MIS Quarterly Executive

Volume 23 | Issue 1 Article 3

March 2024

How a Utility Company Established a Corporate Data Culture for Data-Driven Decision Making

Philipp Staudt

Rainer Hoffmann

MIS QUARTERLY EXECUTIVE

How a Utility Company Established a Corporate Data Culture for Data-Driven Decision-Making

The increasing importance of data, propelled by advanced AI capabilities, offers opportunities to increase profitability and sustainability. But many organizations face challenges in harnessing data's true value. We describe how a German utility company addressed this challenge by establishing a data culture. The transition involved three interlinked transformations: enabling the workforce, improving the data lifecycle, and employee-centered data management. The recommendations derived from this case will enable business leaders in industrial organizations to navigate future data, knowledge and AI-related challenges.[1,2]

Philipp Staudt
Carl von Ossietzky University of Oldenburg (Germany)

Rainer Hoffmann
EnBW AG, Karlsruhe (Germany)

Need for a Data Culture in the Era of Artificial Intelligence

In 2006, British mathematician Clive Humby coined the phrase "data is the new oil," meaning that, just like oil, data has to be processed and refined to derive value. In 2017, *The Economist* published an article describing how data has become the most valuable resource for organizations.[3] This view is supported by more recent analyses. A 2022 Economist Impact study reports that the European Commission estimates that the EU data economy will be worth €829 billion ($875 billion[4]) by 2025, increasing from 2% to 6% of total GDP.[5] The study also states that the value of data sharing in the manufacturing industry is worth €83 billion ($87.6 billion) in process optimization and will lead to an estimated 20% improvement in resource efficiency, underlining how data can improve sustainability efforts in organizations.

One particular area in which data is of growing importance is as the foundation for many applications that enable the energy and manufacturing industries to address climate change issues.[6] Given the complexity of energy supply and distribution systems, the integration of intermittent renewable generation from wind and solar is a data-intensive problem that

SIM WHERE IT LEADERS CONNECT

1 Rick Watson is the accepting senior editor for this article.
2 We sincerely thank the senior editor for his invaluable guidance, editing and recommendations. We also thank the anonymous reviewers for their very constructive and helpful comments that improved the article. Finally, we thank Tobias Krauss and the other interviewees for the insights they provided.
3 "The World's Most Valuable Resource Is No Longer Oil, but Data," The Economist, May 6, 2017.
4 Currency conversion as at October 2023.
5 *The Future of Europe's Data Economy*, Economist Impact, 2022.
6 Various applications are described in Rolnick, D., Donti, P. L., Kaack, L. H., Kochanski, K., Lacoste, A., Sankaran, K. ... and Bengio, Y. "Tackling Climate Change with Machine Learning," *ACM Computing Surveys* (55:2), February 2022, pp. 1-96.

DOI: 10.17705/2msqe.00088

requires corresponding information systems.[7] The World Economic Forum and Boston Consulting Group report that, of 1,300 manufacturing executives surveyed, 80% consider data and analytics important for their businesses, with 72% stating an increase in importance compared to three years prior. However, only 17% report that they captured satisfactory value from their data.[8] This seems to be a common theme. Various case studies have described the difficulty that organizations face in deriving value from data. As a consequence, there is a widening gap between companies that gain insight from data and those that do not.

Researchers have identified various mistakes made by organizations that are failing to derive value from their data.[9] One particular issue is the perceived complexity and promise of artificial intelligence (AI) models paired with insufficient understanding and a lack of proper communication of AI-derived knowledge. Moreover, management is often not seeking the data needed to underpin new projects but simply using the data already available.[10] These analyses indicate the need for an improved management culture around data-driven approaches in terms of both data and machine learning strategy.

Such a management culture needs to encompass what data to collect and how to develop the necessary technical capabilities. It is necessary to strike a balance between storing all available data and storing data that is likely to be useful.[11] Additionally, many organizations are unaware of the technical necessities of making data available for AI-based models.[12] Data is often scattered in various forms that are not machine readable. Managing a data strategy in the era of AI is a complex task that involves decisions at various organizational levels. Management should take responsibility for data but should also involve other stakeholders within the organization. Given these complexities, it is timely to discuss best practice cases on how to create an organization-wide data culture. This is particularly relevant for organizations that do not have a digital business model and for whom large IS projects are particularly challenging.[13]

In this article, we present the case of a large German power utility company and its distribution system operator (DSO). The case shows that establishing a data culture can be achieved by concentrating on three intertwined transformations. First, managers must ensure that large parts of the workforce are integrated into a data-driven organizational transformation. This includes providing opportunities for employees to receive training on data-driven techniques and providing (digital) spaces for employees to exchange ideas and knowledge on data-driven solutions. Second, the data lifecycle needs to be improved. This includes increasing technical data capabilities and quality gates, as well as assigning clear organizational responsibilities for data quality. Third, the transformation has to be employee-centric. This can be achieved by establishing cross-functional teams to improve data management and ensuring that quick wins from an improved data strategy are made visible early on. Based on the findings from this utility company case, we offer five recommendations for establishing a data culture in industrial and manufacturing companies with a specific sustainability mandate.

Introducing the Case Study

The case company is one of the largest energy utilities in Germany, and its DSO is one of the biggest distribution grid operators in both electricity and gas distribution networks.

7 For a detailed discussion of this data-intensive problem, see Watson, R. T., Ketter, W., Recker, J. and Seidel, S. "Sustainable Energy Transition: Intermittency Policy Based on Digital Mirror Actions," *Journal of the Association for Information Systems* (23:3), May 2022, pp. 631-638.

8 *The Data-Driven Journey Towards Manufacturing Excellence*, World Economic Forum in collaboration with Boston Consulting Group, January 2022, available at https://www.weforum.org/whitepapers/the-data-driven-journey-towards-manufacturing-excellence/.

9 Joshi, M. P., Su, N., Austin, R. D. and Sundaram, A. K. "Why So Many Data Science Projects Fail to Deliver," *MIT Sloan Management Review* (62:3), March 2, 2021, pp. 85-89.

10 de Langhe, B. and Puntoni, S. "Leading with Decision-Driven Data Analytics," *MIT Sloan Management Review* (62:3), March 8, 2021, pp. 1-4.

11 Desai, V., Fountaine, T. and Rowshankish, K. "A Better Way to Put Your Data to Work," *Harvard Business Review* (100:4), July-August 2022, pp. 100-107.

12 Vial, G., Jiang, J., Giannelia, T. and Cameron, A. F. "The Data Problem Stalling AI," *MIT Sloan Management Review* (62:2), January 2021, pp. 47-53.

13 See, for example, Gust, G., Flath, C. M., Brandt, T., Strohle, P. and Neumann, D. "How a Traditional Company Seeded New Analytics Capabilities," *MIS Quarterly Executive* (16:3), September 2017, pp. 215-230.

