POLITECNICO DI TORINO

Collegio di Ingegneria Gestionale
Corso di Laurea in Ingegneria Gestionale – Space Economy

Tesi di Laurea di II livello

Artificial Intelligence in the Manufacturing Context: Technologies, Agents, and Lifecycle Integration

Relatore:
Prof. Giulia Bruno

Candidato:
Edoardo Bruno

Anno Accademico 2024-2025

Summary

CHAPTER 0 – INTRODUCTION

Context and Motivation

The accelerating development of Artificial Intelligence (AI) over the past decade has profoundly reshaped technological, economic, and social landscapes. From language and vision models to autonomous agents and adaptive systems, AI has transitioned from a primarily research-driven domain to a strategic enabler of industrial innovation. Within this transformation, the manufacturing sector represents one of the most promising and complex frontiers. As companies increasingly integrate digital technologies, sensors, cloud infrastructures, and data analytics, AI becomes the cohesive layer connecting these elements into intelligent, self-adaptive ecosystems.

The motivation behind this thesis stems from the growing need to understand how AI, and particularly the emerging paradigm of agentic AI, can be effectively leveraged within the manufacturing context. Traditional automation approaches are no longer sufficient to address the volatility, customization demands, and sustainability challenges that characterize contemporary industrial systems. Instead, intelligent, data-driven, and autonomous solutions are required to optimize processes, enhance decision-making, and extend value creation across the entire product lifecycle.

Objectives

The main objective of this research is to provide a comprehensive analysis of Artificial Intelligence and its specific application to the manufacturing domain, with a particular focus on AI agents as a transformative technological paradigm. The thesis aims to:

1) Review the current state of AI technologies, models, and market dynamics, identifying their implications for organizations and Small and Medium-sized Enterprises (SMEs).
2) Examine the technical and conceptual foundations of AI agents, highlighting their architecture, capabilities, and advantages over traditional static models.
3) Investigate the integration of AI and agentic systems throughout the manufacturing product lifecycle, from design and production to usage and end-of-life, emphasizing both benefits and limitations.
4) Derive a synthesized framework illustrating how AI can drive a more intelligent, sustainable, and adaptive manufacturing ecosystem.

Methodology

The methodology adopted combines an extensive literature review with a structured analytical approach. The research draws from academic publications, industrial reports, and institutional sources to capture both the theoretical underpinnings and real-world applications of AI and agentic systems. Each chapter builds upon the previous one, forming a progressive exploration: from the general technological landscape of AI to the emergence of autonomous agents, to their concrete deployment in manufacturing. The selection of case studies follows thematic relevance, prioritizing evidence of practical implementation, scalability, and alignment with Industry 4.0 and sustainability objectives. Figures, tables, and conceptual models are employed to summarize complex findings and to facilitate cross-phase comparison.

Structure of the Thesis

The thesis is organized into three main chapters, followed by a concluding synthesis.

- Chapter 1 introduces the current AI landscape, tracing its evolution, key technologies, and major stakeholders, while addressing the ethical, legal, and societal challenges of adoption.
- Chapter 2 focuses on AI agents, outlining their defining characteristics, architectures, use cases, and advantages for companies, as well as the associated risks and limitations.
- Chapter 3 applies the analysis to the manufacturing context, exploring the role of AI across all stages of the product lifecycle, from design and operations to end-of-life, culminating in a synthesis that highlights the transition towards an AI-driven, sustainable industrial ecosystem.

This structure ensures both conceptual coherence and practical relevance, offering a comprehensive understanding of how AI and agentic systems are reshaping the manufacturing domain.

CHAPTER 1 - ARTIFICIAL INTELLIGENCE TODAY

1.1 A brief history and key definitions

Artificial Intelligence formally emerged in 1956 at the Dartmouth Summer Research Project on Artificial Intelligence, where the term “artificial intelligence” was coined by John McCarthy and his colleagues, including Marvin Minsky, Nathaniel Rochester, and Claude Shannon [1]. However, foundational work began earlier: in 1943, Warren McCulloch and Walter Pitts introduced the first mathematical model of an artificial neuron, demonstrating the logic-based computation abilities of simple neuron networks [2]. In 1950, Alan Turing published the seminal paper "Computing Machinery and Intelligence", proposing the famous Turing Test as an operational measure of machine intelligence and reframing the question "Can machines think?" into a more practical framework [3].

During the symbolic AI era from the late 1950s through the mid-1990s, research focused on expert systems and rule-based knowledge representation, including logic programs such as the Logic Theorist and General Problem Solver, and knowledge-based systems driven by inference engines [4]. Despite early successes, limitations in scalability and adaptability led to two major "AI winters," where progress slowed and funding diminished [4].

Beginning around 2012, a paradigm shift occurred with the rise of machine learning and deep learning, thanks to increased computational power, large datasets, and new architectures like convolutional neural networks and large language models (LLMs). These approaches enabled major breakthroughs in computer vision, speech recognition, and natural language generation [5].

A widely accepted modern definition of AI characterizes it as "a system's ability to correctly interpret external data, to learn from such data, and to use those learnings to achieve specific goals through flexible adaptation" [6]. Another commonly cited definition describes AI as systems capable of tasks such as learning, reasoning, decision-making, creativity, and autonomy-functions traditionally associated with human intelligence [7].

1.2 Current AI technologies and models

Artificial Intelligence encompasses several advanced technologies, each enabling distinct capabilities and applications within industry and research.

First, Natural Language Processing (NLP) has made remarkable strides, especially with the rise of transformer-based architectures such as BERT and GPT. These models now support tasks like machine translation, text summarization, question answering, and more, exhibiting human-like fluency and performance in language understanding and generation [8] [9].

In computer vision (CV), breakthroughs such as convolutional neural networks (CNNs) and, more recently, Vision Transformers (ViTs) have significantly improved performance in image classification, object detection, segmentation and video analysis. Transformers in vision also enable modeling long-range dependencies with fewer inductive biases compared to CNNs and support multi-modal tasks that combine visual and textual input [10].

A central pillar of today's developments is Large Language Models (LLMs), massive neural networks trained on billions or even trillions of parameters. These models, such as GPT-4 o, LLaMA, Claude 3 or Gemini, function as foundation models for a multiplicity of downstream tasks, combining both generative and understanding capabilities at scale [11].

Further, multi-modal LLMs (sometimes called VLMs or MM-LLMs) have emerged as models capable of processing and generating across modalities (text, image, audio, video). These models open doors to applications like image captioning, video generation, and speech-to-text/generation all within one unified architecture [8] [12].

Finally, AI agents are autonomous systems built upon LLMs and enhanced machine learning, are increasingly targeting decision-making, task execution, and multi-step workflows without human prompting. These agentic AI systems extend far beyond static models, enabling proactive planning, tool usage, and real-time adaptation in complex environments [13].

1.3 Main market players

The AI landscape today is dominated by three primary categories of players: Big Tech corporations, innovative startups, and the open-source communities.

1.3.1 Big Tech incumbents

Major technology firms, often referred to as “Big Tech”, play a dominant role across all layers of today’s AI landscape. NVIDIA leads the data-center GPU market with roughly a 92 % global market share, supplying the backbone infrastructure for contemporary AI workloads. Its CUDA software ecosystem further strengthens its position as the de facto platform for AI development [14]. Hyperscalers like Microsoft, Amazon (AWS), and Google control most of the foundation-model and model-management platform market: Microsoft holds about a 39 % share, AWS 19 %, and Google 15 % in 2024. Microsoft’s strategic alliance with OpenAI and its enterprise AI offerings (e.g. Copilot, Azure AI) reinforce its leadership position [15] [16]. AWS couples its cloud dominance with proprietary chips (e.g. Trainium, Inferentia) and investments such as a multi-billion-dollar partnership with Anthropic. Google continues to invest heavily (circa US $20–25 billion in 2025 capital expenditures, with a significant portion allocated to AI) and builds its AI capabilities via Vertex AI and internal models (Gemini series) [15].

Other tech giants like Meta also shape the AI market via open-weight initiatives such as the Llama models and large infrastructure investments (e.g. over $35 billion in AI-related capital spending in 2024–2025) [17]. Meta CTO claims that such openness helps spawn innovation in startups and ecosystems that challenge legacy leadership [18].

1.3.2 Fast-growing startups

A wave of AI startups now rivals Big Tech in innovation and influence. OpenAI, though not a public company, is nearing a valuation of $150–300 billion, thanks to its GPT family and ChatGPT platform [15] [19]. Despite speculation, OpenAI has not released open-weight models; its strategy remains largely closed-source, in contrast to open-source challengers like China's DeepSeek.

Anthropic, also growing rapidly, focuses on safety-centered LLMs and has formed major infrastructure alliances, including with AWS [15]. China-based DeepSeek garnered attention by open-releasing its DeepSeek LLM and DeepSeek-Coder models, prompting significant global downloads and putting pressure on proprietary giants [20].

European startup Mistral AI, based in Paris, has raised over €600 million and offers open-weight, performant LLMs, becoming one of the top global model creators outside the Bay Area [15].

Other notable startups include Hugging Face, which hosts open-source models like BLOOM and a vibrant ecosystem, valued at several billion USD [15], as well as Cohere (Canadian NLP startup offering enterprise-grade LLMs & APIs) and MindsDB (automated machine-learning from business data, with open-source roots). Specialized hardware startups such as Cerebras Systems, Astera Labs, Groq and others are innovating at the chip and data-infrastructure level to compete with NVIDIA and AMD.

1.3.3 Open-source communities and ecosystems

Beyond companies, community-driven open-source ecosystems are central to democratizing AI. Entities such as Hugging Face and Mistral focus on open-weight or open-source model releases, fostering transparency and broader access [15]. U.S. and international initiatives are emerging to support open AI development, including federal funding for open models and compute access [17].

Hardware players like AMD and NVIDIA support open-source standards and tooling, for example, AMD's ROCm platform and NVIDIA's support of PyTorch and Linux Kernel improvements. Initiatives such as interoperability frameworks (e.g., from MLCommons or the Open Model Initiative) aim to improve model portability and safety, though no widely recognized "Model Context Protocol" currently exists [17]. The broader result is an ecosystem where startups and smaller groups can build, modify, and deploy powerful AI models outside closed corporate systems [16].

1.4 Available services and applications

Artificial Intelligence has evolved from a predominantly research-driven field into a broad ecosystem of services accessible to enterprises, institutions, and individual users. These solutions are typically delivered as cloud-based APIs, enterprise platforms, or consumer-facing applications, lowering the barriers to entry and enabling even small and medium-sized enterprises (SMEs) to benefit from advanced capabilities without developing models internally [6]. The main categories of AI services available today include chatbots and conversational agents, intelligent productivity tools, automation and robotic process automation (RPA), predictive analytics, and emerging multimodal generative AI applications.

1.4.1 Chatbots and Conversational Agents

Chatbots and virtual assistants represent one of the most visible applications of Natural Language Processing. Systems such as ChatGPT by OpenAI, Claude by Anthropic, or Gemini by Google, allow users to interact via natural language, providing answers, recommendations, or task execution support [8] [9].

- Strengths: they enable intuitive and human-like interaction, reduce workload in customer service and enterprise support functions, and can continuously improve through model fine-tuning.
- Weaknesses: risks include hallucinations (generation of inaccurate or fabricated content), challenges regarding data privacy and regulatory compliance, and a high dependency on the quality and coverage of training data [11].

1.4.2 Intelligent Assistants and Productivity Tools

Another rapidly growing area is the integration of AI into productivity ecosystems. Examples include Microsoft Copilot for Microsoft 365, Gemini for Workspace by Google, Notion AI, and specialized research assistants such as NotebookLM. These tools act as “co-pilots”, supporting tasks such as summarization, drafting, scheduling, or document creation as summarized in Table 1 [15].

- Strengths: seamless integration with widely adopted enterprise platforms, measurable time savings in routine activities, and adaptability to organizational knowledge bases.

- Weaknesses: their adoption is often limited by subscription costs, potential gaps in domain-specific expertise, and risks of over-reliance, which may reduce users' critical engagement with content [16].

Table 1 - Comparative overwiev

Assistant	Provider	Main functions	Strengths	Weaknesses
Microsoft Copilot	Microsoft	Integration with Office 365 (Word, Excel, PowerPoint, Outlook, Teams); text generation; summarization; data insights	Seamless integration with Microsoft ecosystem; enterprise-ready security; high adoption in business contexts	Subscription cost; limited to Microsoft ecosystem; occasional lack of domain-specific depth
Gemini for Workspace	Google	Integration with Google Workspace (Docs, Sheets, Slides, Gmail, Meet); smart writing; meeting summarization	Strong synergy with Google cloud tools; accessible UI; collaboration-friendly	Still less mature than Copilot in enterprise; tiered access within Google Workspace; data privacy concerns
Notion AI	Notion Labs	Document drafting; content summarization; brainstorming; task management support	Integrated into a popular productivity platform; affordable; good for individuals/SMEs	Narrower enterprise adoption; not as powerful for complex analytics or workflows
NotebookLM	Google Research	Research support; personalized "AI notebook" with citations and summaries of documents	Strong in research and knowledge management; focus on reliable sources and citations	Limited availability (still experimental); not optimized for enterprise workflows

1.4.3 Automation and Robotic Process Automation (RPA)

AI-driven automation extends traditional RPA platforms (e.g., UiPath, BluePrism, Automation Anywhere) by enabling the processing of unstructured data such as emails, forms, and invoices [14]. This technology is particularly suited for industries with repetitive, high-volume back-office tasks, such as banking, insurance, and logistics.

- Strengths: significant efficiency gains, cost reduction, and scalability across operations.
- Weaknesses: implementation often requires substantial upfront investment and complex integration; furthermore, RPA solutions show limited adaptability when workflows or processes change frequently [15].

1.4.4 Predictive Analytics and Decision Support Systems

Predictive systems, built on machine learning and deep learning models, are now widely available through cloud AI platforms (e.g., AWS SageMaker, Google Vertex AI, Microsoft Azure AI). They support applications ranging from demand forecasting and fraud detection to predictive maintenance in manufacturing [5] [11].

- Strengths: these solutions enhance decision-making by providing data-driven insights, offering measurable return on investment when effectively deployed.
- Weaknesses: their reliability depends on the availability of large volumes of high-quality data; biases in datasets may lead to distorted outputs, and their deployment and monitoring often require specialized expertise [6].

Examples of predictive analytics applications across different industries are summarized in Table 2.

Table 2 - Predictive analytics applications across industries

Industry	Use case	Example tool	Main benefit
Manufacturing	Predictive maintenance of machinery	Azure IoT Central; PTC ThingWorx; Siemens MindSphere	Reduced downtime; optimized asset utilization
Healthcare	Patient risk prediction; diagnostic support	Google Cloud Healthcare API; Tempus AI; Medica's AI diagnostics	Improved diagnosis accuracy; proactive treatment
Finance	Fraud detection; credit scoring	FICO; SAS Analytics; Azure AI	Reduced financial losses; enhanced risk management
Retail	Demand forecasting; customer behavior analysis	SAP Predictive Analytics; Amazon Forecast	Better inventory management; personalized offers

1.4.5 Emerging Multi-modal and Generative AI Services

The latest frontier is represented by multi-modal AI models, capable of processing and generating across different modalities (text, image, audio, video). Examples include GPT-4o by OpenAI, Gemini 1.5 by Google, or image/video generation tools such as Stable Diffusion and Runway ML [12].

- Strengths: they open new opportunities in creative industries, education, and marketing by enabling cross-media generation and multimodal interaction.
- Weaknesses: however, they raise significant ethical challenges (deepfakes, misinformation), demand very high computational resources, and pose unresolved questions around intellectual property and copyright [13].

