MIS Quarterly Executive

Volume 22 | Issue 1 Article 4

March 2023

Successfully Organizing AI Innovation Through Collaboration with Startups

Jana Oehmichen

Alexander Schult

John Qi Dong

MIS QUARTERLY EXECUTIVE

Successfully Organizing AI Innovation Through Collaboration with Startups

Artificial intelligence (AI) technologies are increasingly being used by firms for competitive advantage. Collaborating with AI startups can be an effective way to leverage these technologies. Based on the observations of six AI use cases at established firms delivered by two different AI startups, we identify five challenges of collaborating with startups—finding the right AI startup, identifying the right AI use case, agreeing on the right commercial terms and conditions, considering the impact on people and overcoming implementation roadblocks—and provide recommendations for addressing these challenges.[1,2]

Jana Oehmichen
University of Mainz (Germany)
University of Groningen (Netherlands)

Alexander Schult
ONDEWO GmbH (Austria)

John Qi Dong
Nanyang Technological University (Singapore)

Challenges of AI Innovation Through Collaboration with AI Startups

Businesses are increasingly looking to exploit the promise of artificial intelligence (AI) technologies,[3] with 85% of executives surveyed in 2017 confirming that AI technologies will allow them to build and maintain their competitive advantage.[4] AI innovation—i.e., innovation that leverages AI technologies[5]—can impact many organizational tasks, such as logistics,[6]

1 Yolande Chan is the accepting senior editor for this article.
2 The authors thank the senior editor and three anonymous reviewers for their constructive feedback and thoughtful guidance throughout the review process.
3 See, for example, von Krogh, G. "Artificial Intelligence in Organizations: New Opportunities for Phenomenon-Based Theorizing," *Academy of Management Discoveries* (4:4), December 2018, pp. 404-409.
4 Ransbotham, B. S., Kiron, D., Gerbert, P. and Reeves, M. "Reshaping Business with Artificial Intelligence: Closing the Gap between Ambition and Action," *MIT Sloan Management Review* (59), September 2017, pp. 1-38.
5 We follow recent literature and define innovation as "the design, creation, development and/or implementation of new or altered products, services, systems, organizational structures, management practices and processes, or business models;" for information on AI innovation, see 1) Benbya, H., Davenport, T. H. and Pachidi, S. "Artificial Intelligence in Organizations: Current State and Future Opportunities," *MIS Quarterly Executive* (19:4), December 2020, pp. 9-21; and 2) Benbya, H. and Leidner, D. "How Allianz UK Used an Idea Management Platform to Harness Employee Innovation," *MIS Quarterly Executive* (17:2), June 2018, pp. 141-157.
6 von Krogh, G., op. cit., December 2018.

DOI: 10.17705/2msqe.00073

recruitment,[7] marketing[8] and medical diagnostics, among others.[9] In summary, AI innovation can be the key to significantly improving products or services and making firms' business processes faster and more efficient, thereby creating greater value for customers.[10]

However, while some AI vendors (e.g., Amazon, Google, IBM and other digital companies) are introducing powerful and affordable AI solutions, many established firms struggle to successfully leverage AI technologies for innovation. In 2017, only 40% of surveyed firms had an AI strategy,[11] and in 2018, only 27% of firms had received a positive return on their AI investments.[12] To help them to move forward with AI-based innovations, established firms are seeking interorganizational collaboration with AI startups[13] to enable a "make with external support" delivery model. Unfortunately, such AI collaborations create their own set of new challenges. In particular, the three distinguishing characteristics of AI solutions—acting autonomously without human involvement, being able to learn from data and experience and being inscrutable for certain stakeholder[14] —need to be considered and managed in partnerships with AI startups. Hence, the purpose of this article is to shed more light on the research question: *How can established firms successfully organize their collaborations with AI startups?*

The most frequent reason for established firms seeking AI partnerships with startups is their lack of skilled expertise;[15] human talent makes a big difference in creating breakthrough AI innovations. Unfortunately, AI talent is even scarcer than regular IS talent, because AI experts not only need to excel in IS topics, but also in data science, statistics and econometrics.[16] Individuals with the necessary AI expertise are needed regardless of whether firms pursue a "make" or "buy" delivery model for AI solutions.[17] Firms using a "make" model require strong human talent for internal bespoke development, and those using a "buy" model to generate business value require their own skilled expertise for customizing and integrating vendors' AI solutions and for continuous retraining due to the novel nature of AI.[18] Moreover, most AI vendors' solutions are designed and developed in a generic and generalizable manner and are highly scalable.[19] Their off-the-shelf AI solutions are thus typically not a perfect match for their clients' firm-specific business challenges.

The scarcity of AI talent is further fueled by the current trend for those with desirable AI expertise to pursue an entrepreneurial career path via founding or joining AI startups. AI startups often have close ties with universities and attract graduate AI talent by offering attractive stock options. As a consequence, the dearth of AI talent available to established firms creates attractive business-to-business (B2B) opportunities for the growing number of AI startups. Thus, interorganizational collaboration with AI startups can help established firms overcome their lack of AI expertise.

7 van den Broek, E., Sergeeva, A. and Huysman, M. "Hiring Algorithms: An Ethnography of Fairness in Practice," *Proceedings of International Conference on Information Systems*, December 15-18, 2019, pp. 1-9.

8 Davenport, T., Guha, A., Grewal, D. and Bressgott, T. "How artificial intelligence will change the future of marketing," *Journal of the Academy of Marketing Science* (48:1), October 2019, pp. 24-42.

9 Lebovitz, S. "Diagnostic Doubt and Artificial Intelligence: An Inductive Field Study of Radiology Work," *Proceedings of International Conference on Information Systems*, December 15-18, 2019, pp. 1-17.

10 Rai, A. "Explainable AI: From Black Box to Glass Box," *Journal of the Academy of Marketing Science* (48:1), December 2019, pp. 137-141.

11 Ransbotham, B. S., Kiron, D., Gerbert, P. and Reeves, M., op. cit., September 2017.

12 Bean, R. "How Big Data and AI Are Driving Business Innovation in 2018," *MIT Sloan Management Review*, 2018, available at https://sloanreview.mit.edu/article/how-big-data-and-ai-are-driving-business-innovation-in-2018/.

13 Berente, N., Gu, B., Recker, J. and Santhanam, R. "Managing Artificial Intelligence," *MIS Quarterly* (45:3), September 2021, pp. 1433–1450.

