This paper was presented at ISPIM Connects Osaka – Connecting and Empowering Society, Osaka, Japan on 2-4 December 2024.
Event Proceedings: ISBN 978-952-65069-7-5

Identification of Use Cases for Generative Artificial Intelligence in Production

Prof. Dr.-Ing. Günther Schuh

RWTH Aachen University, Laboratory for Machine Tools and Production Engineering (WZL), Steinbachstr. 19, 52074 Aachen, Germany.
E-Mail: g.schuh@wzl.rwth-aachen.de

Leonard Cassel, M.Sc. M.Sc.*

Fraunhofer Institute for Production Technology IPT, Steinbachstr. 17, 52074 Aachen, Germany.
E-Mail: leonard.cassel@ipt.fraunhofer.de

* Corresponding author

Abstract: It is imperative that manufacturing companies enhance their capacity for constant change and adaptation. In particular, the ability to adopt new technologies and adapt them to the physical conditions of complex production chains represents a significant challenge. The advent of generative AI, a term encompassing AI-based technologies adept at creating multimodal content, presents a novel opportunity for small and medium-sized manufacturing companies to enhance their competitive edge. Although use cases for discriminative AI are reaching technological maturity in manufacturing applications, the extent to which generative models can be applied in practice remains unclear. This paper presents a method for a systematic identification of generative AI use cases in production. Therefore, a characteristics and capability profile for generative AI is presented and matched with a literature-based investigation of the challenges inherent to production. Lastly, the defined types are analysed regarding their feasibility for implementation in production systems.

Keywords: Artificial Intelligence, Technology Management, Generative Models, Technology Evaluation, Use Case Design, Manufacturing

1 Introduction

In the currently prevailing VUCA environment, characterized by increasing volatility, uncertainty, complexity, and ambiguity, manufacturing companies are confronted with significant challenges to gain a sustainable competitive advantage (Schuh and Dölle, 2021, p. 11). In light of these intricate external factors influencing production enterprises, the capacity to adapt and evolve is identified as a determinant of corporate success (Upton, 1994, p. 72).

The concept of Industry 4.0, which is defined as "the intelligent networking of machines and processes in industry with the help of information and communication technology"

Event Proceedings: ISBN 978-952-65069-7-5

(Dombrowski and Krenkel, 2021, pp. 5-6), offers promising solutions for mastering the challenges posed by the VUCA environment and for achieving improved adaptability (Schuh and Dölle, 2021, p. 11). To enhance interconnectivity between production machinery, a gradual integration of sensor-based technologies, which gather information about companies' manufacturing processes in the form of extensive process datasets, is predominant. In consequence of the growing amount of data, methods, and tools for analysing these datasets are becoming increasingly important. For example, systems incorporating discriminative artificial intelligence (AI), which are highly capable of analysing and interpretating large datasets, are already used in production. These systems can capture complex correlations in real time and can continuously develop and react to changes in the environment. They are already making significant contributions to the adaptability of manufacturing companies (Mockenhaupt and Schlagenhauf, 2024, p. 272; Herf et al., 2024, pp. 33-37).

Generative AI, on the other hand, is characterized by its ability to generate realistic new data in different modalities, e.g. time-series data, images, or text (Kreutzer, 2023, p. 23). It is designed to process user requests efficiently and features an easy accessibility through natural language interaction. Since the release of the first Large Language Model with a chat-based interface by OpenAI, generative AI has recently received significant public attention (Fritz, 2024, p. 344). The increase in its performance can be attributed to three developments in particular: significant improvements in the computing power of high-performance computers, the availability of substantial amounts of training data and the development of the Transformer model architecture (Löser et al., 2023, pp. 6-7).

The global value creation potential with generative AI is estimated at between 2.6 and 4.4 trillion euros per year, which is roughly equivalent to the UK's gross domestic product – Industrial production alone has a value creation potential of 230 billion euros (Siebel, 2024). However, the theoretical potential of generative AI in production has so far been offset by a limited number of use cases. Only 3% of companies have used at least one application of generative AI in industrial production by 2024 (Keiber, 2024). As there has been insufficient research into how the capabilities of generative AI can be used to overcome these challenges, this paper aims to develop and apply a model for the systematic identification of use cases of generative AI in production.

According to KUBICEK, the formulation of specific questions represents the initial stage of an exploratory research process (Kubicek, 1976, p. 15). Based on the initial situation and the objective of the present research project, the following overarching research question is formulated:

Which use cases of generative AI exist in industrial production?

To increase granularity within the research process, sub-research questions are derived:

- Which characteristics and capabilities characterize generative AI?
- What challenges exist in the implementation of Industry 4.0 and the use of data-based technologies in production?
- How can use cases of generative AI in production be systematically identified?

In section 2 of this research paper, the relevant background of this paper, including fundamentals of production management, Industry 4.0 and Artificial Intelligence are introduced. In section 3 the approach used by the authors for identifying use cases is

elaborated. Finally, the overarching and general research questions are elaborated on in section 4 with a conclusion and further needs for research in section 5.

2 Theoretical Background

Production Management

In Production Management, a distinction between normative, strategic, and operative instances is found in literature (Schuh and Schmidt, 2014, pp. 5-7). Normative production management deals with the reflection and legitimization of entrepreneurial activities in the light of social value orientation and moral intrinsic values. Operational production management, on the other hand, is concerned with controlling the production system and associated production processes. In strategic production management, the focus lies on building sustainable competitive advantages and thus securing the future of a company (Schuh and Schmidt, 2014, pp. 5-7; Rüegg-Stürm, 2005, p. 114). In the context of production, there are various target variables that influence the company's success in their interaction. Historically, however, all manufacturing companies find themselves in a field of tension, represented by the "magic triangle of targets" consisting of the target dimensions of *time*, *costs,* and *quality* (Atkinson, 1999, p. 338). Due to the complex possibilities of environmental changes and external factors with an influence on production, adaptability is a key criterion for a successful long-term pursuit of these targets (Westkämper and Löffler, 2016, p. 50; Dombrowski and Krenkel, 2021, pp. 5-6). The required adaptability of production can be achieved through the characteristics of flexibility and versatility (Westkämper and Löffler, 2016, pp. 64-65). *Flexibility* is the speed and the ability to adapt a production to suitably changed conditions. For the extended understanding of the targets of production, the dimension of *flexibility* must be added to the dimensions of *time*, *costs,* and *quality* (Abele et al., 2024, p. 23; Westkämper and Löffler, 2016, pp. 64-65).

