International Journal of
Information Systems and Project Management
ISSN (print):2182-7796, ISSN (online):2182-7788, ISSN (cd-rom):2182-780X
Available online at ijispm.sciencesphere.org

Failure factors of AI projects: results from expert interviews

Dennis Schlegel
Reutlingen University
Alteburgstr. 150, 72762 Reutlingen
Germany
dennis.schlegel@reutlingen-university.de

Kajetan Schuler
EXXETA AG
Stockholmer Platz 1, 70173 Stuttgart
Germany
kajetan.schuler@exxeta.com

Jens Westenberger
Reutlingen University
Alteburgstr. 150, 72762 Reutlingen
Germany
jens.westenberger@posteo.de

Abstract:
In the last few years, business firms have substantially invested into the artificial intelligence (AI) technology. However, according to several studies, a significant percentage of AI projects fail or do not deliver business value. Due to the specific characteristics of AI projects, the existing body of knowledge about success and failure of information systems (IS) projects in general may not be transferrable to the context of AI. Therefore, the objective of our research has been to identify factors that can lead to AI project failure. Based on interviews with AI experts, this article identifies and discusses 12 factors that can lead to project failure. The factors can be further classified into five categories: unrealistic expectations, use case related issues, organizational constraints, lack of key resources, and technological issues. This research contributes to knowledge by providing new empirical data and synthesizing the results with related findings from prior studies. Our results have important managerial implications for firms that aim to adopt AI by helping the organizations to anticipate and actively manage risks in order to increase the chances of project success.

Keywords:
AI; artificial intelligence; machine learning; ML; failure; success.

DOI: 10.12821/ijispm110302

Manuscript received: 28 March 2022
Manuscript accepted: 18 December 2022

1. Introduction

The concept and term of artificial intelligence (AI) dates back to the 1950s [1]. Since then, the technology has lived through cycles of hype (AI spring) and stagnation (AI winter). Even though AI has become increasingly relevant, no unified definition of the term has emerged. Generally speaking, researchers agree that AI belongs to the field of computer sciences and is about developing independent applications that can solve problems on their own [2]. AI technologies can be divided into the subcategories of strong and weak AI. Applications of weak AI, such as speech recognition or fraud detection, are already available today and are constantly being further developed. The main characteristic of such applications is that they are developed for a special task and are not able to execute other tasks [3]. Distinct from this is strong AI, which attempts to replicate the human brain in order to develop an AI that is not limited to specific tasks [2, 3]. As strong AI is not available today [4], our paper focuses on weak AI. Weak AI is technologically based on machine learning (ML), which includes among others neural networks, and deep learning technologies. The terms ML and AI are often used synonymously, especially in a business context.

In the last years, both the adoption of AI, and the expectations regarding the economic potential of AI have risen sharply. Already in 2017, research has shown an increasing investment in AI by leading tech firms [5]. More recent studies even found that all companies in their sample were at least "actively evaluating use cases for ML applications" [6]. However, according to recent research [7], only few such projects are successful in delivering actual business value. Other sources even report that more than 80% of AI-related projects fail [8].

Hence, it is crucial to understand the factors that lead to failure of AI projects in order to avoid the pitfalls and fully exploit the potential of the technology. A large body of literature exists about the success and failure of Information Systems (IS) projects in general [9–14] or specific application types such as ERP systems [15–17]. However, due to the specific characteristics of AI, such as its algorithmic complexity and the broad and holistic changes that accompany the introduction of AI systems in organizations, these factors need to be revisited and extended to fit the context of AI [18].

A few recent publications have dealt with failure of AI-related projects [7, 8, 19–21]. Additionally, a number of studies have been conducted on related themes such as challenges [22], success factors [6] and organizational readiness for AI [23]. Although some of the previous studies provide interesting results, further research is clearly necessary due to the limited transferability of the findings to this paper's aim, as well as, a number of other limitations in the existing body of literature (see section 2.2).

The aim of this study has therefore been to identify factors that lead to failure of AI projects in a general context. To reach this aim, were have conducted a literature review to synthesize prior findings in a systematic way. Additionally, we have collected new empirical data using qualitative, semi-structured interviews that were analyzed using an inductive coding approach [24–26].

This article makes a contribution to knowledge by providing new empirical data on critical factors leading to failure of AI projects, and synthesizing the findings with prior related studies to provide a more complete picture of the topic. The identified factors provide important insights and guidance for organizations to proactively increase the rate of success of their projects in order to exploit the potential of the AI technology and avoid costly project failure.

The paper is organized as follows. In section 2, a brief overview on the topic of success and failure of information systems (IS) is presented, before related work regarding AI projects is reviewed in detail. Subsequently, the research design is explained in section 3, followed by the presentation of the results in section 4. In section 5, we discuss the results and synthesize our findings with prior literature. Finally, we conclude by summarizing the main findings and our contribution to research, discussing practical implications for business firms, as well as, pointing out limitations of the study and opportunities for future research in this area.

2. Literature review

2.1 Success and failure of IS and AI projects

Success and failure of IS projects in general is a thoroughly studied subject. Before discussing typical causes of IS success and failure, these terms need to be defined. In general, success can be defined as “achieving the goals that have been established for an undertaking” [27]. On the flipside, IS failure can be defined as the perceived inability of the project to meet the requirements or expectations of various stakeholders [28]. The mentioned requirements and expectations can be manifold in this context, for example there are not only functional but also financial or time requirements.

As the goals of IS were not always perfectly clear, the definition of IS success was a challenge. In an attempt to identify dimensions of IS success, DeLone and McLean [29] undertook a literature review of papers published between 1981 and 1987. They identified six interdependent dimensions of IS success (system quality, information quality, use, user satisfaction, individual impact, and organizational impact) and used these in a model to explain IS success [29]. In the following years, this model was expanded and modified numerous times [10, 27, 30, 31]. Furthermore, both the original and the revised models have been validated to be good predictors of IS success [32–34]. In another stream of research, several studies attempted to identify determinants that have an effect on one or more of the stated dimensions [27, 35]. In total, over 50 determinants were identified to correlate with dimensions of IS success.

