# Artificial Intelligence, Big Data, and Supply Chain Resilience: A Survey

www.surveyx.cn


## Abstract

This survey examines the intersection of artificial intelligence (AI), big data, and supply chain resilience, providing a comprehensive framework for enhancing organizational efficiency and adaptability. The integration of AI technologies with the Technology-Organization-Environment (TOE) perspective facilitates intelligent decision-making, aligning technological capabilities with organizational goals. AI and big data analytics offer robust tools for extracting actionable insights, essential for strategic planning. A Human-Centered AI (HCAI) approach is emphasized, particularly in manufacturing, fostering innovation and sustainable practices. The survey highlights the pivotal role of risk management, enhanced by AI and data analytics, in identifying and mitigating potential threats. It calls for comprehensive AI policy frameworks and explores ethical implications, ensuring AI technologies contribute positively to societal well-being. Future research directions include developing standardized datasets, enhancing data governance, and integrating AI with blockchain technologies. Promising results in blockchain and federated learning integration are noted, maintaining data privacy while achieving high accuracy in medical datasets. The survey concludes by suggesting future research in AI's transition to proactive assistance, redefining cybersecurity training, and adapting to technological changes. This integration fosters a resilient and adaptive organizational structure, essential for maintaining a competitive edge in the evolving market.


## 1 Introduction

### 1.1 Significance of Integration

The integration of artificial intelligence (AI), big data, and supply chain resilience is essential for enhancing organizational efficiency and adaptability in the global market's dynamic landscape. This convergence significantly transforms decision-making processes and operational efficiencies across various sectors. For instance, AI's role in intelligent decision-making is illustrated through its application in Cyber-Physical Systems (CPS), where standardized frameworks enable the evaluation of AI-enabled systems against traditional controllers, thereby improving industrial operations [1]. Additionally, the fusion of AI with cloud service optimization highlights its potential in intelligent manufacturing, yielding substantial operational advantages [2].

AI's incorporation into environmental and corporate governance frameworks, such as the Climate AI (CAI) model, automates the extraction and validation of corporate carbon reduction commitments, enhancing efficiency and accuracy in environmental management [3]. This underscores AI's broader applicability in achieving sustainability objectives. Moreover, integrating multiple data layers facilitates timely decision-making in emergency management, thereby improving community resilience during natural disasters [4].

In healthcare, generative artificial intelligence, particularly through Multi-modal Large Language Models (MLLMs), enhances secure data management and sharing, significantly boosting organi-

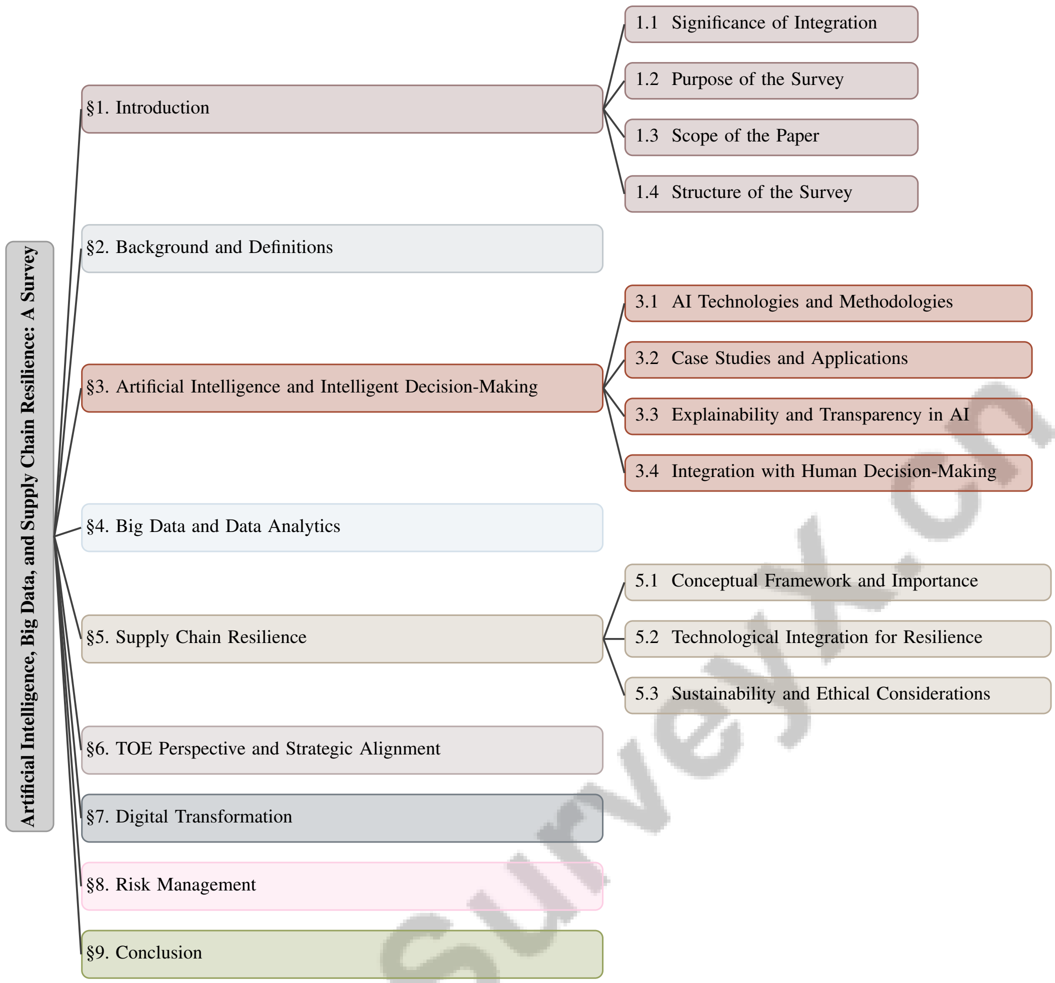



Figure 1: chapter structure

zational efficiency [5]. The emphasis on a human-centered design approach in AI implementation fosters trust and reliability, which are critical for organizational success [6]. A structured classification framework for AI integration, particularly in watershed management, further advances understanding and development in this field [7].

The integration of data-importance aware radio resource management (RRM) with edge machine learning addresses communication bottlenecks, optimizing data flow and enhancing decision-making [8]. In software engineering, AI transforms developer roles and routines, providing enhanced capabilities to organizations [9]. Furthermore, the combination of AI and virtual reality (VR) in cybersecurity training programs increases engagement and effectiveness, showcasing the transformative potential of these technologies [10].

Strategic integration of AI, big data, and supply chain resilience not only strengthens operational capabilities but also equips organizations to navigate the complexities of modern technological landscapes. This integration enhances organizational resilience and adaptability by leveraging advanced technologies such as Edge Computing and AI, which are vital for addressing the challenges of a rapidly evolving market. By embracing dynamic regulatory frameworks and fostering innovation ecosystems, organizations can effectively respond to disruptive technologies and maintain a competitive edge in an uncertain environment [11, 12, 13].

## 1.2 Purpose of the Survey

This survey aims to explore the integration of artificial intelligence, big data, and supply chain resilience, with a focus on their innovative applications and governance mechanisms. A key objective



AI-generated, for reference only.

is to investigate how AI methods can enhance predictions and decision-making in economic policy and governance during significant disruptions, thereby improving organizational adaptability [14]. The survey also seeks to bridge the gap between academic research and practical applications in climate change innovation, emphasizing the role of AI and big data in this context [15].

Additionally, the survey examines the implications of AI reliance, particularly human behavior towards AI systems, and how organizations can effectively leverage AI while enhancing human roles in decision-making. The integration of blockchain technology is proposed to facilitate decentralized intelligence sharing, thereby enhancing trust and transparency in AI systems [16]. Furthermore, the potential of machine learning and AI in advancing the teaching of sustainable engineering practices is explored, contributing to the development of future-ready skills [17].

The survey also addresses communication latency challenges in edge learning by introducing data-importance aware radio resource management (RRM) for edge machine learning [8]. In cybersecurity, it aims to fill the gap in comprehensive studies on the effectiveness of cybersecurity awareness training methods, guiding organizations in selecting suitable strategies to enhance their security posture [10].

### 1.3 Scope of the Paper

This survey delineates the integration of artificial intelligence (AI), big data, and supply chain resilience by focusing on innovative applications and governance mechanisms across various domains. It encompasses AI techniques applied to operating system tasks such as memory management, process scheduling, and security, while excluding unrelated AI applications and operating system functionalities that do not directly involve AI integration [18]. Furthermore, the survey includes climate change innovation projects funded by European framework programmes and peer-reviewed literature from 1979 to 2021, emphasizing the interaction between research outputs and innovation actions [15].