Industry 4.0

The concept of Industry 4.0 and the associated information and communication technologies offer significant potential for resolving the previously stated conflict of objectives and further increase the efficient fulfilment of the target variables of production (Schuh et al., 2020b, p. 175; Schuh and Dölle, 2021, p.11; Schuh et al., 2020c). At the core of Industry 4.0 lies the interconnection of information and communication technologies introduced in the late 20th century. Interconnection and digitalization should enable an exchange of information along the entire value chain. In industrial production, there are a multitude of applications for connecting resources, information, objects, and employees with the objective of increasing productivity (Reinhart, 2017, pp. XXIV-XXXV; (Bauernhansl, 2020, p. 17). This target state can be achieved by means of end-to-end networking of production facilities based on computers, internet connections, real-time sensor measurements, digital assistance systems and automation (Hippmann, Klingner and Leis, 2018, S. 13).

This paper was presented at ISPIM Connects Osaka – Connecting and Empowering Society, Osaka, Japan on 2-4 December 2024.
Event Proceedings: ISBN 978-952-65069-7-5

Artificial Intelligence

The substantial quantity of supplementary production and product data generated by networked production as part of the Industry 4.0 concept also necessitates novel resources for analysis and interpretation. Traditional and manually performed data analyses are often inefficient and time-consuming, as patterns in substantial amounts of data are difficult to access or too complex for human operators. Systems based on AI, a comparatively young field of research that integrates findings from a variety of subject areas such as computer science, mathematics and neuroscience (Russell, 2022, pp. 1-6), can assist in these analyses through the automation of analytical processes, which helps to reduce human error and provide deeper insight into process data (Scholz, 2022, p. 58).

A distinction can be made between generative and discriminative models of AI. While discriminative models distinguish between different classes and categories of data, generative models create new, unique data or content that is similar to the data on which they were trained (Luk, 2023, pp. 10-11). Generative models are trained on complex data distributions to produce content that is ideally indistinguishable from human-made content (GM, 2020, p. 2). Using statistical methods, the models encompass high-dimensional probability distributions from a limited training dataset and create new samples that resemble the underlying training data (Tomczak, 2022, pp. 2-3). Therefore, generative models focus on the underlying data structure and processes for generation, while discriminative models focus on the relationship between input features and target variables (Banh and Strobel, 2023, pp. 3-5).

The rapid increase in the performance of generative AI resulted from the recent development of new model architectures. The underlying architectures determine the learning process of the generative models. The most prominent architectures for generative models are Generative Adversarial Networks (GANs) (Goodfellow et al., 2014), Variational Autoencoders (Kingma and Welling, 2014) and Transformer-Architectures (Vaswani et al., 2017), the latest on which the most prominent commercial application for generative AI, the large language model ChatGPT, is based on.

3 Methodology

In the following section, the authors' approach in identifying use cases is presented. For this, a literature-based analysis of the characteristics of generative AI was conducted, i.e. the inherent characteristics that can be derived either from observations of the state of the art of current models or from the underlying architectures. The capabilities of generative AI are subsequently derived based on these characteristics. The term *capability* refers to the potential of a system to fulfil certain functions or meet certain requirements and achieve a specific outcome. In the context of Industry 4.0, BAYHA ET AL. define capability as an implementation-independent description of the function of a resource (Bayha et al., 2020).

The second step (B) includes a comprehensive analysis of the challenges in production. The common meaning among the various definitions of the term *challenge* is the situation in which the attainment of a goal is in jeopardy (Bauernhansl, ten Hompel and Vogel-Heuser, 2014, p. 103). The literature review considers two key areas: the challenges of implementing Industry 4.0 and the challenges of utilising data-driven technologies, with a particular focus on AI and machine learning, in production processes. By examining the challenges of Industry 4.0, the fundamental and structural challenges inherent to

production are considered. A critical examination of the challenges associated with data-based technologies enables the identification of specific technical obstacles that may be surmounted through the utilisation of generative AI.

The procedure for systematic literature analysis is based on a combination of two analysis methods: the STARLITE methodology (Standards for Reporting Literature Searches) and the PRISMA method (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) (Moher et al., 2009). To this end, a targeted search strategy was employed, using a combination of search terms in scientific databases, including Scopus, Google Scholar, IEEE Xplore, ScienceDirect and SpringerLink. These databases were selected based on their comprehensive coverage of technical and engineering subject areas. By comparing these general capabilities with the challenges in production, specific use cases of generative AI in production are then identified. This process concretises the capabilities of generative AI in specific contexts. The approach is visualized in Figure 2.

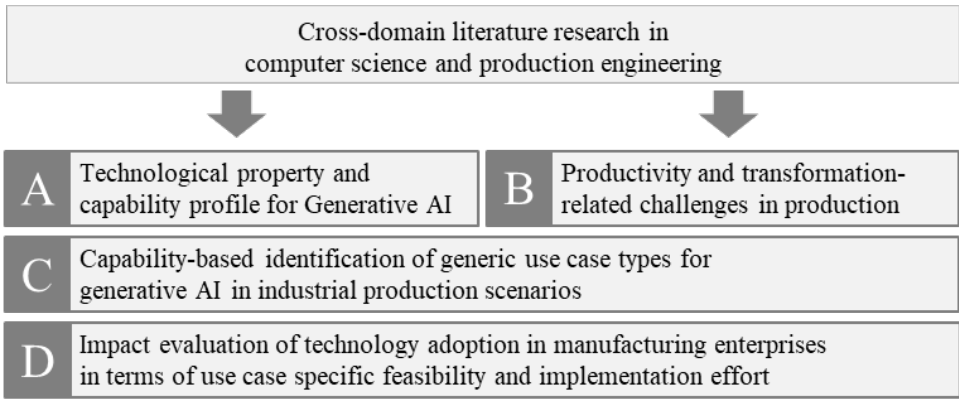

Figure 1: Approach for the Identification of Use Cases for Generative AI in Production

