AMCIS 2022 Proceedings | SIG - Adoption and Diffusion of IT

Aug 10th, 12:00 AM

Beyond MLOps: The Lifecycle of Machine Learning-based Solutions

Tomasz Marcin Mucha
Aalto University, tomasz.mucha@aalto.fi

Sijia Ma
Peking University, sjma@pku.edu.cn

Kaveh Abhari
San Diego State University, kabhari@sdsu.edu

Beyond MLOps: The Lifecycle of Machine Learning-based Solutions

Completed Research

Tomasz Mucha
Aalto University
tomasz.mucha@aalto.fi

Sijia Ma
Peking University
sjma@pku.edu.cn

Kaveh Abhari
San Diego State University
kabhari@sdsu.edu

Abstract

Organizations increasingly use machine learning (ML) to transform their operations. The technical complexity and unique challenges of ML lead to the emergence of ML operations (MLOps) practices. However, the research on MLOps is in its infancy and is fragmented across disciplines. We extend and integrate these conversations by developing a framework that accounts for the technical, organizational, behavioral, and temporal aspects of the overarching ML-based solution lifecycle. We identify the key components of ML-based solution lifecycle and their configuration through an in-depth study of Finland's Artificial Intelligence Accelerator (FAIA) and follow-up semi-structured interviews with experts from multiple international organizations outside FAIA. This study contributes to the recent IS literature concerned with the sociotechnical aspects of ML. We bring new insights into the discussion on organizational learning, conjoined agency, and automation and augmentation. These insights extend and complement MLOps practices, thereby helping organizations better realize the potential of ML technology.

Keywords

Artificial intelligence, machine learning, design, machine learning operations (MLOps), development operations (DevOps), lifecycle.

Introduction

The present wave of artificial intelligence (AI) commercialization builds on the rapid progress of machine learning (ML) advancements (Berente et al. 2021). Organizations can now leverage ML-based solutions to capture and act on tacit knowledge, which used to be exclusively in the domain of humans or beyond their reach altogether (Brynjolfsson and Mitchell 2017). Despite its immense potential, the design and operational implementation of these solutions has proven to be challenging for organizations (John et al. 2021) despite the presence of external ML development tools. This translates to a bigger problem that many ML-based solutions fail to generate meaningful or sustainable impact (Ransbotham et al. 2020) since ML-based solutions face unique design and operational impediments throughout their lifecycle when compared with traditional information systems development practices (Lwakatare et al. 2020; Tamburri 2020). In response to these challenges organizations take different approaches to designing and managing ML-based solution lifecycle, typically by adopting DevOps practices (development-operations). These approaches to streamlining and standardization of ML lifecycle management are referred to as machine learning operations (MLOps). The research on MLOps is, however, in its infancy, despite great interest among practitioners. The scholarly work on MLOps design and operationalization is also fragmented across disciplines. Studies concerned with MLOps in computer science primarily deal with technical aspects (Paterson et al. 2021) and insufficiently account for organizational and social aspects. Hence, Information systems (IS) scholars recognize the need for a sociotechnical perspective on ML (Benbya et al. 2021; Berente et al. 2021; Lyytinen et al. 2020; Teodorescu et al. 2021). Yet they typically zoom in on narrower design problems such as fairness (Teodorescu et al. 2021; van den Broek et al. 2021), organizational learning

(Sturm, Gerlach, et al. 2021) or accountable implementation (Asatiani et al. 2021). This leads to fragmented insights into '*ML-based solution lifecycle*'.

Against this backdrop, we ask the following question: *What is the organizational-level lifecycle of ML-based solutions, and what are its key components?* To answer this question, we conducted an exploratory study with the goal of integrating disjoined views on the sociotechnical aspects of MLOps. In this particular paper, we concentrate on ML-based solutions impacting internal organizational processes and present the initial results of this ongoing study to urge IS scholars to pursue further research on this critical issue. This study contributes to the recent IS literature concerned with understanding the sociotechnical aspects (Sarker et al. 2019) of ML technology (e.g. Baird & Maruping 2021; Benbya et al. 2021; Berente et al. 2021; Murray et al. 2021; Raisch & Krakowski 2021). We also respond to Raisch and Krakowski's (2021) call for complexifying and theorizing ML-based solution design and development to match the reality of their operation and Lyytinen and colleagues' (2020) call for a sociotechnical perspective on ML-based solutions as novel forms of learning capabilities. We bring new insights into the future generation of MLOps practices in the context of organizational learning (Sturm, Gerlach, et al. 2021; Sturm, Koppe, et al. 2021), conjoined agency (Baird & Maruping 2021; Murray et al. 2021; Teodorescu et al. 2021), and automation and augmentation (Raisch & Krakowski 2021). These insights also translate directly into practical implications that extend and complement MLOps practices, thereby helping organizations better realize the potential of ML technology.

