

Secondary Publication



Hopf, Konstantin; Müller, Oliver; Shollo, Arisa; Thiess, Timo

Organizational Implementation of AI : Craft and Mechanical Work

Date of secondary publication: 20.10.2023

Version of Record (Published Version), Article

Persistent identifier: urn:nbn:de:bvb:473-irb-91307

Primary publication

Hopf, Konstantin; Müller, Oliver; Shollo, Arisa; Thiess, Timo: Organizational Implementation of AI : Craft and Mechanical Work. In: California Management Review 66(1) 2023, Thousand Oaks : SAGE Publications, S. 1-25. - DOI: 10.1177/00081256231197445

Legal Notice

This work is protected by copyright and/or the indication of a licence. You are free to use this work in any way permitted by the copyright and/or the licence that applies to your usage. For other uses, you must obtain permission from the rights-holder(s).

This document is made available under a Creative Commons license.



The license information is available online:

<https://creativecommons.org/licenses/by/4.0/legalcode>

Organizational Implementation of AI: CRAFT AND MECHANICAL WORK

California Management Review
1–25© The Regents of the
University of California 2023

Article reuse guidelines:

sagepub.com/journals-permissions

DOI: 10.1177/00081256231197445

journals.sagepub.com/home/cmrv

Konstantin Hopf¹, Oliver Müller², Arisa Shollo³,
and Tiemo Thiess⁴

SUMMARY

Recent years have brought major technological breakthroughs in artificial intelligence (AI), and firms are expected to invest nearly \$98 B in 2023. However, many AI projects never leave the pilot phase, and many companies have difficulties extracting value from their AI initiatives. To explain this contradiction, this article reports on a study of 55 projects implementing AI in organizations. It shows that organizational challenges in implementing AI projects are a result of a paradoxical tension created by two different perspectives on data science work: craft and mechanical work. Executives, managers, and data scientists should actively manage this tension to enable and sustain value creation through AI.

KEYWORDS: artificial intelligence, value creation, organization, workplace, implementing innovation

Imagine a company that is *planning*¹ cutting-edge artificial intelligence (AI)² applications. What likely springs to mind is a picture of structured, predictable, and efficient work: a group of data scientists, highly motivated and perfectly trained, sitting in an office, designing, developing, and implementing AI applications that support or automate current or new business processes. Through standardized software components for machine-learning (ML), large amounts of high-quality data, structured work practices, and frictionless orchestration of internal and external resources, these data scientists can come up with and deploy new AI applications in a short time. At the same time, the new AI applications are seemingly integrated with transactional systems and continuously maintained, upgraded, and released in short cycles.

¹University of Bamberg, Bamberg, Germany

²Paderborn University, Paderborn, Germany

³Copenhagen Business School, Frederiksberg, Denmark

⁴PFA Pension, Copenhagen, Denmark

This imaginary scene is far from current reality in most firms. Apart from a handful of unicorns, organizations are struggling to extract value from their AI initiatives.³ Despite several calls from scholars to further integrate AI in business processes⁴ and in decision-making,⁵ many companies run only ad-hoc pilots or apply AI in just a single business process.⁶ But companies must move from pilots to company-wide programs⁷ to generate sustainable business value from AI.⁸ In current practice, the development of AI applications is characterized by individuals or small teams who compensate for their limited resources with a wide range of experience, a good dose of creativity, experimentation, passion, commitment, and mastery of details.⁹ Formulation of ideas and their implementation are a constant learning process, and the absence of established standard practices in building AI applications dominates. Hence, the current reality of AI application development work is more like *craftwork* than structured or plannable work.

To create and sustain business value from AI, data science work is emerging as an important function in organizations,¹⁰ where data scientists together with other actors are responsible for developing and running AI applications. While it is highly valued, data science work does not come without challenges. The question for corporate leaders is how to manage and overcome the challenges. To narrow this knowledge gap, we investigated 55 projects that had the goal to set up and operate AI applications based on (ML)—the core technology underlying current AI systems. Our main informants were data scientists, the people who carry out data science work to create AI applications.¹¹ They are involved in such projects from start to end, thus, they are key resource orchestrators in AI,¹² and they know both the challenges and best practices from their day-to-day professional lives.

Our inquiry confirmed several challenges of AI development that earlier research works have already identified.¹³ For example, there are false hopes toward the technology, data quality and access problems, legal issues, and technical hurdles in bringing AI into operations. It appears that the challenges persist over time. Searching for underlying causes and explanations for the challenges led us to identify different perspectives on data science work, namely, craft, and mechanical work,¹⁴ each having a different take on the process of developing AI applications. Our data indicate that management often pursues a mechanical work view on AI projects (manifested in the assumptions that AI projects are plannable, structured, relying on commoditized resources, etc.), which is encouraged by the science label attached to “data science” and the automation goal that many AI projects have. Data scientists, in contrast, often have a craftwork perspective on their work (manifested, for example, in their all-roundness, mastery, and dedication).

These two partly conflicting perspectives create a tension because managers apply (conventional) business management methods and mindsets to manage data science work and AI applications (e.g., business proposals, business plans, and deliverables). This often clashes with the data scientists’ methods and their mindset of exploration and experimentation, iterative work, and unfinished products. For

creating business value, however, both mindsets are essential. While data scientists' craftwork perspective facilitates the skillful use of tools to exploit data as their material to *effectively* create AI applications, mechanical work promotes *efficiency* in the exploitation of AI technologies. The exploration–exploitation tension is a well-known paradox in the strategic and management literature.¹⁵

A paradox is constituted by a tension that places competing demands. These demands are contradictory, yet, interdependent and persist over time.¹⁶ As these tensions cannot be fully resolved, organizations need strategies for engaging with and accommodating them.¹⁷ Translating the paradox theory to the context of our study means managers and data scientists who neglect to consider either perspective as important to value creation—and, accordingly, do not develop strategies to address these challenges in planning, developing, and implementing AI applications—will fail to deliver value-adding AI applications.

Our article is intended to inform managers and data scientists about this tension and its potential consequences. If firms accept this tension relationship as paradoxical and pay attention to the competing demands, they can initiate positive developments.¹⁸ Learning how to actively manage the paradox will help managers to overcome their struggles with AI implementation. We find that many organizations are lacking managerial strategies to address the paradoxical tension. Instead, data scientists are the ones who address it by ad-hoc tactics.

Perspectives on Data Science Work: Mechanical Versus Craftwork

Mechanical Work Perspective on Data Science

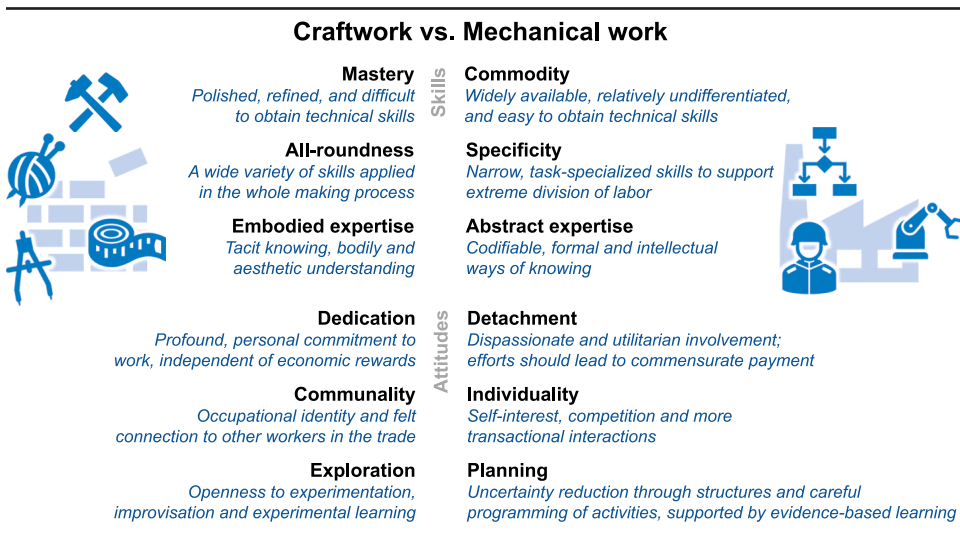
Following the ideas of scientific and bureaucratic management, organizations have for a long time pushed for production and organizational forms that are trimmed toward efficiency and consistency.¹⁹ Hence, since the Industrial Revolution, mechanical work has been the dominant approach to organizing work. Mechanical work is characterized by a planning mindset,²⁰ a form of working that is controlled, highly structured, and predictable. In this way, a mechanical approach to work means that “critical aspects of the process are performed by machines and remaining areas of human involvement are in the form of programmable and marketable tasks”²¹ and “in which the quality of the result is beyond the control” of the individual worker.²² Mechanical work is characterized by worker skills that are easily obtainable (commodity), narrow, and task-specific to support extreme division of labor (specificity). Workers also draw on codifiable knowledge (abstract expertise). In terms of attitudes, they expose a utilitarian involvement with their work (detachment), have rather transactional attitudes to their workplace interactions, and do not follow a strong occupational identity (individuality). Mechanical work is also characterized by established structures and uncertainty reduction through careful programming of activities (planning). We summarize the skills and attitudes of mechanical work in the right part of Figure 1.