Topics related to AI reliance and human behavior towards AI systems are covered, excluding unrelated areas of AI research [19]. In cultural heritage, it encompasses the stages of knowledge dissemination from oral traditions to the digital era, focusing on languages and cultural diversity, while excluding detailed discussions on specific technological implementations [20].

In education, the survey highlights the application of machine learning toolkits in engineering education, particularly in civil and environmental engineering, including specific case studies [17]. AI education methodologies, such as project-based and problem-based learning, are examined, while standalone AI curriculums that do not integrate with other subjects are excluded [21].

The scope also includes cloud service composition, optimization objectives, and algorithms for intelligent manufacturing, while unrelated optimization areas are excluded [2]. In healthcare, the focus is on data management challenges posed by multi-modal healthcare data and the integration of advanced retrieval techniques [5].

Historical and theoretical frameworks of AI are explored, specifically learning models such as 'Once learning', 'One-shot learning', and 'You Only Look Once - YOLO', while unrelated AI applications and technologies that do not align with the proposed classifications are excluded [7]. The survey includes principles, recent advancements, design examples, and research opportunities in data-importance aware radio resource management (RRM), excluding traditional RRM techniques [8].

The integration of AI in the Software Development Life-Cycle (SDLC) implementation phase is included, while areas outside the SDLC are excluded [9]. Additionally, the scope encompasses traditional, technology-based, and innovative training methods for cybersecurity awareness, while excluding non-cybersecurity training methods and non-educational approaches [10].

By establishing clear boundaries, the survey facilitates a focused examination of the interplay between AI, big data, and supply chain resilience. This targeted exploration yields significant insights into their practical applications and governance frameworks, highlighting the potential for these technologies to enhance operational efficiency and security within supply chains. The findings also underscore the importance of understanding the complexities of AI and big data integration and the implications for regulatory measures in a rapidly evolving technological landscape [22, 23, 13].

## 1.4 Structure of the Survey

This survey is structured to provide a comprehensive examination of the integration of artificial intelligence (AI), big data, and supply chain resilience, organized into distinct sections for clarity and depth. The paper begins with an **Introduction**, which discusses the significance of integrating these technologies and outlines the purpose and scope of the survey. Following this, the **Background and Definitions** section delves into core concepts, providing definitions and exploring the interconnections and relevance of AI, big data, and supply chain resilience to the study.

The survey then explores **Artificial Intelligence and Intelligent Decision-Making**, examining AI technologies and methodologies that enhance decision-making processes, supported by case studies and applications. The importance of explainability and transparency in AI systems is discussed, along with the integration of AI with human decision-making processes. The next section, **Big Data and Data Analytics**, investigates the role of big data in analyzing vast datasets, addressing the integration of multiple data sources, security challenges, industry-specific applications, and the impact of IoT on real-time data analytics.

In the **Supply Chain Resilience** section, the focus shifts to maintaining operational continuity through supply chain resilience, exploring strategies and technologies that enhance this resilience, including sustainability and ethical considerations. The survey introduces the **TOE Perspective and Strategic Alignment**, discussing its role in strategic alignment and digital transformation, followed by an exploration of frameworks and models aligning with the TOE perspective.

The section on **Digital Transformation** examines the process of integrating digital technologies into business processes, emphasizing strategic alignment and interdisciplinary approaches. The survey addresses **Risk Management**, highlighting the role of AI and data analytics in identifying and mitigating potential threats, and discussing governance frameworks and ethical considerations.

In the **Conclusion**, we synthesize critical findings from our exploration of various technologies and methodologies, emphasizing their collective impact on fields such as Responsible AI, generative AI in higher education, and Edge Intelligence. We highlight the translational pathways of research into practical applications, noting the significance of patents and code repositories in measuring real-world implications. Additionally, we provide strategic recommendations for future research directions, advocating for a multidisciplinary approach that incorporates diverse expertise to address challenges and opportunities presented by these rapidly evolving technologies [11, 24, 25, 13, 26]. This structured approach ensures a thorough exploration of the integration of AI, big data, and supply chain resilience, offering valuable insights into their applications and governance mechanisms.The following sections are organized as shown in Figure 1.

# 2 Background and Definitions

## 2.1 Core Concepts and Definitions

Artificial Intelligence (AI) is a transformative technology enabling machines to perform tasks requiring human-like intelligence, such as reasoning, learning, and decision-making across various domains [27]. Explainable AI (XAI) enhances transparency and trust by clarifying model predictions and decision-making processes. AI's applications extend to economic policy enhancement and food supply chain management through machine learning and deep learning techniques [14]. This reliance highlights a socio-technical perspective, intertwining social and technological factors within organizations. Incorporating AI into Cyber-Physical Systems (CPS) improves performance by contrasting AI-enabled systems with traditional controllers [1].

Big Data involves vast volumes of structured and unstructured data generated at high velocity, necessitating advanced analytics for valuable insights [28]. The challenges of fragmented data sources require effective integration strategies in AI research [29]. The convergence of big data analytics with AI is crucial for understanding complex networks and spatial-temporal data, especially in public health and epidemic management [30]. Predictive models and shared datasets are vital for enhancing safety outcomes, underscoring data analytics' role in safety predictions [31].

Supply Chain Resilience refers to the ability of supply chains to anticipate, prepare for, and respond to disruptions while maintaining continuous operations [30]. Integrating AI and big data enhances resilience through real-time monitoring and improved decision-making. Blockchain and smart contracts further strengthen supply chain robustness by ensuring secure and transparent transactions.

The Technology-Organization-Environment (TOE) framework is a strategic model for adopting and implementing technological innovations within organizations, emphasizing the alignment of technological capabilities with organizational goals and environmental factors to facilitate digital transformation [32]. This framework highlights embedding digital technologies into business processes to improve efficiency and adaptability [33].

Human Rights Impact Assessment (HRIA) systematically evaluates the potential human rights implications of AI technologies and data-intensive systems [34]. This concept ensures that technological advancements adhere to ethical considerations and positively impact society. Preserving cultural heritage amid rapid digitalization is also crucial, highlighting digital technologies' role in knowledge dissemination and cultural diversity protection [20].

In cybersecurity training, the survey defines traditional methods (passive awareness, classroom-based), technology-based methods (simulation, app-based), and innovative methods (game-based, VR/AR) [10].

These foundational concepts elucidate the intricate interplay between AI, big data, and supply chain resilience, emphasizing their critical interdependencies and significance within contemporary organizational frameworks. By exploring recent advancements and applications, the survey underscores how these technologies collectively address challenges such as data security, algorithmic accountability, and operational efficiency, fostering a more resilient and responsive supply chain ecosystem in the digital era [23, 35, 36, 37, 13].

### 2.2 Interconnections and Relevance

The integration of AI, big data, and supply chain resilience forms a cohesive framework that enhances organizational efficiency and adaptability. The relationship between AI technologies and emergency management exemplifies this integration, where the combination of heterogeneous data sources facilitates superior decision-making during crises [4]. This is further illustrated in healthcare, where secure data sharing and multi-modal data incorporation enhance data freshness and application efficacy [5].

In intelligent manufacturing, optimized cloud service composition demonstrates the synergy between cloud services and manufacturing processes, leading to improved operational efficiency [2]. The Climate AI (CAI) model highlights another critical interconnection by addressing the challenge of extracting structured data from diverse corporate disclosures, linking AI technologies with data extraction in corporate sustainability efforts [3].

AI integration in learning models, such as 'One-shot learning' and 'Few-shot learning', showcases the effectiveness of simpler models in specific applications compared to more complex deep learning models [7]. This emphasizes the necessity of selecting appropriate AI models based on context and application requirements.

In communication, the interconnections between data importance metrics and radio resource management are crucial for enhancing communication efficiency, particularly where data flow optimization is essential [8]. This is mirrored in cybersecurity, where interconnections among various training methodologies are vital for addressing challenges organizations face in implementing effective training strategies [10].

These interconnections highlight the importance of integrating AI, big data, and supply chain resilience to navigate the complexities of evolving technological landscapes. The integration of Edge Computing and AI fosters a robust, efficient, and adaptable organizational framework, essential for maintaining a competitive advantage in today's rapidly changing market, especially as businesses confront disruptive technologies and strive to implement responsible AI practices [37, 12, 13].

