MIS Quarterly Executive

Volume 23 | Issue 2 Article 5

June 2024

How Audi Scales Artificial Intelligence in Manufacturing

André Sagodi

Benjamin van Giffen

Johannes Schniertshauer

Klemens Niehues

Jan vom Brocke

MIS Quarterly Executive

How Audi Scales Artificial Intelligence in Manufacturing

For organizations to realize maximum value from artificial intelligence (AI), they need the capability to scale it and must consider scaling throughout all stages of an AI innovation project. But AI scaling presents significant challenges, especially for manufacturing companies. We describe how Audi, a leading automotive manufacturer, scaled its crack detection AI solution and unlocked long-term business value in manufacturing. Based on lessons learned at Audi, we provide recommendations and actions for CIOs and senior leaders who seek to capture value through scaling AI solutions.[1,2]

André Sagodi
University of St. Gallen (Switzerland)
AUDI AG (Germany)

Benjamin van Giffen
University of St. Gallen (Switzerland)
European Research Center for Information Systems (ERCIS)
University of Liechtenstein (Liechtenstein)

Johannes Schniertshauer
AUDI AG (Germany)

Klemens Niehues
AUDI AG (Germany)

Jan vom Brocke
University of Münster (Germany)
European Research Center for Information Systems (ERCIS)
University of Liechtenstein (Liechtenstein)

Realizing Business Value from Artificial Intelligence Scaling Is Challenging

Artificial intelligence (AI) is becoming increasingly pervasive in contemporary society and is expected to create considerable economic value for both society and industry.[3] AI technologies

1 Varun Grover is the senior accepting editor for this article.
2 The authors thank AUDI AG for sharing its practical experiences in scaling artificial intelligence. We also thank Varun Grover and the review team for their invaluable feedback and advice throughout the review process, which helped us to improve the article with each revision.
3 The McKinsey Global Institute estimates that AI has the potential to create about $13 trillion in additional economic value worldwide by 2030. See Bughin, J., Seong, J., Manyika, J., Chui, M. and Joshi, R. *Notes from the AI Frontier: Modeling the Impact of AI on the World Economy*, McKinsey Global Institute, September 4, 2018, available at https://www.mckinsey.com/featured-insights / artificial-intelligence/notes-from-the-ai-frontier-modeling-the-impact-of-ai-on-the-world-economy.

DOI: 10.17705/2msqe.00094

automate human-like tasks, such as sensing, perceiving, interacting with the environment, problem solving, learning and decision-making.[4] There are various definitions of AI, but many of the current AI deployments are based on machine learning technologies.[5] The steep rise in interest in AI has been fueled by advances in AI-based algorithms and growing computing power. Many companies are launching initiatives to capture AI's value potential for their business. Gartner estimates that by the end of 2024, three quarters of organizations will move from piloting to operationalizing AI.[6]

However, when incorporating new technologies like AI, organizations can encounter unforeseen barriers that hinder value realization or even impair business value. As a result, the business value captured from AI may be below expectations. Executives and industry experts agree that deploying AI at scale is crucial for capturing business value.[7] Particularly in industrial processes, the scaling of AI is seen as a value driver.[8] According to a global study of 1,500 C-suite executives, strategically scaling AI triples the return from AI investments, compared to companies that rely on AI proof of concepts. Indeed, 84% of respondents stated that they believe they must scale AI to meet their growth goals.[9] Moreover, recent research on AI indicates that scaling is essential and that companies must strategically enable the scaling of AI applications to create value.[10] However, it is challenging to scale AI to reap the benefits and achieve a positive return on AI investments.[11,12] For example, IBM reported that 90% of surveyed organizations were struggling to scale AI.[13] Thus, many organizations are at the beginning of their journey when it comes to scaling AI.[14]

In this article, we refer to AI scaling as the organizational capability to innovate and leverage AI-based applications successfully to realize business value.[15] To better understand how organizations can build this capability, we studied the case of Audi to answer the question: *How can organizations realize AI value through scaling?* (Our research methodology is described in Appendix A.)

Many automotive companies face AI implementation challenges, and only a few currently implement AI at scale, preventing most of them from capitalizing on their AI investments.[16] Audi is one exception. We report on Audi's four-year journey to innovate and scale an AI-based quality inspection system in its press shops. We describe the stages Audi traversed as it scaled its AI-based innovation and show how it addressed AI-specific scaling challenges to reap long-term benefits from AI. Based on our analysis of the lessons learned by Audi, we provide recommendations and actions for senior executives and CIOs seeking to leverage

4 Benbya, H., Pachidi, S. and Jarvenpaa, S. L. "Special Issue Editorial: Artificial Intelligence in Organizations: Implications for Information Systems Research," *Journal of the Association for Information Systems* (22:2), 2021, pp. 281-303.

5 Lacity, M. and Willcocks, L. "Becoming Strategic with Intelligent Automation," *MIS Quarterly Executive* (20:2), June 2021, pp. 1-14.

6 *Gartner Identifies Top 10 Data and Analytics Technology Trends for 2020*, Gartner Press Release, June 22, 2020, available at https://www.gartner.com/en/newsroom/press-releases/2020-06-22-gartner-identifies-top-10-data-and-analytics-technolo.

7 Fountaine, T., McCarthy, B. and Saleh, T. "Getting AI to Scale," *Harvard Business Review*, May-June 2021, available at https://hbr.org/2021/05/getting-ai-to-scale.

8 Busch, R. *Industry Is Where the Real Potential of AI Lies*, World Economic Forum, January 21, 2019, available at https://www.weforum.org/agenda/2019/01/industry-is-where-the-real-potential-of-ai-lies/.

9 Reilly, A., Depa, J. and Douglass, G. *AI: Built to Scale*, Accenture, November 19, 2019, available at https://www.accenture.com/us-en/insights/artificial-intelligence/ai-investments.

10 van Giffen, B. and Ludwig, H. "How Siemens Democratized Artificial Intelligence," *MIS Quarterly Executive* (22:1), March 2023, pp. 1-21.

11 *Scaling AI in Manufacturing Operations: A Practitioners' Perspective*, Capgemini Research Institute, 2019, available at https://www.capgemini.com/insights/research-library/scaling-ai-in-manufacturing-operations.

12 "How to Scale AI in Your Organization," *Harvard Business Review*, March 2022, available at https://hbr.org/2022/03/how-to-scale-ai-in-your-organization.

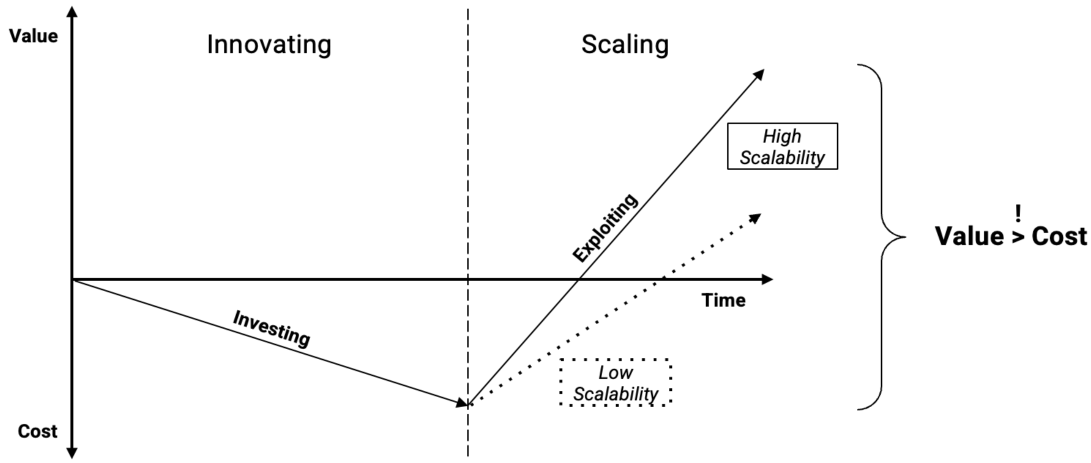
13 *Proven Concepts for Scaling AI*, IBM, 2020, available at https://www.ibm.com/thought-leadership/institute-business-value/en-us/report/scaling-ai.