In summary, the availability of AI services reflects a democratization of access to advanced technologies, accelerating adoption across industries. Their strengths lie in efficiency, scalability, and innovation potential, while their weaknesses highlight challenges of accuracy, ethics, and dependency on external providers. The effective adoption of these tools depends not only on their technical maturity but also on the strategic capacity of organizations to integrate them responsibly within their processes.

1.5 Potential impact on work, business processes, and SMEs

The rapid evolution of Artificial Intelligence (AI) is producing significant implications for the labor market, organizational processes, and SMEs. The diffusion of cloud-based services and foundation models has reduced barriers to adoption, enabling even smaller organizations to leverage advanced technologies that were previously reserved for large corporations [6] [7]. This democratization is reshaping competitive dynamics, requiring both firms and individuals to adapt strategically.

1.5.1 Impact on work and workforce skills.

AI technologies automate repetitive and rule-based tasks, such as back-office operations, routine data analysis, and document management, thus reducing the demand for low-skill labor in these areas [14] [15]. At the same time, they increase the need for high-skill profiles capable of managing, supervising, and co-developing AI systems. Studies highlight the emergence of new professional roles in AI system supervision, AI auditing, data governance, and human-AI workflow design [7] [11]. For many workers, this transformation requires reskilling and upskilling programs to remain competitive, while organizations must ensure ethical and inclusive transitions. Figure 1 illustrates the categories of tasks with the highest potential for automation or augmentation through AI, highlighting routine work that can be fully automated versus strategic tasks that are more likely to benefit from human–AI collaboration.

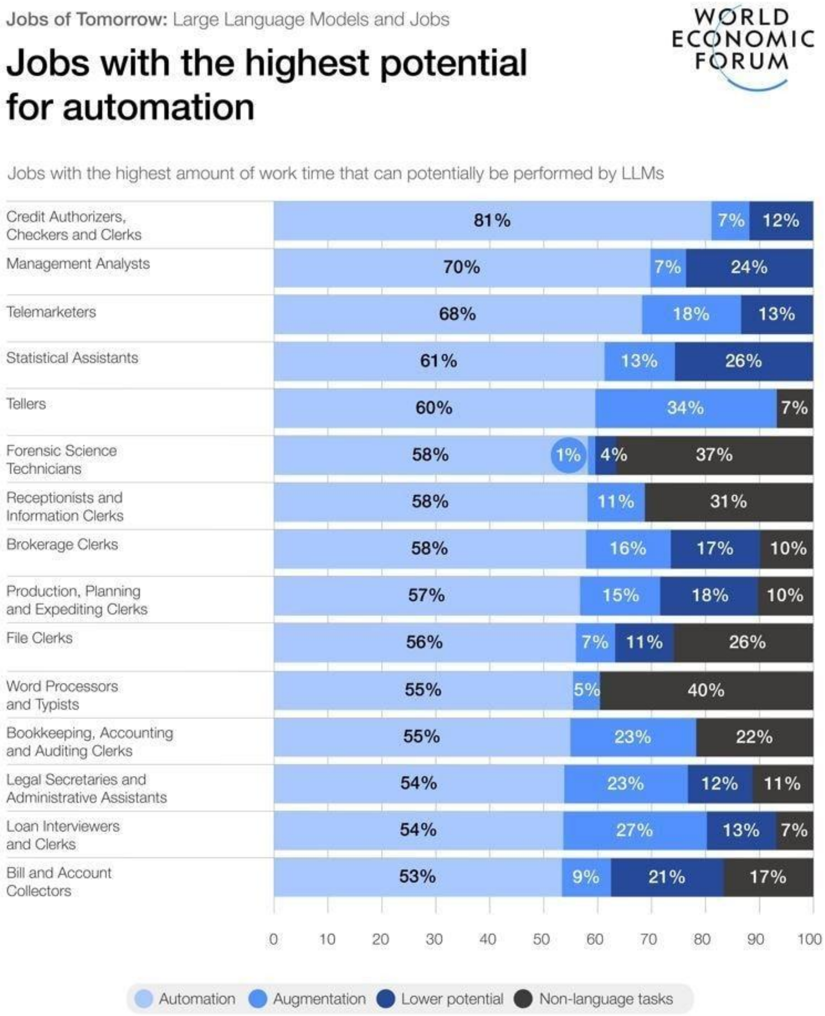

Figura 1 - Categories of tasks most affected by AI: automation versus augmentation

1.5.2 Impact on business processes.

Within organizations, AI acts as a driver of efficiency and innovation across the value chain. Intelligent automation and Robotic Process Automation (RPA) extend beyond traditional workflows, allowing the processing of unstructured data such as invoices, contracts, or emails [14]. Predictive analytics tools support decision-making in areas such as demand forecasting, fraud detection, or predictive maintenance [5] [11]. Furthermore, generative AI and intelligent assistants enable knowledge management, rapid prototyping, and content creation [8] [15]. However, these benefits are balanced by risks: dependency on external providers, data security concerns, and potential over-reliance on algorithmic outputs [6].

Table 3 - Benefits and risks of AI adoption in business processes

Functional area	Benefits	Risks / Challenges
Operations & Automation	Higher efficiency through automation of repetitive tasks; cost reduction; scalability	High implementation costs: limited adaptability when workflows change frequently
Decision-making & Analytics	Enhanced forecasting accuracy; better risk management; data-driven insights	Biased or low-quality data may distort outputs; need for specialized expertise
Customer Interaction	24/7 support via chatbots; personalized recommendations; improved engagement	Risk of hallucinations; data privacy and compliance concerns
Knowledge Management & Creativity	Accelerated content creation; support in research and documentation; innovation in product design	Over-reliance on AI outputs; intellectual property and copyright issues related to training data and output use
SMEs Adoption	Access to advanced tools without proprietary infrastructure; competitive differentiation; reduced entry barriers	Cost sensitivity (especially for premium features); lack of internal expertise and customizations

1.5.3 Impact on SMEs.

SMEs represent a particularly interesting field of application. Historically, limited resources and lack of specialized personnel hindered their ability to adopt emerging technologies. With the advent of AI-as-a-Service, even small firms can now access advanced tools without developing proprietary infrastructures [6] [7]. For example, cloud-based platforms allow SMEs to deploy chatbots for customer support, predictive systems for inventory management, or generative tools for marketing content [11] [12]. These applications contribute to cost reduction, improved customer engagement, and competitive differentiation. Nevertheless, challenges persist: subscription costs, lack of internal expertise, and the risk of adopting generic solutions not aligned with specific business needs [15] [16]. As shown in Figure 2, large enterprises exhibit significantly higher AI adoption rates compared to SMEs, underscoring the structural advantage in resources and implementation capacity.

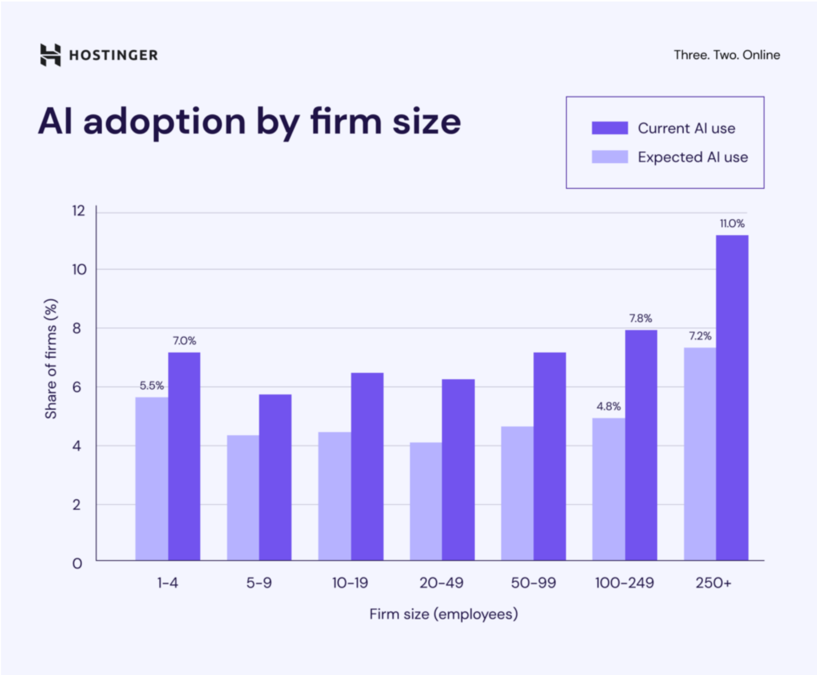

Figura 2 - AI adoption gap between large corporations and SMEs.

1.6 Ethical, legal, and societal challenges

The rapid evolution of Artificial Intelligence (AI) not only brings opportunities for innovation and efficiency but also raises a series of ethical, legal, and societal challenges that must be critically addressed. These challenges span from issues of transparency and accountability to broader questions of social equity, labor dynamics, and human rights. Managing them responsibly is essential to ensure trust, sustainable adoption, and alignment with human values.

1.6.1 Ethical challenges

One of the most debated issues is algorithmic bias: AI systems trained on large datasets often inherit and amplify social, cultural, or gender biases present in the data [6] [11]. This can result in discriminatory outcomes in areas such as recruitment, financial services, or predictive policing. Closely related is the question of explainability. While deep learning and large language models provide impressive performance, their decision-making processes are often opaque (“black-box problem”), limiting users’ ability to understand or contest outcomes [5] [12].
Another central theme is responsibility for harm: when an AI system causes damage, through incorrect predictions, flawed recommendations, or autonomous actions, defining liability between developers, providers, and users becomes legally and ethically complex [7].

Tabella 4 - Ethical principles and technical measures in AI

Ethical principle	Description	Technical / organizational measures
Fairness & non-discrimination	Avoid biased or discriminatory outputs based on gender, ethnicity, age, or other protected attributes.	Bias detection and mitigation tools; balanced datasets; fairness-aware algorithms; continuous monitoring and audits.
Transparency & explainability	Ensure that decisions made by AI systems can be understood and traced by stakeholders.	Explainable AI (XAI) models; interpretable ML techniques (LIME, SHAP); transparent documentation of data training and model design.
Accountability & responsibility	Clearly define who is responsible in case of errors, harms, or misuse of AI.	Governance frameworks; AI ethics boards; impact assessments; traceability logs.
Privacy & data protection	Protect users' personal and sensitive data during collection, processing, and storage.	Privacy-preserving techniques (federated learning, differential privacy, homomorphic encryption); compliance with GDPR and alignment with the EU AI Act.
Safety & robustness	Guarantee that AI systems behave reliably under diverse and unexpected conditions.	Stress-testing and adversarial testing; red-teaming; robust training; ongoing monitoring of deployed systems.
Human oversight	Ensure that final decision-making remains under meaningful human control.	Designs ensuring meaningful human oversight, including "human-in-the-loop" or "human-on-the-loop" mechanisms; override capabilities; training for oversight personnel.

1.6.2 Legal challenges

The deployment of AI also confronts significant legal uncertainties. First, data protection and privacy remain critical concerns. Regulations such as the EU's General Data Protection Regulation (GDPR) impose strict obligations on how personal data is collected, processed, and stored, raising compliance challenges for AI models trained on large-scale datasets [6].
Second, intellectual property (IP) is an unresolved frontier. Generative AI systems produce original content (texts, images, designs) that may draw from copyrighted training data. This creates disputes over authorship, ownership, and potential infringement [12].
Additionally, emerging legislation like the EU AI Act, entered into force in June 2024 and will be fully applicable by 2026, introduces a risk-based framework, categorizing AI systems from minimal to unacceptable risk and imposing requirements on high-risk applications (e.g., biometric identification, healthcare

diagnostics, employment screening). These regulations could significantly influence the design and deployment strategies of companies and SMEs [11].

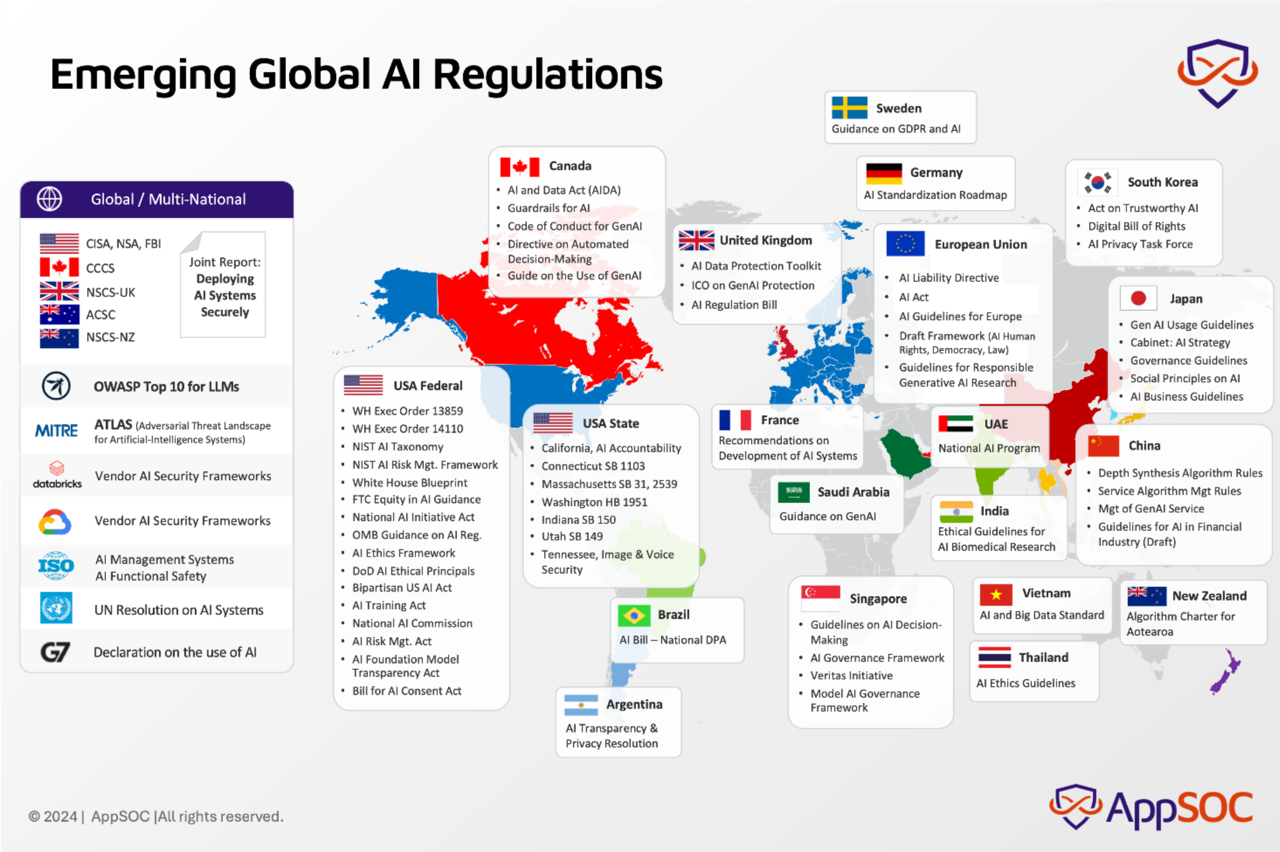

Figure 3 - Keeping up with Global Regulations Around AI: A Complicated Map (august 2024)

1.6.3 Societal challenges

On the societal level, AI raises questions about labor markets, inclusion, and trust. As discussed in Section 1.5, automation displaces routine jobs but creates demand for high-skill roles in AI supervision and governance [14] [15]. However, this transition risks exacerbating inequalities if reskilling programs are insufficient or unevenly distributed.

