Overcoming Challenges for Successful Artificial Intelligence Projects in the Manufacturing Industry

Dissertation

zur Erlangung des Grades eines Doktors der Wirtschaftswissenschaft

der Rechts- und Wirtschaftswissenschaftlichen Fakultät

der Universität Bayreuth

Vorgelegt

von

Daniel Leuthe

aus

Augsburg

Dekan:	Prof. Dr. Claas Christian Germelmann
Erstberichterstatter:	Prof. Dr. Björn Häckel
Zweitberichterstatter:	Prof. Dr. Maximilian Röglinger
Datum der mündlichen Prüfung:	11.10.2024

Copyright Statement

The following sections are partly comprised of content taken from the research articles included in this doctoral thesis. To improve the readability of the text, I omit the standard labeling of these citations.

Abstract

Driven by, on the one hand, recent advances in the field of artificial intelligence (AI), such as the accessibility of large-scale data sources as well as the improvement of AI algorithms, and, on the other hand, growing competitive pressure and changing customer needs, many incumbents in the manufacturing sector have embarked on a journey to realize the business value of AI. In this vein, manufacturers enhance their operational efficiency and optimize existing processes, for example, by using anomaly detection to enable condition monitoring. Even further, AI paves the way to offer disruptive value propositions that push them beyond the limits of existing products or services, such as AI-based prediction and control of machine usage, to offer customized business models. To implement AI and unlock its business value, manufacturers have started to conduct AI projects. However, AI projects pose several challenges that arise on an overarching *organizational level* as well as on a *project level*. Those challenges are conceptualized using the three-dimensional technology-organization-environment framework and extended by a fourth data-centric dimension. First, those challenges include *business challenges*, such as developing an organization-wide AI strategy or identifying the organizational capabilities required for those new AI-based business models. Second, *technical challenges* emerge, such as ensuring the explainability of AI models or managing technology-induced security issues at the organizational level. Third, these challenges encompass *data challenges*, such as providing enough high-quality labeled data or reducing the data dimensionality. Lastly, *sustainability challenges* must be considered, such as increasing fairness during AI projects or promoting responsible AI use. Motivated by these challenges, this cumulative dissertation provides solutions to conduct successful AI projects in the manufacturing sector. It comprises six research articles that deliver research artifacts to assist in overcoming the outlined challenges.

Regarding the *challenges AI projects reveal at an organizational level*, research article #1 provides 24 organization-wide success factors for AI projects, structured along four success dimensions and specified by 93 subordinated success manifestations, laying the foundations on how to plan and execute AI projects successfully. Thereafter, research article #2 addresses what capabilities manufacturers need to implement suitable AI-based and data-driven business models. The result is a maturity model that assists in identifying the organizational capabilities required. Since such AI and data-driven business models increase the risk of information technology security incidents, research article #3 addresses the issue of an organizational-wide incident response management. A maturity model is developed that provides manufacturers with a comprehensive perspective on capabilities for developing effective incident response management since IT security incidents can never be prevented entirely.

Focusing on the *challenges at the project level*, research article #4 develops a data-efficient active learning architecture for anomaly detection in industrial time series data and its instantiation for a real-world robotic screwdriving application. In particular, this helps to overcome the time-consuming anomaly data labeling challenge. Since such advanced AI approaches reduce the explainability of AI models, research article #5 compares and evaluates three frequently used transparent AI models and four different state-of-the-art Explainable AI (XAI) methods by conducting an online survey. The results encourage using XAI methods as the right choice of the methods enables an increase in the measured human-centered explainability by 10%. Lastly, research article #6 presents the sustainable machine learning design pattern matrix to overcome the sustainability challenges. The artifact serves as a diagnostic tool to capture the sustainability status quo and develop a vision regarding the sustainability of AI projects.

In sum, this doctoral thesis strives to empower manufacturers to overcome the business, technical, data, and sustainability challenges of AI projects by presenting and evaluating applicable artifacts that contribute to the existing knowledge of AI in the manufacturing sector.

Table of Contents

I Introduction

1 Motivation

Driven by the development of powerful hardware, the availability of extensive data sources, and the development of new algorithms, artificial intelligence (AI) evokes the interest of researchers and practitioners alike to create value in various domains such as education, finance, and manufacturing (Enholm et al., 2022; M. Kim et al., 2022). In this context, AI is typically associated with the ability of machines to accomplish cognitive functions that reference human intelligence, such as perceiving, reasoning, learning, interacting with the environment, problem-solving, decision-making, and even creativity (Benbya et al., 2021; Berente et al., 2021; Russell & Norvig, 2016). Thus, AI does not describe a specific technology but rather stands for a plethora of algorithmic approaches, methods, and techniques that span across a wide range of application domains (Stone et al., 2022). The prevailing ones include machine learning (ML) (Ågerfalk, 2020; Merhi, 2023). ML is considered the core of present-day AI (Berente et al., 2021; Jordan & Mitchell, 2015). In general, ML comprises capabilities to learn from training data iteratively and to improve their results, solving tasks automatically without explicitly being programmed (Collins et al., 2021). ML is broadly categorized into three types: supervised learning, unsupervised learning, and reinforcement learning (Haenlein & Kaplan, 2019; Jo, 2021). Supervised ML refers to methods associating inputs with target outputs based on labeled data to make classifications or predictions. Unsupervised ML represents methods that infer underlying patterns in unlabeled data without predefined outcomes, while reinforcement learning includes methods that optimize actions toward a predefined goal based on feedback using both penalties and rewards (Bertolini et al., 2021; Goodfellow et al., 2016). In this doctoral thesis, I subsume ML and the three types under the umbrella term AI to keep the wording consistent.

Opening up new avenues for decision-support and problem-solving (Rai et al., 2019; Vial et al., 2023), organizations can utilize AI for a variety of purposes, ranging from analyzing and optimizing internal processes, designing novel product features or devising smart services to invent disruptive AI-based business models (Burström et al., 2021; Sjödin et al., 2021; Stahl et al., 2023). These formidable prospects fuel the ambition of organizations to unlock the underlying business value of AI (Shollo et al., 2022; Vial et al., 2023). The enormous potential of AI in business is evident from the considerable attention it has received in recent years. Global spending on AI reached $154 billion in 2023 across all industries (International Data

Corporation, 2023; Statista, 2024) and is predicted to grow to $298 billion by 2027, especially due to the current generative AI software spending (Gartner, 2024; Maslej et al., 2024). This means an almost doubling of the global AI spending compared to expenditure in 2023, which is in line with the projected compound annual growth rate for AI spendings of 19.1% between 2023 and 2027 (Maslej et al., 2024; McKinsey, 2023).

In this vein, the manufacturing sector – as a leading industry – has embarked on the journey to leverage AI (Bertolini et al., 2021; S. W. Kim et al., 2022a; Stahl et al., 2023). So far, incumbent manufacturers have focused on high-quality physical machinery and equipment as the differentiating feature for decades, but the industry's market conditions have been changing recently (Favoretto et al., 2022). Although Germany still has a vital role in the global supply chain, underpinned by a high export ratio (e.g., Germany accounted for 18% of global machinery exports in 2023), international competition is growing steadily and increasing market pressure on incumbent firms (VDMA, 2024). For example, in 2021, the global machinery production volume amounted to nearly $71.5 billion. Germany accounted for approximately $8.9 billion, placing it second after China with around $21.8 billion, followed by Japan ($8.8 billion) in third place. In 2022, the global production volume amounted to $80.3 billion, with Germany ($9.7 billion) in third place after China with $25.7 billion and Japan with almost $9.9 billion (VDMA, 2023). Therefore, the manufacturing sector is characterized by high competition and thus shrinking profit margins, especially at the core of machinery sales (Björkdahl, 2020). On top of that, factors – such as supply shortages caused by international conflicts, rising energy costs, higher wage settlements, and inflationary gains in raw material prices – are increasing cost pressure and reducing profit margins (Priyono et al., 2020). Additionally, the evolving requirements of both established and emerging customer segments necessitate digital solutions that automate, enhance, or streamline machine operations (Abrell et al., 2016; Loebbecke & Picot, 2015). Further, digital solutions, besides the pure physical machinery, facilitate the entry of new competitors from outside an industry, enabling them to enter new markets and exert pressure on established companies (Ritter et al., 2023; Stahl et al., 2023).

To counteract this, manufacturers are investing more and more in AI and the data collection and processing required. According to the "Global Machinery and Equipment Report" of Bain & Company (2024), 75% of manufacturing executives indicate that adopting and implementing AI is one of their top priorities. This is underlined by a study by the MIT Technology Review (2024) among 300 manufacturers, of which 64% are currently researching or experimenting with AI. In the manufacturing sector, AI is enabled through the broader concepts of cyber-physical systems (CPS) and the Industrial Internet of Things (IIoT) to generate the necessary data basis for AI. In CPS, physical components (e.g., machines and plants) and software components (e.g., in-line monitoring systems) are designed to interact with each other and the surrounding environment by exchanging data (Cui et al., 2020; Häckel et al., 2019). This enables seamless coordination between digital and real-world components. To enable the communication of and within CPS and the needed data exchange, manufacturers must connect their production assets, such as robots and machines, via the IIoT and process the resulting data (Baltuttis et al., 2022). The resulting data value can then be captured via data analytics and AI. On the one hand, this enables them to enhance operational efficiency and optimize existing processes (exploitation) (Holotiuk et al., 2024; Margherita & Braccini, 2023). For example, anomaly detection based on AI is becoming increasingly relevant in the manufacturing sector. Early detection of those anomalies in manufacturing applications, such as condition monitoring, fault diagnosis, or predictive maintenance, avoids economic and environmental losses due to, e.g., maintenance cost reduction, machine fault reduction, increased spare part life, or increased overall production (Bertolini et al., 2021; X. Li et al., 2022; Xiong et al., 2024). An example is Siemens Digital Industries at its production site in Amberg. The highly automated plant manufactures printed circuit boards, and AI is successfully used in two areas. First, they use in-line process data to predict whether or not additional quality testing is necessary, eliminating unnecessary control operations. Second, machine data is leveraged to predict possible machine faults of the printed circuit board assembly lines to smooth production flow (Schmitt et al., 2020; van Giffen & Ludwig, 2023). On the other hand, AI empowers manufacturers to offer their customers novel or adapted value propositions that push them past the limits of an existing product or service core (exploration) (Holotiuk et al., 2024; Knote et al., 2020; Stahl et al., 2023). For example, digital connected machines enable AI-based insight into usage data and allow manufacturers to position themselves with new business models and move toward servitization of business (Favoretto et al., 2022; Gebauer et al., 2021). In this way, the advantages of AI can not only be exploited in the existing core business, such as the manufacturer's production, but can also be marketed to customers in the form of digital and data-driven business models and services. This opens up opportunities for servitization in business-to-business customer relationships, as continuous connections between manufacturing companies and their customers provide manufacturers with contextual and constantly up-to-date data regarding the conditions and uses of intelligent machines and devices (Ardolino et al., 2018; Bertolini et al., 2021). The resulting offerings from manufacturers comprise integrated and intelligent bundles of products and services, often referred to as smart service systems (Beverungen et al., 2021; Heinz et al., 2022). This enables a continuous customer relationship instead of a single transactional relationship in the form of a one-time machine sale. One promising example from the manufacturing sector is WashTec, an incumbent manufacturer and world leader in car wash systems. These systems are operated in more than 80 countries worldwide. They continuously generate structured and unstructured data such as machine status, usage data, time-series sensor data, error messages, or even image data through their IIoT connectivity solution (Ritter et al., 2023). This data, stored centrally in a cloud data warehouse, can then be leveraged to create innovative customer-centric value propositions such as a new digital customer and service portal based on a continuous subscription. This allows the customer journey to be mapped digitally and additional AI-based services to be offered and flexibly monetized. Examples of this include car wash solutions tailored to individual vehicle models based on AI object recognition and a pay-per-use billing for the car wash operators, remote maintenance services depending on the service level agreement, or automated chemical supply services through AI-based prediction of consumption behavior (Häckel et al., 2022; Ritter et al., 2023; WashTec, 2024).

In search of a strategic response to AI´s tremendous potential, as described in the previous paragraph, manufacturers have started to initiate AI projects and thus implement AI to unlock the underlying business value (Shollo et al., 2022; Vial et al., 2023). To execute AI projects, organizations must follow an AI project workflow that entails a series of AI project phases (Benbya et al., 2021; Vial et al., 2023). Generally, to accomplish a project's implementation, companies pass through the four main phases of planning, developing, deploying, and maintaining (Cooper & Zmud, 1990). While the overall discussion of projects, e.g., information technology (IT) projects, is mature, the emergence of AI projects challenges established knowledge owing to the differences between AI and IT projects, such as iterative learning, dependence on data, or unclear possibilities due to the current AI hype (Berente et al., 2021; Merhi, 2023). To address this, the literature elaborates on AI project workflows to guide the execution of AI projects. A comparison of the relevant different AI project workflows and their individual AI project phases is given in Table 1.

Table 1. Overview of AI project workflows

		AI Project Workflows						
		CRISP-DM (Wirth & Hipp, 2000)	KDD (Fayyad et al., 1996)	CRISP-ML (Studer et al., 2021)	AIDOP (Allen et al., 2017)	TDSP (Tabladillo, 2022)	Microsoft ML workflow (Amershi et al., 2019)	ML operation principles (Kreuzberger et al., 2023)
AI Project Phases	Demand Specification	Business understanding	Understanding of the application domain	Business understanding	Plan	Business understanding	Model requirements	MLOps product initiation, Requirements for feature pipeline
	Data Collection and Preparation	Data understanding, Data preparation	Data selection, preprocessing, and transformation	Data understanding, Data preparation	Data preparation, Build ground truth	Data acquisition and understanding, Feature engineering	Data collection, Data labeling, Data cleaning, Feature engineering	Feature engineering pipeline
	Modeling and Training	Modeling, Evaluation	Algorithm selection, Data mining	Modeling, Evaluation	Train model, Evaluate model	Model training, Model evaluation	Model training, Model evaluation	Experimentatio n
	Deployment and Monitoring	Deployment	Interpretation, Evaluation	Deployment, Monitoring and maintenance	Deploy, Monitor, Capture feedback	Deployment	Model deployment, Model monitoring	Automated ML workflow pipeline

Within this dissertation, I reframe the four general project phases (Cooper & Zmud, 1990) by including a more data-centric perspective, accounting for AI specifics, and considering the AI project workflows from Table 1. This both accounts for a comprehensive overview of the whole AI project workflow in four phases, enables an end-to-end consideration of the entire AI project workflow, and forms the basic methodological structure for the research articles in this cumulative dissertation. First, during the planning phase, besides understanding the organizational problems and opportunities, there needs to be a stronger focus on identifying AI model requirements (Amershi et al., 2019; Kreuzberger et al., 2023). Second, there is a need for a data-centric phase prior to model development, as the data basis must be explicitly considered in AI projects besides the pure IT application development (Allen et al., 2017; Papagiannidis et al., 2023). Thus, the data challenges of data understanding, preparation, transformation, feature engineering, labeling, and cleaning must be taken into account (Tabladillo, 2022; Wirth & Hipp, 2000). Third, during the development phase, iterative training, experimentation, and evaluation loops should be conducted to benchmark different AI models (Fayyad et al., 1996; Tabladillo, 2022). This comprises the modeling, training, and evaluation of the ML models based on the previous phase’s data (Amershi et al., 2019). Depending on the AI model evaluation results, adjustments such as hyperparameter optimization can be made (Kreuzberger et al., 2023). Fourth, the deployment and monitoring phases should be considered together in AI projects, as changes such as data and concept drift must be continually monitored and could lead to redeployment (Kreuzberger et al., 2023). Consequently, this consists of deploying the AI model, transitioning the AI model into a software product, and monitoring its predictions and decisions in a real-world environment (Studer et al., 2021). This leads to the four overarching AI project phases (Table 1). These four phases are not strictly sequential: iterations and feedback loops between the phases are both possible and necessary (Singla et al., 2018). The four phases provide a framework-agnostic analysis of the AI project workflow by aggregating different AI project workflow frameworks to a higher abstraction level.

Nevertheless, despite the role of AI to serve as a driver for the exploitation of existing processes or the exploration of innovation and thus as a means for novel value promises in the manufacturing sector (S. W. Kim et al., 2022a; Vial et al., 2023), organizations encounter significant challenges when planning and executing AI projects (M. C. M. Lee et al., 2023; Merhi, 2023). From a theoretical lens, those challenges regarding the introduction of AI in companies are grounded in the technology-organization-environment (TOE) framework (Tornatzky & Fleischer, 1990) as established in previous work (e.g., Alsheibani et al., 2018; Chatterjee et al., 2021; Jöhnk et al., 2021; Pumplun et al., 2019) and widely applied in IS research (Baker, 2012; Wallace et al., 2020). The TOE framework represents a multi-perception theory and is used to investigate the adoption of technologies and innovations from a socio-environmental and technical context at the firm level (Chatterjee et al., 2021). In general, the TOE framework comprises three interrelated dimensions: (a) the technological dimension describes the internal and external technologies available, (b) the organizational dimension focuses on the business characteristics that might influence the adoption process, such as the managerial structure, cultural competencies, and existing resources, and (c) the environmental dimension considers the structure of the industry including the firm's competitors, suppliers, customers, and regulatory environment (Baker, 2012; Tornatzky & Fleischer, 1990). Following Zhu and Kraemer (2005), Baker (2012), and Wallace et al. (2020), the TOE framework can be adopted and extended to the specific technology, in this case, AI. Thus, on the one hand, since the basis of AI is data, the technology dimension is extended by a data-centric dimension. Data-centric AI is a novel paradigm emphasizing that the systematic design and engineering of data are essential for building effective and efficient AI-based systems (Jakubik et al., 2024). On the other hand, to describe technology adoption with a focus on sustainability, recent works have moved to extend the conventional view on environmental factors with a focus on sustainability (Dadhich & Hiran, 2022; Kumar & Krishnamoorthy, 2020). Thus, the TOE framework and its adaptations lead to four specific challenges AI implementation and AI projects oppose, whereby these challenges arise during the four phases of the AI project workflow and occur both on an overarching *organization level* (i.e., where several AI projects are considered together) and on an individual AI *project level* (i.e., where each AI project is considered individually) (Jöhnk et al., 2021; Merhi, 2023; Vial et al., 2023; Weber et al., 2023). These challenges encompass *business challenges* (e.g., developing an organization-wide AI strategy or identifying value-adding AI use cases for individual projects), *technical challenges* (e.g., ensuring explainability of AI algorithms or managing technology-induced security issues at the organizational level), *data challenges* (e.g., providing enough labeled data to learn from or preventing data breaches), and *sustainability challenges* (e.g., promoting responsible AI use throughout the entire organization or increasing fairness during each AI project) (Dennehy et al., 2023; Enholm et al., 2022; Merhi, 2023; Papagiannidis et al., 2023; Weber et al., 2023).

First, the *business challenges* refer to managing the transformation toward a business logic of AI-enabled value creation, delivery, and capture (Ibarra et al., 2018; Lins et al., 2021). Here, an organizational-wide AI strategy development and fit ensures the emergence of an AI strategy aligned with the company’s strategic landscape and vision. An initial task is exploring a target state or archetype that fits existing company processes and capabilities as well as existing and future customers based on established organizational assets like the products or market position (Stahl et al., 2023; Sund et al., 2021). This bundles all corporate efforts towards enabling, executing, and streamlining AI projects. Hence, organizations need to develop a tailored and aligned AI strategy that considers the available core and cultural competencies, as a one-size-fits-all solution to an AI strategy does not work (Keding, 2021; Kolbjørnsrud et al., 2017; Shollo et al., 2022). In AI projects, organizations often focus on a wide array of technical questions without understanding the business problem (Weber et al., 2023). Hence, incumbents must focus on an apparent business problem understanding to comprehend, identify, and scope potential business problems (e.g., customer pain points) to develop productive AI systems (Kinkel et al., 2022; Kreuzberger et al., 2023).

Second, manufacturers also face *technology challenges* when conducting AI projects. Amongst others, AI projects often get stuck in an experimental technical pilot phase without transitioning to productive systems (Benbya et al., 2021; Merhi, 2023). To successfully implement and fully grasp the real impact of AI models, there is a need to consider all AI project phases (Verdecchia et al., 2023; C.-J. Wu et al., 2022). Hence, an end-to-end view of the AI project workflow is crucial for aligning the project to the problem’s requirements, ensuring suitable feature engineering, selecting appropriate AI models, successfully deploying them, and maintaining their performance over time due to concept and data drift (Kreuzberger et al., 2023).

Furthermore, the AI model's technical explainability must be considered during development to ensure understandable and tractable AI models as they often come with the expense of lacking explainability despite their good performance, referred to as the black-box problem (Barredo Arrieta et al., 2020; Vilone & Longo, 2021). Hence, on the one hand, AI models lack transparency about how they process data to derive results, as non-deterministic outcomes occur (G. Miller, 2021). On the other hand, AI models can be manipulated, leading to different results. Thus, model explainability must be guaranteed to some extent to increase transparency, which creates the basis for trust and, therefore, broader AI adoption. This can be done by, e.g., considering opaque models or post-hoc explainability methods (Berente et al., 2021; Jordan & Mitchell, 2015; Vilone & Longo, 2021). Lastly, the ever-rising number of CPS and the resulting step increase in data collection on edge systems and the cloud leads to a steadily growing attack surface for cyberattacks (Häckel et al., 2019). Organizations should implement data security strategies (Bitzer et al., 2023; Böttcher et al., 2022; Leuthe et al., 2024), as, for instance, an infiltration of the systems and machinery is possible, leading to production downtimes or data disclosure which can have harsh financial and reputational consequences (Eitle et al., 2022; Papagiannidis et al., 2023). As a result, IT security has been identified as one of the biggest technology challenges and is a top priority for many IT executives (Kappelman et al., 2020).

Third, *data challenges* arise as data is the key element to enable sophisticated AI projects (Jöhnk et al., 2021). A key challenge is organizational-wide data quality to ensure the accuracy, completeness, consistency, and integrity of data and the associated metadata management (Benbya et al., 2021; Weber et al., 2023). This provides the AI model in individual AI projects with high-quality training, evaluation, and production data, as biased and discriminatory data negatively impact AI systems (M. C. M. Lee et al., 2023; Reis et al., 2020; Vial et al., 2023). Subsequently, companies must ensure high-quality labeled data for AI projects. This, for example, becomes particularly evident in the important task of AI-based anomaly detection in CPS. Supervised learning models often struggle with class imbalances and require high numbers of labeled instances (anomalies) to perform well and correctly detect anomalies (Alaei & Noorbehbahani, 2017; Bertolini et al., 2021; Yuan & Wu, 2021). In manufacturing, these labels typically require domain-specific knowledge, are time-consuming, and are expensive to obtain; thus, they are often unattractive in practice (Xiong et al., 2024; Yu et al., 2021). Unsupervised learning models, by contrast, do not require labeled data and are most conventionally used (Chevrot et al., 2022; Pang et al., 2022). However, they are ineffective in handling high-dimensional data, are susceptible to high false-positive rates, and lack the capability for multi-class anomaly detection and evaluation as well as tracking (Aggarwal, 2017). Semi-supervised learning approaches that combine the former assume that the underlying data do not contain anomalies, making them hardly applicable in practice (Himeur et al., 2021; G. Kim et al., 2023). Thus, new approaches are needed to overcome the data labeling bottleneck and efficiently use data and expert feedback while aiming for improved anomaly detection performance.