Aspects of the mechanical work perspective are evident when looking at data science work. At this, it is important to distinguish between the goal of deployed AI and the process of deploying AI. Many shiny AI showcases, especially end-user applications, create the impression that it is easy to set up corporate ones, as well. This standpoint is supported by existing process models for data science work,²³ which suggest detailed steps and deliverables. They create the impression of plannability. Powerful software tools are available as third-party services²⁴ and vendors are continuously working to make them easier to use.²⁵ This leads to the commoditization of resources. In addition, several algorithmic developments help to automate data science and ML processes: AutoML, which lets algorithms search for ideal parameters and pre-trained models that allow the application of predictive tools to operate in wider ranges of applications. This all helps

to reduce the amount of time and specialized skills required to generate, deploy, and maintain predictive models, by automating the most repetitive steps of the data science [work]. This automation could help data scientists to accelerate the pace of these steps and focus more on other important aspects of analytics.²⁶

Over time, data science skills also become more available and specified. Educational offerings proliferate and expand the pipeline of talent.²⁷ This large base of codified knowledge together with tools that automatically extract information from data let us interpret the contiguity of abstract expertise. While clear evidence has yet to emerge, growing team size and the accompanying specialization within data science teams (which we see in distinct roles for data engineering, visualization, collecting, and cleaning data)²⁸ indicate increased individuality

FIGURE 1. Contraposition of craft and mechanical work according to Kroezen et al.



Note: Jochem Kroezen, Davide Ravasi, Innan Sasaki, Monika Żebrowska, and Roy Suddaby, “Configurations of Craft: Alternative Models for Organizing Work,” *Academy of Management Annals*, 15/2 (July 2021): 502-536.

among employees. Similarly, the current high turnover of data scientists (many stay less than two years at the same company)²⁹ indicates a detachment of these professionals from their employers.

Craftwork Perspective on Data Science

Kroezen et al.³⁰ conceptualize craftwork as an “approach to work that prioritizes human engagement over machine control.” It “implies granting individuals—as ‘makers’—autonomy and control over all facets of a work process, from design to execution.”³¹ Thus, craft relies on refined and difficult-to-obtain skills (mastery) in a broad domain of application (all-roundedness) but exhibits valued skills and has an understanding of aesthetics (embodied expertise). Crafters have a profound and personal commitment to their work (dedication) and are open to experimenting, improvising, and learning by doing (experimentation). They are engaged with their “occupational identity and felt connection to other workers in the trade” (communality).³² We illustrate the contraposition of craft and mechanical work in Figure 1.

While deployed AI applications themselves fit perfectly into the world of mechanical work, the construction of AI seems to have more craft-like features than characteristics of mechanical work. The *all-roundedness* of data scientists becomes visible in their job descriptions, depicting them as a “hybrid of data hacker, analyst, communicator, and trusted adviser,”³³ but also pointing out that “data science is more than analytics/statistics. It also involves behavioral/social sciences (e.g., for ethics and understanding human behavior), industrial engineering ... and visualization.”³⁴ Their *embodied expertise* becomes visible in the purposeful selection of tools to solve problems. Hammerbacher, who created the world’s first data science team at Facebook, portrayed a data scientist as

a team member [who] could author a multistage processing pipeline in Python, design a hypothesis test, perform a regression analysis over data samples with R, design and implement an algorithm for some data-intensive product or service in Hadoop, or communicate the results of our analyses to other members of the organization.³⁵

This description illustrates the *mastery* of their subject. Data Scientists also show high *dedication* with their work, as they engage deeply with data and “often spend a great deal of time decorating simple plots with additional color or symbols.”³⁶ Given the unclear path toward a solution to data-driven problems, data scientists often need to *explore*, as it is their daily business to discover, capture, curate, design, clean, prepare, and analyze data³⁷ in order to develop AI applications. Because of the complexity of the to-be-solved problems, “they need close relationships with the rest of the business. The most important ties for them to forge are with executives in charge of products and services rather than with people overseeing business functions.”³⁸ Our interviewed data scientists also frequently attended meetups and BarCamps, while they are also active participants in several data science communities (online and local industry-based initiatives) where they exchange knowledge with others. Thus, *communality* is important for them.

Five Challenges and Data Scientists' Tactics

In our study of 55 ML-based AI projects, we found evidence of the tension between the managerial (mechanical work) and the data scientists' (craftwork) perspectives on their work. The tension particularly manifests in five challenges that appear along the complete project lifecycle of AI initiatives: During scoping, we found inflated management expectations; during planning, AI projects were often treated like traditional IT projects. During the execution of the projects, missing links to transactional systems and data issues occurred, when initial results were discussed the question of "why" is central. Dynamic environments finally make it difficult to deploy AI models into operations. In the lack of strong organizational strategies to manage the tension, we identify tactics to mitigate each challenge. We summarize all challenges in Table 1 with the manifestation of managers' mechanical work perspective and data scientists' craftwork understanding that are spanning the tension, together with identified tactics to cope with it. In the appendix, we explain our data collection and research method in detail.

Inflated Management Expectations

Many of our interviewees reported that their departments are overrun by requests for AI projects. It just seems that nowadays, "a lot of investors and board members expect to have AI in the process," as one of the interviewed data scientists from a media company stated. Many of these requests contain unrealistic expectations assuming that AI applications consist of distinct parts of purchasable or open-source software that just need to be put together or need to be executed (*commodity* and *specificity*). Such expectations are being fueled by innovations such as AutoML. A data scientist from a global pharmaceutical company complained:

the biggest challenge is the people who want intermediate insights. They [always] want to know what you're up to and what you're finding ... That would be the managers and project leaders. What we then sometimes do, which is actually bad practice, is we work for a whole week on solving the real problems and then on the last day before the weekly catch-up meeting, we do something that they want to see.

He even went so far to say that

the hardest part in this field is not doing the work [developing the ML algorithm and model] ... It is the connection between the Data Scientists and the project managers. That is very frustrating for me ... They want to see green check marks. They're not used to this approach.

This view leads to the impression that software components can be put together in a structured way, which would make such projects *plannable*. Thus, we found that management seems to often go for "very ambitious AI projects,

TABLE I. Identified Challenges, Tensions, and Tactics When Implementing AI in Organizations.

Challenge	Tension		Tactics to Cope with the Tension and Ensure Value Creation
	Managements' View (Mechanical Work Perspective)	Data Scientists' View (Craftwork Perspective)	
C1: Inflated management expectations	<ul style="list-style-type: none"> Expectation to have AI technology in the process Commoditized AI software (purchasable, open source) and automatic approaches (e.g., AutoML) create the illusion of plannability 	<ul style="list-style-type: none"> Dedication to develop good models because of the complex task and variety of tools Communicate to a variety of stakeholders 	<ul style="list-style-type: none"> Establish tools for a quick reality check on what AI can or cannot do Start small, then iterate, and grow Engage with external experts to kickstart AI projects
C2: Managing AI projects like IT projects	<ul style="list-style-type: none"> Assumption that ML models are just computer programs that are easy to create and modify Expectation that a continuous delivery of results is possible 	<ul style="list-style-type: none"> AI applications are learning agents, not deterministically programmed systems Even established agile software development approaches cannot fully account for the exploratory and collaborative nature of AI projects. AI projects are more like research than IT projects 	<ul style="list-style-type: none"> Data scientists train managers on the job Basic AI and ML courses for managers. Appropriate KPIs
C3: Missing data in- and output links to existing systems	<ul style="list-style-type: none"> Interoperability of AI and IT seems to be a matter of interfaces 	<ul style="list-style-type: none"> Maneuvering through the sea of data requires exploration, all-roundness, and mastery of tools Establishing data pipelines is rather plumbing than engineering 	<ul style="list-style-type: none"> Bridges between AI and IT Process-integrated prescriptions instead of predictions Learning apprentice approach Increase data awareness
C4: The question of why	<ul style="list-style-type: none"> Decision makers demand explanations and prescriptions in high-stakes and uncertain situations Legal and ethical requirements demand interpretability Because ML models are codified knowledge (software code), it must be possible to explain their logic 	<ul style="list-style-type: none"> ML models are black boxes and based on correlations only, but they are typically more accurate than traditional methods of prediction 	<ul style="list-style-type: none"> Post-hoc explainability methods Push intrinsically transparent models to the max Causal modeling and rigorous experimental evaluation
C5: Dynamic environments	<ul style="list-style-type: none"> Traditional software, once put into production, runs over years 	<ul style="list-style-type: none"> Predictive accuracy of ML models degrades over time In-house employees do not know how to update or retrain models 	<ul style="list-style-type: none"> Monitoring and detection of data drift Combining AI development and operations (e.g., MLOps)