On the other hand, the failure of IS projects was also widely studied focusing on the discrepancy between actual and expected requirements. Similar to IS success, studies tried to investigate dimensions and determinants of IS failure [36, 37]. For example, Nelson [38] analyzed over 90 IS projects and concluded that there are 36 common mistakes in four categories: process, people, product and technology.

It can be concluded that IS success and failure is an intensely studied subject and existing models have been proven to be good predictors of IS success. However, due to the specific nature of AI projects, it is still largely unclear whether these results can be transferred to the context of AI projects. Due to specific characteristics of the AI technology, it can be assumed that AI has to be regarded separately from other digital technologies, as previously stated in the literature [23]. In the following paragraphs, we will first explain how characteristics of AI differ from other technologies, before addressing particularities of AI projects.

Looking at AI and the underlying technologies it can be seen that technology itself as well as technical characteristics [1, 39] differ from traditional IS/IT projects. As AI is not explicitly programmed to perform a specific task, but it rather learns from previous experience (data), the development and adoption of AI can be seen as a paradigm shift [2]. The shift and therefore the implementation and use of AI requires vastly different skills and is of higher complexity compared to typical software engineering projects. One example is, that most AI algorithms require deep statistical as well as mathematical knowledge. Furthermore, AI is a highly interdisciplinary field that requires not only software engineering and AI skills, but also domain knowledge for example [1].

Indeed, reports by practitioners indicate that AI projects differ from other projects in various characteristics. In his blog post, Mehta [40] presents several dimensions where AI projects differ from traditional IT or software development projects. For instance, the project focus of AI projects is on data exploration and insights instead of application development. Moreover, the goal of AI projects is often to use the technology strategically to transform the business, while traditional IS have more tactical goals. In terms of business knowledge, in traditional projects, business rules are given to programmers to be implemented in the software. In contrast, in AI projects the business data is used to discover the business rules from the data. Due to their more experimental trial and error approach, AI projects are also more difficult to manage to a fixed schedule than traditional projects.

Certainly, more research academic research is necessary to discover and confirm the differences between general IS and AI projects. However, the anecdotal evidence clearly indicates that AI projects do differ substantially from other projects in their goals and approaches. Considering these points, it can be assumed that not only AI projects but also the associated failure factors differ from traditional IS/IT projects and therefore need to be considered separately.

2.2 Related work and research gaps

In our literature review, we have identified several prior studies that have investigated research questions related to failure of AI projects, as well as, success factors, challenges, adoption and organizational readiness regarding AI [6, 8, 20–23, 41]. Table 1 summarizes important characteristics of the related work.

Table 1. Related literature.

Source	Object of analysis	Context	Theme/ construct	Method
Baier et al. [22]	Deployment and operation of ML	General	Challenges	Literature review and interviews
Bauer et al. [6]	ML	SME	Enablers and success factors	Interviews
Jöhnk et al. [23]	AI	General	Readiness	Literature review and interviews
Weiner [8]	AI/data science projects	General	Failure	Anecdotal evidence
Hou [20]	AI projects	Banking	Failure	Anecdotal evidence
Rychtyckyj and Turski [21]	Deployment of AI systems	General	Success and failure	Anecdotal evidence
Ermakova et al. [7]	Data-driven projects	General	Failure	Interviews and questionnaire survey
Reis et al. [19]	AI projects	Healthcare	Failure	Single case study

The comparison of the studies' characteristics shows that there are slightly different objects of analysis (e.g. AI, ML or data-driven projects). At the same time, different themes or constructs have been investigated (e.g. challenges, enablers, failure). Due to the limited number of studies that directly provide answers to our research aim, i.e. explicitly deal with failure of AI projects, we have extended the scope of our literature review to include the above-mentioned related themes. Although these previous results may lead to some interesting first assumptions regarding our research question, it has to be kept in mind that these constructs are different from failure factors. In our understanding, a challenge is defined more broadly as any "hurdle, barrier, concern, or critique" [42] whereas a failure factor leads to actual failure of a project, as defined in the previous section. At the same time, it has to be clearly distinguished between success and failure factors, as a failure factor does not necessarily have to be a success factor and vice-versa.

Besides these papers that deal specifically with AI or ML, we have searched for relevant previous research from less directly related fields, such as Big Data Analytics and Digital Strategy (not included in Table 1). The respective results from seminal articles [43–47] will be discussed in the discussion section of this paper (section 5) in order to compare the similarities and differences between the different fields and discuss possible implications.

Five literature sources were identified that discuss success or failure of AI projects or data-driven projects [7, 8, 19–21]. Two papers [20, 21] summarize "experiences with the development and deployment of AI systems" [21] in a specific company in a practice-oriented format. As these papers have the character of a practice report and lack scientific methodology, as well as, robust research design, they were excluded from a detailed discussion in this paper. In a similar vein, the book by Weiner [8] does provide interesting narratives about failure AI projects, but was excluded from our further analysis as it is not grounded on peer-reviewed research. Reis et al. [19] have conducted research into AI project failure based on a case study in the healthcare sector. The case study is based on a project in a large hospital to introduce a cognitive agent that was intended to assist physicians in their daily work and interact with patients. In this specific project, user resistance was identified as the reason for project failure. Consequently, the authors have conducted detailed analyses of the underlying causes of the non-acceptance and provide recommendations to overcome user resistance. Ermakova et al. [7] use a mixed methods approach to develop and administer an online survey with a sample of 112 experts. The focus of their research is "data-driven projects", i.e. data science in general, including AI and ML. In their approach, they do not distinguish between challenges and failure factors. Instead, the participants were asked to evaluate the perceived impact of challenges on the non-success of projects. Thus, the authors were able to make statements about the criticality of the challenges, i.e. their impact on failure.

Regarding the studies that do not directly regard failure, but related themes, Baier et al. [22] used interviews to analyze challenges particular to the deployment and operation of machine learning models. Another study focuses on general challenges in AI projects in the context of SMEs. To do so, Bauer et al. [6] correlate the identified success factors and challenges to the size and maturity of the companies. The data is collected using a survey approach with mainly CXOs or managing directors of SMEs. Jöhnk et al. [23] focus on AI readiness of companies. The authors collected data with semi-structured interviews focusing on factors that determine the readiness of companies in regard to AI.