Fourth, the *sustainability challenges* of AI usage have increased significantly in recent years (Maslej et al., 2024; Schoormann et al., 2023). Within this dissertation, the concept of environmental, social, and governance (ESG) to operationalize sustainability is used. ESG is chosen as it enables a multidimensional perspective on sustainability and is frequently used in the corporate environment, for example, in the manufacturing sector due to regulatory requirements such as the EU Emission Trading System or the EU’s corporate sustainability reporting directive (Drempetic et al., 2020; Sætra, 2023). Overall, AI’s negative impacts on resource consumption, societal injustice, or even human rights cannot be neglected anymore (Cowls et al., 2023; Dennehy et al., 2023; Koniakou, 2023). For instance, AI holds the unintended risk of reflecting the implicit social bias at the expense of equality, e.g., between genders or ethnic groups (Gupta et al., 2022; van Noorden & Perkel, 2023). Furthermore, the amount of computing power needed to train current AI models has doubled every 3.4 months since 2012 (Amodei & Hernandez, 2018; Debus et al., 2023). The sustainability challenges of AI have therefore become more apparent (Mikalef et al., 2022), leading to calls to work toward the sustainability of AI (SAI) (Schoormann et al., 2023; Schwartz et al., 2020; Tornede et al., 2022). Overall, SAI describes the sustainable design, development, and use of AI through its entire lifecycle and, therefore, across all phases of the AI project workflow (van Wynsberghe, 2021). The environmental impact of the AI development and AI model must be considered in AI projects to reduce energy consumption, as the manufacturing sector, in particular, generates a lot of data due to continuous operation and the large number of machines, the processing of which in AI models is very computationally intensive (Patterson et al., 2022; Veit & Thatcher, 2023; Verdecchia et al., 2023). Additionally, a further focus must be placed on social and ethical aspects of AI as well as increasing fairness during the AI project phases to foster responsible AI (Dennehy et al., 2023; Ferrara, 2023; Mikalef et al., 2022). Lastly, the sustainability challenges AI holds from a governance perspective, such as the EU AI Act, must be considered across the entire organization (Koniakou, 2023; Papagiannidis et al., 2023).

All in all, AI enables manufacturers a strategic response to changing customer requirements and competitive pressure by leveraging the value potential of data through exploitative and explorative AI projects. Nevertheless, due to the previously described four challenges, AI projects often fail to meet the intended outcomes or are even terminated before completion (Shollo et al., 2022; Vial et al., 2023). Specifically, AI projects tend to get stuck in an experimental pilot phase without transitioning from conceptual use cases to productive systems (Benbya et al., 2021; Merhi, 2023). This finding is consistent with studies indicating that about 85% of AI projects have little to no impact (Shollo et al., 2022; Vial et al., 2023). This circumstance turns AI projects into a risky matter, as their failure entails sunk costs and may even jeopardize competitiveness (van Giffen & Ludwig, 2023).

2 Structure of the Thesis and Related Work

Overall, this dissertation aims to address the challenges AI implementation opposes in the manufacturing sector as described in the previous section – i.e., *business*, *technology*, *data*, and *sustainability challenges* – as its main research objective. As different organization-related and project-related capabilities are needed on an overarching *organizational* (i.e., where several AI projects are considered in conjunction) as well as a *project-specific* level (i.e., where each AI project is considered individually) to overcome these four challenges, each research article is assigned to one of the two levels. Furthermore, as those challenges arise during AI projects, the four-phased AI project workflow – i.e., “Demand Specification”, “Data Collection and Preparation”, “Modeling and Training”, and “Deployment and Monitoring” - is used to structure the research articles and the corresponding results. Hence, each research article focuses on one or more of the four challenges along the AI project workflow, either at an *organizational level* or a *project level*. However, a clear allocation of the four challenges to the four AI project phases is not entirely possible due to the phase`s dependencies and their iterative nature and is therefore not pursued in this dissertation. Even if, for example, business challenges primarily occur in the first phase, “Demand Speficiation”, data challenges often arise in the second phase, “Data Collection and Preparation”, and technical challenges often appear in the third and fourth phases, “Modeling and Training” and “Deployment and Monitoring”, sustainability challenges may occur in all phases. Further, business challenges likewise emerge in phases two to four, and data challenges can also occur in the other phases, such as in the “Deployment and Monitoring” phase.

First, concerning the *organizational level*, previous research analyzes AI implementation mainly against the backdrop of generic concepts such as AI adoption (e.g., Alsheibani et al., 2019; Eitle & Buxmann, 2020; Rowland et al., 2022) or AI readiness in particular (e.g., Jöhnk et al., 2021; Pumplun et al., 2019; Uren & Edwards, 2023). Moreover, the literature describes the need to acquire or develop specific capabilities to accomplish AI implementation at an overarching organizational level (e.g., Mikalef et al., 2022; Sjödin et al., 2021; Weber et al., 2023). Especially about the business challenges, previous research conceptualized AI business models and AI use cases as well as generic transformation paths towards AI implementation in the manufacturing sector (Bertolini et al., 2021; Björkdahl, 2020; S. W. Kim et al., 2022b). Although these studies altogether list relevant aspects to conduct AI implementation successfully and thus AI projects, they either do not provide a coherent and conclusive picture of capabilities and success factors required or are not explicitly geared towards the manufacturing sector. In terms of the technology challenges, especially those addressing IT security, these received little attention despite the high strategic relevance of digitalization and AI projects. Nevertheless, the attack surface of entire organizations continuously rises as, for example, more and more data is transferred between machines, the machines and the company, or even between companies (Lallie et al., 2021; Vial, 2019). While the literature suggests solutions, above all technical ones, for proactively enhancing IT security, security incidents cannot always be prevented (Schlette et al., 2021; Van Der Kleij et al., 2022). This increases the need for reactive IT security measures, especially, such as establishing an organizational-wide incident response management, whereby the literature falls short here (Thangavelu et al., 2021). Nevertheless, although existing frameworks for incident response management contain guidelines, practices, and requirements, they barely specify what capabilities organizations need to achieve these standards (Ahmad et al., 2022; De Haes et al., 2013). All in all, there is still a need for research to derive and structure the capabilities and success factors needed, especially regarding business and technology challenges on an organizational level, for successful AI projects in the manufacturing industry.

Second, regarding the *project level*, recent publications propose recommendations and strategies to help manufacturers overcome the challenges that arise in AI projects, drawing on insights from real-world examples (e.g., van Giffen & Ludwig, 2023; Vial et al., 2023; Xiong et al., 2024). Nevertheless, executives seeking to understand better and manage their AI projects can only draw on a limited number of empirically evaluated studies, as AI projects differ from other IS and IT development projects (Vial et al., 2023). Especially regarding the sustainability challenges, AI projects pose completely new hurdles (van Wynsberghe, 2021). Hence, there is an increasing demand from practitioners as well as calls for research to provide comprehensive design approaches and implementable best practices to conduct sustainable AI projects (Dennehy et al., 2023; Pappas et al., 2023; Shneiderman, 2021). Nevertheless, existing literature on this is either scarce or separated across different research areas (Dennehy et al., 2023; Patterson et al., 2022; Verdecchia et al., 2023), such as papers focusing on solutions to reduce the energy consumption of specific AI models (e.g., Patterson et al., 2022; Verdecchia et al., 2023), the social and ethical aspects of AI by increasing fairness during AI projects (e.g., Ferrara, 2023; Mikalef et al., 2022), or the governance perspective of AI development and AI projects (e.g., Koniakou, 2023; Papagiannidis et al., 2023). Concerning the data challenges in AI projects, existing research is mainly engaged with optimizing data input features (Jakubik et al., 2024), whereby the topic of "data-centric AI" and, above all, the provision of labeled data in AI projects, has rarely been addressed to date despite its great relevance in practice, e.g., for in-line quality inspection (Xiong et al., 2024; Zeiser et al., 2023). In manufacturing applications, labeled data is usually provided by domain experts such as machine production employees. However, their time for this is limited and, therefore, often too few data labels are available (Finder et al., 2022; Zeiser et al., 2023). Hence, work needs to be done to efficiently use this data and expert feedback to minimize burdens in practice while aiming for improved performance of the AI models. In this vein of improved AI model performance, better AI models often come with the expense of lacking explainability, referred to as the black-box problem, which leads decision-makers to distrust or even reject them. Explainable AI (XAI) can be leveraged to overcome this technical challenge in AI projects as it helps to comprehend how an AI model decides, predicts, and performs its operations (Barredo Arrieta et al., 2020; Burkart & Huber, 2021; Rai, 2020). Nevertheless, around 70% of papers neglect evaluating XAI methods with users (Brasse et al., 2023) or only emulate the user evaluation (Ali et al., 2023; Ribeiro et al., 2016), leading to inaccurate human-centered insights (Brasse et al., 2023; B. Kim et al., 2020). Hence, research is still needed to conduct human-centered evaluations collecting end-user feedback in the context of specific AI projects and use cases (Ali et al., 2023; Ding et al., 2022a).

As a cumulative dissertation, this work consists of six research articles placed in the four-phase AI project workflow and either contribute on an organizational or project-specific level to address the raised business, technology, data, and sustainability challenges (Figure 1). To achieve their aim, this dissertation builds on the previously described existing research to derive applicable artifacts that deliver prescriptive knowledge for research and practice and an impetus for future research. The research articles use the *Design Science Research (DSR)* paradigm (Hevner et al., 2004) and the *Cross Industry Standard Process for Data Mining (CRISP-DM)* (Wirth & Hipp, 2000) as their methodological approach. As such, the resulting solution-oriented artifacts and their design are built on the extant research knowledge base to tackle real-world problems and are evaluated in real-world world settings, e.g., based on AI projects, AI expert insights, or manufacturing data sets (Sonnenberg & vom Brocke, 2012b; vom Brocke et al., 2020).

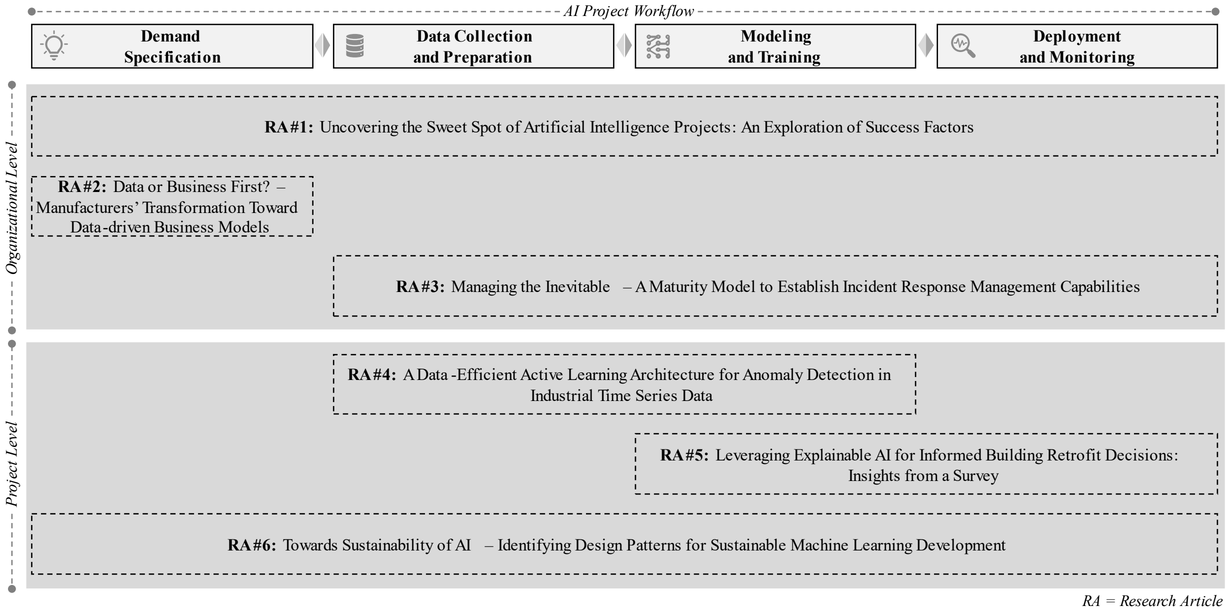

Figure 1. Assignment and structure of the research articles to the topic of this dissertation

In this course, section II.1 – including research articles #1, #2, and #3 – provides insights on assessing and overcoming the challenges at an organizational level. Specifically, research article #1 provides a systematic overview of 24 success factors to be considered across the entire organizational level when conducting AI projects. The success factors are structured along four overarching success dimensions and specified by 93 subordinated success manifestations. Additionally, the success factors are mapped to the four AI project phases. The results were derived by synthesizing extant knowledge from a systematic literature review with emergent insights from an expert interview study. Next, as manufacturers need to determine their suitable target state at the beginning of an AI project, i.e., potential AI-based or data-driven archetypes, and the organizational capabilities required that either already exist or need to be developed, research article #2 provides a maturity model based on five distinct archetypes. Manufacturers can identify the appropriate archetypes and categorize themselves to subsequently identify the capabilities necessary to implement those specific data-driven and AI-based business model archetypes. To ensure an integrated perspective on the entire organizational level, the maturity model levers all organizational layers of the enterprise architecture model by Urbach et al. (2021). As such projects continuously increase the attack surface for IT security incidents, such as data breaches, and security incidents cannot always be prevented, research article #3 focuses on reactive IT security measures by addressing the issue of organizational-wide incident response management (IRM). Such an IRM must be considered early in the AI project phases, i.e., as soon as data is stored and processed. The result of research article #3 is a maturity model for necessary IRM capabilities based on four focus areas and 29 capability dimensions. It is evaluated with seven real-world manufacturing organizations to prove its fidelty with the real-world phenomena and ease of use.

Overcoming the challenges at a project level, section II.2 presents three papers, including research articles #4, #5, and #6. Research article #4 proposes an application-oriented and data-efficient Active Learning Architecture for Anomaly Detection in Manufacturing CPS, called ALMAN, in the course of the data challenges. This work aims to solve the data labeling bottleneck in practice, which is relevant in the “Data Collection and Preparation” and “Modelling and Training” phases. The key motivation for this paper is to efficiently use data and expert feedback to minimize burdens in practice while aiming for improved anomaly detection performance. The developed architecture is demonstrated and validated in a case study of a CPS robotic screwdriving application. Since such efficient but complex architectures reduce the explainability of AI models, research article #5 counteracts this by implementing three transparent AI models and applying four XAI methods to an artificial neural network using a real-world dataset. The explainability is evaluated through a survey with 137 participants considering the human-centered dimensions of explanation satisfaction and perceived fidelity to quantify the added value of XAI methods. To complete the dissertation across all four AI project phases and to overcome the sustainability challenges of AI projects, research article #6 presents the sustainable ML design pattern matrix. This artifact provides 35 design patterns structured along the four AI project phases and subdivides them into the introduced three sustainability dimensions of environmental, social, and governance. The results are grounded in justificatory knowledge from research, refined with naturalistic insights from expert interviews, and validated in three real-world AI case studies using a web-based instantiation.

Section III concludes this dissertation with a summary of key findings, limitations, and provides an outlook on future research. Section IV contains the publication bibliography, and Section V offers additional information on all research articles (V.1), my contributions to each research article (V.2), and the research articles themselves (V.3 - 8).

II Research Overview for Successful AI Projects

1 Challenges and Solutions at an Organizational Level

AI provides manufacturers the basis for exploitative use cases to enhance operational efficiency and optimize existing processes such as condition monitoring and anomaly detection in the production facilities (Bertolini et al., 2021; van Giffen & Ludwig, 2023) as well as explorative use cases to offer novel or adapted value propositions that push them beyond the limits of an existing product or service core such as remote data-driven service solutions on a pay-per-use basis (Favoretto et al., 2022; Sjödin et al., 2021; Stahl et al., 2023), decision-makers must balance the potential and the challenges that appear in the AI projects and their respective AI project phases (Merhi, 2023; Vial et al., 2023; Westenberger et al., 2022). Thus, identifying, assessing, and understanding these challenges to provide solutions on an *organizational level* is essential for leveraging the potential added value that AI projects can create in the manufacturing sector.

Research Article 1# - Uncovering the Sweet Spot of Artificial Intelligence Projects: An Exploration of Success Factors

For organizations that aim to thrive in the trajectory of AI, an overarching systematic understanding of the requirements that drive the successful implementation of AI projects on an organizational level is indispensable (Duan et al., 2019; Dwivedi et al., 2021). In general, such requirements are usually referred to as success factors (SFs) (Bullen & Rockart, 1981). Success factors refer to settings and conditions that directly or indirectly influence the project outcome (Baccarini, 1999; Turner & Müller, 2003). In the literature, the study of the factors that affect the success and failure of IT projects is among the most prominent research streams (Dwivedi et al., 2015). Nevertheless, researchers and practitioners seeking an overview of the SFs for AI projects can only draw on a limited number of studies (Duan et al., 2019; Merhi, 2023). While one could argue that AI projects, as a subset of IT projects, share the same SFs, the SFs for IT projects do not immediately apply to AI projects due to the inherent characteristics of AI and the specific AI project workflow (Berente et al., 2021; Vial et al., 2023). As such, only a few studies list factors that facilitate the successful implementation of AI projects (e.g., Baier et al., 2019; Jöhnk et al., 2021; G. Miller, 2021; van Giffen & Ludwig, 2023; Weber et al., 2023). Furthermore, studies by Lee et al. (2023) and Merhi (2023) provide initial overviews of the antecedents of SFs, but they do not provide a holistic and integrated compilation of SFs, as they do not provide empirical evidence and only consider a limited amount of literature. Hence, research article #1 aims to answer the following research question: *What are the SFs for AI projects?*

The research article follows a three-stage research approach. First, a systematic literature review is conducted to develop an initial set of SFs (Webster & Watson, 2002; Wolfswinkel et al., 2013), following the five-step process: define, search, select, analyze, and present. This ultimately comprised 109 articles relevant for further analysis to iteratively synthesize insights and derive SFs for AI projects (Corbin & Strauss, 1990). Second, an in-depth interview study with 20 subject matter experts is performed, mainly from the manufacturing sector, to validate, refine, and extend the initial results toward a final set of SFs (Bettis et al., 2015; Goldkuhl, 2012). Third, a focus group discussion to situate the SFs in a broader context by mapping them to the four phases of the AI project workflow is conducted (Nyumba et al., 2018; Tremblay et al., 2010). The result is a framework of 24 SFs for AI projects, structured along four overarching success dimensions (i.e., datability, desirability, feasibility, and viability) and specified by 93 subordinated success manifestations. For each SF, the results provide a comprehensive description and depict specific AI characteristics that illustrate their relevance and necessity. Thereby, the four success dimensions specify the key action fields when planning and executing AI projects from a higher-level perspective. In contrast, the 93 success manifestations provide operational support for successful AI projects from a lower-level perspective. An exemplary SF "Data-centric technology stack" indicating its organizational-wide relevance is given below, whereby all SF can be found in the research article:

*"**Data-centric technology stack** aims to integrate data from different sources while remaining scalable and strives to create data repositories such as data lakes that enable mature, non-fragmented data structures* (Hukkelberg & Rolland, 2020; Mikalef et al., 2019). *A mature **data infrastructure** helps organizations leverage their AI systems over the long term* (Shollo et al., 2022). *Since data is never stored in just one place due to existing legacy systems* (Alsheibani et al., 2020), ***data centralization** helps to structure data through unified, centralized data stores and robust data storage systems with sufficient networking capabilities (Brock & Von Wangenheim, 2019; De Silva & Alahakoon, 2022). In this sense, organizations should aim for a robust data storage system as a key concern in AI projects* (Alsheibani et al., 2020; Merhi, 2023). *Furthermore, **data consolidation** helps extract, load, and merge data by combining data from various sources and transforming them according to the underlying use case* (Sjödin et al., 2021). *To do so, organizations must develop a scalable processing and data pipeline system to handle data actuality, velocity, and volume* (J. Li et al., 2021; Ransbotham et al., 2017). *Here,*

data transfer *refers to how data is transferred between different systems. This is mainly because data from legacy systems need to be made usable* (Pumplun et al., 2019; van Giffen & Ludwig, 2023). *Those resulting data transfer pipelines must be secured by creating automated routines for data transmission and transformation* (Merhi, 2023; Sjödin et al., 2021)*."*

Finally, by illustrating how the SFs manifest in the four key phases of the AI project workflow, as shown in Figure 2, the results provide an authoritative instance for a systematic understanding of the scope and scale in which they emerge on an overarching level in AI projects. Hereafter, four key findings, together with associated recommendations for the successful implementation of AI projects on an organizational level, are described.

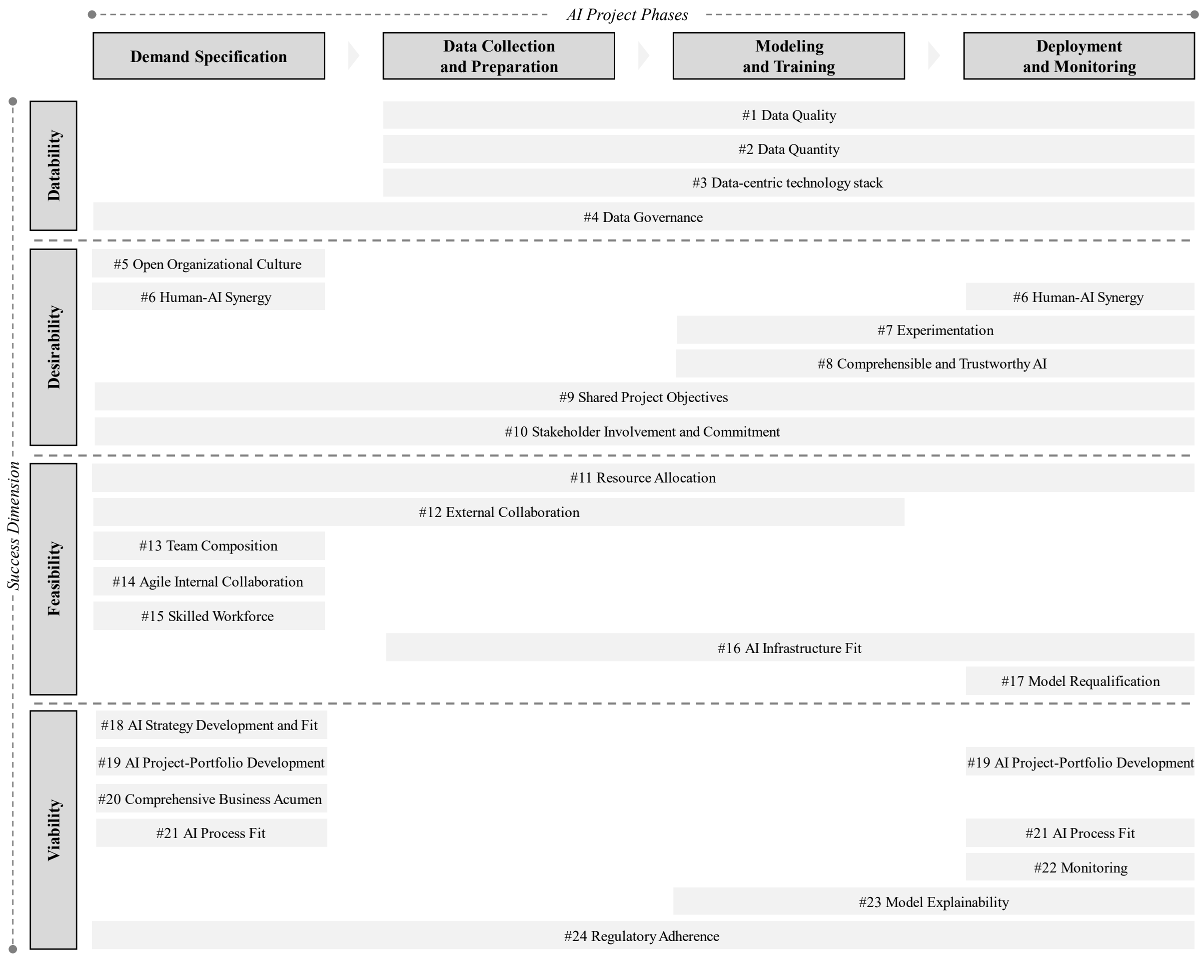

Figure 2. Contextualization of the SFs in the AI project workflow

The results are novel as they are the first to systematically explore the SFs for AI projects by synthesizing extant knowledge from scientific literature with emergent insights into the trajectory of AI through empirical and practical data. From a theoretical perspective, the study of the SFs for AI projects provides three main implications. First, the work complements the existing body of knowledge in IS research on SFs for IT projects, as well as the successful implementation of AI in organizations. Second, the results facilitate further theorizing on the successful implementation of AI in general. Third, the work contributes to recent discussions on the “sweet spot” concept regarding AI projects. While much is known about the interplay of the trifecta of desirability, feasibility, and viability as a foundation of project success, the research article found that AI projects only lead to valuable outcomes when datability is present. From a practical perspective, the study of the SFs for AI projects provides two main implications for managers as organizational decision-makers (e.g., business development representatives and AI project leads). First, managers should leverage the results to structure strategic discussions among various organizational stakeholders on how to advance the successful implementation of AI. Second, managers can build on the results to assess and actively monitor the extent to which their current AI projects cover relevant SFs.