Note: AI = artificial intelligence; IT = information technology; ML = machine-learning; MLOps = machine-learning operations; KPIs: key performance indicators.

where they basically say in one sentence what they want, but nothing is clear at all about what we really need in detail,” reports a data scientist from a heating manufacturer.

Unrealistic expectations of *commoditized* AI software, a kind of mental accounting,³⁹ can lead to disappointments about the return-on-investment of AI pilot projects. This, in turn, can lead to too early withdrawal of planned AI investments—a phenomenon known as de-escalation of commitment⁴⁰—inhibiting any kind of business value to be realized.

The inflated management expectations increase due to misconception of what data science work is about. Instead of just using standardized software components (*detachment*), data scientists are deeply involved in creating meaningful models and display a high level of *dedication*. Our interviewees reported that their work requires *all-roundedness* (e.g., creating a wide variety of models and effectively communicating to a variety of stakeholders), *mastery* (e.g., employing the right tool to the problem at hand), and tacit, hard-to-articulate, knowledge to make AI applications to work (*embodied expertise*). A data scientist at a large IT software firm explained: “Automation is actually one of the challenges [where we have] to manage expectations, because people think much of AI—Fantastic! But it is not really like that ... Models come with errors, you cannot do everything, and you must explain the limitations.”

We found that data scientists as craftsperson, who master their profession, employ the following *tactics* to manage inflated management expectations:

- *Establish tools for a quick reality check on what AI can or cannot do*—Ultimately, data scientists are the ones who may decide which parts of an application can be supported or automated by AI technologies. They employ their mastery of a very diversified toolset, their embodied expertise (their situated knowledge of the culture and business processes), and the holistic approach they follow in tackling problems that allows data scientists to come up with these decisions. Yet, to avoid burdening busy data scientists with the constant evaluation of business ideas for AI suitability, they should develop tools and practices, like documents containing their learning cycles or interactive templates for data quality checks to speed up the decision process whether a problem is solvable or not. In addition, managers can use universally accepted checklists AI application feasibility, for example, the “Checklist for assessing a task’s suitability for supervised ML”⁴¹ for initial assessments,
- *Start small, then iterate, and grow*—Running AI pilot projects and growing these projects in an iterative fashion⁴² in close collaboration with domain experts is another way to manage expectations. It is often more promising to start with “low-hanging fruit” projects that incrementally improve business processes than to directly go for disruptive “moon shots.”⁴³ One interviewee from a global pharmaceutical company described this collaborative and incremental AI journey as follows:

You know, there are problems [AI] can solve, and there are problems it cannot solve. So, I think being clear with that from the start, and how [the project] will be an iterative learning process also for [business] ... is very important. ... [T]hey can't just say, 'Build me an algorithm that does this' and then go away.

- *Engage with external experts to kickstart AI projects*—Another approach to preventing organizations from costly project failures is to consult more experienced or specialized AI experts. These experts do not necessarily have to be large and well-known consultancies but can be small and specialized firms or research institutions that combine AI expertise with deep industry knowledge. We talked to data scientists and managers of such organizations who are active, for example, in the energy retailing, publishing, or engineering industry. Their clients use them to outsource AI tasks that do not belong to their core business (e.g., text generation), to purchase industry-standard solutions (e.g., benchmarks, customer valuation models), enrich their data sets with external data, or to learn from the vendor's project management experience (e.g., in project scoping or evaluation of software frameworks).

Managing AI Like Traditional IT Projects

When it comes to the AI development phase, management's perception that AI models are specific computer programs that can be easily created and modified leads to another problem. Namely, the impression that AI projects can be treated like traditional IT projects, where the focus lies on developing and delivering systems on time, within budget, and within scope (following a linear and planning-oriented, waterfall-like process model at worst). For example, the CEO of a data science company that develops ML-models for the metalworking industry depicts this way of business and management thinking about AI models: "It [the AI model] has got it wrong here, can you program that out?' I say, 'No, you can't program that out.'" In this way, management prioritizes a "workmanship of certainty,"⁴⁴ which assumes that the decision rules of computer models are under full control of data scientists (i.e., the mechanical work perspective).

AI projects are typically much more iterative, exploratory, and open-ended, consisting of steps like framing the problem and asking the right questions, finding, and extracting appropriate data, and running experiments to derive new knowledge or support decisions.⁴⁵ This rather fits the craftwork perspective (i.e., *exploration*). The fuzzy front-end of AI projects, namely the definition of questions and hypotheses, seems to be the most critical phase, as an interviewee from an international bank told us:

I think the most important thing, sometimes the most challenging, is to exactly define the issue business is facing and the outcome they are expecting ... If you don't define the expected outcome very clearly, you can go completely in the wrong direction.

One of our interviewees explained that this way of working is often new to the business, requiring *dedication* by all involved stakeholders:

Business units know how classical IT projects are run: The IT department gets the requirements and then iteratively implements them. But data science projects require much more interaction between domain experts and data scientists, and it is important that the business knows and understands this.

Due to the highly iterative and collaborative nature of AI projects, many companies apply stripped-down versions of agile development practices (e.g., Scrum, Kanban boards) to manage the work of data scientists in AI projects. Such agile project management approaches allow flexibility in that they account for iterative and experimental developments, or project scope changes through new user requirements during project execution. This is meaningful in organizing software development, and some approaches of agile frameworks are helpful for data science projects, as well. Yet, such agile approaches create the illusion of *plannability*. They also assume that at the end of the project (or iteration), a stable release of a software product will be delivered to the client, for example, in the form of shippable software packages, which is rarely the case with data science projects in their early stages. In addition, AI projects face further uncertainties and scope changes due to the nature of the data, which is a constantly changing, editable, and interpretable material.⁴⁶ While new wishes from clients are obviously creating additional effort, data issues are invisible to managers and are uncovered during data analysis. AI systems, thus, are learning agents that are never really finished, and setting them up is rather an *exploration* endeavor. AI applications should be rather seen as a new novice employee or business function, requiring close supervision in the beginning, continuous performance monitoring, and regular retraining (see also Challenge 5 “Dynamic Environments”), which requires *dedication* by all participants to empower them. Our interviews revealed that selecting a too-rigid (agile) project management approach for running AI projects can overload data scientists with administrative tasks (e.g., maintaining task and feature lists, meetings, and shippable prototypes), reducing the time they can spend on actually developing and evaluating data pipelines or ML models. One data scientist working on text generation applications reported:

We will probably try to use Scrum when we grow further, but we are not yet quite sure, because all the people we know in the field of data science or AI who have worked with Scrum have ultimately fallen flat.

Wrong project management tools quickly lead to reduced employee productivity and satisfaction, and ultimately, reduced business value, as the project outcome might be less innovative than it could have been.

We identified three *tactics* for organizations to overcome the challenge of inflated management expectations and the underlying tension between the mechanical and craftwork perspectives on data science work.

- *Data scientists train managers on the job*—Many data scientists find the need to explain or introduce to managers and business collaborators how ML models work or how they should interpret the results. For example, many

of the more technical interviewees in our study mentioned the importance of being a good, data-driven storyteller and being able to provide convincing narratives for complex technical issues. Managers often appreciate a brief explanation of technical details in ML models (e.g., one explanatory slide in a presentation) as it helps them to learn more. Data scientists should keep offering these educational services.

