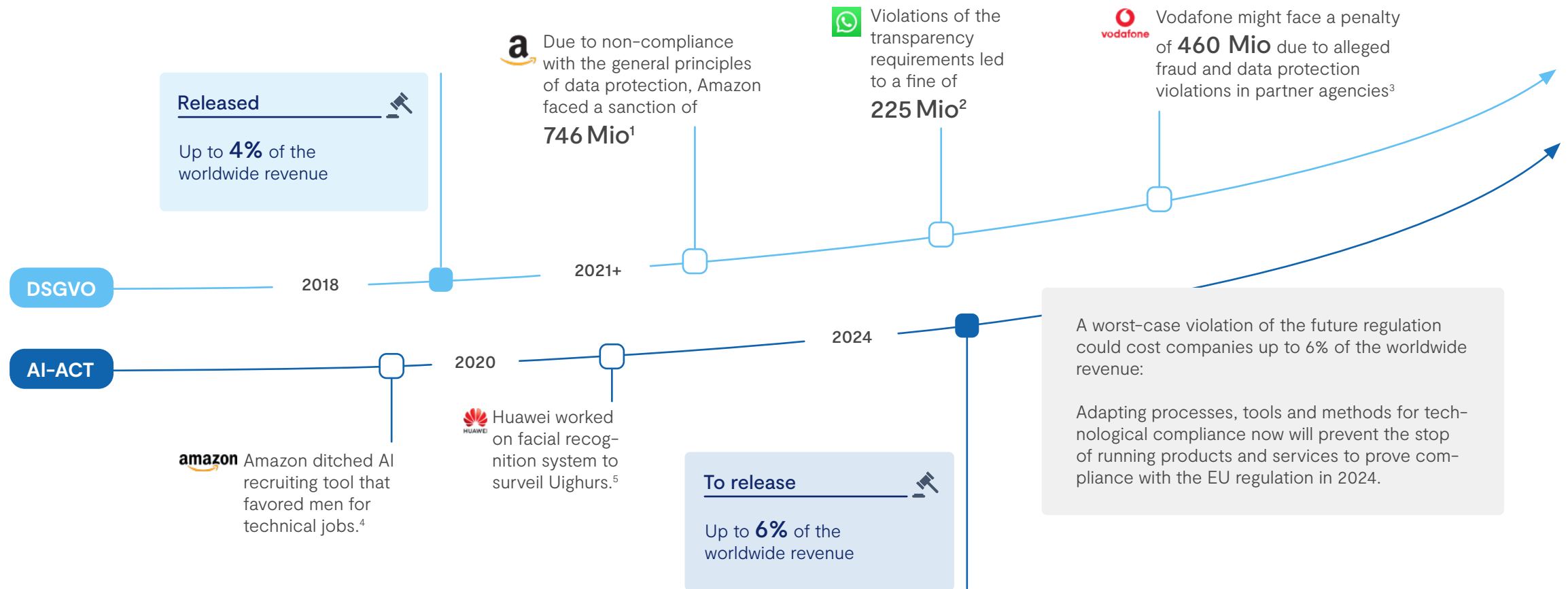
development process is nearing its end, the project will become substantially more difficult; in the worst case, it may even fail.

3. Some companies prohibit the use of certain technologies (machine learning or the utilization of their own data) as a general principle out of concerns regarding the complexity and ambiguity of interpretation. This inevitably reduces process innovation within the company.

Consider privacy policies during development.

Figure 14: Sanctions Imposed by the GDPR

Example: Penalties at DSGVO



¹ FAZ, Amazon soll 746 Millionen Euro als Datenschutz-Strafe zahlen, 2021

² Tagesschau, Rekordstrafe gegen WhatsApp, 2021

³ NTV, Vodafone droht nach Betrug hohes Bußgeld, 2021; Spiegel

⁴ Reuters, Amazon scraps secret AI recruiting tool that showed bias against women, 2018

⁵ Tagesschau, Kamera-Software soll Uiguren erkennen, 2020

7.3 Development of a Compliance Framework

Hypothesis:

The compliance department plays a key role in the process of data-centric innovation. It serves as a bridge and translator between companies and lawmakers. When intelligently anchored in the organization, compliance offers support, translation, and assistance for the development of data-driven innovations.