Table 1. Description of Energy System Technical Terms

Technical Term	Description
Distribution grid	Part of the electricity grid that serves households and small businesses
Distribution system operator (DSO)	Manages the distribution grid; maintains constantly manned control rooms to monitor the grid, among other things
Electrification	Describes the transition from using other forms of energy (such as fossil fuels) to using electricity (e.g., from internal combustion engines to electric vehicles)
Energy transition	The process of replacing fossil fuels in the energy economy with renewable generation, often from wind and solar
Substation	A central node in the electricity grid from which electricity is distributed
Hydrogen	A form of energy storage that is eventually expected to replace natural gas
Intermittency	Variation in wind and solar generation according to weather patterns
(Peak) Load	Power, measured in kW, supplied to or drawn from the grid at a point in time, with the peak being the maximum over time; energy is load supplied over time (kWh).
Photovoltaics (PV)	Technology used to transform solar energy into electricity
Power flow	Power measured in kW, transmitted over a grid element at a point in time

The utility company's DSO manages medium- and low-voltage electricity grids, which involves integrating locally installed renewable generation capacity, such as wind and solar power. Additionally, because of the ongoing electrification of energy (for heating and transportation), the DSO also needs to manage the integration of charging stations for electric vehicles and heat pumps. This will require improved strategies for increasing engineering efficiency, as current labor-intensive practices do not scale well with the increased workload. The need to address these challenges was the trigger for the utility's transformation to a more data-driven company, which was achieved by establishing an organizational data culture. Our comprehensive description of how it achieved this makes it a compelling case study on establishing a corporate data culture.

Some of the terminology used in describing this case is specific to energy generation and distribution systems. To assist in understanding the case, Table 1 describes the most frequently used technical terms used in this industry.

The case description is based on interviews with seven of the utility's domain experts, one of whom is a co-author. These experts cover all technical departments directly involved with the DSO's data strategy and the utility's overall data practices and were therefore able to provide a comprehensive account of the company's data culture. All seven are either directly responsible for ensuring data quality, developing new data-driven solutions with their teams, or supporting the overall structure of introducing data management and data analytics into the company. The interviews were semi-structured and focused on six areas:

1. The opportunities for a data-driven transformation
2. Current operational challenges for grid operation
3. The historical development of the importance and prioritization of data within the company
4. The corresponding data management strategy

The drag pointer is part of a gauge and is depicted by the red pointer in the pictogram. The black pointer shows the currently measured value. When the black pointer goes beyond the maximum value measured since the last reading (where the red drag pointer currently is), then the red pointer is "dragged" with the black pointer to the new maximum. The black pointer later retreats when measured values decrease, but the red drag pointer stays in place.

5. Use cases implemented by the interviewees' departments
6. The outlook for the role of data in future grid operation.[14]

Like similar studies,[15,16,17] we analyzed this specific revelatory case in depth to reveal previously unknown phenomena. Below, we first provide an overview of the need for transformation within the utility and the drivers of this transformation. We then draw on the parallels between establishing a data culture and approaches to knowledge management before describing the main levers of establishing a data culture. Finally, we provide an overview of the resulting transformation and, based on the case findings, offer recommendations for establishing a data culture. Throughout the article, we try to generalize our findings from the utility case company to comparable industrial organizations.

Transformation Driven by the Need to Respond to Energy-Transition Challenges

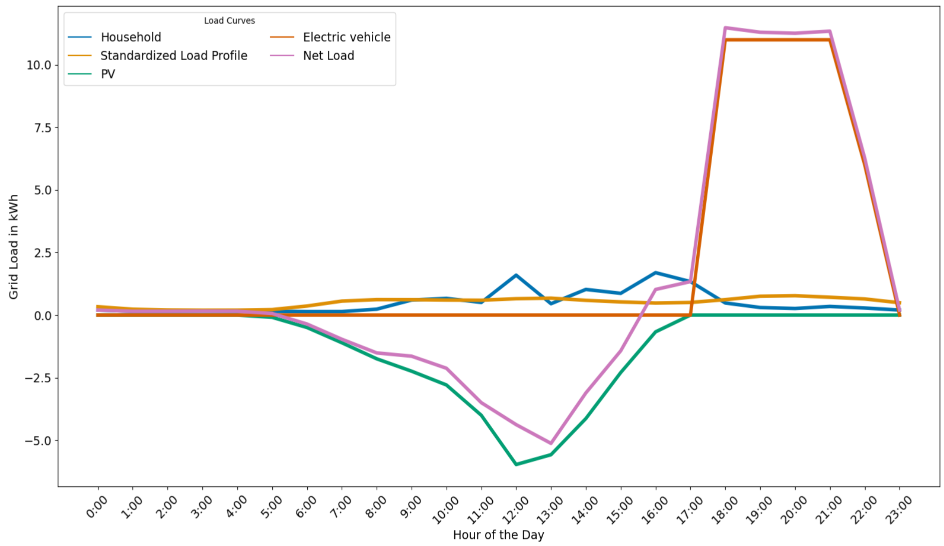
Traditionally, electricity utility companies' planning and decision-making have been largely based on heuristics relying, for example, on the so-called "drag pointer" (described in the box above). The drag pointer captures the maximum load at a substation over a period of time, and this single value is only recorded every few years. The drag pointer does not provide any additional information, such as when the peak occurred or under what circumstances.

In the past, the drag pointer was sufficient for most of the planning for the distribution grid expansion of traditional energy supply systems. The drag pointer approach represents a typical nondigitalized planning process. Broad heuristics are used to oversize the grid, ensuring that capacity constraints do not become binding. However, this approach is no longer appropriate now that electricity grids have to accommodate intermittent renewable generation and increasing demand from new high-powered appliances such as electric vehicles or heat pumps. Grids now have to integrate control algorithms that activate and deactivate appliances on demand and in real time, and planning processes have to take account of a huge increase in consumer requests. This energy transition means that energy utilities are having to adapt quickly, abandoning the traditional planning process and replacing it with data-driven planning approaches.

Previous research has shown that traditional industrial companies such as energy utilities are struggling to build up analytical capabilities.[18] As in many other industries, processes for handling large data volumes and data-driven decision-making in energy utilities have, at best, been mastered in specific departments (such as power trading or sales, where an analytical focus has always been prevalent) or, at worst, are seen as unchartered territory and not part

14 Interviews were conducted based on Schultze, U. and Avital, M. "Designing Interviews to Generate Rich Data for Information Systems Research," *Information and Organization* (21:1) January 2011, pp. 1-16.
15 Seidel, S., Recker, J. and vom Brocke, J. "Sensemaking and Sustainable Practicing: Functional Affordances of Information Systems in Green Transformations," *MIS Quarterly* (37:4), December 2013, pp. 1275-1299.
16 Gust, G., Flath, C. M., Brandt, T., Strohle, P. and Neumann, D., op. cit., September 2017.
17 Kranz, J., Fiedler, M., Seidler, A., Strunk, K. and Ixmeier, A. "Unexpected Benefits from a Shadow Environmental Management Information System," *MIS Quarterly Executive* (20:3), September 2021, pp. 235-256.
18 Gust, G., Flath, C. M., Brandt, T., Strohle, P. and Neumann, D., op. cit., September 2017.