- *Basic AI and ML courses for managers*—Some companies also started training their project managers in ML techniques. Not necessarily to be able to create ML models themselves, but to gain a better understanding about the differences to traditional software development. Similarly, more and more business schools start to teach business administration students ML basics to foster their AI knowledge. The contents of such training should include (among others) the working principle of ML, competencies to assess ML model quality, and typical pitfalls of AI projects in practice.
- *Appropriate key performance indicators (KPIs)*—Successful AI projects also do not end with the go-live of a system but when the system generates new insights or employees use it to make data-driven decisions. These things are hard to monitor and, hence, “it’s necessary to spend a lot of time on defining how ... to measure success,” as one interviewee emphasized. An example of good measurement comes from an insurance company we interviewed. They defined interlinked criteria for the success of its new ML-based churn prediction system (see Table 2) and measured predictive accuracy at the data scientist level, the development of actual churn rates at the business level, and the customer lifetime value of prevented churners as KPIs for the project success at the executive level.

Missing Data Input and Output Links to Existing Systems

Corporate AI applications heavily rely on data that either is necessary to construct models using ML approaches or AI generates data that is further used in business processes. Data availability and quality is a serious challenge during development to which researchers and practitioners alike have pointed.⁴⁷

TABLE 2. Example of a Set of Interrelated Metrics for Measuring the Success of a Churn Prediction System at an Insurance Company.

Stakeholder	Key Question	Metric
Data scientist	How accurately can I predict whether a customer will churn in the next 3 months?	Predictive accuracy of a machine-learning model for churn prediction.
Business unit	How can we prevent customers from leaving us?	Development of churn rates after deployment of the system.
Executive management	How can we maximize the lifetime value of our customers?	Predicted net profit from all future transactions with customers who could be prevented from churning.

“Data science is often described as data-driven, comprising unambiguous data and proceeding through regularized steps of analysis, ... [which] focuses more on abstract processes, pipelines, and workflows.”⁴⁸ Many firms are seeing their data lakes filled and they can also draw on many open data sources. This can create the impression of data as a *commodity* that just needs to be used in a prescribed way using bespoke tools (*specificity, abstract expertise*). Yet, our interviewees underlined that data scientists spend significant time identifying, getting access to, and preparing the data before they can start with the original AI engineering. They also have a hard time getting the results of their efforts integrated into existing IT systems so that operational processes can use them. Both are prime examples of why AI applications are more *exploration* than *plannable* projects and why data scientists need a fair amount of *dedication* but also diverse skills to get there, as we exemplify for the input and output side of AI applications.

In practice, creating the data input side of AI applications involves crafting data pipelines that associate more with plumbing than engineering as a data scientist involved in natural language processing explains:

For example, the pipeline that we had in mind when we spoke has changed as we moved forward. This is where all brainstorming comes after because we run into things that we need to reevaluate. So, in the beginning, you have a high-level understanding of how it might work, and you think it’s feasible, and then you need to read about it and actually improve.

Another problem here is that data stored in IT systems were often not recorded against the background of analyzing them to obtain predictions. Thus, increasing the quality of AI-based predictions might also require changes to the data-generating processes.

The challenge becomes also visible with the output side of AI systems, it almost seems that there is an insurmountable barrier between AI systems that should support or automate decision-making and those systems that are used to execute actions. One of our interviewees phrased it as “Our ERP system is my biggest headache in the whole world. It takes six months to get data out of it ... and, yeah, we can’t put data back into it.” The problem of information exchange between newly developed AI systems and systems in place seems to be simply a matter of missing interfaces (*specificity*). But developing those is usually more complex than initially thought or sometimes not even feasible at all. Often, not only the IT systems must change, but whole business processes need to be re-engineered to be able to handle data-driven insights, for example, including (semi) automatic decision-making. One data scientist from the pharmaceutical industry complained:

the first step in embedding it in the processes is getting it embedded in the systems that run the processes, and that’s what we can’t do ... we have all these nice models for when they get a new planning system, and we can deploy our own things within it, one day.

Large international firms are often running their business processes with enterprise systems that are ten or 20 years old and have been developed on a completely different technology stack. Figuring out necessary changes to such IT systems and realizing them is a crafty rather than a mechanical task, given the broad and abstract required skill set, experience, and *all-roundness*.

Data scientists reported on several *tactics* to handle the challenge of missing links to existing IT systems and data scarcity:

- *Bridges between AI and IT*—A common solution is to build a bridge between new AI systems and legacy IT systems is the implementation of middleware infrastructure offering interfaces and adapters to various third-party systems, thereby enabling systems to talk to each other. These systems are, however, often expensive. A less-immersive, but also less-robust, solution is robotic process automation (RPA). RPA systems watch the user executing a business process in the graphical user interface of, for example, an enterprise resource planning (ERP) system and are then able to automate the process by repeating the tasks performed by the user directly in the graphical user interface. The RPA system could, for example, copy and paste the outputs of an AI system into an ERP system, while simultaneously checking simple rules and customizing the process execution depending on predictions made by the AI system.
- *Process-integrated prescriptions instead of predictions*—To overcome the user acceptance problem, organizations need to find new ways of presenting the outputs of AI systems to end users. For example, predictions or probabilities—that is, the raw output of ML models—need to be transformed into actionable prescriptions and presented to employees at the right place and the right point in time. In the context of churn prediction, for example, an electricity provider fed the churn probability estimates into their call center management system. Based on the predicted churn scores, the system suggests different questions to the call center agent to be asked during the phone call. So, instead of showing call center agents static questions, they gave them personalized and actionable instructions, which are more digestible but also difficult to work around.
- *Learning apprentice approach*—Organizations are often faced with limited data for ML applications.⁴⁹ They often need to wait a considerable amount of time, sometimes years, before they have collected sufficiently large and representative amount of data. In other cases, the tasks to be automated are not completely supported by a single IT system, or the performed steps are not logged in the required level of detail. In such cases, it is difficult to collect the required training data. A strategy to obtain training data in such situations is the “learning apprentice” approach.⁵⁰ Here, the AI system acts as an apprentice watching the human experts performing their tasks and recording all relevant input data and outcomes. After observing several thousands of repetitions of the same task, ideally performed by different individuals, the

system is then able to learn the function required to correctly transform the input data into the outcomes; often better than any human expert did before, because ML can learn from many more examples in a short time compared to humans.

- *Increase data awareness*—Another tactic to overcome this challenge is to start creating awareness of the importance of data quality and the willingness to share and reuse data across departments. This results from the fact that—besides technical problems with extracting, transforming, and loading the data from the source systems—many informants had problems interpreting the correct meaning of existing data. For example, many sales departments create sales orders in their ERP systems to make ad-hoc stock reservations for customers. Once the customer has ordered the material, or does not need it anymore, they cancel these sales orders again. Including such “shadow” orders in training sets for recommender systems or price optimization models can easily lead to biases in ML models. The problem in such a case is not that the required training data do not exist, have missing or wrong values, but that the data do not represent what they seem to. Data scientists create data awareness by constantly pointing to the relevance of original data points, given that employees often do “not imagine that someone else would use the data afterward,” as one interviewee explained.

The Question of Why

The latest generation of ML algorithms demonstrates remarkable predictive power. Deep neural networks have already achieved (super-)human performance in tasks as diverse as playing games, performing image-based medical diagnoses, or translating texts from one language to another.⁵¹ However, in many situations, high predictive accuracy alone is not sufficient to convince decision makers to accept and implement the recommendations of an algorithm. Especially in highly uncertain environments and in situations where the stakes are high, decision makers often want to understand how or why an AI system made a specific prediction. This requirement of ML methods is often referred to as interpretability or explainability,⁵² but practitioners also describe it as transparency or traceability. For example, they want to see drivers of customer behavior (“Why do[n’t] people buy our product?”), the causes of a machine failure, or receive an explanation for why an ML model that showed high predictive power in a laboratory test underperforms in field use. Especially decision makers trained in an engineering or scientific environment demand that the outputs of ML systems can be comprehended by humans (*abstract expertise, planning*). After all, one core reason to implement AI is to move toward more rational and evidence-based decision-making processes. If the output of AI applications is opaque, however, the goal of more rational decision-making can hardly be realized. The need for explanations is also necessary in cases of failures. Managers who are accountable for the decisions made by an AI system are unlikely to follow its recommendations if they cannot explain why something went wrong (*individuality, planning*). In addition, legal requirements in several countries require that

operators of AI applications can produce easy-to-understand explanations for why a model made a specific prediction (*abstract expertise, planning*). For example, the European Union General Data Protection Regulation (GDPR) contains a “right for explanation,” obliging data processors who make algorithmic decisions about individuals (e.g., online credit application, e-recruiting practices) to provide them on demand with “meaningful information about the logic involved” (Art. 13 and 14 GDPR).⁵³ In the United States, similar rights exist in the context of credit scoring.