By juxtaposing these general capabilities with the production challenges identified, the derivation of specific use cases of generative AI in production is realized. The term *use case* is not clearly defined in research literature and is subject to a context-dependent understanding. In the context of information technology, a use case is understood as a description of sequences of interactions between a system under consideration and various actors (Cockburn, 2006). Furthermore, *use cases* can be understood as practical or research examples that show the greatest need for action, potential, solution approaches or even challenges in connection with new technologies (Kohl et al., 2019).

This process concretises the capabilities of generative AI in specific contexts and thus assigns them in a practice-oriented context within the domain of production. As shown in Figure 2, the relationship between technological characteristics, resulting capabilities, domain-specific challenges and resulting target criteria of industrial production is thus formed for use case determination.

Lastly, an evaluation of the use cases, with a particular focus on the implications of their implementation, is applied. Based on the general description of the use cases, an assessment of the feasibility and effort involved in their implementation is conducted. The feasibility dimension pertains to the availability of the requisite data basis. The effort dimension assesses the use cases in terms of the availability and necessary adaptations of

Event Proceedings: ISBN 978-952-65069-7-5

existing generative AI models, as well as the feasibility of implementation without the involvement of external partners.

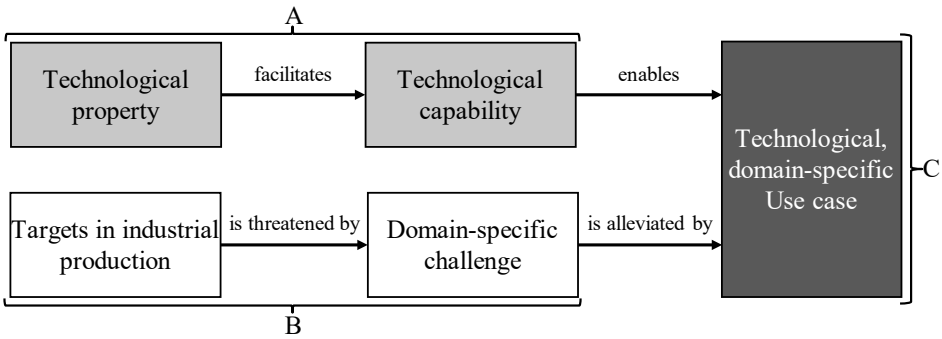

Figure 2: The relational context of the use case building elements

4 Results

A – Technological characteristic and capability profile for Generative AI

According to the definitions laid out in Section 2, the basic characteristics of generative AI and their interdependencies are analysed. As a result, a total of 9 distinct characteristics have been identified:

1. *Scalability* is the characteristic of a generative AI model that enables the efficient processing of large amounts of high-dimensional data (Bommasani et al., 2021, pp. 16, pp. 76-77).
2. The principle of *emergence* describes the phenomenon whereby basic models acquire new capabilities or behaviors without having been explicitly trained to do so (Kintz et al., 2023, p. 15). According to WEI ET AL., emergent capabilities are those capabilities that occur only in models of larger size (Wei et al., 2022, p. 2).
3. *Homogenization* signifies the trend toward consolidation from numerous task-specific models to a limited number of comprehensive general models. The enhanced performance of these general models enables them to undertake a diverse range of activities while being cost-effectively adapted to the respective context with minimal company-, industry- or task-specific data.
4. The characteristic *compositionality* provides a means of integrating previously acquired knowledge and conceptual frameworks, thereby facilitating a more profound comprehension or a more exact reproduction of content. It has already been established as a pivotal element within the foundational structure.
5. *Variability* describes the generation of varying outputs from a singular input (Banh and Strobel, 2023, pp. 4-5). The variability or deviations in the outputs consist mainly in the way the output is expressed. The content, on the other hand, remains consistent (Chan and Colloton, 2024, p. 49).
6. *Real-time behavior* represents the ability of generative AI models to process inputs with almost no time delay and generate outputs in real time. The self-attention mechanism of the transformer enables the parallel consideration and processing of all positions of an input sequence (Vaswani et al., 2017), pp. 6-7).
7. *Contextuality* describes the ability of generative AI to recognize and consider context and thus the meaning of individual input sequences in order to generate

precise and relevant output based on this. Generative AI models are able to recognize relationships that are difficult to identify and mutual influences in data points of an input sequence that are far apart from each other (Kintz et al., 2023, pp. 16-17).

8. *Customizability* describes the ability of basic models to be adapted to specific requirements (Chan and Colloton, 2024, p. 49, p. 51).
9. *Adaptivity* describes the ability of large generative AI models to adapt independently to new tasks, contexts or requirements. In contrast to customizability, these adaptations are made by the model itself and do not require retraining and therefore no adjustment of the parameters (Luk, 2023, p. 21).