Figure 1. Exemplary Daily Load Curves for a Four-Person Household in Berlin in March



of utilities' DNA.[19] However, the sudden increase in pressure from the accelerating and sustained energy transition faced by the case study DSO meant it had to respond quickly to the challenges of growing electrification of energy demand and the changing structure of supply and demand. It responded to these challenges by establishing an organization-wide data culture in a relatively short time frame.

Electrification of Energy Demand

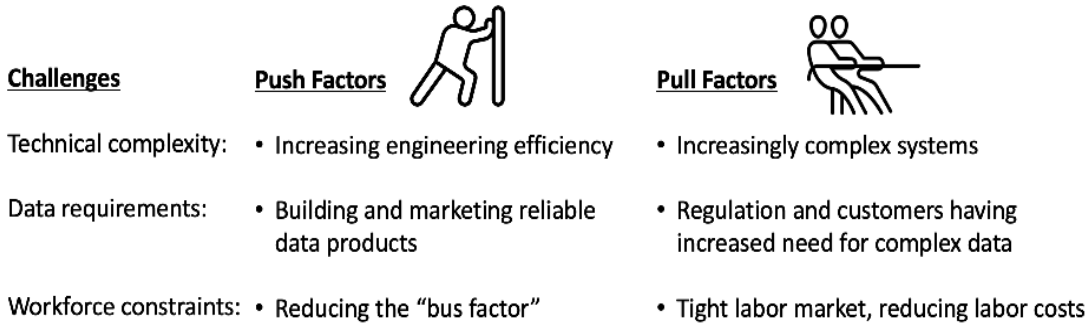
Up to 2019, the DSO received about a dozen monthly requests to install private charging stations for electric vehicles. Both registering and approving the stations require action. A few dozen requests per month was manageable, but by 2022, the DSO was receiving between 2,000 and 3,000 requests per month. One interviewee noted: "We started preparing early for the integration of electric mobility, but it remains a challenge." The DSO also has to approve every new residential PV installation. Whereas the drag pointer previously provided all the necessary information, the system dynamics have changed so that engineers can no longer be sure whether the maximum value displayed by the drag pointer is power consumed behind the substation or power delivered to the substation by local PV installations.

Changing Structure of Supply and Demand

Much of the energy transition to decentralized renewable generation happens in the distribution grid. This means that DSOs are the first to be affected by the energy transition. New appliances are added at their voltage level, such as electric vehicle charging stations or heat pumps, and new local generation is added, such as rooftop PV installations, which generate electricity that is fed into the grid at the distribution system level. To illustrate the related challenges faced by DSOs, Figure 1 shows a daily load profile of a four-person household in Berlin, Germany, in March compared to an assumed standard load profile for a four-person household, the generation from a modest 10kW PV installation in Berlin in March, and the charging profile of an empty electric vehicle with a common 50kWh battery and an 11kW charging station. The figure also shows the resulting net grid load.

19 For a general review of barriers to data science strategies, see Reddy, R, C., Bhattacharjee, B., Mishra, D. and Mandal, A. "A Systematic Literature Review towards a Conceptual Framework for Enablers and Barriers of an Enterprise Data Science Strategy," *Information Systems and e-Business Management* (20:1), March 2022, pp. 223-255.

Figure 2. Push and Pull Factors for a Data-Driven Culture



Responding to New Challenges

Existing grid assets need to perform in a changing environment. For instance, gas distribution pipelines will physically react differently to natural gas that is enriched with hydrogen.[20] The increasing system dynamics of a decentralized energy system are a challenge for DSOs because their systems are not equipped with the necessary sensors to monitor the grid state. As previously mentioned, large parts of the grid are monitored by a drag pointer that is consulted only every few years. Meanwhile, the unmonitored electricity grid needs to service higher loads because of the trend to electrify energy demand.[21] Thus, the traditional approach of simply expanding the grid capacity to match local needs is no longer feasible. Simultaneous expansions throughout the grid will be required, which are prohibitively expensive and take a long time. To better manage existing grid capacities, more data and better real-time knowledge of the grid state will be necessary.

The challenges described above are examples of the trends in the energy utility industry. In this business environment, utilities have to transform their data management quickly and at various levels. Data has to be processed more automatically. Data that was previously siloed is suddenly important for various departments. Errors in data input have to be identified and corrected more quickly. New types of data have to be managed, and additional sensors have to be installed and integrated. These transformations make a DSO a valuable case study of establishing a data culture.

External and Internal Pressure Demanding a Data Culture

We now describe the challenges faced by the case study DSO in more detail at the business level, differentiating between internal push and external pull factors and generalizing them appropriately. A classification of organizational internal push and external pull factors is depicted in Figure 2. This is not a comprehensive list of such factors; it merely describes the factors that were uncovered in the case study.

Push Factors

The utility company's management organized a company-wide workshop aimed at improving business processes. During the workshop, employees identified data quality and access as a pain point. Data was scattered across various SAP databases and engineers were forced to gather data from different sources, which not only caused inefficiencies but also led to inconsistencies when different data in different databases was not synchronized.[22] This made

20 For more information, see Saedi, I., Mhanna, S. and Mancarella, P. "Integrated Electricity and Gas System Modelling with Hydrogen Injections and Gas Composition Tracking," *Applied Energy* (303:4), December 2021.

21 For an overview of electrification in the industrial sector, see Sorknæs, P., Johannsen, R. M., Korberg, A. D. and Nielsen, T.B. "Electrification of the Industrial Sector in 100% Renewable Energy Scenarios," *Energy* (254:1), May 2022.

22 For a broader discussion on this issue, see Vial, G., Jiang, J., Giannelia, T. and Cameron, A. F., op. cit., January 2021.

it difficult to communicate consistent data to external partners.

For instance, there was a problem with a dashboard that municipalities acquired from the DSO, which showed electric vehicle charging stations within their premises. Updating this dashboard was previously handled by a single employee with considerable domain knowledge, which became apparent when the employee switched positions within the company. He still had to be involved because no one else had the required domain knowledge. This issue raised questions about knowledge management. Like other industrial organizations, the so-called "bus factor,"[23] which describes reliance on siloed knowledge of individual employees, was rather high within the DSO.