Research Article 2# - Data or Business First? – Manufacturers´ Transformation toward Data-driven Business Models

The manifold possibilities for AI-based and data-driven value propositions available to manufacturers further complicate the business challenge at the beginning of AI projects to explore a target state or archetype that fits existing company processes as well as existing and future customers. Here, different value propositions characterize different archetypes of data-driven business models (DDBMs) (Hunke et al., 2022), leading to significant variations in the organizational capabilities required (Vial, 2019). Against this backdrop, research article #2 guides manufacturing companies in, on the one hand, the identification of DDBMs by using DDBM archetypes that can be adopted (Müller & Buliga, 2019) and, on the other hand, the transforming toward the archetypes using a maturity model to address the following research question: *What capabilities do manufacturers require to transform toward distinct archetypal data-driven business models?*

The methodological approach follows the procedure model of Becker et al. (2009) that specifies the DSR methodology for maturity models regarding their design and evaluation. Based on interviews with practitioners, the work outlines three key requirements for the maturity model to be developed: First, the model should integrate established business model archetypes to offer clear guidance on the target state of transformation. Second, the model should allow comprehensive coverage of socio-technical capabilities on all organizational layers. Third, the model must include complete capability descriptions to enhance the model’s prescriptive value and usability for practice. Based on these requirements, the data-driven business model maturity model (DDBM3) was developed using the archetypal data-driven business models of Hunke et al. (2022) (i.e., data provider, insight provider, recommendation provider, and digital-solution provider) as maturity levels (columns). To cover all organizational levels, it uses the layered enterprise architecture model of Urbach and Röglinger (2019) to structure the model’s 22 capability dimensions (rows) in five major focus areas. Finally, as a continuous maturity model, it offers capability descriptions for every cell in the resulting matrix.

The archetypal DDBMs serve as strategic orientation for the transformation by systematically characterizing different configurations of DDBMs. In the first archetype data provider, manufacturers provide customers with (product) data beyond the physical product. The data is only moderately processed, provided in a standardized form, and subject to descriptive analytics (Hartmann et al., 2016), such as aggregated reports on machine utilization. The insight provider business model delivers diagnostic and supportive insights, such as digital alerts triggered when machines or processes malfunction. The customization increases further with the recommendation provider business model, which provides customized recommendations based on predictive analytics, such as AI-driven root cause analysis or automatic course of action (Hunke et al., 2022; Sarker, 2021). Finally, manufacturers can act as digital solution providers, opening up novel ways of doing business, e.g., by turning into a smart data platform provider. In this context, value is created through unique digital information, and the physical product recedes into the background (Beverungen et al., 2021).

Figure 3 illustrates the DDBM3, encompassing the five focus areas based on the enterprise architecture mode and the five archetypes. The first focus area, business model, includes four capabilities dimensions (i.e., “value proposition,” “customer interaction,” “monetization and pricing,” and “sales and channel management”). These outward-faced capability dimensions are essential for manufacturers to define, market, and monetize data-driven business models based on the archetypal offerings. The second focus area, business processes, includes four capability dimensions (i.e., “strategy and vision for data-based business,” “data-centric process management,” “knowledge sharing and management,” and “product life cycle management”). It especially covers specific processual capabilities that create, deliver, and capture the value of data-driven services and describes how to manage the needed activities. The third focus area, people and applications, includes cultural aspects (“recognition and mindset”), soft and hard skills (“methods,“ “data analytics competencies”), responsibilities (“roles and responsibilities”), and tools (“data analytics tooling”) for data-driven business models at the employee level and seeks to empower the staff members. The fourth focus area, data and information, comprises four capability dimensions (i.e., “applied forms of analytics,” “data management, “ “data governance and quality,“ and “horizontal and vertical data integration”). Thus, it covers data management, the integration of data sources (horizontal and vertical), the establishment of governance and quality mechanisms, and the analytics applied to data. The last focus area, infrastructure, is comprised primarily of technological enablers that organizations need to provide DDBMs. The associated capabilities include the secure and scalable operation of software and hardware. It comprises five capability dimensions (i.e., “data analytics software management and operations, “ “data-driven service integration and deployment,” “data architecture and scaling,“ “cybersecurity and -privacy,” and “cyber-physical systems and connectivity”).

In line with the development procedure of Becker et al. (2009), the DDBM3 was evaluated using artificial and naturalistic settings. The artificial evaluation is based on an academic focus group (Nyumba et al., 2018; Tremblay et al., 2010). The naturalistic evaluation involves applying the model with executives of two manufacturing companies, Alpha and Beta, to assess their status quo and target the state of transformation toward DDBM (Sonnenberg & vom Brocke, 2012b) (Sonnenberg & vom Brocke, 2012). The results show that the two manufacturers took different approaches to transform toward data-driven business models, namely “data first” and “business first”, as presented in Figure 3. For Alpha, a “business first” approach was identified. Alpha’s transformation was mainly driven from the business side as customers demanded data delivery from connected machinery. Hence, Alpha demonstrates higher maturity levels in most capabilities dimensions relating to the business model and business processes focus areas in the upper levels of the DDBM3. Against this, shortcomings were identified in the bottom capability dimensions, for example, in the focus areas of data and information as well as infrastructure. In contrast, Beta, with the “data first” approach, exhibits more mature capabilities in the technological and data-related areas of the DDBM3 (e.g., “cyber-physical systems” or “data analytics software management and operations”). However, Beta’s capabilities are less developed in the business-related capability dimensions, such as value proposition and digital channel management.

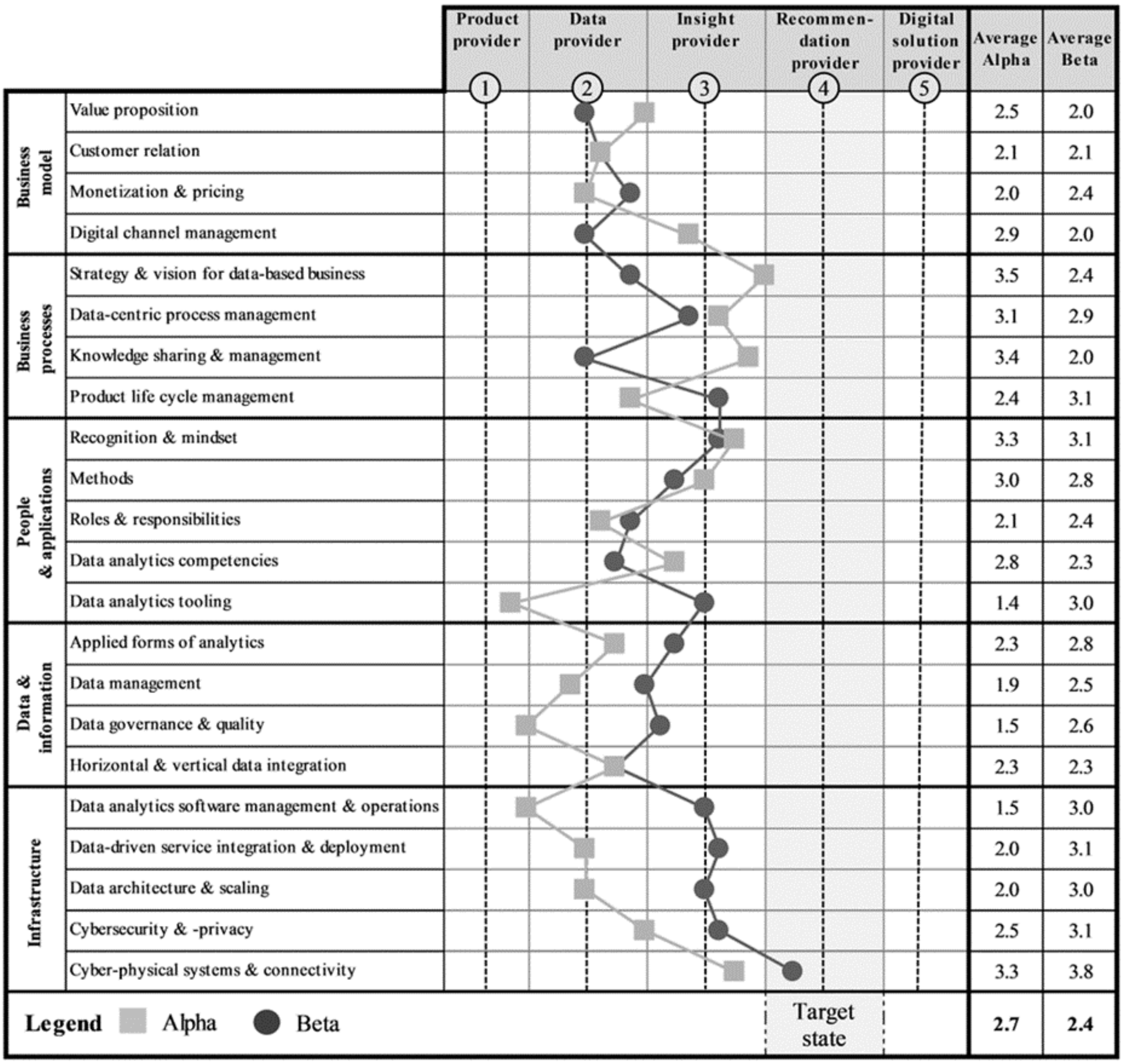

Figure 3. The data-driven business model maturity model (DDBM3) and its application at two manufacturers

All in all, the DDBM3 enables the exploration of a target state or archetype for DDBM. The contribution of this research article is twofold: First, the model can serve as an ‘analytical lens’ that allows an investigation of the progress of the DDBM transformation across all organizational levels (Pöppelbuß & Röglinger, 2011). Thus, research can identify patterns and maturation paths that emerge from its application. This may unravel success factors or impediments associated with distinct paths of maturation. Second, the naturalistic demonstration of the DDBM´s proves applicability and usefulness for practitioners. In this vein, the DDBM3 allows a status quo assessment and derives fields of action to develop the capabilities required for the aspired DDBM, especially at the beginning of the respective projects.

Research Article #3 – Managing the Inevitable – A Maturity Model to Establish Incident Response Management Capabilities

Nevertheless, the introduction of new digital capabilities and technologies through AI projects causes the attack surface of organizations to expand continuously. In this vein, the global cost of cyber-attacks is estimated to grow by 15% per year, from $3 trillion in 2015 to $10.5 trillion annually by 2025 (Microsoft, 2016; Morgan, 2021). Nowadays, around two-thirds of organizations are affected by cyber-attacks yearly, some even several times (Barreuther et al., 2022; Cyberedge Group, 2021). Despite the best security measures, security incidents cannot be completely averted. AI projects, in particular, are susceptible to security incidents as they process large volumes of data that are often highly critical because, for example, they allow conclusions to be drawn about the machines and their intellectual property. Hence, the mitigation of these incidents plays a decisive role in reducing the extent of damage and restoring the operability of systems as quickly as possible across the entire organizational structure (Ahmad et al., 2022).

Therefore, Incident Response Management (IRM) has been established as an effective tool for reactive IT security (Van Der Kleij et al., 2022). IRM aims to maintain the business processes, minimize the impacts of security incidents, and respond effectively to them (Ruefle et al., 2014; Wegener et al., 2016). Thereby, IRM can be seen as a process including the phases of incident preparation, incident detection, incident resolution, and post-incident activity (Ab Rahman & Choo, 2015; Ahmad et al., 2022; Cichonski et al., 2012). Although IRM has great significance within most companies (Ahmad et al., 2012; Ruefle et al., 2014), it is often not well developed. Often, IRM is seen as a cost-center because it creates resourcing constraints (Ahmad et al., 2021) and management awareness is missing (Van Der Kleij et al., 2022). Furthermore, organizations face challenges like limited resources (Kuypers et al., 2016) since IRM is a timely and cost-intensive task. Nevertheless, IRM can be considered crucial for organizations as incidents can escalate into emergencies and lead to reputational or financial losses, besides disrupting business continuity (Farahmand et al., 2003; Thangavelu et al., 2021). In March 2024, for instance, the battery manufacturer Varta had to shut down its entire five production sites and administration for around four weeks after a ransomware attack (Varta, 2024). According to the IBM 2022 Cost of a Data Breach Report (IBM, 2022), the average data breach cost for manufacturers was around $4.47 million. To address this problem, managers need to know what IRM capabilities are required to allocate their tight budget appropriately. However, many approaches do not apply to organizations with immature IRM capabilities, especially for small and medium-sized enterprises when the IRM maturity is low. Thus, research article #3 answers the following research question: *Which capabilities do organizations need to approach incident response management?*

To tackle this research question, the central artifact of research article #3 is an IRM maturity model (IRM3) closely aligned with practical requirements. The work follows the well-established approach of Becker et al. (2009), building on the DSR principles by Hevner et al. (2004) and the evaluation patterns by Sonnenberg and vom Brocke (2012b) (i.e., Eval 1-4).

First, five expert interviews were conducted to derive the need for the IRM3 and elicit three design requirements: First, applicability to organizations with immature IRM. Second, consideration of the social-technical perspective. And third, extensive evaluation in practice. Afterward, the research gap was justified by identifying related MMs screening existing models based on the interviews and a literature search (Eval 1). This revealed that the existing models (e.g., the Security Incident Management Maturity Model 'SIM3' of Stikvoort (2019)) could not fulfill or only partly fulfill all three requirements. Thus, building on and enhancing existing literature, a new MM was developed (Becker et al., 2009). The resulting IRM3 possesses four focus areas (i.e., organization, human, tools, and processes) and a total of 29 capability dimensions (Figure 4). The capabilities describe the as-is situation of IRM in a particular stage and enable organizations to find themselves in one of the capabilities. With regard to a focus level MM-design, the number of maturity capabilities varies between two and five and changes in terms of quality or quantity over the stages (Van Steenbergen et al., 2010). At the end of the development process, the IRM3's design was evaluated regarding fidelity with the real-world phenomena, completeness, and internal consistency (Eval 2) (Sonnenberg & vom Brocke, 2012a) using an academic focus group (Tremblay et al., 2010). First, the focus area organization contains seven dimensions describing pre-defined interaction of humans, resources, infrastructures, and processes. It is about specific and strategic goals related to IRM. It includes fundamental principles and organizational measures to structure and implement IRM. The realization of these organizational aspects requires the involvement of decision-makers. Second, the focus area human consists of six dimensions that describe how employees work together to realize organizational goals. This focus area considers the collective values and behaviors of individuals or teams and, thus, the human factor. Consequently, the area covers dimensions that affect or require employee participation to respond appropriately to incidents. Third, the focus area tools contains eight dimensions and concentrates on the applications, programs, services, and other parts of equipment to conduct incident response. These tools enable the company to achieve the goals described in the focus area organization. With their help, an organization can improve its IRM regarding time, granularity, or quality. Fourth, the focus area processes consists of eight dimensions and defines IRM procedures carried out by tools or humans. The procedures support the incident management or services, which are part of the incident response process. To increase the effectiveness of IRM, procedures need to be repeatable, measurable, adaptable, and documented. For Eval3, the IRM3 was transferred into an online survey tool to enable a straightforward assessment. This online tool was pre-tested by the authors' team and then validated by two practitioners. As this yielded minor corrections to the model, the authors decided to move forward to an application of the IRM3 in a real-world context. Hence, in Eval 4, the IRM 3 was applied to seven German organizations from the manufacturing sector to provide a naturalistic evaluation and assess its practical value (Sonnenberg & vom Brocke, 2012a). After conducting seven applications, which no longer led to any new adjustments, the model's effectiveness, applicability, and generality were proven.

The IRM3 and the design process contribute to theory and practice. From a practical point of view, the work provides a highly applicable and accessible tool that enables management to approach IRM in a structured way in AI projects. In doing so, the IRM3 leverages existing knowledge from other works in the IRM domain to provide a holistic scope of IRM capabilities. The model's descriptive value helps organizations capture the status quo to identify weaknesses in IRM. With this assessment of their current IRM practice, organizations can prioritize further measures and decide which capabilities need to be considered in their context. From an academic perspective, the IRM3's aim and scope enrich the existing knowledge base on IRM. By targeting an industry-independent IRM capability assessment for low-maturity organizations, IRM3 addresses a relevant but unfilled research gap. In contrast to existing works, the IRM3 in this paper is developed in a practical context, thus going through the entire artifact development and evaluation process (Becker et al., 2009). On the one hand, this increases the validity of the IRM results and, on the other hand, provides relevant insights into the development of IRM capabilities in different organizations.

Focus Area	Capability Dimension	Maturity Capabilities				
Organization	Management Support	De-prioritization & Negligence	Management Awareness	+ Active Management Support	+ Sufficient Resources	
	Service Description	No Service Description	Implicit Service Description	Explicit Service Description	+ Regular Reviewing of Service Description	
	Responsibility	Responsibility Unclear	Clearly Assigned Responsibilities	+ Regular Review & Adjustment of Responsibilities		
	Emergency Availability	Undefined Contact Points	Defined Contact Points	+ Publicity of Contact Points	+ Ensured Reachability of Contact Points	
	Incident Classification	Ad-hoc & Intuitive Classification	Case-by-Case Classification	Developed Classification Scheme	+ Institutional Application of Classification Scheme	+ Permanent Improvement of Classification
	Security Policy	Focus on Damage Avoidance	+ Adherence to Legal Guidelines	+ Defined Internal Guidelines	+ Adaption towards Internal Guidelines	+ Consideration of Changing Requirements
	External Collaboration	No Collaboration	Case-by-Case Collaboration	+ Involvement of External Service Providers	+ Acquisition of Cyber Insurance	
Human	Security Awareness	No IT Security Awareness	Awareness of Existence	+ Awareness of Importance	+ Training and Awareness Raising	
	Communication Culture	Reactive Communication	Proactive Communication	+ Know-how-Transfer	+ Competence Sharing	
	In-House Cooperation	Occasional Interaction	Proactive Interaction & Teamwork	+ Cooperation across Departments		
	Personnel Resilience	Insufficient IT Workforce	Insufficient Security Workforce	Dedicated Security Workforce	+ Absence-resilient Workforce	
	Personnel Characteristics	Pro-forma Assignment in IT	Dedicated Competence Profile	+ Skilled & Trained Personnel	+ Experienced Personnel	
	Training Opportunities	No Training Opportunities	Demand-oriented Training	Training for Person in Charge	+ Training for IT Members	+ Company-Wide Training
Tools	IT Resources	No IT Resource & Asset Management	Regular Inspection of Resources	+ Classification of Resources	+ Documentation & Description of Critical Resources	
	Work Equipment (Systems)	Vulnerable IT Systems	Protected IT Systems	+ Implementation of Redundancies	+ Highly Resilient and Scalable Systems	
	Work Equipment (Personnel)	Vulnerable Work Equipment	Protected Work Equipment	+ Equipment for Replacement	+ Seamless Recovery	
	Prevention Toolset	Unestablished Prevention Tools	Integrated Antivirus Programs & Firewalls	+ Access Control & Management	+ Tools for Patch Management	
	Detection Toolset	No Detection Tools	Unestablished Detection Tools	Tools for NetFlow & Traffic Analysis	+ Tools for Endpoint Detection and Response	
	Tracking System	No Tracking of Incidents	Manual Tracking of Incidents	Incident Tracking System		
	Resolution Toolset	Unestablished Resolution Tools	Integrated Logging Tools	+ Tools for Configuration & Backup Management	+ Established Forensic Tools	+ Integration of Sandboxing Tools
	Documentation System	Unsystematic Collection of Data & Knowledge	System for Knowledge Management	+ Encouragement of Documentation	+ Time for Documentation & Reflection	
Processes	Incident Prevention	No Prevention Measures	+ Defined Prevention Measures	+ Established Prevention Measures	+ Permanently Reviewed Prevention Measures	
	Incident Detection	Unsystematic Detection Measures	+ Defined Detection Measures	+ Established Detection Measures	+ Permanently Reviewed Detection Measures	
	Escalation	Unsystematic Escalation	Implicit Escalation	Systematic Escalation	+ Defined Escalation Criteria	+ Continuous Process Improvement & Learning
	Incident Resolution	Unsystematic Resolution Measures	Defined Resolution Measures	+ Established Resolution Measures	+ IT Business Continuity Management	+ Permanently Reviewed Resolution Measures
	Incident Reflection	No Reflection of Incidents	Case-by-Case Learning	+ Change Management Process	Permanent Improvement	
	Audit	No Revision Measures	Self-Assessment & Learning	+ Integrated Audit Management	+ Internal & External Assessment	+ Certification
	Knowledge Acquisition	Implicit Knowledge Acquisition	Systematic Research	+ Exchange & External Networking		
	External Communication	Negligence for Security	Unsystematic PR Activity	Market Observation	+ Systematic & Proactive PR Activity	

Zeta | Eta | Theta | Jota | Kappa | Lambda | My

Figure 4. The Incident Response Management Maturity Model (IRM3) and its application

All in all, to conclude section II.1, the three research articles #1, #2, and #3 contribute to overcoming the challenges AI projects pose at an organizational level. Research article #1 focuses on an overarching view of challenges and solutions across all four AI project phases by deriving practical-grounded SFs. Research article #2 is primarily concerned with the first AI project phase, "Demand Specification", with the identification of meaningful target states in the form of DDBMs and the structuring of the organization-wide capabilities required for this. Research article #3 focuses on the later AI project phase and the necessary IRM throughout the entire organizational layers, as AI projects increase the cyber attack surface, such as the unwanted outflow of machine data through the introduction of new digital technologies.