Theoretical Background

While smart agents enabled by advanced ML algorithms (ML-based solutions) can provide benefits to organizations, there are numerous obstacles to their design, seamless integration, and flawless operation (Ransbotham et al. 2020). There are three main reasons behind this claim. Firstly, these ML-based solutions are fundamentally different from conventional information systems (Lwakatare et al. 2020) due to a different development lifecycle (Ågerfalk 2020; Benbya et al. 2021; Berente et al. 2021; Lyytinen et al. 2020). Secondly, unlike typical information systems, they require sufficient data to be built, trained, and tested before being deployed, and to be maintained after integration (Makinen et al. 2021). Thirdly, these agents can act or react autonomously on behalf of humans (Baird & Maruping 2021). Therefore, the integration of these agents requires continuous evaluation and refinement as well as real-time feedback, especially when these agents share responsibilities with humans (Rai et al. 2019). Addressing this necessity, prior studies mainly focused on the development challenges associated with data, model training, or performance (e.g., Zhang et al. 2020). However, the implementation challenges of ML-based solutions go beyond training and testing ML models. Hence, researchers call for more human-centered approaches to the design, development, and operation of these solutions (Martínez-Fernández et al. 2021).

From DevOps to MLOps

DevOps is a popular practice for developing and operating modern information systems in an integrated fashion (Alnafessah et al. 2021; Gall and Pigni 2021; Moreschini et al. 2022). The increased adoption of ML-based solutions encourages researchers and practitioners alike to develop MLOps, as a new approach to DevOps to take full benefit from these solutions (Alla and Adari 2021; Goyal 2020; John et al. 2021; Makinen et al. 2021; Moreschini et al. 2022). MLOps can be characterized as either a DevOps cycle with ML components (Moreschini et al. 2022) or a DevOps cycle dedicated to ML-based solution development (Lwakatare et al. 2020; Martínez-Fernández et al. 2021). Regardless, the MLOps cycle includes both DevOps' development cycle (planning, coding, building, and testing) and its operation cycle (release, deploy, operate and monitor). The difference, however, is how different authors integrated ML in these two cycles. For example, Moreschini et al., (2022) introduced a parallel cycle of ML development with four components, planning, coding, debugging, and validation. Martínez-Fernández et al. (2021) proposed two new integrated practices for the development and evaluation of context-specific AI models. Similarly, Lwakatare et al., (2020) introduced two new cycles for data management and ML modeling. While these studies investigated different aspects of MLOps, they did not provide a clear picture of the integration and operation of ML-based solutions as an integral part of organizational processes (Colantoni et al. 2020). Furthermore, while these MLOps models theoretically deliver or optimize ML-based solutions for organizations, they are limited in explaining the human-machine dynamics, from delegation to negotiation.

MLOps Operational Limitations

Research showed that the operationalization of MLOps at scale requires the integration of the development process with the organizational processes that may be directly or indirectly affected by ML-based solutions (Makinen et al. 2021). However, MLOps practices do not sufficiently account for the possible interaction scenarios between human and non-human agents to realize the full potential of these solutions. Recent studies argued that the MLOps cycle, unlike typical DevOps, cannot be simply integrated with organizations due to various reasons, from the absence of a systematic ML pipeline (Moreschini et al. 2022) with context-specific solutions (Garcia et al. 2018) to the lack of knowledge on governing ML-based solutions (Lwakatare et al. 2020) and delegation mechanisms (Baird & Maruping 2021). Moreover, there are some sociotechnical challenges concerning the integration of MLOps such as establishing users' trust and facilitating self-adaptation (Martínez-Fernández et al. 2021) mainly due to limited empirical work investigating the human aspects of managing ML-based solutions in organizations (Asatiani et al. 2021; Keding 2021). MLOps also suffers from the typical DevOps limitations related to overlooking human factors in technology development, adoption, and diffusion (Riungu-Kalliosaari et al. 2016). Some of these limitations stem from the modularity of MLOps components. ML-based solutions are typically developed in an isolated environment (Tamburri 2020). This separation leads to unclear definitions of rules and responsibilities and insufficient assessment of integrability. This limitation may result in unsuccessful delegation, insufficient coordination, and unnecessary negotiations between human and smart agents with a negative impact on the overall operation (Asatiani et al. 2021). Furthermore, post-implementation, ML-based solutions need to be carefully monitored by human actors (Lyytinen et al. 2020) due to their error-prone nature (Asatiani et al. 2021). This monitoring and when necessary, the intervention processes require a significant amount of coordination and operational employees' involvement beyond MLOps' reach (Pääkkönen et al. 2020).

MLOps from a Sociotechnical Perspective

Recent studies argue that the notion of MLOps should be revisited from a sociotechnical perspective for mainly four reasons. Firstly, the development of ML-based solutions is not independent of the organizational stakeholders since the outcome of MLOps can radically change how these stakeholders perform different tasks and in turn how the organization fulfills its operational goals (Asatiani et al. 2021). Secondly, the operation of ML-based solutions requires an organizational commitment across multiple levels and functions (Raisch & Krakowski 2021). Thirdly, current MLOps models fall short in considering the temporality of ML-based solution integration at different levels of adoption and operation (Grønsund & Aanestad 2020; Strich et al. 2021; van den Broek et al. 2021). Lastly, ML-based solution development is not a linear workflow to follow but a transformational lifecycle to master (Ågerfalk 2020; Baird & Maruping 2021; Lyytinen et al. 2020; Raisch & Krakowski 2021). These studies altogether emphasize the importance of an overarching framework that systematically accounts for the technical, organizational, behavioral, and temporal aspects of ML-based solutions. To this end, this study is an attempt to offer a nuanced understanding of 'ML-based solution lifecycle', that helps the development and integration of ML technology in organizations with a larger and long-term impact in mind. While we are not able to address all MLOps' limitations, we can offer a new perspective to revisit MLOps. This new perspective is not solution/technology-specific, and therefore, it can serve as an integrative paradigm for future research.