These challenges can be generalized to other industrial organizations. Increasing the efficiency of well-trained employees is generally desirable. Data products are becoming increasingly important, and it is generally advisable to avoid concentrating knowledge within a few specialized employees.

Pull Factors

To calculate the effects of added infrastructure on electricity distribution grids, such as PV installations or electric vehicle charging stations, engineers have to combine data from various databases to construct a digital twin of a local distribution grid. One interviewee noted that this process took around 30 minutes per calculation and had to be repeated hundreds of times daily—thus a considerable opportunity for potential efficiency gains and increased productivity.

Furthermore, data consistency for distribution grid assets is becoming more important as regulators request an increasing amount of data from DSOs to assess their efficiency, which partly determines regulated payments. Though improving data consistency can be addressed by better data integration, the sheer volume of requests to install charging stations remains a challenge. Recruiting sufficient talent to keep the established manual processes working is difficult and highlights the need for further progress in digitalization and data-driven decision-making.

These pull factors are common because many industrial organizations must deal with increasingly complex systems relating to vendors, customers and supply chains. At the same time, stakeholders often have an increased need for data, such as sustainability scores. Finally, reducing costs and the load on well-trained employees is a general overall objective.

To respond to the challenges posed by the push and pull factors, management of the energy utility and its DSO set about transforming to create more data-driven processes and increase automation, including digitalized end-to-end processes and integrated data management systems.[24] The DSO established a data governance team with a mandate to focus on the processes through which data is generated and stored. One interviewee commented that "given the importance of data quality and integration for management and many employees, our data governance team got a very broad mandate." The management of the DSO committed to becoming a data-driven company by 2030. To achieve this, employees had to subscribe to the idea that data is now a valuable organizational asset.

The transformation succeeded through the establishment of a data culture. Though data cultures are mostly rooted in data management processes, there are parallels with knowledge management concepts. However, AI-based knowledge generation approaches are now challenging some of these concepts. Below, we describe both the parallels and differences between data culture and knowledge management concepts.

Relevance of Knowledge Management Insights for Establishing a Data Culture

Managers tasked with establishing an organization-wide data culture can find valuable insights in the existing literature on knowledge management. The knowledge-based view of a firm posits that knowledge is a scarce resource

23 The rather morbid reason for this term is the risk that an indispensable employee might be hit by a bus. For a discussion of the concept, see Cosentino, V., Izquierdo, J. L. C. and Cabot. J. "Assessing the Bus Factor of Git Repositories," *Proceedings of IEEE 22nd International Conference on Software Analysis, Evolution, and Reengineering* (SANER), March 2015.

24 For a broader discussion on opportunities of data-driven automation in organizations, see Benbya, H, Pachidi, S. and Jarvenpaa, S. "Special Issue Editorial: Artificial Intelligence in Organizations: Implications for Information Systems Research," *Journal of the Association for Information Systems* (22:2), March 2021.

that is needed to ensure competitiveness.[25] We believe this resource perspective might now shift to data and models as they become the foundation for creating knowledge through machine learning.

Knowledge has been described as the capacity to use information. This relates to the common differentiation between data as a foundation from which information can be derived, which is then interpreted in context to create knowledge.[26] Knowledge is part of the overall organizational memory. When not made explicit in organizations through processes or documents, it is personalized and implicit, which is why communication between individuals is key to ensuring a knowledge base.[27] Thus, creating or sharing knowledge is, in essence, equivalent to increasing organizational and human capital,[28] implying that the culture of an organization influences its level of knowledge.[29]

Four levels of knowledge are commonly differentiated—know-what, know-how, know-why and care-why.[30] With data-based knowledge generation using machine learning, the know-why dimension becomes less important. If decisions are based on models and data, instead of communicating domain knowledge per se, employees need to communicate data and the corresponding machine learning models. These components can be made explicit using well-established data management approaches (for instance, a data catalog and corresponding databases). This implies that knowledge about data and machine learning models increases in relative importance over explicit domain knowledge.[31]

As a result, organizations are faced with the challenge of ensuring that data-based machine learning decisions are interpreted correctly, which in turn requires a well-trained workforce that has the knowledge of how to interact with data and machine learning tools.[32] In other words, knowledge is no longer necessary to interpret data but is needed to interpret decisions made by machines based on data. Well-managed data management then becomes a prerequisite for effective knowledge management. The focus of organizational learning therefore shifts from transferring knowledge to transferring data and the corresponding models.[33] This requires organizations to adjust their culture toward data and model management—that is, to establish a data culture.

Establishing a Data Culture

The push and pull factors described above motivated the utility company to embrace data as a resource and to take various steps toward becoming a more data-driven organization. In 2020, AppliedAI published a broad AI maturity assessment of the utility.[34] We build on that study by describing the company's integrated efforts to create a data culture, including a workforce that perceives data as an inherent value generator and the establishment of the corresponding digital infrastructure. The development of the utility company's data culture involved three overarching transformations: 1) enabling the workforce to participate in the data-driven transformation, 2) improving the data lifecycle by moving to asset-centered lifecycle management, and 3) transitioning to employee-centered data management.

Enabling the Workforce

Achieving widespread buy-in from employees was prioritized early in the transformation process. The approaches taken to enable the workforce to participate in the data-driven

25 Conner, K. R. and Prahalad, C. K. "A Resource-Based Theory of the Firm: Knowledge versus Opportunism," *Organization Science* (7:5), October 1996, pp. 477-501.
26 Watson, R. T. Data Management, *Databases and Organizations*, John Wiley & Sons, 2008.
27 Alavi, M. and Leidner, D. E. "Knowledge Management Systems: Issues, Challenges, and Benefits," *Communications of the Association for Information Systems* (1), 1999, pp. 1-28.
28 For a detailed discussion on various capital-creation mechanisms, see Watson, R. T. *Capital, Systems, and Objects: The Foundation and Future of Organizations*, Springer Nature, 2020.
29 Alavi, M. and Leidner, D. E. "Knowledge Management and Knowledge Management Systems: Conceptual Foundations and Research Issues," *MIS Quarterly* (25:1), March 2001, pp. 107-136.
30 For more information, see Quinn, J. B. and Anderson, P. "Leveraging Intellect," *Academy of Management Executive* (10:3), November 1996, pp. 7-27.
31 Benbya, H., Pachidi, S. and Jarvenpaa, S., op. cit., March 2021.
32 Berente, N., Gu, B., Recker, J. and Santhanam, R. "Managing Artificial Intelligence," *MIS Quarterly* (45:3), September 2021, pp. 1433-1450
33 Safadi, H. and Watson, R. T. "Knowledge Monopolies and the Innovation Divide: A Governance Perspective," *Information and Organization* (33:2), June 2023, Article 100466.
34 *EnBW: On the Way to Enterprise-Wide Scaling of AI*, AppliedAI, 2020, available at https://www.appliedai.de/en/hub-en/enbw-enterprise-wide-scaling-ai.