14 Berente, N., Gu, B., Recker, J. and Santhanam, R. "Managing Artificial Intelligence," *MIS Quarterly* (45:3), September 2021, pp. 1433–1450.

15 Lou, B. and Wu, L. "AI on Drugs: Can Artificial Intelligence Accelerate Drug Development? Evidence from a Large-Scale Examination of Biopharma Firms," *MIS Quarterly* (45:3), June 2021, pp. 1451-1482.

16 Benbya, H., Davenport, T. H. and Pachidi, S., op. cit., December 2020.

17 Zhang, Z., Nandhakumar, J., Hummel, J. T. and Waardenburg, L. "Addressing the Key Challenges of Developing Machine Learning AI Systems for Knowledge-Intensive Work," *MIS Quarterly Executive* (19:4), December 2020, pp. 221-238.

18 Berente, N., Gu, B., Recker, J. and Santhanam, R., op. cit., September 2021.

19 Belgaum, M. R., Alansari, Z., Musa, S., Alam, M. M. and Mazliham, M. S. "Role of Artificial Intelligence in Cloud Computing, IoT and SDN: Reliability and Scalability Issues," *International Journal of Electrical and Computer Engineering* (11:5), October 2021, pp. 4458-4470.

Based on an insider-outsider ethnography[20] across six AI use cases involving two different AI startups (three each from Startup 1 and Startup 2), we identified five challenges of organizing collaboration for AI innovation between established firms and AI startups:

- Challenge 1: Finding the right AI startup
- Challenge 2: Identifying the right AI use case
- Challenge 3: Agreeing on the right commercial terms and conditions
- Challenge 4: Considering the impacts on people
- Challenge 5: Overcoming implementation roadblocks.

To enable firms to address these challenges, we provide the following five recommendations:

- Recommendation 1: Overcome the inscrutability of AI startups by assessing credible quality signals
- Recommendation 2: Leverage the outside-in perspective of AI startups
- Recommendation 3: Track AI-based value creation and let the AI startups participate to align incentives
- Recommendation 4: Consciously consider the right degree of AI autonomy for each AI use case
- Recommendation 5: Work with or around stakeholders to ensure the prerequisites for AI use case success are in place.

The challenges and our recommendations for addressing them are described in detail below. First, however, we provide a brief overview of the history of AI-based innovation.

Overview of the History of AI-Based Innovation

In IS and management research, AI is defined as "the ability of a machine to perform cognitive functions that we associate with human minds, such as perceiving, reasoning, learning, interacting with the environment, problem solving, decision-making, and even demonstrating creativity."[21] AI can be seen as a collection of computer-assisted systems that encompass task input (such as sound, image, text and numbers), task processes (such as algorithms) and task outputs in the form of solutions and decisions.[22] The most prominent types of AI technologies are natural language processing, machine learning and computer vision.[23]

AI has a long history. As long ago as 1971, Nilsson defined AI as "machines performing cognitive functions usually associated with human minds, such as learning, interacting, and problem-solving."[24] Businesses have experienced earlier waves of AI technologies, such as expert systems in the 1980s,[25] which came to nothing because these systems had a too strong problem orientation that hindered efficient implementation.[26] However, business interest in AI technologies and solutions has been recently revived due to technological progress, such as decreased costs of hardware and cloud-based technologies.[27]

The ever-evolving nature of AI brings its own challenges that differ from traditional IT solutions. According to Berente et al., AI is "the frontier of computational advancements that references human intelligence in addressing ever more complex decision-making problems."[28] This frontier constantly shifts as the three distinguishing characteristics of AI—autonomy, learning and inscrutability—evolve. Autonomy relates to AI systems increasingly acting on their own as agents or reducing the action range of involved humans. Learning relates to complex algorithms improving together with humans or by themselves based on large-scale training data. Inscrutability relates to aspects such as opacity, transparency, explainability (the ability

20 Insider-outsider ethnography is an observational research approach via participating in the research setting and with the subject being studied.

21 Rai, A., Constantinides, P. and Sarker, S. "Next-Generation Digital Platforms: Toward Human-AI Hybrids," *MIS Quarterly* (43:1), February 2019, pp. iii–ix.

22 von Krogh, G. op. cit., December 2018.

23 Brynjolfsson, E. and McAfee, A. *The Second Machine Age: Progress, and Prosperity in a Time of Brilliant Technologies*, W. W. Norton & Company, 2014.

24 Nilsson, N. J. *Problems-Solving Methods in Artificial Intelligence*, McGraw Hill, 1971.

25 Gill, T. G. "Early Expert Systems: Where Are They Now?" *MIS Quarterly* (19:1), March 1995, pp. 51-70.

26 Davenport, T. H. and Harris, J. G. "Automated Decision Making Comes of Age," *MIT Sloan Management Review* (46:4), July 2005, pp. 83-89.

27 von Krogh, G. op. cit., December 2018.

28 Berente, N., Gu, B., Recker, J. and Santhanam, R., op. cit., September 2021.

to explain how AI systems arrive at a decision or recommendation) and the interpretability of complex or self-evolving AI algorithms. The evolving nature of these three characteristics means that AI-based innovation and the related challenges differ from traditional IT projects that do not have to take account of an ever-shifting frontier.

Introduction to the Six AI Use Cases and our Case Analysis

In our research for this article, we conducted an insider-outsider ethnography.[29] This action research approach, sometimes called "opportunistic ethnography,"[30] is a common approach in IS research[31] and other fields (e.g., sociology).[32] One author of this article is an AI practitioner involved in two AI startups (Startup 1 and Startup 2), where he has full-time strategic and operational leadership roles, including co-founder, business development and project management. The recommendations in this article are based on the real-life AI use cases in 2019-2021 from his work. The other two authors are academics.