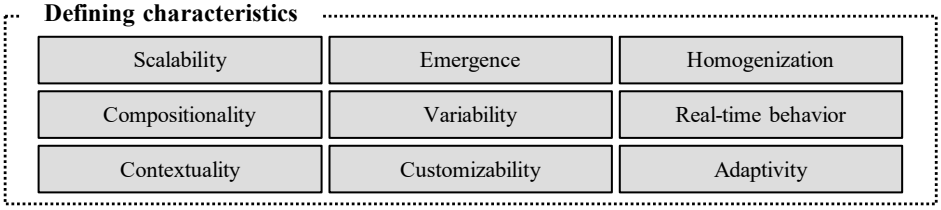

Figure 3: Identified technological characteristics and capabilities of generative AI models

Resulting from these characteristics, six technological capabilities were derived and summarized in Figure 4:

1. *Content generation* is so fundamental to generative AI that it not only represents its central capability, but is also decisive for its naming. It refers to the ability to create new and unique content or data that is indistinguishable from the training data in terms of its form and content. Content generation can be further subdivided into content *completion*, *modification*, *transformation* and *aggregation* (Hofmann, Rückel and Urbach, 2021, p. 5507; Shi et al., 2023, pp. 12-13; Wecke, B. (2024), p.6).
2. *Multimodality* describes the ability of generative AI to receive and understand input requests in different modalities and to generate output in different modalities. For example, multimodal models can process voice input and create corresponding output in text form and with supporting images (Banh and Strobel, 2023, pp. 4-5).
3. *Generalization* describes the capability of generative AI to cope with a variety of different tasks. It is not limited to individual tasks or subject areas, but can be used flexibly in different areas and tasks (Löser et al., 2023, p. 13; Bommasani et al., 2021, p. 76).
4. *Dialogue capability* describes the capability of generative AI to conduct human-like, contextually relevant and coherent conversations. Both, the understanding of user input and the creation of situationally adapted, comprehensible and high-quality output make it possible to have natural conversations with generative AI. Communication between humans and machines in any language makes the interaction user-friendly and intuitive (Leible et al., 2024, p. 13; Obrenovic et al., 2024, p. 8).
5. *Role adaptation* describes the capability of generative AI to adapt flexibly to different perspectives or roles. This ability allows generative AI to put itself in different contexts and act accordingly. For example, depending on the application scenario, it can take on a strategic, creative, advisory or informative role (Leible et al., 2024, p. 14).

This paper was presented at ISPIM Connects Osaka – Connecting and Empowering Society, Osaka, Japan on 2-4 December 2024.
Event Proceedings: ISBN 978-952-65069-7-5

6. *Fault tolerance* describes the capability of generative AI to generate coherent and relevant outputs despite noisy, incorrect incomplete or inconsistent inputs. This is enabled by the training on extremely large datasets, which increases the robustness of generative models (Brown et al., 2020, pp. 29-30).

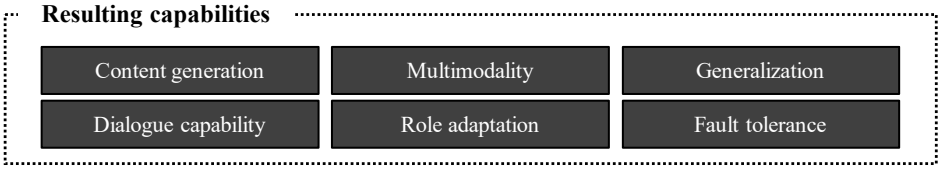

Figure 4: Summary of identified capabilities of Generative AI

B Productivity and transformation-related challenges in production

In the second literature analysis, challenges inherent to the production domain were identified. The results have been structured in four main categories: *Organizational, data-related, structural* and *interaction-based challenges. Organizational challenges* refer to the internal difficulties and obstacles within an organization's structure and processes, aimed at adapting to the dynamic production environment. These include factors like the availability of skilled labor shortage, skills, and the ability to transfer knowledge (Nayyar and Kumar, 2020, p. 200). *Data-related challenges* involve the difficulties in processing and utilizing data. Data is crucial for interconnected and adaptive production in the context of Industry 4.0. These challenges are divided into quality-related and quantity-related aspects. High data quality is essential for reliable results and informed decisions (Dogan and Birant, 2021, p. 16), while large data volumes require significant technical resources and infrastructure (Bajic et al., 2021, p. 550; Bakhtari et al., 2020, p. 2386). *Structural challenges* concern the fundamental characteristics of production, independent of specific technologies. These include industry-specific requirements such as reliability, data confidentiality, and process specialization. Companies must transfer established safety standards to digital technologies and ensure high data confidentiality. Furthermore, manufacturing companies face the challenge of high data and information confidentiality which inhibits the technological development of production. *Interaction-based challenges* relate to problems arising from the integration and interaction of different technologies and their interaction with humans. The integration and compatibility of technologies, components, and systems represent a significant challenge, as the dynamics and rapid pace of technological developments exceed the adaptability of the infrastructure. Linking established and new software and hardware technologies is necessary for intelligent and interconnected production but is often difficult due to inadequate digital infrastructures (Bajic et al., 2021, p. 550).
The identified challenges are summarized in Table 1.

Table 1: Consolidated challenges in industrial production

Organizational challenges	Data challenges	Structural challenges	Challenges of technology interaction
• Shortage of skilled labor • Skills • Knowledge transfer	• Quality related • Quantity related	• Reliability • Confidentiality • Specialization	• Combatability & integration • Human-machine communication

C Capability and challenge-based identification of use cases

The identified use cases for generative AI in production are summarized in Table 2, indicating the addressed challenges and capabilities used. Subsequently, each use case is briefly described in the following section:

Table 2: Identified use cases of generative AI in production

Use case for generative AI in production	Needed technological capabilities	Alleviated challenges in industrial production
Synthetic training data	• Generation	• Quantity related • Reliability
Knowledge and training platform	• Generation • Aggregation • Role adaptation	• Skills • Knowledge transfer
Engineering assistant	• Generation • Modification • Transformation • Multimodality • Dialogue cability	• Professionals • Specialization • Human-machine communication
Computed Aided Manufacturing (CAM) assistant	• Generation • Completion • Modification • Transformation • Multimodality • Dialogue cability	• Professionals • Skills • Reliability • Specialization • Human-machine communication
Machine assistant	• Generation • Aggregation • Dialog cability	• Skills • Reliability • Specialization • Combatability & integration • Human-machine communication
Assembly assistant	• Generation • Modification • Multimodality • Generalization • Dialogue cability	• Knowledge transfer • Reliability • Specialization • Human-machine communication
Technology enabler	• Multimodality • Generalization • Dialogue cability	• Professionals • Combatability & integration • Human-machine communication

To generate **synthetic training data** generative AI mimics real data in structure and variety, which can be used to train discriminative AI models, particularly for visual inspection in industrial quality control, where a large amount of training data is needed. Due to high quality demands in production, there are often few instances of defective parts, leading to unbalanced data and model distortions. Generative AI can supplement these datasets by generating synthetic images, improving model performance and reliability in detecting defects. The manufacturing company BOSCH is already using generative AI for the provision of faulty part images (Westfall, 2024).