2 Challenges and Solutions at a Project Level

Besides overcoming the challenges of AI projects at an overarching organizational level, manufacturers are confronted with the specific challenges in individual AI projects. In each AI project, the AI project phases are characterized by sequential dependencies, feedback loops, and an indefinite number of data exploration as well as AI model experimentation cycles (Amershi et al., 2019; Kreuzberger et al., 2023; Vial et al., 2023). The results of the AI model's performance and, thus also, the success of the AI projects is greatly influenced by the underlying data basis and, in particular, the data labels provided, which makes it difficult to plan and manage the AI project outcome (Merhi, 2023; Vial et al., 2023). Additionally, especially in real-world AI projects whose results are delivered to or used by end users, the explainability of the AI models must be considered to create trust and acceptance (Barredo Arrieta et al., 2020; Brasse et al., 2023). This explainability needs to be implemented technically and, furthermore, addresses the increasingly relevant sustainability challenges of AI projects. The growing widespread use of ever-larger AI models means that the environmental, social, and governance challenges must additionally be taken into account in each phase of an AI project (Papagiannidis et al., 2023; Veit & Thatcher, 2023; Verdecchia et al., 2023). Hence, motivated by these challenges, the following three research articles seek to provide solutions at a *project level*.

Research Article #4 – A Data-Efficient Active Learning Architecture for Anomaly Detection in Industrial Time Series Data

A large number of AI projects in the manufacturing sector deal with anomaly detection, especially due to the rising number of CPS and the amount of data available (Bertolini et al., 2021; Pang et al., 2022). Anomalies are described as instances that significantly deviate from other observations, allowing conclusions to be made about unexpected machine behavior (Z. Li et al., 2021). Hence, anomaly detection is increasingly recognized to reveal early indicators of machine failures and to enable condition monitoring, fault diagnosis, or predictive maintenance (Bertolini et al., 2021; Feng & Tian, 2021). However, with the increasing complexity of today's CPS, traditional rule-based approaches for anomaly detection are insufficient, and AI-based approaches are gaining importance (Barbado et al., 2022; Yuan & Wu, 2021). Since CPS are often operated continuously and multiple sensors are used, multivariate time series data is a starting point for anomaly detection (Zhao et al., 2020). Although supervised and unsupervised AI approaches reveal successful results for multivariate time series, barriers exist in practice (Z. Li et al., 2021; Saqlain et al., 2023). Supervised learning models often struggle with class imbalances and require high numbers of labeled instances (anomalies). These labels typically require domain-specific knowledge and are time-consuming to obtain (Alaei & Noorbehbahani, 2017; Yuan & Wu, 2021). Unsupervised learning models do not require labeled data and are most conventionally used (Chevrot et al., 2022; Pang et al., 2021). However, they are ineffective in handling high-dimensional data and are susceptible to high false-positive rates (Aggarwal, 2017).

To overcome this bottleneck, research article #4 proposes an Active Learning Architecture for Anomaly Detection in Manufacturing CPS called ALMAN. The key motivation is to efficiently use data and expert feedback to minimize burdens in AI projects while aiming for improved anomaly detection performance. Active learning systems aim to label previously unlabeled instances by querying an oracle, i.e., a human expert, thereby enhancing accuracy with as few labeled data as possible, thus minimizing the costs of labeling (Das et al., 2016; Finder et al., 2022; Ren et al., 2022). The goal is to select an optimal number of unlabeled data instances that are annotated by a domain expert, maximizing the learning ability of the AI model (X. Wu et al., 2021). An active learning framework consists of two components: a query engine that selects the data instances from the unlabeled pool and an oracle that provides the corresponding labels (Finder et al., 2022; Settles, 2010). The query process is repeated until adequate performance is achieved (Settles, 2010). Das et al. (2016) introduce a budget of B queries indicating the capacity of the expert system. The goal is to maximize the total number of true anomalies within the budget B.

To realize the ALMAN's architecture, this work combines an unsupervised deep-learning model with a supervised deep-learning model involving a human feedback module. Thus, the following three components are defined: an unsupervised model, a feedback module, and an active anomaly detection model (Figure 5).

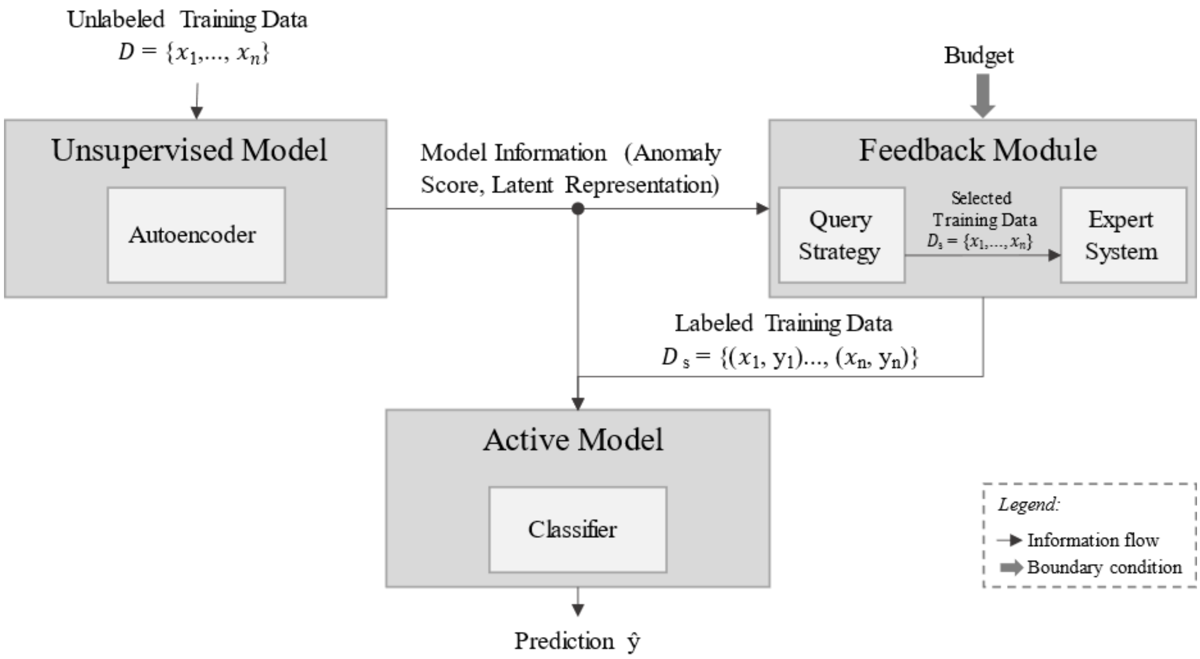

Figure 5. Components and information flow of the active learning architecture

The appropriate selection of the unsupervised model builds on the taxonomy of deep learning approaches for anomaly detection (Pang et al., 2022). Hence, an autoencoder (AE) is chosen that is trained to reconstruct a given input by finding a latent representation of the data (Chevrot et al., 2022; Goodfellow et al., 2016). Anomaly detection with an AE relies on the assumption that normal cases can be reconstructed more accurately than anomalies. The unlabeled data and the previously described unsupervised model serve as input for the feedback module. The module's output represents a set of labeled data (x, y) as depicted in Figure 5. Internally, the feedback module consists of a query strategy and an expert system. The query strategy selects a subset of examples from the training data that are presented to the expert system for classification. Subsequently, the architecture of the active supervised model is based on a feedforward neural network to finally classify the anomalies.

To evaluate and interpret ALMAN's performance, the research article #3 uses the real-world universal robot screwdriving anomaly detection data set (AURSAD) (Leporowski et al., 2021). The data set describes the repeated execution of an assembly process by a collaborative robot type with an average execution time of 15 s. The data set includes 1,420 records representing normal operations and 625 records indicative of various fault states, such as damaged screws or damaged plate threads. With the goal of a data-efficient approach, the AURSAD data is divided into two data sets for binary anomaly detection. Data set I contains four features of energy consumption data, and data set II contains seven features of mechanical properties. After a five-stage processing of this data, three different experiments are conducted. First, a sensitivity analysis is applied to identify the optimal query strategy for ALMAN. Second, the approach is compared with the solely unsupervised model to evaluate the performance of active learning. Third, the data efficiency of ALMAN is analyzed. Within those experiments, ALMAN´s data efficiency was proven as exemplary shown in Figure 6. While the active learning approach detects anomalies with comparable high performance for both data sets (F1 scores of 0.865 for energy data and 0.849 for mechanical process data), the unsupervised model trained based on mechanical process data fails to reliably detect the anomalies (F1 score of 0.286) but performs sufficiently well with the help of active learning (F1 score of 0.545). Consequently, in both approaches, the use of energy consumption data considering the F1 score is more reasonable.

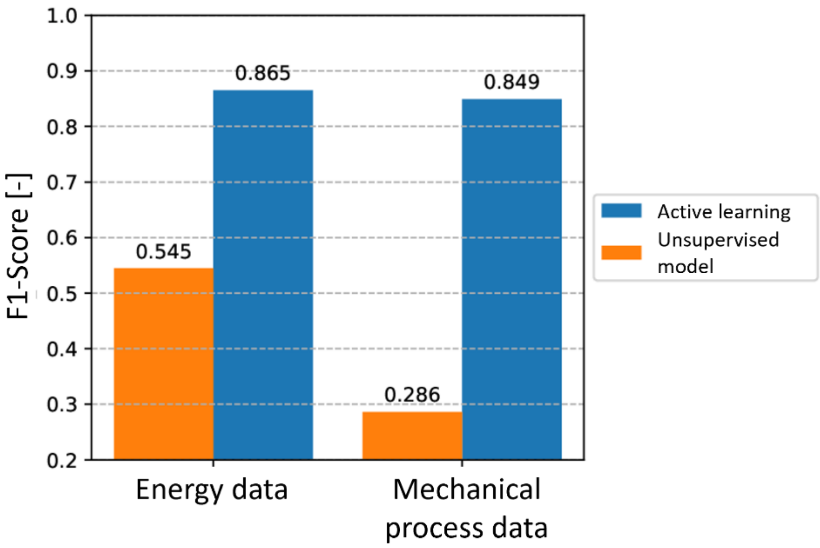

Figure 6. F1 score comparison of the active learning system with the unsupervised model for energy and mechanical process data

In summary, the research article #4 makes the following four contributions. First, it proposes ALMAN to overcome the limitations and drawbacks of existing supervised and unsupervised approaches for anomaly detection in manufacturing applications. Second, it demonstrates and validates the developed active learning architecture in a case study of a cyber-physical robotic screwdriving application. The results indicate that the active learning system outperforms a state-of-the-art unsupervised model by 59% in F1 score. Third, the work investigates the data efficiency potential of using energy consumption data only for anomaly detection instead of common, hardly accessible mechanical process data. Findings emphasize the data efficiency potential of energy consumption data with F1 scores similar to those using mechanical process data. Fourth, the active learning architecture enables costly expert feedback to be used efficiently and, thus, reduces concerns about the limitations of existing anomaly detection approaches or even adds an active learning model to an already existing unsupervised model in manufacturing applications.

Research Article #5 – Leveraging Explainable AI for Informed Building Retrofit Decisions: Insights from a Survey

Although previous work shows that such efficient but complex AI models as, for example, presented in the previous research article #4, can often achieve more accurate predictions than conventional rule- or physical-based methods, these models come with the expense of lacking explainability, referred to as the black-box problem, which leads decision-makers to distrust or even reject them (Barredo Arrieta et al., 2020; Burkart & Huber, 2021). Indeed, comprehending why a model makes certain decisions is often as important as its prediction accuracy (B. Kim et al., 2020; Shin, 2021). XAI can be leveraged to create this understanding as it helps to comprehend how a model decides, predicts, and performs its operations. Hence, research article #5 implements three common transparent models (Linear Regression, Decision Tree, QLattice) and applies four prevailing XAI methods (Partial Dependency Plots, Accumulated Local Effects, Local Interpretable Model-Agnostic Explanations, Shapley Additive Explanations) to an artificial neural network (ANN) to evaluate the effectiveness of these XAI methods. Measuring the effectiveness of XAI methods can either be done by using quantitative objective metrics such as sensitivity measures (Kindermans et al., 2019; T. Miller, 2017; Vilone & Longo, 2021) or by conducting human-centered evaluations collecting end-user feedback (Ali et al., 2023; Ding et al., 2022b). Regarding the last, either qualitative questions (i.e., open-ended survey) aimed at achieving deeper insights or quantitative questions (i.e., close-ended survey) aimed to be statistically analyzed can be used (Ali et al., 2023; K. Lee et al., 2022). Nevertheless, around 70% of research articles neglect evaluating XAI methods with potential users (Brasse et al., 2023) or only emulate the user evaluation (Ali et al., 2023), leading to inaccurate human-centered insights (Brasse et al., 2023; S. W. Kim et al., 2022b). To conduct this human-centered evaluation, research article #5 uses a real-world dataset of 25,000 residential buildings as the energy consumption prediction in buildings remains a challenge to fighting climate change with widely reported inaccuracies in prediction, known as the energy performance gap. Especially the building sector accounts for 36% of total global energy consumption and, therefore, faces a need for decarbonization (Ahlrichs et al., 2022; Visscher et al., 2016). Also, the manufacturing sector, which has a large number of old production halls combined with decreasing demolition rates (Saffari & Beagon, 2022), necessitates both an increase in the stagnating rate (Mayer et al., 2022) and depth of retrofits to reduce energy consumption effectively (Tsoka et al., 2022; Yalcintas, 2008). In addition to the environmental aspect, adequate retrofit measures are often cost-effective (Adisorn et al., 2020; Ahlrichs & Rockstuhl, 2022). Hence, these issues lead to the first research question of research article #5: *What is the perceived degree of explainability of explainable artificial intelligence methods in building energy consumption forecasting?*

As explainability is typically viewed as a trade-off with prediction accuracy (Barredo Arrieta et al., 2020), it is of interest to investigate this trade-off in the case of building energy consumption forecasting, leading to the second research question: *To what extent does explainability affect the prediction accuracy of machine learning models in building energy consumption forecasting?*

To answer these two research questions, the work follows a three-step approach (Figure 7). First, the four ML models and the four XAI methods on a real-world dataset of German one- and two-family residential buildings are implemented. The model's prediction accuracy is assessed with three commonly used prediction accuracy metrics, i.e., metrics Mean Absolute Error (MAE), Root Mean Square Error (RMSE), and Mean Absolute Percentage Error (MAPE), for predicting annual building energy consumption (Amasyali & El-Gohary, 2018). Second, the degree of explainability of these models and XAI methods is evaluated by conducting an online survey with 137 participants, thereby addressing the first research question. The survey is conducted among mostly non-AI-experts based on the two human-centered dimensions of Explanation Satisfaction and Perceived Fidelity (Hoffman et al., 2018; Löfström et al., 2022) using a seven-point Likert scale. Third, the second research question is examined by evaluating the prediction accuracy of the AI models while taking into account the explainability based on the survey. Hence, the results are combined and analyzed to address the trade-off between prediction accuracy and explainability in data-driven building energy consumption forecasting and derive implications.

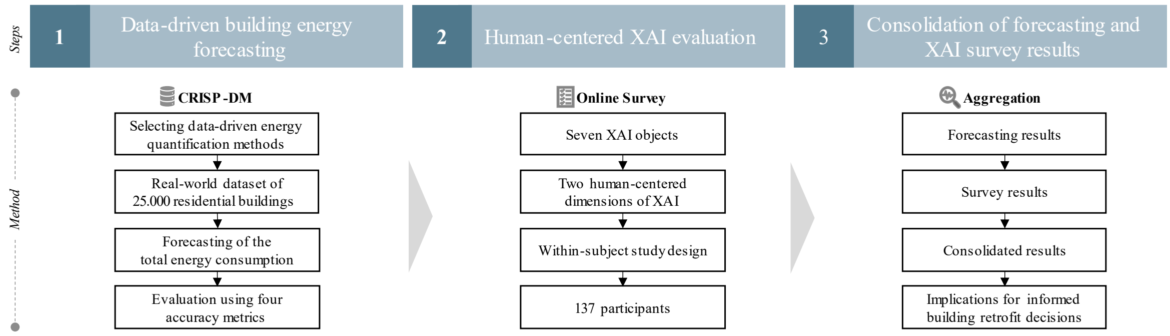

Figure 7. Methodological three-step approach of research article #5

The results of the three steps are shown in compact form in Figure 8. First, the results of the prediction accuracies, i.e., the first step, are presented on the left-hand side. The final results of the models on the testing set confirm the findings of previous works that the ANN achieves better prediction accuracy results than the transparent models (Dosilovic et al., 2018). When looking at the MAE and the RMSE (MAE = 32.94, RMSE = 43.67), the ANN achieves a better value by about 4% than the transparent models on average. To statistically test these observations, a Wilcoxon-Signed-Rank test (Siegel, 1956) with a 1% significance level is applied. The test statistically confirmed the assumption that there is a difference between the prediction accuracy of the ANN and each of the transparent models, but not within the transparent models. Second, the XAI methods elevate the ANN to a comparable level of explainability as the transparent models. The decision tree achieves the best results in terms of explainability with a score of 5.21, followed by linear regression with a score of 5.07. Thus, the two common transparent models fare the best. However, they are closely followed by the XAI methods PDP (5.01) and SHAP (4.83), with some differences not even being statistically significant. Next in order, with a little distance, are ALE (4.55) and LIME (4.68). Thus, the right choice of the post-hoc XAI methods based on the well-performing ANN enables an increase in the Explainability by 10% (i.e., when considering ALE with 4.55 to the comparable global post-hoc method PDP with 5.01). Third, the results indicate that for the transparent model's linear regression (RMSE = 45.55) and decision tree (RMSE = 45.33), the higher Explainability is accompanied by poorer Prediction Accuracy compared to the opaque model ANN (RMSE = 43.67), as shown in the right-hand side of Figure 8.

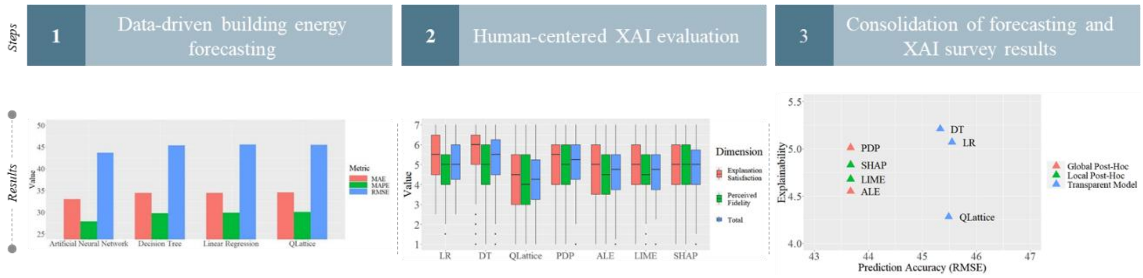

Figure 8. Results: 1) Comparison of prediction accuracies 2) Explainability evaluation from the online survey 3) Trade-off between explainability and prediction accuracy

In sum, research article #5 contributes to existing research in five ways. First, it closes the existing research gap of the lack of evaluation of XAI methods by real end users, i.e., potential property owners, which leads to meaningful research results that can be applied in practice. Second, various XAI methods are applied to the prediction of the long-term energy performance of buildings with the aim of explaining the prediction mechanisms, considering the influence of numerous input features. These XAI methods, on the one hand, reduce complexity while maintaining accuracy by removing less important input features and, on the other hand, provide guidance for decision-makers by revealing the key factors to focus on when determining appropriate retrofit measures (Pham et al., 2020; Rai, 2020). Third, it demonstrates a practical approach for a human-based measurement and evaluation of the degree of Explainability of XAI methods based on two dimensions, which can be transferred to other fields (Löfström et al., 2022). Fourth, by addressing the research gaps and providing an analysis of the application of XAI methods to an ANN, which has been done insufficiently in the residential energy context (Machlev et al., 2022; Tsoka et al., 2022). Fifth, the results are transferred into implications and recommendations for research, policy, and decision-makers based on the quantified trade-off between prediction accuracy and explainability.

Research Article #6 – Towards Sustainability of AI – Identifying Design Patterns for Sustainable Machine Learning Development

As more and more AI projects are conducted (Benbya et al., 2021; Berente et al., 2021; Bertolini et al., 2021), AI’s negative impacts on resource consumption, societal injustice, or even human rights cannot be neglected anymore (Cowls et al., 2023; Dennehy et al., 2023; Koniakou, 2023), leading to calls to work toward the sustainability of AI (SAI) (Schoormann et al., 2023; Schwartz et al., 2020; Tornede et al., 2022). SAI describes the sustainable design, development, and use of AI through its entire lifecycle, i.e., across all phases of the AI project workflow (van Wynsberghe, 2021). Nevertheless, previous work is fragmented across several streams, leading to overlapping recommendations and difficulties, especially for practitioners, to comprehensively assess possible measures to conduct more sustainable AI projects. At the same time, there is an increasing call for research to shift from pure principles to comprehensive design approaches and implementable best practices for SAI, for instance, to avoid involuntary exclusion or unnecessary resource consumption (Dennehy et al., 2023; Pappas et al., 2023; Shneiderman, 2021; Vassilakopoulou & Hustad, 2023). Here, design patterns (DPs) have been proven to be valuable, as they capture best practices, guidelines, and recommendations and are a common tool to provide methodological support (Gamma, 1995; Goel et al., 2023). They have the advantage of being specific to solve a problem but also generic enough to address future similar problems, as they provide simple entry points and are easy to understand (Gregor et al., 2020). Thus, research article #6 poses the following research question: *What are design patterns that ML development stakeholders can incorporate to increase the sustainability of the ML development process?*

To answer this research question, the overarching artifact of the research article #6 is a comprehensive framework, namely the Sustainable Machine Learning Design Pattern Matrix (SML-DPM), that provides researchers and practitioners with recommendations to increase the sustainability of the ML development process and thus the sustainability of AI projects. The SML-DPM provides 35 DPs structured along four phases of the AI project workflow and subdivides them into three sustainability dimensions. The work follows the DSR paradigm to develop the SML-DPM in close alignment with four literature-grounded key requirements (Hevner et al., 2004; Peffers et al., 2007): 1) End-to-end consideration of the ML development process, 2) Holistic view on sustainability, 3) Applicability of the design patterns for ML development stakeholder, and 4) Clear assignment of the ML development stakeholders involved. The first set of DPs is derived from 41 multivocal references (Garousi et al., 2019). To evaluate and iterate on these DPs, the criteria developed by Sonnenberg and vom Brocke (2012a) are used. Thus, the work first assesses their applicability and usefulness through focus groups and semi-structured interviews with industry experts. Thereafter, a web-based prototype to evaluate the intentions of users, levering the SML-DPM based on a case study in three real-world AI projects, is developed.

The final SML-DPM (Figure 9) is divided into the ESG dimensions on the vertical axis and the AI project phases on the horizontal axis. The environmental dimension encompasses 14 DPs, the social dimension 12 DPs, and the governmental dimension 9 DPs. The DPs are assigned to the ML development stakeholders. The work opted for this explicit and non-overlapping one-to-one allocation, which makes it clear when which DP is relevant and clearly applicable in an AI project. The evaluation of the SML-DPM revealed four global, pattern-agnostic insights for AI projects: 1) The relationship between today's application of design patterns and increases in revenue, 2) Environmental sustainability in ML implies cost reductions, 3) Context-dependency and its focal points for sustainability, and 4) The bigger, the better does not hold true for sustainability in ML development.