14 Organizations will likely acquire and develop the necessary competencies stepwise to fully leverage the capability to scale AI, as research has found for IT capabilities such as IT architecture competencies. See Ross, J. W. "Creating a Strategic IT Architecture Competency: Learning in Stages," *MIS Quarterly Executive*, (2:1), March 2008, pp. 31-43.

15 Starting from a use case perspective, we focus on the ability to leverage the value of a particular AI use case in the context of manufacturing. Previous research has focused on the ability to leverage the value of a digital portfolio. See, for example, Alfaro, E., Bressan, M., Girardin, F., Murillo, J., Someh, I. and Wixom, B. H. "BBVA's Data Monetization Journey," *MIS Quarterly Executive* (18:2), June 2019, pp. 117-128.

16 According to the Capgemini Research Institute, in 2019 only 10% of automotive companies had implemented AI at scale. See *Accelerating Automotive's AI Transformation: How Driving AI Enterprise-Wide Can Turbo-Charge Organizational Value*, Capgemini Research Institute, available at https://www.capgemini.com/insights/research-library/accelerating-automotives-ai-transformation/

Figure 1: Conceptual Model of Scaling and Scalability of Digital Innovations



the value potential of AI through scaling. These recommendations will enable leaders to organize the scaling of AI in their organizations. First, however, we define what we mean by AI scaling and AI scalability.

Defining AI Scaling and AI Scalability

Scaling and scalability are commonly used terms in the information systems community. They are often associated with lower marginal costs—e.g., for scaling digital infrastructure[17] or the user base of digital initiatives.[18] According to Gartner: "Scalability is the measure of a system's ability to increase or decrease in performance and cost in response to changes in application and system processing demands."[19] A scalable database can, for example, handle more queries without performance drops. In this sense, scalability is achieved by decoupling outputs (e.g., value realized) from inputs (e.g., resources used), with greater scalability implying greater decoupling.

It is important to distinguish between innovating and scaling (see Figure 1, which depicts our conceptual model of scaling and the scalability of digital innovations). During the *innovating* stage, investments in new products, services or business models are being made, with the inherent risk of failure and uncertainties related to resource consumption, return on investment and innovation success. During the *scaling* stage, an innovation is exploited in a value-oriented manner. A steeper (or exponential) value realization trajectory is associated with higher scalability.

The core of sustainable AI-based value generation is a scalable AI system.[20] To achieve this, organizations first invest in AI-based innovations built for scalability from the outset and then exploit the scalability of these innovations to create value. Technically, AI scaling is the continual increase in data volumes, AI model size and data processing capacity, while

17 Henfridsson, O. and Bygstad, B. "The Generative Mechanisms of Digital Infrastructure Evolution," *MIS Quarterly* (37:3), September 2013, pp. 907-931.

18 Huang, J. C., Henfridsson, O., Liu, M. J. and Newell, S. "Growing on Steroids: Rapidly Scaling the User Base of Digital Ventures Through Digital Innovation," *MIS Quarterly* (41:1), March 2017, pp. 301-314.

19 *Gartner Glossary*, available at https://www.gartner.com/en/information-technology/glossary/scalability.

20 Wixom, B. H., Someh, I. A. and Gregory, R. W. *Scaling AI To Generate Better and Different Outcomes*, MIT Sloan School of Management Center for Information Systems Research, Research Briefing XXI-12, December 16, 2021, available at https://cisr.mit.edu/publication/2021_1201_ScalingAI_WixomSomehGregory.

Table 1: AI Scaling Challenges and Practical Implications

Description	Practical implications
Challenge 1: Learning Requirements	
• AI algorithms learn from large amounts of data. • The underlying data quality determines the AI system's performance. • The AI training data required depends on the AI task complexity and the number of features in the dataset.	• Collecting training data from distributed deployments of the application • Ensuring data quality reflects the business understanding of the entire user base • Verifying that the training data is representative of the complexity of the scaled AI task
Challenge 2: Probabilistic Reasoning	
• AI performance is initially uncertain, and performance estimates evolve from vague to increasingly robust statements. • AI systems produce probabilistic outcomes; decision quality can vary depending on contextual conditions.	• Exploring the achievable AI performance for different deployments of a solution without knowing the contextual conditions • Safeguarding AI decision quality for every deployment of a scaled AI system during real-world operations
Challenge 3: Data Processing Requirements	
• AI systems need to be integrated into legacy IT infrastructures and processes. • AI system development and operations are data- and compute-intensive.	• Establishing operational AI solutions • Providing resources for large-scale data processing and algorithm training

the resources used remain stable (or are even reduced).[21]

The high scalability of an AI system enables high value-cost decoupling in the scaling stage and appears in three different forms: First, it increases the value contribution of a single deployment of an AI system without increasing the cost of that deployment (e.g., improved system performance). Second, it increases an AI application's cumulative value contribution through more deployments of that application. Third, it reduces the cost of an AI application through synergies between multiple deployments of the application (e.g., distribution of operating costs).

21 In this study, we focus on the challenges organizations encounter when scaling AI. Other research has recognized that successful digital transformation depends crucially on organizations establishing a robust (technical) backbone for efficiency and operational excellence as well as digital platforms that enable business agility and rapid innovation. For more information, see Sebastian, I. M., Ross, J. W., Beath, C., Mocker, M., Moloney, K. G. and Fonstad, N. O. "How Big Old Companies Navigate Digital Transformation," *MIS Quarterly Executive* (16:3), September 2017, pp. 197-213.

Challenges of Realizing AI Value Through Scaling

To derive business value and benefits from AI technology, organizations must address three AI-specific scaling challenges, which are summarized in Table 1 and described below.

Challenge 1: Learning Requirements

AI systems require large amounts of data in order to successfully learn. Data quality, both during learning and operations, must be sufficient to reach and maintain the desired level of AI prediction performance. The training data required depends on the AI task complexity, the algorithm choice and the underlying dataset features. These learning requirements have three implications for scaling AI:

1. Scaling requires a considerable volume of data to enhance an AI system's prediction performance. This requires collecting and using data from diverse sources to enable further learning and improvement of the AI model.

2. Data quality becomes increasingly critical as an AI system is scaled. Consistent validation, careful labeling and rigorous data management across an increasing user base are vital to maintaining data quality and facilitating the scalability of AI systems.
3. Continuous user engagement is critical to driving AI model refinement and preventing model degeneration. Feedback loops empower users to report issues, providing further insights.

Operating and scaling AI systems depends on collecting new data and using it to further reinforce the predictive quality of the AI model. In addition, involving the user base in improving the AI model ensures that the scaled AI system matches the organization's requirements and expectations.

Challenge 2: Probabilistic Reasoning

AI systems generate probabilistic outcomes. At the outset, developers rely on uncertain assumptions and hypotheses about how an AI system will perform. Initial experiments, data collection and AI model training and fine-tuning are required to gain confidence and gradually turn initial uncertainties into increasingly robust statements about performance. As AI algorithms learn from historical data and draw conclusions based on actual data in real-world contexts, the quality of their decisions will vary, depending on the contextual conditions in which they operate. We refer to this as "AI context sensitivity." But because AI algorithms learn and improve over time by absorbing more data, their prediction performance will generalize better. However, this characteristic can also lead to operational performance fluctuations. As a consequence, AI's probabilistic reasoning has two practical implications for scaling:

1. The achievable performance of an AI system, and thus its value potential, must be explored and evaluated for different deployments and without knowing the precise future contextual conditions in which the system will or could be scaled.
2. The decision quality must be safeguarded for every deployment of a scaled AI system to ensure that value is generated.