Event Proceedings: ISBN 978-952-65069-7-5

The **knowledge and training platform** leverages generative AI to facilitate knowledge transfer and employee training in manufacturing companies. A generative AI enhanced platform is able to store expert knowledge and experiences, such as notes and instructions, making them accessible to inexperienced employees. By personalizing training content, the platform can tailor practical training to match an employee's skills and requirements. Generative AI is able to aggregate internal documents, data, and empirical values to impart knowledge and create individualized training materials. In industry, several vendors offering Learning Management Systems (LMS) have announced plans to integrate AI-based functions into their portfolio (Bersin, 2024).

Assistant use cases, in which employees are assisted by a generative AI based systems in are perceivable daily tasks ranging from manual assembly to engineering:

The **engineering assistant** use case involves using generative AI to support and optimize design processes. Generative AI can autonomously create 3D design proposals, modify existing designs, and assist in transferring designs to CAD models. It also enables faster test iterations by quickly generating and evaluating multiple design variants, speeding up the development process. This is especially efficient for small batch sizes where design specifications take more time relative to overall value creation. There are several developments actively pursued by industry, e.g. the *Industrial Copilot* by the German company SIEMENS (Siemens AG, 2024).

Similarly, the **computer aided manufacturing (CAM) assistant** use case involves using generative AI to generate and optimize automation code for controlling production systems and manufacturing processes. Comparable to the engineering assistant, models can create automation code through natural language input and based on pretrained CAD models, reducing time and error potential. It also improves existing programs via automated error correction.

The **machine assistant** use case uses generative AI to support machine operators and maintenance technicians. The assistant accesses relevant documentation, manuals, guidelines, and process and sensor data, and communicates with users in natural language. It prepares relevant operating information, identifies critical machine and process parameters early, and communicates their causes to prevent downtimes. In case of machine errors, it provides interactive, step-by-step rectification, reducing downtimes.

The **assembly assistant** use case involves using generative AI to support assembly line workers. Generative AI simplifies the creation of assembly instructions by automatically adapting existing instructions for new product variants, saving time. It provides real-time visual and auditory step-by-step instructions, reducing complexity and error risk. Workers can ask questions and receive real-time responses from the assistant, aiding in the assembly process, without the need for touch- or type-based interaction. The Fraunhofer IPT is currently exploring this use case in cooperation with a consortium of industrial partners.

Technology enabler uses generative AI to supplement, optimize, and enhance the functionality of other technologies rather than acting as an independent solution. Generative AI functions as a supporting enabler, improving the performance of technologies like humanoid robots. It enhances their communication skills, allowing them to interpret complex human language and execute detailed instructions (Torka and BMW Group, 2024; Gibney, 2024, pp. 22-24). In the developing field of humanoid robots, it was demonstrated that the integration of generative models has increased the common-sense reasoning and problem solving of such robots.

D Evaluation of use case specific feasibility and implementation effort

The following chapter serves to characterise the use cases regarding the implications of their implementation. According to COOPER the evaluation of implementation projects is ideally done multidimensionally, including the likelihood of technological success next to their general impact (Cooper, 2007, p. 26). As the business impact of the presented use cases relies mainly on the overall business model and product portfolio of the applying company, a general assessment of the use cases feasibility and effort is applicable. Building on the general description of the use cases, an assessment indication about the feasibility and effort of their implementation is provided in this section.

The feasibility dimension refers to the availability of the required data basis. The effort dimension evaluates the use cases in terms of the availability and necessary adaptations of existing generative AI models, as well as the feasibility of implementation without the involvement of external partners. Other dimensions, such as the skills required for implementation, are not considered. The criteria are briefly described in Table 3, whereas the evaluation results are shown in Table 4.

Table 3: Evaluation criteria for feasibility and implementation effort

Dimension	Rating	Description
Feasibility	Low	Data required for the use case type is mostly available outside the company and in decentralized systems.
	Medium	Data required for the use case type is partly available within the company and in centralized systems.
	High	Data required for the use case type is mostly available within the company and in centralized systems.
Effort	Low	The implementation of the use case type requires minor adaptations to existing generative AI models and can be realized without external support.
	Medium	The implementation of the use case type requires moderate adaptations to existing generative AI models and can be realized with moderate support.
	High	The implementation of the use case type requires considerable adjustments to existing generative AI models and cannot be realized without external support.

Due to the high quality of modern manufacturing, there is a lack of image data of potential defect types, which hinders the training of the models. By generating **synthetic training data**, these gaps can be filled by using defect images of similar products as a basis. GANs are suitable due to their ability to generate high-quality images. Manufacturing companies can implement these systems indepdently, as no complex interfaces are required in addition to a small amount of available fault images. The integration of even a few synthetic data points can significantly improve the performance of quality control systems.

Knowledge and training platforms require an extensive data base, including e.g. internal company documents, worker performance data, SOPs, manuals, and technical specifications. Creating personalized training is based on this data and requires strict compliance with data protection regulations. Transformer models, such as a company-specific GPT model, are suitable for context-based text data preparation.

Event Proceedings: ISBN 978-952-65069-7-5

The **engineering assistant** use cases requires data base of existing 3D models, material information, and manufacturing processes, which can be sourced from CAD databases, PLM, and ERP systems. User inputs are also crucial to account for specific preferences and technical requirements. GANs are suitable due to their ability to create high-quality 3D models. Although, the implementation of such a use case probably relies on the interfaces provided by the CAD software manufacturers.