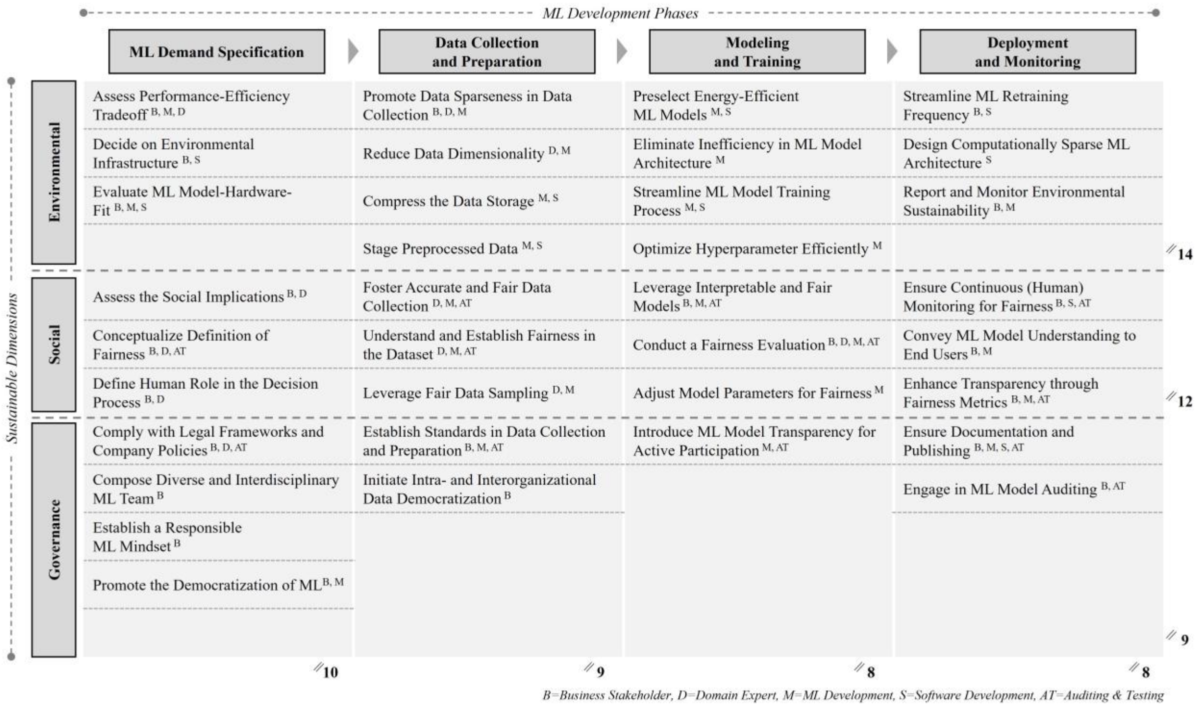

Figure 9. The Sustainable Machine Learning Design Pattern Matrix (SML-DPM)

The design, content, and evaluation of the SML-DPM and the associated DPs led to three main contributions. First, the SML-DPM bridges the gap between the ESG sustainability concept and the end-to-end ML development process. Second, the 35 DP with justificatory knowledge from expert insights are provided, which enables to increase the sustainability of the AI project phases within AI projects. Third, the work contributes by providing extensive naturalistic insights into the SML-DPM's application based on its web-based prototype.

The research has two primary theoretical implications. First, the work has opened a new discussion on how to structure SAI and, subsequently, what SAI comprises regarding clear and implementable practices. It specifically investigated the relationship between the end-to-end AI project workflow and the three sustainability dimensions of environmental, social, and governance. Thus, the results shed light on the end-to-end process view of ML by opening a discussion about the different AI project phases and the unique sustainability challenges faced in each of these (Papagiannidis et al., 2023). Second, by presenting the 35 DPs and validating them with subject matter experts, the article have responded to calls for research into merging hitherto fragmented theoretical knowledge and validating it with practitioner views, facilitating theorizing toward sustainable AI (Veit & Thatcher, 2023; Verdecchia et al., 2023). From a practical perspective, the SML-DPM holds two primary implications. First, the different stakeholders can lever the SML-DPM to capture the status quo and to develop a vision regarding the sustainability of the AI project workflow. Second, the SML-DPM guides the different stakeholders in implementing DPs for the sustainable development of ML. The DPs act as a simple point of entry, as they are easy to understand. Thus, ML development stakeholders can use the SML-DPM to identify DPs that fit their role (e.g., business stakeholder), the current project phase (e.g., Modeling and Training), and the sustainability focus (e.g., environmental).

In sum, to conclude section II.2, the three research articles #4, #5, and #6 contribute to overcoming the challenges AI projects pose at a project level. Research article #4 develops a data-efficient active learning architecture for anomaly detection in industrial time series data to tackle the data challenges. Research article #5 compares and evaluates three frequently used transparent AI models and four different XAI methods to overcome the technical black-box challenge. Lastly, research article #6 introduces the SML-DPM to overcome the sustainability challenges in AI projects.

III Summary and Limitations

1 Summary

Due to, on the one hand, recent advances in the field of AI, such as the accessibility of large-scale data sources as well as the improvement of AI algorithms, and, on the other hand, growing global competitive pressure and changing customer needs, the manufacturing sector has embarked on a journey to lever the potential business value of AI (Berente et al., 2021; S. W. Kim et al., 2022b; Merhi, 2023). Many manufacturers enhance or plan to enhance their operational efficiency and optimize existing processes, for example, by leveraging historical error messages to provide automatic maintenance suggestions for new machine faults (Bertolini et al., 2021; van Giffen & Ludwig, 2023). Furthermore, AI enables them to adapt or even offer novel value propositions, for example, replacing hardware-based measurement techniques with software-based AI control solutions that are less susceptible to maintenance (Favoretto et al., 2022; Ritter et al., 2023; Stahl et al., 2023).

In response to these promising opportunities, manufacturers have initiated AI projects and thus unlock AI´s business value (Shollo et al., 2022; Vial et al., 2023). Nevertheless, AI projects challenge established knowledge due to AI specifics such as iterative learning, the dependency on data, the interdependencies of the AI project phases, or unclear possibilities due to the current AI hype. Consequently, many AI projects often get stuck in an experimental pilot phase without transitioning to productive systems (Benbya et al., 2021; Merhi, 2023). This finding goes along with studies indicating that about 85% of AI projects have little to no impact (Shollo et al., 2022; Vial et al., 2023). Thus, implementing AI poses manufacturers with several challenges that arise on an overarching *organization level* as well as on an individual AI *project level* (Jöhnk et al., 2021; Vial et al., 2023; Weber et al., 2023). Those challenges encompass *business challenges* (e.g., developing an organization-wide AI strategy), *technical challenges* (e.g., ensuring explainability of AI algorithms), *data challenges* (e.g., providing enough labeled data), and *sustainability challenges* (e.g., promoting responsible AI use) (Dennehy et al., 2023; Enholm et al., 2022; Merhi, 2023; Papagiannidis et al., 2023; Weber et al., 2023; Westenberger et al., 2022). Motivated by these challenges, this doctoral thesis seeks to provide solutions to conduct successful AI projects in the manufacturing sector. To structure the results of the six included research articles, the four-phased AI project workflow is used, and each research article focuses on one or more of the four challenges, either at an *organizational level* or a *project level*.

Concerning the *challenges AI projects reveal at an overarching organizational level*, Section II.1 provides an initial entry point to the topic of this cumulative dissertation by presenting an overview of 24 organization-wide success factors for AI projects, structured along four success dimensions, i.e., datability, desirability, feasibility, and viability, and specified by 93 subordinated success manifestations. The SFs are further situated in a broader context by mapping them to the four phases of the AI project workflow, i.e., “Demand Specification”, “Data Collection and Preparation”, “Modeling and Training,” and “Deployment and Monitoring”, and four recommendations to capture a hands-on perspective on the requirements for successful AI projects are derived (research article #1). Thereafter, this thesis provides a maturity model for the capabilities necessary to identify and implement suitable data-driven and AI-based business model archetypes, especially relevant at the beginning of AI projects, i.e., the “Demand Specification” phase. The maturity model builds on Hunke et al. (2022) archetypal data-driven business models (i.e., data provider, insight provider, recommendation provider, and digital solution provider) and is structured along the entire organizational architecture using the five-layered enterprise architecture model of Urbach and Röglinger (2019) (research article #2). As introducing AI and data-driven business models increases the organization's attack surface for cyberattacks such as data breaches and security incidents cannot always be prevented, Section II.1 concludes by presenting the IRM3. The IRM3 is an IRM maturity model closely aligned with practice expectations under a socio-technical perspective and consists of four focus areas (i.e., organization, human, tools, and processes) and 29 capability dimensions. The IRM3 takes a comprehensive view on the entire organization to answer the research question of which capabilities organizations require to approach IRM. The maturity model is applied to seven different organizations to investigate their status quo and target state of IRM capabilities (research article #3).

Regarding the *challenges at the specific project level*, Section II.2 provides three in-depth investigations to overcome AI projects' challenges. First, a data-efficient active learning architecture for anomaly detection (AIMAN) in industrial time series data and its instantiation for a real-world robotic screwdriving application representing a CPS is proposed. Overcoming the time-consuming anomaly data labeling challenge in practice, particularly relevant in the AI project phases of “Data Collection and Preparation” as well as “Modeling and Training”, the approach optimizes an unsupervised model based on an autoencoder with budgeted expert feedback using four different strategies for querying the unlabeled data. The results demonstrate that the active learning approach outperforms a state-of-the-art unsupervised model by 59% in

F1-Score (research article #4). Second, as such advanced AI architectures increase the technical black box problem (Barredo Arrieta et al., 2020), three different transparent AI models (Linear Regression, Decision Tree, QLattice) and four different XAI methods (Partial Dependency Plots, Accumulated Local Effects, Local Interpretable Model-Agnostic Explanations, Shapley Additive Explanations) are compared based on a real-world dataset to overcome the lack of explainability. Their human-centered explainability is evaluated by conducting an extensive online survey with 144 participants in the domain of building energy prediction. The results quantify the explainability and accuracy in predicting building energy consumption and encourage using XAI methods as the right choice of the XAI method enables an increase in the measured explainability by 10% compared to the poorer performing transparent AI models and the other XAI methods (research article #5). Third, to conclude section II.2, the SML-DPM is presented to overcome the sustainability challenges, i.e., environmental, social, and governance, throughout all AI project phases. Hence, the SML-DPM embraces 35 DPs for increased sustainability in the AI project workflow. The SML-DPM was developed following the DSR paradigm in close alignment with four literature-grounded key requirements (Hevner et al., 2004; Peffers et al., 2007) and evaluated with industry experts and in three AI project case studies based on a web-based prototype. The SML-DPM serves as a diagnostic tool for different AI project stakeholders to capture the sustainability status quo and develop a vision regarding the sustainability of their AI project workflow in their current and future AI projects (research article #6).

2 Limitations and Future Research

The results of this doctoral thesis need to be reflected against limitations that provide an impetus for future research. While the six research articles contain a detailed perspective on the limitations of this research endeavor (see Appendix V.3 to V.8), this section provides an aggregated overview of the limitations. Thus, the following presents three overarching limitations and avenues for future research to overcome challenges for successful AI projects in the manufacturing industry.

First, the research results in this thesis build on existing knowledge to contribute novel artifacts for research and practice. Thus, the results were established inductively based on qualitative interview studies and multivocal literature reviews. Drawing on the research methodologies of the DSR paradigm (Hevner et al., 2004; Peffers et al., 2007) and the CRISP-DM (Wirth & Hipp, 2000), the underlying problem statements, design objectives, and the applicability, completeness, and consistency of the artifacts are assessed through evaluations with both practitioners and academics, for example, based on AI projects, AI expert insights, or real-world data sets. Accordingly, some limitations are inherent in the nature of these methodological approaches despite their rigor. In this vein, in research article #1, the validation and refinement of the SFs were conducted through an interview study with 20 subject matter experts. While expert interviews provide the opportunity to explore an emerging phenomenon in depth, the perspectives of individual participants are subjective (Etikan, 2016). Thus, future research should use a confirmatory study (e.g., Delphi study) to substantiate the findings. Similarly, while the maturity models regarding DDBM and IRM of research articles #3 and #4 are developed and evaluated based on data from multiple organizations, researchers could study a larger sample of organizations to challenge the findings' consistency. Regarding the resulting technical architecture in research article #4, the active learning approach was evaluated on a single data set, limiting the generalization. The same as for the data set holds true for the conducted data split into an energy and a mechanical data set to evaluate the data efficiency. Thus, future research could extend the approach to other manufacturing applications. To do so, the paper provides the code as an open-source GitHub repository. Last, regarding research article #5, the evaluation reduced the explainability to the human-centered explainability with the two dimensions of explanation satisfaction and perceived fidelity. Future research could include the model's inherent complexity and technical factors, such as the number of variables in the form of a multi-dimensional study. Nevertheless, the practical-oriented approaches and real-world data included can be considered a strength of this doctoral thesis, and the selected research methods are supposed to serve as blueprints to address the described limitations in future research.

Second, the developed maturity models of research articles #2 and #3, as well as the technical artifacts of research articles #4 and #5, either use simplifying assumptions or consider input parameters as deterministic or expected values. Research article #2 builds on established archetypes of data-driven business models (Hunke et al., 2022) and, therefore, uses a primarily deductive approach (Bhattacherjee, 2012). While established as structural frameworks, these archetypes come with some limitations. The consideration of different archetypal business models could fall short in accounting for their suitability within the company's specific context, such as its size or regional focus. This presents an opportunity for future research to scrutinize the qualitative appropriateness of specific archetypes within a company's context. Moreover, archetypal business models could oversimplify the reality, limiting the potential for customization while facilitating the identification of a target business model, thus opening avenues for future research. Research article #4 assumes that the expert system always correctly labels the data, which may not be true in practice. Experts working under time and quality pressures in a heavily efficiency-driven production environment may make human errors, distorting the data labels. Given this, future research can build on Zhu and Yang (2019), who developed a concept that distinguishes between expert systems of different levels of reliability. Lastly, as research article #5 focuses on data-based research, the work is limited by the dataset used and the model optimization conducted. For instance, the dataset is missing information about the insulation of specific components of the buildings and occupant behavior influencing energy consumption. Further, other approaches exist to optimize each AI model, such as choosing a different cross-validation split. Future studies could address both aspects by collecting the necessary data and enhancing model optimization before XAI analysis. Nevertheless, it should be mentioned that assumptions must be made in all methodological procedures and that the assumptions made in this work have been either validated theoretically or practically.

Third, this thesis takes an organizational level perspective (Section II.1) and a project level perspective (Section II.2) to overcome the four challenges, i.e., business, technical, data, and sustainability challenges, AI projects entail. This structure has been stimulated by previous research and confirmed by the results of this thesis. As such, it complements overarching organizational management contributions (e.g., on organizational readiness for AI, AI adoption, and AI capabilities development) and existing AI project case studies (e.g., insights from real-world AI projects and AI project workflows). However, the results presented are expected to, on the one hand, overlap with findings from these related research streams and, on the other hand, overlap with themselves as there are dependencies from the organizational level to the project level and vice versa. Accordingly, future research is encouraged to further advance knowledge synthesis between the different research angles to conduct successful AI projects. Finally, given the relation of AI to ML and data science, this dissertation does not assert sole ownership of the insights presented herein over AI. Instead, it seeks to inspire cross-disciplinary learning and to derive specific interpretations in the context of AI and AI projects.

In sum, this work contributes to the existing knowledge of AI in the manufacturing sector by presenting artifacts and approaches that help tackle the business, technical, data, and sustainability challenges of AI projects. Therefore, notwithstanding the above limitations, I hope this doctoral thesis will support researchers and practitioners in navigating the opportunities and challenges of AI in the manufacturing industry.

IV Publication Bibliography

Ab Rahman, N. H., & Choo, K.-K. R. (2015). A survey of information security incident handling in the cloud. *Computers & Security*, *49*, 45–69. https://doi.org/10.1016/j.cose.2014.11.006

Abrell, T., Pihlajamaa, M., Kanto, L., Vom Brocke, J., & Uebernickel, F. (2016). The role of users and customers in digital innovation: Insights from B2B manufacturing firms. *Information & Management*, *53*(3), 324–335. https://doi.org/10.1016/j.im.2015.12.005

Adisorn, T., Tholen, L., Thema, J., Luetkehaus, H., Braungardt, S., Huenecke, K., & Schumacher, K. (2020). Towards a More Realistic Cost–Benefit Analysis—Attempting to Integrate Transaction Costs and Energy Efficiency Services. *Energies*, *14*(1), 152. https://doi.org/10.3390/en14010152

Ågerfalk, P. J. (2020). Artificial intelligence as digital agency. *European Journal of Information Systems*, *29*(1), 1–8. https://doi.org/10.1080/0960085X.2020.1721947

Aggarwal, C. C. (2017). *Outlier Analysis*. Springer International Publishing. https://doi.org/10.1007/978-3-319-47578-3

Ahlrichs, J., & Rockstuhl, S. (2022). Estimating fair rent increases after building retrofits: A max-min fairness approach. *Energy Policy*, *164*, 112923. https://doi.org/10.1016/j.enpol.2022.112923

Ahlrichs, J., Wenninger, S., Wiethe, C., & Häckel, B. (2022). Impact of socio-economic factors on local energetic retrofitting needs—A data analytics approach. *Energy Policy*, *160*, 112646. https://doi.org/10.1016/j.enpol.2021.112646

Ahmad, A., Hadgkiss, J., & Ruighaver, A. B. (2012). Incident response teams – Challenges in supporting the organisational security function. *Computers & Security*, *31*(5), 643–652. https://doi.org/10.1016/j.cose.2012.04.001

Ahmad, A., Maynard, S. B., Desouza, K. C., Kotsias, J., Whitty, M. T., & Baskerville, R. L. (2021). How can organizations develop situation awareness for incident response: A case study of management practice. *Computers & Security*, *101*, 102122. https://doi.org/10.1016/j.cose.2020.102122

Ahmad, A., Maynard, S., & Baskerville, R. (2022). Editorial. *Computers & Security*, *112*, 102530. https://doi.org/10.1016/j.cose.2021.102530

Alaei, P., & Noorbehbahani, F. (2017). Incremental anomaly-based intrusion detection system using limited labeled data. *2017 3th International Conference on Web Research (ICWR)*, 178–184. https://doi.org/10.1109/ICWR.2017.7959324

Ali, S., Abuhmed, T., El-Sappagh, S., Muhammad, K., Alonso-Moral, J. M., Confalonieri, R., Guidotti, R., Del Ser, J., Díaz-Rodríguez, N., & Herrera, F. (2023). Explainable Artificial Intelligence (XAI): What we know and what is left to attain Trustworthy Artificial Intelligence. *Information Fusion*, *99*, 101805. https://doi.org/10.1016/j.inffus.2023.101805

Allen, J., Freed, A., & Chandrasekaran, S. (2017). *Adapt DevOps to cognitive and artificial intelligence systems*. https://developer.ibm.com/articles/cc-devops-artificial-intelligence-cognitive/

Alsheibani, S. A., Cheung, Y., & Messom, C. (2018). Artificial Intelligence Adoption: AI-Readiness at Firm-Level. *PACIS 2018 Proceedings*, *37*, 1–9. https://core.ac.uk/download/pdf/301376079.pdf

Alsheibani, S. A., Cheung, Y., Messom, C., & Alhosni, M. (2020). Winning AI Strategy: Six-Steps to Create Value from Artificial Intelligence. In *AMCIS 2020 Proceedings* (Vol. 1). https://aisel.aisnet.org/amcis2020/adv_info_systems_research/adv_info_systems_research/1

Alsheibani, S. A., Cheung, Y. P., & Messom, C. H. (2019). Factors Inhibiting the Adoption of Artificial Intelligence at Organizational-Level: A Preliminary Investigation. *Americas Conference on Information Systems*.

Amasyali, K., & El-Gohary, N. M. (2018). A review of data-driven building energy consumption prediction studies. *Renewable and Sustainable Energy Reviews*, *81*, 1192–1205. https://doi.org/10.1016/j.rser.2017.04.095

Amershi, S., Begel, A., Bird, C., DeLine, R., Gall, H., Kamar, E., Nagappan, N., Nushi, B., & Zimmermann, T. (2019). Software Engineering for Machine Learning: A Case Study. *2019 IEEE/ACM 41st International Conference on Software Engineering: Software Engineering in Practice (ICSE-SEIP)*, 291–300. https://doi.org/10.1109/ICSE-SEIP.2019.00042

Amodei, D., & Hernandez, D. (2018). *AI and compute*. OpenAI. https://openai.com/research/ai-and-compute

Ardolino, M., Rapaccini, M., Saccani, N., Gaiardelli, P., Crespi, G., & Ruggeri, C. (2018). The role of digital technologies for the service transformation of industrial companies. *International Journal of Production Research*, *56*(6), 2116–2132. https://doi.org/10.1080/00207543.2017.1324224

Baccarini, D. (1999). The Logical Framework Method for Defining Project Success. *Project Management Journal*, *30*(4), 25–32. https://doi.org/10.1177/875697289903000405

Baier, L., Jöhren, F., & Seebacher, S. (2019). *Challenges in the Deployment and Operation of Machine Learning in Practice*. https://doi.org/10.5445/IR/1000095028

Bain & Company. (2024). *Global Machinery & Equipment Report 2024*. https://www.bain.com/globalassets/noindex/2024/bain_report_machinery_and_equipment_report_2024.pdf

Baker, J. (2012). The Technology–Organization–Environment Framework. In Y. K. Dwivedi, M. R. Wade, & S. L. Schneberger (Eds.), *Information Systems Theory* (Vol. 28, pp. 231–245). Springer New York. https://doi.org/10.1007/978-1-4419-6108-2_12

Baltuttis, D., Häckel, B., Jonas, C. M., Oberländer, A. M., Röglinger, M., & Seyfried, J. (2022). Conceptualizing and Assessing the Value of Internet of Things Solutions. *Journal of Business Research*, *140*, 245–263. https://doi.org/10.1016/j.jbusres.2021.10.063

Barbado, A., Corcho, Ó., & Benjamins, R. (2022). Rule extraction in unsupervised anomaly detection for model explainability: Application to OneClass SVM. *Expert Systems with Applications*, *189*, 116100. https://doi.org/10.1016/j.eswa.2021.116100

Barredo Arrieta, A., Díaz-Rodríguez, N., Del Ser, J., Bennetot, A., Tabik, S., Barbado, A., Garcia, S., Gil-Lopez, S., Molina, D., Benjamins, R., Chatila, R., & Herrera, F. (2020). Explainable Artificial Intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI. *Information Fusion, 58,* 82–115. https://doi.org/10.1016/j.inffus.2019.12.012

Barreuther, P., Wanke, K., Cichon, F., Klüh, S., & Halusa, S. (2022). *Cyber-Security-Risk-Report 2021*. MHP Management- und IT Beratung GmbH.

Becker, J., Knackstedt, R., & Pöppelbuß, J. (2009). Developing Maturity Models for IT Management: A Procedure Model and its Application. *Business & Information Systems Engineering*, *1*(3), 213–222. https://doi.org/10.1007/s12599-009-0044-5

Benbya, H., Deakin University, Melbourne, Australia, Pachidi, S., Cambridge JudgeBusiness School, University of Cambridge, United Kingdom, Jarvenpaa, S. L., & McCombs School of Business, University of Texas at Austin, U.S.A. (2021). Special Issue Editorial: Artificial Intelligence in Organizations: Implications for Information Systems Research. *Journal of the Association for Information Systems*, *22*(2), 281–303. https://doi.org/10.17705/1jais.00662

Berente, N., Gu, B., Recker, J., & Santhanam, R. (2021). Managing Artificial Intelligence. *MIS Quarterly*, *45*, 1433–1450.