AI also impacts democratic processes and information ecosystems. The rise of generative models capable of producing deepfakes and synthetic media raises concerns over misinformation, manipulation, and erosion of public trust in information sources [13]. This has led to calls for watermarking, content provenance systems, and stricter digital governance [17].

Finally, societal acceptance is crucial: surveys show that users' trust in AI is fragile, heavily influenced by perceptions of safety, fairness, and transparency [6]. Without public confidence, even technically advanced solutions may encounter resistance to adoption.

1.6.4 Towards responsible AI

Addressing these ethical, legal, and societal issues requires a multi-stakeholder approach, involving policymakers, companies, academia, and civil society. International organizations, such as UNESCO and the OECD, have proposed global principles of trustworthy AI, stressing human oversight, robustness, fairness, and transparency [6].
From a corporate perspective, leading firms are beginning to adopt AI governance frameworks, including bias audits, ethics boards, and standardized reporting on AI models [15]. However, implementation remains uneven, and SMEs often lack the resources to establish such structures independently.

In conclusion, ensuring that AI is developed and deployed responsibly is as critical as advancing its technical capabilities. Only by integrating ethical reflection, legal compliance, and societal dialogue can AI deliver sustainable benefits without undermining fundamental rights and social cohesion.

CHAPTER 2 - FOCUS: AI AGENTS

2.1 Definition of AI agent

The concept of AI agents represents one of the most relevant recent evolutions in the field of Artificial Intelligence. While traditional models, often referred to as static models, are designed to perform specific tasks when prompted by human users, AI agents go beyond this paradigm by introducing autonomy, proactivity, and interaction with external environments [11].

A static model is typically characterized by a fixed set of capabilities: it receives an input, processes it according to its trained parameters, and produces an output. Large Language Models (LLMs), for example, can generate text or answer questions with high accuracy, but they remain reactive tools: without human prompting, they do not initiate actions or adapt to dynamic tasks [9] [11].

In contrast, an AI agent is an autonomous system built upon advanced models, often foundation models such as LLMs, combined with additional layers that enable reasoning, memory, planning, and interaction with tools or environments [13]. This shift allows agents not only to respond but also to act: they can break down complex goals into sub-tasks, decide which tools to use (e.g., search engines, APIs, robotic systems), and adapt their strategies according to feedback from the environment [11] [13].

One of the most cited definitions describes an AI agent as “a system situated in and interacting with an environment, capable of perceiving it through sensors and acting upon it through effectors, with the goal of achieving specific objectives” [7]. This definition underscores the agent’s continuous interaction loop, which distinguishes it from the one-shot functioning of static models.

2.1.1 Comparative dimensions

The main differences between static models and AI agents can be synthesized along several dimensions: autonomy, adaptability, goal orientation, and tool use. While static models are limited to executing predefined tasks, agents possess the capacity for self-directed operation and multi-step workflows [13].
Table 5 below provides a structured comparison between the two paradigms.

Tabella 5 - Comparison between static AI models and AI agents

Dimension	Static AI Models	AI Agents
Input/Output	Single input → single output	Continuous perception-action loop
Autonomy	Primarily reactive; requires explicit prompts or external orchestration for complex tasks	Proactive and autonomous in initiating tasks

Dimension	Static AI Models	AI Agents
Adaptability	Limited, fixed behavior	Dynamic, capable of learning from environmental feedback
Goal orientation	No intrinsic goals, task-bound	Oriented towards objectives defined by human or context
Tool use	None or limited (e.g., APIs if connected)	Can select, combine, and execute external tools
Memory	Typically, absent or very short-term	Equipped with long-term and working memory capabilities
Examples	GPT for text generation, ResNet for image classification	LangChain-based research assistants, robotic systems, autonomous software agents (e.g., using LangGraph or Microsoft Autogen

2.1.2 Towards agentic AI

This transition reflects a broader move from passive AI systems to agentic AI. The latter encompasses decision-making, multi-step reasoning, and adaptive behavior in real-world contexts [13]. By integrating perception, memory, and tool-use, agents embody a form of operational intelligence closer to human-like problem-solving than static models ever could.

To visually reinforce this distinction, Figure 4 illustrates the evolution from static models to agent-based architectures, showing how added layers (memory, planning, autonomy) transform a reactive tool into an active, context-aware system.

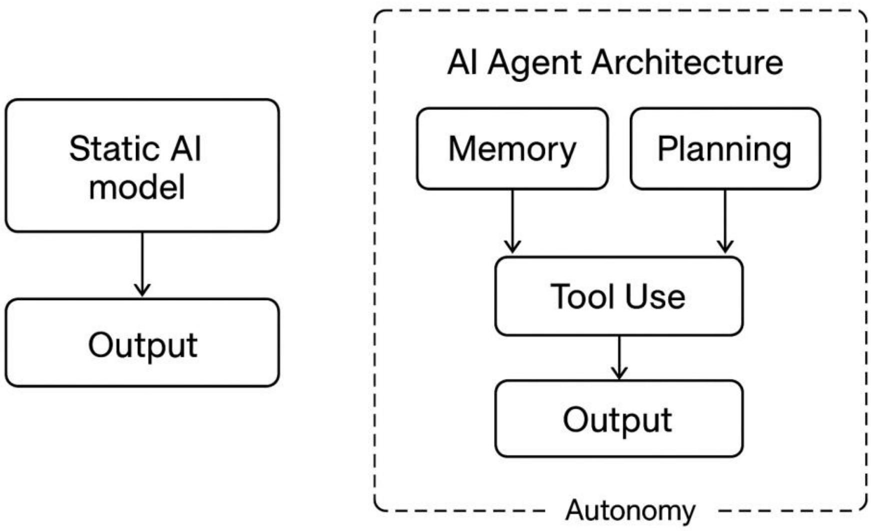

Figure 4 - Evolution from static AI models to agent-based architecture. Original elaboration by the author, based on [7], [9], [11], [13]

2.2 How they work: architecture and basic principles

The functioning of AI agents is based on an architectural framework that extends the paradigm of large models into a dynamic, interactive system capable of perceiving, reasoning, and acting in complex environments. Unlike static AI models, which operate in a single input–output cycle, agents are designed as continuous loops of perception, decision, and action, mediated by four foundational principles: planning and reasoning, memory, tool use, and autonomy [11] [13].

2.2.1 Prompting and reasoning mechanisms

Prompting represents the primary interface between the agent and its environment or human users. While in static models prompts serve only as queries for one-shot responses, in agentic systems prompting becomes iterative and context aware. Agents are capable of decomposing high level goals into sub-prompts, generating reasoning chains, and reformulating queries to external tools when needed [9] [13]. This mechanism introduces a form of meta-reasoning, whereby the agent can evaluate intermediate outputs, refine them, and redirect its course of action.

2.2.2 Memory: short-term, long-term, and episodic

Another core distinction is the presence of memory layers. Static LLMs are constrained by short context windows, limiting their ability to "remember" beyond a few thousand tokens [11]. Agents, by contrast, integrate:

- Short-term (working) memory, used to store the immediate conversational or task state.
- Long-term memory, supported by external databases, vector stores, or knowledge graphs, allowing the agent to retrieve past interactions and contextual information [13].
- Episodic memory, enabling the system to recall sequences of experiences and adjust its strategy over time.

The integration of memory fosters continuity and adaptation, qualities essential when deploying agents in business workflows or manufacturing contexts, as anticipated in Chapter 1.5 and further developed in Chapter 3.

2.2.3 Tool use and environment interaction

A defining characteristic of AI agents is their capacity to use tools. These may range from software APIs and search engines to robotic actuators in physical environments. Through frameworks like LangChain, LangGraph, or Microsoft Autogen, agents can dynamically select and orchestrate tools according to the problem at hand [13]. This capability transforms them from closed systems into open, modular ecosystems, where knowledge is extended beyond the model's internal parameters. Table 6 provides an overview of the main categories of tools that agents can integrate, emphasizing their role in extending the agent's capabilities beyond the core model.

Tabella 6 - Categories of tools in AI agents

Category	Examples	Role in agent workflows
Data retrieval	APIs, vector databases, knowledge graphs	Access to external knowledge; long-term memory support
Analytics	Python execution, ML models, BI tools	Data processing, predictive modeling, decision support
Communication	Email, messaging APIs, speech synthesis	Interaction with humans or other systems
Physical actuators	Robotic arms, IoT devices, drones	Execution of tasks in real-world environments
Information retrieval	Search engines, web scraping tools	Access to up-to-date, real-time information from the internet

2.2.4 Autonomy and self-directed behavior

Perhaps the most distinctive feature of agents lies in their autonomy. By combining prompting, memory, and tool use, they can plan multi-step tasks, monitor their execution, and adjust strategies in response to feedback from the environment [7] [13]. Autonomy does not mean independence from human oversight: as highlighted in Chapter 1.6 on ethical and societal challenges, meaningful human control remains central. Instead, autonomy refers to the ability to operate proactively, reduce human micromanagement, and act as co-pilots in decision making and operational contexts. This anticipates their role in intelligent automation and task execution for enterprises (Chapter 2.4) and their transformative potential in manufacturing processes (Chapter 3.4).

2.2.5 Towards an integrated architecture

Synthesizing these elements, the architecture of an AI agent can be visualized as a layered system:

- Core model (often an LLM or multimodal foundation model) providing baseline reasoning and generative capabilities [11].
- Control layer managing prompting strategies, goal decomposition, and reasoning chains.
- Memory layer ensuring context continuity and adaptive learning.
- Tool interface enabling external actions, both digital and physical.
- Orchestration layer integrating planning, memory, and tools to enable autonomous, goal-driven behavior (autonomy is an emergent property).

In summary, the architecture of AI agents rests on four foundational elements: prompting, memory, tool use, and autonomy. Which, when combined, transform static models into adaptive and continuous operating systems. These principles should not be seen as isolated blocks but as interconnected components that shape the agent's functionality and overall level of sophistication. Understanding these mechanisms provides the groundwork for the following sections, where real-world applications, business advantages, and current limitations will be explored.

2.3 Real-world and experimental use cases

The growing interest in AI agents has led to a rapid proliferation of both experimental prototypes in research laboratories and commercially available solutions on the market. These systems embody the architectural principles described in Section 2.2, reasoning, memory, tool use, and autonomy, while translating them into practical applications that extend beyond static large models. In this sense, real-world deployments of agents represent a natural evolution of the applications introduced in Chapter 1.4, where chatbots, intelligent assistants, and predictive systems were still bounded by reactive paradigms [8] [11].

A first category of use cases concerns <u>research-oriented experimental agents</u>. Academic institutions and industrial research labs have been at the forefront of testing agentic architectures in controlled environments. For instance, initiatives based on LangChain or LangGraph provide frameworks where agents can autonomously query databases, generate structured research reports, or design experiments by orchestrating multiple tools [13]. Projects such as AutoGPT and BabyAGI served as early proof-of-concepts for autonomous workflows and inspired further development in agentic AI, despite their limitations in reliability and scalability. While still limited by reliability and cost, these experiments have highlighted the potential of agents in domains such as scientific discovery and autonomous coding [9] [13].

In the commercial sphere, <u>enterprise-oriented co-pilot agents</u> are among the most visible applications. Building upon the productivity assistants discussed in Section 1.4.2, companies are now releasing enhanced versions with agentic capabilities. Microsoft Copilot and Gemini for Workspace, for example, have introduced features that go beyond one-shot text generation, enabling task orchestration across applications, proactive workflow management, and integration with enterprise knowledge bases [15]. These systems illustrate how the architectural principles of memory and tool use (see Section 2.2.2 and 2.2.3) can be concretely leveraged to create digital co-workers rather than simple assistants.

A further line of development involves autonomous agents in customer-facing contexts. Companies in e-commerce and finance are experimenting with agents capable of not only answering queries but also executing transactions, adjusting offers dynamically, and escalating issues based on contextual awareness. Compared to traditional chatbots (Section 1.4.1), these agents demonstrate higher levels of autonomy and adaptive reasoning, bridging the gap between conversational interfaces and operational decision-making [11] [13].

Another promising domain is AI-driven robotics and IoT integration. Here, agents are connected to physical actuators, enabling real-time perception–action loops in dynamic environments. In manufacturing and logistics, for example, prototypes of robotic agents can coordinate fleets of autonomous vehicles, adjust production lines, or carry out predictive maintenance by accessing sensor data. These use cases anticipate the applications that will be analyzed in Chapter 3, where manufacturing contexts provide fertile ground for embedding agentic intelligence [11] [13]. A graphical synthesis of these categories of use cases is provided in Figure 5, which situates agents along a continuum from experimental prototypes to enterprise-ready solutions.

In addition, multi-agent simulation environments are emerging as valuable experimental grounds. Researchers have developed digital societies of agents capable of exhibiting emergent behaviors, cooperation, and competition, offering insights into collective intelligence and organizational dynamics [13]. Such simulations not only advance theoretical understanding but also provide testbeds for evaluating risks, such as coordination failures or bias propagation, which are critical for real-world deployment (further discussed in Section 2.5).

Finally, specialized industry use cases demonstrate how agents can generate tangible value in high-stakes domains. In healthcare, experimental diagnostic agents assist doctors by autonomously retrieving medical literature, proposing differential diagnoses, and scheduling follow-ups, all while integrating memory of prior cases [11]. In finance, agents are being tested for autonomous trading strategies and real-time fraud detection, although strict regulatory oversight remains necessary [14]. These examples connect back to the broader applications of predictive analytics and decision support systems (Section 1.4.4), showing how agents transform them into continuous, adaptive, and autonomous processes.

In summary, the landscape of agentic AI applications is highly heterogeneous, ranging from open-source experimental prototypes to enterprise-ready copilots and robotics integrations. What unites these initiatives is the progressive realization of the principles discussed in Section 2.2 and their convergence toward the needs of companies, as will be explored in Section 2.4. To reinforce this, Table 7 provides a comparative overview of selected real-world and experimental use cases, highlighting the transition from research prototypes to market-ready solutions.

Tabella 7 - Selected real-world and experimental use cases of AI agents

Domain	Example	Key Features	Connection with agent principles
Research prototypes	AutoGPT, BabyAGI (research prototypes)	Autonomous reasoning chains, task decomposition	Prompting, memory, autonomy
Enterprise copilots	Microsoft Copilot, Gemini for Workspace	Task orchestration, integration with apps	Tool use, memory, autonomy
Customer interaction	Experimental e-commerce agents for personalized purchasing	Execute transactions, adaptive offers	Reasoning, autonomy
Robotics & IoT	Autonomous warehouse robots	Sensor data integration, fleet coordination	Perception–action loop, tool use
Simulation environments	Multi-agent societies	Emergent collaboration, coordination	Multi-agent reasoning, autonomy
Specialized industries	Healthcare diagnostic agents, financial trading bots	Case memory, decision support	Long-term memory, reasoning

2.4 Specific advantages for companies

The introduction of AI agents into corporate contexts offers tangible advantages that go beyond the incremental improvements provided by traditional AI systems. While static AI models already enabled significant progress in areas such as chatbots, predictive analytics, and productivity tools (see Chapter 1.4), the agentic paradigm extends these capabilities by adding autonomy, memory, and adaptive reasoning (as discussed in Chapter 2.2). For companies, this evolution translates into a set of strategic benefits that impact efficiency, decision-making, and innovation across the value chain.

2.4.1 Intelligent Automation

One of the clearest advantages of AI agents for enterprises lies in intelligent automation. Building upon the foundations of RPA (see Section 1.4.3), agents enhance automation by incorporating the ability to process unstructured data, adapt workflows in real time, and interact with heterogeneous systems. For example, whereas traditional RPA could extract structured data from invoices, an

AI agent can autonomously classify non-standard documents, query external databases for missing information, and escalate ambiguous cases to human supervisors [11] [13].