Method

We conducted an in-depth qualitative study of ML-solution lifecycle in the context of large organizations through a combination of participatory observation of Finland's Artificial Intelligence Accelerator (FAIA) and follow-up semi-structured interviews with experts from multiple international organizations outside FAIA. FAIA was initiated and originally funded by the Finnish government in 2018 to facilitate collaboration and stimulate ML adoption within the participating companies while extracting the key lessons and informing a broader audience of organizations in Finland. Unlike start-up accelerators, FAIA focused its efforts on established organizations. Firms participating in FAIA included some of the largest Nordic companies, such as Elisa (telecom operator), Nordea (bank), Posti (Finnish national postal services), S-group (retail chain), Telia (telecom operator), YLE (Finnish national broadcasting company). Thus, companies participating in FAIA provided us a convenience sample representative of large organizations investing in ML across various industries in Finland. The organizations participating in FAIA formed several semi-formal groups (batches), typically with 4-8 members each. Each batch focused on specific types of ML-based solutions or ML-related practices. The batches met regularly throughout the acceleration period for six months. The role of FAIA's team was to facilitate and catalyze the collaboration by creating a community of practice with peer support and peer pressure. The data collection window covered the first 3.5 years of FAIA's activities (August 2018 to December 2021). The first author engaged in observation and carried out the field study by participating in weekly internal meetings of the FAIA team and workshops with companies participating in the accelerator. The first author also had unrestricted access to internal documentation of FAIA and was included in part of the email correspondence between the FAIA team and the participating companies. To complement this data, additional confirmatory interviews were conducted with non-FAIA affiliated companies, consulting firms, and researchers based in Europe, China, and the United States. The companies and questions selected for these additional interviews allowed us to verify findings in the context of large organizations outside Finland. To the extent possible, the meetings, workshops, and interviews (events) were recorded and transcribed (alternatively, notes were taken during the events). Most of the events were in English, but some were in Finnish or Chinese (transcribed and translated to English). Overall, our data covers 149 events with an approximate total length of 175 hours. As a screening mechanism during the interviews, we first discussed the informants' roles in their organizations and their involvement in ML-related activities. During our data collection, we focused on developing an understanding of the ML-solution lifecycle, its integral components, and their relationships.

Given the complexity of the phenomenon we studied, our approach followed an abductive reasoning process (Levallet et al. 2021) informed by engaged scholarship practice (Van de Ven 2007). This type of approach has been recommended for studying IS phenomena that are not "amendable to simple explanations" (Levallet et al. 2021) and understanding complex processes (Graebner et al. 2012). Furthermore, a high degree of engagement with practitioners and participatory research has been advocated for understanding sociotechnical aspects of ML-based solutions (Lyytinen et al. 2020). Overall, our abductive approach allowed us to cycle between empirical material and literature, as well as zoom in and out between levels of analysis, framework components, their relationships, and theoretical framings. Consequently, we were able to identify nuanced insights that complement a broader understanding of ML-based solution lifecycle.

Results

This section presents the main findings and empirical insights from our analysis. We focused on how ML-based solutions served the participating organizations, how these organizations managed the design, development, operation, and maintenance of ML-based solutions, and what were the challenges these organizations faced in this process. The results highlighted the role of management in ML development and operation. Despite differences in sizes and responsibilities, the <u>management team consistently played a pivotal role</u> in coordinating, updating, and monitoring the development process of ML-based solutions.

> "*Our hybrid model aims to achieve the best of a centralized and distributed organization. [Our] centrally coordinated Platform Team provides common capabilities to Service and Production Teams. Common capabilities refer to technical components and solutions that can be utilized by multiple teams in an organization or that require highly specialized expertise to produce. Good examples are the data models maintained by the Platform Team and the ready-made solutions on the infrastructure side. This approach supports both the autonomous work of the teams and a clear division of ownership and responsibility.*" ~Owner of Data Platforms and Technologies, Meeting notes

However, the <u>disconnect between the management and developers</u> was a challenge frequently reported.

> "*[...] this morning a colleague told me that the latest dashboards look very, very weird. [...] he went to R&D and found out that they changed the logic of how [data is] transmitted from the machines. [...] it's totally, totally blew out the logic on the other side. But that's the daily life.*" ~Director, Data Driven Services, Interview transcript.

Apart from the inherent need to continuously reflect on the ML design, development, and operation process, <u>ML organizations were also dynamic and evolving.</u>

"*Organizing [ML] is a balance between hierarchical structures and self-directed ways of working both have their pros and cons. Underlying this is the idea that it is a matter of balancing the extremes [...]. When you want to develop the artificial intelligence functions and capabilities of your own organization, you should look at the end of the year, for example, and think about what kind of organizational form you should move towards next.*" ~AI consultant, FAIA meeting field notes.

Furthermore, both management and development teams needed to actively engage with people from across the organization to collect feedback, as well as to contribute to business process renewal or reengineering.

"Make it a habit to spend a day now and then in the front lines. Experiencing problems firsthand is often the best way for getting a non-biased view on any issue. [...] If underlying process is rubbish, process should first be optimized to ensure valuable automation." ~Director of Data and Automation, Workshop field notes.