Hence, we have selected the remaining fiver papers [6, 7, 19, 22, 23] for a detailed analysis as they were most relevant for the aim of this paper. For a better overview, the factors from these three studies are summarized in Table 2. The listed factors are abstracted to categories and may contain more than one individual factor presented in the studies, as well as, changes in wording due to the mapping to a uniform terminology. Additionally, the underlying theme or construct (challenges, failure factors etc.) was again mentioned in Table 2 to highlight the need for a careful interpretation when comparing the studies.

Table 2. Factors identified in previous studies.

Source	**Baier et al. [22]**	**Bauer et al. [6]**	**Ermakova et al. [7]**	**Jöhnk et al. [23]**	**Reis et al. [19]**
Theme/ construct	Challenges	Enablers and success factors	Challenges / failure factors	Readiness factors	Failure factors
Data	X	X	X	X	
Know how	X	X	X	X	
Infrastructure	X	X	X	X	
Project Management		X	X	X	
Communication	X		X	X	
Customer relation	X		X	X	
Acceptance	X		X		X
Ethics & Legal Issues	X		X	X	
Commitment		X		X	
Result validation	X	X			
Business Impact	X			X	
User friendliness	X				
Deployment	X				
Security			X		
Budget			X		

As shown in the table, some categories are mentioned by four out of five studies. One of the most prominent categories that was mentioned numerous times is *data*. As data is seen as the fuel for AI [41], it is not surprising that factors such as data quality, availability and governance are mentioned as important factors [6, 7, 22, 23]. In this category of our overview, we have summarized several factors that are dealt with in detail by the previous studies. For instance, Jöhnk et al. [23] list data flow as an interesting factor besides commonly mentioned factors like data availability, data accessibility and data quality. According to the authors, a good data flow enables AI professionals to “move data from its source to its use” by means of extract-transform-load processes, as well as, data pipelines and data streams. The work by Baier et al. [22] points toward important challenges in the area of data, such as imbalanced data or encrypted training data. Another factor that has been mentioned by most of the studies is *know how*. In the context of SME, Bauer et al. [6] see a lack of dedicated ML experts as an important size-related challenge experienced by this type of companies. At the same time, they state that an existing team in the area of business intelligence or data science can be an important success factor when it comes to the adoption of ML. This might be related to the fact that the use of ML models can be very challenging for employees that do not have a background in the field of data science. For this reason, other authors emphasize the need for user-friendly tools that enable non-technical employees to apply ML models [22]. Besides the obvious requirement of having a certain expertise in order to be able to work with AI-related technologies, an "AI awareness" also helps employees to have adequate expectations toward AI [23]. While many researchers focus on technical knowledge that is necessary for successful AI projects, one paper [22] also notices that domain knowledge can be crucial and can have an important implication on data quality. Ermakova et al. [7] see both, soft skills and hard skill, as important challenges for data-driven projects. Furthermore, different *infrastructure*-related factors seem to be a relevant challenges, for example regarding computational power. Depending on whether the necessary infrastructure is available in-house, time-consuming and complex investment decisions have to be made. However, due to cloud technologies, this problem can be mitigated [6, 22]. Factors that were mentioned by three of the studies are *communication, customer relation, acceptance*, as well as, *ethics and legal issues*.

While these prior studies provide seminal findings for their respective research aims, a number of research gaps remain with regard to the specific aim of our study. First, it has to be clearly stated that the results are not completely transferrable to our aim. As can be seen from Table 2, most of the prior studies have regarded different themes than failure. As previously stated, challenges, readiness factors or success factors can only be an approximation of failure factors. The study by Reis et al. [19] does deal with project failure. However, the research has been conducted in the specific context of healthcare and focuses on non-acceptance by users as one failure factor. The research by Ermakova et al. [7] does analyze failure in a more general context and is the most similar to this study. However, their object of analysis is data-driven projects as opposed to AI projects. It remains unclear, whether there is a significant difference in the definition that will have an impact on the results. Second, due to the limited number of studies and the mixed results, further research to corroborate the findings is necessary. The table shows that not all categories are mentioned in every study. The main reason for this might be the different focus of the studies. For example, Bauer et al. [6] focus on different company sizes while Jöhnk et al. [23] analyze readiness factors. Therefore, it can be concluded that collecting new empirical data for the specific question of AI project failure in a general context is clearly necessary. Finally, a comparison and synthesis of the related studies is required. The literature review in this section makes a first step toward an integrated discussion of the different papers.

3. Research design

For this study, a qualitative research design based on semi-structured expert interviews was chosen [24, 25]. Interviews are a common method in the IS discipline and have also been used as a method in prior related work (e.g. [6, 23]). The rationale behind choosing a qualitative methodology for this study is that the purpose of our study was to identify factors, as opposed to quantitatively testing them. In order to ensure the rigor of the qualitative research process, several measures were implemented [48]. These include critical discussion and reflection of methods and results throughout the different phases of the research process, as well as, redundant data analysis by different members of the team of authors, in order to minimize subjectivity and bias.

Following the recommendations for semi-structured interviews from the literature [49], we have developed an interview guide that consisted of a number of predetermined questions. However, the interviewer was also allowed to change the wording of questions, make clarifications and probe beyond the answers to the questions. As suggested by the methods literature, we have considered the objectives of our research, the type of data we were aiming to collect, as well as, conceptual areas from the literature review in the development of the instrument. Following common themes from the literature, we also included questions about challenges and success factors, besides our main concern, the question about project failure. Additionally, general introductory questions about project experiences and use cases were asked. The reason for this broad set of question was to stimulate an open discussion that will generate many aspects and ideas to be further discussed between the interviewer and respondents. However, for our framework of failure factors, as presented in the results section of this article, statements of the respondents were only considered if they explicitly referred to project failure. This was important in order to clearly distinguish between failure and challenges, as well as, between failure and success factors. All other statements were discarded for the purpose of this study in order to be precise in the measurement of the concept of failure.