A data scientist from an insurance company explained how management expects that the reasoning of ML models is codifiable: “we cannot provide a black box prediction model to our insurance brokers ... brokers need to know which data point has which influence on the likelihood of churn.” This opinion was echoed by many of our interviewees across industries; when it comes to supporting knowledge workers in their decision-making processes through AI, one of the interviewees noted that “traceability [interpretability] is the most important criterion” for user acceptance of algorithmic decision-making.

The best-performing predictive models are, however, often black boxes, which easily contain millions of internal parameters that jointly define the function for translating inputs into outputs. It is nearly impossible, even for experts, to comprehend and interpret how these models make predictions (*mastery*). In other words, many AI systems have superhuman predictive capabilities but are unable to explain the why and how behind their predictions (*embodied expertise*). Some voices advocate putting more effort into model development rather than trying to make black-box models explainable afterward.⁵⁴ To do this, for example, conceptual models, expert knowledge, and human cognition could be used to develop meaningful input variables for ML models that make good predictions but also allow reasoning to be explained, as examples from medicine⁵⁵ and energy⁵⁶ demonstrate. The fast life of data science work, however, often does not allow such deep digging into the data (see the example cited in the challenge “Inflated management expectations”). Data scientists, thus, need to carefully balance between the requirements and constraints of their work (accuracy versus interpretability versus available time). In addition, they require their *all-roundness* to tune models for good predictions, need to know approaches to make models interpretable, engineer variables, and do this in a short time.

In addition to the accuracy versus interpretability tradeoff, there is another crucial issue that must be considered when applying ML to model real-world phenomena, namely, the difference between correlation and causality. There is no guarantee that the variables, which an ML algorithm detects as good predictors of a phenomenon (i.e., they are correlated with the target variable), are actually the causes of this phenomenon. If prediction is the primary goal of an AI system, and the system was trained on carefully selected training data, it makes accurate and consistent predictions. In such a case, the lack of an underlying causal model might not represent a problem. If, however, an organization or decision maker wants to place real-world interventions, that is, to alter the environment with algorithmically prescribed actions (e.g., lower the price for an existing product), a lack of causal relationships in a model can be a major

problem. Consider the following example from one of the cases we analyzed in our study. A digital marketing agency used deep neural networks to predict the success of Instagram posts that involve product placement by Internet celebrities. They found that product placement in front of mountain scenery is associated with higher numbers of likes and comments by fans. Should they recommend that all their clients shoot pictures with mountains in the background? Probably not. It just happened that some of their most famous clients specialized in marketing sports and fashion products for the winter season. So, the real cause for the success of their posts was their huge base of fans and followers, and not the background of the images they posted. This example nicely demonstrates that, when it comes to making decisions informed by AI systems and translating these decisions into real-life actions, it is indispensable to know whether an identified pattern is a real cause-and-effect relationship or just a spurious correlation, as intervening on variables which are spuriously correlated with an intended outcome will have no effect. This requires *embodied expertise* of data scientists as they bring technical as well as business knowledge in the AI development process.

We identified three technical *tactics* to overcome the accuracy versus interpretability versus available time tradeoff.⁵⁷

- *Post-hoc explainability methods*—One solution is to apply methods that try to explain the logic of black box predictive models (for an example see Figure 2).

FIGURE 2. Example of the outputs of a post-hoc explainability model providing an explanation for why a convolutional neural network has classified two animals in a picture as a dog and a cat (the pixels that the model relied on are highlighted).



Note: M. T. Ribeiro, S. Singh, and C. Guestrin, "Why Should I Trust You? Explaining the Predictions of Any Classifier," (Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 2016), doi:10.1145/2939672.2939778.

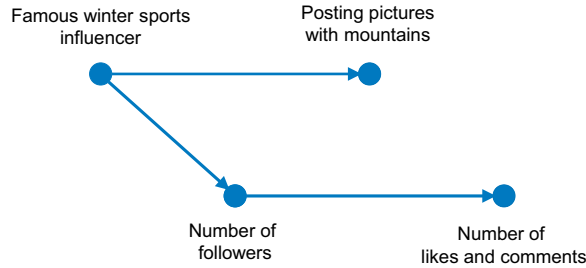
By building an interpretable model on top of a complex black box model, this strategy tries to combine the predictive accuracy of modern ML algorithms with the interpretability of classical statistical models. However, developing so-called post-hoc explainability methods that possess high fidelity and end-user friendliness is still an ongoing research effort.⁵⁸

- *Push intrinsically transparent models to the max*—As noted above, a complex deep learning model typically outperforms simple additive and linear models when fed with the same raw data. This is due to the ability of deep neural networks to automatically discover and represent relevant features and model non-linearities and interactions. However, experiments⁵⁹ have demonstrated that in many situations, especially those involving structured numerical input data, careful modeling, especially feature engineering, informed by high domain expertise can result in inherently interpretable models that possess the same predictive accuracy as black box ML algorithms. Of course, tuning these models to the maximum extent is associated with extra time and costs for understanding the problem domain and developing a solution, but it avoids relying on black boxes or unfaithful post-hoc explanations in high-stakes decision problems.
- *Causal modeling and rigorous experimental evaluation*—The gold standard in science for making causal inferences is to conduct randomized controlled experiments. In the form of A/B tests, the experimental method is also gaining more and more popularity in industry, especially at Internet companies. Uber and Amazon are good examples of how firms can combine ML techniques with classical behavioral science methods to overcome the causality challenge.⁶⁰ Whenever possible, they conduct randomized controlled trials to investigate whether hypotheses generated through data science methods hold in real life. In many situations, obtaining experimental data through randomized and controlled experiments is simply not possible. For example, when introducing a product or service innovation, it is often not possible to compare its effect to an appropriate control group. In such situations, one has no other option than to work with existing observational data. One way to investigate causal questions with observational data is to model the causal process that has generated the data as thoroughly as possible and use appropriate statistical methods and tools to control for confounding factors.⁶¹ This not only supports modeling, but also a careful formulation of the investigated business problem (see Challenge 2 “Managing AI Projects like Traditional IT Projects”). Discovering the data-generating causal graph and investigating the nature and strength of its relationships requires close collaboration between data scientists and domain experts, who have substantive knowledge of the business problem in question. Figure 3 shows a possible causal data-generating graph for the influence marketing example.

Dynamic Environments

ML algorithms learn a function that is able to correctly map inputs into outputs; for example, a function to output the correct English translation for a given French input text, or a function that can predict the number of likes or

FIGURE 3. Example of a possible data-generating causal graph showing the causal relationships between variables in the influencer marketing example.



Note: Being a famous winter sports influencer is causally related to posting pictures with mountains and having many followers. Furthermore, a larger number of followers is causally related to the number of likes and comments a posted picture receives. If one does not control for the number of followers and being a famous winter sports influencer, it seems as if a mountain in a picture causes the number of likes and comments to rise.

comments (output) of an Instagram fashion post (input). Once the function has been learned with sufficient accuracy, it can be used to automate tasks that have traditionally been performed by human experts such as interpreters or marketing managers. This constitutes a manufacturing mindset, where data scientists produce, implement, and deliver *specific* software parts in *plannable* cycles. Following the mindset of a software engineering and production process, the developer would not be involved in the operations of the IT system once it is completed. This leads to a *detachment* of the creator from the product. This approach assumes, however, that the function to predict outputs based on inputs is stable over time.⁶² In the example of translating French to English, this assumption is largely met. In the social media marketing example, however, the factors that determine a popular post (i.e., in terms of user reactions) change at least as fast as fashion and popular culture trends change. In other words, what used to be hip on Instagram six months ago, may not trigger half as many “likes” today. Sensing such structural changes and assessing their consequences for predictions and decisions made by ML-based systems is a challenge that is often overlooked by companies, especially those that are new to ML. Data scientists need to have high *dedication* and continuous involvement to *explore* the functioning of ML-based software artifacts, as one informant from an international jewelry manufacturer and retailer explained:

You don’t put something into production and then just keep it running. It’s very much about continuous monitoring and figuring out that there’s a drift in the data ... Traditional software keeps functioning in the same way over time. But the performance of ML models might degrade when the data changes.