The results suggested that an ML-based solution lifecycle involved data, infrastructure, models, and other software artifacts that need to be managed, updated, and coordinated by the management team.

"*From my side of things, Digital Engineering, what we end up working with a lot is some of the infrastructure pieces that are needed to support data engineering or data scientists' work*" ~Digital Engineering Lead, Interview transcript.

However, most participating organizations were benefiting from centralized data services.

"*Everything what we collect, [whether] it comes from the equipment or from the business system, it goes to Amazon AWS data lake and it's accessible by the data scientist.*" ~Director, Data Driven Services, Interview transcript.

Irrespective of the level of maturity of the ML-related practices, ML technology development and operation encompassed multiple iterations to achieve the desired outcomes and higher levels of automation.

"*Once the [NLP-based contract analysis] tool achieves good enough performance level, [...] periodic retraining might be needed to adopt to new regulations, such as introduction of GDPR recently, or changes in contract drafting styles or standards.*" ~Corporate Legal Counsel, Workshop field notes.

The participants also emphasized the importance of improving the efficiency of ML-based solution developments by including the development teams in monitoring and maintaining the ML-based solutions.

"*It would be nice to move to a model like in [name of another company participating in the workshop] where Data Science Team manages the whole process, rather than needing to send tickets to fix some things.*" ~Business Development Manager, Robotics & AI, Workshop field notes.

The participants emphasized that ML technology was of little value by itself. ML technology could unlock value only when they aligned the technology with their business processes and strategic priorities.

"*The fact that there's AI in the middle of [solution name] is almost unimportant. The point is now [users/customers] they've got a capability they didn't have before, which we've made work by understanding the technicalities of AI, the way that the AI works, all the complexities of actually embedding it in the system, getting the human in the loop, putting in front of the customer in a way that makes sense to the customer.*" ~Digital Engineering Lead, Interview transcript.

The participants frequently noted that the understanding of business processes and users' needs, and preferences were the critical success factors of ML-based solution lifecycles. Yet, communication and getting to a common understanding within the organization were often challenging.

"*The management may be willing and mentally ready to start dealing with AI, but [...] no one in the management can name a single application. [The] operative employees understand what data they have and don't have, but they don't necessarily have the capability of assessing how it affects their competitiveness if they made bigger investments. This conflict is constant. [...] We're talking about a technology that requires cooperation from many parties in the organization: the one who understands what's valuable in the business, then maybe someone who's more of a visionary—what's worth*

> *pursuing in a certain number of years–and then some technological people who can actually do that stuff, those who know the data, and so on. It requires cooperation from so many fronts."* ~Managing Director of an AI consulting company, Workshop transcript.

This leads to surprising insights into the <u>dynamics of responsibilities</u> between employees and ML-based solutions when integrating into organizational tasks—what is referred to as <u>conjoined agency</u> in this study.

> "*[We] built a machine learning-based tool that enables [company name] airline to predict possible disruptions to air traffic more accurately. Currently, they have a few guys who have been watching the weather forecasts and monitoring the air traffic for years, but they are going to retire soon. This tool captures their know-how and helps younger employees learn and do the job, as if they had years of experience.*" ~Head of Operations, AI consulting company, Workshop field notes.

> "*Some users were feeling ashamed that our bot couldn't understand what they meant. This negatively impacted the usage rate.*" ~Senior Data Scientist, Workshop field notes.

ML solution design needed to evolve as the requirements, experience, and understanding of users had changed. Hence, <u>the necessity of individual users' involvement</u> in the improvement process was frequently noted.

> "*The end users should be involved already in the design phase of the solution. [Later, they] are often the party ultimately responsible for the further development of the solution.*" ~AI consultant, FAIA internal meeting notes.

We also observed that many participants, more than expected, depended on <u>vendors for ML tools</u>.

> "*You know, vendors have a lot of very fancy material they will throw at you. And they will try it high in the organization, so that they can get the top leadership really super excited and then [...] you need to do a pilot with this vendor.*" ~Head of Digital Product Development, Interview transcript.

Finally, an ML-based solution might be influenced by factors from outside of the organization. These <u>external influences</u> might be positive, for example, might speed up the development by "build[ing] the things up from existing pieces" (Senior advisor on Data & AI). However, we noted many ways in which the external environment exerted negative pressure on both management and development teams. Despite external environment seems exogenous to the lifecycle of ML-based solutions, our study revealed that the recognition of these external influences as an important component of the overall configuration.

> "*We were the first customer for [vendor name] in our market and we trained their intent recognition model with our data. After some time, it's turned out that they packaged that model into their service offering and now they have a pre-trained model ready for use by our competitors. This was an important lesson. We are currently in the process of changing the architecture to be more vendor-independent and to maintain higher control over things.*" ~Senior Data Scientist, Workshop field notes.

The results helped us to identify the key components of the ML-based solution lifecycle (Table 1) and summarize their relationships, the configuration of key components within a proposed ML-based solution lifecycle (Figure 1A). The resulting definitions of and relationships between these components emerged from abductive comparison of empirical findings and the extant literature. As discussed in the next section, our model recognized two defining roles, *ML Organization* and *ML Technology,* and how their collaboration leads to the establishment of Conjoined Agency in executing tasks. We also acknowledge that this process is not independent of organizational key processes and therefore, it should be seamlessly integrated with proper feedback loops.