AI systems do not crash or stop operating; rather, they produce incorrect results, which is why monitoring AI is more complex than traditional rule-based systems and requires approaches to systematically detect and remediate deviations from the expected performance of each deployment of the AI solution.

Challenge 3: Data Processing Requirements

Operating AI at scale requires substantial data processing capabilities and computational resources to process and extract information from data during learning. This learning, typically carried out using high-performance graphics processing units and specialized hardware, is used to develop and iteratively refine accurate AI models that involve parameter-rich mathematical operations on vast datasets. These data processing requirements have two implications for AI scaling:

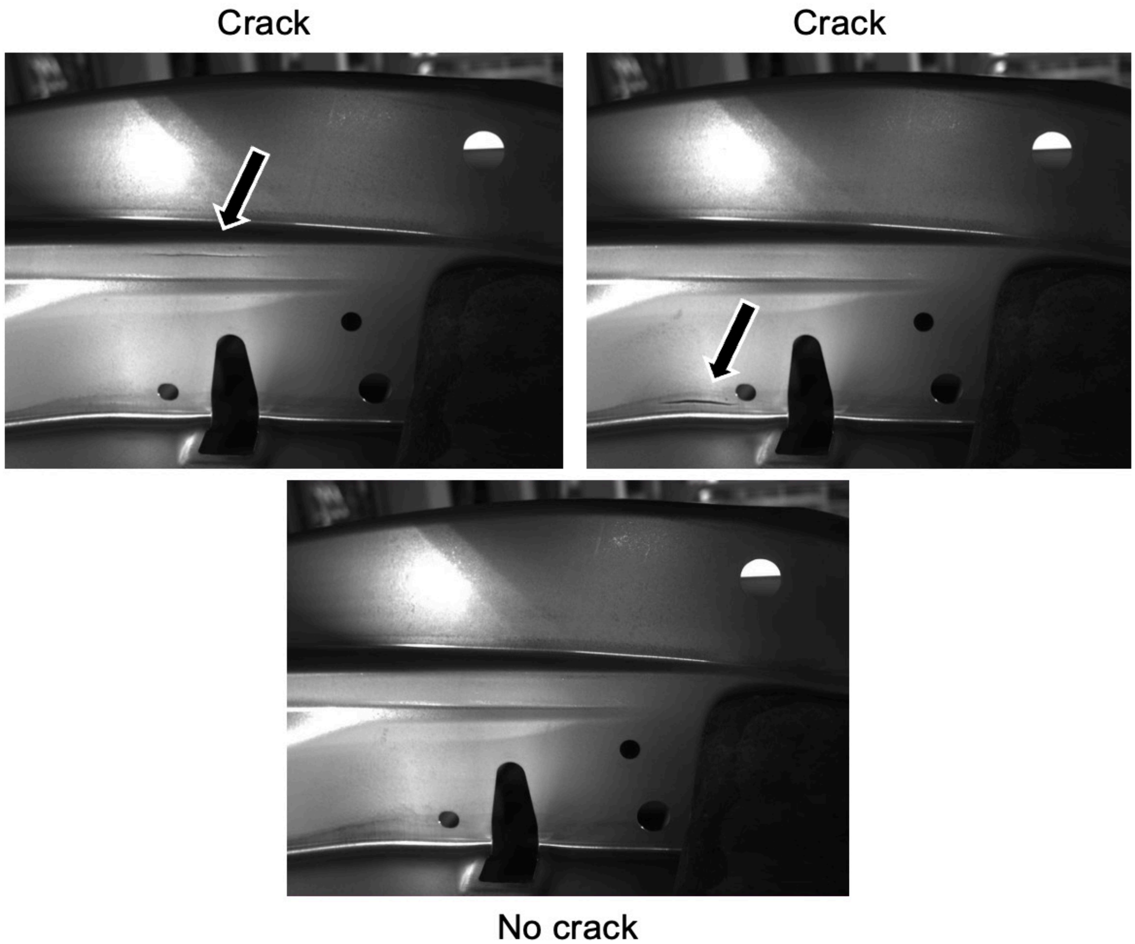
1. Organizations must integrate and adapt AI technologies to work with legacy systems while also ensuring compatibility and compliance with established processes.
2. Organizations need to allocate the required resources for large-scale data processing and the training of complex algorithms to develop and operate AI systems effectively.

Corporate Background of AUDI AG

AUDI AG is headquartered in Ingolstadt, Germany, and has been part of the Volkswagen Group since 1966. It has more than 87,000 employees and delivered over 1.9 million vehicles, including a growing number of fully electric models, to customers in 2023. It has an operating margin of 9.0% and a net cash flow of €4.7 billion ($5.0 billion).[22,23]

AUDI AG manufactures cars worldwide at 21 production sites in 12 countries and estimates that digitalization, including AI, can reduce factory costs by up to 30%.[24,25] With today's ever shorter product life cycles, flexibility[26] and adaptability[27] have become even more critical capabilities in the company's production system.

Figure 2: Sample Images of Cracked and Non-Cracked Sheet Metal Parts



The Four Stages of Audi's AI Scaling Journey

AUDI AG, one of the world's leading premium car manufacturers, uses AI-based visual inspection to automate the detection of cracks in sheet metal parts in its press shops, a crucial step in achieving zero-defect quality standards in its press shops. An overview of AUDI AG is provided in the text box below.

Given that a single press shop can produce 3.1 million parts each year and that roughly one in every 1,000 parts has a crack, crack detection is a highly significant part of ensuring zero defects. Figure 2 shows images of cracked and non-cracked sheet metal parts. Audi's AI-based crack detection system has been implemented and scaled at its production sites in Ingolstadt and Neckarsulm with several deployments at each site.

22 *After a Solid Fiscal Year 2023: Audi Strengthens and Expands Its Product Portfolio*, AUDI AG, May 19, 2024, available at https://www.audi-mediacenter.com/en/press-releases/after-a-solid-fiscal-year-2023-audi-strengthens-and-expands-its-product-portfolio-15957.

23 Currency conversion as of May 2024.

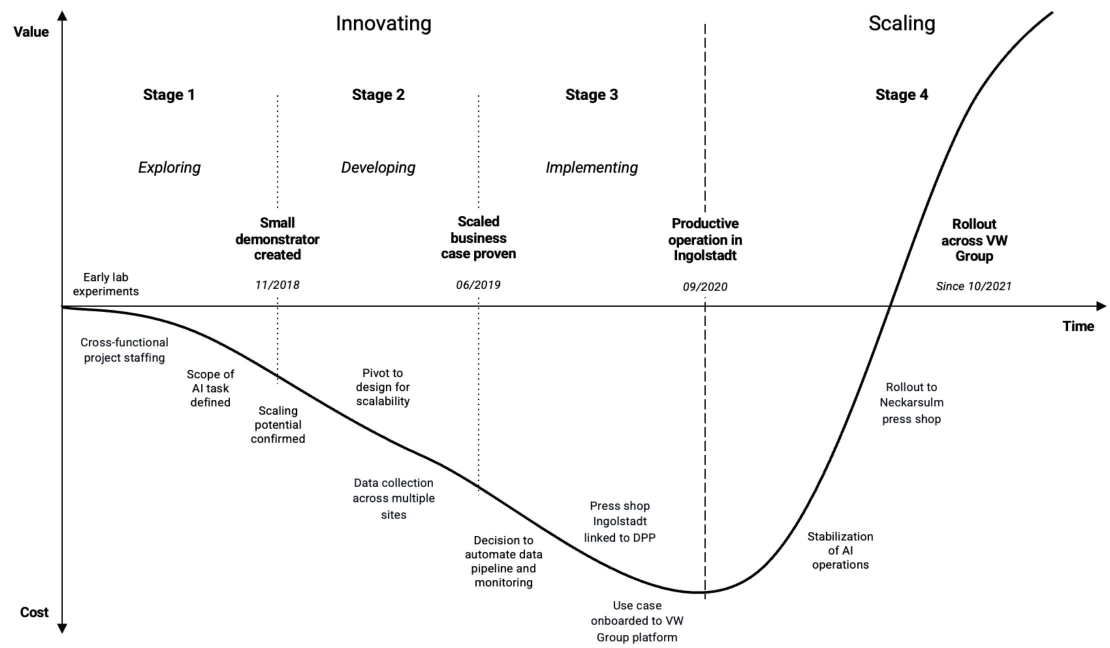
24 *Locations of the Audi Group*, AUDI AG, 2023, available at https://www.audi.com/en/company/profile/locations.html.