Providing departments with a secure framework for the practical application of laws and guidelines when working with innovative technologies requires the establishment of a holistic compliance framework from the very beginning.

In the compliance framework of [Figure 15](#), laws and standards serve as the foundation for the levels of “Technology” and “Mindset & Culture” on which the separate initiatives are built.

The **“Mindset & Culture” level** focuses on imparting and deepening the required knowledge of pertinent

laws and values. Principles such as human centricity, transparency, and a sense of accountability are essential. There are several fields of action at this level.

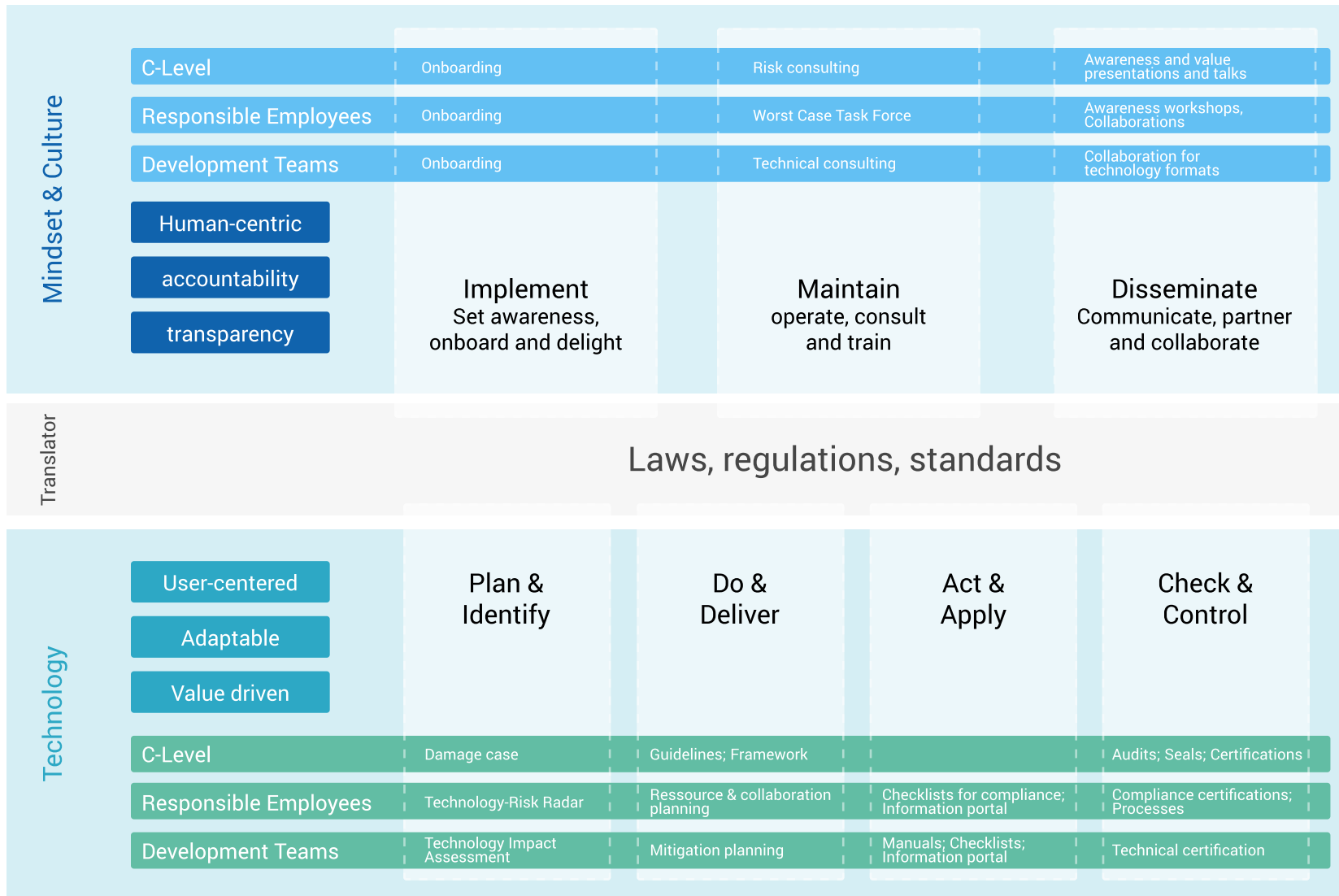
- *Creating (responsibility) awareness and building knowledge:*
Developments must remain within the boundaries of the legal framework. This message must be clear at every level of the companies, and employees should be aware of their individual responsibility. This necessity is the reason behind the concrete offers of onboard training at many companies at which time new employees are taught practical basic knowledge of the laws right from the start of their work. Their internalization of values and regulations also creates a clear profile to the outside world.
- *Initiating measures and supporting consulting services:*
Whenever new technologies are used in companies, questions that can rarely be resolved by the developers alone often arise, which is why it is important to provide unbureaucratic consulting services with compliance experts so that obstacles can be quickly eliminated.