Discussion and Contributions

Our study confirms the need to consider ML-based solutions as a sociotechnical system (Lyytinen et al. 2020). Accordingly, we define an ML-based solution lifecycle as a polycyclic process with three key components, *ML Organization cycle, ML Technology cycle,* and *Conjoined Agency cycle* that collectively allows the use of ML-based IS artifact in executing an organizational task or simply for "job to be done" (see

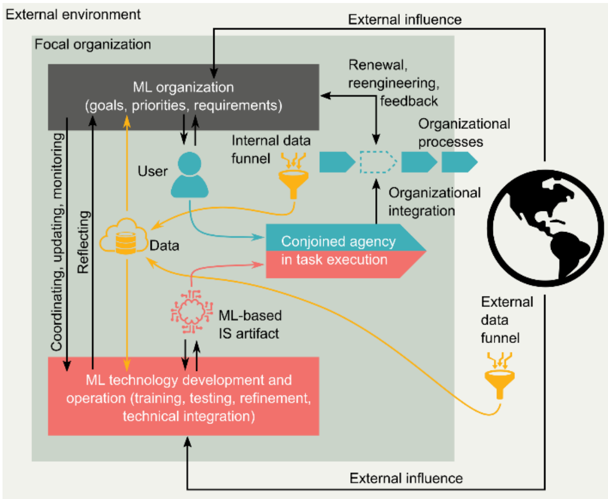

Figure 1A. Configuration of key components within ML-based solution lifecycle

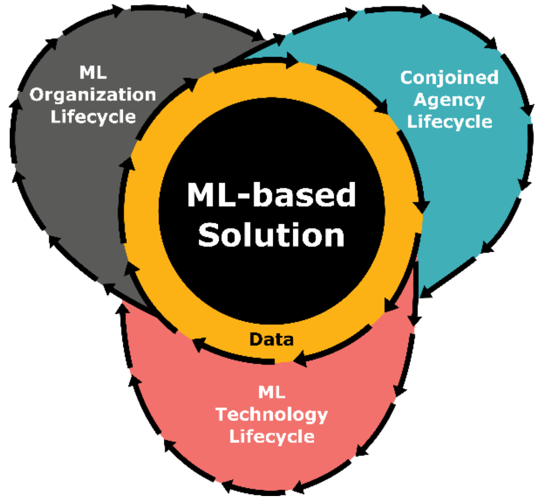

Figure 1B. The ML-based solution lifecycle framework

COMPONENT	DEFINITION AND EXAMPLES
ML Organization	**Definition:** A temporary or permanent social subsystem formed under the umbrella of the focal organization that controls resources, enacts processes, sets rules and objectives, and possesses ML-related competencies for coordinating, updating, and monitoring ML development, implementation, and operation. **Examples:** Data science team, vendor, external developers, internal task force coordinating the work.
ML Technology	**Definition:** The physical infrastructure, software and ML tools, and artifacts created internally or externally and used by ML organization in ML model development (training, testing, verification, refinement). **Examples:** Pre-trained ML model; GPUs (Graphics Processing Units) used for ML model training.
Conjoined Agency (in task execution)	**Definition:** A shared capacity between humans and ML-based Solutions to exercise intentionality by leveraging resource endowments, preferences, and delegation. **Examples:** An ML-based solution generating control parameters for a manufacturing process where human operators are needed to implement these settings; A job candidate screening carried out by an HR professional supported by ML-based recommendation engine.
User	**Definition:** The organizational agent who directly interacts with or benefits from ML-based solutions in organizational task execution. **Examples:** Loan consultant using an ML-based Solution to evaluate the credibility of customers applying for loans; Purchasing manager using an ML-based Solution to estimate product inventory levels for the upcoming quarter.
ML-based Solution	**Definition:** An ML-based IS artifact (smart agent) consisting of an ML model integrated with other IS or non-IS technologies, and potentially other information or social artifacts. **Examples:** Corporate intranet chatbot answering common employee questions related to HR (human resource) issues; Manufacturing quality inspector using image recognition to monitor production quality.
Organizational Processes	**Definition:** Sets of tasks carried out within or across organizational boundary and transforming inputs into outputs in a goal-oriented fashion, affecting material world, digital world, or both. **Examples:** Insurance claim handling process; Planning raw material and component purchases.
Focal Organizational Task	**Definition:** An individual task within the organizational process where ML-based solutions with a conjoined agency are deployed and integrated with the intention of improving task performance or task outcomes. **Examples:** Estimating the final dollar value of an insurance claim; Classifying a support ticket for handling by a specific customer support agent.
Data	**Definition:** Raw facts in digital form, which are (could be) used by ML organizations in ML model development (training, testing, verification, refinement) or by ML-based solutions to generate output. Examples: Video frames from a camera filming self-service checkout of a customer in a grocery retail store; PDF documents with realized sales contracts.
External Environment	**Definition:** The broader context outside of the focal organization's boundary, which includes social, economic, environmental, political, regulatory, technological, and competitive forces influencing or being influenced by organizational actions, processes, and outcomes. **Examples:** Customer demand for smart product recommendations. Use privacy regulations

Table 1. Components of ML-based solution lifecycle configuration.