Startup 1 was founded in Austria and has developed natural language processing algorithms with speech-to-text, intent recognition, entity recognition and text-to-speech algorithms that enable firms to automate customer conversations via voice and text. The three AI use cases from established firms collaborating with Startup 1 reported in this article are:

- *Use Case 1: Text-based conversation recommendations for call center agents*. Startup 1 customized its natural language processing algorithms so that a Finnish telecoms firm could detect intents and entities of incoming chat messages from customers and to suggest likely answers to call center agents. The agents can send the suggested answer directly or choose to modify the message before sending it. By removing the need for agents to repeatedly type the answers, the introduction of AI doubled call center agent productivity. The AI solution was integrated into the existing operational IT system and workflow so that call center agents do not lose time switching between systems.
- *Use Case 2: Outbound call automation for lead generation*. Startup 1 built a voice bot for an established Austrian logistics firm. The voice bot handles outbound calls fully autonomously and calls B2B customers that have previously used the firm's logistics services. The AI bot identifies a customer's interest in using the logistics service again, summarizes the lead in an email to the firm's sales representative who would be responsible for closing the deal, and updates contact information with changes—for example, if customers indicate that they do not want to talk to a voice bot again. Due to the AI bot's capability to conduct calls in parallel, this AI use case enabled the firm to reach over 1,500 customers within hours and sell short-term freight capacity that previously went unused.
- *Use Case 3: Support for 911 emergency dispatchers*. Startup 1 collaborated with an established European software firm that produces software for fire brigade and rescue control centers to build a module that listens to and understands emergency calls and prefills critical pieces of information into the dispatching software. The AI module was integrated into the existing software solution as an innovative feature extension. This use case increases quality as well as dispatching speed because it helps the dispatchers

29 Compare, e.g.: 1) van den Broek, E., Sergeeva, A. and Huysman, M. "When the Machine Meets the Expert: An Ethnography of Developing AI for Hiring," *MIS Quarterly* (45:3), September 2021, pp. 1557-1580; 2) Orlikowski, W. J. "Using Technology and Constituting Structures: A Practice Lens for Studying Technology in Organizations," *Organization Science* (11:4), July 2000, pp. 404-428; and 3) Tracey, P., Phillips, N. and Jarvis, O. "Bridging Institutional Entrepreneurship and the Creation of New Organizational Forms: A Multilevel Model," *Organization Science* (22:1), February 2011, pp. 60-80.

30 Riemer, J. W. "Varieties of Opportunistic Research," *Journal of Contemporary Ethnography* (5:4), January 1977, pp. 467-477.

31 See: 1) Aakhus, M., Ågerfalk, P. J., Lyytine, K. and Te'eni, D. "Symbolic Action Research in Information Systems," *MIS Quarterly* (38:4), December 2014, pp. 1187-1200; and 2) Baskerville, R. L. and Myers, M. D. "Special Issue on Action Research in Information Systems: Making IS Research Relevant to Practice: Foreword," *MIS Quarterly* (28:3), September 2004, pp. 329-335.

32 Tracey, P., Phillips, N. and Jarvis, O., op. cit., February 2011.

focus on the conversation and reduces the need for typing.

Startup 2 was originally incorporated in the U.S. and quickly expanded into several European markets. In Europe, it serves global clients, such as insurance companies and online retailers, focusing on data-driven topics, such as dynamic pricing, fraud detection, case management, inventory management, product portfolio optimization and lead generation. AI use cases in these application areas are based on individually developed, case-specific machine-learning approaches. The three AI use cases from established firms collaborating with Startup 2 reported in this article are:

- *Use Case 4: Car fleet insurance pricing*. Startup 2 created a novel claims model for a large global insurance company to predict future claims. The purpose of this AI-based solution was to support a team of underwriters in their pricing decisions for enterprise insurance clients. In the past, the personal judgment of underwriters had a major influence on the insurance prices for enterprises' large vehicle fleets, as traditional rule-based IT tools could not effectively reflect the complexity of the pricing decisions. The novel AI-based claims prediction model required accessing and preparing data from various internal and external sources. In addition, a new operational IT system with a web-based user interface and a back-end server solution had to be created to provide access for the underwriters to the new pricing model. AI-based pricing yielded a threefold return on investment in the first year alone.
- *Use Case 5: AI-based claims cost reduction*. The goal of this AI use case is to help a global insurance company reduce claims costs for auto repairs after car accidents. A machine-learning model predicts if claims handlers can influence a customer's repair shop choice and if doing so is financially attractive. The AI system also recommends whether claims handlers should offer financial incentives to a customer to pick the insurance company's preferred repair shop. Recommendations from the machine-learning model are displayed in the established legacy operational IT system that claims handlers use during customer calls. This legacy system was modified so that case steering attempts and success rates could be tracked at the team level. The overall success rate of steering customers to the preferred repair shop was below expectations but varied significantly among teams, indicating that additional nontechnical factors, such as claims handlers' skills and incentives, business processes and managerial control, had a major influence on the overall outcome of this use case.
- *Use Case 6: Dynamic pricing*. Startup 2 created a pricing model based on machine learning for a large global firm in the financial services sector to optimize the trade-off between margin improvements and the likelihood of customer churn. The pricing model was used to generate optimized prices that were made available via a scalable server solution that could answer online pricing inquiries in under 200 milliseconds. By predicting price elasticities based on customer characteristics, this AI use case helped increase the financial product's margin by over 2%.

Based on these six AI use cases, the authors used a four-step approach to systematically distill the AI-specific challenges of collaborating with startups and identify the actions for addressing the challenges. First, both academic authors interviewed the practitioner author with the open questions listed in the Appendix to generate a list of unfiltered observations for each use case. The questions aimed to combine the perspectives of the AI startups and established firms to ensure that observations covered both sides of the collaboration. Second, both academic authors independently classified these observations according to their relationships with the three distinguishing characteristics of AI—autonomy, learning and inscrutability. Third, all authors compared the individual classifications and jointly resolved inconsistencies. This step involved outlining each author's rationale for the classification and further in-depth discussions and clarification of unfiltered observations where needed. Fourth, all authors dropped the