- *Communication and partnerships:*
Transparency in the presentation of values and compliance is an important buying factor for customers. Communication campaigns based on this awareness can strengthen values and create competitive advantages. Partnerships aid in maintaining best practices or securing a common strategic position. When legislative actions such as the AI Act are pending, the exchange of ideas at an early stage of realization is highly valuable. The results of these cooperative activities can even have a significant impact on the final form of the legislation.

The **Technology level** should orchestrate technological adoption in a user-oriented, adaptable, and value-oriented manner. It is characterized by cooperation with the respective specialist departments that supervise or use a specific technology.

- *Identification and analysis of potential risks of new technologies and procedures to liability or reputation:*
Technologies that will play an important strategic role for the company in the future, but that are subject to

Figure 15: Example of a Compliance Framework



significant regulation can be identified by radar tools or impact maps. For instance, the interplay of social and political factors influencing a technology can lead to the application of certain legal requirements. The use of artificial intelligence in HR is one example. The legislative bill of the AI Act stipulates that its use in this department is subject to very narrow and strict conditions.

- *Definition of guidelines and principles:* The technologies or practices identified in the first step are now analyzed and a target vision for dealing with them is defined. This target vision outlines principles derived from the legal provisions, but incorporates as well the company's own values. For example, many companies set for themselves binding guidelines for dealing with artificial intelligence that often go beyond the scope of the (future) regulations.

· *Design of user-friendly and practical measures, tools, and methods:*

In the next step, the guidelines and principles are realized in practical actions clearly reflecting their purpose for each level of responsibility and in the context of any given job. The selection of topics (such as the handling of data relating to artificial intelligence) as well as the method should be developed in collaboration with users with the aim of maximizing acceptance. Creative methods such as a chatbot that can be consulted rather than a manual to find quickly answers to various questions are also conceivable.

· *Development of review measures:*

Finally, tools, measures, processes, and methods to verify that the previously defined guidelines have realized in the related technology are developed. An example here is the use of an ethics seal certifying adherence with the previously established guidelines on working with artificial intelligence.

Companies need a change in their way of thinking.

7.4 Challenges

Despite the numerous advantages achieved from the implementation of these developed tools and methods, it is extraordinarily important not to stop with the intensification of the culture and mindset related to compliance with the company; the organization must itself become more compliance friendly as the cooperation between the business departments and the compliance department heightens the user friendliness of the processes and the general attitude toward compliance measures, significantly impacting the success and acceptance of the products.

7.5 Result

Authorities have only recently begun to enforce fully the penalties established in the GDPR, actions that can have far-reaching consequences for both companies and individuals. If they are to secure their competitiveness, companies need a change in their way of thinking, innovative processes, and methods that will integrate the laws into their operations; most companies still use complicated static processes to ensure the legal compliance of their own products or services or leave the review to external consultants or even their own devel-

opment teams. Regardless of the approach chosen to take the GDPR into account, we believe that any support provided for the implementation and translation of laws and regulations will enable the pertinent users to develop legally compliant, data-driven innovations, to protect the company from liability claims and damage to its reputation, and in the long run conserve resources by means of standardization within the company.

This approach must be firmly established at an early stage to assure compliance with future regulations as well; the next proposed legislation for the territory of the EU, which will have a massive impact on data-centric companies and data-driven business models, is already under discussion: the bill governing trustworthy artificial intelligence.

7.6 Recommendations – User-centric Approach to the Legally Compliant Use of Technologies

In the best case, a win-win situation that both guarantees compliance with the law and creates added value for users (e.g., with certificates for online use) can be achieved.