Figure 1B). ML Organization changes throughout its lifecycle not only in terms of maturity of its practices and capabilities (John et al. 2021) but also with respect to the centralization of ML-related decisions (Fountaine et al. 2019). ML Technology lifecycle reflects the process of ML design, development, integration, and improvement into a fully-fledged ML-based IS artifact (cf. Paterson et al. 2021). The

Conjoined Agency lifecycle represents how over time users and ML-based IS artifacts share the capacity to exercise intentionality in task execution and autonomy in the negotiation of responsibilities (Murray et al. 2021). These subordinate lifecycles are deeply interwoven with each other through data, human interaction, and material aspects of technology. This dynamic could be affected by internal events such as changes in strategic priorities or external events such as technological advancements. It is also important to note that the term "lifecycle" here is not restricting the process to a linear development of prescribed sequences or workflow. Rather, we recognize that the ML-based solution lifecycle is a complex process characterized by multiple iteration loops, equifinality, and emergent behavior. We also emphasize the temporal dimension of ML-based lifecycle and its subordinate lifecycles as a progression of events over time that stabilizes and legitimizes a specific ML-based solution in an organizational context. While this study was exploratory and limited to the context of large organizations, the persistence of the identified patterns indicates likely similar ML-based solution lifecycle and design challenges in small and medium-sized organizations. This, however, will require further investigation.

Theoretical Contributions

Our study contributes to IS literature in three notable ways. Firstly, our proposed ML-based solution lifecycle provides nuanced insight into the context of organizational learning. ML-based solutions significantly differ from human-only learning with respect to learning speed, timing, scope, and scale (Lyytinen et al. 2020). Scholars studying organizational learning recognize that ML technologies will play an increasingly important role in that context and that organizational and social aspects influence the potential impact of ML (Argote et al. 2021). Joint human-machine learning has the potential to mitigate bias (Rai et al. 2019) or improve trading task performance (Sturm, Koppe, et al. 2021). Our study suggests that learning in an ML-based solution lifecycle takes place at least on four levels (a) ML Technology learns from training example datasets and subsequent retraining, (b) users of ML-based solutions learn how to work with (around) these tools in their task execution, (c) ML Organization learns about how to continuously develop and operate ML technology, and (d) ML Organization and users learn how to identify and refine organizational use of ML to extract new insights or value by business process renewal or reengineering. The presence of learning in ML Organization aligns with the reflective organizational functions (Lyytinen et al. 2020) that ensure the long-term utility of ML-based solutions (Ransbotham et al. 2020). The vitality and potential for value generation from organizational learning within an ML-based solution life cycle depend on the ML Organization's ability to collect user feedback, understand business, develop ML competencies, and monitor internal and external events.

Secondly, our results bring a new perspective to the discussion on conjoined agency (Murray et al. 2021) and agentic IS artifacts (Baird & Maruping 2021). Thus far, this research has been primarily concerned with the consequences of organizational use of ML on, for example, fairness (Teodorescu et al. 2021) or routines (Murray et al. 2021). Baird and Maruping (2021) have also proposed a framework for delegation, which recognizes agencies for ML-based IS solutions. Complementing these efforts, this study emphasizes the lifecycle view of the conjoined agency. While Murray and colleagues (2021) distinguish different types of conjoined agency in protocol development and action selection, their perspective overlooks the lifecycle of conjoined agency preceding protocol development. Our results indicate that even if the locus of agency in protocol development rests with ML Technology, ML Organization has an active role in shaping and refining how ML-based solutions arrive at that protocol. In other words, even if protocol design comes from an ML Technology, human actors are still in charge of designing and refining the related metaprotocol (protocol for designing protocol). Our findings also suggested that human actors, not necessarily developers and users, can react to conjoin agency and facilitate its integration and refinement. Conjoined agencies can also attract and engage new human actors. That means organizational use of ML-based solutions may redistribute agency between people in potentially new and unexpected ways.

Finally, ML-based solution lifecycle theorization elaborates our understanding of how organizations can manage the automation-augmentation paradox (Raisch & Krakowski 2021). Automation is commonly associated with machines taking over human tasks, while augmentation designates humans and machines collaborating with and/or complementing each other (Benbya et al. 2021; Raisch & Krakowski 2021; Teodorescu et al. 2021). Our study confirms that focusing on either of these approaches is counterproductive and may lead to negative consequences. Instead, organizations should proactively balance these across space and time. Our analysis suggests that managing the transitions between automation and augmentation, as well as deciding on where to leverage these requires awareness of three interlinked lifecycles of ML Organization, ML Technology, and Conjoined Agency. Each of these cycles is a function of temporal factors and thereby they might progress or iterate at a different pace. These "differential clock speeds" mean that organizations need to decide to what extent they want to leverage ML-based solutions where the subordinate lifecycles are relatively easy to synchronize versus those where substantial effort might be required. For example, the iteration cycle for technology might be dramatically faster than the pace at which users take an ML-based solution into use and develop the surrounding task routines. The "differential clock speeds" of these cycles might render disconnection, conflicts, or confusion. This can take place, for example when an ML-based solution changes behavior or performance at a much higher pace than what users can detect or adjust to.