This level of adaptability reduces the brittleness of conventional automation solutions and ensures scalability across dynamic environments.

Moreover, intelligent automation driven by agents allows companies to achieve not only cost reduction but also resilience. By continuously monitoring execution and adapting to changes (e.g., regulation updates, fluctuating supply chains), agents align with the broader need for organizational agility emphasized in Chapter 1.5.2. In this sense, they act as dynamic enablers of process innovation rather than as rigid substitutes for human effort.

2.4.2 Co-pilot Agents and Augmented Decision-Making

A second area of corporate advantage comes from the deployment of co-pilot agents. These extend the concept of intelligent assistants introduced in Section 1.4.2 by embedding proactive orchestration and memory. In practice, co-pilot agents operate as digital colleagues: they manage multi-step tasks, integrate enterprise knowledge bases, and provide contextual suggestions to human workers [15].

For instance, in project management, a co-pilot agent can autonomously track progress across multiple documents, flag inconsistencies, suggest resource reallocations, and even simulate alternative scenarios through predictive models [5] [11]. In knowledge-intensive industries such as consulting or engineering, agents become repositories of organizational memory, capable of retrieving past projects and adapting them to new contexts, thus reducing redundancy and accelerating time-to-solution.

This augmentation resonates with the discussion in Chapter 1.5.1 about the transformation of workforce skills: rather than replacing human professionals, co-pilot agents amplify their cognitive capacity, enabling more strategic focus and reducing cognitive overload.

2.4.3 Task Execution and Operational Autonomy

Beyond automation and assistance, AI agents offer companies the possibility of delegating entire tasks with minimal supervision. This capability builds on the autonomy principle outlined in Section 2.2.4 and is already visible in real-world deployments (see Chapter 2.3). Agents can, for example:

- Execute end-to-end procurement tasks, from supplier identification to preliminary contract drafting.

- Manage customer interactions not only by answering questions but also by completing transactions, offering personalized recommendations, and adjusting policies in real time [13] [14].
- Coordinate IoT devices and robotic systems in logistics or manufacturing, dynamically reallocating resources according to operational data streams [11].

Such task execution potential connects directly to the manufacturing scenarios that will be analyzed in Chapter 3. Here, the integration of agents with industrial IoT and predictive maintenance systems can close the loop between perception, reasoning, and action, unlocking a new generation of adaptive production processes.

Importantly, operational autonomy does not imply a lack of oversight. As emphasized in Chapter 1.6, meaningful human control remains essential for safety, accountability, and trust. In corporate settings, this translates into hybrid workflows where agents handle execution while humans retain responsibility for supervision and high-stakes decisions.

2.4.4 Strategic Implications for Enterprises

From a strategic perspective, the adoption of AI agents brings advantages at multiple levels:

- Efficiency gains: through the reduction of repetitive workload and minimization of human errors in structured and unstructured processes.
- Innovation and agility: by enabling proactive, adaptive solutions that can adjust to dynamic environments, supporting resilience in uncertain markets.
- Workforce transformation: by freeing human capital for higher-value activities, fostering human–AI collaboration, and encouraging reskilling initiatives.
- Integration across the lifecycle: anticipating the role of AI agents in manufacturing (Chapter 3), where their capacity for memory and continuous adaptation aligns with the requirements of Industry 4.0 ecosystems.

2.5 Current risks and limitations

Despite the remarkable progress described in Sections 2.2–2.4, the deployment of AI agents is still constrained by a set of risks and limitations that companies, researchers, and policymakers must carefully consider. These challenges echo with broader issues already highlighted in Chapter 1.6 on ethical, legal, and societal aspects, but acquire new nuances due to the autonomous and adaptive nature of agents.

A first limitation concerns reliability and hallucinations. While large models have already shown a tendency to generate incorrect or fabricated outputs [11] [13], the problem is amplified in agentic contexts. An agent that relies on erroneous information may autonomously trigger incorrect actions or escalate flawed decisions. Unlike static chatbots, agents operate within multi-step workflows (see Section 2.2.1), where a single hallucination can propagate through an entire task chain. This increases operational risks in high-stakes environments such as finance, healthcare, or industrial manufacturing.

Closely related is the issue of control and oversight. As noted in Section 2.2.4, autonomy is a defining feature of agents; however, this autonomy challenges traditional paradigms of accountability. Without robust "human-in-the-loop" mechanisms [6] [7], agents risk executing actions beyond their intended scope, raising questions of liability in case of errors. This is particularly critical when agents interact with IoT devices or robotic systems (see Section 2.3), where decisions are directly translated into physical operations.

Another limitation is scalability and cost-efficiency. Although agents promise resilience and adaptability (see Section 2.4.1), they require continuous orchestration of memory layers, external tools, and reasoning loops [13]. This leads to significant computational overhead compared to static AI services. For SMEs, which already face resource constraints (Chapter 1.5.3), the adoption of agentic architectures may be hindered by infrastructure costs, integration complexity, and subscription models of third-party platforms [15].

Security vulnerabilities also represent a crucial risk. By design, agents rely on external APIs, databases, or connected devices (Table 6). Each interface becomes a potential attack vector, exposing the system to adversarial manipulation, data leakage, or malicious tool injection. Research in adversarial AI has shown how even small perturbations can mislead models [5], and in agentic workflows these vulnerabilities can be compounded by the agent's ability to act autonomously in digital or physical environments.

From a societal perspective, bias and ethical concerns remain unresolved. As emphasized in Chapter 1.6.1, models often inherit biases from training data [6] [11]. When embedded in agents, such biases do not remain confined to single outputs but may drive entire decision-making sequences, potentially reinforcing discriminatory practices in hiring, credit scoring, or customer service. Moreover, the opacity of reasoning chains in agents makes explainability even more challenging than in static models [12].

Another risk is coordination in multi-agent systems. As discussed in Section 2.3, experimental environments with multiple agents show promising emergent behaviors, but also highlight risks of misalignment, cooperation breakdown, and the amplification of systemic errors [13]. These scenarios anticipate challenges in organizational contexts, where multiple agents may need to collaborate across departments or supply chains.

Finally, regulatory uncertainty complicates adoption. The EU AI Act, which introduces obligations for high-risk AI systems [11], does not yet fully account for the distinct dynamics of agentic AI, leaving open questions about compliance, responsibility, and certification. This regulatory gap increases legal risks for companies experimenting with advanced agentic solutions, especially in sensitive sectors. To synthesize these aspects, Table 8 provides an overview of the main risks and limitations of AI agents, structured along technical, organizational, and societal dimensions.

Tabella 8 - Main risks and limitations of AI agents

Dimension	Limitation	Example / Implication
Technical	Hallucinations and error propagation	Incorrect outputs in multi-step workflows, leading to flawed decisions
Technical	Computational cost and resource demands	High computational overhead due to memory and tool orchestration
Organizational	Coordination failures	Emergent misalignment or systemic errors in collective behaviors
Security	Insecure API integrations and external tool dependencies	Risk of adversarial attacks or malicious API/tool manipulation
Organizational	Oversight and accountability	Ambiguity in human–agent control and liability assignment
Societal	Bias and discrimination	Agents reinforcing existing data biases in automated decisions
Technical	Explainability gap	Difficulty in tracing multi-step reasoning and ensuring transparency
Legal	Regulatory uncertainty	The EU AI Act provides a framework for high-risk AI, including agents, but implementation, enforcement, and global alignment remain challenging

In conclusion, while the advantages of agents for companies are undeniable (see Section 2.4), their limitations highlight the need for cautious adoption and robust governance frameworks. Addressing risks such as hallucinations, security vulnerabilities, or ethical concerns is not only a technical requirement but also a strategic imperative. These limitations also serve as a bridge to Chapter 3, where the application of AI agents in manufacturing contexts will be explored. In that domain, ensuring safety, reliability, and compliance becomes even more critical, as digital decisions increasingly interact with physical processes and complex industrial ecosystems.

CHAPTER 3 – ARTIFICIAL INTELLIGENCE IN THE MANUFACTURING CONTEXT

3.1 AI in the Product Lifecycle: an integrated view

The industrial landscape is undergoing a profound transformation driven by the integration of Artificial Intelligence (AI) and Industrial Digital Technologies (IDTs), necessitating a holistic, integrated perspective on the product lifecycle. This integration is critical for addressing the inherent complexity and enhancing efficiency, quality, and sustainability across all phases, from initial design through to eventual end-of-life (EOL). Understanding the comprehensive application of AI requires mapping specific techniques to the unique challenges presented at each stage of the product's existence, ensuring that data generated in one phase provides actionable insights for subsequent phases [21].

The role of AI begins in the Design Phase (Section 3.2), where it extends capabilities beyond traditional Computer-Aided Design (CAD). Here, AI techniques are crucial for facilitating inspiration, supporting concept generation, streamlining shape synthesis, and executing complex Topology Optimization (TO) tasks. Deep generative models and Generative Adversarial Networks (GAN) are prominent in this domain, providing a flexible framework to explore unconventional and complex design spaces. Such approaches enable engineers to consider multiple conflicting objectives, leading to more efficient, functional, and sustainable product designs from the outset [21].

Moving into the Manufacturing Context (Section 3.3), AI becomes indispensable for process control, optimization, and quality assurance. Whether optimizing process parameters in Additive Manufacturing (AM) or conventional machining operations (such as turning and milling), AI/Machine Learning (ML) techniques are employed to predict machinery parameters like cutting forces or surface roughness, and to monitor tool condition. Techniques frequently utilized include Artificial Neural Networks (ANNs), Support Vector Regression (SVR), and various evolutionary algorithms (EAs) (e.g., Genetic Algorithms (GA) and Particle Swarm Optimization (PSO). Moreover, complexity in the supply chain requires intelligent solutions for tasks such as inventory management, where AI heuristic approaches like Monte Carlo Tree Search (MCTS), fuzzy logic, and Reinforcement Learning (RL) help minimize overall inventory costs by managing stochastic factors. The diverse application of AI within this phase is evidenced by the wide array of techniques utilized, spanning image and pattern recognition (CNN, KNN, Multi-Layer Perceptron (MLP)) for quality control, to optimization (GA, PSO, Gaussian Process (GP)) [21].

The Usage Phase (Section 3.4) centers on ensuring continuous operation and equipment reliability through intelligent maintenance strategies. The paradigm shift toward Predictive Maintenance (PdM) is enabled by the availability of vast amounts of data collected throughout the equipment's lifecycle and the integration of advanced diagnostic and prognostic algorithms. PdM utilizes data-driven approaches, relying on AI models trained on historical data to anticipate and diagnose faults in complex components like engines, gearboxes, and turbines. Techniques such as Support Vector Machine (SVM), Artificial Neural Networks (ANN), and Long Short-Term Memory (LSTM) models are essential for fault diagnosis and prognosis (e.g., Remaining Useful Life estimation) respectively, leveraging their strength in classification and sequential data analysis. This phase relies heavily on foundational concepts, such as the Digital Twin (DT), which acts as a crucial system-of-systems by providing a real-time, bi-directional link to the physical asset, enabling adaptive operations and real-time anomaly detection [21] [22].

Finally, the End-of-Life (EOL) Phase (Section 3.5) links back to the overarching goal of sustainability and the circular economy. The successful implementation of reuse, recycling, and remanufacturing strategies depends on AI for generating optimal disassembly sequences (Disassembly Sequence Planning - DSP) and estimating the quality and residual value of components. The challenges posed by complex product geometries and the need for non-destructive disassembly are often tackled using meta-heuristic algorithms like GA and Ant Colony Optimization (ACO), or sophisticated robotic Task and Motion Planning (TAMP). This integrated view confirms that AI methodologies and techniques underpin every stage of the Industrial Equipment Lifecycle [23] [24].

For the development of this chapter, a structured literature analysis was conducted to ensure comprehensive coverage of each stage of the manufacturing process described above. Initially, ten sources were selected from the Scopus database for each phase of the product lifecycle. Subsequently, the focus was refined by prioritizing the most relevant publications, those in which the key topics of interest were most comprehensively addressed. Some figures from these sources were reproduced or adapted to visually communicate the main concepts discussed in this thesis, enhancing clarity and cohesion. Sources that were ultimately deemed less relevant were excluded from the final reference list to maintain conciseness and thematic consistency.

3.2 Design and Development Phase

The design and development phase of the product lifecycle is arguably where the most significant value is created, as decisions made here profoundly impact everything from manufacturing complexity and cost to in-field performance and end-of-life sustainability [21]. Traditionally a human-centric domain driven by creativity and engineering expertise, this phase is now undergoing a profound transformation powered by Artificial Intelligence. Building on the concepts of agentic AI discussed in Chapter 2, intelligent systems are no longer just passive tools but are becoming active partners in innovation, capable of augmenting human creativity, automating complex optimization tasks, and integrating data from the entire product lifecycle to inform better design choices [25].

This section explores the application of AI techniques across the key stages of product design and development. We will examine how AI is being leveraged for concept generation and innovation, detailed design synthesis, and holistic process optimization, ultimately enabling the creation of more efficient, sustainable, and competitive industrial equipment.

3.2.1 AI for Conceptual Design and Innovation

The initial stage of design, often referred to as the "fuzzy front end," is characterized by ambiguity and a high degree of creativity [26]. It is here that AI demonstrates significant potential to augment human ingenuity, primarily by analyzing vast datasets to uncover insights and by generating novel concepts that might elude human designers.

One of the primary challenges in conceptual design is design fixation, a cognitive barrier that limits designers to familiar solutions and hinders innovation. AI algorithms, particularly Machine Learning (ML) and Deep Learning (DL), help overcome this by systematically exploring a vast design space. For instance, Genetic Algorithms (GA) can explore thousands of configurations for a complex system, such as a powertrain, often outperforming human experts in identifying optimal arrangements. Similarly, Artificial Neural Networks (ANN) have been used for product configuration, determining ideal design features based on customer preferences extracted from market data [21].

Furthermore, AI-driven tools can analyze extensive datasets, such as market analyses, customer feedback, patent databases, and academic journals, to provide valuable guidance for creating user-centric and innovative products. Natural Language Processing (NLP) techniques, as referenced in Chapter 1, can be employed to systematically search and analyze textual and visual sources like customer reviews and complaints to extract and categorize explicit and hidden customer needs [21]. For example, a combination of supervised algorithms like fastText and unsupervised algorithms like VADER have been successfully used to identify and classify customer needs from large volumes of unstructured text [21].

More advanced generative models, such as Generative Adversarial Networks (GANs) and Variational Autoencoders (VAEs), are revolutionizing concept generation by creating novel product options from visual or parametric data [21]. By training on existing designs, these models can generate a wide range of new design alternatives for mechanical structures, such as wheels or airfoils, even with limited data input. This generative capability, which will be detailed further in the next sub-section, allows designers to move beyond incremental improvements and explore truly innovative forms and functionalities.

3.2.2 AI in Detailed Design and Optimization

Following the conceptual phase, AI provides powerful tools for detailed design, where specifications are finalized, and performance is optimized. This includes shape synthesis, topology optimization, and the integration of physical properties into digital models.

Shape Synthesis and Generative Design Generative design leverages AI algorithms to autonomously produce a multitude of design options that meet specified constraints [26]. Techniques like Autoencoders (AE/VAE), Deep Convolutional Networks (DCN), and various GAN architectures play a critical role in shape synthesis, offering diverse avenues for generating innovative and functional forms. These models are not just creating aesthetically pleasing shapes; they are optimizing for performance criteria such as weight, material usage, and structural integrity. For example, by integrating human-in-the-loop frameworks, researchers have used electroencephalography (EEG) signals to capture a designer's preferences, which then guide a GAN to generate product concept images that align with unspoken design requirements [21].