Figure 3. Reinforcing Cycle for Promoting a Data-Driven Culture among the Workforce

transformation can be divided into four intertwined areas: upskilling, enabling use case identification, knowledge sharing and communication. The interaction between these areas creates a reinforcing cycle, which is depicted in Figure 3.

Upskilling: Buy-in of the existing workforce is central to establishing a data culture. A critical component of the data strategy was therefore to upskill willing and capable employees. The utility company provided training opportunities in the areas of data analytics and machine learning. A key objective of this training was to build the self-efficacy and data literacy of employees in relation to data-related problems and tasks.[35] Employees were encouraged to become "Data and AI Multipliers" (see below) in their departments. Though only 20 employees per year go through the top-level training program, many more have the opportunity to improve their basic data skills in an entry-level program.

Enabling use case identification: One major challenge of introducing a data culture is identifying use cases where data and corresponding tools can quickly add value. However, it is difficult for data experts to identify use cases without domain knowledge.[36] To overcome this difficulty, the utility company trained data and AI multipliers. In effect, these employees are ambassadors of data-driven approaches within their departments. They support the identification and implementation of relevant use cases that can add value to the department and can be implemented using data-driven approaches. Note that the data and AI multiplier role is similar to approaches for promoting knowledge management that encourage exchange between employees from different departments.

Knowledge sharing: The growing body of employees trained in handling data and analytics proposed more and more use cases. In response, the utility company established a use case library that enabled employees to share their ideas and solutions centrally. The library ensured that similar ideas could be bundled and tackled jointly. Additionally, contributing and interested employees are being connected through an annual internal week-long conference, which allows them to network, broadcast new training opportunities, and discuss problems and solutions. This conference provides a forum for sharing knowledge on data and models.

Communication: Changing the company's attitude toward data was one of the major challenges of establishing a data culture. This challenge was addressed by management and employees discussing the envisioned transition and corresponding roadmap and clearly communicating the aims and progress of the transition across the company. This communication also included a description of the new vision for the company.

Effort and costs of enabling the workforce: The four intertwined approaches involve effort and costs. The entire process requires resources

35 For a discussion on self-efficacy as a construct in information systems research, see Compeau, D. R. and Christopher A. H. "Computer Self-Efficacy: Development of a Measure and Initial Test," *MIS Quarterly* (19:2), June 1995, pp. 189-211.

36 For recommendations on overcoming these difficulties, see Desai, V., Fountaine, T. and Rowshankish, K., op. cit., July-August 2022.

both from participating employees and from management. The top-level training program requires employees to spend up to 20% of their time on training for an entire year. Participants in this program also do mini-internships over a few days in other departments to cross-fertilize ideas for data-related tools (as previously proposed for knowledge management). The entry-level program requires employees to undertake at least 20 hours of training overall. Additionally, all employees have access to online courses paid for by the company. Overall, there has been widespread participation in the efforts to enable the workforce to participate in the data-driven transformation. The broad managerial buy-in to the concept and a voluntary bottom-up implementation process have so far prevented employee discontent.

Moving to Asset-Centered Data Lifecycle Management

A crucial element of the utility company's efforts to establish a data culture was to overhaul its data management capabilities and approach to data lifecycle management. The company started to treat data as a resource and assigned corresponding responsibilities to the newly created data governance team. This team examined the complete data lifecycles for individual physical assets. The lifecycle stages comprise data generation, curation, maintenance, analysis and exploitation. The team began to analyze how data is recorded and transmitted to find possible origins of errors and possibilities for efficiency gains, focusing on interfaces with manual data transmission. One interviewee stated: "We are trying to replace manual steps along digital processes to increase the speed and quality of our reporting."

Recording new forms of data: Traditionally, data management for a DSO has been mostly limited to processing static data. This data relates mainly to grid topology and assets. However, with the requirement for better real-time monitoring of the grid state, the case study DSO now needs to record and process dynamic real-time data and make it available for real-time calculations and decision support. One interviewee said that "traditionally, this amount of data was simply not necessary. Distribution grid operation was possible without processing it." Another stated: "High-resolution data is essential for the grid state estimation." The challenge of recording dynamic real-time data can be generalized to many industrial organizations. For example, manufacturing firms need to increasingly monitor their shop floors in real time to ensure efficient machine scheduling and to implement predictive analyses.

Recording dynamic data in real time requires an improved data infrastructure architecture. To achieve this, the DSO now operates two different databases. The first one records dynamic data in real time and makes it available for processing. Once this data is no longer up-to-date and does not need to be quickly available, it is moved to the second static database. This database does not allow for the same processing speed but has more storage space.[37]

The second database serves as a data lake and is fed from all available data sources, such as the company's various SAP systems. It provides a one-stop shop for all data needs throughout the company. This facilitates data access and makes working with various databases and sources obsolete.

Data integration: To describe the importance of data integration for the DSO, we consider the process of approving a new private electric vehicle charging station. The engineer has to acquire technical data on the charging station to be installed, the geo-position of the charging point and all other major appliances that might be installed along the same power distribution line. Furthermore, the engineer has to consider the "simultaneity factors" of all installed appliances (i.e., the extent to which they are expected to draw power simultaneously) and the local grid topology. This information has traditionally been scattered across several databases, with some of the data in pdf files. Industrial organizations face similar data integration challenges when adding a new order to the manufacturing queue, integrating new vendors or upgrading the factory floor with new machines or shifts.

Based on the collected data, the engineer evaluating the request to install a new electric

37 The DSO's data architecture is a variation of a lambda architecture. For further information, see Kiran, M., Murphy, P., Monga, I., Dugan, J. and Baveja, S. S. "Lambda Architecture for Cost-Effective Batch and Speed Big Data Processing," *Proceeding of 2015 IEEE International Conference on Big Data (Big Data)*, IEEE, October 2015, pp. 2785-2792.

vehicle charging station creates a model that calculates the time series of load values. This model is used to determine whether a new charging station can be installed safely. These calculations can take quite a long time. But even if they were easy to do, the growing number of requests for new charging stations would remain a challenge. It is therefore important to make data access as easy as possible so the engineer can focus on the core task.

To facilitate easy data access, the utility company took various data integration initiatives. An early initiative was the creation of a data catalog that integrated data from all required available sources. Moreover, individual project teams created applications that further facilitated access to relevant data using intuitive interfaces.

Data quality improvement: Data access and integration do not add value if the quality of the data is poor. Part of the utility company's early effort to improve data management was to establish data quality criteria. The DSO created over 3,000 rules that are applied to individual data when it is being processed. These rules are constantly reevaluated, and new rules are added. For example, it is now possible to see how often a certain rule is triggered. These rules have also led to a cultural shift because they made data quality more visible and objective.