Figure 1: Summary of the Challenges and our Recommendations

Challenge	Challenge Components for Established Firms	Recommendation	Recommendation Details
1 Finding the right AI startup	1(i) Inscrutability of startup's true capabilities 1(ii) Inability to assess the quality and potential of the startup's employees	**Overcome the inscrutability of startup's capabilities and talent by systematically assessing related credible quality signals from trusted sources**	1(a) Leverage firm-level quality signals from other institutions 1(b) Draw conclusions from non-AI achievements of startup employees
2 Identifying the right use case	2(i) Inability to understand relevance of algorithmic activity in various application areas 2(ii) Inability to understand the potential of algorithmic activity to generate value	**Leverage the outside-in perspective of AI startups to generate and prioritize AI use cases based on value creation potential above-and-beyond data availability**	2(a) Engage in joint use case idea-generation workshops 2(b) Jointly prioritize use cases based on impact and feasibility
3 Agreeing the right commercial terms and conditions	3(i) Lack of transparency on implementation progress 3(ii) Unfamiliarity with the option to objectively measure AI performance	**Track AI-based value creation and let startups participate to align incentives**	3(a) Create incentives with performance-based or usage-based compensation 3(b) Use the startup to build impact measuring capabilities
4 Considering the impact on people	4(i) Difficulty of fine-tuning AI autonomy to optimize the automation vs. user acceptance tradeoff 4(ii) Potential workforce resistance due to users' fear of the role-changing potential of AI	**Consciously consider the right degree of AI autonomy for each use case and position collaboration with the startup accordingly**	4(a) Include the startup's view on how to adjust AI autonomy over time 4(b) Clarify the role of the startup in the change process
5 Overcoming implementation roadblocks	5(i) Multitude of stakeholders on critical path toward AI use case success 5(ii) Multiple prerequisites for meeting AI impact expectations	**Depending on internal organizational maturity and startup's capabilities, work with or around stakeholders to ensure pre-requisites for use case success are in place**	5(a) Ensure that the startup successfully collaborates with stakeholders 5(b) Use the startup to bypass roadblocks

observations that did not stem from the current frontier of AI and jointly structured the remaining AI-related findings that were common across multiple use cases along the lifecycle of the partnership between established firms and AI startups.

Overview of the Challenges of Collaborating with AI Startups and our Recommendations

The four-step approach led us to identify five challenges that established firms face when collaborating with AI startups, each of which has two components. These challenges and components are shown in Figure 1, along with a summary of our recommendations for overcoming them. For each use case, Table 1 summarizes the links between the challenge components and our recommendations. For example, in Use Case 1, the way in which the Finnish telecoms firm addressed Challenge 1(i) (inscrutability of AI startup's true capabilities) helped us to formulate Recommendation 1(a) (leverage firm-level signals from other institutions). The challenges and our recommendations for addressing them are now described in detail.

Challenge 1: Finding the Right AI Startup for Collaboration

The challenge for an established firm of finding the right AI startup for collaboration has two components: 1) the inscrutability of an AI startup's true capabilities at a firm level, and 2) the difficulty of assessing the quality and potential of a startup's employees at an individual level.

The first of these challenge components results from AI being overly used as a buzzword by many startups to attract attention. AI is currently a hot topic and many "wannabee AI startups" claim to be using AI to make their products more appealing to the market. The established firms in our AI use cases found that newly founded AI startups usually provided "black boxes" and had limited market track records, such as reference customers or recurring revenues. Those that are successful quickly become unattractive as collaboration partners (e.g., because they are acquired by big tech companies aiming to cash in on the "winner takes it all" market dynamics).

The second component of the challenge of initializing a collaboration is that many established firms cannot assess whether an AI startup's employees have the necessary

Table 1: Overview of AI Use Cases Findings that Led to the Recommendations

	Use Case 1	Use Case 2	Use Case 3	Use Case 4	Use Case 5	Use Case 6
General information						
Case description	Text-based conversation recommendations for call center agents	Outbound call automation for lead generation	Support for 911 emergency dispatchers	Car fleet insurance pricing	AI-based claims cost reduction	Dynamic pricing
Goal	Increase call center capacity	Increase sales	Improve dispatching quality	Increase profit margin	Increase profit margin	Increase profit margin
Impact lever	Increase number of customer interactions	Increase number of sales calls	Reduce administrative work for dispatchers	Improve claims prediction	Reduce repair shop costs	Optimize trade-off between churn and price increase
Established firm's industry	Telecoms	Logistics	Software	Insurance	Insurance	Financial Services
AI Startup	1	1	1	2	2	2
Observations and Findings by Challenge (Observed and/or Mitigated Challenges → Related Recommendations)						
Finding the right AI startup	• 1(i) → 1(a)	• n/a	• 1(ii) → 1(b)	• 1(i) → 1(a)	• 1(i) → 1(a)	• 1(ii) → 1(b)
Identifying the right AI use case	• 2(i) → 2(a) • 2(ii) → 2(b)	• n/a	• n/a	• 2(i) → 2(a) • 2(ii) → 2(b)	• 2(ii) → 2(b)	• n/a
Agreeing on the right commercial terms and conditions	• 3(i) → 3(a)	• n/a	• n/a	• 3(i) → 3(a) • 3(ii) → 3(b)	• 3(i) → 3(a)	• 3(ii) → 3(b)
Considering the impact on people	• 4(i) → 4(a) • 4(ii) → 4(b)	• 4(ii) → 4(b)	• 4(i) → 4(a) • 4(ii) → 4(b)	• 4(i) → 4(a) • 4(ii) → 4(b)	• 4(i) → 4(a)	• 4(ii) → 4(b)
Overcoming implementation roadblocks	• 5(i) → 5(a)	• 5(ii) → 5(b)	• n/a	• 5(i) → 5(a)	• 5(i) (not resolved)	• 5(ii) → 5(b)

skills for a successful collaboration. There is a very limited number of experts capable of doing this assessment. Established firms are often unfamiliar with individual AI-specific achievements, such as AI coding challenges or public contributions of open-source packages. This makes it difficult for established firms to judge human talent within AI startups.

Our analysis of the six use cases shows that the best approach to finding the right AI startup for collaboration is for established firms to make use of credibility signals at both the individual and firm levels. The alternative of assessing a startup's products and offerings is often difficult because trying them out will likely incur costs in terms of both time and money. Even shorter proof-of-concept approaches may well require significant resource commitments. This is especially the case for AI technologies because their opaque algorithms are typically harder to assess than conventional IT products.

Recommendation for Addressing Challenge 1

Our overall recommendation for finding the right AI startup for collaboration is to *overcome the inscrutability of AI startups' capabilities and talent by systematically assessing related credible quality signals from trusted sources.*

Implementing this recommendation will require practitioners in established firms to systematically assess credibility signals at the firm level and individual level.

Recommendation 1(a): Leverage Firm-Level Quality Signals from Other Institutions. At the firm level, an AI startup's track record can serve as an indicator of its AI offerings' quality. In the absence of reference customers or showcase clients, the track record can include contests won, public grants and private investments. Such signals usually indicate that other stakeholders with AI-specific knowledge (e.g., specialized investors) have assessed the technical capabilities of the AI startup, which creates legitimacy.