To select the interview candidates, we have applied purposeful sampling in order to collect information-rich data that will help to illuminate the research questions [50]. In the selection of the candidates, our main focus was to include a diverse range of respondents in order to be able to identify as many relevant determinants of project failure as possible. Therefore, the sample includes AI experts that have heterogeneous professional backgrounds in terms of industry and company sizes, but also career levels and roles in their organization (see Table 3). In the sample, several candidates from consulting and software development firms are included that have worked in projects with different companies. Such experts are of less interest as an single case, but rather represent a more comprehensive source of knowledge based on cases in many firms. Hence, we were able to obtain sufficient data with a relatively low number of interviews. Following the concept of data saturation [51], we have not pre-determined a sample size. Instead, conducting new interviews was discontinued at a point when no new concepts had emerged from the data anymore. All of the interviews were conducted as audio or video calls between January and February 2021 by one of our authors, except one interview that was delivered in written form.

Table 3. Interview candidates.

#	Industry	Position	Focus/ expertise
1	Plant engineering	Team leader	Robotics and visual recognition
2	Software development	Founder and CEO	Visual recognition
3	Consulting	Senior consultant	AI in general
4	Software development	Developer	Natural Language Processing
5	Automotive	Development engineer	AI in sensor fusion
6	Automotive	Middle management	Driver assistance systems

With the consent of the participants, the interviews were recorded and subsequently transcribed using AI-based speech recognition. Subsequently, we used an inductive approach based on established methods for the analysis of qualitative data [25, 26] to derive a hierarchical coding structure. In a first step, the transcripts were inductively coded in an open coding approach. Finally, the codes were aggregated to several levels of higher-level categories based on their similarity in order to derive the factors presented in the results section of this paper. In order to avoid subjectivity, this analysis was first done independently by all authors and revised several times in an iterative process, before the consolidated version was finalized.

4. Results

4.1 Overview

Using the data from the interviews, a total of 12 factors that can lead to failure of AI projects were identified [18]. Based on our inductive method, these factors were further aggregated into the following five categories: Unrealistic expectations, use case related issues, organizational constraints, lack of key resources and technological issues (see Table 4).

Table 4. Categories and factors identified in the interviews.

Category	Factor
Unrealistic expectations	Misunderstanding of AI capabilities
	Thinking too big
Use case related issues	Missing value or cost-benefit ratio
	Complexity
	Low error tolerance

Category	Factor
Organizational constraints	Budget too low
	Regulations
Lack of key resources	Lack of employees with expertise
	Data availability
Technological issues	Model instability
	Lack of transparency (black box problem)
	Possible result manipulation

4.2 Unrealistic expectations

Factors regarding the expectations of AI projects are summarized in the category unrealistic expectations. Stakeholders and project members are often not fully aware of AI capabilities. This can lead to misunderstandings about technologies to be used. AI projects are sometimes only entitled as AI but are, in fact, not using any AI-related technologies. As interviewee 1 states, “most people have no idea what AI is actually supposed to be”, resulting in the situation that large, rule-based systems with human-made, pre-defined rules are programmed that are not based on AI technology. The expert even goes so far as to say that these systems “are not AI, but fake”. Such projects can clearly be considered as failure since they do not really lead to AI adoption.

Another factor related to unrealistic expectations is “thinking too big”. If expectations rise and managers become overly ambitious, projects scopes are getting wider and wider, until it is mostly impossible to make the projects work due to the lack of focus. A more successful approach to AI adoption, according to one expert, would be to “think in small steps” instead, in order to incrementally develop workable solutions. The root cause of too big expectations might often be linked to “too large promises” (Interviewee 3) that have been made.

4.3 Use case related issues

In general, use case related issues can also lead to project failure. The adoption of AI is sometimes done without value-adding use cases. In order to achieve a return on investment, value has to be generated, for instance by automating tasks that have previously been done by humans. If there are no additional revenues or cost savings, only expenses to introduce and operate the AI system, these projects fail in the sense of not delivering any economic benefits. One interviewee even stated, that “most use cases do not provide any value” (Interviewee 2).

Another failure factor is the use case complexity. If the complexity of a project surpasses the capabilities of the internal development teams, project can be “impossible to accomplish” (Interviewee 1). This means that project expectations and capabilities need to be aligned to prevent failure.

In special use cases, like autonomous driving, low error tolerance can lead to project failure. These use cases rely on precise and correct predictions and results, as error can have fatal outcomes. In AI projects, since the fidelity of results is only achieved after the models have been created, projects must be started first to verify accuracy. If the targeted and required accuracy is not achieved, projects are often discontinued.

4.4 Organizational constraints

Factors in the category organizational constraints represent external impacts on projects from within the company or the environment. Projects involving AI often represent a risk due to the uncertainty of the outcome. Therefore, often insufficient resources are allocated, leading to premature termination as they are running out of budget. However, the fact that often too low budgets are assigned is not only due to a reluctancy to invest, but also due to enormous budget requirements of AI projects. The budgets and resources are not only used to hire experts, but also to pay for training data and the training itself. Especially acquiring labelled datasets can be very expensive, as the generation of these datasets often requires a lot of human work in the first place. When these data are subsequently used to train AI models, also this next step is a huge effort, according to Interviewee 6. Interestingly, it was also mentioned that cost for hardware is not a relevant factor, as required machines or devices have become relatively inexpensive.

Additionally, regulations, internal or external, can cause issues for AI projects. One interviewee said, that there were “bureaucratic hurdles to even only attach a Raspberry Pi to an industrial machine” (Interviewee 2). However, the extent of this factor presumably depends on the country and company.

4.5 Key resources

Key resources, or the lack of those, were often described as a major influence on AI project failure. Three of the interviewees said that the lack of expertise was a key reason for the failure of AI projects. For example, Interviewee 6 mentioned that projects “sometimes fail because of the competencies of the employees, to be honest”. This problem can be related to other issues, like low budgets, as one interviewee mentioned: “If you put the wrong person, a person without enough knowledge, on an AI project, it is possible that the budget gets blown without any outcomes” (Interviewee 1).

As AI models strongly depend on the quantity and quality of training data, data availability is a factor that influences the project outcome. As an interviewee from the automotive sector mentioned, AI projects fail because correctly labelled training data is often not available or too expensive. This factors is “maybe even the most important one”, according to Interviewee 6’s opinion.