25 *AI in production at Audi: A perfect field of application*, AUDI AG, 2022, available at https://www.audi.com/en/innovation/development/ai-in-production.html.

26 In this context, "flexibility" is a system's ability to handle a high degree of variance in production operations—for example, different vehicle models being built on the same production line.

27 "Adaptability" is the ability to reconfigure a system with little time and financial effort when new product types or part variants are added.

Figure 3: Timeline and Key Events of Audi’s AI Scaling Journey



Audi innovated the AI system right from the outset for scalability. In addition to successfully implementing its AI-based crack detection system in the two manufacturing plants, the company chose to integrate it into a strategically initiated computer vision platform within the Volkswagen Group to facilitate scaling its use beyond Audi. Figure 3 depicts the timeline of Audi’s AI scaling journey, indicating key events throughout the three key innovation stages and the long-term scaling stage.[28]

Stage 1: Exploring AI-Based Crack Detection for Scalability

In 2017, Audi began to extensively engage with AI technologies, primarily for autonomous driving, and showcased AI-based traffic detection during internal technology days. During one of these events, experts from the press shop recognized the immense value potential of AI-based image analysis for the detection of cracks in sheet metal parts. As recalled by a press shop engineer: “I thought to myself if one can detect and locate various objects in urban traffic, why not a crack on deep-drawn sheet metal parts?” Audi’s exploration of AI-based crack detection, which lasted from late 2017 until November 2018, culminated with a proof of concept that demonstrated the technical feasibility of using AI for crack detection.

Next, experts from Audi’s press shops and AI specialists jointly specified the requirement for a minimum viable product, which entailed defining the key spots where crack detection was needed and classifying images either as “no crack” or “crack.” This requirement was based on two key considerations. First, AI experts emphasized the importance of starting with a small and narrowly defined AI task. Second, the detection of cracks in sheet metal parts is an essential task in the deep-

28 The cost-value curve shown in the figure was independently reviewed for plausibility with internal project stakeholders at Audi.

Table 2: Exploring Crack Detection AI for Scalability at Audi*

Case-Specific AI Scaling Challenges	AI Scaling Actions	L	P	D
Identifying AI applications that are scaling-relevant for the business	Domain and AI experts pinpointed the crack detection task as the minimum AI project scope, balancing business relevance and AI feasibility estimates.	X	X	
	An experienced press shop expert temporarily moved from the press shop to the Audi Production Lab to co-create the prototype with a machine learning engineer.	X	X	
Envisioning a scaled AI application based on a local AI prototype	The project team created a small demonstrator to illustrate the value proposition of an AI-based crack detection system throughout the organization.		X	
	A group-wide press shop committee confirmed the business need for a scaled AI solution and approved a uniform set of requirements.	X		

* AI scaling challenges: (L) Learning requirements, (P) Probabilistic reasoning, (D) Data processing requirements

drawing process that must be revised every time a new vehicle is designed.[29]

To gain a more comprehensive understanding of the idea of using AI for crack detection and to co-create the initial pilot, a press shop expert temporarily transferred from the press shop to the Audi Production Lab,[30] a dedicated technology development department. Audi also assigned a machine learning engineer from the corporate IT organization to the innovation project to provide essential AI expertise and guidance.

The project team evaluated and compared the AI-based crack detection approach against conventional non-AI systems. This comparison, which is detailed in Appendix B, indicated that a scalable AI solution could yield substantial long-term benefits. Furthermore, the team developed a proof of concept to demonstrate AI's crack detection capability to a broader user base. This pilot application was well received and led to a strong demand for AI-based crack detection by the managers of the individual press plants. A press shop committee validated the paramount importance of a crack detection AI solution and approved the scope and requirements of the application in collaboration with the project team.

In summary, the activities in the *exploring* stage of Audi's AI scaling innovation project can be summarized under the two headings shown in Table 2. The table also shows which of the three AI scaling challenges each action addresses.

Stage 2: Developing AI for Scalability

Given the commitment of the press shop committee to an AI solution, the project team next designed and developed a scalable AI-based solution that could be deployed multiple times in one or more press lines at Audi's Ingolstadt press shop.[31] (This development stage lasted from November 2018 to June 2019.) The team wanted to build one universal AI model for crack detection to avoid the emergence of a fragmented AI model landscape and to leverage data

29 Design is one of the main reasons people buy a vehicle. Progressive vehicle design requires increasingly high degrees of forming sheet metal parts with sharp contours. This requirement maxes out the technical possibilities, resulting in narrow process windows for deep drawing. The sporadic appearance of defects cannot be ruled out entirely.

30 *Audi Production Lab: The Link between an Idea and High-Volume Production*, AUDI AG, August 8, 2022, available at https://www.audi-mediacenter.com/en/press-releases/audi-production-lab-the-link-between-an-idea-and-high-volume-production-14782.

31 The project team planned to integrate up to eight cameras in one press line.

Table 3: Developing Crack Detection AI for Scalability at Audi*

Case-Specific AI Scaling Challenges	AI Scaling Actions	L	P	D
Leveraging the maximum AI development synergies for the organization	Internal machine learning engineers decided to build a single AI model for multiple deployments to leverage data network effects and avoid a fragmented AI model landscape.	X		X
	A data labeling platform facilitated efficient workforce engagement in both data collection and labeling and enabled consistent resolution of discrepancies.	X		
	The design of the crack detection AI solution focused on reusability and ensured a clear distinction between general AI functions and the specifics of the solution.	X	X	X
	The crack detection AI solution provided a blueprint for future AI scaling endeavors, in particular by providing best practices for AI-based computer vision.	X	X	X
Evaluating the value of the scaled AI solution by building on local AI performance	The project team co-defined “error slip” and “pseudo defects” as the key AI performance metrics from which press shop experts can derive the business value from scaling the AI solution.		X	
	The project team made a detailed profitability assessment for the crack detection AI solution when implemented at scale, aiming to justify the required upfront investment.			X

* AI scaling challenges: (L) Learning requirements, (P) Probabilistic reasoning, (D) Data processing requirements.

network effects.[32] Critical factors for enhancing the scalability of the AI model were AI task generalization and prediction accuracy. These factors were crucially important for including data from various sources during early-stage AI model training.

In addition to collecting training data from the Ingolstadt press shop, the project team collected images from multiple sources at different plants to create an extensive database.. Access to the multiple data sources was facilitated by project team members tapping into their existing networks of subject matter experts from previous projects.[33] Because the incidence of cracked parts is very small in a highly optimized manufacturing environment, this approach was beneficial in creating a balanced and sufficient AI training dataset. As explained by the crack detection project leader, factory workers were committed to and supported the search for defective parts: “After explaining [to factory workers] what AI does and why the data collection is so important, they were willing to support our project. They even fished defective, cracked parts out of the trash to take pictures of cracks.”

32 Data network effects describe a virtuous cycle facilitated by AI: More users generate more data, which leads to a better AI solution and therefore more value for the user. See, for example, Gregory, R. W., Henfridsson, O., Kaganer, E. and Kyriakou, H. “The Role of Artificial Intelligence and Data Network Effects for Creating User Value,” *Academy of Management Review* (46:3), July 2021, pp. 534-551.