The **CAM assistant** use case needs a comprehensive data base that includes machine data, process flows, and production data, as well as access to central control systems. Transformer models are suitable due to their ability to generate code and understand natural language. The complex control logic must be thoroughly tested to avoid operational disruptions. Close collaboration with machine manufacturers is required.

Similarly, the **manufacturing assistant** use case requires the integration of machine and process data in real-time, as well as Standard Operation Procedures (SOPs) and maintenance logs. The multimodality of generative AI enables intuitive interaction. Transformer-based models are suitable due to their ability to interact multimodally and process real-time data. Close collaboration with machine manufacturers is necessary.

The **assembly assistant** requires the CAD model of the finished product, the product structure, and individual components, as well as information from previous assembly processes. Multimodal models are necessary to generate step-by-step instructions and provide guidance. Real-time integration of process data and assembly steps is needed.

The **technology enabler** application cannot be generally assessed as it heavily depends on the technology addressed. The extent of enabling a technology through generative AI determines the required data base. The complexity of the system landscape influences the effort required for implementation. In complex system landscapes, extensive adjustments are necessary to effectively integrate generative AI.

Table 4: Initial evaluation of use case feasibility and implementation effort

Application Case Type	Feasibility	Effort
Synthetic Training Data Generation	High	Low
Knowledge and Training Platform	Medium	Low
Engineering assistant	Medium	High
CAM assistant	Medium	High
Machine assistant	Medium	High
Assembly assistant	High	Medium
Technology enabler	Dependant on technology	Dependant on technology

5 Conclusion and Outlook

A literature review revealed and initially assessed seven distinct use cases for generative AI in production. The use cases were constructed through an analysis of nine technological characteristics and six capabilities of generative AI. The aforementioned use cases were designed to address a set of ten challenges identified in the context of production. Subsequently, the use cases were subjected to an evaluation process in order to ascertain their feasibility and implementation efforts. These use cases can be employed by manufacturing enterprises to integrate generative AI as a novel technology within the production domain. Although the technology has only recently begun to be applied in a broader range of fields beyond computer science and is still undergoing rapid development,

the variety of use cases demonstrates considerable promise. The authors propose further research in different areas, which they believe to be particularly worthy of investigation:

One significant challenge associated with generative AI is the possibility of 'hallucinations,' in which faulty outputs are presented as convincingly formulated result. This risk has not been sufficiently addressed and could impact the reliability of results, especially in deployment with risk-averse decision-makers. Secondly, market-leading large language models required significant resources for their development, making it unfeasible for resource-limited companies to pursue independent developments. Instead, an investigation into collaboration models should be focused, with participation from both software and manufacturing industry to develop domain-specific models.

In conclusion, domain-specific research on the applications of generative AI is still in its early stages. The rapid advancements in model performance are outpacing the development of suitable domain use cases, leading to a lag in practical implementation. Consequently, there is a significant need for further research from the scientific community to accelerate this process.

References and Notes

Abele, E.; Metternich, J.; Tisch, M. and Kreß, A. (2024): *Learning Factories. Featuring New Concepts, Guidelines, Worldwide Best-Practice Examples*. 2nd edn. Berlin: Springer.

Atkinson, R. (1999): Project management: cost, time and quality. two best guesses and a phenomenon, its time to accept other success criteria. *International Journal of Project Management*. 17th edn. No. 6, pp. 337-342. https://doi.org/10.1016/S0263-7863(98)00069-6

Bajic, B.; Rikalovic, A.; Suzic, N.; Piuri, V. (2021): *Industry 4.0 Implementation Challenges and Opportunities: A Managerial Perspective.* IEEE Systems Journal 15 (1), pp. 546–559.

Bakhtari, Ahmad Reshad; Kumar, Vineet; Mohammad Maqbool Waris; Sanin, Cesar; Szczerbicki, Edward (2020): Industry 4.0 Implementation Challenges in Manufacturing Industries: An Interpretive Structural Modelling Approach. *Procedia Computer Science*, 176th edn., pp. 2384-2393. https://doi.org/10.1016/j.procs.2020.09.306.

Banh, L. and Strobel, G. (2023): *Generative artificial intelligence*. Electron Markets 33, 63. https://doi.org/10.1007/s12525-023-00680-1

Bauernhansl, T. (2020): *Management in der Produktion. Fabrikbetriebslehre 1*. Berlin, Heidelberg: Springer Vieweg.

Bauernhansl, T.; ten Hompel, M.; Vogel-Heuser, B. (2014): *Industrie 4.0 in Produktion, Automatisierung und Logistik. Anwendung, Technologien, Migration*. Wiesbaden: Springer Vieweg.

Bayha, A.; Bock, J.; Boss, B.; Diedrich, C.; Malakuti, S. (2020): *Describing Capabilities of Industrie 4.0 Components*. Federal Ministry for Economic Affairs and Energy (BMWi).

Bengesi, Staphord; El-Sayed, Hoda; Sarker; MD Kamruzzaman; Houkpati, Yao; Irungu, John and Oladunni, Timothy (2024): *Advancements in Generative AI: A Comprehensive Review of GANs, GPT, Autoencoders, Diffusion Model, and Transformers.* IEEE Access 12, pp. 69812–69837.

Bersin, Josh (2024): *Autonomous Corporate Learning Platforms: Arriving Now, Powered by AI*. Available at https://joshbersin.com/2024/02/autonomous-corporate-learning-platforms-arriving-now-powered-by-ai/, updated on 23rd February 2024, accessed 22th October 2024.

Bommasani et al. (2021): *On the Opportunities and Risks of Foundation Models*. Center for Research on Foundation Models (CRFM) Stanford Institute for Human-Centered Artificial Intelligence (HAI) Stanford University. ArXiv.

Brunnbauer, M.; Piller, G. and Rothlauf, F. (2021): *idea-AI: Developing a Method for the Systematic Identification of AI Use Cases.* AMCIS 2021 Proceedings.