The following steps are essential for a user-centric approach.

1. Define the goal and desired impact of the initiatives:

Define the long- and short-term impact you want to achieve with the activities. Work on an equal footing with the users and develop user-friendly tools, methods, and processes that enrich the user journey.

2. Determine the technology focus (Plan & Identify):

The starting point is the identification of technology and procedural errors that could trigger liability or damage to reputation. A structured analysis from the perspective of the company's own strategy, business models, and use cases determines what possible threats are present and require a controlled approach.

3. Determine the scope of action (Do & Deliver):

The next step displays the actions regarding the different systems or technologies and data categories. For instance, the principles of processing within the sense of Art. 5 GDPR must be observed during the processing of personal data. Before beginning to use artificial intelligence with an ethical orientation, an individual value catalog of a company can map the framework for action.

4. Define the users:

This is followed by the identification of all stakeholders and users. User interviews or surveys can later help to decide or to prioritize specific initiatives before classifying the user groups. Different classifications can point to the right tools for a group. Examples of gradations of user groups could be various levels of responsibility or the involvement of groups in differing product life cycles of the technology.

5. Understand the needs for each user group:

Now is the time to define the behavior per user group and to examine all aspects that could support or hinder users' behavior (processes, product life cycles, etc.). Owing to the varying role models, conscious perception of an individual sense of responsibility for handling data is not possible.

Further problem areas are also evident:

1. The lack of easy-to-understand and practical information.
2. Understanding whether and to what extent the GDPR or other regulations and their specific requirements must be applied to specific systems.
3. Uncertainty about the implementation of the laws during the relevant phase of the development life cycle.
4. Elaborate and complicated audits and verification systems shortly before the system is launched.

6. Challenge for the needs of a user group:

A tried and tested instrument to clarify the needs of a user group are so-called challenge workshops. A challenge is designed to define the goal, the area, and the scope of measures. Define the challenge as precisely as necessary, but as openly as possible to give scope for the subsequent step of finding a solution.

A challenge is designed to define the goal.

7. Finding solutions: –
Five steps to mastering the challenges:



Select the user group for which you want to solve a problem in the subsequent process, present their needs and context, their everyday work, and processes.



Analyze the regulation and determine what it means in terms of the users and the technology.



Generate ideas about the type of translation help, tool, or method that could aid the selected user group in their development and operation of the products and services in compliance with the law.



Design prototypes at the earliest possible stage and test them with the user group. Check to see whether their needs have been met and whether the users can carry on with the initiative independently.



Prioritize rapid development and launch of the initiative, seeking initially to satisfy primarily the basic needs. Always seek feedback proactively and continuously improve processes and content.

_ 8 The Authors



Steffen Kuhn

Steffen Kuhn is the founder of the Digital Engineering Center at Detecon. Having had more than 20 years of professional experience, Steffen is a highly competent senior advisor to CXO and a recognized expert in disruptive technologies, innovative strategies, and future business models.

As he is a pioneer exploring the future of the consulting profession, Steffen's mission is to combine consulting and innovation with high-tech and implementation expertise with the objective of ensuring the transfer from the innovation ecosystem to practical application. His firm conviction is that all organizations, their customers, and their employees will benefit from digital transformation.



Dr. Volker Rieger

Dr. Volker Rieger is in charge of the corporate and digital strategy consulting activities. He has been dealing with the changes in business models resulting from digitalization for more than 20 years. He is proud to have been one of the first to coin the term data-centric business models more than ten years ago.

Volker studied physics in Germany, the USA, and Japan and earned his doctorate in quantum optics.

Drawing on this experience, he served as a consultant supporting the transformation of the telecommunications industry for about ten years before turning to the same challenge in the energy sector.



Marcus Berlin

Marcus Berlin is Senior Manager at Detecon's Digital Engineering Center and oversees the work in data analysis and artificial intelligence. During his projects, he supports large companies in the fields of data strategy and governance and in the design of data-driven new business. This latter facet is also based on his extensive experience in the application of data-centric design thinking methods (data thinking).