Data governance: The utility company also improved its organizational data processes. The processes were defined for individual physical assets, which made them more tangible for employees. This was particularly innovative because, for the first time, individuals were assigned responsibility for data beyond specific data applications. This represented a major shift in priorities. Data became the central entity rather than being associated with the digital application for which the data is needed.

The newly formed data governance team appointed data officers within the organization who are responsible for further advancing data quality within their departments.[38] These appointments enabled responsibility for data quality to be shared across the organization. Data officers can further appoint data stewards. These employees work within the line organization and play a role along the data lifecycle of a specific asset, such as updating or recording specific dates or changes. Data stewards become more aware of their responsibility in the data lifecycle and the role establishes a sense of responsibility for data quality. It also increases the awareness of the value and importance of data within the line organization. One interviewee remarked: "Data officers and data stewards ensure that data governance is secured in functional teams as well." As with the upskilling of employees, there are costs associated with this approach. Employees have to dedicate time to the maintenance of data quality instead of performing their regular duties.

In addition to appointing data officers and data stewards, the DSO is trying to reduce the possibility of human error affecting data quality. It intends to automate digital and data processes so that: 1) process knowledge is explicitly coded and 2) human errors are minimized. One interviewee stated: "We want to ensure that domain knowledge doesn't get lost. Ideally, it's encoded in end-to-end processes." This process automation makes previously implicit knowledge explicit.[39]

Overview and summary of data lifecycle activities: An overview of the activities along the data lifecycle is shown in Figure 4, which shows that data governance is a pillar of data lifecycle management. The figure also shows the specific functional requirements identified by the utility company for recording new types of data, data integration and data quality improvement, together with the capabilities developed to meet those requirements. First, a new database was created to record and process real-time data. Second, there is now a single point of access to data, which facilitates employee access and ensures transparency of the available data. Third, data quality is ensured through rule-based quality gates and by assigning clear responsibilities for data quality along asset classes. In combination, these capabilities facilitated the more general development of a data culture.

38 Note that the same concept is often called "data stewards" in the literature. However, the DSO differentiates between data officers and data stewards. For more information, see Desai, V., Fountaine, T. and Rowshankish, K., op. cit., July-August 2022.

39 For a discussion on implicit and explicit knowledge, see Alavi, M. and Leidner, D. E., op. cit., March 2001.

Figure 4. Overview of Activities along the Data Lifecycle

	Data Lifecycle Management		
Data Need	**Recording New Forms of Data**	**Data Integration**	**Data Quality Improvement**
Functional Requirement	Dynamic monitoring of grid state	Increasing order volume and engineering load	New reporting requirements and products
Capability Development	Development of capabilities to process and store dynamic real-time data	Creating single-access points to data and a data catalog	Rule-based checks of data correctness and consistency
Data Governance Introducing asset-based data responsibilities, data officers and data stewards; increasing explicitly encoded processes			

Transitioning to Employee-Centered Data Management

The data management approach followed by the DSO sheds some light on the change management needed to establish a data culture. Such a transition cannot simply be mandated or practiced by a few select employees; instead, the established company culture has to change.[40] To ensure a successful transition, the case company therefore used an employee-centric data management approach, which involved the five activities described below. These activities are not specific to the DSO and can easily be adapted for industrial organizations.

Establishing new roles: As mentioned earlier, the data officer and data steward roles are part of the DSO's data strategy. An additional advantage of these new roles is that data stewards provide feedback on observations regarding data processes. This further improves the activities of the central data governance team and contributes to disseminating a data culture throughout the organization.

Forming cross-functional teams: The data governance team was formed using an interdisciplinary approach,[41] with team members having backgrounds in both engineering and computer science. They emphasized transparent and regular communication with employees throughout the organization in diverse roles and with diverse business problems. This communication established trust in the data governance team and led to more acceptance of the changed approach toward managing data. The exchange between team members on the challenges of communicating with more reserved employees was actively encouraged. This was particularly important because the business challenges of the target groups varied widely. The groups range, for example, from specialists with Ph.D. degrees in electrical engineering who have to decide on future grid designs, to field technicians who repair power lines and need to find the cause of a line fault.

Cross-functional teams were also formed for any data-related project. This approach increased

40 This is the same as the need for a change in company culture for knowledge management, as discussed in ibid.

41 This approach is similar to that proposed by Desai, V., Fountaine, T. and Rowshankish, K., op. cit., July-August 2022.

the commitment of specific departments and ensured alignment with business needs, and also broadened the data culture throughout the entire company.[42]

Showcasing quick wins: Another success factor was showcasing quick wins that demonstrated to employees the benefit of improved data management. For instance, the data governance team automated the extensive reporting on electricity pole replacements in the distribution grid. This report was originally published three times per year and preparing it required one person-month from each of two employees. After the data overhaul, the data required for the report can be monitored in real time. The availability of the automated report reinforced top management buy-in and demonstrated the value of improved data management. Our interviewees saw this as a success factor, with one remarking: "One of our success factors is demonstrating tangible progress along the development process."

Boosting credibility with external support: Based on the initial results achieved by improving data governance, the data governance team began to propose further use cases. For instance, the team developed a proof of concept for predictive maintenance of the gas infrastructure. To boost credibility within the organization, this concept was developed together with a group of researchers from a large German university and the approach was published in a peer-reviewed journal.[43]

Communicating transparently and consistently: Another success factor was transparent communication and frequent interactions between employees and developers. In some cases, an extensive shadowing and research phase was added before a project's operational kick-off. This was particularly helpful in cases where many previous digitalization efforts had failed, leading to frustration among employees. One interviewee noted: "We promoted a very communicative culture within the data governance team and with all our stakeholders." Based on this extensive preparation, developers established "click-dummies" (clickable prototypes that enable early feedback from users) with the necessary functionality, creating employees' trust in the renewed effort.

Summary of transitioning to employee-centered data management: The five activities involved in transitioning to employee-centered data management were accompanied by training opportunities that allowed employees in the line organization to become more proficient in using data for analytical tasks. Overall, these activities and the associated training can be depicted as a waterfall model (see Figure 5).

Building Data-Driven Decision Support Systems

The transition to a data culture has led to various improvements throughout the utility company. Below, we describe three successful data-driven decision support systems implemented or under development at the DSO that contribute to better integration of renewable energy generation. These use cases show how an organization-wide data culture can be leveraged to improve sustainability. We relate these DSO-specific use cases to more general applications in industrial organizations.

The energy transition requires DSOs to become more innovative along their entire value chain and over different time horizons. At the energy utility's DSO, this can be broken down into three distinct areas: 1) grid operation—i.e., real-time operations to assure grid stability; 2) grid maintenance—i.e., day-to-day repair and replacement of components; and 3) grid strategy—i.e., the long-term management of grid expansion. Figure 6 provides an overview of these three use cases and relates them to the time horizon, which ranges from short term to long term. It also describes equivalent use cases in industrial organizations.