The statement highlights two points: First, the importance of continuously monitoring the performance of predictive models with a sense for peripheral changes that might affect the models’ outcomes in the future—in the language of a traditional craftsman this means that they are conscious about their material (i.e., data).⁶³ This is often more difficult than it sounds, as it requires access to up-to-date ground-truth data as a benchmark for the algorithmic predictions.

Second, it sets the focus on an important make-or-buy decision. Simply buying a standard software package or hiring external consultants to build a predictive model (*commodity*) is often not sufficient.

Successfully employing an ML-based AI system is not a one-time project, but a continuous effort that requires specific in-house expertise, *dedication*, and hard-to-specify knowledge (*embodied expertise*). Even a company like Google learned this the hard way, when their famous Google Flu Trend service, which was able to predict the spread of the flu based on what users type into Google's search box, started to wildly overpredict flu levels.⁶⁴ One of the main reasons was that Google's engineers never updated the underlying ML models, although both the nature of influenza spread and the logic and usage of its search engine changed considerably over the years. The Google Flu Trends example—after an initial hype, the service has been turned off after it started overpredicting the flu—illustrates how dynamic environments challenge the creation of effective AI applications. Inaccurate predictions not only lead to suboptimal decisions, but sometimes can even damage the reputation of whole companies (external intangible value). Crafting machines that learn entails continuous work and adjustments, where data scientists remain engaged with the ML model even after production and they maintain, improve, renovate—take care of the AI applications—to sustain creating business value through AI.

Inspired by the practices of interviewees and current research on ML, we find two *tactics* to address this challenge:

- *Monitoring and detection of data drift*—Since data are subject to constant change (be it that data-generating processes change, e.g., due to the introduction of a new IT system, or be it that human behavior changes, for example, human habits change due to COVID-19), ML models degrade over time and lose predictive power. ML research is very active in proposing approaches to detect and monitor input data drift. These methods enable to identify occasions when productive models need to be updated to avoid a drop in performance.
- *Combining AI development and operations*—Beyond the mere monitoring of ML models goes the approach of MLOps.⁶⁵ Inspired by practices from software development to reduce the time between committing a change to a system and the change being successfully placed into normal production (DevOps), MLOps closely dovetails the processes of AI development and AI operation to enable rapid updating of ML models in productive AI systems. It aims at improving the quality of ML models in production by, on one hand, increasing automation in monitoring and improving the performance of these models, and, on the other hand, implementing organizational structures for ensuring compliance with business and regulatory requirements. Concrete practices include, for example, the use of drift detection techniques, conducting A/B tests to detect degradation or side-effects in the real-world impact of models, or interdisciplinary teams spanning technical and business experts to ensure quality and minimize the risk of AI systems.

Concluding Remarks

Organizations continue to struggle with AI value creation. Our analysis of 55 ML-based AI projects revealed that five challenges (summarized in Table 1) are the result of a tension created by two contradictory, yet complementary, perspectives to data science work: craft and mechanical work. The reality of developing and maintaining AI applications takes place in a paradoxical situation where both perspectives, mechanical work and craft, are necessary. The crafting approach better captures the process by which *effective* AI applications come to be. This approach appears to be embraced by skilled data scientists who demonstrate mastery of the AI tools, all-roundedness in addressing business problems through AI, and embodied expertise when they engage with the data and AI models. They exhibit dedication and a strong sense of exploration and communality in producing the best possible results. The mechanical approach to data science work, long popular in the management and software engineering literature, better captures the process by which organizations *efficiently* develop and implement AI applications. This approach to data science work is primarily embraced by management that sees data science skills as easily obtainable, specifiable, and codifiable while their attitude toward AI projects is characterized by expectations of commensurable results, transactional interactions among project members, and treating AI projects as plannable entities.

A large portion of the challenges that organizations encounter when trying to extract value from AI are caused by ignoring the paradoxical tension inherent in data science work. Tensions, if they are paradoxical, are always active. Satisfying only one of the two demands exacerbates the need for the other⁶⁶ and can start what Smith and Lewis call a vicious cycle.⁶⁷ Yet, we see evidence that some experienced organizations manage to industrialize AI development without sacrificing innovation. While these organizations engage in standardization (e.g., setting up a standard data pipeline, developing ML modeling best practices) and knowledge specialization (e.g., specialized data science roles) at the operational level, they—at the same time—try to familiarize managers with the craftwork nature of data science work by broadening the knowledge and skills of mid- and top-level project managers and executives. In these organizations, managers across functions and hierarchies know the basics about how ML algorithms work, what data are required to train them, and how to assess their performance. This is not to enable them to develop AI systems themselves but to be able to understand which business processes can be supported by ML and to communicate and collaborate with data science teams. Vice versa, senior data scientists at these organizations are not only responsible for leading technical teams but are actively engaging in innovating business models and processes.

Managing paradoxical tensions means enabling a virtuous cycle that accentuates acceptance rather than defensiveness.⁶⁸ Built on acceptance and resolution, organizational tensions can be considered both as a call for and a source of novel creative solutions⁶⁹ and should correspondingly be nurtured and managed. The tactics that we have reported here can be used as a starting point for developing more complete strategies in managing the paradoxical nature of data science work and creating business value from AI applications.

Appendix

Method

To generate rich data about the challenges companies are facing when trying to implement AI applications, we interviewed data scientists—professionals with an interdisciplinary background and holistic view on AI and its application in organizations,⁷⁰ but also C-level representatives of small to medium-sized firms with data-driven business models. Data scientists are involved from start to finish in AI projects and, in many cases, end up leading and driving AI value realization.⁷¹ As a point of departure, our data collection process included two rounds. In the first round (between October 2018 and January 2020), we conducted 40 semistructured interviews. Our interview partners were located mainly in Germany and Denmark, but also in other European countries, the United States, and Singapore. The interview guide focused on concrete AI projects (some interviews covered more than one project). We structured our interview guide along a unified model of the process of extracting insights from data and translating these insights into decisions and actions.⁷² All interviews were recorded and transcribed. The data from this first phase formed the primary foundation for our data analysis and development of findings. A second round of data collection (in January and February 2020) focused on the validation of our findings. For this, we organized a workshop with data scientists with 13 participants (seven of them were interview partners, while six were new participants). In this focus group, we presented our findings and asked for feedback.

Our qualitative data analysis followed the approach of Gioia et al.⁷³ It started with applying open coding on the interview transcripts to identify first-order concepts, which stay close to the informants' original statements. We linked these original concepts together and formulated second-order themes. Finally, we further grouped the themes together and came up with the five challenges that we presented here. After this first round of analysis, we realized that the identified challenges were similar to challenges mentioned in previous studies of organizational implementation of AI. This made us wonder why these challenges persist. What causes them? Our search for an explanation of the challenges led us to appreciate different perspectives on data science work. In our data, we noticed that the way data scientists see their work is different from the way management sees it. These differences made us take a work perspective. Inspired also by a current discourse in organizational science about different approaches to work—such as craft and mechanical work⁷⁴—we performed another round of analysis in coding to find reasons for these challenges in the way data science work was perceived. This second round of analysis led us to understand that each challenge was the result of a paradoxical tension created by two different perspectives on data science work: craft and mechanical work.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

Author Biographies

Konstantin Hopf is a Postdoctoral Researcher and Lecturer at the Chair of Information Systems and Energy Efficient Systems at the University of Bamberg (email: konstantin.hopf@uni-bamberg.de).

Oliver Müller is a Full Professor at Paderborn University, where he holds the Chair of Management Information Systems and Data Analytics (email: oliver.mueller@uni-paderborn.de).

Arisa Shollo is an Associate Professor at the Department of Digitalization at the Copenhagen Business School (email: ash.digi@cbs.dk).

Tiemo Thiess is a Senior Data Scientist at Denmark's largest commercial pension company PFA and holds a Ph.D. in Business IT from the IT University of Copenhagen (email: t.thiess@outlook.com).