Topology Optimization (TO) is a computational method that optimizes material layout within a given design space for a given set of loads and boundary conditions. AI has significantly enhanced this field. Deep Generative Models (DGMs), especially modified GAN variants like Wasserstein GAN (WGAN) and Conditional GAN (CGAN), have become a major focus of research. These models can be trained on datasets generated through traditional TO to learn how to produce designs that are both technically sound and aesthetically superior without direct human intervention. For instance, researchers have used a WGAN to automatically generate automobile wheel designs that possess significant technical value. Other approaches combine VAEs with techniques like style transfer to enhance the quality of generated designs or use Variational Deep Embedding (VaDE) to identify and fill unexplored regions in the design space, thereby expanding the range of feasible solutions [21].

A key advancement is the integration of physical properties into the generative process. Some models use spatial fields of physical parameters, such as von Mises stress, displacement, and strain energy density, as inputs to a CGAN, training the network to generate topologies that are inherently compliant with physical laws. This results in more realistic and manufacturable designs.

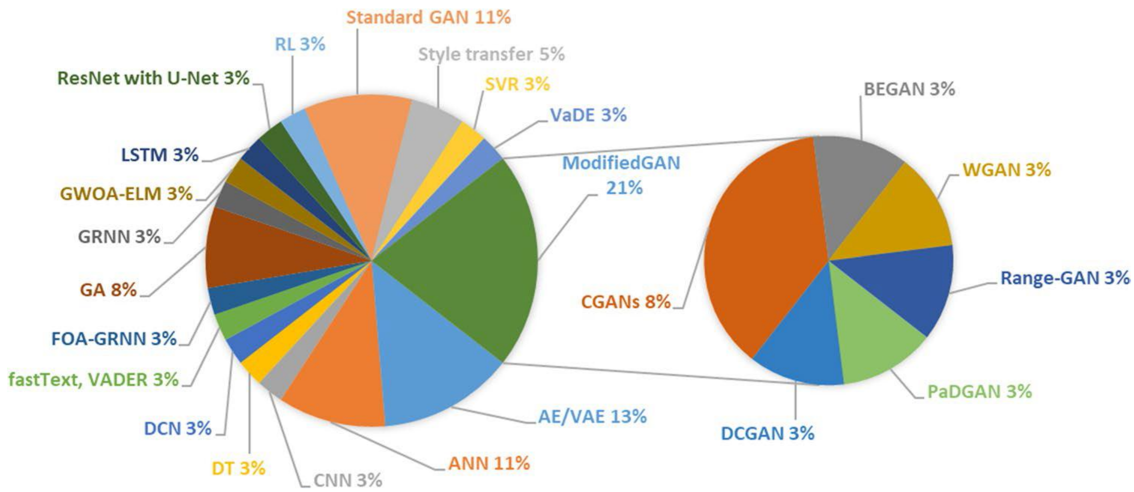

Figure 5 - AI techniques used at product design phase

The chart illustrates the relative frequency of AI techniques applied in product design. The larger pie (79%) groups the most widely adopted methods (e.g., VAE, GAN, Autoencoders, DCNs), while the smaller segment (21%) highlights modified GAN architectures with enhanced loss functions, representing emerging approaches [21].

3.2.3 Integrating Data and Models for a Holistic Design Process

The true power of AI in the design phase is realized when it is integrated into a broader data ecosystem, creating a continuous feedback loop throughout the product lifecycle. Concepts like the Internet of Production (IoP) and Digital Shadows are central to this vision. The IoP proposes a globally interconnected system where data from all phases of the product cycle, development, production, and usage, is shared to fuel innovation [27].

A Digital Shadow is a purpose-driven, aggregated dataset that includes not only raw data but also models and metadata providing context about the product or process [19]. In the design phase, AI agents, as discussed in Chapter 2, can leverage Digital Shadows to make data-driven decisions. For example, data collected from sensors on in-field equipment (IoT) can be fed back to the design stage, allowing AI models to identify performance bottlenecks or common failure modes. This information can then be used to inform redesigns or create the next generation of products [28].

This approach enables what is known as model-integrated AI, where data-driven AI methods are combined with domain-specific engineering models (e.g., CAD, Finite Element Analysis) [27]. AI can analyze simulation results or real-world performance data to automatically refine design parameters, test new material compositions, or optimize for sustainability criteria such as recyclability and energy consumption [29] [28]. This creates an AI-driven design loop where products are continuously improved based on real-world evidence, moving beyond static, one-off design processes [27].

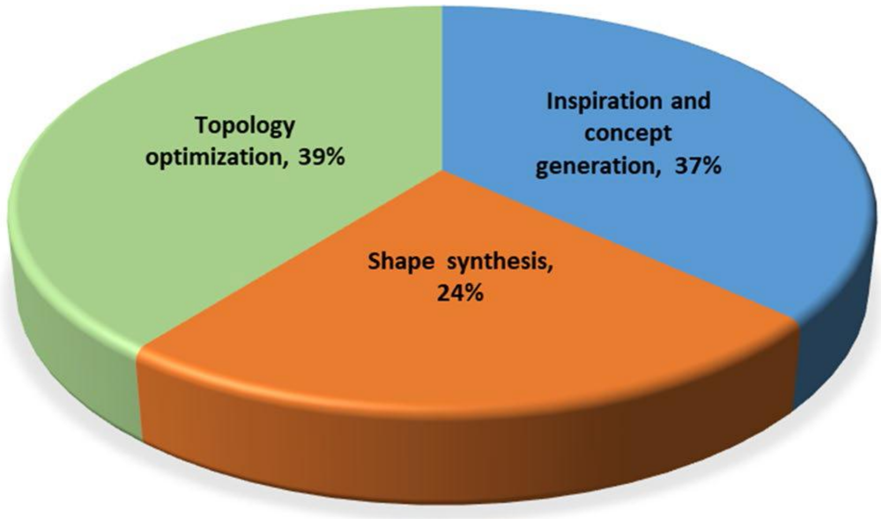

Figure 6 - AI techniques and their prevalence in different design stages

The diagram compares AI techniques across different stages of product design, from conceptual exploration to detailed optimization. It highlights, for example, the dominance of GANs in generative concept design and the growing role of Digital Shadows and model-integrated AI in holistic optimization.

3.2.4 Case Studies and Best Practices

The application of AI in the design and development phase is not merely theoretical; numerous case studies demonstrate its practical value across various industries.

In the **automotive sector**, generative AI is used for component design optimization. As mentioned, Wasserstein GANs (WGANs) have been successfully used to automatically generate aesthetically superior and technically valuable wheel designs without human intervention. Other projects use Conditional GANs (CGANs) combined with human input to interactively explore and refine structural designs, merging the designer's expertise with the generative power of AI. These approaches lead to lighter, stronger, and more efficient components [21].

In the **aerospace industry** AI models like GANs and Conditional Variational Autoencoders (cVAEs) have been employed to generate a wide array of design options for complex mechanical structures such as airfoils, even with limited

initial design data [1]. This accelerates the exploration of novel aerodynamic profiles, which is critical for improving aircraft performance and fuel efficiency.
In **manufacturing**, AI contributes to the ‘Design for X’ paradigm, where X may represent manufacturability, assembly, or disassembly. For example, AI algorithms like Genetic Algorithms (GA) and Ant Colony Optimization (ACO) are used to generate optimal disassembly sequences for End-of-Life (EoL) products [23]. This information is crucial during the design phase to create products that are easier to repair, remanufacture, or recycle, aligning with circular-economy principles [28]. The Internet of Production (IoP) framework enables a "Worldwide Lab" where a company can retrieve machine parameters from a different production site to adapt its design in response to material variations, demonstrating a best practice in cross-domain collaboration [27].
While outside traditional manufacturing, the **pharmaceutical industry** provides a powerful analogy. Here, AI models, including Graph Neural Networks (GNNs), are used for new drug designs by modeling molecular structures and predicting their properties [30]. This parallels how AI is used in manufacturing to design novel materials or components with specific functional characteristics.

From these cases, several best practices emerge for integrating AI into the design and development phase:

- Integrate Human-in-the-Loop: combine AI's computational power with human expertise and intuition. Interactive frameworks where designers guide or refine AI-generated outputs lead to more practical and preferred solutions [21].
- Create a Data Feedback Loop: establish systems (like Digital Shadows) to collect and feed data from the entire product lifecycle back into the design process. This ensures that designs are continuously improved based on real-world performance and usage patterns [21].
- Design for the Full Lifecycle: employ AI to optimize not just for performance but also for sustainability, manufacturability, and end-of-life considerations. AI-driven "Design for Disassembly" is a key enabler of the circular economy [23] [28].
- Adopt Model-Integrated AI: instead of treating AI as a "black box," integrate it with established engineering models (CAD, FEA). This ensures that AI-generated designs are grounded in physical principles and are more reliable and trustworthy [21] [27].

By adopting these practices, manufacturing companies can harness the full potential of AI to create superior products more efficiently and sustainably, paving the way for an AI-driven product lifecycle.

3.3 Production and Operations Phase

Building upon the digitally conceived and optimized product designs from the previous phase, the Production and Operations phase is where the product is physically realized, and the supporting ecosystem is managed. The integration of Artificial Intelligence in this stage marks a pivotal shift from traditional, rigid production lines to smart, adaptive, and data-driven manufacturing environments, a concept at the core of Industry 4.0. This transformation is not merely about automation but involves creating a deeply interconnected system where data flows seamlessly from machines, processes, and supply chains to inform real-time decision-making. This vision aligns perfectly with the Internet of Production (IoP) framework introduced in section 3.2, where Digital Shadows of production processes provide the data necessary for AI agents to optimize operations dynamically [21] [27].

AI technologies are applied across a spectrum of activities within this phase, including core manufacturing processes, quality control, predictive maintenance, and the orchestration of the broader supply chain.

3.3.1 AI in Core Manufacturing Processes

The manufacturing environment is characterized by a variety of processes and tasks, including production planning, procurement, and the core production operations such as machining and assembly. The integration of AI has brought about a transformative shift, particularly by enabling the intelligent integration of subtractive manufacturing (SM) and additive manufacturing (AM) processes. AI algorithms analyze design constraints, material properties, and production requirements to determine the optimal combination of these heterogeneous processes, leading to improved product quality and increased speed [24].

In **Additive Manufacturing (AM)**, commonly known as 3D printing, AI/ML methodologies are fundamentally reshaping optimization techniques. Given the advantages AM offers, such as creating components with intricate geometries and minimizing material waste, the focus of AI application is centered on controlling the variability inherent in the layer-by-layer process. AI techniques are primarily utilized for process and parameter optimization, which involves regulating variables like laser power and scanning speed to achieve precise geometric specifications, such as a particular melt pool geometry or bead dimensions. In fact, among the various manufacturing stages, Process Parameter Optimization commands the most significant share of AI technique utilization, standing at 28%. This distribution, emphasizing the importance of fine-tuning operational inputs, is adeptly visualized in a pie of the pie-chart illustrating the distribution of AI techniques among manufacturing stages, which highlights the prominent reliance on AI techniques for optimizing critical manufacturing parameters. Techniques like Artificial Neural Networks (ANN) are applied in processes such as Directed Energy Deposition (DED) to accurately identify the necessary process parameters to achieve desired outcomes [21].

Conversely, traditional **Subtractive Manufacturing (SM)** processes, which historically faced issues related to process control and inefficiency, are being revolutionized by AI to achieve smart machining. This involves the capability to autonomously adapt machining parameters during operation to achieve optimal results. While MLR provides useful statistical baselines, AI-driven methods such as SVR extend predictive power for machining optimization in conventional processes like turning. For more intricate operations such as boring, Support Vector Machines (SVM) analyze vibration signals extracted via discrete wavelet transforms, classifying the operational state (stable, transition, or chatter) to mitigate chatter-related issues and ensure superior surface finish quality [21].

3.3.2 In-Process Monitoring and Quality Control

Ensuring product quality is a continuous process that AI enhances through real-time monitoring and post-production inspection.

In-Process Monitoring: during manufacturing, especially in AM, AI models analyze data from various sensors (e.g., infrared cameras, acoustic sensors) to detect anomalies in real-time. For instance, unsupervised machine learning techniques can be used on plume images from an infrared camera to automatically detect unstable melt pools during the SLM process. Similarly, models like Deep Belief Networks (DBN) and Convolutional Neural Networks (CNN) can classify different melting states based on acoustic signals, identifying defects such as balling or cracking as they occur. This real-time feedback loop, a practical implementation of the IoP's Digital Shadow, allows for immediate corrective actions, reducing waste and improving yield [21].

Post-Production Inspection and Defect Detection: after production, AI-powered computer vision systems automate the inspection process. CNN-based systems are particularly effective for the visual inspection of defects in products like solar panels or nanoscale transistors. These systems can analyze electroluminescence (EL) images or post-build Computed Tomography (CT) scans to detect flaws with remarkable accuracy, even with limited training data. This automated approach not only accelerates quality control but also enhances its reliability compared to manual inspection [21].

3.3.3 Predictive Maintenance (PdM)

A critical component of the operations phase is ensuring the reliability and uptime of manufacturing equipment. Predictive Maintenance (PdM) leverages AI to shift from a reactive or scheduled maintenance model to a proactive, predictive one. The core objective of PdM is to forecast the Remaining Useful Life (RUL) of a component or machine, allowing maintenance to be scheduled precisely when needed, thus minimizing downtime and costs.

Data-driven techniques, including various machine learning models, are employed to analyze historical data, sensor readings (e.g., vibration, temperature, acoustic emissions), and operating conditions to predict the RUL of critical components. For example, Long Short-Term Memory (LSTM) networks, a type of RNN, are highly effective for handling time-series data from sensors to predict failures in complex systems like airplane turbofan engines. Other commonly used techniques include ANNs, SVMs, Random Forests (RF), and Deep Belief Networks (DBN), which have been applied to diagnose faults and predict the RUL of components in engines, bearings, gears, and EV batteries. Hybrid physics-informed ML models are emerging as best practices in predictive maintenance, complementing purely data-driven approaches [21].

A key advancement in this area is the integration of eXplainable AI (XAI), which addresses the "black box" nature of many complex models. By providing human-intelligible narratives and justifications for its predictions, XAI empowers maintenance technicians to understand why a certain component is predicted to fail, thereby increasing trust and enabling more effective human-in-the-loop decision-making. This fosters a more sustainable and collaborative maintenance environment where human expertise is augmented, not replaced, by AI [21].

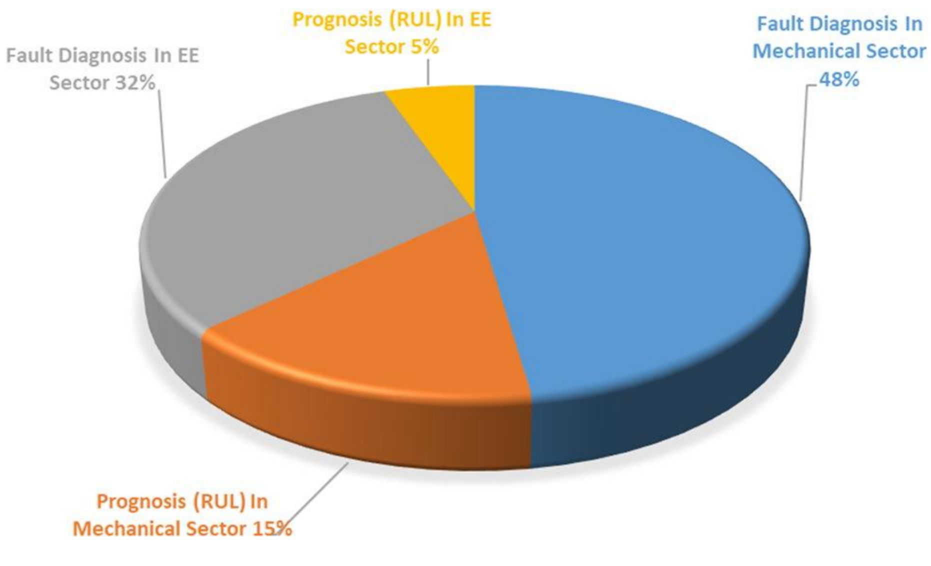

Figure 7 - Incorporation of AI for fault diagnosis and prognosis in mechanical and EE sector

Figure 7 provides a comprehensive overview of how different AI techniques are incorporated for fault diagnosis and prognosis across both mechanical and electrical/electronic (EE) sectors, illustrating the breadth of PdM applications [21].