Bertolini, M., Mezzogori, D., Neroni, M., & Zammori, F. (2021). Machine Learning for industrial applications: A comprehensive literature review. *Expert Systems with Applications*, *175*, 114820. https://doi.org/10.1016/j.eswa.2021.114820

Bettis, R. A., Gambardella, A., Helfat, C., & Mitchell, W. (2015). Qualitative empirical research in strategic management. *Strategic Management Journal*, *36*(5), 637–639. https://doi.org/10.1002/smj.2317

Beverungen, D., Kundisch, D., & Wünderlich, N. (2021). Transforming into a platform provider: Strategic options for industrial smart service providers. *Journal of Service Management*, *32*(4), 507–532. https://doi.org/10.1108/JOSM-03-2020-0066

Bhattacherjee, A. (2012). *Social science research: Principles, methods, and practices* (Second edition). Anol Bhattacherjee.

Bitzer, M., Häckel, B., Leuthe, D., Ott, J., Stahl, B., & Strobel, J. (2023). Managing the Inevitable – A Maturity Model to Establish Incident Response Management Capabilities. *Computers & Security*, *125*, 103050. https://doi.org/10.1016/j.cose.2022.103050

Björkdahl, J. (2020). Strategies for Digitalization in Manufacturing Firms. *California Management Review*, *62*(4), 17–36. https://doi.org/10.1177/0008125620920349

Böttcher, T., Weber, M., Weking, J., Hein, A., & Krcmar, H. (2022). Value Drivers of Artificial Intelligence. In *AMCIS 2022 Proceedings* (Vol. 5).

Brasse, J., Broder, H. R., Förster, M., Klier, M., & Sigler, I. (2023). Explainable artificial intelligence in information systems: A review of the status quo and future research directions. *Electronic Markets*, *33*(1), 26. https://doi.org/10.1007/s12525-023-00644-5

Brock, J. K.-U., & Von Wangenheim, F. (2019). Demystifying AI: What Digital Transformation Leaders Can Teach You about Realistic Artificial Intelligence. *California Management Review*, *61*(4), 110–134. https://doi.org/10.1177/1536504219865226

Bullen, C. V., & Rockart, J. F. (1981). A primer on critical success factors. *Center for Information Systems Research, Massachusetts Institute of Technology*, *69*, 1–64.

Burkart, N., & Huber, M. F. (2021). A Survey on the Explainability of Supervised Machine Learning. *Journal of Artificial Intelligence Research*, *70*, 245–317. https://doi.org/10.1613/jair.1.12228

Burström, T., Parida, V., Lahti, T., & Wincent, J. (2021). AI-enabled business-model innovation and transformation in industrial ecosystems: A framework, model and outline for further research. *Journal of Business Research*, *127*, 85–95. https://doi.org/10.1016/j.jbusres.2021.01.016

Chatterjee, S., Rana, N. P., Dwivedi, Y. K., & Baabdullah, A. M. (2021). Understanding AI adoption in manufacturing and production firms using an integrated TAM-TOE model. *Technological Forecasting and Social Change*, *170*, 120880. https://doi.org/10.1016/j.techfore.2021.120880

Chevrot, A., Vernotte, A., & Legeard, B. (2022). CAE: Contextual auto-encoder for multivariate time-series anomaly detection in air transportation. *Computers & Security*, *116*, 102652. https://doi.org/10.1016/j.cose.2022.102652

Cichonski, P., Millar, T., Grance, T., & Scarfone, K. (2012). *Computer Security Incident Handling Guide: Recommendations of the National Institute of Standards and Technology* (NIST SP 800-61r2; p. NIST SP 800-61r2). National Institute of Standards and Technology. https://doi.org/10.6028/NIST.SP.800-61r2

Collins, C., Dennehy, D., Conboy, K., & Mikalef, P. (2021). Artificial intelligence in information systems research: A systematic literature review and research agenda. *International Journal of Information Management*, *60*, 102383. https://doi.org/10.1016/j.ijinfomgt.2021.102383

Cooper, R. B., & Zmud, R. W. (1990). Information Technology Implementation Research: A Technological Diffusion Approach. *Management Science*, *36*(2), 123–139. https://doi.org/10.1287/mnsc.36.2.123

Corbin, J. M., & Strauss, A. (1990). Grounded theory research: Procedures, canons, and evaluative criteria. *Qualitative Sociology*, *13*(1), 3–21. https://doi.org/10.1007/BF00988593

Cowls, J., Tsamados, A., Taddeo, M., & Floridi, L. (2023). The AI gambit: Leveraging artificial intelligence to combat climate change—opportunities, challenges, and recommendations. *AI & SOCIETY*, *38*(1), 283–307. https://doi.org/10.1007/s00146-021-01294-x

Cui, Y., Kara, S., & Chan, K. C. (2020). Manufacturing big data ecosystem: A systematic literature review. *Robotics and Computer-Integrated Manufacturing*, *62*, 101861. https://doi.org/10.1016/j.rcim.2019.101861

Cyberedge Group. (2021). *2021 Cyberthreat Defense Report*. Cyberedge Group.

Dadhich, M., & Hiran, K. K. (2022). Empirical investigation of extended TOE model on Corporate Environment Sustainability and dimensions of operating performance of SMEs: A high order PLS-ANN approach. *Journal of Cleaner Production*, *363*, 132309. https://doi.org/10.1016/j.jclepro.2022.132309

Das, S., Wong, W.-K., Dietterich, T., Fern, A., & Emmott, A. (2016). Incorporating Expert Feedback into Active Anomaly Discovery. *2016 IEEE 16th International Conference on Data Mining (ICDM)*, 853–858. https://doi.org/10.1109/ICDM.2016.0102

De Haes, S., Van Grembergen, W., & Debreceny, R. S. (2013). COBIT 5 and Enterprise Governance of Information Technology: Building Blocks and Research Opportunities. *Journal of Information Systems*, *27*(1), 307–324. https://doi.org/10.2308/isys-50422

De Silva, D., & Alahakoon, D. (2022). An artificial intelligence life cycle: From conception to production. *Patterns*, *3*(6), 100489. https://doi.org/10.1016/j.patter.2022.100489

Debus, C., Piraud, M., Streit, A., Theis, F., & Götz, M. (2023). Reporting electricity consumption is essential for sustainable AI. *Nature Machine Intelligence*, *5*(11), 1176–1178. https://doi.org/10.1038/s42256-023-00750-1

Dennehy, D., Griva, A., Pouloudi, N., Dwivedi, Y. K., Mäntymäki, M., & Pappas, I. O. (2023). Artificial Intelligence (AI) and Information Systems: Perspectives to Responsible AI. *Information Systems Frontiers*, *25*(1), 1–7. https://doi.org/10.1007/s10796-022-10365-3

Ding, W., Abdel-Basset, M., Hawash, H., & Ali, A. M. (2022a). Explainability of artificial intelligence methods, applications and challenges: A comprehensive survey. *Information Sciences*, *615*, 238–292. https://doi.org/10.1016/j.ins.2022.10.013

Ding, W., Abdel-Basset, M., Hawash, H., & Ali, A. M. (2022b). Explainability of artificial intelligence methods, applications and challenges: A comprehensive survey. *Information Sciences*, *615*, 238–292. https://doi.org/10.1016/j.ins.2022.10.013

Drempetic, S., Klein, C., & Zwergel, B. (2020). The Influence of Firm Size on the ESG Score: Corporate Sustainability Ratings Under Review. *Journal of Business Ethics*, *167*(2), 333–360. https://doi.org/10.1007/s10551-019-04164-1

Duan, Y., Edwards, J. S., & Dwivedi, Y. K. (2019). Artificial intelligence for decision making in the era of Big Data – evolution, challenges and research agenda. *International Journal of Information Management*, *48*, 63–71. https://doi.org/10.1016/j.ijinfomgt.2019.01.021

Dwivedi, Y. K., Hughes, L., Ismagilova, E., Aarts, G., Coombs, C., Crick, T., Duan, Y., Dwivedi, R., Edwards, J., Eirug, A., Galanos, V., Ilavarasan, P. V., Janssen, M., Jones, P., Kar, A. K., Kizgin, H., Kronemann, B., Lal, B., Lucini, B., … Williams, M. D. (2021). Artificial Intelligence (AI): Multidisciplinary perspectives on emerging challenges, opportunities, and agenda for research, practice and policy. *International Journal of Information Management*, *57*, 101994. https://doi.org/10.1016/j.ijinfomgt.2019.08.002

Dwivedi, Y. K., Wastell, D., Laumer, S., Henriksen, H. Z., Myers, M. D., Bunker, D., Elbanna, A., Ravishankar, M. N., & Srivastava, S. C. (2015). Research on information systems failures and successes: Status update and future directions. *Information Systems Frontiers*, *17*(1), 143–157. https://doi.org/10.1007/s10796-014-9500-y

Eitle, V., & Buxmann, P. (2020). Cultural Differences in Machine Learning Adoption: An International Comparison between Germany and the United States. *Proceedings of the 28th European Conference on Information Systems (ECIS)*. https://aisel.aisnet.org/ecis2020_rp/138/

Eitle, V., Zöll, A., & Buxmann, P. (2022). *Organizational Readiness Concept for AI: A Quantitative Analysis of a Multi-stage Adoption Process from the Perspective of Data Scientists*. *Proceedings of the 30th European Conference on Information Systems (ECIS)*.

Enholm, I. M., Papagiannidis, E., Mikalef, P., & Krogstie, J. (2022). Artificial Intelligence and Business Value: A Literature Review. *Information Systems Frontiers*, *24*(5), 1709–1734. https://doi.org/10.1007/s10796-021-10186-w

Etikan, I. (2016). Comparison of Convenience Sampling and Purposive Sampling. *American Journal of Theoretical and Applied Statistics*, *5*(1), 1. https://doi.org/10.11648/j.ajtas.20160501.11

Farahmand, F., Navathe, S. B., Enslow, P. H., & Sharp, G. P. (2003). Managing vulnerabilities of information systems to security incidents. *Proceedings of the 5th International Conference on Electronic Commerce - ICEC '03*, 348–354. https://doi.org/10.1145/948005.948050

Favoretto, C., Mendes, G. H. D. S., Filho, M. G., Gouvea De Oliveira, M., & Ganga, G. M. D. (2022). Digital transformation of business model in manufacturing companies: Challenges and research agenda. *Journal of Business & Industrial Marketing*, *37*(4), 748–767. https://doi.org/10.1108/JBIM-10-2020-0477

Fayyad, U., Haussler, D., & Stolorz, P. (1996). KDD for science data analysis: Issues and examples. *Proceedings of the Second International Conference on Knowledge Discovery and Data Mining*.

Feng, C., & Tian, P. (2021). Time Series Anomaly Detection for Cyber-physical Systems via Neural System Identification and Bayesian Filtering. *Proceedings of the 27th ACM SIGKDD Conference on Knowledge Discovery & Data Mining*, 2858–2867. https://doi.org/10.1145/3447548.3467137

Ferrara, E. (2023). *Fairness And Bias in Artificial Intelligence: A Brief Survey of Sources, Impacts, And Mitigation Strategies* (arXiv:2304.07683). arXiv.

Finder, I., Sheetrit, E., & Nissim, N. (2022). A time-interval-based active learning framework for enhanced PE malware acquisition and detection. *Computers & Security*, *121*, 102838. https://doi.org/10.1016/j.cose.2022.102838

Gamma, E. (Ed.). (1995). *Design patterns: Elements of reusable object-oriented software*. Addison-Wesley.

Garousi, V., Felderer, M., & Mäntylä, M. V. (2019). Guidelines for including grey literature and conducting multivocal literature reviews in software engineering. *Information and Software Technology*, *106*, 101–121. https://doi.org/10.1016/j.infsof.2018.09.006

Gartner. (2024). *Invest Implications: Forecast Analysis: Artificial Intelligence Software, 2023-2027, Worldwide*. https://www.gartner.com/en/documents/4925331

Gebauer, H., Paiola, M., Saccani, N., & Rapaccini, M. (2021). Digital servitization: Crossing the perspectives of digitization and servitization. *Industrial Marketing Management*, *93*, 382–388. https://doi.org/10.1016/j.indmarman.2020.05.011

Goel, K., Fehrer, T., Röglinger, M., & Wynn, M. T. (2023). Not Here, But There: Human Resource Allocation Patterns. In C. Di Francescomarino, A. Burattin, C. Janiesch, & S. Sadiq (Eds.), *Business Process Management* (Vol. 14159, pp. 377–394). Springer Nature Switzerland. https://doi.org/10.1007/978-3-031-41620-0_22

Goldkuhl, G. (2012). Pragmatism vs interpretivism in qualitative information systems research. *European Journal of Information Systems*, *21*(2), 135–146. https://doi.org/10.1057/ejis.2011.54

Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep learning*. The MIT press.

Gregor, S., Kruse, L., & Seidel, S. (2020). Research Perspectives: The Anatomy of a Design Principle. *Journal of the Association for Information Systems*, *21*, 1622–1652. https://doi.org/10.17705/1jais.00649

Häckel, B., Hänsch, F., Hertel, M., & Übelhör, J. (2019). Assessing IT availability risks in smart factory networks. *Business Research*, *12*(2), 523–558. https://doi.org/10.1007/s40685-018-0071-5

Häckel, B., Karnebogen, P., & Ritter, C. (2022). AI-based industrial full-service offerings: A model for payment structure selection considering predictive power. *Decision Support Systems*, *152*, 113653. https://doi.org/10.1016/j.dss.2021.113653

Haenlein, M., & Kaplan, A. (2019). A Brief History of Artificial Intelligence: On the Past, Present, and Future of Artificial Intelligence. *California Management Review*, *61*(4), 5–14. https://doi.org/10.1177/0008125619864925

Hartmann, P. M., Zaki, M., Feldmann, N., & Neely, A. (2016). Capturing value from big data – a taxonomy of data-driven business models used by start-up firms. *International Journal of Operations & Production Management*, *36*(10), 1382–1406. https://doi.org/10.1108/IJOPM-02-2014-0098

Heinz, D., Benz, C., Bode, J., Hunke, F., & Satzger, G. (2022). Exploring the Potential of Smart Service Systems: A Multi-Actor View on Affordances and Their Actualization. *Journal of Service Management Research*, *6*(2), 132–146. https://doi.org/10.5771/2511-8676-2022-2-132

Hevner, March, Park, & Ram. (2004). Design Science in Information Systems Research. *MIS Quarterly*, *28*(1), 75. https://doi.org/10.2307/25148625

Himeur, Y., Ghanem, K., Alsalemi, A., Bensaali, F., & Amira, A. (2021). Artificial intelligence based anomaly detection of energy consumption in buildings: A review, current trends and new perspectives. *Applied Energy*, *287*, 116601. https://doi.org/10.1016/j.apenergy.2021.116601

Hoffman, R. R., Mueller, S. T., Klein, G., & Litman, J. (2018). Metrics for Explainable AI: Challenges and Prospects. *Technical Report, DARPA Explainable AI Program*. https://doi.org/10.48550/arXiv.1812.04608

Holotiuk, F., Beimborn, D., & Hund, A. (2024). Mechanisms for Achieving Ambidexterity in the Context of Digital Transformation: Insights from Digital Innovation Labs. *Journal of the Association for Information Systems, 136*. https://doi.org/10.17705/1jais.00885

Hukkelberg, I., & Rolland, K. (2020). Exploring Machine Learning in a Large Governmental Organization: An Information Infrastructure Perspective. *Proceedings of the 28th European Conference on Information Systems (ECIS).*

Hunke, F., Heinz, D., & Satzger, G. (2022). Creating customer value from data: Foundations and archetypes of analytics-based services. *Electronic Markets*, *32*(2), 503–521. https://doi.org/10.1007/s12525-021-00506-y

Ibarra, D., Ganzarain, J., & Igartua, J. I. (2018). Business model innovation through Industry 4.0: A review. *Procedia Manufacturing*, *22*, 4–10. https://doi.org/10.1016/j.promfg.2018.03.002

IBM. (2022). *Cost of a Data Breach Report 2022*. https://www.ibm.com/downloads/cas/3R8N1DZJ

International Data Corporation. (2023). *Worldwide AI and Generative AI Spending Guide*. https://www.idc.com/getdoc.jsp?containerId=IDC_P33198

Jakubik, J., Vössing, M., Kühl, N., Walk, J., & Satzger, G. (2024). Data-Centric Artificial Intelligence. *Business & Information Systems Engineering*. https://doi.org/10.1007/s12599-024-00857-8

Jo, T. (2021). *Machine Learning Foundations: Supervised, Unsupervised, and Advanced Learning*. Springer International Publishing. https://doi.org/10.1007/978-3-030-65900-4

Jöhnk, J., Weißert, M., & Wyrtki, K. (2021). Ready or Not, AI Comes—An Interview Study of Organizational AI Readiness Factors. *Business & Information Systems Engineering*, *63*(1), 5–20. https://doi.org/10.1007/s12599-020-00676-7

Jordan, M. I., & Mitchell, T. M. (2015). Machine learning: Trends, perspectives, and prospects. *Science*, *349*(6245), 255–260. https://doi.org/10.1126/science.aaa8415

Kappelman, L., L. Johnson, V., Maurer, C., Guerra, K., McLean, E., Torres, R., Snyder, M., & Kim, K. (2020). The 2019 SIM IT Issues and Trends Study. *MIS Quarterly Executive*, *19*(1), 69–104. https://doi.org/10.17705/2msqe.00026

Keding, C. (2021). Understanding the interplay of artificial intelligence and strategic management: Four decades of research in review. *Management Review Quarterly*, *71*(1).

Kim, B., Park, J., & Suh, J. (2020). Transparency and accountability in AI decision support: Explaining and visualizing convolutional neural networks for text information. *Decision Support Systems*, *134*, 113302. https://doi.org/10.1016/j.dss.2020.113302

Kim, G., Choi, J. G., Ku, M., & Lim, S. (2023). Developing a semi-supervised learning and ordinal classification framework for quality level prediction in manufacturing. *Computers & Industrial Engineering*, *181*, 109286. https://doi.org/10.1016/j.cie.2023.109286

Kim, M., Saad, W., Mozaffari, M., & Debbah, M. (2022). On the Tradeoff between Energy, Precision, and Accuracy in Federated Quantized Neural Networks. *ICC 2022 - IEEE International Conference on Communications*, 2194–2199. https://doi.org/10.1109/ICC45855.2022.9838362

Kim, S. W., Kong, J. H., Lee, S. W., & Lee, S. (2022a). Recent Advances of Artificial Intelligence in Manufacturing Industrial Sectors: A Review. *International Journal of Precision Engineering and Manufacturing*, *23*(1), 111–129. https://doi.org/10.1007/s12541-021-00600-3

Kim, S. W., Kong, J. H., Lee, S. W., & Lee, S. (2022b). Recent Advances of Artificial Intelligence in Manufacturing Industrial Sectors: A Review. *International Journal of Precision Engineering and Manufacturing*, *23*(1), 111–129. https://doi.org/10.1007/s12541-021-00600-3

Kindermans, P.-J., Hooker, S., Adebayo, J., Alber, M., Schütt, K. T., Dähne, S., Erhan, D., & Kim, B. (2019). The (Un)reliability of Saliency Methods. In W. Samek, G. Montavon, A. Vedaldi, L. K. Hansen, & K.-R. Müller (Eds.), *Explainable AI: Interpreting, Explaining and Visualizing Deep Learning* (Vol. 11700, pp. 267–280). Springer International Publishing. https://doi.org/10.1007/978-3-030-28954-6_14

Kinkel, S., Baumgartner, M., & Cherubini, E. (2022). Prerequisites for the adoption of AI technologies in manufacturing – Evidence from a worldwide sample of manufacturing companies. *Technovation*, *110*, 102375. https://doi.org/10.1016/j.technovation.2021.102375

Knote, R., Janson, A., Söllner, M., & Leimeister, J. M. (2020). Value Co-Creation in Smart Services: A Functional Affordances Perspective on Smart Personal Assistants. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.3923706

Kolbjørnsrud, V., Amico, R., & Thomas, R. J. (2017). Partnering with AI: How organizations can win over skeptical managers. *Strategy & Leadership*, *45*(1), 37–43. https://doi.org/10.1108/SL-12-2016-0085

Koniakou, V. (2023). From the “rush to ethics” to the “race for governance” in Artificial Intelligence. *Information Systems Frontiers*, *25*(1), 71–102. https://doi.org/10.1007/s10796-022-10300-6

Kreuzberger, D., Kühl, N., & Hirschl, S. (2023). Machine Learning Operations (MLOps): Overview, Definition, and Architecture. *IEEE Access*, *11*, 31866–31879. https://doi.org/10.1109/ACCESS.2023.3262138

Kumar, A., & Krishnamoorthy, B. (2020). Business Analytics Adoption in Firms: A Qualitative Study Elaborating TOE Framework in India. *International Journal of Global Business and Competitiveness*, *15*(2), 80–93. https://doi.org/10.1007/s42943-020-00013-5

Kuypers, M. A., Maillart, T., & Paté-Cornell, E. (2016). *An Empirical Analysis of Cyber Security Incidents at a Large Organization*. Stanford University. https://fsi.stanford.edu/publication/empirical-analysis-cyber-security-incidents-large-organization

Lallie, H. S., Shepherd, L. A., Nurse, J. R. C., Erola, A., Epiphaniou, G., Maple, C., & Bellekens, X. (2021). Cyber security in the age of COVID-19: A timeline and analysis of cyber-crime and cyber-attacks during the pandemic. *Computers & Security*, *105*, 102248. https://doi.org/10.1016/j.cose.2021.102248

Lee, K., Ayyasamy, M. V., Ji, Y., & Balachandran, P. V. (2022). A comparison of explainable artificial intelligence methods in the phase classification of multi-principal element alloys. *Scientific Reports*, *12*(1), 11591. https://doi.org/10.1038/s41598-022-15618-4

Lee, M. C. M., Scheepers, H., Lui, A. K. H., & Ngai, E. W. T. (2023). The implementation of artificial intelligence in organizations: A systematic literature review. *Information & Management*, *60*(5), 103816. https://doi.org/10.1016/j.im.2023.103816

Leporowski, B., Tola, D., Hansen, C., & Iosifidis, A. (2021). *AURSAD: Universal Robot Screwdriving Anomaly Detection Dataset* (arXiv:2102.01409). arXiv. http://arxiv.org/abs/2102.01409

Leuthe, D., Weiß, F., Dersch, J., & Bitzer, M. (2024). Towards Secure Cloud-Computing in FinTechs – An Artefact for Prioritizing Information Security Measures. *Proceedings of the 57th Hawaii International Conference on System Sciences*.