33 It is important to achieve buy-in from end users, in this case, shop floor workers. If end users perceive little or no value for themselves in the new AI technologies but rather an additional effort and a loss of autonomy, resistance to AI adoption will be considerably higher. For more information, see Kellogg, K. C., Sendak, M. and Balu, S. “AI on the Front Lines,” *MIT Sloan Management Review*, May 2022, available at https://sloanreview.mit.edu/article/ai-on-the-front-lines/.

The project team dedicated significant efforts to the acquisition of image data.[34] It initiated a tedious and labor-intensive data collection and labeling process, aiming to ensure the utmost quality in the training data. This process involved fine-tuning the distinction between genuine cracks and non-cracked parts in the training dataset.[35] To streamline these efforts and involve experts who were distributed over different sites, the project team developed a data labeling manual to provide guidance on annotating and segmenting crack images. Multiple experts reviewed the same data, with any disagreements resolved through organized discussions of borderline cases.

In parallel, the team established a centralized labeling platform in preparation for scaling the AI solution. This platform provides a well-defined workflow for continuous data labeling and issue management. It also facilitates seamless communication between press shop experts and the AI team, streamlining the process of resolving emerging labeling issues and technical queries. Any labeling discrepancies or differences in expert opinions are systematically documented and resolved in a structured manner.

The project team designed the crack detection AI solution with a strong emphasis on reusability and maintained a clear distinction between general AI functions and the specific requirements of the solution. It adopted a modular approach to the software functions, avoiding stand-alone developments and ensuring that as many requirements as possible were designed for platform integration. For example, software modules, including a general user interface and camera connectors, were designed to be general and not specific to the crack detection AI solution.

In collaboration with press shop experts, the project team defined two key performance metrics aimed at quantifying the prospective business value of scaling the AI solution across various locations. First, the “pseudo defect rate” (i.e., false positives) measures the rate at which the AI model incorrectly identifies parts as defective. These cases generate unnecessary waste on the shop floor. Second, the “error slip rate” (i.e., false negatives) signifies the rate at which the AI system fails to detect cracks even though the parts are defective, resulting in costs and delays later in the production process.

Based on these two metrics, the project team conducted an exhaustive performance assessment for the scaled implementation of the crack detection AI solution, with the objective of justifying the required upfront investments for scaling. The evaluation approach was grounded in the analysis of localized data and enabled a more precise estimate of local value potentials and the identification of additional learning requirements. In close collaboration with press shop stakeholders, machine learning engineers tested the trained AI model on local data, extracting performance metrics and evaluating the potential local business value. Prior to the acceptance of the scaled business case, they rigorously assessed the scalability of different solution approaches in coordination with the press shops. This iterative process involved ongoing refinements of the business case calculations for the scaled operation.

In summary, the activities in the *developing* stage of Audi’s AI scaling innovation project can be summarized under the two headings shown in Table 3. The table also shows which of the three AI scaling challenges each action addresses.

Stage 3: Implementing AI for Scalability

The *implementing* stage started in June 2019 and ran until the first AI-based crack detection system became operational at the Ingolstadt press shop in September 2020. During this stage, the project team seamlessly integrated the Ingolstadt press shop into Volkswagen’s Digital Production Platform (DPP), which provides essential IT infrastructure facilitating the use case alignment with the Volkswagen industrial cloud ecosystem.[36] As an IT requirements specialist noted: “Of course, the AI part was one focus of

34 The project team used images from existing image analysis systems (see Appendix B) and additional images, such as cell phone images of defective parts. In September 2020, the labeled dataset included more than 14,000 images.

35 According to the domain experts we interviewed, recognizing cracks in images is not an easy task. For example, slight (noncritical) indentations or oil residues can appear to be cracks.

36 Subsequently, Audi also connected its Neckarsulm site to the DPP. The DPP brings together the data of all machines, plants and systems for all the plants across the Volkswagen group. See *Audi Launches Initiative for Digital Factory Transformation in Heilbronn*, Audi Media Center, April 30, 2021, available at https://www.audi-mediacenter.com/en/press-releases/audi-launches-initiative-for-digital-factory-transformation-in-heilbronn-13945.

Table 4: Implementing Crack Detection AI for Scalability at Audi*

Case-Specific AI Scaling Challenges	AI Scaling Actions	L	P	D
Facilitating AI scaling within legacy IT systems and infrastructure	Audi integrated its press shop located in Ingolstadt into the Volkswagen Digital Production Platform (DPP).			X
	The crack detection AI solution was onboarded to Volkswagen Group's computer vision platform, enabling it to be rolled out throughout the VW Group.			X
Establishing efficient AI operating models and partnerships	Audi's machine learning engineers invested in an automated data pipeline and AI monitoring system that facilitates the deployment of scalable AI solutions with minimal manual effort.	X	X	X
	Audi leveraged collaboration within the VW Group to expand the scope of the AI solution and outsourced the continuous data labeling to external service providers.	X		X

* AI scaling challenges: (L) Learning requirements, (P) Probabilistic reasoning, (D) Data processing requirements

attention, but I would estimate that the actual AI algorithm accounts for 20% of the effort. The main effort was related to infrastructure, operating model, role definitions, etc."

Audi chose a cloud-based approach to data processing and AI model training because it provided the flexibility to adjust computing and data storage capacities to align with the specific demands of each deployment of the crack detection AI solution. In addition, deploying an AI solution in multiple locations requires significant diligence in monitoring data drift and AI model performance. As explained by a machine learning engineer:

> *"Running an AI model in manufacturing requires algorithm monitoring and updating, continuous integration and development pipelines or data versioning. If you want to use an AI model at many different locations, you must find solutions with non-proportional maintenance and monitoring overheads as it scales."*

To ensure that additional deployments of the crack detection AI solution would not increase staffing requirements, the project team instituted automated pipelines for managing machine learning workloads efficiently.[37] These pipelines provide a comprehensive AI monitoring system that not only automates model deployment but also allows Audi to revert to previous model states, thus enhancing transparency and providing reproducibility. An "uncertainty" pipeline continuously monitors the AI model's confidence in detecting cracks, while a "drift" pipeline detects changes in input image data, such as when new parts are introduced to the press line. Because the AI operations logic is automated, machine learning engineers no longer have to manually detect data drift and supervise AI model performance.

The experience gained from exploring the scalability of the crack detection AI solution has prepared Audi for future AI scaling initiatives in two ways. First, implementing the solution has fostered collaboration within the VW Group, optimizing cost efficiencies by capitalizing on platform synergies, such as software function reuse. Second, it has fostered collaboration in developing the business cases for additional deployments of the crack detection AI solution

37 Audi leveraged its AI scaling capabilities through collaboration with external partners. Recent research has investigated AI-based innovation challenges when collaborating with startups. For more information, see Oehmichen, J., Schult, A. and Qi Dong, J. "Successfully Organizing AI Innovation Through Collaboration with Startups," *MIS Quarterly Executive* (22:1), March 2023, pp. 23-38.

Figure 4: Camera Installed Underneath a Door Side Panel in the Press Line[38]

across Audi's production sites, thus enhancing the inspection capabilities of the solution. This approach also augments the value derived from AI-based inspections.

Finally, during the implementing stage, Audi streamlined operations by outsourcing data labeling to external service providers. This step not only reduced costs associated with manual data labeling but also ensured the effective management of continuous data labeling processes, supporting the organization's broader AI scaling initiatives.

In summary, the activities in the implementing stage of Audi's AI scaling innovation project can be summarized under the two headings in Table 4. The table also shows which of the three AI scaling challenges each action addresses.

Stage 4: Scaling AI

Audi has been using the crack detection AI solution since September 2020. Since the initial implementation in the Ingolstadt press line, Audi has scaled the solution by deploying it multiple times at a number of manufacturing sites. This scaling has been achieved without dramatically increasing operational expenses, thus decoupling value and cost.