Notes

1. Thirty-five years ago, Mintzberg urged the management audience to think about two different images of strategies: planning and crafting. In a similar fashion, we ask the audience to imagine planning and crafting AI applications. H. Mintzberg, "Crafting Strategy," *Harvard Business Review*, 65/4 (July/August 1987): 66-75, <https://hbr.org/1987/07/crafting-strategy>.
2. Although the term AI was coined many decades ago, its exact definition is still a matter of debate (see, for example, S. Legg and M. Hutter, "A Collection of Definition of Intelligence," in *Advances in Artificial General Intelligence: Concepts, Architectures and Algorithms*, ed. B. Goerzl and P. Wang (Washington, DC: IOS Press, 2007), p. 17. In this article, we use the term AI to refer to ML-based applications that learn how to perform tasks in a data-driven way instead of being explicitly programmed. Typical examples of such ML-based applications include recommendation agents, predictive maintenance, or machine translation.
3. T. Fountaine, B. McCarthy, and T. Saleh, "Building the AI-Powered Organization," *Harvard Business Review*, 97/4 (July/August 2019): 62-73; J. K. U. Brock and F. von Wangenheim, "Demystifying AI: What Digital Transformation Leaders Can Teach You About Realistic Artificial Intelligence," *California Management Review*, 61/4 (Summer 2019): 110-134.
4. E. E. Makarius, M. Debmalya, J. D. Fox, and A. K. Fox, "Rising with the Machines: A Sociotechnical Framework for Bringing Artificial Intelligence into the Organization," *Journal of Business Research*, 120 (November 2020): 262-273, doi:10.1016/j.jbusres.2020.07.045.
5. N. Balasubramanian, Y. Ye, and M. Xu, "Substituting Human Decision-Making with Machine Learning: Implications for Organizational Learning," *Academy of Management Review*, 47/3 (July 2022): 448-465, doi:10.5465/amr.2019.0470; Y. R. Shrestha, S. M. Ben-Menahem, and G. von Krogh, "Organizational Decision-Making Structures in the Age of Artificial Intelligence," *California Management Review*, 61/4 (Summer 2019): 66-83, doi:10.1177/0008125619862257.
6. Fountaine et al. (2019), op. cit.; T. H. Davenport and D. D. D'Augelli, "Artificial Intelligence for the Real World," *Harvard Business Review*, 96/1 (January/February 2018): 108-111, <https://hbr.org/2018/01/artificial-intelligence-for-the-real-world>.
7. Fountaine et al. (2019), op. cit.
8. Worldwide spendings on AI systems are skyrocketing and expected to reach nearly \$154B in 2023 (According to New IDC Spending Guide, <https://www.idc.com/getdoc.jsp?containerId=prUS50454123>).

9. M. Tarafdar, C. M. Beath, and J. W. Ross, "Using AI to Enhance Business Operations," *MIT Sloan Management Review*, 60/4 (Summer 2019): 37-44; T. H. Davenport and D. J. Patil, "Data Scientist: The Sexiest Job of the 21st Century," *Harvard Business Review*, October 1, 2012, <https://hbr.org/2012/10/data-scientist-the-sexiest-job-of-the-21st-century/>; M. P. Joshi, N. Su, R. D. Austin, and A. K. Sundaram, "Why So Many Data Science Projects Fail to Deliver," *MIT Sloan Management Review*, 62/3 (Spring 2021): 85-89, <https://sloanreview.mit.edu/article/why-so-many-data-science-projects-fail-to-deliver/>.
10. Joshi et al. (2021), op. cit.
11. Joshi et al. (2021), op. cit.; the Kaggle's 2020 *State of Data Science and Machine Learning* survey demonstrates the important role of Data Scientists in setting up AI applications (<https://www.kaggle.com/kaggle-survey-2020>); A. Shollo, K. Hopf, T. Thiess, and O. Müller, "Shifting ML Value Creation Mechanisms: A Process Model of ML Value Creation," *The Journal of Strategic Information Systems*, 31/3 (September 2022): 101734, doi:10.1016/j.jsis.2022.101734.
12. Ibid.
13. Joshi et al. (2021), op. cit.; L. Baier, F. Jöhren, and S. Seebacher, "Challenges in the Deployment and Operation of Machine Learning in Practice," (ECIS 2019 Proceedings, 2019), https://aisel.aisnet.org/ecis2019_rp/163.
14. J. Kroezen, D. Ravasi, I. Sasaki, M. Żebrowska, and R. Suddaby, "Configurations of Craft: Alternative Models for Organizing Work," *Academy of Management Annals*, 15/2 (July 2021): 502-536, doi:10.5465/annals.2019.0145.
15. S. L. Nielsen, K. Klyver, M. R. Ewald, and T. Bager, *Entrepreneurship in Theory and Practice: Paradoxes in Play*, 3rd ed. (Cheltenham, UK: Edward Elgar Publishing, 2021).
16. W. K. Smith and P. Tracey, "Institutional Complexity and Paradox Theory: Complementarities of Competing Demands," *Strategic Organization*, 14/4 (November 2016): 455-466.
17. W. K. Smith, M. Erez, S. Jarvenpaa, M. W. Lewis, and P. Tracey, "Adding Complexity to Theories of Paradox, Tensions, and Dualities of Innovation and Change: Introduction to Organization Studies Special Issue on Paradox, Tensions, and Dualities of Innovation and Change," *Organization Studies*, 38/3-4 (April 2017): 303-317, doi:10.1177/0170840617693560.
18. J. Schad, M. W. Lewis, S. Raisch, and W. K. Smith, "Paradox Research in Management Science: Looking Back to Move Forward," *Academy of Management Annals*, 10/1 (2016): 5-64, doi:10.5465/19416520.2016.1162422; S. Raisch and S. Krakowski, "Artificial Intelligence and Management: The Automation-Augmentation Paradox," *Academy of Management Review*, 46/1 (January 2020): 192-210.
19. R. H. Coase, "The Nature of the Firm," *Economica*, 4/16 (1937): 386-405; F. W. Taylor, "Scientific Management," *The Sociological Review*, 7/3 (1914): 266-269; M. Weber, *Economy and Society* (Berkeley, CA: University of California Press, 1978); O. E. Williamson, *The Economic Institutions of Capitalism* (New York, NY: Macmillan, 1985); Kroezen et al. (2021), op. cit.
20. Kroezen et al. (2021), op. cit.
21. Kroezen et al. (2021), op. cit., p. 11.
22. Kroezen et al. (2021), op. cit., p. 12.
23. C. Shearer, "The CRISP-DM Model: The New Blueprint for Data Mining," *Journal of Data Warehousing*, 5/4 (2000): 13-22; U. Fayyad, G. Piatetsky-Shapiro, and P. Smyth, "The KDD Process for Extracting Useful Knowledge from Volumes of Data," *Communications of the ACM*, 39/11 (1996): 27-34, doi:10.1145/240455.240464.
24. S. Akter, R. Bandara, U. Hani, S. F. Wamba, C. Foroapon, and T. Papadopoulos, "Analytics-Based Decision-Making for Service Systems: A Qualitative Study and Agenda for Future Research," *International Journal of Information Management*, 48 (2019): 85-95, doi:10.1016/j.ijinfomgt.2019.01.020.
25. Davenport and Patil (2012), op. cit.
26. A. Ebadi, Y. Gauthier, S. Tremblay, and P. Paul, "How Can Automated Machine Learning Help Business Data Science Teams?" (2019 18th IEEE International Conference on Machine Learning and Applications [ICMLA], 2019), pp. 1186-1191, doi:10.1109/ICMLA.2019.00196.
27. Davenport and Patil (2012), op. cit.
28. J. S. Saltz and N. W. Grady, "The Ambiguity of Data Science Team Roles and the Need for a Data Science Workflow Framework," in *2017 IEEE International Conference on Big Data (Big Data)*, (2017), pp. 2355-2361, doi:10.1109/BigData.2017.8258190; Kaggle (2020), op. cit.
29. "Who Is a Data Scientist in 2021? A Research on 1,001 Data Scientists," *365 Data Science*, October 6, 2021, <https://365datascience.com/career-advice/career-guides/data-scientist-2021/>.