Real-Time Operation

DSOs use the System Average Interruption Duration Index (SAIDI) to measure the average power outage duration within a distribution grid area. This index is used to benchmark DSOs against each other and therefore has

42 A broader discussion of the interaction of teams during the development of information systems can be found in Sawyer, S., Guinan, P. J. and Cooprider, J. "Social Interactions of Information Systems Development Teams: A Performance Perspective," *Information Systems Journal* (20:1), January 2010, pp. 81-107.

43 Betz, W., Papaioannou, I., Zeh, T., Hesping, D., Krauss, T. and Straub, D. "Data-Driven Predictive Maintenance for Gas Distribution Networks," *ASCE-ASME Journal of Risk and Uncertainty in Engineering Systems*, Part A: Civil Engineering (8:2), June 2022.

Figure 5. Waterfall Model of the Employee-Centric Data Management Approach

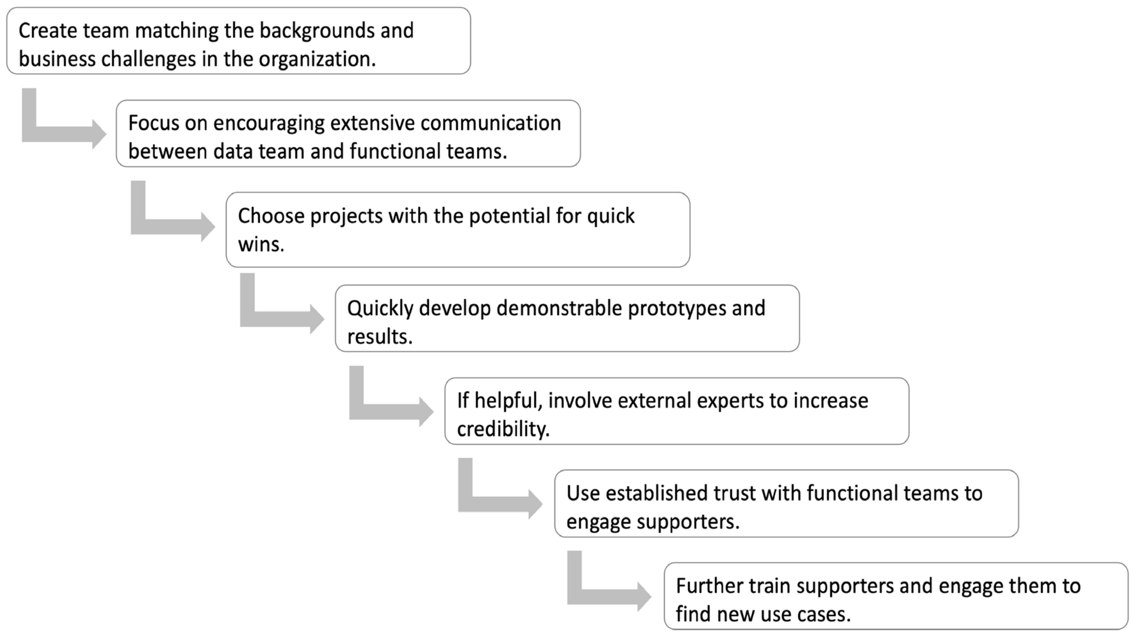


Figure 6. Three DSO Use Cases, Their Time Horizons and Equivalent Industrial Use Cases

Operational Level	Real-time Operation	Maintenance	Planning
Business Challenge	Increasing system dynamics require real-time knowledge of grid state	Changing requirements of components	Very dynamic development of local loads
Heuristic	Build grid to withstand expected load	Static replacement schedules	Periodic consultation of drag pointer and corresponding grid expansion
Data-Driven Solution	Monitor grid state using strategically installed sensors	Predictive maintenance based on dynamic measured values	Data-driven decision support based on known local grid structure
Equivalent Industrial Use Cases	*Managing system complexity through oversized capacity vs. optimizing system dynamically in real time.*	*Machine maintenance schedule through regular scheduled inspection vs. predictive maintenance*	*Strategic decisions on system based on static key performance indicators vs. detailed digital twin*
	Short Term	Time Horizon	Long Term

consequences for the bottom line. Thus, not having access to real-time data about the state of the distribution grid and the impact of large intermittent power sources can have financial repercussions. To increase visibility into the state of the distribution grid, the case study DSO deploys sensors within the grid that monitor relevant state values, essentially building a digital twin of the power grid. One interviewee stated: "It is our objective to equip a small portion of the grid with sensors allowing us to monitor the entire grid." In other words, the DSO is aiming to

  |

monitor the grid as efficiently as possible because equipping every substation with sensors would be prohibitively expensive. To achieve this aim, the DSO has built a machine learning system for state estimation based on a small number of strategically distributed sensors. For industrial organizations, the equivalent problem is real-time factory floor monitoring, where various machines and components would have to be equipped with sensors.

Maintenance

To minimize power outages caused by component failure, the DSO uses a maintenance and replacement strategy. The importance of having such a strategy was brought to public attention in 2018 when a century-old C-hook broke in Pacific Gas and Electric Co.'s California grid, causing a wildfire that killed dozens of people.[44] Traditionally, the DSO's strategy would be to simply replace equipment after a certain time in service, but even this strategy requires accurate data records. Today, however, in the changing environment in which DSOs now operate, such a static maintenance approach may no longer be suitable. An interviewee said: "Our predictive maintenance approach also helps us to monitor assets under conditions with which we have less experience." The equivalent industrial application is the predictive maintenance of existing and newly added machines.

Planning

There is a broad consensus that power distribution grids have to be expanded in response to the transition from hydrocarbon fuels to mostly electrified energy consumption, which will be a significant challenge for DSOs. Traditional heuristic approaches to planning and expanding the grid are no longer sustainable for this challenge. Periodic consultation of drag pointers and the corresponding lengthy planning processes, with the only possible response being capacity expansion, are inadequate in an energy system with large intermittent energy generation sources. The utility company's DSO has already laid the groundwork to better understand where the grid needs to be reinforced and where other smart solutions can be deployed. By providing easily accessible grid data, the DSO can support engineers to better simulate and plan the integration of new appliances and resources at the distribution grid level. This leaves more time for engineers to assess complicated cases and prioritize necessary grid expansion. For industrial organizations, the equivalent is the strategic management of complex systems such as factory floors and supply chains.

Successful Deployment of Data-Driven Decision Support Systems

In summary, the DSO's newly developed data culture has led to data-driven processes that have replaced former heuristics-based decision-making. This cultural shift continues to pay dividends in terms of data products that make processes more efficient. This is not only beneficial for customers but is paramount for ensuring that the electrification of energy demand and the integration of volatile renewable generation is successful, leading to a more sustainable energy system.