## 3 Artificial Intelligence and Intelligent Decision-Making

In recent years, the intersection of artificial intelligence (AI) and intelligent decision-making has garnered significant attention across various fields, highlighting the profound impact of AI technolo-

| Category | Feature | Method |
|---|---|---|
| **AI Technologies and Methodologies** | Language and Text Processing<br>Data Integration and Analysis | HRMMF[5], CAI[3]<br>AI-ERT[4] |
| **Explainability and Transparency in AI** | Knowledge-Based Interpretability | CBR$_E$[38], $DKIE$[39], $E-KELL$[40] |
| **Integration with Human Decision-Making** | Human-AI Collaboration | RAS[41], XAI-Heatmap[42] |

Table 1: This table presents a comprehensive overview of artificial intelligence (AI) technologies and methodologies that are pivotal in enhancing decision-making processes across various sectors. It categorizes AI methods into three main areas: AI Technologies and Methodologies, Explainability and Transparency in AI, and Integration with Human Decision-Making, highlighting specific features and methods employed within each category. The table serves as a valuable resource for understanding the diverse applications and approaches of AI in optimizing decision-making and fostering collaboration between AI systems and human decision-makers.

gies on enhancing decision-making processes. As organizations increasingly adopt AI solutions, it becomes essential to explore the underlying technologies and methodologies that facilitate this integration. Table 1 provides a detailed summary of the AI technologies and methodologies that are instrumental in optimizing decision-making processes across diverse sectors, emphasizing the importance of explainability, transparency, and integration with human decision-making. Additionally, Table 4 provides a comparative analysis of key AI methodologies, illustrating their roles in enhancing decision-making processes across diverse sectors. Figure 2 illustrates the hierarchical structure of AI's impact on decision-making processes, categorizing AI technologies, case studies, explainability, and integration with human decision-making. This figure not only highlights the diverse applications of AI across various sectors but also emphasizes the importance of transparency and collaboration between AI systems and human decision-makers. The following subsection will delve into the specific AI technologies and methodologies that are instrumental in optimizing decision-making across diverse sectors, illustrating their applicability and effectiveness in real-world scenarios.

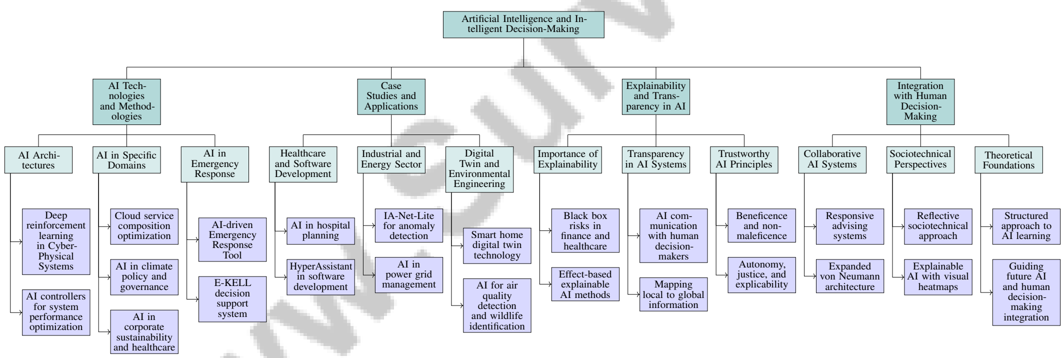



Figure 2: This figure illustrates the hierarchical structure of AI's impact on decision-making processes, categorizing AI technologies, case studies, explainability, and integration with human decision-making. It highlights the diverse applications of AI across various sectors and emphasizes the importance of transparency and collaboration between AI systems and human decision-makers.

### 3.1 AI Technologies and Methodologies

| Method Name | Decision-Making Enhancement | Methodological Diversity | Trust and Transparency |
|---|---|---|---|
| CAI[3] | Increased Accuracy Efficiency | Nlp Techniques Llms | - |
| HRMMF[5] | Data Retrieval Accuracy | Reinforcement Learning Algorithm | Contract Theory Model |
| AI-ERT[4] | Machine Learning Algorithms | Semantic Segmentation Techniques | - |
| E-KELL[40] | Evidence-based Decision-making | Knowledge Graphs | Reliable Outputs |

Table 2: Comparison of AI methodologies in decision-making, methodological diversity, and transparency. This table highlights various AI models and their contributions to enhancing decision-making processes, illustrating the diversity of methods employed, and examining trust and transparency aspects in AI systems.

Artificial Intelligence (AI) technologies and methodologies are instrumental in revolutionizing decision-making processes by integrating machine intelligence with human insights, thereby enhanc-



AI-generated, for reference only.

ing accuracy and efficiency across a myriad of sectors. The categorization of AI methods into distinct architectures, such as deep reinforcement learning in Cyber-Physical Systems (CPS), exemplifies how AI technologies enhance decision-making by training AI controllers to optimize system performance [1]. These architectures provide a robust foundation for developing AI systems tailored to address specific decision-making challenges.

In the realm of cloud service composition, AI's integration is underscored by the categorization of optimization indicators and algorithms into heuristic and reinforcement learning approaches, detailing methodologies that enhance decision-making in cloud environments [2]. Similarly, a novel framework in climate change combines AI methods and network science to analyze and quantify the relationship between research and innovation, thereby enhancing decision-making in climate policy and governance [15].

AI's transformative potential in corporate sustainability efforts is further illustrated through the use of large language models and natural language processing techniques, as demonstrated in the Climate AI (CAI) model, which employs a four-stage pipeline to process context, search relevant texts, extract metrics, and validate data [3]. In healthcare, the Hybrid RAG-empowered Medical MLLMs Framework (HRMMF) integrates hybrid RAG techniques with Multi-modal Large Language Models (MLLMs) to enhance decision-making in healthcare data management, ensuring secure and efficient data handling [5].

The AI-driven Emergency Response Tool (AI-ERT) utilizes machine learning algorithms and geospatial data to improve decision-making for evacuation routing and resource allocation during emergencies [4]. Additionally, E-KELL, a decision support system, utilizes knowledge graphs constructed from emergency standards and regulations to guide large language models in reasoning and generating reliable outputs for emergency decision-making [40].

The survey categorizes research into various fields, including decision-making approaches and transparency mechanisms related to AI reliance, emphasizing the importance of transparency and trust in AI systems [19]. This aligns with the principles of trustworthy AI, which can be achieved through adherence to foundational principles such as beneficence, non-maleficence, autonomy, justice, and explicability [43]. The benchmark study innovatively combines task difficulty and transparency as variables to study their effects on user behavior and AI interaction, diverging from previous benchmarks that focused on simpler metrics [44].

Furthermore, a transdisciplinary perspective on AI education advocates for an integrated approach that connects AI learning with various disciplines and real-world contexts, thereby promoting comprehensive decision-making strategies [21]. Collectively, these methodologies underscore the transformative potential of AI in improving decision-making processes, offering diverse tools and approaches suited to various needs and contexts. As AI technologies continue to advance, they present unprecedented opportunities for enhancing decision-making across multiple sectors, driving efficiency and innovation in organizational practices.

As shown in Figure 3, this figure illustrates the hierarchical categorization of AI technologies and methodologies, focusing on decision-making, applications, and trust and education. Each category highlights key AI methods and their applications across various domains, emphasizing their transformative potential in enhancing decision-making processes. Understanding the intricacies of AI technologies and methodologies is crucial for advancing both theoretical and practical applications. The examples provided in the figure contribute uniquely to the broader understanding of the field, further supporting the discussion of AI's capabilities in diverse contexts [45, 46, 47]. In addition, Table 2 presents a detailed comparison of AI methodologies, focusing on their impact on decision-making enhancement, methodological diversity, and trust and transparency, which are crucial for understanding the role of AI in various domains.

### 3.2 Case Studies and Applications

The successful implementation of artificial intelligence (AI) in decision-making processes is evident across various domains, demonstrating the transformative potential of AI technologies. In hospital planning, a case study illustrates the application of AI technologies, underscoring the practical benefits of integrating AI into healthcare management systems. This integration enhances operational efficiency and facilitates informed decision-making, thereby improving patient care outcomes [6].

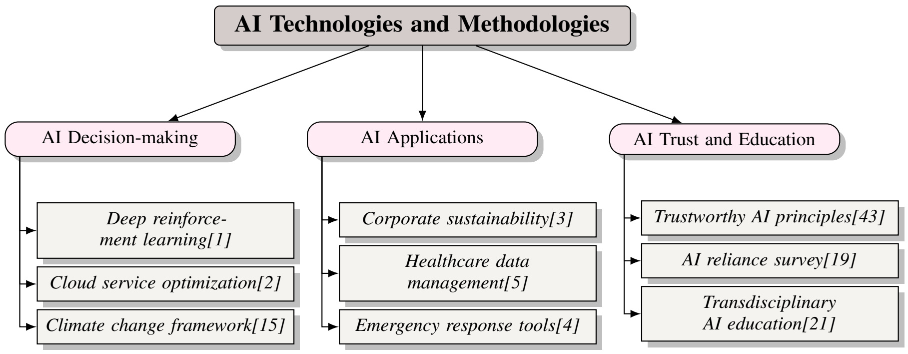



Figure 3: This figure illustrates the hierarchical categorization of AI technologies and methodologies, focusing on decision-making, applications, and trust and education. Each category highlights key AI methods and their applications across various domains, emphasizing their transformative potential in enhancing decision-making processes.