30. Kroezen et al. (2021), op. cit., p. 2.
31. Kroezen et al. (2021), op. cit., p. 11.
32. Kroezen et al. (2021), op. cit., p. 18.
33. Davenport and Patil (2012), op. cit.
34. W. M. P. van der Aalst, "Data Scientist: The Engineer of the Future," in *Enterprise Interoperability VI*, Proceedings of the I-ESA Conferences, ed. K. Mertins, F. Bénaben, R. Poler, and J. P. Bourrières, Vol. 7 (Cham, Switzerland: Springer, 2014), pp. 13-26, doi:10.1007/978-3-319-04948-9_2.
35. As cited by M. Loukides, "What Is Data Science?" *O'Reilly*, June 2, 2010, <https://www.oreilly.com/radar/what-is-data-science/>.
36. D. Donoho, "50 Years of Data Science," *Journal of Computational and Graphical Statistics*, 26/4 (2017): 745-766, doi:10.1080/10618600.2017.1384734.
37. M. Müller, I. Lange, D. Wang, D. Piorkowski, J. Tsay, Q. V. Liao, C. Dugan, and T. Erickson, "How Data Science Workers Work with Data: Discovery, Capture, Curation, Design, Creation," in *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, May 2019, pp. 1-15, doi:10.1145/3290605.3300356.
38. Davenport and Patil (2012), op. cit.
39. R. H. Thaler, "Mental Accounting Matters," *Journal of Behavioral Decision Making*, 12/3 (1999): 183-206.
40. C. Heath, "Escalation and De-Escalation of Commitment in Response to Sunk Costs: The Role of Budgeting in Mental Accounting," *Organizational Behavior and Human Decision Processes*, 62/1 (April 1995): 38-54, doi:10.1006/obhd.1995.1029.
41. E. Brynjolfsson and T. Mitchell, "What Can Machine Learning Do? Workforce Implications," *Science*, 358/6370 (2017): 1530-1534, doi:10.1126/science.aap8062.
42. Shollo et al. (2022), op. cit.
43. Davenport and Ronanki (2018), op. cit.
44. Pye (1995), as cited in Kroezen et al. (2021), op. cit., p. 9.
45. D. A. Marchand and J. Peppard, "Why IT Fumbles Analytics," *Harvard Business Review*, 91/1-2 (January/February 2013): 104-112, <https://hbr.org/2013/01/why-it-fumbles-analytics>.
46. E. Parmiggiani, T. Østerlie, and P. G. Almklov, "In the Backrooms of Data Science," *Journal of the Association for Information Systems*, 23/1 (2022): 139-164.
47. A. L'Heureux, K. Grolinger, H. F. Elyamany, and M. A. M. Capretz, "Machine Learning with Big Data: Challenges and Approaches," *IEEE Access*, 5 (2017): 7776-7797, doi:10.1109/ACCESS.2017.2696365; Baier et al., op. cit.
48. Müller et al. (2019), op. cit.
49. K. Hopf, A. Weigert, and T. Staake, "Value Creation from Analytics with Limited Data: A Case Study on the Retailing of Durable Consumer Goods," *Journal of Decision Systems*, 32/2 (2022): 289-325, doi:10.1080/12460125.2022.2059172.
50. T. M. Mitchell, S. Mabadevan, and L. I. Steinberg, "LEAP: A Learning Apprentice for VLSI Design," in *Machine Learning*, ed. Y. Kodratoff and R. S. Michalski (San Mateo, CA: Morgan Kaufmann Publishers, Elsevier, 1990), pp. 271-289, doi:10.1016/B978-0-08-051055-2.50016-X.
51. "Measuring Trends in Artificial Intelligence," *The AI Index Report*, <https://aiindex.stanford.edu/report/>.
52. O. Biran and C. Cotton, "Explanation and Justification in Machine Learning: A Survey," (2017), http://www.cs.columbia.edu/~orb/papers/xai_survey_paper_2017.pdf; Z. C. Lipton, "The Mythos of Model Interpretability," *Queue*, 16/3 (May/June 2018): 31-57, doi:10.1145/3236386.3241340.
53. The EU General Data Protection Regulation (GDPR), <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:32016R0679>.
54. C. Rudin, "Stop Explaining Black Box Machine Learning Models for High Stakes Decisions and Use Interpretable Models Instead," *Nature Machine Intelligence*, 1/5 (May 2019): 206-215, doi:10.1038/s42256-019-0048-x.
55. R. Caruana, Y. Lou, J. Gehrke, P. Koch, M. Sturm, and N. Elhadad, "Intelligible Models for HealthCare: Predicting Pneumonia Risk and Hospital 30-Day Readmission," in *Proceedings of the 21st ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (August 2015), pp. 1721-1730, doi:10.1145/2783258.2788613.
56. K. Hopf, "Predictive Analytics for Energy Efficiency and Energy Retailing," *University of Bamberg Press*, 36 (2019): 43, doi:10.20378/irbo-54833.

57. M. Du, N. Liu, and X. Hu, "Techniques for Interpretable Machine Learning," *Communications of the ACM*, 63/1 (2019): 68-77, doi:10.1145/3359786; M. T. Ribeiro, S. Singh, and C. Guestrin, "Why Should I Trust You? Explaining the Predictions of Any Classifier," in *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, <https://github.com/marcotcr/lime>.
58. J. Wastensteiner, T. M. Weiss, F. Haag, and K. Hopf, "Explainable AI for Tailored Electricity Consumption Feedback—An Experimental Evaluation of Visualizations," in *ECIS 2021 Research Papers*, 2021, https://aisel.aisnet.org/ecis2021_rp/55.
59. Rudin (2019), op. cit.
60. "Using Causal Inference to Improve the Uber User Experience," *Uber Blog*, <https://eng.uber.com/causal-inference-at-uber/>; "Causal Inference," *Amazon Science*, <https://www.amazon.science/tag/causal-inference>.
61. J. Pearl and D. Mackenzie, *The Book of Why: The New Science of Cause and Effect* (New York, NY: Basic Books, 2018).
62. J. Gama, I. Žliobaitė, A. Bifet, M. Pechenizkiy, and A. Bouchachia, "A Survey on Concept Drift Adaptation," *ACM Computing Surveys*, 46/4 (April 2014): Article 44, pp. 1-37, doi:10.1145/2523813.
63. Kroezen et al. (2021), op. cit.
64. D. Lazer, R. Kennedy, G. King, and A. Vespignani, "The Parable of Google Flu: Traps in Big Data Analysis," *Science*, 343/6176 (2014): 1203-1205, doi:10.1126/science.1248506.
65. D. Kreuzberger, N. Kühn, and S. Hirschl, "Machine Learning Operations (MLOps): Overview, Definition, and Architecture," *arXiv:2205.02302*, May 14, 2022, doi:10.48550/arXiv.2205.02302.
66. M. W. Lewis, "Exploring Paradox: Toward a More Comprehensive Guide," *Academy of Management Review*, 25/4 (2000): 760-776; C. Sundaramurthy and M. Lewis, "Control and Collaboration: Paradoxes of Governance," *Academy of Management Review*, 28/3 (July 2003): 397-415.
67. W. K. Smith and M. W. Lewis, "Toward a Theory of Paradox: A Dynamic Equilibrium Model of Organizing," *Academy of Management Review*, 36/2 (April 2011): 381-403.
68. *ibid.*
69. N. Beech, H. Burns, L. de Caestecker, R. MacIntosh, and D. MacLean, "Paradox as Invitation to Act in Problematic Change Situations," *Human Relations*, 57/10 (October 2004): 1313-1332; M. Gaim, "On the Emergence and Management of Paradoxical Tensions: The Case of Architectural Firms," *European Management Journal*, 36/4 (August 2018): 497-518; E. Miron-Spektor, F. Gino, and L. Argote, "Paradoxical Frames and Creative Sparks: Enhancing Individual Creativity through Conflict and Integration," *Organizational Behavior and Human Decision Processes*, 116/2 (November 2011): 229-240.
70. Davenport and Patil (2012), op. cit.; E. Plastino and M. Purdy, "Game Changing Value from Artificial Intelligence: Eight Strategies," *Strategy & Leadership*, 46/1 (2018): 16-22, doi:10.1108/SL-11-2017-0106; van der Aalst (2014), op. cit.
71. Shollo et al. (2022), op. cit.
72. A. Abbasi, S. Sarker, and R. Chiang, "Big Data Research in Information Systems: Toward an Inclusive Research Agenda," *Journal of the Association for Information Systems*, 17/2 (2016): 1-32, <https://aisel.aisnet.org/jais/vol17/iss2/3/>; R. Sharma, S. Mithas, and A. Kankanhalli, "Transforming Decision-Making Processes: A Research Agenda for Understanding the Impact of Business Analytics on Organisations," *European Journal of Information Systems*, 23/4 (2014): 433-441, doi:10.1057/ejis.2014.17.
73. D. A. Gioia, K. G. Corley, and A. L. Hamilton, "Seeking Qualitative Rigor in Inductive Research: Notes on the Gioia Methodology," *Organizational Research Methods*, 16/1 (January 2013): 15-31, doi:10.1177/1094428112452151.
74. Kroezen et al. (2021), op. cit.