Brown, Tom; Mann, Benjamin; Ryder, Nick; Subbiah, Melanie; Kaplan, Jared D.; Dhariwal, Prafulla et al. (2020): Language Models are Few-Shot Learners. DOI: 10.48550/arXiv.2005.14165.

Chan, C.; Colloton, T. (2024): *Generative AI in Higher Education*; *The ChatGPT Effect*. Abingdon: Routledge

Cockburn, A. (2006): *Writing effective use cases*. Boston: Addison-Wesley (The Agile software development series, 16th print).

Cooper, Robert (2007): Managing technology development projects. *Engineering Management Review, IEEE* 35. DOI: 10.1109/EMR.2007.329141.

Dogan, A.; Birant D. (2021): *Machine learning and data mining in manufacturing. In Expert Systems with Applications* (166). Available online at https://www.sciencedirect.com/science/article/pii/S095741742030823X, accessed 10th October 2024.

Dombrowski, U. and Krenkel, P. (2021): *Ganzheitliches Produktionsmanagement. Strategischer Rahmen und operative Umsetzung*. Berlin, Heidelberg: Springer Vieweg.

Foster, D. (2023): *Generative Deep Learning. Teaching machines to paint, write, compose, and play*. 2nd edn. Sebastopol, CA: O'Reilly Media Incorporated.

Fritz, J. (2024): Generative KI in der Industrie. Chancen der generativen künstlichen Intelligenz (KI) verstehen und deren Einsatz zielgerichtet implementieren. *Zeitschrift für wirtschaftlichen Fabrikbetrieb,* 119th edn., no. 5, pp. 344-348. https://doi.org/10.1515/zwf-2024-1056.

Gibney, Elizabeth (2024): The AI revolution is coming to robots: how will it change them? *Nature* 630 (8015), pp. 22–24. DOI: 10.1038/d41586-024-01442-5.

GM, H., Gourisaria, M.K., Pandey, M. und Rautaray, Swarup, S. (2020): A comprehensive survey and analysis of generative models in machine learning. *Computer Science Review*. 38th edn. https://doi.org/10.1016/j.cosrev.2020.100285.

Goodfellow, I.J., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S. et al. (2014): *Generative Adversarial Networks*. https://doi.org/10.48550/arXiv.1406.2661

Herf, Muriel; Hager, Nora; Schreiber, Tobias (2024*): Industry 4.0 Barometer 2024*. MHP Management- und IT-Beratung GmbH.

Hippmann, S.; Klingner, R.; Leis, M. (2018): *Digitalisierung. Anwendungsfelder und Forschungsziele*. Berlin, Heidelberg: Springer Berlin Heidelberg

Hofmann, P.; Jöhnk, J.; Protschky, D.; Stähle, P.; Urbach, N.; Buck, C. (2020): *KI-Anwendungsfälle zielgerichtet identifizieren.* Wirtschaftsinformatik & Management (3). DOI:10.1365/s35764-020-00257-z

Hofmann, P.; Rückel, T.; Urbach, N. (2021*): Innovating with Artificial Intelligence*. Proceedings of the 54th Hawaii International Conference on System Sciences. DOI:10.24251/HICSS.2021.669

Keiber, M. (2024): *Generative künstliche Intelligenz erhält Einzug in Fertigung*. Springer Professional. Available at: https://www.springerprofessional.de/kuenstliche-intelligenz/unternehmensberatung/generative-kuenstliche-intelligenz-erhaelt-einzug-in-fertigung/27067918, accessed 23rd October 2024.

Kingma, D.; Welling, M. (2014): *Auto-Encoding Variational Bayes. https://doi.org/10.48550/arXiv.1312.6114*

Kohl, M.; Malik, B.; Lopitzsch, J.; Fottner, J. (2019): Auswahl- und Priorisierungsmodell für Industrie 4.0-Use Cases in der Logistik. *Logistics Journal.* Vol. 2019. https://doi.org/10.2195/lj_Proc_kohl_de_201912_01

Kintz, M. et al. (2023): *Potenziale generativer KI für den Mittelstand*. Available at: https://www.digital.iao.fraunhofer.de/de/publikationen/Potenziale-Generativer-KI-fuer-den-Mittelstand.html. Accessed: 10th October 2024.

Krauß, J.; Hülsmann, T.; Leyendecker, L.; Schmitt, R. H. (2023): *Application Areas, Use Cases, and Data Sets for Machine Learning and Artificial Intelligence in Production*. DOI:10.1007/978-3-031-18318-8_51

Kreutzer, R. (2023): *Künstliche Intelligenz verstehen. Grundlagen – Use-Cases – unternehmenseigene KI-Journey*. 2nd edn. Wiesbaden, Heidelberg: Springer Gabler.

Kubicek, H. (1976): *Heuristische Bezugsrahmen und heuristisch angelegte Forschungs-design als Elemente einer Konstruktionsstrategie empirischer Forschung*. Berlin (Arbeitspapier / Institut für Unternehmungsführung im Fachbereich Wirtschaftswissenschaft der Freien Universität Berlin, 16).

Leible, S.; Gücük, G.-L.; Simic, D.; Brackel-Schmidt, C. von; Lewandowski, T. (2024): *Zwischen Forschung und Praxis: Fähigkeiten und Limitationen generativer KI sowie ihre wachsende Bedeutung in der Zukunft* (61). https://doi.org/10.1365/s40702-024-01050-x

Liewald, M.; Verl, A.; Bauernhansl, T.; Möhring, H.-C. (2023): *Production at the leading edge of technology*. Proceedings of the 12th Congress of the German Academic

Association for Production Technology (WGP), University of Stuttgart, October 2022. Cham: Springer International Publishing; Imprint Springer (Lecture notes in production engineering).