3.3.4 Supply Chain and Logistics Optimization

The production phase is intrinsically linked to the broader supply chain. AI optimizes these interconnected processes by enhancing decision-making, inventory control, and logistics. Reinforcement Learning (RL), a concept related to the agentic behaviors discussed in Chapter 2, is particularly effective for dynamic optimization problems. For instance, RL models are used to minimize overall inventory costs by learning optimal ordering policies in complex supply chain systems [31] [32].

In logistics, AI techniques contribute to solving complex planning problems. Decision Support Systems (DSS) and heuristics are used to optimize container assignment and loading problems, while AI combined with technologies like Radio Frequency Identification (RFID) enhances the responsiveness and traceability of the logistics workflow [21].

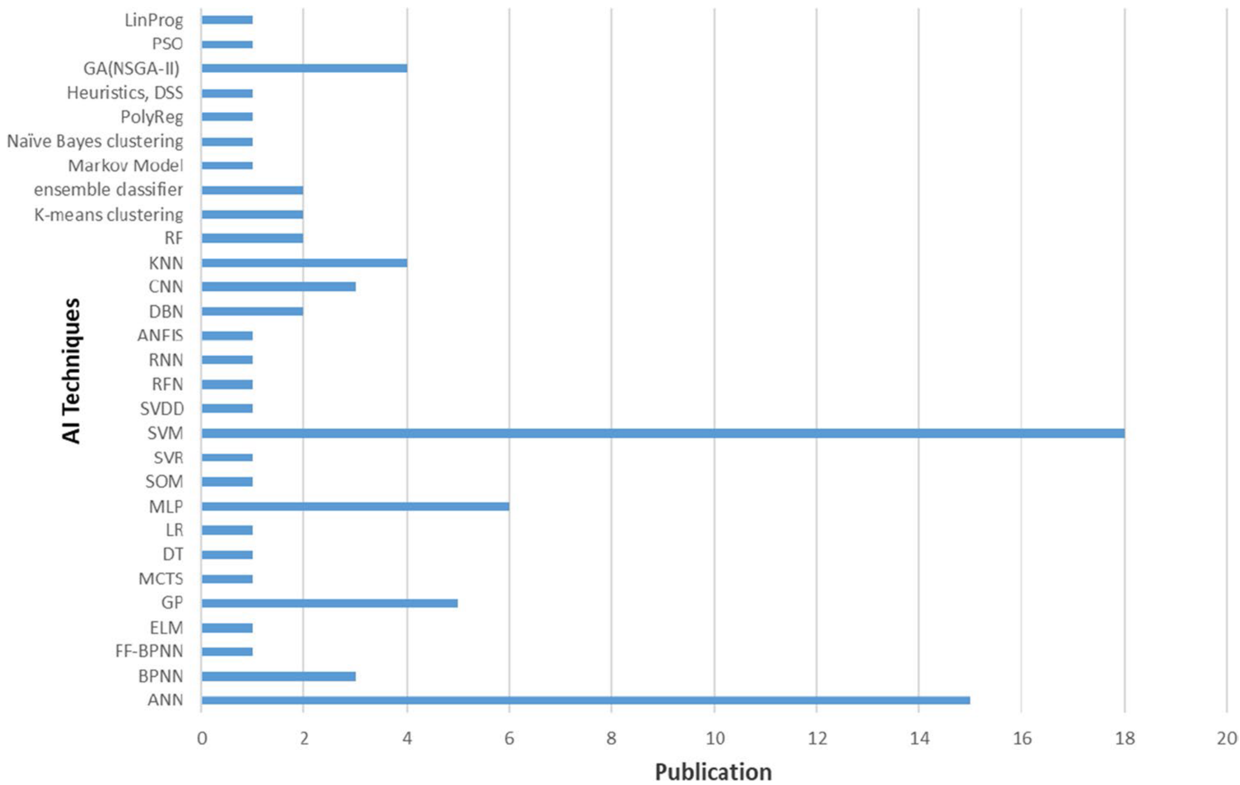

Figure 8 - Popularity of different AI techniques in manufacturing phase

The prevalence and importance of various AI techniques in revolutionizing the manufacturing landscape are visually summarized in the bar chart in Figure 8, which showcases the distribution of 29 distinct AI techniques across numerous publications. Noteworthy is the widespread use of traditional models like ANNs and SVMs, which appear in 15 and 18 publications respectively within the reviewed literature.

By integrating AI into these core operational areas, manufacturing companies can move towards the adaptive, resilient, and efficient production systems envisioned by Industry 4.0. This data-driven ecosystem sets the stage for the next phase of the product lifecycle, where data collection continues and provides further opportunities for AI-driven services and insights.

3.3.5 Case Studies and Best Practices

The application of AI in the production and operations phase extends across a wide range of industries, where it consistently demonstrates best practices in enhancing efficiency, quality, and sustainability.

In the **automotive industry**, AI plays a dual role in production optimization and quality control. Support Vector Machine (SVM) and Naive Bayes (NB) classifiers have been successfully employed to analyze acoustic data from automobile gearboxes, enabling accurate defect diagnosis. Predictive maintenance is equally relevant: Random Forest classifiers have been applied to vehicle maintenance records and operational data, allowing early detection of failures such as air compressor malfunctions in trucks and buses. This proactive approach has improved both scheduling efficiency and fleet reliability [21].

The **aerospace sector** provides high-stakes examples of AI in predictive maintenance. A well-established best practice is the use of Long Short-Term Memory (LSTM) networks to estimate the Remaining Useful Life (RUL) of aircraft turbofan engines. By leveraging multi-sensor time-series data, including the NASA C-MAPSS dataset, LSTM models can accurately anticipate engine degradation and optimize maintenance schedules, thereby ensuring operational safety and continuity.

In the **electronics and semiconductor industry**, AI has become a cornerstone of defect detection and quality assurance. Convolutional Neural Networks (CNNs), for example, are employed to inspect solar panels by analyzing Electroluminescence (EL) images, effectively identifying micro-cracks and other flaws that would escape manual inspection. Similarly, in the semiconductor industry, machine learning models accelerate failure analysis of nanoscale transistors, reducing downtime and improving throughput in production lines.

The **pharmaceutical and healthcare sector** applies AI to optimize both production processes and final product performance. Artificial Neural Networks (ANNs) combined with Monte Carlo simulations are used to refine drug formulations and predict release kinetics, ensuring that dosage forms meet strict performance criteria. On the manufacturing side, AI-driven predictive maintenance systems, such as deep transformer models, are deployed to minimize downtime in production lines, reinforcing consistency and cost efficiency.

Finally, the **agriculture sector (Agriculture 4.0)** provides a particularly innovative field of application. Unmanned Aerial Vehicles (UAVs), equipped with multispectral or hyperspectral cameras, generate high-resolution imagery of crops. Computer vision algorithms analyze this data to detect pests, diseases, and nutrient deficiencies. The outcome is precision agriculture: resources such as water, fertilizers, and pesticides are applied only where necessary, improving sustainability and boosting yields. As illustrated in Figure 9, UAV-based AI research spans diverse agronomic practices, with significant attention devoted to yield estimation, crop mapping, and targeted agrochemical application.

Across these industries, a unifying theme emerges: the most successful implementations of AI are those that integrate domain-specific expertise with advanced data-driven models. Rather than serving merely as automation tools, AI systems enhance human decision-making, optimize resource allocation, and build more resilient and sustainable operational ecosystems.

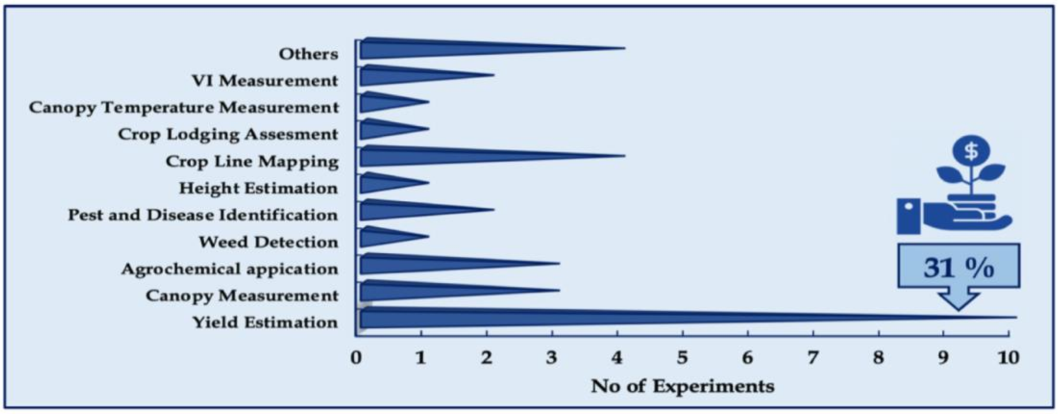

Figure 9 - Number of UAV research and publications in sugarcane crops conducted for different agronomic practices

As illustrated in Figure 9, the application of UAVs and AI in agriculture is diverse, with a significant portion of research focused on yield estimation, crop line mapping, and agrochemical application, showcasing the technology's versatility in optimizing farm operations.

These case studies illustrate a common theme: the successful application of AI in production and operations relies on the integration of domain-specific knowledge with data-driven models. The best practices consistently involve using AI not just to automate tasks, but to augment human decision-making, improve resource efficiency, and create more resilient and sustainable operational systems.

3.4 Usage Phase: In-field Operations and Customer Experience

The transition from the Production and Operations Phase (3.3), where the physical asset is manufactured and initially tested, to the Usage Phase marks a critical shift from internal optimization to continuous in-field validation and customer interaction. The operational reliability and efficiency achieved through predictive maintenance models during production must be sustained and evaluated in real-world scenarios, leveraging technologies that monitor and adapt the product or service throughout its lifespan [22]. This necessitates the sophisticated handling and classification of generated usage data and integrating complex human factors into the technological framework.

3.4.1 Data Characteristics and Classification for In-Field Analytics

A successful usage phase hinges on deriving high-quality insights from field data to support ongoing product planning. Unlike the manufacturing environment, usage data is highly heterogeneous and context-dependent [29] [33]. The research identifies five key categories of use phase data from a product planning view: usage data (describing how the product is used), user behavior data (summarizing user interaction), service data (dealing with problems and quality), product behavior data (showing performance during operation), and status data (describing product "health").

The complexity and variety of this data require specialized classification, moving beyond simply grouping data sources towards analytically relevant characteristics. Data may present as real-time time series (such as sensor data, control signals, hardware states) or as text data (including warnings, complaints, ratings, and login information). The effectiveness of subsequent analytics techniques relies heavily on this smart classification, which informs preprocessing and algorithm selection. These methodologies define data based on factors such as volume, velocity, complexity, dimensionality, and distribution, allowing industrial stakeholders to match usage data to generalized analytics workflows [33].

To illustrate the necessary structure and rigor applied to managing this vast array of field data, it is informative to review the formal process for establishing data classification artifacts. Figure 10 provides a visual representation of how empirical and conceptual approaches are iteratively combined to define data characteristics and dimensions that are analytically relevant for subsequent processing [33].

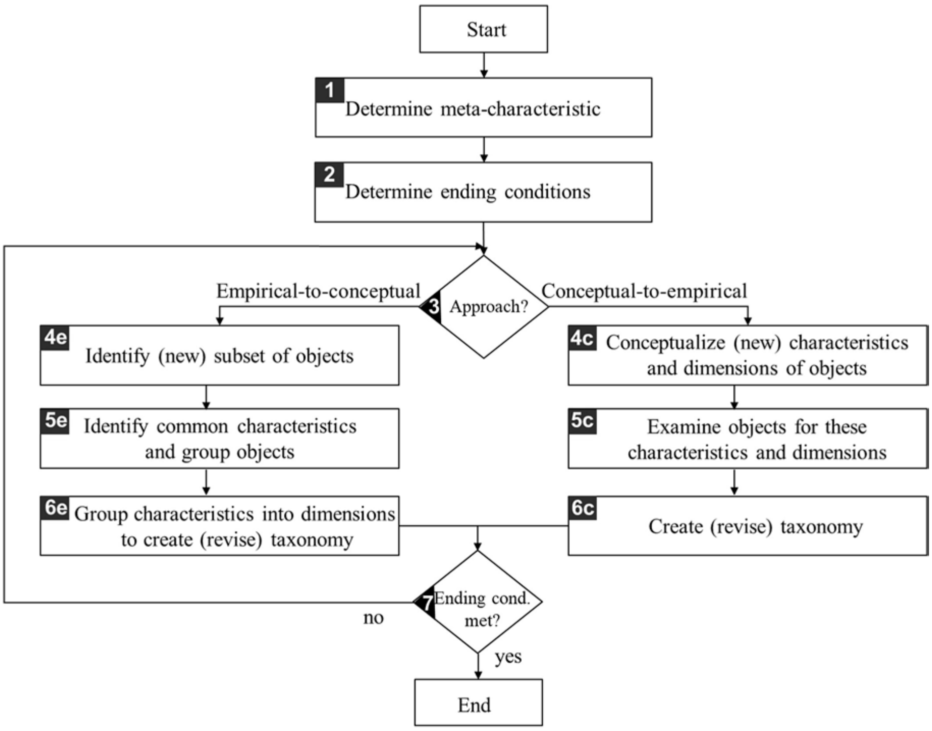

Figure 10 - the Taxonomy development method according to Nickerson et al. (2013)

3.4.2 Leveraging Digital Twins for Adaptive Operations

The Digital Twin (DT) paradigm acts as a foundational system for the Usage Phase, extending the connectivity established during the Production Phase (3.3) into continuous operation. The DT is conceptualized as a system of systems, built by melding several enabling technologies to create an intelligent virtual representation of a physical entity. Unlike Digital Shadows, which aggregate historical data, Digital Twins enable real-time synchronization with the physical asset. The DT framework is characterized by three core functional blocks: the physical asset, its virtual counterpart, and the two-way communication medium that binds them together in a symbiotic relationship. This connection is typically enabled by the Industrial Internet of Things (IIoT), facilitating the real-time exchange of the heterogeneous usage data described previously [22].

DTs are crucial because they support various services and applications dependent on the specific use case, ranging from intelligent prognostics for predicting remaining useful life to continuous lifetime monitoring. Given that the DT facilitates the continuous flow of data to and from the real asset, rigorous data security measures (including privacy, authentication, integrity, and traceability) are paramount [8]. The architectural models, such as Tao's 5D model, emphasize how DT implementation details must be highly dependent on the use-case requirements. Tao's 5D model encompasses physical entities, virtual models, services, connections, and data, which together define the full structure of a Digital Twin system.