In the Ingolstadt factory's lead press shop, the crack detection AI solution is used to inspect various deep-drawn parts, including interior door parts, tailgates, side frames and transmission tunnels. Up to eight cameras are mounted in each press line with magnetic mounts (see Figure 4), each using the same AI model. This setup enables swift adjustments based on manufacturing requirements, ensuring high accuracy without the need for manual setup and calibration.

Audi has scaled the crack detection AI solution by deploying it in the Neckarsulm press shop and the Volkswagen press shop in Wolfsburg. In the latter location, a distinct configuration uses the AI model to inspect sheet metal parts through

Table 5: Scaling Crack Detection AI at Audi*

Case-Specific AI Scaling Challenges	AI Scaling Actions	L	P	D
Securing and leveraging the value-cost decoupling effectiveness for AI scaling	Following a larger expansion plan, Audi scaled the AI solution to the Neckarsulm and Wolfsburg plants.	X	X	X
	Local AI model performance was safeguarded for each deployment of the solution before releasing an updated model.	X	X	
	Audi performed post-implementation reviews to quantify the savings through the automation of functional services.	X	X	

* AI scaling challenges: (L) Learning requirements, (P) Probabilistic reasoning, (D) Data processing requirements

a camera tunnel. The scaling is part of a long-term expansion strategy that includes future implementations at additional manufacturing sites, including the broader VW manufacturing network.

AI-based crack detection is based on a learning system designed to continuously ingest new data for retraining the AI model for optimal performance. Audi therefore must ensure that the AI model performance is safeguarded for each deployment throughout its lifecycle. Thus, before each new deployment, the model is tested on a larger subset of data from the particular press line and then run in shadow mode. As a machine learning engineer explained: "While the AI model may perform well on a distributed dataset on average, it can still lead to business value destruction if it fails to perform in a particular location."

A further crucial consideration for Audi was ensuring that the expected value from its AI scaling efforts was realized. It therefore conducted regular post-implementation reviews to measure the savings generated through the automation of functional services. These reviews revealed substantial cost reductions in operational expenses, stemming mainly from the automation of critical functionalities required for AI-based operations. Key factors contributing to these savings included automated monitoring, a streamlined deployment pipeline, efficient metadata management and an effective labeling platform.[38]

The activities in the scaling stage of Audi's AI innovation project are summarized in Table 5. The table also shows which of the three AI scaling challenges each action addresses.

Summary of Realizing AI Value Through Scaling at Audi

Audi's AI scaling journey showcases the value potential of scaling AI successfully. The focus on innovating AI systems for scalability is reflected in three different ways in the Audi case. First, Audi leveraged data network effects to improve the performance of an individual system. Second, it leveraged the cumulative value of the AI model by deploying it in multiple press lines at multiple sites. Third, it leveraged synergies across the deployments of the crack detection AI solution—for example, by automating common functions and distributing operating costs.

The tangible value derived from Audi's scaling efforts is evident in the increased output of its crack detection AI solution. This was achieved by augmenting the volume of quality inspections performed in a press line and by replicating the solution across different contexts, such as a second press line in the Ingolstadt plant and multiple replications in the Neckarsulm plant.

38 Source: *Audi Optimizes Quality Inspections in the Press Shop with Artificial Intelligence*, Audi MediaCenter, October 15, 2018, available at https://www.audi-mediacenter.com/en/press-releases/audi-optimizes-quality-inspections-in-the-press-shop-with-artificial-intelligence-10847.

Table 6: Summary of Recommendations and Actions for AI Scaling

Recommendation 1: Make AI Scaling a Strategic Priority
Action 1.1. Prioritize AI scaling use cases that support the business strategy. **Action 1.2.** Ensure scalability in AI use case development. **Action 1.3.** Focus investments on long-term, scaled AI value creation.
Recommendation 2: Foster Collaboration and Shared Ownership When Scaling AI
Action 2.1. Establish interdisciplinary product teams with strong ownership to find scalable AI value potentials. **Action 2.2.** Involve AI experts in evaluating and determining AI scaling pathways. **Action 2.3.** Include the scaled user base in defining acceptance criteria before funding AI scaling.
Recommendation 3: Streamline Operations to Scale AI Efficiently
Action 3.1. Prioritize the automation and orchestration of data and AI pipelines to increase operational efficiency. **Action 3.2.** Invest in building AI safeguarding capabilities to mitigate operational risks. **Action 3.3.** Establish robust governance frameworks to oversee AI initiatives.

Key factors that facilitated the decoupling of value and cost include a centralized approach to data collection. In addition, AI model training was further enhanced by establishing a central data labeling platform, which streamlined data labeling activities across press shop experts from different shifts and locations, ensuring high data quality and resolving borderline cases.

On the cost side, Audi's AI scaling focused on achieving synergies across the multiple deployments of the AI solution, ensuring that costs remained stable or were even reduced as outputs increased. These synergies and cost benefits were realized through integrating diverse images from different parts, locations and presses to increase the AI model's ability to correctly detect cracks. This integration ensured low costs during the ramp-up phase of an additional deployment due to the AI model's improved performance over time.

Centralizing the management of platform connectivity, streamlining the usage of resources for exploring and developing additional deployments of the AI solution, and modularizing the solution's functionalities contributed to decoupling value and cost. Moreover, switching to a platform-oriented development approach and connecting to the Volkswagen Group's computer vision platform further amplified the potential of AI scaling.

Recommendations and Actions for AI Scaling

The Audi case demonstrates the central role that AI scaling plays in driving value because scaling decouples value from cost. The case also highlights the critical strategic organizational capabilities needed to successfully scale AI, including the importance of purposefully developing AI use cases from the outset. This imperative is rooted in the three unique challenges of scaling AI solutions—learning requirements, probabilistic reasoning and data processing requirements.

Drawing from our analysis of the Audi case, we provide three recommendations, each with three recommended actions, for CIOs and senior business executives seeking to harness AI value at scale (see Table 6). By following these recommendations, leaders can effectively leverage AI scaling to advance their business objectives.

Recommendation 1: Make AI Scaling a Strategic Priority

Scaling is a key driver for realizing AI value. Executives and CIOs should therefore ensure that the strategy for selecting AI use cases focuses on scalable business problems with the potential to deliver substantial business value. The early consideration of technical needs and prospective application contexts during scaling enables value-oriented AI development right from start. Executives and CIOs should take three actions to make AI scaling a strategic priority.

Action 1.1: Prioritize AI scaling use cases that support the business strategy. Though many manufacturing organizations have well-established idea-generation processes or are learning about AI applications through industry peers, a key challenge that executives often face is selecting and prioritizing the "right" AI use cases. Implementing localized AI solutions can provide benefits, but the greater value potential lies in industrial AI applications that scale.

However, given the significant effort needed to develop the business case for a scalable AI solution, executives and CIOs should focus on areas where AI can offer substantial business value and prioritize AI use cases accordingly. They should also communicate how a scaled AI solution use case is connected to the business strategy and establish key performance indicators to continually assess the impact of AI scaling. This approach provides transparency into the expected and realized benefits and helps executives communicate the broader implications of AI scaling to the organization's stakeholders.

At Audi, decision makers closely aligned business value and technical success metrics, showcasing the importance of designing a scalable AI solution. Audi also incorporated AI-based quality inspection into its manufacturing growth and quality assurance strategy to reduce factory costs and is now scaling its crack detection AI solution to additional Audi factories.

Action 1.2: Ensure scalability in AI use case development. Executives and CIOs should ensure that AI scalability is based on clearly defined objectives and embedded as a fundamental concept at each innovation stage. During the exploring stage, these objectives guide the selection and design of scalable AI use cases. During the developing stage, technical decisions are made to ensure that the use case is transferable and replicable. During the implementing stage, the AI operating model and IT architecture are designed to leverage the value-cost dynamics at the scaling stage.

Audi embedded scalability into its crack detection AI solution by creating a common AI model that can be deployed across various contexts. Platform-centric feature development in the implementing stage facilitated seamless integration with the VW Group platform, which set the stage for scaling the solution across the group.