Li, J., Li, M., Wang, X., & Bennett Thatcher, J. (2021). Strategic Directions for AI: The Role of CIOs and Boards of Directors. *MIS Quarterly*, *45*(3), 1603–1644. https://doi.org/10.25300/MISQ/2021/16523

Li, X., Ran, Y., Wan, F., Yu, H., Zhang, G., & He, Y. (2022). Condition-based maintenance strategy optimization of meta-action unit considering imperfect preventive maintenance based on Wiener process. *Flexible Services and Manufacturing Journal*, *34*(1), 204–233. https://doi.org/10.1007/s10696-021-09407-w

Li, Z., Zhao, Y., Han, J., Su, Y., Jiao, R., Wen, X., & Pei, D. (2021). Multivariate Time Series Anomaly Detection and Interpretation using Hierarchical Inter-Metric and Temporal Embedding. In F. Zhu (Ed.), *Proceedings of the 27th ACM SIGKDD Conference on Knowledge Discovery & Data Mining* (pp. 3220–3230). Association for Computing Machinery. https://doi.org/10.1145/3447548.3467075

Lins, S., Pandl, K. D., Teigeler, H., Thiebes, S., Bayer, C., & Sunyaev, A. (2021). Artificial Intelligence as a Service: Classification and Research Directions. *Business & Information Systems Engineering*, *63*(4), 441–456. https://doi.org/10.1007/s12599-021-00708-w

Loebbecke, C., & Picot, A. (2015). Reflections on societal and business model transformation arising from digitization and big data analytics: A research agenda. *The Journal of Strategic Information Systems*, *24*(3), 149–157. https://doi.org/10.1016/j.jsis.2015.08.002

Löfström, H., Hammar, K., & Johansson, U. (2022). A Meta Survey of Quality Evaluation Criteria in Explanation Methods. In J. De Weerdt & A. Polyvyanyy (Eds.), *Intelligent Information Systems* (Vol. 452, pp. 55–63). Springer International Publishing. https://doi.org/10.1007/978-3-031-07481-3_7

Machlev, R., Heistrene, L., Perl, M., Levy, K. Y., Belikov, J., Mannor, S., & Levron, Y. (2022). Explainable Artificial Intelligence (XAI) techniques for energy and power systems: Review, challenges and opportunities. *Energy and AI*, *9*, 100169. https://doi.org/10.1016/j.egyai.2022.100169

Margherita, E. G., & Braccini, A. M. (2023). Industry 4.0 Technologies in Flexible Manufacturing for Sustainable Organizational Value: Reflections from a Multiple Case Study of Italian Manufacturers. *Information Systems Frontiers*, *25*(3), 995–1016. https://doi.org/10.1007/s10796-020-10047-y

Maslej, N., Fattorini, L., Perrault, R., Parli, V., Reuel, A., Brynjolfsson, E., Etchemendy, J., Ligett, K., Lyons, T., Manyika, J., Niebles, J. C., Shoham, Y., Wald, R., & Clark, J. (2024). *The AI Index 2024 Annual Report*. Stanford University.

Mayer, Z., Volk, R., & Schultmann, F. (2022). Analysis of financial benefits for energy retrofits of owner-occupied single-family houses in Germany. *Building and Environment*, *211*, 108722. https://doi.org/10.1016/j.buildenv.2021.108722

McKinsey. (2023). *The state of AI in 2023: Generative AI's breakout year*. https://www.mckinsey.com/capabilities/quantumblack/our-insights/the-state-of-ai-in-2023-generative-ais-breakout-year#/

Merhi, M. I. (2023). An evaluation of the critical success factors impacting artificial intelligence implementation. *International Journal of Information Management*, *69*, 102545. https://doi.org/10.1016/j.ijinfomgt.2022.102545

Microsoft. (2016). *The Emerging Era of Cyber Defense and Cybercrime*. https://www.microsoft.com/security/blog/2016/01/27/the-emerging-era-of-cyber-defense-and-cybercrime/

Mikalef, P., Conboy, K., Lundström, J. E., & Popovič, A. (2022). Thinking responsibly about responsible AI and 'the dark side' of AI. *European Journal of Information Systems, 31*(3), 257–268. https://doi.org/10.1080/0960085X.2022.2026621

Mikalef, P., Fjørtoft, S. O., & Torvatn, H. Y. (2019). Developing an Artificial Intelligence Capability: A Theoretical Framework for Business Value. In W. Abramowicz & R. Corchuelo

(Eds.), *Business Information Systems Workshops* (Vol. 373, pp. 409–416). Springer International Publishing. https://doi.org/10.1007/978-3-030-36691-9_34

Miller, G. (2021). *Artificial Intelligence Project Success Factors: Moral Decision-Making with Algorithms*. 379–390. https://doi.org/10.15439/2021F26

Miller, T. (2017). *Explanation in Artificial Intelligence: Insights from the Social Sciences*. https://doi.org/10.48550/ARXIV.1706.07269

MIT Technology Review. (2024). *Taking AI to the next level in manufacturing*. https://www.technologyreview.com/2024/04/09/1090880/taking-ai-to-the-next-level-in-manufacturing/

Morgan, S. (2021). *Special Report: Cyberwarfare In The C-Suite*. https://cybersecurityventures.com/cybercrime-damage-costs-10-trillion-by-2025/

Müller, J., & Buliga, O. (2019). Archetypes for data-driven business models for manufacturing companies in Industry 4.0. *International Conference on Information Systems 2019 Special Interest Group on Big Data Proceedings*, *2*.

Nyumba, T., Wilson, K., Derrick, C. J., & Mukherjee, N. (2018). The use of focus group discussion methodology: Insights from two decades of application in conservation. *Methods in Ecology and Evolution*, *9*(1), 20–32. https://doi.org/10.1111/2041-210X.12860

Pang, G., Shen, C., Cao, L., & Hengel, A. V. D. (2022). Deep Learning for Anomaly Detection: A Review. *ACM Computing Surveys*, *54*(2), 1–38. https://doi.org/10.1145/3439950

Pang, G., Van Den Hengel, A., Shen, C., & Cao, L. (2021). Toward Deep Supervised Anomaly Detection: Reinforcement Learning from Partially Labeled Anomaly Data. *Proceedings of the 27th ACM SIGKDD Conference on Knowledge Discovery & Data Mining*, 1298–1308. https://doi.org/10.1145/3447548.3467417

Papagiannidis, E., Enholm, I. M., Dremel, C., Mikalef, P., & Krogstie, J. (2023). Toward AI Governance: Identifying Best Practices and Potential Barriers and Outcomes. *Information Systems Frontiers*, *25*(1), 123–141. https://doi.org/10.1007/s10796-022-10251-y

Pappas, I. O., Mikalef, P., Dwivedi, Y. K., Jaccheri, L., & Krogstie, J. (2023). Responsible Digital Transformation for a Sustainable Society. *Information Systems Frontiers*, *25*(3), 945–953. https://doi.org/10.1007/s10796-023-10406-5

Patterson, D., Gonzalez, J., Holzle, U., Le, Q., Liang, C., Munguia, L.-M., Rothchild, D., So, D. R., Texier, M., & Dean, J. (2022). The Carbon Footprint of Machine Learning Training Will Plateau, Then Shrink. *Computer*, *55*(7), 18–28. https://doi.org/10.1109/MC.2022.3148714

Peffers, K., Tuunanen, T., Rothenberger, M. A., & Chatterjee, S. (2007). A Design Science Research Methodology for Information Systems Research. *Journal of Management Information Systems*, *24*(3), 45–77. https://doi.org/10.2753/MIS0742-1222240302

Pham, A.-D., Ngo, N.-T., Ha Truong, T. T., Huynh, N.-T., & Truong, N.-S. (2020). Predicting energy consumption in multiple buildings using machine learning for improving energy efficiency and sustainability. *Journal of Cleaner Production*, *260*, 121082. https://doi.org/10.1016/j.jclepro.2020.121082

Pöppelbuß, J., & Röglinger, M. (2011). What makes a useful maturity model? A framework of general design principles for maturity models and its demonstration in business process management. *Proceedings of the 19th European Conference on Information Systems (ECIS 2011), 28*.

Priyono, A., Moin, A., & Putri, V. N. A. O. (2020). Identifying Digital Transformation Paths in the Business Model of SMEs during the COVID-19 Pandemic. *Journal of Open Innovation: Technology, Market, and Complexity*, *6*(4), 104. https://doi.org/10.3390/joitmc6040104

Pumplun, L., Tauchert, C., & and Heidt, M. (2019). A New Organizational Chassis for Artificial Intelligence-Exploring Organizational Readiness Factors. *Proceedings of the 27th European Conference on Information Systems (ECIS), Stockholm & Uppsala, Sweden*.

Rai, A. (2020). Explainable AI: From black box to glass box. *Journal of the Academy of Marketing Science*, *48*(1), 137–141. https://doi.org/10.1007/s11747-019-00710-5

Rai, A., Constantinides, P., & Sarker, S. (2019). Editor's Comments: Next-Generation Digital Platforms: Toward Human–AI Hybrids. *MIS Quarterly*, *43*(1).

Ransbotham, S., Kiron, D., Gerbert, P., & Reeves, M. (2017). Reshaping Business with Artificial Intelligence: Closing the Gap Between Ambition and Action. *MIT Sloan Management Review*. https://sloanreview.mit.edu/projects/reshaping-business-with-artificial-intelligence/

Reis, L., Maier, C., University of Bamberg (Germany), Mattke, J., University of Bamberg (Germany), Creutzenberg, M., University Hospital Regensburg (Germany), Weitzel, T., & University of Bamberg (Germany). (2020). Addressing User Resistance Would Have Prevented a Healthcare AI Project Failure. *MIS Quarterly Executive*, 279–296. https://doi.org/10.17705/2msqe.00038

Ren, P., Xiao, Y., Chang, X., Huang, P.-Y., Li, Z., Gupta, B. B., Chen, X., & Wang, X. (2022). A Survey of Deep Active Learning. *ACM Computing Surveys*, *54*(9), 1–40. https://doi.org/10.1145/3472291

Ribeiro, M., Singh, S., & Guestrin, C. (2016). "Why Should I Trust You?": Explaining the Predictions of Any Classifier. *Proceedings of the 2016 Conference of the North American Chapter of the Association for Computational Linguistics: Demonstrations*, 97–101. https://doi.org/10.18653/v1/N16-3020

Ritter, C., Oberländer, A. M., Stahl, B., Häckel, B., Klees, C., Koeppe, R., & Röglinger, M. (2023). How WashTec Explored Digital Business Models. *MIS Quarterly Executive, 22*(3). https://aisel.aisnet.org/misqe/vol22/iss3/3

Rowland, M., Carroll, N., & Conboy, K. (2022). Examining the Adoption of Artificial Intelligence for Digital Transformation. *Proceedings of the 30th European Conference on Information Systems (ECIS)*. https://aisel.aisnet.org/ecis2022_rp/140/

Ruefle, R., Dorofee, A., Mundie, D., Householder, A. D., Murray, M., & Perl, S. J. (2014). Computer Security Incident Response Team Development and Evolution. *IEEE Security & Privacy*, *12*(5), 16–26. https://doi.org/10.1109/MSP.2014.89

Russell, S. J., & Norvig, P. (with Davis, E., & Edwards, D.). (2016). *Artificial intelligence: A modern approach* (Third edition, Global edition). Pearson.

Sætra, H. S. (2023). The AI ESG protocol: Evaluating and disclosing the environment, social, and governance implications of artificial intelligence capabilities, assets, and activities. *Sustainable Development*, *31*(2), 1027–1037. https://doi.org/10.1002/sd.2438

Saffari, M., & Beagon, P. (2022). Home energy retrofit: Reviewing its depth, scale of delivery, and sustainability. *Energy and Buildings*, *269*, 112253. https://doi.org/10.1016/j.enbuild.2022.112253

Saqlain, M., Ali, S., & Lee, J. Y. (2023). A Monte-Carlo tree search algorithm for the flexible job-shop scheduling in manufacturing systems. *Flexible Services and Manufacturing Journal*, *35*(2), 548–571. https://doi.org/10.1007/s10696-021-09437-4

Sarker, I. H. (2021). Data Science and Analytics: An Overview from Data-Driven Smart Computing, Decision-Making and Applications Perspective. *SN Computer Science*, *2*(5), 377. https://doi.org/10.1007/s42979-021-00765-8

Schlette, D., Vielberth, M., & Pernul, G. (2021). CTI-SOC2M2 – The quest for mature, intelligence-driven security operations and incident response capabilities. *Computers & Security*, *111*, 102482. https://doi.org/10.1016/j.cose.2021.102482

Schmitt, J., Bönig, J., Borggräfe, T., Beitinger, G., & Deuse, J. (2020). Predictive model-based quality inspection using Machine Learning and Edge Cloud Computing. *Advanced Engineering Informatics*, *45*, 101101. https://doi.org/10.1016/j.aei.2020.101101

Schoormann, T., Strobel, G., Möller, F., Petrik, D., & Zschech, P. (2023). Artificial Intelligence for Sustainability—A Systematic Review of Information Systems Literature. *Communications of the Association for Information Systems*, *52*, 199–237. https://doi.org/10.17705/1CAIS.05209

Schwartz, R., Dodge, J., Smith, N. A., & Etzioni, O. (2020). Green AI. *Communications of the ACM*, *63*(12), 54–63. https://doi.org/10.1145/3381831

Settles, B. (2010). Active Learning Literature Survey. *University of Wisconsin, Madison*, *52*. https://minds.wisconsin.edu/handle/1793/60660

Shin, D. (2021). The effects of explainability and causability on perception, trust, and acceptance: Implications for explainable AI. *International Journal of Human-Computer Studies*, *146*, 102551. https://doi.org/10.1016/j.ijhcs.2020.102551

Shneiderman, B. (2021). Responsible AI: Bridging from ethics to practice. *Communications of the ACM*, *64*(8), 32–35. https://doi.org/10.1145/3445973

Shollo, A., Hopf, K., Thiess, T., & Müller, O. (2022). Shifting ML value creation mechanisms: A process model of ML value creation. *The Journal of Strategic Information Systems*, *31*(3), 101734. https://doi.org/10.1016/j.jsis.2022.101734

Siegel, S. (1956). *Nonparametric Statistics for the Behavioral Sciences*. McGraw-Hill Book Co. http://www.ru.ac.bd/wp-content/uploads/sites/25/2019/03/303_11_Siegel-Nonparametric-statistics-for-the-behavioral-sciences.pdf

Singla, K., Bose, J., & Naik, C. (2018). Analysis of Software Engineering for Agile Machine Learning Projects. *2018 15th IEEE India Council International Conference (INDICON)*, 1–5. https://doi.org/10.1109/INDICON45594.2018.8987154

Sjödin, D., Parida, V., Palmié, M., & Wincent, J. (2021). How AI capabilities enable business model innovation: Scaling AI through co-evolutionary processes and feedback loops. *Journal of Business Research*, *134*, 574–587. https://doi.org/10.1016/j.jbusres.2021.05.009

Sonnenberg, C., & vom Brocke, J. (2012a). Evaluation Patterns for Design Science Research Artefacts. In M. Helfert & B. Donnellan (Eds.), *Practical Aspects of Design Science* (Vol. 286, pp. 71–83). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-33681-2_7

Sonnenberg, C., & vom Brocke, J. (2012b). Evaluations in the Science of the Artificial – Reconsidering the Build-Evaluate Pattern in Design Science Research. In K. Peffers, M. Rothenberger, & B. Kuechler (Eds.), *Design Science Research in Information Systems. Advances in Theory and Practice* (Vol. 7286, pp. 381–397). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-29863-9_28

Stahl, B., Häckel, B., Leuthe, D., & Ritter, C. (2023). Data or Business First?—Manufacturers' Transformation Toward Data-driven Business Models. *Schmalenbach Journal of Business Research*, *75*(3), 303–343. https://doi.org/10.1007/s41471-023-00154-2

Statista. (2024). *Estimated worldwide spending on artificial intelligence (AI)-centric systems in 2023*. https://www.statista.com/statistics/1446052/worldwide-spending-on-ai-by-industry/

Stikvoort, D. (2019). *SIM3: Security Incident Management Maturity Model* (Issue SIM3 mkXVIIIc1). Open CSIRT Foundation (OCF) et al. http://opencsirt.org/wp-content/uploads/2019/12/SIM3-mkXVIIIc.pdf

Stone, P., Brooks, R., Brynjolfsson, E., Calo, R., Etzioni, O., Hager, G., Hirschberg, J., Kalyanakrishnan, S., Kamar, E., Kraus, S., Leyton-Brown, K., Parkes, D., Press, W., Saxenian, A., Shah, J., Tambe, M., & Teller, A. (2022). *Artificial Intelligence and Life in 2030: The One Hundred Year Study on Artificial Intelligence* (arXiv:2211.06318). arXiv. http://arxiv.org/abs/2211.06318

Studer, S., Bui, T. B., Drescher, C., Hanuschkin, A., Winkler, L., Peters, S., & Müller, K.-R. (2021). Towards CRISP-ML(Q): A Machine Learning Process Model with Quality Assurance Methodology. *Machine Learning and Knowledge Extraction*, *3*(2), 392–413. https://doi.org/10.3390/make3020020

Sund, K. J., Bogers, M. L. A. M., & Sahramaa, M. (2021). Managing business model exploration in incumbent firms: A case study of innovation labs in European banks. *Journal of Business Research*, *128*, 11–19. https://doi.org/10.1016/j.jbusres.2021.01.059

Tabladillo, M. (2022). *The Team Data Science Process lifecycle*. https://learn.microsoft.com/en-us/azure/architecture/data-science-process/lifecycle

Thangavelu, M., Krishnaswamy, V., & Sharma, M. (2021). Impact of comprehensive information security awareness and cognitive characteristics on security incident management – an empirical study. *Computers & Security*, *109*, 102401. https://doi.org/10.1016/j.cose.2021.102401

Tornatzky, L. G., & Fleischer, M. (1990). *The processes of technological innovation* (4. print). Lexington Books.

Tornede, T., Tornede, A., Hanselle, J., Mohr, F., Wever, M., & Hüllermeier, E. (2022). Towards Green Automated Machine Learning: Status Quo and Future Directions. *Journal of Artificial Intelligence Research*. https://doi.org/10.1613/jair.1.14340

Tremblay, M. C., Hevner, A. R., & Berndt, D. J. (2010). Focus Groups for Artifact Refinement and Evaluation in Design Research. *Communications of the Association for Information Systems*, *26*. https://doi.org/10.17705/1CAIS.02627

Tsoka, T., Ye, X., Chen, Y., Gong, D., & Xia, X. (2022). Explainable artificial intelligence for building energy performance certificate labelling classification. *Journal of Cleaner Production*, *355*, 131626. https://doi.org/10.1016/j.jclepro.2022.131626

Turner, J. R., & Müller, R. (2003). On the nature of the project as a temporary organization. *International Journal of Project Management*, *21*(1), 1–8. https://doi.org/10.1016/S0263-7863(02)00020-0

Urbach, N., & Röglinger, M. (Eds.). (2019). *Digitalization Cases: How Organizations Rethink Their Business for the Digital Age*. Springer International Publishing. https://doi.org/10.1007/978-3-319-95273-4

Urbach, N., Röglinger, M., Kautz, K., Alias, R. A., Saunders, C., & Wiener, M. (Eds.). (2021). *Digitalization Cases Vol. 2: Mastering Digital Transformation for Global Business*. Springer International Publishing. https://doi.org/10.1007/978-3-030-80003-1

Uren, V., & Edwards, J. S. (2023). Technology readiness and the organizational journey towards AI adoption: An empirical study. *International Journal of Information Management*, *68*, 102588. https://doi.org/10.1016/j.ijinfomgt.2022.102588

Van Der Kleij, R., Schraagen, J. M., Cadet, B., & Young, H. (2022). Developing decision support for cybersecurity threat and incident managers. *Computers & Security*, *113*, 102535. https://doi.org/10.1016/j.cose.2021.102535

van Giffen, B., & Ludwig, H. (2023). How siemens democratized artificial intelligence. *MIS Quarterly Executive*, *22*(1), 3.

Van Steenbergen, M., Bos, R., Brinkkemper, S., Van De Weerd, I., & Bekkers, W. (2010). The Design of Focus Area Maturity Models. In R. Winter, J. L. Zhao, & S. Aier (Eds.), *Global Perspectives on Design Science Research* (Vol. 6105, pp. 317–332). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-13335-0_22

van Wynsberghe, A. (2021). Sustainable AI: AI for sustainability and the sustainability of AI. *AI and Ethics*, *1*(3), 213–218. https://doi.org/10.1007/s43681-021-00043-6

Varta. (2024). *VARTA affected by cyber attack*. https://www.varta-ag.com/fileadmin/varta_ag/publications/ad-hoc_announcements/20240213_VARTA_AG_Ad_hoc_cyber_attack_EN.pdf

Vassilakopoulou, P., & Hustad, E. (2023). Bridging Digital Divides: A Literature Review and Research Agenda for Information Systems Research. *Information Systems Frontiers*, *25*(3), 955–969. https://doi.org/10.1007/s10796-020-10096-3

VDMA. (2023). *Marktbericht 2022: Die deutsche Werkzeugmaschinenindustrie und ihre Stellung im Weltmarkt*. https://vdw.de/wp-content/uploads/2023/07/VDW-Marktbericht-2022.pdf

VDMA. (2024). *Marktbericht 2023: Die deutsche Werkzeugmaschinenindustrie und ihre Stellung im Weltmarkt*. https://vdw.de/wp-content/uploads/2024/06/VDW-Marktbericht-2023.pdf

Veit, D. J., & Thatcher, J. B. (2023). Digitalization as a problem or solution? Charting the path for research on sustainable information systems. *Journal of Business Economics*, *93*(6–7), 1231–1253. https://doi.org/10.1007/s11573-023-01143-x

Verdecchia, R., Sallou, J., & Cruz, L. (2023). A systematic review of Green AI. *WIREs Data Mining and Knowledge Discovery*, *13*(4), e1507. https://doi.org/10.1002/widm.1507

Vial, G. (2019). Understanding digital transformation: A review and a research agenda. *The Journal of Strategic Information Systems*, *28*(2), 118–144. https://doi.org/10.1016/j.jsis.2019.01.003

Vial, G., Cameron, A., Giannelia, T., & Jiang, J. (2023). Managing artificial intelligence projects: Key insights from an AI consulting firm. *Information Systems Journal*, *33*(3), 669–691. https://doi.org/10.1111/isj.12420

Vilone, G., & Longo, L. (2021). Notions of explainability and evaluation approaches for explainable artificial intelligence. *Information Fusion*, *76*, 89–106. https://doi.org/10.1016/j.inffus.2021.05.009

Visscher, H., Sartori, I., & Dascalaki, E. (2016). Towards an energy efficient European housing stock: Monitoring, mapping and modelling retrofitting processes. *Energy and Buildings*, *132*, 1–3. https://doi.org/10.1016/j.enbuild.2016.07.039

vom Brocke, J., Winter, R., Hevner, A., & Maedche, A. (2020). Special Issue Editorial – Accumulation and Evolution of Design Knowledge in Design Science Research: A Journey Through Time and Space. *Journal of the Association for Information Systems*, *21*(3), 520–544. https://doi.org/10.17705/1jais.00611

Wallace, S., Green, K. Y., Johnson, C. M., Cooper, J. T., & Gilstrap, C. M. (2020). An Extended TOE Framework for Cybersecurity Adoption Decisions. *Communications of the Association for Information Systems*, *47*, 338–363. https://doi.org/10.17705/1CAIS.04716

WashTec. (2024). *WashTec & Exasol: Glänzende Aussichten für die Datenzukunft*. https://www.exasol.com/de/casestudy/washtec/

Weber, M., Engert, M., Schaffer, N., Weking, J., & Krcmar, H. (2023). Organizational Capabilities for AI Implementation—Coping with Inscrutability and Data Dependency in AI. *Information Systems Frontiers*, *25*(4), 1549–1569. https://doi.org/10.1007/s10796-022-10297-y

Webster, J., & Watson, R. (2002). Analyzing the Past to Prepare for the Future: Writing a Literature Review. *MIS Quarterly*, *26*. https://doi.org/10.2307/4132319

Wegener, C., Milde, T., & Dolle, W. (2016). *Informationssicherheits-Management: Leitfaden für Praktiker und Begleitbuch zur CISM-Zertifizierung*. Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-662-49167-6

Westenberger, J., Schuler, K., & Schlegel, D. (2022). Failure of AI projects: Understanding the critical factors. *Procedia Computer Science*, *196*, 69–76. https://doi.org/10.1016/j.procs.2021.11.074

Wirth, R., & Hipp, J. (2000). CRISP-DM: Towards a Standard Process Model for Data Mining. *Proceedings of the 4th International Conference on the Practical Applications of Knowledge Discovery and Data Mining*.