For example, the Finnish telecoms firm in Use Case 1 had no experience in natural language processing or any other AI approach and thus did not possess the necessary capabilities to assess whether Startup 1's AI offerings could be beneficial. Instead, it used external quality signals to address the inscrutability challenge and to predict the potential impact of collaborating with Startup 1. These signals included Startup 1's success in winning smaller contests, the granting of significant public funds from the state of Austria for its natural language processing platform and its success in attracting first investors. Such visible, positive signals from trustable institutions that have the skills to judge AI capabilities made it easier for the Finnish telecoms firm to determine whether to partner with AI Startup 1.

In contrast, a global insurance company (Use Cases 4 and 5) assessed Startup 2 through a series of smaller initial proof of concepts of AI solutions to replace its legacy rule-based price calculations and rule-based claims handling. The visible results from these proof of concepts provided a sound basis for building and growing a long-term partnership. Though this approach was significantly slower and more costly compared to Use Case 1, it helped to overcome the inscrutability challenge and to build a solid foundation for jointly deploying AI-based solutions for a range of business challenges across functions and product lines.

Recommendation 1(b): Draw Conclusions from Non-AI Achievements of Startup Employees. At the AI startup's individual employee level, established firms should go beyond merely screening the technical knowledge of AI talent. They should also include more broadly interpretable signals such as knowledge in adjacent topics (such as strategy and IT), prior work experience (e.g., in well-known and successful companies) and academic achievements (e.g., published research papers and Ph.D. degrees). Signals such as these can inform the decision on whether to collaborate with an AI startup because they provide a rounded view of the startup's human talent.

For example, in Use Case 3 the software firm used the academic achievements of the other co-founder of AI Startup 1 as a proxy for his AI capabilities. The co-founder's list of academic publications indicated his outstanding position within the natural language processing research field and thus served as a legitimate quality signal for the established firm. Likewise, the financial services firm in Use Case 6 intensively screened the CVs of Startup 2's employees, which revealed the startup's approach to hiring.

It not only hired people with relevant skills needed to successfully deliver AI use cases (e.g., econometric, statistical and machine-learning skills, as well as IT engineering, strategy, project management and change expertise) but also with visible career successes (e.g., top schools, top grades, Ph.D. degrees supervised by renowned academics, papers in top publications and top-brand previous employers). This hiring approach made it possible for the established firm to judge the quality of Startup 2's AI talent and to decide to partner with this startup.

Challenge 2: Identifying the Right AI Use Case

Our six use cases indicate that established firms can benefit from the horizon-expanding outside-in perspective of AI startups by engaging in joint AI use case idea-generation workshops and then prioritizing use cases according to their expected impact. The challenge for established firms in identifying the right AI use case has two components: 1) their inability to understand the relevance of algorithmic activity in various application areas, and 2) their inability to understand the potential of algorithmic activity to generate value.

The established firms in all six AI use cases addressed the first component by coupling their push toward AI with an organization-wide "data orientation" paradigm. The strong focus on organizational data mostly stemmed from mainstream wisdom (e.g., "data is the new oil"), competitor pressure or capital market pressure (e.g., to expand into new data-driven business models), rather than the need for AI use cases needing vast amounts of training data. But by seeking to answer the question of "What can we do with our data?," these established firms tended to miss the question of "How can a new technology like AI improve our business?," and thus limited their ideas for AI use cases.

Regarding the second component, we found that established firms tended to prioritize AI use cases predominantly on feasibility considerations (e.g., data availability questions that are without doubt important for successful implementation) without first considering their organizations' goals. Without answering crucial business questions like "How can we generate the most impact with AI?," they run the risk that their data-inspired AI projects quickly fall victim to corporate cost-cutting. Thus, prioritizing AI use cases primarily on feasibility can create a bias toward use cases with limited impact potential.

Our observations of the six use cases showed that established firms benefitted from structured interactions with the AI startups during idea generation for AI use cases and their prioritization. We therefore recommend that established firms follow a structured process for AI use case selection that harnesses external inspiration (e.g., via showcases from other firms) and with a strong focus on creating impact. This is especially relevant for AI use cases because the data orientation arising from the need for vast amounts of training data can become predominant and even mask the need to consider impact.

Recommendation for Addressing Challenge 2

Our overall recommendation for identifying the right AI use case is to *leverage the outside-in perspective of AI startups to generate and prioritize AI use cases based on value creation potential above and beyond data availability.* As described below, this recommendation requires that established firms take two actions.

Recommendation 2(a): Engage in Joint AI Use Case Idea-Generation Workshops. Practical steps for AI use case idea generation should include identifying and prioritizing business goals, quantifying a baseline for value potential (e.g., understanding in which business units or product lines AI use cases could build on the largest datasets) and pushing the AI startup to show how its AI use cases can move economic impact levers.

For example, the established insurance company in Use Case 4 had a long history and thus a narrow view of how car fleet insurance premiums need to be calculated. The initial idea was to simply add additional variables to the existing rule-based approach. During a joint idea-generation workshop, Startup 2 presented a novel type of AI algorithm it had used in a different industry and that would be well-suited to improve risk prediction for car fleets.

Likewise, the established telecoms firm in Use Case 1 had a large dataset on local service providers' and callers' service needs. Its initial use case idea was to create an AI-based service recommender. The final idea for the use case was developed in a joint workshop: Startup 1 first outlined how natural language processing can create value in the telecoms industry and then invited participants to jointly generate impactful application ideas for their business units.

Recommendation 2(b): Jointly Prioritize AI Use Cases Based on Impact and Feasibility. Established firms can also leverage their AI startup partners' capabilities in estimating the impact potential and assessing feasibility. As outlined above, because Use Case 1's goal was to improve workforce efficiency, it was prioritized over the established firm's original use case idea, which would have been the first choice based on a feasibility ranking. The collaboration with Startup 1 helped the established firm to focus on the value perspective in several ways. First, Startup 1 highlighted the value contribution of its AI solution as a unique selling proposition and had detailed technical benchmarks on its natural language processing performance to justify its costs. Second, the startup pushed for a joint idea-generation and prioritization workshop with senior business representatives to ensure that all relevant decision makers were present. This workshop focused mainly on value discussions because of the predominantly business background of the established firm's participants.