Results of the Transformation and Future Challenges

The utility case company now operates more than 50 data-driven solutions across several business units. The central team to drive the transition to a data culture has more than 15 employees in roles such as data strategists, designers or change coaches.

Moving forward, data and AI-based decision-making now play such a crucial role in the utility that it is involved in discussions on the forthcoming EU AI Act.[45] The increased use of AI-based tools also causes the need for organizational adaptation. Electricity utilities are part of a country's critical infrastructure. Engineers are therefore required to work with particular prudence. With more and more decisions being automated, the responsibility for these decisions has to be attributed. Can engineers justify their actions with the results of AI-based decision support tools or are they still responsible, and if so, what does this mean for the

44 "Revealed: the PG&E hook that started the Camp Fire," *San Francisco Chronicle*, December 4, 2019, available at https://www.sfchronicle.com/california-wildfires/article/Attorneys-say-this-photo-shows-the-PG-E-hook-that-14882924.php.

45 *EU AI Act: First Regulation on Artificial Intelligence*, European Parliament, June 14, 2023, available at https://www.europarl.europa.eu/news/en/headlines/society/20230601STO93804/eu-ai-act-first-regulation-on-artificial-intelligence.

design of the tools? The need for this discussion will be relevant in many industrial organizations over the next decade. Moreover, the EU AI Act might add further requirements for operators of critical infrastructure and AI technology.

Recommendations for Business Leaders

Based on the utility company's transition to a data culture, we provide five recommendations for industrial organizations, relating them to the three transformations the utility undertook as it established its data culture. The first recommendation relates to enabling the workforce, the next two to asset-centered data lifecycle management and the final two to employee-centered data management.

1. Ensure Broad Company Participation

One factor of the utility company's successful transition to a data culture was offering training to a broad group of employees. This ensured that they would not feel left behind and were able to keep track of the transformation. Based on this upskilling, they became involved in use case identification and participated in (virtual) spaces where they could interact with other like-minded employees and exchange ideas, which created a dynamic, positive environment for developing new ideas. This environment fostered a positive image of the company's transformation and encouraged more employees to participate.

2. Improve Capabilities Along the Data Lifecycle

The purpose of an organizational data culture is to avoid employee frustration when working with data. This means the data lifecycle must be closely examined to better understand how corresponding processes can be improved. A good understanding of the data lifecycle will further ensure that data quality issues can be clearly identified and localized, which enables data management teams to target actual rather than perceived problems. In part, improving capabilities along the data lifecycle also means adding technical capabilities to data management teams, such as data engineering, database management and data analysis experts. These experts can then assess critical interfaces along the data lifecycle to ensure correct data recording, data consistency and integration, and quality improvement.

3. Establish Data Quality Responsibilities

When a company treats data as a resource, it has to establish defined responsibilities for assuring data quality. The utility company's DSO achieved this by defining individual data lifecycles for physical assets, which allowed the DSO to assign responsibilities based on asset classes. Other companies might choose to establish data responsibilities according to processes. However, ensuring individual accountability for data quality establishes ownership for data-generating processes and increases awareness of the role of data as an asset.

4. Form a Cross-Functional Data Governance Teams

Data needs might differ greatly across a company's departments depending on business problems or use cases, which means it is difficult for an independent data governance team to ensure that data integration and processes correspond to broad organizational needs. We therefore recommend that companies staff data governance teams cross-functionally, with representatives from multiple departments. This will not only broaden the team's knowledge base but will also ensure a broader acceptance of data governance activities across the different departments. Additionally, a cross-functional team can facilitate communication between the different departments.

5. Showcase Quick Wins

Any successful transformation requires the broad support of the workforce. A good way of ensuring broad acceptance of a data culture is to showcase quick wins. For instance, regular reports could be automated with online access to current data. This would ensure buy-in from those managers who value the increased transparency. Early prototypes could motivate departments that might have been disappointed by digitalization projects in the past. Showcasing quick wins will promote moves toward improved data integration and quality, allowing the development of more sophisticated machine learning solutions.

Concluding Comments

The increasing capabilities of AI-based models make data a valuable resource for organizations seeking to increase profit and become more sustainable. DSOs face many challenges in terms of resource management and the integration of new forms of energy, and these challenges are data intensive and require a well-maintained data basis. More generally, industrial organizations still struggle with deriving value from data. To maximize the value derived from data, organizations need to establish a data culture.

The utility company case described in this article shows that, though establishing a data culture is rooted in data management, there are similarities to the implementation of organizational knowledge management. The case also demonstrates how the essence of knowledge management changes with the integration of machine learning, from sharing knowledge directly to sharing knowledge about data, models and corresponding results. In other words, the extent of know-why knowledge within an organization is diminished.

We have described how external and internal pressures forced the utility company to adapt and how it established a data culture as it transitioned its decision-making from heuristics to (partly) automated data-driven approaches. The transition involved taking actions in three intertwined areas: 1) enabling the workforce, to ensure broad participation in the process; 2) improving the data lifecycle, to ensure better technical capabilities and accountability for data quality; and 3) transitioning to employee-centered data management, to avoid employee frustration with the data-driven transformation. The five recommendations derived from the case can help business leaders in industrial organizations successfully establish their own data cultures.

Finally, the case shows that even early preparation for anticipated challenges does not always guarantee that an organization will stay in front of the curve. It is therefore also a cautionary tale for business leaders, who should prepare for changing requirements for managing data, knowledge and machine learning models.

About the Authors

Philipp Staudt

Philipp Staudt (philipp.staudt@uol.de) is a professor of environmental and sustainable information systems at the Carl von Ossietzky University of Oldenburg, Germany. His research focuses on various aspects of digital technology and sustainability, including the possibilities for using digital technology for energy-related decision-making and the corresponding impacts on energy markets, the role of software for sustainability strategies in organizations, and sustainable digital platforms. He received his Ph.D. from the Karlsruhe Institute of Technology in 2019 and subsequently worked as a postdoctoral researcher at KIT before joining MIT on a Marie Skłodowska-Curie Fellowship.

Rainer Hoffmann

Rainer Hoffmann (ra.hoffmann@enbw.com) is chief data officer at EnBW AG, one of Germany’s largest utility companies. Previously, he was responsible for implementing the data and AI strategy throughout the company and establishing data-driven methodologies in the trading department. Focusing on the intersection of technology and organizational structures, Dr. Hoffmann emphasizes the human dimension of increasingly relying on AI and machine learning. He holds a Ph.D. in operations research from the Karlsruhe Institute of Technology and a post-graduate certificate in machine learning and artificial intelligence from MIT.