Wolfswinkel, J. F., Furtmueller, E., & Wilderom, C. P. M. (2013). Using grounded theory as a method for rigorously reviewing literature. *European Journal of Information Systems*, *22*(1), 45–55. https://doi.org/10.1057/ejis.2011.51

Wu, C.-J., Raghavendra, R., Gupta, U., Acun, B., Ardalani, N., Maeng, K., Chang, G., Behram, F. A., Huang, J., Bai, C., Gschwind, M., Gupta, A., Ott, M., Melnikov, A., Candido, S., Brooks, D., Chauhan, G., Lee, B., Lee, H.-H. S., … Hazelwood, K. (2022). Sustainable AI: Environmental Implications, Challenges and Opportunities. *Proceedings of Machine Learning and Systems 4*. https://doi.org/10.48550/ARXIV.2111.00364

Wu, X., Zheng, W., Chen, X., Zhao, Y., Yu, T., & Mu, D. (2021). Improving high-impact bug report prediction with combination of interactive machine learning and active learning. *Information and Software Technology*, *133*, 106530. https://doi.org/10.1016/j.infsof.2021.106530

Xiong, F., Kühl, N., & Stauder, M. (2024). Designing a computer-vision-based artifact for automated quality control: A case study in the food industry. *Flexible Services and Manufacturing Journal*. https://doi.org/10.1007/s10696-023-09523-9

Yalcintas, M. (2008). Energy-savings predictions for building-equipment retrofits. *Energy and Buildings*, *40*(12), 2111–2120. https://doi.org/10.1016/j.enbuild.2008.06.008

Yu, Z., Kaplan, Z., Yan, Q., & Zhang, N. (2021). Security and Privacy in the Emerging Cyber-Physical World: A Survey. *IEEE Communications Surveys & Tutorials*, *23*(3), 1879–1919. https://doi.org/10.1109/COMST.2021.3081450

Yuan, S., & Wu, X. (2021). Deep learning for insider threat detection: Review, challenges and opportunities. *Computers & Security*, *104*, 102221. https://doi.org/10.1016/j.cose.2021.102221

Zeiser, A., Özcan, B., Van Stein, B., & Bäck, T. (2023). Evaluation of deep unsupervised anomaly detection methods with a data-centric approach for on-line inspection. *Computers in Industry*, *146*, 103852. https://doi.org/10.1016/j.compind.2023.103852

Zhao, H., Wang, Y., Duan, J., Huang, C., Cao, D., Tong, Y., Xu, B., Bai, J., Tong, J., & Zhang, Q. (2020). Multivariate Time-Series Anomaly Detection via Graph Attention Network. *2020 IEEE International Conference on Data Mining (ICDM)*, 841–850. https://doi.org/10.1109/ICDM50108.2020.00093

Zhu, K., & Kraemer, K. L. (2005). Post-Adoption Variations in Usage and Value of E-Business by Organizations: Cross-Country Evidence from the Retail Industry. *Information Systems Research*, *16*(1), 61–84. https://doi.org/10.1287/isre.1050.0045

Zhu, Y., & Yang, K. (2019). Tripartite Active Learning for Interactive Anomaly Discovery. *IEEE Access*, *7*, 63195–63203. https://doi.org/10.1109/ACCESS.2019.2915388

V Appendix

1 Index of Research Articles

Research Article #1: Uncovering the Sweet Spot of Artificial Intelligence Projects: An Exploration of Success Factors

Leuthe, Daniel; Meierhöfer, Simon; Häckel, Björn; Kolbeck, Thomas. Uncovering the Sweet Spot of Artificial Intelligence Projects: An Exploration of Success Factors. *Submitted Working Paper*

Research Article #2: Data or Business First? – Manufacturers´ Transformation toward Data-driven Business Models

Stahl, Bastian; Häckel, Björn; Leuthe, Daniel; Ritter, Christian. Data or Business First- Manufacturer´s Transformation toward Data-driven Business Models. *Schmalenbach Journal of Business Research (SBUR) (2023).* DOI: 10.1007/s41471-023-00154-2

(VHB-Jourqual 3: B | VHB-Jourqual 4: B | Impact Factor: -)

Research Article #3: Managing the Inevitable – A Maturity Model to Establish Incident Response Management Capabilities

Bitzer, Michael; Häckel, Björn; Leuthe, Daniel; Ott, Joshua; Stahl, Bastian; Strobel, Jacqueline. Managing the Inevitable – A Maturity Model to Establish Incident Response Management Capabilities. *Computers & Security (2023).* DOI: 10.1016/j.cose.2022.103050

(VHB- Jourqual 3: - | VHB-Jourqual 4: B | Impact Factor: 5.6)

Research Article #4: A Data-Efficient Active Learning Architecture for Anomaly Detection in Industrial Time Series Data

Holtz, David; Kaymakci, Can; Leuthe, Daniel; Wenninger, Simon; Sauer, Alexander. A Data-Efficient Active Learning Architecture for Anomaly Detection in Industrial Time Series Data. *Flexible Services and Manufacturing Journal (2025).* DOI: 10.1007/s10696-024-09588-0

(VHB-Jourqual 3: B | VHB-Jourqual 4: B | Impact Factor: 2.5)

Research Article #5: Leveraging Explainable AI for Informed Building Retrofit Decisions: Insights from a Survey

Leuthe, Daniel; Mirlach, Jonas; Wenninger, Simon; Wiethe, Christian. Leveraging Explainable AI for Informed Building Retrofit Decisions: Insights from a Survey. *Energy and Buildings (2024).* DOI: 10.1016/j.enbuild.2024.114426

(VHB-Jourqual 3: - | VHB-Jourqual 4: - | Impact Factor: 6.6)

Research Article #6: Towards Sustainability of AI – Identifying Design Patterns for Sustainable Machine Learning Development

Leuthe, Daniel; Meyer-Hollatz, Tim; Plank, Tobias; Senkmüller, Anja. Towards Sustainability of AI – Identifying Design Patterns for Sustainable Machine Learning Development. *Information Systems Frontiers (2024).* DOI: 10.1007/s10796-024-10526-6

(VHB-Jourqual 3: B | VHB-Jourqual 4: B | Impact Factor: 6.9)

I also co-authored further research papers throughout the dissertation, which are not part of this doctoral thesis. An excerpt of the articles can be found in the following:

Leuthe, Daniel; Weiß, Florian; Dersch Julian; Bitzer, Michael. **Towards Secure Cloud-Computing in FinTechs: An Artefact for Prioritizing Information Security Measures**. *Proceedings of the 57th Hawaii International Conference on System Sciences (2024).*

Fabri, Lukas; Leuthe, Daniel; Schneider, Lars-Manuel; Wenninger, Simon. **Fostering Non-Intrusive Load Monitoring for Smart Energy Management in Industrial Manufacturing: An Active Machine Learning Approach**. *Working Paper (2025).*

Dormehl, Julian; Leuthe, Daniel; Karnebogen, Philip; Häckel, Björn; Wenninger, Simon. **Beyond mere Prediction Performance – Deriving Criteria for the Selection of Supervised Machine Learning Algorithms**. *Working Paper (2023).*

Beck, Sebastian; Fabri, Lukas; Häckel, Björn; Leuthe, Daniel; Meierhöfer, Simon. **Unraveling Digital Transformation Projects: A Taxonomy and Prevalent Archetypes**. *Working Paper (2024).*

Plank, Tobias; Leuthe, Daniel; Meyer-Hollatz, Tim. **Unlocking Artificial Intelligence for Sustainability: A Taxonomy and Seven Archetypes for Achieving the SDGs**. *Working Paper (2025).*

2 Individual Contribution to the Research Articles

This cumulative dissertation comprises six research articles representing the main body of work. All articles were developed in teams with multiple co-authors. This section details the respective research settings and highlights my individual contributions to each research article.

Research article #1: I co-authored this research article with Simon Meierhöfer, Björn Häckel, and Thomas Kolbeck. Overall, the development of the article and its research idea were mainly driven by Simon Meierhöfer and myself. As the paper was developed in a three-phased methodical approach, I closely engaged in all three phases to derive, structure, and evaluate the paper’s main findings. Regarding the development of the research article, I co-developed the initial draft of the research paper, and I was mainly engaged in identifying and structuring the success factors as well as their integration into the core artifact. While, to a large extent, this article reflects the work of Simon Meierhöfer and myself, all co-authors promoted the advancement of the paper throughout the entire project.

Research article #2: This research article was developed by a team of four co-authors (Bastian Stahl, Björn Häckel, Daniel Leuthe, and Christian Ritter). Together, we developed the maturity model for archetypes and capabilities required for distinct data-driven business models in the manufacturing sector. My contributions included specifying the research method and deriving the structure and content of the final artifact, especially with regard to the technical dimensions based on the identified related work. Furthermore, I conducted several interviews with industry experts to evaluate the maturity model. I engaged in the initial draft of the paper and its further textual elaboration throughout the revisions. Bastian Stahl is the lead author of this paper.

Research article #3: I co-authored this research article with Michael Bitzer, Björn Häckel, Joshua Ott, Bastian Stahl, and Jacqueline Strobel. All six co-authors jointly developed the incident response management maturity model. Hence, all co-authors contributed equally to the article’s content and supported the project throughout its duration. In this vein, I was especially involved in the steps of the conceptualization, the structure of the methodology, and writing the original draft. Additionally, I engaged in the further development and its additional textual and content-related refinement throughout the two revisions.

Research article #4: This research article was developed by a team of five co-authors (David Holtz, Can Kaymakci, Daniel Leuthe, Simon Wenninger, and Alexander Sauer). Our collaborative effort resulted in developing the active learning architecture and its validation in a case study. My contributions especially included outlining the overall storyline of the research article, writing its textual content such as the introduction, theoretical background, experiment results, and, in particular, the methodical approach. Thus, I engaged in the initial draft of the paper and led its further textual and technical elaboration throughout the submissions and revisions. All co-authors contributed equally to the article’s content.

Research article #5: This research article was developed by a team of four co-authors (Daniel Leuthe, Jonas Mirlach, Simon Wenninger, and Christian Wiethe). As the leading author, I developed the artifact's basic research idea and concept and contributed significantly to the design of the three-step research methodology - of both technical model development and data provision. Further, I contributed to the structure of the overarching storyline, evaluation, and writing all sections of the manuscript. Additionally, I was in charge of preparing the article’s refinement and preparing it for submission. While, to a large extent, this article reflects my work, all co-authors promoted the advancement of the paper throughout the entire project.

Research article #6: I co-authored this research article with Tim Meyer-Hollatz, Anja Senkmüller, and Tobias Plank. In particular, Tim Meyer-Hollatz and I played a crucial role in the entire process, from the creation and conceptualization of the research idea, investigation, development, visualization, and evaluation of the results to writing all chapters of the original manuscript draft. Furthermore, together with Tim Meyer-Hollatz, I contributed significantly to preparing the research article for submission and extensively revising the paper after receiving feedback during the review process. All four co-authors contributed to the article’s content and supported the project throughout its duration.

3 Research Article #1

Uncovering the Sweet Spot of Artificial Intelligence Projects: An Exploration of Success Factors

Authors: Leuthe, Daniel; Meierhöfer, Simon; Häckel, Björn; Kolbeck, Thomas

Submitted Working Paper

Extended Abstract[1]:

Organizations across industries aim to disseminate AI through respective projects. Nevertheless, despite the role of AI to serve as a driver for innovation, organizations encounter significant pitfalls when planning and executing AI projects (Merhi 2023). As a result, AI projects often fail to live up to the intended outcomes or are terminated before completion. This circumstance turns AI projects into a risky matter, as their failure entails sunk costs and may jeopardize competitiveness (Vial et al. 2023).

Hence, for organizations that aim to thrive in the trajectory of AI, a systematic understanding of the requirements that drive the successful implementation of AI projects is indispensable. This research article refers to such antecedents as success factors (SFs) (Bullen and Rockart 1981). In the literature, the number of studies that deal with the successful implementation of AI in organizations has grown remarkably in recent years. Here, scholars discuss AI implementation mainly against the backdrop of concepts such as AI adoption in general or AI readiness in particular (Jöhnk et al. 2021; Weber et al. 2023). Moreover, the literature points to the need to develop or acquire specific capabilities to accomplish AI implementation. Studies by Lee et al. (2023) and Merhi (2023) provide initial overviews of SFs, but they do not provide a holistic compilation of SFs, as they do not provide empirical evidence and only consider a limited amount of literature. Hence, this research article aims to answer the following research questions: What are the SFs for AI projects?

To answer the research question, a three-stage research approach is conducted (i.e., systematic literature review, in-depth interview study, focus group discussions). The result is a framework of 24 SFs for AI projects, structured along four overarching success dimensions (i.e., datability, desirability, feasibility, and viability) and specified by 93 subordinated success

[1] At the time of writing, this research article is under review for publication in a scientific journal. Therefore, an extended abstract, taken from the research article, is provided here.

manifestations. For each SF, a comprehensive description and the specific AI characteristics are provided. Finally, by illustrating how the SFs manifest in the four key phases of the AI project workflow, the results provide an authoritative instance for a systematic understanding of the scope in which they emerge in AI projects.

The results are novel as they systematically explore the SFs for AI projects by synthesizing extant knowledge from literature with insights into the trajectory of AI through empirical data. In this way, we not only lay the foundation for researchers to advance knowledge on how to conceptualize and operationalize AI projects, but also provide empirical groundwork for further theorizing on the successful implementation of AI in general. Further, the results provide organizational stakeholders with a coherent and conclusive picture of the SFs and their tangible SMs contextualized in the AI project workflow that help them plan and execute AI projects successfully.

Keywords: Artificial Intelligence, Artificial Intelligence Project, Artificial Intelligence Project Workflow, Project Management, Project Success, Success Factors

References:

Bullen CV, Rockart JF (1981) A primer on critical success factors. Center for Information Systems Research, Massachusetts Institute of Technology 69:1–64

Jöhnk J, Weißert M, Wyrtki K (2021) Ready or Not, AI Comes— An Interview Study of Organizational AI Readiness Factors. Bus Inf Syst Eng 63:5–20. https://doi.org/10.1007/s12599-020-00676-7

Lee MC, Scheepers H, Lui AK, Ngai EW (2023) The implementation of artificial intelligence in organizations: A systematic literature review. Information & Management 60:103816. https://doi.org/10.1016/j.im.2023.103816

Merhi MI (2023) An evaluation of the critical success factors impacting artificial intelligence implementation. International Journal of Information Management 69:102545. https://doi.org/10.1016/j.ijinfomgt.2022.102545

Vial G, Cameron A-F, Giannelia T, Jiang J (2023) Managing artificial intelligence projects: Key insights from an AI consulting firm. Information Systems Journal 33:669–691. https://doi.org/10.1111/isj.12420

Weber M, Engert M, Schaffer N, Weking J, Krcmar H (2023) Organizational Capabilities for AI Implementation—Coping with Inscrutability and Data Dependency in AI. Inf Syst Front. https://doi.org/10.1007/s10796-022-10297-y

4 Research Article #2

Data or Business First? – Manufacturers´ Transformation toward Data-driven Business Models

Authors: Stahl, Bastian; Häckel, Björn; Leuthe, Daniel; Ritter, Christian

Published in: *Schmalenbach Journal of Business Research (SBUR) (2023)*

Abstract: Driven by digital technologies, manufacturers aim to tap into data-driven business models, in which value is generated from data as a complement to physical products. However, this transformation can be complex, as different archetypes of data-driven business models require substantially different business and technical capabilities. While there are manifold contributions to research on technical capability development, an integrated and aligned perspective on both business and technology capabilities for distinct data-driven business model archetypes is needed. This perspective promises to enhance research's understanding of this transformation and offers guidance for practitioners. As maturity models have proven to be valuable tools in capability development, we follow a design science approach to develop a maturity model for the transformation toward archetypal data-driven business models. To provide an integrated perspective on business and technology capabilities, the maturity model leverages a layered enterprise architecture model. By applying and evaluating in use at two manufacturers, we find two different transformation approaches, namely 'data first' and 'business first'. The resulting insights highlight the model's integrative perspective's value for research to improve the understanding of this transformation. For practitioners, the maturity model allows a status quo assessment and derives fields of action to develop the capabilities required for the aspired data-driven business model.

Keywords: Data-driven business models, data-driven services, data analytics, manufacturing, enterprise architecture

5 Research Article #3

Managing the Inevitable – A Maturity Model to Establish Incident Response Management Capabilities

Authors: Bitzer, Michael; Häckel, Björn; Leuthe, Daniel; Ott, Joshua; Stahl, Bastian; Strobel, Jacqueline

Published in: *Computers & Security (2023)*

Abstract: Although the ongoing digital transformation offers new opportunities for organizations, more emphasis on information security is needed due to the evolving cyber-threat landscape. Despite all preventive measures, security incidents cannot entirely be mitigated. Organizations must establish incident response management to treat inevitable incidents in a structured manner and under considerable time pressure. If not handled, incidents can result in reputational or financial losses and disrupt business continuity. Especially organizations that have not addressed incident response management extensively need to understand which capabilities are required to develop their incident response management. However, research still lacks a practice-grounded and socio-technical conceptualization of those capabilities and their development. For such challenges, maturity models have proven valuable in practice and research. This paper follows a design science research approach to develop an incident response management maturity model (IRM3) closely aligned with practice requirements under a socio-technical lens. Iteratively applying and evaluating the IRM3 with seven real-world organizations leverages its comprehensive view based on four focus areas and 29 capability dimensions to understand which capabilities organizations need to approach incident response management. Building on existing research, this work provides a comprehensive perspective on incident response management and its associated capabilities. For practitioners, especially in organizations with initial incident response maturity, the IRM3 offers descriptive value when used as a status quo assessment tool and prescriptive value by outlining capabilities for successful incident response management.

Keywords: Design science research, Incident response management, Information security, Maturity model, Socio-technical

6 Research Article #4

A Data-Efficient Active Learning Architecture for Anomaly Detection in Industrial Time Series Data

Authors: Holtz, David; Kaymakci, Can; Leuthe, Daniel; Wenninger, Simon; Sauer, Alexander

Published in: *Flexible Services and Manufacturing Journal (2025)*

Abstract: Anomaly detection is becoming increasingly important and has found its way into manufacturing applications. The potential is seen in use cases such as maintenance cost reduction, machine fault reduction, or increased overall production based on industrial time series data. However, obstacles arise in practice. Supervised algorithms lack limited and expensive labeled training data, and unsupervised algorithms do not have the capabilities for evaluation and tracking. We propose a data-efficient architecture for anomaly detection using energy consumption time series data to address these limitations. To do so, we design an active learning model that optimizes an unsupervised model by integrating budgeted expert feedback. Our solution builds on an autoencoder to leverage latent space representations for an additional supervised feedforward network trained with expert knowledge labels to distinguish between normal data and anomalies. Four different strategies for querying the still-unlabeled data are compared so that the expert's resources are used efficiently. We validate our concept in an industrial robotic screwdriving application based on energy data for condition monitoring. Findings for the application tested indicate that anomaly detection performance can be significantly increased by 59 % for the F1 score with active learning compared to unsupervised models. Furthermore, models trained only on energy consumption data exhibit the same performance as models trained on difficult-to-obtain mechanical process data, thus confirming the practicality of our proposed approach and data efficiency for the use of easily accessible energy data in manufacturing applications. While our approach enables an active learning model to be added to an existing unsupervised model, it allows for straightforward benchmarking and extension to other manufacturing applications.

Keywords: Anomaly Detection, Active Learning, Data Efficient, Manufacturing System, Multivariate Time Series

7 Research Article #5

Leveraging Explainable AI for Informed Building Retrofit Decisions: Insights from a Survey

Authors: Leuthe, Daniel; Mirlach, Jonas; Wenninger, Simon; Wiethe, Christian

Published in: *Energy and Buildings (2024)*

Abstract: Accurate predictions of building energy consumption are essential for reducing the energy performance gap. While data-driven energy quantification methods based on machine learning deliver promising results, the lack of Explaina-bility prevents their widespread application. To overcome this, Explainable Artificial Intelligence (XAI) was intro-duced. However, to this point, no research has examined how effective these explanations are concerning decision-makers, i.e., property owners. To address this, we implement three transparent models (Linear Regression, Decision Tree, QLattice) and apply four XAI methods (Partial Dependency Plots, Accumulated Local Effects, Local Interpreta-ble Model-Agnostic Explanations, Shapley Additive Explanations) to an Artificial Neural Network using a real-world dataset of 25,000 residential buildings. We evaluate their Prediction Accuracy and Explainability through a survey with 137 participants considering the human-centered dimensions of explanation satisfaction and perceived fidelity. The results quantify the Explainability-Accuracy trade-off in building energy consumption forecasting and how it can be counteracted by choosing the right XAI method to foster informed retrofit decisions. For research, we set the foundation for further increasing the Explainability of data-driven energy quantification methods and their human-centered evaluation. For practice, we encourage using XAI to reduce the acceptance gap of data-driven meth-ods, whereby the XAI method should be selected carefully, as the Explainability within the methods varies by up to 10%.

Keywords: Building energy performance; Energy efficiency; Energy quantification methods; Explainability-accuracy trade-off; Explainable artificial intelligence; Survey

8 Research Article #6

Towards Sustainability of AI – Identifying Design Patterns for Sustainable Machine Learning Development

Authors: Leuthe, Daniel; Meyer-Hollatz, Tim; Plank, Tobias; Senkmüller, Anja

Published in: *Information Systems Frontiers (2024)*

Abstract: As artificial intelligence (AI) and machine learning (ML) advance, concerns about their sustainability impact grow. The emerging field "Sustainability of AI" addresses this issue, with papers exploring distinct aspects of ML's sustainability. However, it lacks a comprehensive approach that considers all ML development phases, treats sustainability holistically, and incorporates practitioner feedback. In response, we developed the sustainable ML design pattern matrix (SML-DPM) consisting of 35 design patterns grounded in justificatory knowledge from research, refined with naturalistic insights from expert interviews and validated in three real-world case studies using a web-based instantiation. The design patterns are structured along a four-phased ML development process, the sustainability dimensions of environmental, social, and governance (ESG), and allocated to five ML stakeholder groups. It represents the first artifact to enhance each ML development phase along each ESG dimension. The SML-DPM fuels advancement by aggregating distinct research, laying the groundwork for future investigations, and providing a roadmap for sustainable ML development.

Keywords: Artificial Intelligence, Design Patterns, ESG, Machine Learning, Sustainability of AI