The insurance company in Use Cases 4 and 5 went even further and engaged AI Startup 2 to build a firm-wide AI use case portfolio with detailed assessments of potential AI use cases alongside multiple impacts and feasibility categories.

Challenge 3: Agreeing on the Right Commercial Terms and Conditions

The inscrutability of AI makes it both necessary and challenging for an established firm to find the right incentive scheme for its collaboration with an AI startup. The necessity arises from the information asymmetry between established firms and AI startups, which makes it hard for an established firm to understand how much effort an AI startup partner will need to implement and deliver a well-performing AI use case. As the saying goes "If you are not part of the solution, there's good money in prolonging the problem," which makes it tempting for some AI startups to exploit the information asymmetry on implementation progress and project steering capabilities.

Additionally, some established firms are not fully aware that AI performance can be objectively measured and tracked, despite algorithmic inscrutability. Objective measurements can open up attractive options for creating new incentive schemes for AI collaborations. So-called A/B tests can help measure AI impacts by applying an AI innovation only to a representative subgroup of a population. Such tests are also commonly used to further improve AI innovations by comparing different versions of AI algorithms. However, there are costs associated with A/B tests: missed benefits for those not in the subgroup and the costs of technical modifications.

Our observations of the six use cases show that, for an AI solution to fully realize its intended impact, the collaboration partners need to align their incentives. This implies that practitioners in established firms should consider investing in the measurability of AI solution performance to reduce the information asymmetry between the firm and its AI startup partner, as both can benefit from successful AI use case implementation.

Recommendation for Addressing Challenge 3

Our overall recommendation for agreeing on the right commercial terms and conditions is to *track AI-based value creation and let the AI startups participate to align incentives*. The two actions that established firms need to take to implement this recommendation are described below.

Recommendation 3(a): Create Incentives with Performance-Based or Usage-Based Compensation. The data-driven learning and algorithmic improvement aspect of AI drives the need for compensation schemes, such as impact sharing. Setting up impact-sharing deals requires both parties to define clear key performance indicators and respective target levels. For example, in Use Cases 4 and 5 the established insurance company agreed to a payment scheme based on sharing long-term achieved impact: more than 50% of Startup 2's compensation depended on the improvement of financial figures—e.g., increased earnings or reduced claims costs. This impact-sharing deal led to a quick delivery in Use Case 4 and a quick decision to abandon implementation in Use Case 5.

Use Case 4 required Startup 2 to build a fundamentally new AI-based pricing algorithm based on additional data with novel methods. Nevertheless, it delivered the first prototype of the pricing algorithm in the first few months of the one-and-a-half-year-long project and later fine-tuned and optimized the prototype. It took an additional year to implement the associated IT system and integrate it with the AI-based solution and to roll out the system, conduct user tests and training, prepare documentation, gain internal approval, officially go live and hand over responsibility. Given the long duration of these tasks, Startup 2 could have easily used this as an excuse for taking longer to implement the AI algorithm. Many firms also promote a "culture of failure," which increases their tolerance of delays and budget increases.

In Use Case 5, management of the established insurance firm observed unexplainable performance differences between different claims handling teams and set up an extended testing period to gather sufficient data that could be used to subsequently adapt the AI algorithm. Even though only Startup 2 had enough insights to judge whether continued adaptations would deliver further benefits, it did not use this information advantage and called for a timely go/no-go decision on whether to cancel the project.

The established telecoms firm in Use Case 1 used an alternative approach to creating incentives for its AI-based call center system. Once the AI system was in production, it provided Startup 1 with financial incentives in the form of a pay-per-use scheme and thus prevented the startup from exploiting information asymmetries during the initial development stage. Instead, Startup 1 was compensated at this stage through a traditional flat fee.

Recommendation 3(b): Use the AI Startup to Build Impact-Measuring Capabilities. Established firms should invest in building impact-measuring capabilities to track value generation from AI use cases. Setting up an economically and ethically feasible impact-measurement process includes agreeing on how often the impact is measured, how responsibilities are distributed and how measurement costs are split. AI startup partners are often clearly able to build the desired capabilities but established firms should not underestimate the effort.

For example, the financial services firm in Use Case 6 decided that A/B tests were economically and ethically feasible. These tests required two different pricing schemes to run in parallel, but Startup 2 took three months to modify the IT system to enable this to happen. In contrast, A/B tests were not an option for the insurance company in Use Case 4, because market pressure required a quick and completely optimized pricing scheme for all car fleet insurance contracts. Instead, the insurance company and Startup 2 built tools for simulation, impact tracking and reporting to understand the impact of the AI-based solution.

Challenge 4: Considering the Impact on People

The autonomy characteristic of AI can lead to the creation of autonomous agents that are able to act without human intervention and guidance. Such agents significantly alter the interactions between AI and people.[33] AI use cases can have a much higher degree of autonomy, compared to other technologies and, as a consequence, can have a more radical impact on users. As user acceptance is often critical for realizing the envisioned impact, established firms face two specific challenges: 1) uncertainty about how to optimize the trade-off between automation gains and user acceptance, and 2) overcoming potential workforce resistance arising from fears about the role-changing potential of AI.

First, our observations of the six use cases indicate that many established firms do not understand the different degrees of AI autonomy that are possible. Firms may have a specific degree in mind, but sometimes do not systematically reflect whether more or less autonomy would yield even better results. The AI startup partners in the six use cases were often able to help optimize AI use toward high user acceptance and thus greater impact.

33 See Berente, N., Gu, B., Recker, J. and Santhanam, R. op. cit., September 2021.

Second, regardless of the chosen degree of AI autonomy, the mere possibility of a fully autonomous AI algorithm triggers intense user emotions, such as the fear of being replaced by AI. While traditional IT systems are now mostly seen as tools that support human activities, many employees are worried that AI-based systems could eventually take over human tasks entirely and thus make human workers obsolete. Moreover, the presence of AI startups can further increase such fears among employees, as employees often associate external implementation partners with either cost-cutting management consultants or IT providers interested in outsourcing deals. This likely makes achieving user acceptance of AI solutions developed in collaboration between established firms and AI startups even more difficult.

Established firms and their AI startup partners need to carefully consider two aspects of change management to balance the trade-off between automation gains and user acceptance: on the one hand, they need to decide how much change they want to create by fine-tuning how much autonomy to give to the AI solution. On the other hand, they need to jointly design a delivery mode that ensures that users are fully on board.