4.6 Technology

The technology itself is also a factor that can lead to project failure. Although Interviewee 2 mentioned that the technical implementation is usually not a reason for project failure in his context, several other interviewees did mention technology-related issues that can be critical.

One mentioned aspect is model instability. Companies rely on consistent results when it comes to AI algorithms. As the algorithms and systems are updated, there is “no guarantee that the systems work exactly like the last one and gives the same results” (Interviewee 4). This unpredictable behavior can lead to the termination of projects.

Furthermore, AI algorithms lack transparency as of how the algorithms ended up getting the result. This issue is especially relevant for results of neural networks and referred to as the so-called black box problem.

Furthermore, models can be manipulated to produce different results, e.g. if street signs are manipulated with stickers, there might be a wrong result interpreting it. The possible error introduced by manipulation can be too high to safely use the AI, depending on the context.

5. Discussion

Our results show that there are a variety of factors that can lead to failure of AI projects. A closer look at the factors reveals some interesting insights. First, it can be seen that technological issues can be one reason for failure. However, the statements of the candidates have shown that failure often seems to occur because of non-technical factors such as false expectations or lack of resources. Especially the lack of expertise or competent employees was emphasized by several interview candidates. Second, many factors or their detailed characteristics can hardly be anticipated before the start of an AI project and therefore cannot be appropriately considered in the planning of such a project. This can be observed, for example, in the factor *possible result manipulation.* At the beginning of a project, it is impossible to predict all possible ways how a result can be manipulated. Other factors, like the actual *complexity* of a use case or *model instability* can be equally difficult to estimate or anticipate. Therefore, it seems difficult to completely avoid possible project failure due to such reasons or to manage these risks as they often only emerge in the course of the project. On the other hand, some of the factors can be anticipated and managed in advance. For example, the needed know how for an AI project can be evaluated and actions can be taken. Furthermore, it can be checked if sufficient data is available to start an AI project.

An important contribution of our research is to distinguish failure factors from related constructs that have been discussed in prior work, such as challenges, readiness factors or success factors. By comparing our results (Table 4) with the prior results from related work (Table 2), we are able to draw the following conclusion: The factors *know how* [6, 7, 22, 23], *business impact* [22, 23] and *result validation* [6, 22] can be confirmed as being not only a challenge, but indeed critical for AI failure. Also *data* is a critical factor, when it comes to availability of suitable data. It can thus be seen that some of the already known challenges can also be concrete reasons for failure. On the other hand, some prominent factors from previous studies seem to be not as important for failure. These include the factors infrastructure, communication, deployment, user friendliness and customer relation. A possible explanation for the lower relevance with regard to failure is that these factors might indeed be relevant challenges in AI projects, but problems can be resolved if they occur and thus do not lead to project failure. For example, in the case of infrastructure, it is likely that problems related to this category can be resolved by investing in new on-site infrastructure or using cloud-based solutions.

Our research has additionally uncovered factors that have not been previously identified as failure factors or related constructs. These include *unrealistic expectations* and the specific *technological issues* of model instability and possible result manipulation. Overall, it can be summarized, that our study partially confirms prior results and also contributes new failure factors to the body of knowledge. Especially, regarding the prior study that is most similar to ours in terms of the research question [7], it seems that the results complement each other. However, due to the different methodology, measurement and classification of the factors, it is difficult to directly compare the results.

While we already have outlined previous related studies in the context of AI in section 2.2, it is also interesting to compare and synthesize our results with further findings from a broader context in order to discuss similarities and differences between different, related fields. Before the individual factors will be discussed in the subsequent paragraph, an overview of related context, as well as, seminal papers from the respective fields, is given. The first related context is the formulation of digital business strategies. For example, Holotiuk and Beimborn [43] have developed their Digital Business Strategy Framework based on a review of industry reports on digitalization. They have derived 40 critical success factors that are sorted into eight dimensions: sales and customer experience, culture and leadership, capabilities and HR competencies, foresight and vision, data and IT, operations and organization. Schuler and Schlegel [45] present a framework for corporate AI strategy formulation based on a systematic literature review that is supposed to outline important considerations when approaching AI adoption in a holistic approach. Based on inductive coding of factors extracted from the literature, they state that companies need to think about their capabilities, use cases, data, infrastructure and organization, as well as, ethical/legal constraints and managerial processes. The second stream of literature that can be considered as a related context is Big Data, being defined as data having high volume, velocity and variety, coming from different sources such as social media and video [47]. In his seminal paper, Watson [47] outlines success factors that organizations should consider in order to exploit the potential of big data analytics. According to the author, the factors include “a clear business need, strong committed sponsorship, alignment between the business and IT strategies, a fact-based decision-making culture, a strong data infrastructure, the right analytical tools, and people skilled in the use of analytics” [47]. In a similar vein, Saltz and Shamshurin [44] discuss key factors for a project’s success in the context of Big Data team process methodologies. They find a large number of success factors that are categorized into the categories data, governance, process, objectives, team and tools. Finally, Phillips-Wren and Hoskisson [46] have conducted case study research in order to identify management challenges when it comes to formulating a big data strategy in the context of customer relationship management (CRM) in mid-sized hospitality industry firms. According to their results, the dimensions customer, CRM process, organizational alignment and CRM outputs need to be considered. They also identify common challenges such as inconsistent and unstandardized data, relevant data not known, leadership, finding people with relevant skills.

Comparing these results from a broader context and synthesizing them with our own research, it turns out that especially two factors that we have summarized as “key resources” in our research seem to be universally important, as they can be found in all of the studies in related fields: first, *employees with relevant skills*, and second, *data-related factors*.