3.4.3 Enhancing Customer Experience and Human-Machine Interaction

In the Usage Phase, the focus shifts from the reliability of the industrial equipment to the seamless interaction between the user and the smart product, aligning with the principles established in the product's initial design (section 3.2). User Experience (UX) is defined by the feelings and perceptions users derive from interacting with a system, aiming to maximize usability, usefulness, and desirability [34]. AI-enabled systems introduce adaptability, which requires clear communication to the user to ensure transparency and trust, especially in complex domains like autonomous vehicles (AVs) [29].

In vehicles, AI is employed in interactive explanation concepts for adaptive systems, such as navigation, driving modes, and well-being features. These systems propose adaptations based on learned user preferences and the current context. As automation levels increase, the traditional definition of the human component broadens from just the driver to include passengers, pedestrians, and cyclists, necessitating a holistic approach to UX evaluation and HMI design. UX evaluation systems are crucial for measuring quality in a systematic way, considering factors such as the user's internal state (e.g., mood, motivation), system characteristics (e.g., complexity), and the environmental context. Ultimately, the goal is to develop predictive models, informed by continuous field data, that can anticipate user attributes, behavior, and demands, forming a basis for personalized product functions and service recommendations [29].

3.4.5 Case Studies and Best Practices

The integration of AI, IoT, and DTs in the Usage Phase demonstrates best practices in achieving operational longevity, efficiency, and superior customer satisfaction, representing the practical culmination of processes defined in preceding chapters (3.1, 3.2, 3.3).

One significant area is smart retrofitting, which extends the life of industrial equipment, complementing the sustainability goals addressed in later phases (3.5An Industry 4.0-oriented approach uses machine-learning models, such as a two-layer neural network (NN), to estimate thermal characteristics of injection molding machines, enabling adaptive control and improved energy efficiency. A comparison of various AI techniques employed in this domain shows that Artificial Neural Networks (ANN) have the highest adoption rate, featuring in four relevant publications, confirming ANN's effectiveness in leveraging operational data for targeted system upgrades [21].

In the transportation sector, the implementation of Big Data Analytics (BDA) and AI in ridesharing platforms reveals the necessity of balancing technological efficiency with user perception. The resulting research model uses the concepts of perceived benefits (utility/usefulness) and perceived risks (uncertainty and invasion of privacy) to explain passenger participation in ridesharing. This emphasizes that successful in-field operations must move beyond technical efficacy to address the cognitive and social dimensions of technology acceptance [35].

A further best practice involves the shift towards service-centric business models, such as Component-as-a-Service (CaaS) for Electrical and Electronic Equipment (EEE). This model inherently improves product circularity and sustainability. Future research recommendations highlight that integrating data management strategies, including IoT, digital twins, and data analytics, directly into the design guidelines (3.2) is necessary to guide industry stakeholders in implementing CaaS effectively [36].

In complex manufacturing scenarios, such as the tea industry, a case study demonstrates the use of Digital Twins in monitoring semi-automated systems that involve both machines and human operators. This DT tracks several sequential steps from blending to packaging, providing crucial oversight of activities in the conveyor belt. Observing the snapshots of different activities in the tea manufacturing conveyor belt process, such as adding tea and herbs, dosage, blending, and packaging, reveals the intricate details necessary for operational oversight.

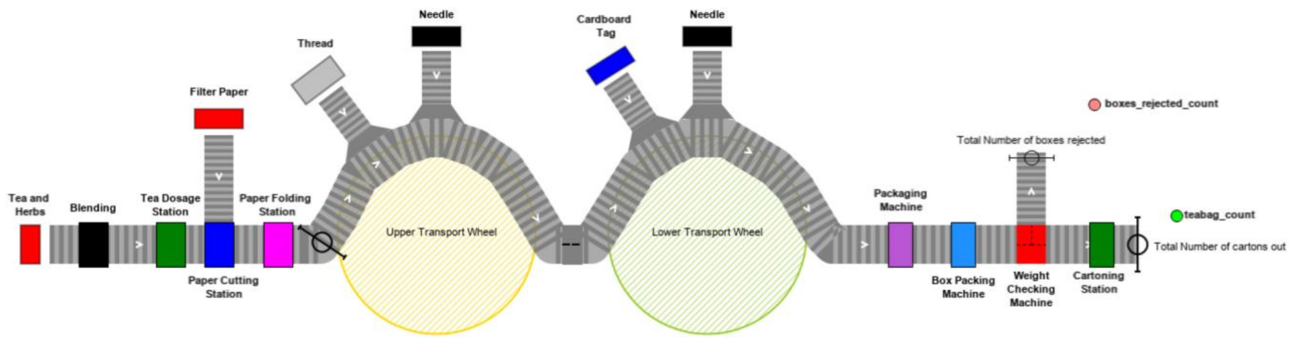

Figure 11 - Snapshot of the different activities in the conveyor belt of the tea manufacturing company

Figure 11 schematically represents the Digital Twin architecture, highlighting the continuous bidirectional data flow between the physical asset and its virtual counterpart through an IIoT communication layer [22].

3.5 End-of-Life and Sustainability

The End-of-Life (EoL) phase represents the culmination of the product lifecycle within the manufacturing ecosystem, focusing on essential circular strategies such as reuse, recycling, remanufacturing, and smart retrofitting. The successful implementation of these strategies is critically dependent upon the effective utilization of AI and Industrial Digital Technologies (IDTs) to achieve both resource efficiency and alignment with circular economy (CE) principles [36]. This strategic shift requires linking information and capabilities generated during the design (3.1), manufacturing (3.3), and usage (3.4) phases to inform efficient EoL decision-making. Specifically, the extensive collection and classification of Usage Phase data (as detailed in Section 3.4) becomes invaluable for accurately diagnosing the status and viability of components for recovery processes [33] [37].

Contemporary literature highlights IDTs as crucial enablers for managing the inherent complexity of Sustainable Supply Chain Management (SSCM) within a circular economy framework. These technologies enhance both the physical and cyber capabilities of the supply chain, facilitating closed-loop lifecycle management and ensuring that operations adhere to sustainable requirements. Comprehensive theoretical models, such as the CAB2IN framework, integrate key emerging technologies like Cloud Service (CS), AI, Big Data Analytics (BDA), Blockchain Technology (BT), and the Internet of Things (IoT) to ensure supply chain information is efficiently transformed into actionable knowledge. The holistic nature of such frameworks allows stakeholders to enhance their capability to resist risks and sustain long-term profitability while verifying that the entire chain operates without environmental or societal damage through precise and traceable decision-making [28].

3.5.1 AI for Complex Component Management and Diagnostics

The challenges faced during the EoL phase are particularly pronounced for complex or high-value components, such as Lithium-Ion Batteries (LIBs) used in Electric Vehicles (EVs). The core processes for battery recovery, including remanufacturing, necessitate accurate diagnostic and screening procedures to assess component viability. This aligns conceptually with the predictive modeling introduced in Section 3.4, extending the focus from operational prediction to EoL evaluation [37].

The literature identifies key research streams concerning the sustainable management of EV battery remanufacturing, specifically mentioning disassembly procedures, diagnostics and screening, data sharing, and supply chain design. The advancement of these procedures is inherently linked to technological and digitalization progress. For example, AI/Machine Learning (ML) techniques are essential for diagnostics and screening, determining the current state and recovery potential of the asset. Concurrently, data sharing and transparency are critical components of a circular strategy, although they introduce security concerns and technological limitations that remain active areas of research. The rigorous review protocol used in these highly specialized studies highlights the meticulous approach needed to consolidate knowledge in this emerging domain [37]. To ensure a rigorous and evidence-based discussion on AI applications in end-of-life battery management, a structured literature review methodology was adopted. Figure 12 illustrates the article selection process. Starting from an initial pool of identified studies, successive screening phases based on titles, abstracts, and full texts were applied to select only those works directly relevant to AI-driven diagnostics, disassembly, and remanufacturing of Lithium-Ion Batteries (LIBs). This transparent selection process ensures the quality and focus of the sources underpinning the analysis in this section.

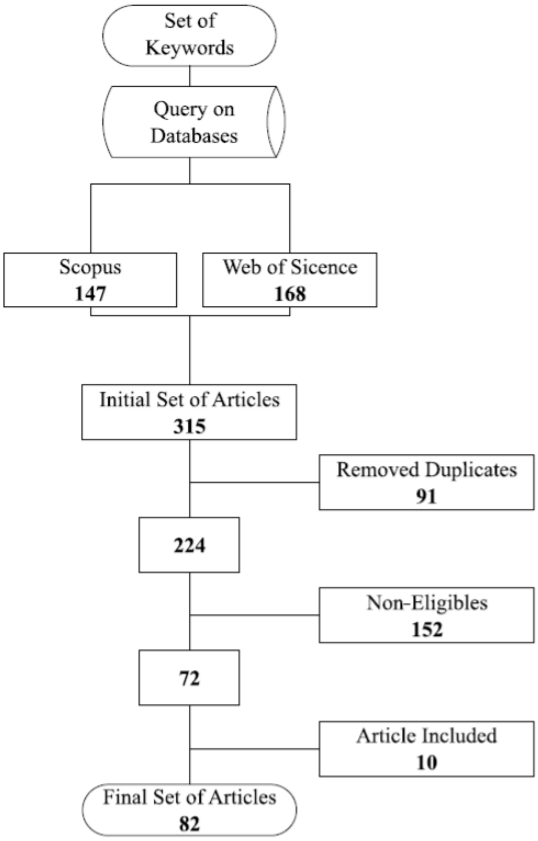

Figure 12 - Selection of articles

3.5.2 Smart Retrofitting and Predictive Integration

The concept of smart retrofitting directly addresses the extension of product life, bridging the Usage Phase (3.4) with EoL considerations. This practice involves incorporating AI and ML assistance, often within Digital Twin (DT) frameworks, to enhance monitoring and anomaly detection for predictive maintenance, thereby transforming older assets into connected systems [21].

While detailed information on specific AI/ML techniques used in general retrofitting frameworks can sometimes be lacking in the literature, certain applications clearly mention the deployment of Convolutional Neural Networks (CNN), Artificial Neural Networks (ANN), and their variants. These techniques are often employed in applications that are tightly coupled with the DT (as introduced in 3.4.2) or even the more advanced Digital Triplet systems.

Furthermore, ML techniques, specifically two-layer neural networks, have been used in thermal design methodologies within smart retrofitting applications to predict heat transfer parameters and support lumped parameter simulations for embedded components. In industrial settings, frameworks focusing on Industry 4.0 paradigms for retrofitting old process plants have been proposed, utilizing supervised machine learning algorithms for anomaly detection to improve safety and maintainability [21].

The effective execution of both diagnostics and retrofitting procedures relies on a holistic view of the component's characteristics and the evolution of research priorities across the lifecycle. The maturation of EoL topics often follows the initial breakthroughs in design and core technical development.

3.5.3 Optimization and Environmental Sustainability

The goal of sustainability extends beyond prolonging asset life to actively mitigate the negative environmental impact associated with industrial operations and the computational demands of AI itself, a concept often termed Green AI. Large AI models, particularly in Machine Learning (ML), are known for their high energy demand during both training and inference, leading to proportional carbon dioxide emissions [36] [38].

In response to this challenge, strategies focused on reducing the computational footprint are gaining traction. A key tactic is the use of lightweight AI models, often achieved through techniques like model distillation. These models facilitate the shift toward sustainable energy by remaining consistent with circular economy principles. This approach seeks to define a holistic understanding of ML sustainability, balancing trade-offs related to energy efficiency, particularly between computational accuracy, model complexity, and energy consumption [36] [39].

Furthermore, the environmental pillar of sustainability can be formally assessed using advanced analytical methods such as Retrospective Life Cycle Assessment (LCA). LCA processes, which calculate the potential environmental impacts of a product across its lifespan, are becoming integrated with AI techniques, particularly in complex material domains like composite manufacturing. Although the primary focus of this assessment is often environmental, it exerts indirect influence on the economic and social pillars of sustainability. Achieving widespread success in deploying resource-conscious AI solutions is often contingent on the establishment of appropriate managerial and policy frameworks that incentivize carbon footprint reductions in digital operations [39] [40].

3.5.4 Case Studies and Best Practices

The application of AI and IDTs in the EoL phase demonstrates several best practices, particularly regarding complex logistical and automated tasks.

Robotic Disassembly and Automation A crucial aspect of materials recovery and component remanufacturing is the efficient disassembly of EoL products, which is increasingly managed through advanced robotics and AI. Disassembly processes, classified as stream in specialized literature, are fundamentally determined by the initial product design and evolve alongside digitalization. This highly technical area focuses on methodological and application advancements in robotic disassembly technology, distinct from purely economic or non-technical evaluations [24] [37].

Intelligent Maintenance and Circular Systems Best practices in circular systems often involve the creation of intelligent support tools for maintenance technicians, integrating human-centric Industry 5.0 concepts. For instance, frameworks exist that combine Discrete Event Simulation (DES), text mining, and AI to integrate disparate IoT-generated data sources and information systems. The Auto-Circular Simulator concept exemplifies this approach by providing semantic descriptions to identify and highlight linkages between data and knowledge for the worker, enabling sustainable treatment of automotive components. Context awareness is vital for the continued development of explainable AI (XAI) in this domain, providing a structure for communicating requisite narratives that support the technician. As shown in Figure 13, the simulator exemplifies how explainable and context-aware AI supports maintenance technicians in interpreting system outputs and making informed, sustainable operational choices [41].

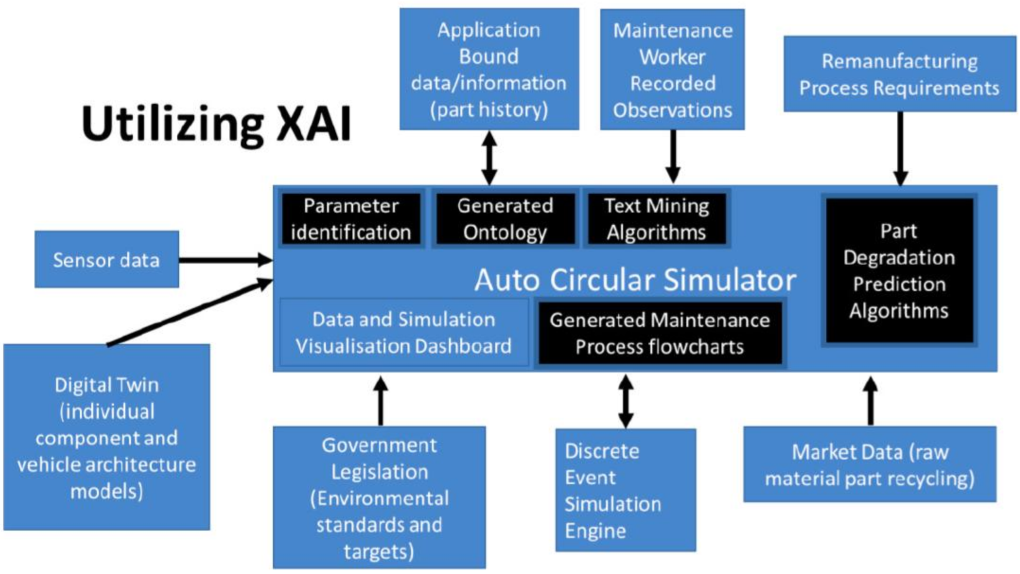

Figure 13 - Auto Circular Simulator Utilization of XAI: Adapted from Turner et al. (2022)

3.6 Benefits and Challenges for Manufacturing Companies

The progressive integration of Artificial Intelligence (AI) across all stages of the product lifecycle, from design to end-of-life, has transformed the manufacturing sector into a data-driven, adaptive, and highly interconnected ecosystem. While these technological advances offer unprecedented opportunities for efficiency, innovation, and sustainability, they also introduce new layers of complexity in data management, human–machine collaboration, and organizational governance. This section outlines the main benefits and challenges that manufacturing companies face in adopting and scaling AI-driven solutions.