Matthias Fischer

Matthias Fischer is Senior Consultant in the field of digital strategy and innovation, where he focuses on business model analysis and development, enabling him to support companies as they strategically realign their own business models to meet the challenges of digitalization. In his previous projects, he advised infrastructure providers in the energy, telecommunications, and mobility sectors. Matthias studied industrial engineering in Germany and New Zealand.

development of innovative products. In her project work, she provides guidelines for successful data-driven product development for clients from various industries ranging from manufacturing, telecommunications, automotive to public transportation and beyond.



Isabell Neubert

Isabell Neubert is Consultant in the Co-Innovation department of the Digital Engineering Center. Her consulting activities focus on digital ethics, trustworthy artificial intelligence, value-based innovation management, and crossover topics such as co-innovation methods and digital service design. She holds a bachelor's degree in industrial design and a master's degree in Strategic Innovation in products and services. After earning these degrees, she completed a program of graduate study at Deutsche Telekom.



Elena Fomenko

Elena is Manager at Detecon's Digital Engineering Center and is in charge of AI and data-driven product development. She brings to her work a unique combination of technical skills in data sciences and vast experience in the conceptual



Philipp Maier

Philipp is Consultant in the field of analytical intelligence at the Digital Engineering Center and a member of our Munich office. He focuses on the topics of data strategy, data management, and data analytics. Before coming to Detecon, he gained expertise while working at various consulting firms, where he focused on technology and corporate venture capital. He earned his master's degree in computer science and entrepreneurial finance at the Technical University of Munich; his studies included time in Cambridge (USA) and Grenoble.



Johannes David Pfeffer

Johannes Pfeffer is Detecon Consultant in the field of digital transformation and specializing in value creation and strategy. In his project work, he steers his clients as they chart their course for the creation of

value on volatile digital markets by developing strategies that work end-to-end and by designing digital business models and business ecosystems.



Malte Schröder

Malte Schröder is Manager in the Global Chapter Services & Industries at Detecon Consulting. He has been working as a (multi-)project manager in the energy, transportation and logistics, and public and health care sectors since 2016 and has managed major projects for clients in the GAS region and all across the EU.

In addition to his technical specialization, Malte designs customer-centric, digital, and data-driven products, services, and business models on a functional level with his clients, orchestrates complex rollout projects, and support clients as they prepare to tackle the growing challenges related to IT security and data governance.

He is also passionate about advancing the skills of organization so that they are better equipped to meet the demands of the digital age.

Furthermore, we would like to express our sincere thanks for their cooperation in this publication to

Sai Keerthana Nallapula Ananthraj, Katarzyna Bojarska, Britta Cornelius, Akshay Narendra Shewatkar and Xin Yu.

_ 9 The Company

Management consulting with pronounced technology expertise

Detecon is the leading, globally operating technology management consulting company with headquarters in Germany and has been combining classic management consulting with high technological competence for over 40 years. The focus of its activities is on digital transformation: Detecon supports companies from all areas of business during the adaptation their business models and operational processes in response to the competitive conditions and customer requirements of the digitalized, globalized economy with state-of-the-art communication and information technology. Detecon's expertise bundles the knowledge from the successful conclusion of management and ICT consulting projects in more than 160 countries.

Detecon is a subsidiary of T-Systems International, one of the world's leading vendor-independent providers of digital services and subsidiary of Deutsche Telekom. Detecon represents the consulting pillar in the business strategy of T-Systems.

From concept to implementation

Detecon is driving forward its consulting approach Beyond Consulting, a significant evolutionary step forward in traditional consulting methods adapted to meet the demands of digitalization today and in the future. The concept features top consulting that covers the entire spectrum from innovation to implementation. Ground-breaking digital consulting demands ever greater technology expertise and a high degree of agility that incorporates flexible, but precisely tailored networks of experts

for complex, digital ecosystems in particular. At the same time, it is more and more important in digital consulting to accompany clients from innovation to prototyping to implementation.

This factor prompted Detecon to found the Digital Engineering Centers for Cyber Security, Analytical Intelligence, Co-Innovation, and Industrial IoT in Berlin in 2017 as facilities that extend the added-value chain of consulting and accelerate the realization of digital strategies and solutions by means of prototypes and proofs of concept.

DETECON

CONSULTING

Detecon International GmbH

Bayenwerft 12–14

50678 Köln

Phone: +49 221 9161 0

Detecon Digital Innovation

Winterfeldtstraße 21

10781 Berlin

E-Mail: info@detecon.com

Internet: www.detecon.com