(1) Employees with relevant skills: When it comes to digital business strategy, capabilities and competencies that will be required in the future, do not only encompass technological skills, but also the capability to redesign value chains and business models [43]. In the context of Big Data, Phillips-Wren and Hoskisson [46] explain the necessity to combine domain knowledge with analytical skills in order to provide business insights and improve decision-making. Not only on individual employee level, but also when it comes to team work, multidisciplinarity is stressed as a success factor in other studies [44]. Watson [47] states that different types of big data analytics users need have different roles which require different skillsets. On one end of the continuum, there are end users that need to have an understanding of the data's business impact without having to know the detailed functionality of algorithms. On the other hand, there are highly-trained data scientists who search for patterns in the data [47]. Despite the obvious importance of employees' skills and competencies, according to some authors, top managers in many firms have "not yet worked out strategies for recruiting and training the talent needed to get the most value from intelligent systems." [52]. It is therefore recommended that managers identify employees who are "both willing and able to collaborate with smart machines" [52]. An interesting question with regard to hiring and training is whether existing internal employees that become obsolete due to digital transformation can be reskilled and trained into highly-required digital profiles, or whether these skills need to be hired externally. Based on an analysis of job profiles in the context of robotic process automation (RPA), one study highly doubts the reskilling hypothesis due to the specific nature of the technical skills that are required in this technology [53], which certainly can also be transferred to the field of AI. Other authors [23] see upskilling as an important organizational necessity in order to enable staff to develop new AI-related skills.

(2) Data-related factors: In the context of Big Data, the literature highlights the importance of a strong data infrastructure: "When a strong data infrastructure is in place, applications can often be developed in days. Without a strong data infrastructure, applications may never be completed." [47]. In his article, Watson [47] discusses different relevant technological developments that have taken place in recent years, including CPU improvements, in-database analytics and columnar databases. Other authors focus less on the technical infrastructure and more on the data itself. Based on their case study in the hospitality sector, Phillips-Wren and Hoskisson [46] report day-to-day challenges when dealing with data, for example that users are not aware of the original source of data that is delivered by the IT department which leads to trust issues as these data are often also inconsistent. Several authors [43, 47] suggest using data and information from one central source in order to rely on one "single version (or source) of the truth for decision support data" [47]. Finally, further success factors related to data that have been identified in prior research are data quality management and ownership, as well as, data integration and security [44].

This discussion shows that there are indeed both, similarities, and differences between our results and the related prior research, as well as, the AI field and related contexts such as Big Data. The main similarity is certainly the importance of the general themes of people and know how, as well as, data-related factors. However, when it comes to the data-related factors, it has to be acknowledged that this is a very broad theme and the factors indeed do differ substantially when having a closer look. As previously noted, general aspects of data infrastructure and data management were emphasized in the literature in both the AI and Big Data field. However, our research has shown that these aspects are not critical to failure. Instead, the mere availability of labelled training data for the AI models is a key constraint. In a similar vein, some of the categories we have identified in our research are highly specific to AI. These include for example the problem of unrealistic expectations based on misunderstanding of AI capabilities and thinking too big. But also use case related issues such as the high complexity in AI projects, as well as, domain-specific technological issues including model instability and the black box problem, are specific to AI. On the other hand, it might be possible to transfer some findings from related fields to our context in order to give more specific guidance for the proactive management of failure factors. For example, regarding the management of skills and competencies in the firm by hiring and training employees, the existing body of literature from related contexts can be consulted to get further advise.

6. Conclusions

6.1 Summary and contribution to knowledge

The evidence from this study suggests that there are several factors that can lead to success or failure of AI projects. On the one hand, these factors include technological issues such as model instability or the black box problem. On the other hand, especially non-technological factors seem to play an important role, including misunderstanding of AI capabilities, or missing economic value of projects. Moreover, the lack of two types of key resources, employees with relevant expertise and adequate data, often lead to project failure. A comparison with prior studies from the context of AI and related field shows that these two key resources seem to common challenges in AI projects, as well as, Big Data and digital strategy contexts.

Our research makes a number of important contributions to the field. First of all, our research has underlined the importance of distinguishing between general challenges and failure factors of AI projects. Based on new empirical data, our study contributes to knowledge by making this distinction for previously known factors. For example, having adequate infrastructure to develop and employ AI applications, which has previously been identified as a challenge, is not a critical factor for project failure. Second, our new empirical data contributes to knowledge by identifying new factors such as unrealistic expectations. Finally, our article has synthesized and compared prior results from related work, as well as, embedded the results into the wider context of digital strategy and Big Data.

6.2 Implications

The findings of our research have important managerial implications for organizations that are planning to adopt AI. While some of the failure factors are hard to anticipate and manage, the relevance of other typical factors for a particular organization can easily be clarified in advance. Managers are advised to have clear and honest look at their organizations' capabilities and resources, as well as, their own expectations and understanding of AI, before starting an AI project. It is also recommended to conduct a systematic feasibility analysis before starting specific AI projects. After an evaluation of potential critical risks, appropriate measures can be taken to mitigate these risks.

If the risk of failure is estimated as too high, an honest acknowledgement of the organization's lack of AI readiness, combined with a mid-term roadmap to improve the capabilities, might be a better advise than rushing into disaster with one's eyes open. In order to improve their organizations' readiness for AI, especially the two key resources employees and data should be developed in the medium term by investing in upskilling and recruiting of high-profile employees, as well as, data infrastructure and management.

6.3 Limitations and further research

Our work may have some limitations. Given the qualitative approach and small sample size of our study, caution must be used in generalizing the findings or transferring them to other contexts. Additionally, due to the dynamic nature of the topic, we regard the results as a snapshot of current failure factors that has been taken in a certain moment and may evolve over time. Therefore, the results might have to be updated on an ongoing basis. However, the discussion of this study's results has shown that the results are overall plausible when comparing them to related studies which underlines the trustworthiness and credibility of our research.

Despite the limitations, we believe that our work lays the ground for further research in this area. We propose that further quantitative studies should be conducted to corroborate our findings and generate representative results based on the categories and factors identified in this study. For example, survey research design can be used to generate quantitative results on project failure, taking our identified factors, supplemented by other similar studies [7] as a basis for the design of the survey instrument. Additionally, future projects could deal with the question, how project failure can be avoided by systematically evaluating the risk factors found in this study.

References

[1] S. Russell and P. Norvig, *Artificial Intelligence: A Modern Approach (3rd Edition)*: Pearson, 2009.

[2] P. Buxmann and H. Schmidt, *Künstliche Intelligenz: Mit Algorithmen zum wirtschaftlichen Erfolg*. Berlin: Springer Gabler, 2019.

[3] C. Pennachin and B. Goertzel, “Contemporary Approaches to Artificial General Intelligence,” in B. Goertzel and C. Pennachin, Eds.: Springer, 2007, 1–30.