Action 1.3: Focus Investments on Long-Term, Scaled AI Value Creation. In many cases, the investment required to establish AI scalability will rarely pay off in the short term. Given the commitment of human, technical and financial resources, executives should make investment decisions based on longer-term business case calculations. Audi prioritized scalability during the developing stage and based the justification for related investments on the scaled business case.

Focusing on short-term returns of AI investments can be a major obstacle to prioritizing and funding AI solutions with the potential for scaling in the longer term. To improve their decision-making with respect the AI scalability, executives and CIOs should therefore ask: Can a non-scalable AI use case ensure immediate tangible results to justify the required investment? Should we make investments that will enable the solution's future scalability? Is there sufficient business justification for funding the investments required for developing a scalable solution? Answering these questions, however, is inherently challenging because of the difficulty of assessing the benefits of scalability at an early stage. Nevertheless, the Audi case shows that finding answers to these questions can be crucial for informed AI portfolio decision-making.

Recommendation 2: Foster Collaboration and Shared Ownership When Scaling AI

Scaling an AI solution requires the joint exploration of AI opportunities by domain and AI experts and its development and implementation across multiple contexts. Executives should foster collaboration and shared ownership among the scaled user base to ensure that the emerging challenges are addressed effectively and that opportunities for AI-based value generation are realized throughout the AI scaling journey by taking three actions.

Action 2.1: Establish interdisciplinary product teams with strong ownership to find scalable AI value potentials. Executives should encourage employees to embrace AI scalability and promote the full exploration and understanding of relevant business problems, ensuring that investments in AI scalability are justified by significant business opportunities. A business unit's sustained commitment, ownership and value focus is critical to ensuring that an AI project remains dedicated to solving scalable business problems.

At Audi, the active involvement of press shop engineers and business experts played a critical role in validating the AI scaling assumptions and aligning them with the organization's core objectives, challenges and growth opportunities, as reflected in the following statement from the crack detection project leader:

> *"Solving a problem with AI may sound easy, but many people don't realize the challenges behind it. ... The effort [to scale AI] is significant, and we only endured it because we knew the [impact] of scaling would be huge. ... It was important to involve the department experts from the beginning to create a shared sense of responsibility. Only then you will see how much pain they are feeling today and whether they are willing to take the effort."*

Action 2.2: Involve AI experts in evaluating and determining AI scaling pathways. The collaborative approach should focus on validating real, scalable business opportunities and assessing the suitability of AI solutions to effectively address them. Executives should foster open communication and provide incentives for cross-functional collaboration to harness the collective intelligence of their workforce, overcome AI implementation barriers and drive widespread AI adoption. The Audi case illustrates the success of such collaborative efforts, where the extensive experience and knowledge of a core project team was instrumental in shaping a comprehensive AI scaling perspective over several years of collaboration, particularly between AI experts, business stakeholders and IT specialists.

Action 2.3: Include the scaled user base in defining acceptance criteria before funding AI scaling. Executives in organizations embarking on AI scaling efforts should ensure that the scaled user base is included in defining the acceptance criteria for the scaled solution before allocating resources and funding for scalability. By emphasizing the centrality of users and the business problem from the outset, organizations can ensure that AI solutions resonate with end users, drive adoption of the solutions and maximize value realization. Incorporating the perspectives and requirements of the scaled user base in the acceptance criteria not only increases the relevance and usability of AI solutions but also promotes business and process alignment across contexts. This approach facilitates collaboration, drives business process alignment and ultimately prepares organizations to successfully scale AI.

Recommendation 3: Streamline Operations to Scale AI Efficiently

The efficient management of AI operations is critical to decoupling value and cost in scaled AI applications. Efficient management ensures that operational costs do not increase proportionally as AI applications scale in volume and across different environments. Robust and stable AI operations serve as a backbone throughout the lifecycle of AI applications, ensuring seamless model deployment and effective monitoring of probabilistic outcomes to ensure optimal AI performance. Given the context sensitivity of AI solutions and potential variations in performance across different environments, implementing robust and cost-efficient AI operations is fundamental to scaling AI applications. Executives should take the following three actions.

Action 3.1: Prioritize the automation and orchestration of data and AI pipelines to increase operational efficiency. Workflows can be optimized by automating repetitive tasks and procedures. Executives should therefore prioritize the deployment of automation tools tailored for key AI tasks such as data preprocessing, model deployment and performance monitoring, thereby reducing manual efforts and errors. Audi's IT function deployed automated tools for monitoring key performance indicators such as model accuracy, latency and resource usage, and the deployment of Audi's AI models is supported by automated pipelines. By migrating the crack detection AI solution to the Volkswagen DPP, Audi has further

Table 7: Key Characteristics of Conventional and AI-Based Image Analysis

Conventional Image Analysis	AI-Based Image Analysis
Key characteristics • Camera experts configure and adjust each system individually based on pre-engineered features. • Systems require lengthy (re)calibration when contextual conditions change. • Local hardware and image processing needed for each system, i.e., a complete system. • Local system optimizations for each system without scaling economies.	**Key characteristics** • Domain experts label a comprehensive dataset and AI experts train the AI algorithm. • Systems are robust to changing contextual conditions and do not require (re)calibration. • Centralized inference logic with low-cost cameras/ components needed for each system. • AI model optimizations can be deployed across multiple deployments, enabling compound effects.
→ Low scaling potential	**→ High scaling potential**

streamlined access to preexisting AI algorithms, software tools and frameworks.

Action 3.2: Invest in building AI safeguarding capabilities to mitigate operational risks. This action involves developing internal AI competencies, ensuring employees are comprehensively qualified and leveraging advanced technologies like explainable AI and AI auditing tools. By nurturing internal expertise to build AI safeguarding capabilities, organizations can deploy AI technologies confidently and can proactively mitigate emerging risks such as algorithmic bias that can lead to poor AI performance. To continuously advance data management capabilities that enable AI scaling, IT executives should foster company-wide data integration.

Audi internally validates the crack detection AI models in a shadow environment before deploying them in a live press line, adding extra training data if necessary to mitigate the risk of biased decisions. Borderline cases are continuously reviewed in consultation with press shop experts to further minimize false positives and negatives and improve the performance of the AI solution across multiple real-world environments.

Action 3.3: Establish robust governance frameworks to oversee AI initiatives. Executives should ensure that the governance frameworks provide clear guidance through policies, methods and standards for AI development and deployment. By defining roles, responsibilities and decision-making protocols, executives and CIOs can mitigate risks and avoid legal ramifications associated with AI deployments. Comprehensive data governance frameworks, together with stringent quality controls and meticulous lifecycle management processes, ensure the accuracy, reliability and accessibility of data for AI applications. In the Audi case, a labeling guideline was defined to specify how data used for training the crack detection AI solution is named and classified. In addition, Audi standardized technology components such as camera connectors to enable easy integration and connection to the VW platform ecosystem.

Concluding Comments

Harnessing the full value potential of AI requires an entire organization to embrace a transformative journey and prioritize AI scaling as a strategic imperative. Given the distinct nature of AI compared to traditional information technologies, organizations need to cultivate a nuanced understanding of AI-based innovations and prepare for the unique demands of scaling AI solutions. Embracing AI scaling empowers organizations to navigate the complexities of decoupling value and cost at the strategic, technical and operational levels.

In this article, we have described how Audi designed a scalable AI-based quality inspection system in manufacturing, thus ensuring long-term AI value realization at scale. Audi's AI scaling journey comprised three innovation stages followed by an ongoing scaling stage, during which the scaled AI solution provided increasing cost benefits. Based on our analysis of the Audi case, we offer three recommendations, coupled with three actions each, that can help executives:

- Address and overcome the challenges of efficiently scaling AI applications and realizing the transformative potential of AI technologies
- Cultivate an organizational environment conducive to AI scaling by fostering innovation and continuous learning so that AI scaling permeates the entire organization
- Drive long-term success with AI by strategically positioning their organizations for sustained success and unlocking the full value potential of their AI investments.