In the realm of software development, HyperAssistant, an AI tool, supports developers by addressing mental health, code quality, and team dynamics. This tool exemplifies AI's role in enhancing decision-making processes by providing real-time feedback and insights, which contribute to improved software quality and team collaboration [9].

In industrial settings, frameworks like IA-Net-Lite showcase significant advancements in low-latency anomaly detection, reducing service latency and enhancing decision-making in real-time applications. This framework underscores the transformative potential of artificial intelligence (AI) in optimizing industrial operations by enhancing efficiency and minimizing operational risks. It emphasizes the significant impact of AI-mediated knowledge access systems on workplace dynamics, highlighting the importance of understanding the associated risks to workers' value, power, and well-being. Additionally, it addresses the necessity for explainability in AI applications, particularly in safety-critical environments, to foster trust and accountability among users and stakeholders. By identifying the specific mechanisms through which AI systems influence operational processes and worker safety, this framework serves as a vital tool for practitioners aiming to design and implement AI solutions that not only optimize performance but also safeguard employee interests and promote ethical practices within industrial settings. [48, 45, 49]. Similarly, in the energy sector, AI applications in power grids utilize both supervised and unsupervised learning techniques to optimize operations, tested against historical data and real-time measurements, thereby improving decision-making processes in energy management.

In digital twin technology, experiments conducted in smart home environments demonstrate AI's role in enhancing decision-making by creating digital twins at various capability levels. This application of AI facilitates real-time monitoring and management of home systems, leading to improved energy efficiency and user convenience. AI's influence on environmental engineering is illustrated by innovative projects like air quality detection systems and automated wildlife identification tools. These initiatives highlight the practical applications of machine learning in tackling pressing environmental issues, such as monitoring pollutants and conserving biodiversity. For instance, machine learning algorithms can analyze data from sensors to assess air quality in real-time, while automated identification systems can monitor wildlife populations and their habitats, assisting in conservation efforts. Such applications not only streamline data processing but also enhance the accuracy and efficiency of environmental assessments, demonstrating the transformative potential of AI in promoting sustainable engineering practices. [17, 25]

Furthermore, in the educational domain, the AI curriculum at Neom Community School incorporates project-based and problem-based learning, allowing students to engage with AI concepts through real-world challenges. This approach enhances decision-making skills by equipping students with hands-on experience in utilizing AI technologies, such as generative AI, to tackle complex problems, thereby fostering critical thinking and problem-solving abilities essential for navigating the evolving landscape of higher education and various industries. [50, 11, 51, 24, 41]

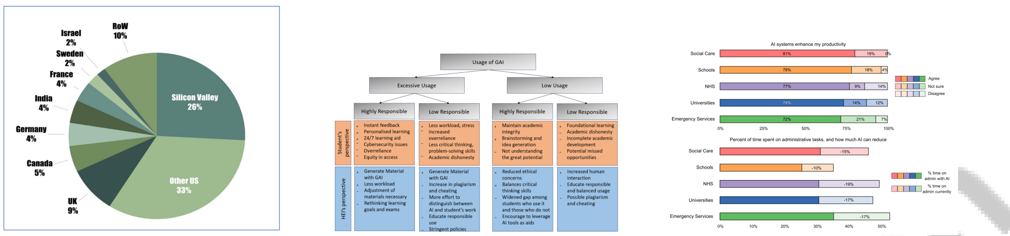
The case studies and applications presented collectively demonstrate the varied and significant ways in which artificial intelligence (AI) has been effectively utilized to enhance decision-making across





AI-generated, for reference only.

multiple industries, including healthcare, finance, manufacturing, and retail. These implementations not only address specific challenges within each sector but also showcase measurable improvements in operational efficiency, innovation opportunities, and overall societal welfare, while also considering ethical implications and the future trajectory of AI development. [51, 24]. They underscore AI's potential to revolutionize decision-making processes across various sectors, driving innovation and efficiency in organizational practices.




(a) Global Innovation Landscape: A Comparative Analysis[12]

(b) The Usage of GAI in Higher Education: A Comparative Analysis[11]

(c) AI systems enhance my productivity[52]

Figure 4: Examples of Case Studies and Applications

As shown in Figure 4, In the realm of artificial intelligence and intelligent decision-making, the integration of AI technologies across various sectors has led to diverse applications and case studies that highlight its transformative potential. The provided examples encapsulate this diversity, showcasing how AI is reshaping innovation, education, and productivity. The first case study, "Global Innovation Landscape: A Comparative Analysis," offers a visual representation of the distribution of global innovation, emphasizing the significant contributions from regions like Silicon Valley and other parts of the United States. This highlights the pivotal role these areas play in driving technological advancements. The second example, "The Usage of GAI in Higher Education: A Comparative Analysis," delves into the adoption of Generative Artificial Intelligence within academic institutions, presenting insights from both students and higher education institutions (HEIs). This analysis underscores the varying levels of responsibility perceived by these stakeholders in leveraging AI for educational purposes. Lastly, the third case study, "AI systems enhance my productivity," presents a bar chart that examines the sentiment towards AI's impact on productivity across different sectors, such as social care, schools, and universities. This example illustrates the varied acceptance and perceived benefits of AI, reflecting its nuanced influence on enhancing efficiency in diverse professional environments. Together, these case studies offer a comprehensive view of AI's capabilities and its evolving role in shaping intelligent decision-making processes. [**?** ]fenwick2024businessregulatoryresponsesartificial,krause2024evolutionlearningassessingtransformative,bright2024generativeaiwides

### 3.3 Explainability and Transparency in AI

| Method Name | Explainability Importance | Transparency Practices | Domain-Specific Applications |
|---|---|---|---|
| DKIE[39] | Better Explainability Needed | Ethical Concerns Compliance | Finance And Healthcare |
| CBR_E[38] | High Interpretability | Transparency And Trust | Financial Risk Detection |
| XAI-Heatmap[42] | High-stakes Environments | Standardized Methods | Manufacturing And Medicine |
| E-KELL[40] | Interpretable Decision-making | Standardized Methods | Emergency Management |

Table 3: The table presents a comparative analysis of various AI methods focusing on explainability, transparency practices, and domain-specific applications. It highlights the importance of explainability in enhancing trust and interpretability in AI systems across different sectors, including finance, healthcare, manufacturing, and emergency management. The table also underscores the role of standardized transparency practices in ensuring ethical compliance and effective decision-making.

Explainability and transparency are pivotal components in the deployment of artificial intelligence (AI) systems, especially as they permeate intricate sociotechnical environments. These attributes ensure that AI systems are not only effective but also interpretable and trustworthy, providing users with a comprehensive understanding of the decision-making processes involved. The 'black box' nature of advanced AI models, particularly in sensitive domains such as finance and healthcare, presents significant risks due to the lack of interpretability [39]. Integrating domain knowledge into AI models can enhance explainability by substituting complex features with simpler, more comprehensible ones [39].

Table 3 provides a detailed overview of different AI methods, emphasizing their explainability, transparency practices, and specific domain applications, which are crucial for fostering trust and understanding in AI systems. The necessity of explainability for fostering trust in AI systems aligns with the broader requirement for standardized practices that enhance transparency across various stages of AI deployment [53]. Effect-based explainable AI (XAI) methods, particularly when employed with robust models like LightGBM, effectively approximate true sensitivities, thereby overcoming the limitations of additive methods that often fail to capture intricate dynamics [48]. In regulated environments, such as financial risk prediction, explainable case-based reasoning (CBR) systems offer enhanced transparency and interpretability, addressing the challenges posed by traditional black-box algorithms [38].

The importance of transparency is further underscored in AI systems used for decision-making, where clear communication between AI systems and human decision-makers is crucial [19]. This necessity is particularly evident in scenarios where AI systems must provide actionable insights that are understandable to users, thereby fostering trust and informed decision-making [42]. The mapping of local information (known knowledge) to global information (new learned knowledge) is a theoretical approach that aids in understanding learning processes within AI systems, contributing to their transparency and interpretability [54].