[4] V. C. Müller and N. Bostrom, “Future Progress in Artificial Intelligence: A Survey of Expert Opinion,” in *Fundamental Issues of Artificial Intelligence*, V. C. Müller, Ed., Cham: Springer International Publishing, 2016, pp. 555–572.

[5] J. Bughin *et al.*, *Artificial intelligence: the next digital frontier?* [Online]. Available: https://www.mckinsey.com/~/media/McKinsey/Industries/Advanced%20Electronics/Our%20Insights/How%20artificial%20intelligence%20can%20deliver%20real%20value%20to%20companies/MGI-Artificial-Intelligence-Discussion-paper.ashx

[6] M. Bauer, C. van Dinther, and D. Kiefer, “Machine Learning in SME: An Empirical Study on Enablers and Success Factors,” in *AMCIS 2020 Proceedings*, 2020, pp. 1–10.

[7] T. Ermakova, J. Blume, B. Fabian, E. Fomenko, M. Berlin, and M. Hauswirth, “Beyond the hype: why do data-driven projects fail?,” in *Proceedings of the 54th Hawaii International Conference on System Sciences*, p. 5081.

[8] J. Weiner, *Why AI/Data Science Projects Fail*: Morgan & Claypool Publishers, 2020.

[9] Y. K. Dwivedi, K. K. Kapoor, M. D. Williams, and J. Williams, “RFID systems in libraries: An empirical examination of factors affecting system use and user satisfaction,” *International Journal of Information Management*, vol. 33, no. 2, pp. 367–377, 2013.

[10] K. R. T. Larsen, “A Taxonomy of Antecedents of Information Systems Success: Variable Analysis Studies,” *Journal of Management Information Systems*, vol. 20, no. 2, pp. 169–246, 2003, doi: 10.1080/07421222.2003.11045768.

[11] D. Avison and D. Wilson, “It Failure and the Collapse of One.Tel,” in *Information Systems: The e-Business Challenge*, R. Traunmüller, Ed., Boston, MA: Springer US, 2002, pp. 31–46.

[12] P. Beynon-Davies, “Information systems ‘failure’: the case of the London Ambulance Service’s Computer Aided Despatch project,” *European Journal of Information Systems*, vol. 4, no. 3, pp. 171–184, 1995.

[13] W. Bussen and M. D. Myers, “Executive information system failure: a New Zealand case study,” *Journal of Information Technology*, vol. 12, no. 2, pp. 145–153, 1997, doi: 10.1080/026839697345152.

[14] G. Gable, D. Sedera, and T. Chan, “Re-Conceptualizing Information System Success: The IS-Impact Measurement Model,” *Journal of the Association of Information Systems*, vol. 9, no. 7, pp. 1–18, 2008.

[15] C. Barth and S. Koch, “Critical success factors in ERP upgrade projects,” *Industrial Management & Data Systems*, vol. 119, no. 3, pp. 656–675, 2019, doi: 10.1108/IMDS-01-2018-0016.

[16] S. Finney and M. Corbett, “ERP implementation: a compilation and analysis of critical success factors,” *Business Process Management Journal*, vol. 13, no. 3, pp. 329–347, 2007, doi: 10.1108/14637150710752272.

[17] C. Holland and B. Light, “A Critical Success Factors Model for ERP Implementation,” *IEEE Software*, vol. 16, no. 3, pp. 30–36, 1999, doi: 10.1109/52.765784.

[18] J. Westenberger, K. Schuler, and D. Schlegel, “Failure of AI projects: understanding the critical factors,” *Procedia Computer Science*, vol. 196, pp. 69–76, 2022, doi: 10.1016/j.procs.2021.11.074.

[19] L. Reis, C. Maier, J. Mattke, M. Creutzenberg, and T. Weitzel, “Addressing User Resistance Would Have Prevented a Healthcare AI Project Failure,” *MIS Quarterly Executive*, vol. 19, no. 4, 2020.

[20] F. Hou, “Analyzing the Performance of the Unmanned Bank to Explore the Failure Reasons for AI Projects,” in *AIAM2020 - 2nd International Conference on Artificial Intelligence and Advanced Manufacture*, 2020.

[21] N. Rychtyckyj and A. Turski, “Reasons for success (and failure) in the development and deployment of AI systems,” in *AAAI 2008 workshop on What Went Wrong and Why*, 2008.

[22] L. Baier, F. Jöhren, and S. Seebacher, “Challenges in the Deployment and Operation of Machine Learning in Practice,” in *27th European Conference on Information Systems (ECIS), Stockholm & Uppsala, Sweden*, 2019.

[23] J. Jöhnk, M. Weißert, and K. Wyrtki, “Ready or Not, AI Comes— An Interview Study of Organizational AI Readiness Factors,” *Bus Inf Syst Eng*, vol. 63, no. 1, pp. 5–20, 2021, doi: 10.1007/s12599-020-00676-7.

[24] B. J. Oates, *Researching information systems and computing*. Los Angeles, London New Delhi, Singapore, Washington, DC: Sage, 2006. [Online]. Available: http://www.loc.gov/catdir/enhancements/fy0713/2005927557-d.html

[25] P. Mayring, *Qualitative Inhaltsanalyse: Grundlagen und Techniken*, 12th ed. Weinheim, Basel: Beltz, 2015.

[26] J. Saldaña, *The coding manual for qualitative researchers*. Los Angeles: Sage, 2009.

[27] S. Petter, W. DeLone, and E. R. McLean, “Information Systems Success: The Quest for the Independent Variables,” *Journal of Management Information Systems*, vol. 29, no. 4, pp. 7–62, 2013, doi: 10.2753/MIS0742-1222290401.

[28] K. Ewusi-Mensah, *Software Development Failures*: MIT Press, 2003. [Online]. Available: https://books.google.de/books?id=cWde_yxJorEC

[29] W. DeLone and E. McLean, “Information Systems Success: The Quest for the Dependent Variable,” *Information Systems Research*, vol. 3, no. 4, pp. 60–95, 1992, doi: 10.1287/isre.3.1.60.