In summary, this article emphasizes the paramount importance of embracing a strategic and thorough approach to AI-based innovation, with Audi's AI scaling journey providing a comprehensive illustration of this imperative. By recognizing AI scaling as a crucial value driver and proactively tackling the associated challenges, CIOs and senior executives can effectively harness the long-term value potential of AI, laying the groundwork for operational excellence and sustained success in the age of AI.

Appendix A: Research Methodology

Our research aimed to explore how organizations can realize value through scaling AI solutions. We conducted a single in-depth case study of Audi's journey to scale its AI solution for detecting cracks in sheet metal parts. We studied this case for three reasons:

1. To gain the maximum benefits from its crack detection AI solution, Audi needed to scale it across multiple press lines at multiple sites. This pioneering deployment of a scaled AI solution allowed us to surface AI-induced scaling challenges and investigate how Audi overcame them.
2. From the outset, Audi intended to scale the solution, allowing us to observe and describe AI scaling-specific activities during all innovation stages.
3. We had broad access to the organization, allowing us to interact strongly with those involved in the AI scaling journey and obtain unique insights from the field.

We collected primary data from March 2020 to December 2021 through interviews with the various stakeholders of the AI solution. Throughout our research, two authors were deeply involved in developing, deploying and scaling the crack detection AI solution, providing us with a unique source of insight.[39] We followed interpretative research principles and triangulated our data with additional data sources, including interview and meeting notes.[40] We also accessed extensive secondary data, including internal documents and public statements.

We followed a stepwise coding approach consisting of open, axial and selective coding to identify AI scaling challenges and how Audi managed to overcome them.[41] During open coding, codes emerged through case descriptions and summaries, which we used to summarize our data and get an initial overview of all the case data.[42] Due to the topic's novelty, we initially developed codes inductively. In the axial coding phase, we condensed the data and categorized it according to the four stages of an AI innovation project: *exploring*, *developing*, *implementing* and *scaling*. We summarized the emerging codes to identify recurring themes, resulting in the identification of four AI scaling-specific challenges. Finally, we iterated through our case data and mapped Audi's case evidence to our coding scheme.

During the coding process, we obtained detailed findings from the primary and secondary data analysis by continually comparing and triangulating them with internal information (e.g., project presentations) and public case material (e.g., press releases). In doing this, we avoided any potential bias resulting from the two authors' close involvement in the crack detection

39 Patton, M. Q. *Qualitative Evaluation and Research Methods*, SAGE Publications, 1990.

40 Klein, H. K. and Myers, M. D. "A Set of Principles for Conducting and Evaluating Interpretive Field Studies in Information Systems," *MIS Quarterly* (23:1), 1999, pp. 67-93.

41 See: 1) Strauss, A. and Corbin, J. M. Grounded Theory in Practice, SAGE Publications, 1997; and 2) Corbin, J. and Strauss, A. "Grounded Theory Research: Procedures, Canons, and Evaluative Criteria," *Qualitative Sociology* (13:1), 1990.

42 Yin, R. K. *Case Study Research Design and Methods—Third Edition*, Applied Social Research Methods Series (5), SAGE Publications, 2003.

AI solution. Moreover, we relied on iterative discussions among all researchers after each round of coding to validate our findings.

Appendix B: Comparison of Conventional and AI-Based Crack Detection Systems

Audi's Ingolstadt press shop in Germany manufactures deep-drawn sheet metal parts that are later assembled in the body shop. The deep-drawing process involves up to six steps, with each step taking up to four seconds. During these forming steps, occasional cracks can develop at critical points on the sheet metal parts. These cracks come in various sizes, widths and directions, making their detection a complex task. Previously, workers performed sporadic manual inspections of parts due to the short cycle times in the forming process. However, manual inspections were only to uncover a minority of defects, causing inefficiencies when missed defects were identified downstream. In its quest for a suitable automated solution for the crack detection task, Audi compared conventional and AI-based image analysis approaches (see Table 7).

Conventional image analysis systems operate by relying on pre-engineered, manually configured features. Experts set up image processing software to inspect images for consistent, fixed or repetitive patterns. As a consequence, these systems require recalibration whenever significant changes occur in the manufacturing process that impact inspection tasks, such as alterations in the color gradients of the inspected parts. Recalibrating the system entails additional integration costs, either by engaging external service providers or by placing extra demands on internal personnel. Additionally, conventional image analysis systems run on localized hardware and software, requiring manufacturing companies to procure and individually configure a complete inspection system for each application.

The training data for an AI-based image analysis system comprises exemplary images. Such an AI system is trained to perform a human-like task independent of discrete features; it doesn't rely on specific features but interprets images based on its learning from the training data. In essence, the AI algorithm learns to understand the "concept" of a crack, just like a press shop expert, and the algorithm can be applied to similar crack detection tasks. Moreover, when changes or deviations occur in the manufacturing process, the need for reconfigurations and recalibrations is largely obsolete. Another advantage of AI-based image analysis systems is that they centralize the inference logic, leading to synergies across multiple deployments of the systems and reduced time for system integrations or adjustments.

About the Authors

André Sagodi

André Sagodi (andre.sagodi@audi.de) is a manufacturing and innovation expert at AUDI AG, Germany. He has extensive experience in production planning and innovation management. At the Audi Production Lab (P-Lab), he explores and tests state-of-the-art technologies to increase efficiency, precision, and quality in Audi's global production network. André is also a doctoral researcher in information management and a member of the Management of Artificial Intelligence Research Lab at the University of St. Gallen, Switzerland. His research focuses on AI-based digital innovation in manufacturing.

Benjamin van Giffen

Benjamin van Giffen (benjamin.vangiffen@unisg.ch) is an assistant professor of information management and head of the Artificial Intelligence Research Lab at the University of St. Gallen, Switzerland. He is an associate researcher with the European Research Center for Information Systems (ERCIS) and the University of Liechtenstein. His research focuses on AI adoption, the responsible and value-oriented design and management of AI, digital platforms and ecosystems and on how to leverage design thinking for artificial intelligence. He collaborates with leading firms in Europe and the U.S., advising on AI strategies for customer and business value.

Johannes Schniertshauer

Johannes Schniertshauer (johannes.schniertshauer@audi.de) is an accomplished machine learning engineer at AUDI AG in Germany, where he leads the development of numerous AI solutions deployed across the

company's factories. He also serves as the IT lead for the Industrial Computer Vision Team at the Production Data-Factory, a specialized unit focusing on leveraging data to enhance manufacturing processes. Johannes contributes to research initiatives in MLOps[43] and AI safeguarding, focusing on enhancing the reliability and safety of AI applications. His research is dedicated to pioneering advanced methods that ensure operational excellence.

Klemens Niehues

Klemens Niehues (klemens.niehues@audi.de) is a production engineer at AUDI AG, Germany, focusing on technology and automation. As product owner and project leader, he is responsible for different computer vision projects in the production area, focusing especially on AI. Klemens received his Ph.D. from the Institute for Machine Tools and Industrial Management (iwb) at the Technical University of Munich, Germany.

Jan vom Brocke

Jan vom Brocke (jan.vom.brocke@uni-muenster.de) is the director of the European Research Center for Information Systems (ERCIS). He is a professor and holds the chair of Information Systems & Business Process Management at the University of Münster in Germany. He is a visiting professor at the University of Liechtenstein and has been named a Fellow of the Association for Information Systems (AIS), a Fellow of the ESCP Center for Design Science in Entrepreneurship, a Schoeller Senior Fellow at Friedrich Alexander University (FAU) in Germany and a Distinguished Professor at the National University of Ireland, Maynooth University.

43 MLOps is a machine learning culture and practice that unifies machine learning (ML) application development and system deployment and operations (Ops).