The study by Ha et al. concluded that task difficulty significantly influences user reliance on AI suggestions, although transparency levels do not have a measurable impact on trust or suggestion usage [44]. Furthermore, the E-KELL system exemplifies the advantage of providing reliable, contextually relevant decision support through structured knowledge and logical reasoning, addressing the limitations of traditional large language model applications [40].

The emphasis on explainability and transparency in artificial intelligence (AI) systems is crucial for fostering trust among users and stakeholders, as it enables a better understanding of AI decision-making processes and outcomes. This transparency is not only essential for responsible AI deployment but also aligns with the foundational principles of Trustworthy AI, which include beneficence, non-maleficence, autonomy, justice, and explicability. By implementing user-centered transparency frameworks, such as the Artificial Intelligence Disclosure (AID) Framework, and adopting a multidisciplinary approach to explainability that considers contextual factors, we can ensure that AI technologies are developed and utilized in a manner that is ethical, accountable, and beneficial to society. [55, 56, 43, 45]. By addressing these aspects, AI systems can achieve greater acceptance and integration into society, ultimately enhancing their effectiveness and reliability.

### 3.4 Integration with Human Decision-Making

The integration of artificial intelligence (AI) with human decision-making processes is a pivotal aspect of enhancing decision quality and efficiency across various domains. A responsive advising system, as proposed by Noti et al., exemplifies this integration by interacting with human decision-makers and offering guidance only when it is likely to improve decision outcomes [41]. This approach ensures that AI complements human judgment rather than replacing it, fostering a collaborative environment where both human intuition and machine intelligence contribute to optimal decision-making.

The expanded von Neumann architecture, which includes components for creativity and knowledge sharing, provides a theoretical framework for integrating AI with human cognitive processes [57]. This architecture supports the development of AI systems that can engage in creative problem-solving and knowledge exchange, aligning closely with human cognitive abilities and enhancing collaborative decision-making.

A reflective sociotechnical approach is crucial for understanding the interplay between technical systems and human factors, as highlighted by Ehsan et al. [58]. This perspective emphasizes the need to consider the social and organizational contexts in which AI systems are deployed, ensuring that AI technologies are designed to support human decision-makers in a manner that is both effective and ethically sound.

The use of explainable AI, particularly visual heatmaps, has been shown to enhance human decision-making by allowing domain experts to validate AI predictions and improve their overall task perfor-



AI-generated, for reference only.

mance [42]. This integration of AI with human decision-making processes ensures that AI systems provide transparent and interpretable insights, enabling users to make informed decisions based on a clear understanding of AI-generated recommendations.

As illustrated in Figure 5, the integration of AI with human decision-making encompasses key methods such as Responsive Advising Systems, Explainable AI, and Reflective Sociotechnical Approaches, each playing a significant role in enhancing decision quality and efficiency.

Furthermore, the structured approach to understanding learning in AI, as proposed by Wu et al., offers a necessary condition that can guide future research in integrating AI with human decision-making [54]. By providing a theoretical foundation for AI learning processes, this approach facilitates the development of AI systems that can effectively collaborate with human decision-makers, enhancing the overall decision-making process.

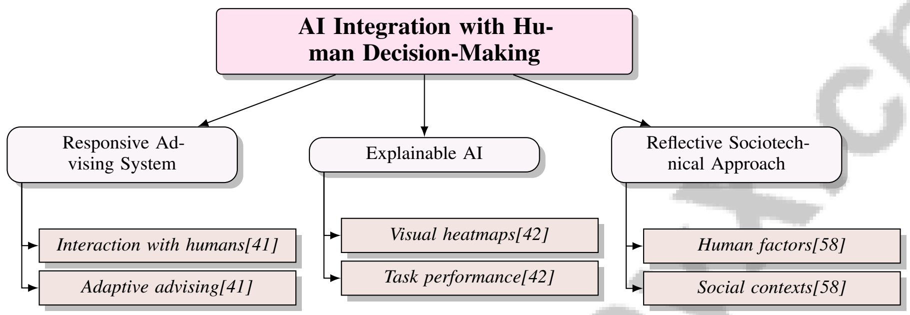



Figure 5: This figure illustrates the integration of AI with human decision-making through key methods: Responsive Advising Systems, Explainable AI, and Reflective Sociotechnical Approaches, highlighting their roles in enhancing decision quality and efficiency.

| Feature | Deep Reinforcement Learning | Climate AI Model | Hybrid RAG-empowered MLLMs |
|---|---|---|---|
| **Applicability** | Cyber-Physical Systems | Climate Policy | Healthcare |
| **Explainability** | Limited | Moderate | Moderate |
| **Integration** | System Performance | Data Validation | Data Management |

Table 4: Comparison of AI Methodologies in Decision-Making Across Various Sectors. This table highlights the applicability, explainability, and integration focus of three distinct AI models: Deep Reinforcement Learning, Climate AI Model, and Hybrid RAG-empowered MLLMs. It provides insights into how these methodologies are tailored to optimize decision-making in cyber-physical systems, climate policy, and healthcare, respectively.

# 4 Big Data and Data Analytics

The synergy between big data and data analytics is pivotal for extracting valuable insights from extensive datasets, emphasizing the integration of diverse data sources. This integration underpins effective analysis, allowing organizations to harness comprehensive datasets for informed decision-making. The following subsection delves into the importance of integrating multiple data sources to bolster data analysis capabilities and yield actionable insights.

## 4.1 Integration of Multiple Data Sources

Integrating multiple data sources is fundamental for comprehensive analysis, especially in big data contexts where diverse datasets must be unified to derive actionable insights. This process involves harmonizing heterogeneous data types, such as social, contextual, and financial data. For example, a dataset from a repository of accident reports shared by nine companies demonstrates how integrating diverse data sources enhances safety predictions and decision-making [31].

In visual analytics, data from the 2011 Visual Analytics Science and Technology (VAST) Challenge illustrates the use of social media data to identify illness-related content, improving epidemic tracking and response [44]. Moreover, AI-generated visual heatmaps provide domain experts with insights into factors affecting AI predictions, thus enhancing interpretability and decision-making [42].

These examples underscore the necessity of integrating diverse data sources to enhance data analysis and overcome challenges posed by fragmented data formats, which can hinder AI applications. Utilizing a broad range of data from various domains—including finance, life sciences, and social media—enhances data readiness and processing efficiency, advancing AI and machine learning initiatives [29, 36, 59]. Advanced methodologies and tools enable deeper insights, fostering innovation and maintaining a competitive edge in the data-driven landscape.

### 4.2 Security and Sensitive Data Handling

Secure handling of sensitive data in big data environments presents significant challenges, particularly concerning data breaches and privacy. A primary concern is the risk associated with shared hardware resources in cloud environments, where compromised hypervisors can lead to data breaches [60]. This vulnerability necessitates robust security measures to ensure data integrity and confidentiality.

In federated learning, data reconstruction attacks exploit distributed systems to reconstruct private data, highlighting the need for leakage-resilient aggregation methods to protect privacy [61]. Addressing these challenges requires a comprehensive approach integrating advanced encryption techniques and secure communication protocols to mitigate exposure risks.

Ensuring AI systems are understandable and explainable complicates the secure handling of sensitive data. In contexts where AI decisions have significant implications, maintaining transparency and security is crucial for user trust and regulatory compliance [58]. Implementing explainable AI (XAI) methods that accurately assess input feature sensitivity enhances the transparency and robustness of AI models in industrial processes [48].

Moreover, cybersecurity requirements for platforms handling sensitive data must be addressed to prevent unauthorized access and breaches. This entails establishing comprehensive security frameworks that include data encryption, access controls, and continuous monitoring to detect and respond to potential threats [62]. By tackling these challenges, organizations can secure sensitive data handling, protect user privacy, and maintain the integrity of big data systems.

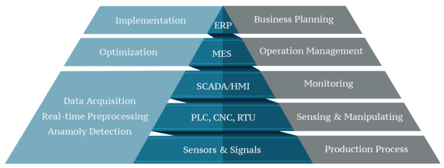
### 4.3 Industry-Specific Applications

Big data applications are integral across various industries, driving innovation and enhancing decision-making through vast datasets. In construction, the development of generic models utilizing data from multiple companies highlights big data's transformative potential, improving prediction accuracy and operational efficiencies [31].

In cloud computing, a secure architecture has been proposed to package and schedule workflows while ensuring data security during transit, in use, and at rest. This architecture addresses the critical need for confidentiality and integrity in cloud environments, facilitating secure processing of sensitive data across sectors [60].