Löser, A., Tresp, V. et al (2023): Große Sprachmodelle. Grundlagen, Potenziale und Herausforderungen für die Forschung. *Plattform Lernende Systeme*. Munich. https://doi.org/10.48669/pls_2023-3

Luk, M. (2023): *Generative AI. Overview, Economic Impact, and Applications in Asset Management*. DOI:10.2139/ssrn.4574814

Mockenhaupt, A.; Schlagenhauf, T. (2024): *Digitalisierung und Künstliche Intelligenz in der Produktion. Grundlagen und Anwendung*. 2*nd* edn. Wiesbaden, Heidelberg: Springer Vieweg.

Moher, David; Liberati, Alessandro; Tetzlaff, Jennifer; Altman, Douglas G. (2009): *Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement.* In BMJ 339. https://doi.org/10.1371/journal.pmed.1000097

Nayyar, Anand; Kumar, Akshi (2020): *A Roadmap to Industry 4.0: Smart Production, Sharp Business and Sustainable Development*. 1st edn. 2020. Cham: Springer (Springer eBooks Earth and Environmental Science).

Obrenovic, B. et al. (2024): Generative AI and human–robot interaction: implications and future agenda for business, society and ethics. *AI & Soc*. https://doi.org/10.1007/s00146-024-01889-0

Reinhart, G. (2017): *Handbuch Industrie 4.0. Geschäftsmodelle, Prozesse, Technik*. München: Hanser.

Rüegg-Stürm, J. (2005): *Das neue St. Galler Management-Modell. Grundkategorien einer integrierten Managementlehre; der HSG-Ansatz*. 8th edn. Bern: Haupt Verlag.

Russell, S. J.; Norvig, P. (2022): *Artificial intelligence. A modern approach*. 4th edn. Boston: Pearson (Always learning).

Scholz, P. (2022): *Applikation MLS. Applikation maschinell lernender Systeme in der vernetzten adaptiven Produktion*. Aachen: Apprimus-Verlag.

Schuh, G. and Schmidt, C. (2014): *Produktionsmanagement*. 2nd edn. Berlin: Springer Vieweg (VDI-Buch, 5).

Schuh, G., Anderl, R., Dumitrescu, R., Krüger, A. and ten Hompel, M. (2020a): *Industrie 4.0 Maturity Index. Managing the Digital Transformation of Companies – UPDATE 2020.* acatech STUDY.

Schuh, G.; Kozielski, S.; Weber, W.; Gützlaff, A.; Sauermann, F.; Wlecke, S.; Maetschke, J. (2020b): *Digital Lean Lighthouse Factory. Nachhaltig effiziente Produktion.* Zeitschrift für wirtschaftlichen Fabrikbetrieb, No. 115. DOI:10.3139/104.112251

Schuh, G.; Prote, J.-P.; Molitor, M.; Cremer, S. (2020c): *Internet of Production. Steigerung des Wertschöpfungsanteils durch domänenübergreifende Kollaboration*. Heidelberg: Springer Berlin Heidelberg

Schuh, G. and Dölle, C. (2021): Sustainable Innovation. Nachhaltig Werte schaffen. 2nd edn. Berlin, Heidelberg: Springer Vieweg.

Shi, J.; Jain, R.; Doh, H.; Suzuki, R.; Ramani, K. (2023): *An HCI-Centric Survey and Taxonomy of Human-Generative-AI Interactions*. https://doi.org/10.48550/arXiv.2310.07127

Siebel, T. (2024): *GenAI sorgt für Produktivität und macht abhängig*. Available at: https://www.springerprofessional.de/kuenstliche-intelligenz/fertigungstechnik/genai-sorgt-fuer-produktivitaet---und-macht-abhaengig/27333848. accessed 24th September 2024.

Siemens AG (2024): *Unleashing the power of generative AI in industry*. Siemens AG. https://www.siemens.com/global/en/products/automation/topic-areas/tia/future-topics/industrial-copilot.html, accessed 18th October 2024.

Strobel, G.; Banh, L.; Möller, F.; Schoormann, T. (2024): *Exploring Generative Artificial Intelligence: A Taxonomy and Types.* Conference: Hawaii International Conference on System Sciences (HICSS)At: Honolulu, Hawaii.

Sturm, T.; Fecho, M.; Buxmann, P. (2021): *To Use or Not to Use Artificial Intelligence? A Framework for the Ideation and Evaluation of Problems to Be Solved with Artificial Intelligence.* Proceedings of the 54th Hawaii International Conference on System Sciences. Hawaii International Conference on System Sciences: Hawaii International Conference on System Sciences, 2021. DOI:10.24251/HICSS.2021.023

Tomczak, J. (2022): *Deep generative modeling*. Cham: Springer.

Torka, Benedikt; BMW Group (2024): *Successful test of humanoid robots at BMW Group Plant Spartanburg.* Munich/Spartanburg, USA. Available at https://www.press.bmwgroup.com/global/article/detail/T0444265EN/successful-test-of-humanoid-robots-at-bmw-group-plant-spartanburg?language=en, updated on 6th August 2024, accessed 22nd October 2024.

Upton, D. M. (1994): The Management of Manufacturing Flexibility. *California Management Review,* 36(2)th edn., pp. 72–89. DOI: 10.2307/41165745.

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A.N. et al. (2017): *Attention is All you Need*. https://doi.org/10.48550/arXiv.1706.03762

Wecke, B. (2024): *Wachstum durch den Einsatz von Generativer KI. Funktionsweise und Anwendungsgebiete im Marketing*. Wiesbaden, Heidelberg: Springer Gabler (essentials).

Wei, J. et al. (2022): *Emergent Abilities of Large Language Models*. https://doi.org/10.48550/arXiv.2206.07682

Westfall, C. (2024): OpenAI-Powered Humanoid Robot Fills Spot at BMW Assembly Plant, *Forbes,* 7th July 2024. Available at

https://www.forbes.com/sites/chriswestfall/2024/07/07/openai-powered-humanoid-robot-fills-spot-at-bmw-assembly-plant/, accessed 18th October 2024.

Westkämper, E. and Löffler, C. (2016): *Strategien der Produktion. Technologien, Konzepte und Wege in die Praxis*. Berlin, Heidelberg: Springer Vieweg.