Recommendation for Addressing Challenge 4

Our overall recommendation for considering the impact on people is to *consciously consider the right degree of AI autonomy for each AI use case and position the collaboration with AI startups accordingly*. The two actions that established firms should take to implement this recommendation are described below.

Recommendation 4(a): Include the AI Startup's View on How to Adjust AI Autonomy over Time. Established firms should understand all possible options within the spectrum of autonomy, choose an initial degree of autonomy that fits the AI impact aspirations and user acceptance and, from the outset, plan future adaptations to the degree of autonomy. The AI startup's perspective should be included in this process because it will often have a deeper technical understanding of AI autonomy options and possible transition paths between options.

A good starting point is to initially implement an AI-based solution with low autonomy because this will create a showcase that reduces fear and becomes well-accepted by the workforce and clients, especially if their buy-in is critical for success. With growing user acceptance, there may be opportunities in the future for step-wise increases in AI autonomy that result in the realization of primary impact goals, such as efficiency gains.

For example, the telecoms firm in Use Case 1 wanted to demonstrate that AI and employees can efficiently work together. Startup 1 showed how its AI solution could first be used to support call center agents and later, if needed, take over the whole process. The firm thus chose to initially use AI in a low autonomy role so that the role of call center agents only slightly changed from being fully in charge of customer interactions to moderating a conversation recommended by the AI system. Overall, this initial deployment of AI led to 50% efficiency gains while meeting customer expectations and fully ensuring employee acceptance. This initial deployment also laid the foundation for potential greater autonomy in the future.

Likewise, Use Case 3 was consciously given low AI autonomy. Supporting emergency dispatchers by filling out forms in the background but keeping them ultimately responsible was the optimal setting at that moment to improve call quality in life-critical situations and ensure user acceptance.

User acceptance of AI solutions with medium or high degrees of autonomy can be more difficult to achieve. Although Use Cases 4 and 5 had similar degrees of autonomy, the outcome and user acceptance were completely different. Use Case 4 was a well-accepted AI-based tool for underwriters that kept them in control of the final prices. The key to this acceptance was that underwriters were involved in the design process from the beginning. During the design process, the insurance company and Startup 2 took full account of the underwriter's initial worries. They chose to give underwriters certain options for price adjustments so that underwriters could still modify price recommendations from the AI system based on their extensive and deep market experience. Apart from requiring a more complex front end, this choice made little difference for the implementation activities of Startup 2 and the overall impact of the AI use case but greatly preserved the role of the underwriters. Recent tracking of underwriters' activity showed that they rarely modify the recommended prices. Thus, there is future potential to further increase AI autonomy in the pricing process if desired.

In contrast, the claims handlers in Use Case 5 perceived that AI's involvement interfered with their ways of working and they largely ignored the AI-based recommendations. As a consequence, AI Use Case 5 was canceled.

Recommendation 4(b): Clarify the Role of the AI Startup Partner in the Change Process. Our analysis of the six use cases shows there are two reasonable, practical options for positioning the collaboration between an established firm and an AI startup: positioning the startup either as a helpful knowledge provider or as a neutral solutions provider. We found that with increasing autonomy of the AI solution, the startup positioning tended more toward that of knowledge provider.

For example, the financial services firm in Use Case 6 engaged Startup 2 to train and coach users to perform well according to the requirements of the AI solution. Once users had been taught how to work in a new development environment, how to continue improving AI algorithms and how to handle AI algorithms training and deployment, they were prepared to take over responsibility for the new AI solution. Most users welcomed the upskilling that came with the role change. Likewise, in Use Case 4, employees from Startup 2 were embedded with the team of underwriters and actuaries to jointly design and build the new AI-based pricing system.

In contrast, the software company in Use Case 3 leveraged Startup 1 as a neutral research partner to support its internal research and development teams. Likewise, in Use Case 2, Startup 1 acted as a neutral solution provider. Despite the high degree of autonomy in this use case, the positioning worked well because the AI solution wasn't automating an existing business process and thus did not require much change by the software company's employees.

Challenge 5: Overcoming Implementation Roadblocks

The success of AI use cases depends on 1) a multitude of stakeholders involved in AI solutions, and 2) a large number of prerequisites that have to be in place if expectations are to be met. The learning characteristic of AI means that a wide variety of stakeholders in the organization are involved in AI use cases and have to collaborate in complex cross-functional teams with diverse knowledge backgrounds.[34] IT engineers now need to collaborate with data scientists who can create algorithms with opaque and probabilistic logic that are able to learn and with data engineers who can handle large-scale training data sets.[35] All of these stakeholders have to collaborate with domain experts to improve the overall performance of the AI solution.[36] This creates more dependencies and interaction needs between all involved stakeholders and makes it even more challenging for AI startups to navigate within the complex intra-organizational networks and hierarchies of an established firm. For example, involving hundreds of claims handlers in Use Case 5 was more challenging compared to the small team of underwriters in Use Case 4.

In addition to the larger circle of involved stakeholders, AI use case success depends on multiple prerequisites for meeting AI impact expectations. Success depends not only on supportive management practices, well-working business processes and adaptable IT systems, but also on the availability of training data and strong algorithmic performance. Moreover, all of these prerequisites must be in place at the same time.

Overall, our use case analysis indicates that established firms should assess whether their organizational maturity is sufficient for AI to positively affect existing organizational units, processes and IT systems. If this is not the case, building new and independent organizational units, processes and IT systems with the help of an AI startup may increase the chances of successful AI use cases. When using such a greenfield approach that doesn't rely on existing organizational legacy processes and systems, practitioners should also consider the additional complexity that a parallel solution creates and plan how to gradually shift work from existing

34 Li, J., Li, M., Wang, X. and Thatcher, J. B. "Strategic Directions for AI: The Role of CIOs and Boards of Directors," *MIS Quarterly* (45:3), September 2021, pp. 1603-1643.

35 Berente, N., Gu, B., Recker, J. and Santhanam, R., op. cit., September 2021.

36 van den Broek, E., Sergeeva, A. and Huysman, M., op. cit., September 2021.

organizational units, processes and IT systems to the new AI solution.