These examples illustrate big data's wide-ranging applications across industries, including finance, healthcare, and retail, demonstrating its capacity to streamline operations, enhance predictive analytics, and bolster data security measures, ultimately driving innovation and economic growth while addressing industry-specific challenges [29, 36, 51]. As industries continue to adopt big data technologies, they achieve greater efficiencies and maintain a competitive edge in the rapidly evolving digital landscape.

As shown in Figure 6, industry-specific applications within Big Data and Data Analytics offer transformative potential by tailoring technological advancements to meet unique sector demands. The first image illustrates a pyramid structure that delineates hierarchical levels within an industrial automation system, emphasizing structured data flow and optimization for seamless operations and strategic planning. The second image focuses on human-centered AI applications in the forest industry, highlighting AI's role in enhancing safety, efficiency, and sustainability in forestry operations. These diverse applications underscore how industry-specific adaptations of big data and analytics can drive innovation and operational excellence, fostering a more intelligent and responsive industrial ecosystem [63, 64].




(a) The image represents a pyramid structure illustrating the hierarchy of different levels in an industrial automation system.[63]


(b) Human-centered AI in the forest industry: Enhancing safety, efficiency, and sustainability[64]

Figure 6: Examples of Industry-Specific Applications

### 4.4 IoT and Real-Time Data Analytics

The integration of Internet of Things (IoT) devices is pivotal for enabling real-time data analytics, offering transformative potential across various domains through immediate data collection and processing. The Edge Intelligence-based Traffic Monitoring System (EI-TMS) exemplifies IoT application in traffic monitoring, providing crucial insights for effective management by continuously monitoring traffic conditions [65].

This integration is visually represented in Figure 7, which illustrates the incorporation of IoT in real-time data analytics across various domains, including smart cities, industrial applications, and healthcare. The figure highlights key applications such as traffic monitoring, predictive maintenance, and patient monitoring, emphasizing the transformative potential of IoT technologies in enhancing operational efficiency, resource allocation, and personalized healthcare.

IoT devices are essential for developing smart environments, where they facilitate the collection of extensive data that can be analyzed in real-time to optimize system performance. In smart city initiatives, IoT sensors monitor environmental conditions, energy usage, and public safety, providing city planners with actionable insights to enhance urban living. Real-time analytics from IoT systems empower cities to make dynamic adjustments to infrastructure, significantly improving resource allocation and service delivery through advanced technologies like edge computing and artificial intelligence [66, 65].

In industrial settings, IoT-enabled real-time data analytics supports predictive maintenance by continuously monitoring equipment performance and identifying potential failures before they occur. This proactive approach leverages advanced IoT technologies and 5G networks to minimize downtime and enhance operational efficiency, showcasing their transformative role in modern smart factories. By integrating intelligent network devices capable of in-network processing, manufacturers can reduce service latency by up to 40

The integration of IoT with real-time data analytics is also transforming healthcare by utilizing IoT devices to continuously monitor patient health metrics, such as vital signs and activity levels. This real-time monitoring allows immediate responses to critical health events, enhancing patient safety and enabling proactive interventions. The fusion of IoT with artificial intelligence (AI) facilitates the analysis of vast amounts of patient-generated health data (PGHD) and contextual information, supporting personalized health management strategies that improve overall health outcomes [67, 68, 66]. This integration enhances patient care by providing healthcare providers with up-to-date information for timely medical interventions.

The role of IoT in enabling real-time data analytics is crucial for driving innovation and efficiency across sectors. By leveraging IoT capabilities, organizations can gain real-time operational insights, enhance decision-making, and secure a competitive advantage in a data-driven landscape. The integration of advanced technologies such as edge computing and artificial intelligence enables efficient processing of vast data generated at the network edge, facilitating applications in healthcare, disaster recovery, and immersive communication. Furthermore, the combination of 5G networks and edge intelligence allows for low-latency, high-capacity data transmission, essential for developing sophisticated applications like personalized healthcare solutions and autonomous systems. This strategic utilization of IoT technologies supports improved operational efficiency and fosters innovative approaches to data management and analysis, ensuring organizations can adapt and thrive amidst evolving market demands [8, 69, 66, 67, 13].

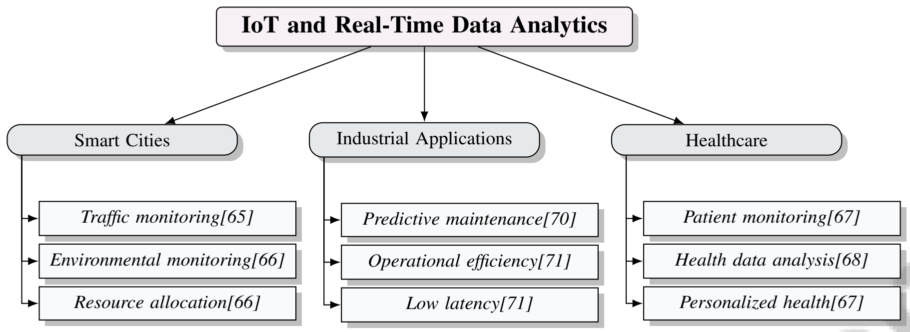



Figure 7: This figure illustrates the integration of IoT in real-time data analytics across various domains, including smart cities, industrial applications, and healthcare. It highlights key applications such as traffic monitoring, predictive maintenance, and patient monitoring, emphasizing the transformative potential of IoT technologies in enhancing operational efficiency, resource allocation, and personalized healthcare.

# 5 Supply Chain Resilience

## 5.1 Conceptual Framework and Importance

Supply chain resilience is defined by the capacity to anticipate, prepare for, and respond to disruptions while maintaining operational continuity. This framework is vital in managing the complexities and risks inherent in modern supply chains. Real-time analysis and decision-making are crucial during disruptions, such as pandemics, as evidenced in economic policy and governance contexts [14]. The integration of AI and big data technologies significantly bolsters community resilience during emergencies by ensuring timely access to reliable information [4]. The E-KELL system exemplifies this by using knowledge graphs to structure emergency-related information, guiding large language models (LLMs) for accurate decision support [40].

In Figure 8, the figure illustrates the conceptual framework of supply chain resilience, highlighting key components such as real-time analysis, AI and big data technologies, and performance assessments in Cyber-Physical Systems (CPS). Each component is supported by relevant studies, demonstrating their roles in enhancing decision-making and operational efficiency in modern supply chains.

In Cyber-Physical Systems (CPS), assessing AI controllers' performance is essential for ensuring safety, reliability, and robustness—key elements of supply chain resilience [1]. Evaluating knowledge flows within complex systems, such as the pharmaceutical industry's drug pipeline, is crucial for developing resilient supply chains capable of withstanding disruptions [30]. This framework combines real-time analysis, AI and big data technologies, and robust performance assessments in CPS, empowering organizations to navigate modern supply chain challenges. By leveraging technologies like AI and knowledge graphs, organizations can enhance decision-making and operational efficiency, preparing for uncertainties and strengthening resilience against evolving complexities [40, 10].

## 5.2 Technological Integration for Resilience

Integrating advanced technologies, particularly AI and big data, is crucial for enhancing supply chain resilience, enabling organizations to effectively anticipate, respond to, and recover from disruptions. Optimizing cloud services significantly contributes to this integration, improving resilience in intelligent manufacturing platforms by ensuring service reliability and efficiency [2]. The Climate AI (CAI) model illustrates how AI can automate carbon reduction commitment extraction, aiding organizations in meeting sustainability goals and enhancing resilience [3].

In healthcare, the hybrid Retrieval-Augmented Generation (RAG) framework with Multi-modal Large Language Models (MLLMs) enhances resilience in data management by ensuring data freshness and

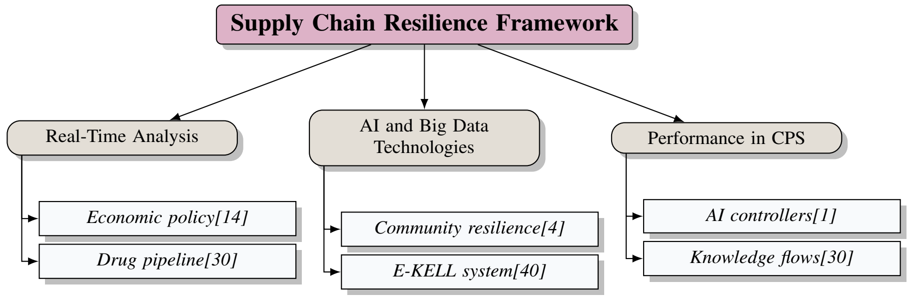



Figure 8: This figure illustrates the conceptual framework of supply chain resilience, highlighting key components such as real-time analysis, AI and big data technologies, and performance assessments in Cyber-Physical Systems (CPS). Each component is supported by relevant studies, demonstrating their roles in enhancing decision-making and operational efficiency in modern supply chains.

reliability [5]. A structured approach to AI implementation that prioritizes stakeholder involvement and trust is crucial for effectively integrating AI systems into organizational practices [6]. A structured classification of AI guides future research and development, enhancing resilience through a better understanding of AI applications and impacts [7]. Understanding user behavior in AI-assisted data exploration can optimize decision-making processes and enhance resilience [44].