Practical Implications

This study offers the following implications for practitioners engaged in developing, managing, and advising on organizational efforts relating to MLOps. Firstly, we stress the importance of broadening the scope of MLOps to better reflect the sociotechnical nature of these activities. The present focus on the ML technology lifecycle is the natural first step in managing the ML-based solution lifecycle. However, to fully realize the potential of ML technology and, at the same time, mitigate the associated risks, MLOps practitioners need to fully appreciate the ML organization and conjoined agency lifecycles as constituent parts of the ML-solution lifecycle. Secondly, learning is a crucial mechanism that MLOps need to activate to ultimately deliver value to organizations. Our research points the attention of MLOps practitioners to different levels of learning within ML-based solutions. Apart from the obvious MLOps learning cycle, practitioners need to reflect on how ML Technology and ML Organization lifecycles can be improved by observing the Conjoined Agency lifecycle. MLOps practitioners should recognize that such a process directly and ultimately affects the users and how they learn to share responsibilities with ML-based solutions. Thirdly, the more holistic understanding of the ML-based solution lifecycle reveals that MLOps practitioners might (un)knowingly assume part of the responsibility for determining protocols (guidelines and rules) regarding how to seamlessly integrate ML-based solutions with work practices while considering strategical and ethical consequences of such integration. Finally, we call MLOps practitioners to not overoptimize the ML technology cycle speed at the expense of hindering the overall cycling pace of ML-based solution integration. This requires MLOps teams to devote efforts to improving the pace at which they can reflect and develop insights into ML-based solution integration, as well as the speed with which users accommodate ML-based solutions into routines supporting organizational learning or task execution.

References

Ågerfalk, P. J. (2020). Artificial intelligence as digital agency. European Journal of Information Systems, 29(1), 1–8.

Alla, S., and Adari, S. K. 2021. "What Is MLOps?," in Beginning MLOps with MLFlow.

Alnafessah, A., Gias, A. U., Wang, R., Zhu, L., Casale, G., and Filieri, A. 2021. "Quality-Aware DevOps Research: Where Do We Stand?," IEEE Access (9). (https://doi.org/10.1109/ACCESS.2021.3064867).

Argote, L., Lee, S., & Park, J. (2021). Organizational Learning Processes and Outcomes: Major Findings and Future Research Directions. Management Science, 67(9), 5399–5429. https://doi.org/10.1287/mnsc.2020.3693

Asatiani, A., Malo, P., Nagbøl, P. R., Penttinen, E., Rinta-Kahila, T., & Salovaara, A. (2021). Sociotechnical Envelopment of Artificial Intelligence: An Approach to Organizational Deployment of Inscrutable Artificial Intelligence Systems. Journal of the Association for Information Systems, 22(2), 28.

Baird, A., & Maruping, L. M. (2021). The Next Generation of Research on Is Use: A Theoretical Framework of Delegation to and from Agentic Is Artifacts. MIS Quarterly, 45(1), 315–341. https://doi.org/10.25300/MISQ/2021/15882

Benbya, H., Davenport, T. H., & Pachidi, S. (2020). Special Issue Editorial: Artificial Intelligence in Organizations: Current State and Future Opportunities. MIS Quarterly Executive, 19(4), ix–xxi. Business Source Complete.

Benbya, H., Pachidi, S., & Jarvenpaa, S. (2021). Special Issue Editorial: Artificial Intelligence in Organizations: Implications for Information Systems Research. Journal of the Association for Information Systems, 22(2), 10. http://dx.doi.org/10.17705/1jais.00662

Berente, N., Bin Gu, Recker, J., & Santhanam, R. (2021). Managing Artificial Intelligence. MIS Quarterly, 45(3).

Brynjolfsson, E., & Mitchell, T. (2017). What can machine learning do? Workforce implications. Science, 358(6370).

Colantoni, A., Berardinelli, L., and Wimmer, M. 2020. "DevOpsML: Towards Modeling DevOps Processes and Platforms," in Proceedings - 23rd ACM/IEEE International Conference on Model Driven Engineering Languages and Systems, MODELS-C 2020 - Companion Proceedings.

Fountaine, T., McCarthy, B., & Saleh, T. (2019). Building the AI-Powered Organization. Harvard Business Review, 97(4), 13.