[30] P. B. Seddon, D. Staples, R. Patnayakuni, and M. J. Bowtell, “Dimensions of Information Systems Success,” *Communications of AIS*, vol. 2, no. 20, pp. 1–61, 1999.

[31] N. Urbach, S. Smolnik, and G. Riempp, “The State of Research on Information Systems Success,” *Bus Inf Syst Eng*, vol. 1, no. 4, pp. 315–325, 2009, doi: 10.1007/s12599-009-0059-y.

[32] V. McKinney, K. Yoon, and F. Zahedi, “The Measurement of Web-Customer Satisfaction: An Expectation and Disconfirmation Approach,” *Information Systems Research*, vol. 13, no. 3, pp. 296–315, 2002, doi: 10.1287/isre.13.3.296.76.

[33] U. Kulkarni, S. Ravindran, and R. Freeze, “A Knowledge Management Success Model: Theoretical Development and Empirical Validation,” *Journal of Management Information Systems*, vol. 23, no. 3, pp. 309–347, 2007, doi: 10.2753/NUS0742-1222230311.

[34] J. Iivari, “An Empirical Test of the DeLone-McLean Model of Information System Success,” *ACM SIGMIS Database*, vol. 36, no. 2, pp. 8–27, 2005, doi: 10.1145/1066149.1066152.

[35] C. Iriarte and S. Bayona, “IT projects success factors: a literature review,” *International Journal of Information Systems and Project Management*, vol. 8, no. 2, pp. 49–78, 2020, doi: 10.12821/ijispm080203.

[36] H. C. Lucas, *Why Information Systems Fail*: Columbia University Press, 1975.

[37] C. Sauer, *Why Information Systems Fail: A Case Study Approach*, Alfred Waller, 1993.

[38] R. Nelson, “IT Project Management: Infamous Failures, Classic Mistakes, and Best Practices,” *MIS Quarterly Executive*, vol. 6, no. 2, pp. 67–78, 2007.

[39] K. Haller, *Managing AI in the Enterprise: Succeeding with AI Projects and MLOps to Build Sustainable AI Organizations*, 1st ed. Erscheinungsort nicht ermittelbar, Boston, MA: Apress; Safari, 2022.

[40] V. Mehta, *How AI/Machine Learning projects are different from traditional IT/Software Development.* [Online]. Available: https://www.linkedin.com/pulse/how-aimachine-learning-projects-different-from-itsoftware-mehta/

[41] L. Pumplun, C. Tauchert, and M. Heidt, "A New Organizational Chassis for Artificial Intelligence-Exploring Organizational Readiness Factors," in *27th European Conference on Information Systems (ECIS), Stockholm & Uppsala, Sweden*, 2019, pp. 1–15.

[42] L. Kostyak, D. M. Shaw, B. Elger, and B. Annaheim, "A means of improving public health in low- and middle-income countries? Benefits and challenges of international public–private partnerships," *Public Health*, vol. 149, pp. 120–129, 2017, doi: 10.1016/j.puhe.2017.03.005.

[43] F. Holotiuk and D. Beimborn, "Critical Success Factors of Digital Business Strategy," *13th International Conference on Wirtschaftsinformatik (WI), St. Gallen*, 2017.

[44] J. S. Saltz and I. Shamshurin, "Big data team process methodologies: A literature review and the identification of key factors for a project's success," in *2016 IEEE International Conference on Big Data (Big Data)*, 2016, pp. 2872–2879.

[45] K. Schuler and D. Schlegel, "A Framework for Corporate Artificial Intelligence Strategy," in *Digital Economy. Emerging Technologies and Business Innovation*, 2021, pp. 121–133.

[46] G. Phillips-Wren and A. Hoskisson, "An analytical journey towards big data," *Journal of Decision Systems*, vol. 24, no. 1, pp. 87–102, 2015, doi: 10.1080/12460125.2015.994333.

[47] H. J. Watson, "Tutorial: Big Data Analytics: Concepts, Technologies, and Applications," *Communications of the Association for Information Systems*, vol. 34, pp. 1247–1268, 2014, doi: 10.17705/1CAIS.03465.

[48] K. Saumure and L. M. Given, "Rigor in Qualitative Research," in *The Sage encyclopedia of qualitative research methods*, L. M. Given, Ed., Los Angeles Calif.: Sage Publications, 2008.

[49] B. L. Berg and H. Lune, *Qualitative research methods for the social sciences*. Boston: Pearson, 2017.

[50] M. Q. Patton, *Qualitative research & evaluation methods: Integrating theory and practice*. Thousand Oaks California: SAGE Publications Inc, 2015.

[51] J. M. Corbin and A. L. Strauss, *Basics of qualitative research: Techniques and procedures for developing grounded theory*, 3rd ed. Los Angeles, California: Sage Publ, 2008.

[52] V. Kolbjørnsrud, R. Amico, and R. Thomas, "Partnering with AI: how organizations can win over skeptical managers," *Strategy & Leadership*, vol. 45, no. 1, pp. 37–43, 2017, doi: 10.1108/SL-12-2016-0085.

[53] D. Schlegel and P. Kraus, "Skills and competencies for digital transformation – a critical analysis in the context of robotic process automation," *International Journal of Organizational Analysis*, vol. 31 no. 3, pp. 804–822, 2023, doi: 10.1108/IJOA-04-2021-2707.

Biographical notes

Dennis Schlegel
Dennis Schlegel is a Professor of Business Informatics (Information Systems) at Reutlingen University, Germany. His current research interest lies in the business and societal implications of emerging information technologies. After graduating with a PhD from Leeds Beckett University in 2013, Dennis Schlegel has gained many years of practical experience at a Big Four consultancy firm, most recently at Senior Manager level, before returning to academia.

Kajetan Schuler
Kajetan Schuler is a consultant in data strategy and data science at Exxeta AG. Before, he graduated with a master's degree in Business Informatics (Information Systems) at Reutlingen University, Germany. His research interest lies in artificial intelligence and how to adopt it successfully on a project as well as strategic level in companies.

Jens Westenberger
Jens Westenberger is a product manager in the field of practical artificial intelligence. Managing not only the tech aspect but also external teams of service providers from four different companies with over 20 employees, he deals with project risks on a regular basis. His current research interests include risk minimization in high technology projects.