In edge learning environments, data-importance aware radio resource management (RRM) mitigates communication latency, bolstering operational continuity and resilience [8]. AI tools in software development streamline workflows and improve software quality, further contributing to resilience [9]. Innovative training methods, including AI and virtual reality (VR), are critical for enhancing organizational resilience against cyber threats, ensuring robust supply chains in the face of evolving security challenges [10]. These technological integrations illustrate AI and big data's transformative potential in fortifying supply chains, enabling organizations to achieve greater resilience and adaptability in a complex market environment.

### 5.3 Sustainability and Ethical Considerations

Sustainability and ethical considerations are crucial in developing resilient supply chains, requiring a comprehensive approach that addresses environmental, social, and governance challenges associated with AI and related technologies. While AI integration offers significant opportunities for enhancing efficiency and sustainability, it also presents ethical challenges requiring proactive management. A balanced approach to data sharing is essential, incorporating organizational, legal, and technical considerations to enhance privacy without impeding research progress [72].

The Human Rights Impact Assessment (HRIA) methodology provides a structured framework for assessing the human rights implications of AI applications, crucial for building supply chains that respect human rights and contribute to societal well-being [34]. Addressing the challenges posed by the volume and complexity of data generated during innovation processes is vital for ensuring sustainable and ethically sound supply chains [32]. Enhanced collaboration and operationalization of research findings in areas like bioproducts and climate-smart agriculture emphasize sustainability and ethical considerations [15]. Integrating machine learning (ML) into engineering education is essential for achieving sustainability goals, highlighting the need for interdisciplinary approaches that incorporate ethical considerations into the curriculum [17].

Future research should explore AI's ethical implications in society, fostering a deeper understanding of societal impacts and ensuring alignment with ethical principles through cohesive AI curricula integrating transdisciplinary approaches [21]. Building resilient cultural frameworks necessitates addressing the loss of languages and oral traditions due to technological barriers, underscoring the importance of sustainable and ethical approaches to cultural preservation [20]. Additionally, cybersecurity requirements for platforms handling sensitive data must be addressed to prevent unauthorized access and data breaches, ensuring the integrity and sustainability of supply chains [62].

By proactively addressing sustainability and ethical considerations, organizations can create resilient supply chains that endure disruptions while fostering positive societal and environmental impacts. This approach ensures long-term resilience and ethical alignment by integrating responsible practices and governance frameworks prioritizing privacy, fairness, and accountability, especially in the context of advancing technologies like AI. Insights from industry practitioners can help organizations identify effective strategies and structures that enhance the efficacy of their responsible initiatives [73, 37].

# 6 TOE Perspective and Strategic Alignment

## 6.1 Introduction to the TOE Perspective

The Technology-Organization-Environment (TOE) framework is pivotal in analyzing technological innovation adoption within organizations, emphasizing strategic alignment and digital transformation. It highlights the integration of AI technologies with traditional systems to achieve strategic objectives [18]. The adaptive systems focus within TOE necessitates frameworks that elucidate AI's relationship with operating systems, promoting adaptive architectures for optimized performance [18]. This adaptability enhances operational efficiency and responsiveness to environmental changes.

Trust and transparency in AI systems are integral to the TOE framework, aligning with trustworthy AI principles. It involves analyzing tensions between technological innovation and ethical considerations, suggesting distributed ledger technology (DLT) as a solution [43]. By embedding trust and transparency, organizations can foster acceptance and integration of AI technologies, ensuring alignment with organizational values and societal expectations.

The TOE perspective also examines sociotechnical dynamics in AI adoption, recognizing the interplay of technological, organizational, and environmental factors. It emphasizes human-centered approaches and ethical considerations throughout the AI lifecycle, navigating ethical dilemmas across sectors and fostering responsible innovation aligned with societal values [74, 75, 76, 73, 37].

## 6.2 Frameworks and Models Aligning with TOE

Frameworks and models aligned with the TOE perspective are crucial for strategic alignment and digital transformation. Dellermann et al.'s taxonomy organizes methods based on task types and learning paradigms, enhancing human-AI collaboration [46]. This taxonomy ensures AI systems complement human capabilities and drive innovation. In cloud service optimization, standardized definitions align with the TOE perspective, supporting business objectives [2].

LLM-based AIOps methods, as highlighted by Zhang et al., offer superior capabilities over traditional approaches, enhancing technological capabilities and supporting organizational resilience [77]. Alowais et al. emphasize AI's role in enhancing clinical decision-making, aligning AI technologies with organizational processes for strategic objectives [68]. Addressing research fragmentation, as noted by Wilkens, is essential for cohesive development and strategic alignment [74].

AI certification frameworks categorize programs based on governance potential and ethical alignment [75]. Cultural influences on AI systems enhance understanding of technology's cultural dimensions, ensuring innovations are culturally sensitive and strategically aligned [78]. These frameworks highlight the need for harmonizing technological capabilities with organizational structures and environmental contexts, facilitating strategic coherence and digital transformation. Collaborative ecosystems, dynamic regulation, and human-centered design approaches ensure technology development aligns with user needs and ethical considerations [12, 79, 37, 35, 6].

## 6.3 Technological Capabilities and Organizational Goals

Aligning technological capabilities with organizational goals is essential for strategic governance, integrating insights from diverse sectors and anticipating challenges posed by AI and ML. This alignment fosters innovation and enhances operational efficiency. Human-Centered AI (HCAI) capabilities in manufacturing exemplify technological advancements driving innovation while aligning with strategic objectives [80]. Cross-functional collaboration integrates responsible AI practices into organizational cultures, ensuring alignment with organizational values [37].

Ethical frameworks for technology implementation in human resources address limitations and explore HR's evolving role in a tech-driven landscape, fostering a culture of innovation and ethical responsibility [81]. AI optimization in airline reservation systems illustrates the importance of aligning technological capabilities with organizational goals, reducing latency and enhancing efficiency [82].

Future research should develop governance frameworks addressing AI's benefits and risks, ensuring alignment with organizational goals and societal expectations. This includes regulatory frameworks accommodating AI innovations, enhancing data privacy, and educational programs bridging energy and AI disciplines [83]. In education, frameworks ensure responsible AI use aligns with educational goals and ethical standards [84]. Addressing biased data and algorithm accountability is crucial for alignment [26].

Yu et al.'s theory on collective learning and decentralized intelligence, facilitated by blockchain, aligns technological capabilities with organizational goals [16]. This is relevant in federated learning, where research focuses on improving algorithm efficiency for edge devices [13]. Schulz et al. emphasize aligning ML and AI with educational goals in engineering curricula, integrating these technologies into educational frameworks [17]. Weigang outlines gaps in integrating AI theories with practical applications, informing alignment with organizational goals [7]. Addressing these gaps and fostering collaboration enables organizations to harness technological capabilities for strategic objectives, enhancing transparency and resource optimization.

# 7 Digital Transformation

## 7.1 Strategic Alignment in Digital Transformation

Strategic alignment is crucial for effective digital transformation, harmonizing technological advancements with organizational goals to ensure both efficacy and sustainability. The integration of AI and ML into business operations exemplifies this alignment by fostering innovation and improving operational efficiency [37]. Achieving strategic alignment requires a comprehensive approach that includes governance frameworks, ethical considerations, and stakeholder engagement. Organizations must develop governance structures that facilitate AI integration while ensuring adherence to regulatory standards and ethical principles [26]. This involves creating policies that address data privacy, algorithmic accountability, and transparency to build stakeholder trust.

A human-centered design approach is vital in aligning technological solutions with user needs, thereby enhancing user experiences [80]. Involving end-users in the design and implementation phases ensures that technological capabilities align with user expectations and organizational objectives. Culturally, organizations need to foster a mindset of innovation and adaptability to overcome digital transformation challenges. This includes promoting continuous learning and collaboration, enabling teams to leverage digital technologies effectively in pursuit of strategic goals [81]. Investment in skill development is also essential, providing training and educational programs that equip employees with the competencies needed to engage with advanced technologies, thereby aligning human capital with technological capabilities [17].