Gall, M., and Pigni, F. 2021. "Taking DevOps Mainstream: A Critical Review and Conceptual Framework," Https://Doi.Org/10.1080/0960085X.2021.1997100, Taylor & Francis.
Garcia, R., Sreekanti, V., Yadwadkar, N., Crankshaw, D., Gonzalez, J. E., and Hellerstein, J. M. 2018. "Context: The Missing Piece in the Machine Learning Lifecycle," in ACM CMI, London.
Goyal, A. 2020. "MLOps - Machine Learning Operations," International Journal of Information Technology Insights & Transformations.
Graebner, M. E., Martin, J. A., & Roundy, P. T. (2012). Qualitative data: Cooking without a recipe. Strategic Organization, 10(3), 276–284. https://doi.org/10.1177/1476127012452821
Grønsund, T., & Aanestad, M. (2020). Augmenting the algorithm: Emerging human-in-the-loop work configurations. The Journal of Strategic Information Systems, 29(2), 101614. https://doi.org/10.1016/j.jsis.2020.101614
John, M. M., Olsson, H. H., & Bosch, J. (2021). Towards MLOps: A Framework and Maturity Model. 2021 47th Euromicro Conference on Software Engineering and Advanced Applications (SEAA), 1–8.
Levallet, N., Denford, J. S., & Chan, Y. E. (2021). Following the MAP (Methods, Approaches, Perspectives) in Information Systems Research. Information Systems Research, 32(1), 130–146.
Lwakatare, L. E., Crnkovic, I., and Bosch, J. 2020. "DevOps for AI - Challenges in Development of AI-Enabled Applications," in International Conference on Software, Telecommunications and Computer Networks, SoftCOM 2020, Institute of Electrical and Electronics Engineers Inc., September 17.
Lyytinen, K., & Newman, M. (2008). Explaining information systems change: A punctuated socio-technical change model. European Journal of Information Systems, 17(6), 589–613. https://doi.org/10.1057/ejis.2008.50
Lyytinen, K., Nickerson, J. V., & King, J. L. (2020). Metahuman systems = humans + machines that learn. Journal of Information Technology, 0268396220915917. https://doi.org/10.1177/0268396220915917
Makinen, S., Skogstrom, H., Laaksonen, E., and Mikkonen, T. 2021. "Who Needs MLOps: What Data Scientists Seek to Accomplish and How Can MLOps Help?," in Proceedings - 2021 IEEE/ACM 1st Workshop on AI Engineering - Software Engineering for AI, WAIN 2021. (https://doi.org/10.1109/WAIN52551.2021.00024).
Martínez-Fernández, S., Franch, X., Jedlitschka, A., Oriol, M., and Trendowicz, A. 2021. "Developing and Operating Artificial Intelligence Models in Trustworthy Autonomous Systems," in Lecture Notes in Business Information Processing (Vol. 415 LNBIP), Springer Science and Business Media Deutschland GmbH, pp. 221–229.
Moreschini, S., Hästbacka, D., Taibi, D., and Lomio, F. 2022. "MLOps for Evolvable AI Intensive Software Systems," IEEE International Conference on Software Analysis, Evolution and Reengineering.
Murray, A., Rhymer, J., & Sirmon, D. G. (2021). Humans and Technology: Forms of Conjoined Agency in Organizations. Academy of Management Review, 46(3), 552–571. https://doi.org/10.5465/amr.2019.0186
Paterson, C., Calinescu, R., & Ashmore, R. (2021). Assuring the Machine Learning Lifecycle: Desiderata, Methods, and Challenges. ACM Computing Surveys.
Rai, A., Constantinides, P., and Sarker, S. 2019. "Editor's Comments: Next-Generation Digital Platforms: Toward Human–AI Hybrids," Management Information Systems Quarterly (43:1).
Raisch, S., & Krakowski, S. (2021). Artificial Intelligence and Management: The Automation–Augmentation Paradox. Academy of Management Review, 46(1), 192–210. https://doi.org/10.5465/amr.2018.0072
Ransbotham, S., Khodabandeh, S., Kiron, D., Candelon, F., Chui, M., & LaFountain, B. (2020). Expanding AI's Impact With Organizational Learning. MIT Sloan Management Review and Boston Consulting Group. https://sloanreview.mit.edu/projects/expanding-ais-impact-with-organizational-learning/
Riungu-Kalliosaari, L., Mäkinen, S., Lwakatare, L. E., Tiihonen, J., and Männistö, T. 2016. "DevOps Adoption Benefits and Challenges in Practice: A Case Study," in Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics) (Vol. 10027 LNCS).
Sarker, S., Chatterjee, S., Xiao Xiao, & Elbanna, A. (2019). The Sociotechnical Axis of Cohesion for the Is Discipline: Its Historical Legacy and Its Continued Relevance. MIS Quarterly, 43(3), 695-A5.
Strich, F., Mayer, A.-S., & Fiedler, M. (2021). What Do I Do in a World of Artificial Intelligence? Investigating the Impact of Substitutive Decision-Making AI Systems on Employees' Professional Role Identity. Journal of the Association for Information Systems, 22(2), 9. http://dx.doi.org/10.17705/1jais.00663
Sturm, T., Gerlach, J. P., Pumplun, L., Mesbah, N., Peters, F., Tauchert, C., Nan, N., & Buxmann, P. (2021). Coordinating Human and Machine Learning for Effective Organizational Learning. MIS Quarterly, 45(3).
Sturm, T., Koppe, T., Scholz, Y., & Buxmann, P. (2021). The Case of Human-Machine Trading as Bilateral Organizational Learning. 18.
Tamburri, D. A. 2020. "Sustainable MLOps: Trends and Challenges," in Proceedings - 2020 22nd International Symposium on Symbolic and Numeric Algorithms for Scientific Computing, SYNASC 2020.
Teodorescu, M. H. M., Morse, L., Awwad, Y., & Kane, G. C. (2021). Failures of Fairness in Automation Require a Deeper Understanding of Human-Ml Augmentation. MIS Quarterly, 45(3), 1483–1499.
Zhang, Z., Nandhakumar, J., Hummel, J. T., and Waardenburg, L. 2020. "Addressing the Key Challenges of Developing Machine Learning AI Systems for Knowledge-Intensive Work," MIS Quarterly Executive (19:4), Indiana University, pp. 221–238. (https://doi.org/10.17705/2msqe.00035).
Van de Ven, A. H. (2007). Engaged scholarship: A guide for organizational and social research. Oxford University Press on Demand.
van den Broek, E., Sergeeva, A., & Huysman, M. (2021). When the Machine Meets the Expert: An Ethnography of Developing Ai for Hiring. MIS Quarterly, 45(3), 1557–1580. https://doi.org/10.25300/MISQ/2021/16559