Recommendation for Addressing Challenge 5

Our overall recommendation for overcoming implementation roadblocks is to *work with or around stakeholders to create prerequisites for AI use case success, depending on internal organizational maturity and AI startups' capabilities. Two actions are needed to implement this recommendation.*

Recommendation 5(a): Ensure that the AI Startup Successfully Collaborates with Stakeholders. AI Use Case 4 took one-and-a-half years to complete. The most time-consuming part was not building the AI algorithm itself, but rather aligning with all other stakeholders in the company for delivery integration. This alignment involved convincing management to adjust its steering logic, persuading business representatives to adapt processes, involving users to generate buy-in, agreeing with the central IT department on a new operating model with shared product ownership, requesting the business unit IT department to build a novel type of interface to existing IT systems and going through endless loops with the data security and IT architecture teams to get their approval. All these stakeholders had their own goals and their established ways of dealing with change.

To facilitate stakeholder interactions, the insurance company in Use Case 4 created a joint cross-functional team with Startup 2 and ensured that internal team members were always available to effectively handle stakeholders. However, the central IT department's unwillingness to assume product ownership for the AI part of the solution almost became a showstopper in the final project phase. The IT department was hesitant because it had no experience or understanding of the AI retraining and deployment processes and timelines.[37] The insurance company resolved this challenge by making the business units' decentralized IT department the owner of the AI part and by engaging Startup 2 to ensure knowledge transfer.

Likewise, the telecoms firm in Use Case 1 leveraged the well-defined technical interfaces of Startup 1's natural language processing solution and thus reduced the IT change effort in the operational systems to a handful of person-days. Both this use case and Use Case 4 show that overcoming implementation roadblocks becomes easier if the AI startup has all the necessary skills, not only AI and data experts but also business experts and IT engineers, to smoothly collaborate with internal stakeholders.

Recommendation 5(b): Use the AI Startup to Bypass Roadblocks. An alternative approach to overcoming implementation roadblocks when organizational maturity is low is to reduce dependencies on existing internal stakeholders and work around them. Adopting a greenfield approach and leveraging the AI startup's capabilities may be the best way forward if existing processes, IT systems or organizational units cannot be changed with reasonable effort. For example, the logistics firm in Use Case 2 had to overcome the roadblock caused by the lack of training data by leveraging its suppliers because no internal stakeholder of the logistics firm could commit to creating training data.

Similarly, integrating Use Case 6 into the end-of-lifecycle legacy operational IT systems would have been a multiyear project. Instead, the financial services firm leveraged Startup 2's software engineering capabilities and commissioned the startup to build a new IT system capable of delivering the new prices without the support of the firm's IT teams. This reduction of dependencies on internal stakeholders ensured much quicker delivery of the AI use case.

Concluding Comments

In this article, we have identified several crucial challenges for organizing AI innovation in collaboration with startups and provided recommendations for addressing these challenges. Specifically, we recommend looking out for credible signals of AI startups' capabilities, identifying AI use cases based on value creation potential instead of data availability, tracking AI-based value creation and allowing partners to participate in and align incentives. Established firms should also consider the degree of autonomy each AI use

37 Our observations here are in line with prior research that provides a more detailed account of AI deployment challenges. See, for example, Benbya, H., Davenport, T. H. and Pachidi, S., op. cit., December 2020.

case should have, and carefully check whether they are sufficiently mature to realize AI value creation. We believe that our findings will help practitioners in established firms successfully manage collaborations with AI startup partners. We also hope that our findings will encourage future research that develops more detailed theoretical insights for overcoming the challenges of collaborating with startups for AI innovation.

Appendix: Interview Questions

Step 1: Open Questions Used for Structuring the Initial Observation Gathering

A: AI startup's perspective on the collaboration:

- Which challenges did the AI startup face?
- How did the partner firm perceive these challenges?
- How did the AI startup try to overcome these challenges?
- How did the partner firm provide support for overcoming these challenges?
- What made the chosen mitigation actions effective?

B: Established firm's perspective on the collaboration:

- Which challenges did the firm face?
- How did the AI startup perceive these challenges?
- How did the firm try to overcome these challenges?
- How did the AI startup provide support for overcoming these challenges?
- What made the chosen mitigation actions effective?

Step 2: Classification Scheme to Determine Relationship to the Three Distinguishing Characteristics of AI

A: Observations not related to any of the three characteristics

B: Observations related to challenges related to at least one of the three distinguishing characteristics of AI—e.g., triggered or intensified by at least one of the three characteristics:

- Autonomy: Challenges related, e.g., to AI solutions increasingly acting on their own as agents or reducing the action range of involved humans
- Learning: Challenges related, e.g., to the improvement of AI algorithms based on large volumes of data, and to AI and people learning together
- Inscrutability: Challenges related, e.g., to opacity, transparency, explainability and interpretability of complex or self-evolving AI algorithms

About the Authors

Jana Oehmichen

Jana Oehmichen (j.oehmichen@uni-mainz.de) is a full professor of organization, HR and management studies at the University of Mainz and an honorary professor of leadership and governance at the University of Groningen. Her work has been published in journals including *Strategic Management Journal*, *Organizational Behavior and Human Decision Processes*, *Leadership Quarterly*, *Global Strategy Journal* and *Journal of Management Studies*. Her research is focused on digital leadership and strategy as well as on corporate governance. She serves as an associate editor for *Schmalenbach Journal of Business Research* and as an editorial board member for Corporate Governance: *An International Review.*

Alexander Schult

Alexander Schult (alexander.schult@ondewo.com) is an entrepreneur and consultant specializing in artificial intelligence, data analytics and digital business models. Prior to his current involvement in multiple startups (including ONDEWO GmbH, which he co-founded), he worked for almost nine years at McKinsey & Company, focusing on digital transformations and strategy for insurance, healthcare and technology clients. He holds a doctoral degree in management. His work has been published in journals such as *Human Resource Management* and *Schmalenbach Business Review*, and in the proceedings of the International Conference on Information Systems.

John Qi Dong

John Qi Dong (john.dong@ntu.edu.sg) is an associate professor with tenure at Nanyang Business School, Nanyang Technological University. His research bridges information systems and strategic management and is focused on digital innovation and technology strategy. His work has been published in *MIS Quarterly*, *Strategic Management Journal*, *Journal of Management, Journal of Management Information Systems*, *Journal of the Association for Information Systems*, *European Journal of Information Systems*, *Information Systems Journal* and *Journal of Strategic Information Systems*. He serves as an associate editor for *MIS Quarterly* and *Journal of the Association for Information Systems*.