Benefits

AI significantly enhances operational efficiency by automating repetitive tasks, optimizing production parameters, and reducing waste. Advanced control systems based on machine learning, such as Artificial Neural Networks (ANNs) in Additive Manufacturing or Support Vector Machines (SVMs) for chatter detection in machining, adjust process conditions in real time, ensuring higher throughput and consistent product quality. Predictive analytics enable proactive decision-making, while intelligent scheduling and robotics improve resource allocation and reduce unplanned downtime, collectively contributing to leaner operations and lower manufacturing costs.

Furthermore, AI enables a shift from reactive or scheduled maintenance to condition-based and predictive maintenance (PdM). By leveraging sensor data and models such as Long Short-Term Memory (LSTM) networks or Random Forests, companies can accurately forecast the Remaining Useful Life (RUL) of critical components, thereby minimizing unexpected failures, extending equipment lifespan, and optimizing maintenance budgets. This approach not only improves operational reliability but also enhances worker safety in high-risk environments.

AI also fosters an evidence-based approach to manufacturing. Through the integration of Digital Twins and the Internet of Production (IoP), data from design, production, and usage phases converge into a unified decision space. This enables real-time monitoring, enhanced traceability, faster root-cause analysis, and improved demand forecasting and strategic planning. Moreover, AI drives product innovation by combining generative design, simulation, and multi-objective optimization. Techniques such as Generative Adversarial Networks (GANs), Variational Autoencoders (VAEs), and reinforcement learning allow engineers to explore novel design spaces that balance performance, manufacturability, and sustainability. AI-driven personalization further enables manufacturers to deliver tailored products and services, enhancing customer satisfaction and competitive differentiation.

Finally, AI supports sustainability and circular economy practices. From early-stage design for disassembly (DfD) to end-of-life diagnostics and smart retrofitting, intelligent systems optimize energy consumption, minimize waste, and facilitate reuse, remanufacturing, and recycling, aligning industrial operations with environmental regulations and corporate social responsibility goals.

Challenges

Despite these benefits, manufacturing companies face significant challenges in implementing AI at scale. One primary obstacle is managing the vast and heterogeneous data generated across production environments. Integrating information from legacy systems, IoT sensors, and cloud platforms requires standardized taxonomies and interoperable data architectures; inconsistent or poor-quality data can severely compromise model performance and decision reliability.

Additionally, industrial AI applications demand substantial computational resources and low-latency processing. Real-time inference in robotics or additive manufacturing often necessitates high-performance edge computing and secure connectivity, entailing high costs for infrastructure, software, and model deployment, particularly burdensome for SMEs already facing resource constraints (see Chapter 1.5.3). This computational burden is further amplified in agentic AI systems, where continuous orchestration of reasoning, memory, and tool use increases overhead (see Section 2.5).

Human and organizational factors also pose significant barriers. Transitioning to AI-enabled manufacturing requires new competencies in data engineering, AI supervision, and digital ethics. Workforce resistance, insufficient training, and organizational inertia may hinder adoption, while effective human–AI collaboration demands clear role definitions and robust oversight mechanisms, such as “human-in-the-loop” (HITL), to ensure AI augments rather than replaces human expertise.

Trust, explainability, and safety remain critical concerns, especially in high-risk industrial settings. The opacity of complex AI models challenges accountability and decision validation. While emerging approaches like Explainable AI (XAI) aim to provide interpretable justifications for predictions (see Section 3.3.3), their integration remains limited to pilot applications. Cybersecurity vulnerabilities, including adversarial attacks and data manipulation, can propagate rapidly through interconnected production networks, threatening both operational integrity and worker safety.

Finally, economic and regulatory uncertainty presents additional hurdles. Although AI promises long-term gains, initial investments are substantial and short-term ROI is often difficult to quantify. Evolving legal frameworks, such as the EU AI Act, impose obligations on high-risk AI systems but do not yet fully address the unique dynamics of autonomous agents interacting with physical environments, a gap that increases compliance risks for companies operating in sensitive sectors (see Sections 1.6.2 and 2.5).

Synthesis

In conclusion, AI offers manufacturing companies transformative benefits, including increased efficiency, predictive capabilities, sustainable innovation, and enhanced competitiveness. However, these advantages are counterbalanced by persistent challenges in data governance, technical scalability, workforce adaptation, cybersecurity, and regulatory compliance. The successful adoption of AI in manufacturing therefore requires a carefully balanced strategy that integrates technological readiness, organizational maturity, workforce development, and ethical responsibility. These considerations form the foundation for the following section, which synthesizes the overall implications of an AI-driven product lifecycle.

4. Conclusions

The preceding chapters have mapped, at different levels of granularity, how Artificial Intelligence and related Industrial Digital Technologies can be progressively embedded across the entire product lifecycle, from conceptual design to end-of-life, and how agentic paradigms, digital shadows and digital twins enable a continuous, data-driven feedback loop between phases. This synthesis draws together those threads and articulates a coherent view of the opportunities, the tensions and the pragmatic levers that emerge from the analysis presented in Chapters 1–3.

At the core of an AI-driven lifecycle lies an integrated data and modelling ecosystem. Design decisions can no longer be treated as isolated artifacts: Digital Shadows, aggregated datasets enriched with context from the Internet of Production (IoP), enable the transformation of in-field signals, production traces, and simulation outputs into reusable knowledge that informs successive design iterations, process parameter tuning and maintenance strategies. Model-integrated AI, i.e. the coupling of data-driven methods with domain engineering models (CAD, FEA), is therefore a necessary enabler for solutions that are both technically sound and operationally deployable. Likewise, Digital Twins provide the real-time, bidirectional interface that closes the loop between physical assets and their virtual counterparts.

Agentic AI and co-pilot paradigms amplify the lifecycle value proposition by turning passive analytics into active orchestration. Agents endowed with memory, tool-use and planning capabilities can sequence multi-step workflows (for example: diagnose, propose repair, schedule intervention), coordinate heterogeneous tools (analytics engines, scheduling systems, robotic actuators) and maintain contextual continuity across phases. Used within well-defined human-in-the-loop schemes, such agents raise the productivity and resilience of manufacturing operations while preserving meaningful human oversight.

The benefits that flow from this integrated vision are multiple and mutually reinforcing: improved design quality and faster innovation through generative and model-integrated design; higher operational efficiency and reduced downtime via predictive maintenance and real-time process control; enhanced product-service personalization and evidence-based decision-making through converged data; stronger sustainability outcomes when Design-for-X and end-of-life diagnostics are embedded into the same information architecture. These gains, extensively documented across the chapter case studies, underpin the strategic rationale for AI investment in manufacturing.

At the same time, the analysis highlights persistent and interdependent challenges that temper immediate, indiscriminate adoption. Chief among these are data governance and interoperability (heterogeneous, legacy systems and inconsistent taxonomies), the computational and integration costs of advanced/agentic architectures (a barrier for many SMEs), human and organizational readiness (skills, trust, role redefinition), explainability and safety (the need for XAI and robust HITL practices), cybersecurity and adversarial exposure, and regulatory uncertainty (notably in the treatment of high-risk autonomous systems). The synthesis in section 3.6 captures these trade-offs and stresses that technological potential must be matched with governance and capability building.

From these premises follow a set of pragmatic, internally consistent recommendations for firms and practitioners seeking to deploy an AI-driven lifecycle:

- Design the information backbone first. Prioritize interoperable data schemas, Digital Shadows and agreed taxonomies so that insights generated in one phase are actionable in others.

- Adopt a model-integrated approach. Combine physics-based engineering models with data-driven learners to ensure outputs are physically credible and manufacturable.
- Start with modular pilots that embed human-in-the-loop controls. Validate reliability, XAI narratives and human oversight mechanisms at small scale before scaling to mission-critical operations.
- Invest in workforce transition and governance. Reskilling, clear accountability frameworks, and ethics/audit capabilities are as important as technical components.
- Factor sustainability and computational footprint into architectural choices. Lightweight models, model distillation and Green-AI considerations should be part of design and deployment decisions.

In closing, the thesis demonstrates that an AI-driven product lifecycle is not a single technology project but an organizational transformation: it requires aligning data architecture, modelling paradigms, operational processes and human capabilities around a continuous learning loop. When implemented with measured pilots, rigorous governance, and a clear focus on interoperability and explainability, the integrated approach outlined in this work can deliver substantial gains in efficiency, innovation and sustainability. Yet, as highlighted throughout the analysis, realizing this promise demands conscious trade-offs and institutional commitment: AI must be embedded responsibly, with human agency, safety and legal compliance always preserved as foundational constraints.

References

[1] J. McCarthy, M. Minsky, N. Rochester and C. Shannon, "Proposal for the Dartmouth Summer Research Project on Artificial Intelligence," *Dartmouth Summer Research Project,* 1956.

[2] W. S. McCulloch and W. Pitts, "A Logical Calculus of the Ideas Immanent in Nervous Activity," *Bulletin of Mathematical Biophysics,* 1943.

[3] A. Turing, "Computing Machinery and Intelligence," *Mind (Oxford University Press),* 1950.

[4] D. Crevier, "AI: The Tumultuous History of the Search for Artificial Intelligence," *Book, Basic Books,* 1993.

[5] C. Janiesch, P. Zschech and K. Heinrich, "Machine learning and deep learning," *Electronic Markets (Springer),* 2021.

[6] A. Kaplan and M. Haenlein, "Siri, Siri, in my hand: Who's the fairest in the land? On the interpretations, illustrations, and implications of artificial intelligence," *Business Horizons (Elsevier),* 2019.

[7] IBM, "What Is Artificial Intelligence (AI)?," *Corporate web article, IBM Research,* 2024.

[8] A. P. Wibawa, "Recent advancements in Natural Language Processing," 2024.

[9] N. Hafiz, "A Comprehensive Overview of Large Language Models," *arXiv preprint (Cornell University Library),* 2023.

[10] K. Salman, N. Muzammal, H. Munawar, S. W. Zamir, F. S. Khan and M. Shah, "Transformers in Vision: A Survey," *arXiv preprint,* 2021.

[11] X. Tu, Z. He, Y. Huang, Z. Zhang, M. Yang and J. Zhao, "An overview of large AI models and their applications," *Information Fusion (Elsevier),* 2024.

[12] Kilian Carolan, Laura Fennelly and Alan F. Smeaton, "A Review of Multi-Modal Large Language and Vision Models," *arXiv preprint,* 2024.

[13] Z. Xi, W. Chen, X. Guo, W. He, Y. Ding, B. Hong, M. Zhang and J. Tang, "The Rise and Potential of Large Language Model Based Agents: A Survey," *Science China Information Sciences,* 2023.

[14] I. Analytics, "The leading generative AI companies," *Industry report, Intelligence Analytics,* 2025.

[15] I. Analytics, "Generative AI foundational models and platforms," *Industry report, Intelligence Analytics,* 2025.

[16] M. Baccardax, "The AI Trade Is More Than Just Nvidia and AMD. 2 Other Stocks to Consider.," *TheStreet (Financial news site),* 2025.

[17] I. Fried, "Tech giants ramp up spending as AI starts to deliver," *CNBC (Tech news),* 2025.

[18] B. Insider, "Meta CTO breaks down who in Big Tech he thinks has an advantage in the AI race (and it's not Google)," *Business Insider (Tech section),* 2025.

[19] F. Times, "OpenAI in talks for share sale valuing ChatGPT maker at US $500 bn," *Financial Times,* 2025.

[20] T. Guardian, "OpenAI takes on Meta and DeepSeek with free and customisable AI models," *The Guardian (Technology),* 2025.

[21] E. Mahboob, O. A. Samuel, L. M. L. Jose and A. P. G. Jose, "A comprehensive literature review of the applications of AI techniques," *Discover Artificial Intelligence,* 2023.

[22] M. Stefan, Y. Mahnoor and e. al., "Digital Twins: A Survey on Enabling Technologies,," *IEEE Communications Surveys & Tutorials,* 2022.

[23] M. G. Bala and K. P. Sai, "Optimal disassembly sequence generation and disposal of parts using," *Sådhanå,* 2021.

[24] E. A. Mohammed and R. Alireza, "Robotic disassembly for end-of-life products focusing on task and motion," *ournal of Manufacturing Systems,* 2024.

[25] X. Wei and Y. Zhaohui, "Edge Learning for B5G Networks With Distributed Signal Processing: Semantic Communication, Edge Computing, and Wireless Sensing," *IEEE Transactions on Wireless Communications,* 2023.

[26] K. Lukas, A. Antonino and N. Georg, "Boosting Innovation by Artificial Intelligence (AI): Effects and Potentials of AI in the Process of Innovating," *Portland International Conference on Management of Engineering and Technology,* 2022.

[27] B. Philipp, D. Manuela and J. Matthias, "A Computer Science Perspective on Digital Transformation in Production," *ACM Transactions on Internet Things,* 2021.

[28] L. Lingdi, S. Wenyan and L. Yang, " Leveraging digital capabilities toward a circular economy," *Computers & Industrial Engineering,* 2023.

[29] L. Shuran, W. Chengwei, R. Liying, W. Yuwei and W. Zhiqiang, "Towards Designing Smart Public Spaces: A Framework for Designers to Leverage AI and IoT Technologies," *Late Breaking Papers,* 2023.

[30] L. Vora, A. Gholap and K. Jetha, "Artificial Intelligence in Pharmaceutical Technology and Drug Delivery Design," *Pharmaceutics,* 2023.

[31] M. Rezwanul and A. M. Mohammad, "A Comprehensive Review on Artificial Intelligence/Machine Learning Algorithms for Empowering the Future IoT Toward 6G Era," *IEEE Access,* 2022.

[32] K. D. Prasanta, C. Soumyadeb and A. Amelie, "Artificial intelligence-driven supply chain resilience in Vietnamese manufacturing small end medium sized enterprises," *International Journal of Production,* 2024.

[33] P. Melina, V. E. Sebastian and M. Maurice, "Characterization of Usage Data with the Help of Data Classifications," *Journal of the Knowledge Economy,* 2024.

[34] Y. Luo, "A Systematic Literature Review on the Use, Development, and Perception of AI-Enabled UX Design Tools," *Advances in Human-Computer Interaction,* 2025.

[35] C. Xusen, S. Linlin and L. Xin, "The Good, the Bad, and the Uglyugly: impact of analytics and artificial intelligence-enabled personal information collection on privacy and participation in ridesharing," *European Journal of Information,* 2021.

[36] Y. Murat, R. Andreas and G. Christian, "Considerations About Integration of GenAI into Products and Services from an Ethical and Legal Perspective," *Communications in Computer and Information Science,* 2024.

[37] N. Alessandro, A. B. Maria and G. Rita, "Sustainable management of electric vehicle battery remanufacturing: A systematic literature review and future directions," *Journal of Manufacturing Systems,* 2024.

[38] G. Stefanos, K. Maria and S. Tushar, "Green AI: Do Deep Learning Frameworks Have Different Costs?," *International Conference on Software Engineering,* 2022.

[39] M. T., D. I., K. E. and K. P., "The Role of Lightweight AI Models in Supporting a Sustainable Transition to Renewable Energy: A Systematic Review," *Energies,* 2025.

[40] Y. Vahid and K. Baris, "Retrosynthetic Life Cycle Assessment: A Short Perspective on the Sustainability of Integrating Thermoplastics and Artificial Intelligence Into Composite Systems," *Advanced Sustainable Systems,* 2024.

[41] R. M. B. University of Surrey, "Robust facility layout design under uncertain product demands," *Proc CIRP.,* 2017.