## 7.2 Interdisciplinary Approaches and Ethical Considerations

Interdisciplinary approaches are essential in digital transformation, integrating diverse perspectives to enhance the development and implementation of digital technologies. The convergence of computer science, social sciences, and ethics is crucial for addressing complex challenges, ensuring technological advancements are robust and socially responsible. This is especially relevant in sectors like healthcare and finance, where frameworks must address AI implications, emphasizing principles such as beneficence, non-maleficence, and justice in AI development and deployment [25, 43, 73, 37, 26].

A primary focus of these interdisciplinary approaches is the development of human-interpretable AI systems. Future research should aim to create AI systems whose outputs are interpretable by humans, thereby enhancing transparency and trustworthiness in AI decisions [85]. Aligning AI outputs with human cognitive processes is critical for fostering acceptance and trust in these technologies, which is essential for successful digital transformation. Ethical considerations are paramount, addressing the societal impacts and moral implications of advanced technology deployment. Integrating domain knowledge into AI systems enhances interpretability and aligns AI outputs with established concepts, such as financial principles, thereby increasing trust and ensuring ethical operation [39]. By embedding ethical considerations into the design and implementation of digital technologies, organizations can mitigate risks and ensure their digital transformation efforts yield positive societal contributions.

Interdisciplinary approaches and ethical considerations are also crucial for developing governance frameworks that guide responsible digital technology use. These frameworks must address data privacy, algorithmic accountability, and transparency, ensuring adherence to ethical principles and reflecting societal values. Recent discussions advocate for practical approaches to AI ethics, moving from theoretical principles to actionable practices. As AI increasingly integrates into sectors like healthcare, finance, and public services, addressing the complexities of implementing fairness, trustworthiness, and interpretability in AI systems becomes imperative. Balancing individual privacy protection with societal benefits, such as life-saving medical advancements, necessitates robust governance mechanisms that facilitate ethical AI deployment while mitigating risks. This is vital for fostering public trust and maximizing the positive impact of digital transformation [73, 76, 86]. By promoting interdisciplinary collaboration and embedding ethical considerations into digital transformation strategies, organizations can effectively navigate the complexities of the digital landscape and achieve sustainable, responsible innovation.

# 8 Risk Management

## 8.1 AI and Data Analytics in Risk Management

AI and data analytics are integral to modern risk management, offering advanced tools for identifying, assessing, and mitigating risks across various sectors. These technologies enhance decision-making by analyzing large datasets, as demonstrated in financial risk management, where AI systems like CBR_E improve interpretability and predictive accuracy [38]. However, there is a notable reliance on AI recommendations even when accuracy is compromised, highlighting AI's role in risk management during data exploration [44]. In Cyber-Physical Systems (CPS), AI controllers sometimes outperform traditional ones, yet they can falter in complex situations, necessitating robust evaluation metrics [1].

The Human Rights Impact Assessment (HRIA) methodology is crucial for identifying and mitigating human rights risks linked to AI, ensuring ethical deployment and societal benefit [34]. AI models like Climate AI (CAI) play a significant role in sustainability-related risk management by providing structured data on corporate carbon commitments [3]. In emergency management, AI systems such as E-KELL enhance decision support, surpassing baseline models in accuracy and standards adherence [40]. In healthcare, Multi-modal Large Language Models (MLLMs) and hybrid Retrieval-Augmented Generation (RAG) techniques facilitate secure data sharing and improve risk management [5].

AI and data analytics also revolutionize cybersecurity training by offering personalized, immersive experiences that enhance risk management capabilities [10]. In software development, while current AI tools provide basic risk management support, future systems like HyperAssistant are anticipated to offer advanced, context-aware assistance [9]. The integration of these technologies into risk management frameworks offers transformative opportunities for organizations to better identify and mitigate risks. Recent advancements, particularly in large language models, may reshape enterprise knowledge access but introduce new risks related to worker value and power dynamics. A comprehensive governance framework is essential to address these risks, emphasizing understanding AI systems' operational mechanisms and moral implications, including commodification and power concentration. Organizations must align their culture and structure with responsible AI initiatives to ensure effective governance and accountability. As AI capabilities evolve, particularly in autonomous systems, it is crucial for organizations to develop adaptive governance mechanisms to manage the extreme risks associated with advanced AI technologies [87, 37, 49]. Leveraging these technologies can enhance resilience and adaptability in risk management, ensuring effective and ethical decision-making across diverse sectors.

## 8.2 Sector-Specific Risk Management Strategies

In healthcare, AI and data analytics improve risk management by enhancing patient safety and operational efficiency. Technologies like MLLMs and hybrid RAG techniques enable secure data management, mitigating risks of data breaches and ensuring regulatory compliance [5]. These innovations allow healthcare providers to proactively identify risks, streamline workflows, and improve patient outcomes.

In finance, AI-driven systems enhance the interpretability and predictive accuracy of risk assessments. Systems like CBR_E offer transparent, data-driven insights into financial threats, aiding institutions in making informed decisions and reducing risk exposure [38].

Manufacturing benefits from predictive maintenance strategies using AI and data analytics, which employ IoT devices and real-time analytics to monitor equipment performance and predict failures. This proactive approach minimizes operational downtime and boosts production efficiency, addressing challenges in service latency and failure management through innovative AIOps strategies [77, 71, 2].

In the energy sector, AI applications enhance power grid operations through supervised and unsupervised learning techniques, optimizing decision-making and energy distribution [1].

Cybersecurity employs innovative training methods, including AI and virtual reality (VR), to bolster resilience against cyber threats. These immersive training experiences provide personalized learning, enabling organizations to effectively manage cybersecurity risks [10].

### 8.3 Governance Frameworks and Ethical Considerations

Establishing governance frameworks and ethical considerations is vital for addressing the challenges of AI integration in risk management. In bioengineering, significant AI-related risks necessitate updated governance frameworks to mitigate these risks [88]. Translating ethical principles into practical governance measures remains a complex task in AI development [89].

Ethical issues such as data privacy, bias, accountability, and transparency are central to risk management discourse. Comprehensive governance frameworks must address these concerns to ensure responsible AI deployment [24]. In AI risk management, emphasizing complementary metrics alongside training compute provides a holistic approach to regulatory oversight and ethical governance [90].

In education, AI raises specific ethical concerns regarding data privacy and the implications of deploying AI technologies in learning environments, underscoring the need for robust governance frameworks to safeguard student privacy while promoting ethical AI use [84].

Developing robust governance frameworks and ethical considerations in AI risk management is crucial for responsible deployment across sectors, including healthcare, finance, and policing. This approach is essential for mitigating potential harms and maximizing societal benefits. Effective governance involves addressing empirical questions about AI risks and benefits while navigating normative challenges to align AI practices with societal values such as privacy, fairness, and autonomy. Developing AI certification programs can enhance ethical practices by ensuring the implementation of these principles, fostering trust and accountability in AI systems. Advancing ethical AI governance is vital for promoting innovation while safeguarding individual and community interests [73, 89, 75]. Addressing these challenges enables organizations to foster trust and accountability in AI systems, enhancing their effectiveness and societal acceptance.

## 9 Conclusion

This survey underscores the transformative impact of integrating artificial intelligence (AI), big data, and supply chain resilience, highlighting their role in enhancing organizational efficiency and adaptability across diverse sectors. By aligning with the Technology-Organization-Environment (TOE) framework, AI technologies facilitate intelligent decision-making processes that synchronize technological capabilities with organizational objectives. The synergy between AI and big data analytics provides robust tools for deriving actionable insights, essential for strategic planning and informed decision-making.

Incorporating a Human-Centered AI (HCAI) approach is particularly crucial in sectors like manufacturing, where it fosters innovation and sustainable practices. This aligns with the broader goal of enhancing human capabilities through HCAI principles, ensuring a seamless transition to advanced technological frameworks. The survey emphasizes the importance of early and continuous engagement in information governance throughout the AI lifecycle, highlighting the need for stakeholder collaboration to address the complexities of information governance.

Risk management is identified as a pivotal aspect, with AI and data analytics offering sophisticated tools for threat identification and mitigation. The development of comprehensive AI policy frameworks and the exploration of ethical implications are vital to ensuring AI technologies contribute positively to societal well-being. Future research should focus on creating standardized datasets and



AI-generated, for reference only.

preprocessing pipelines to improve machine learning applications, thereby enhancing the adaptability of emergency decision-making models.

The survey advocates for coordinated policy approaches across institutions, suggesting future research directions aimed at strengthening data governance frameworks, developing interpretable AI models, and examining the ethical dimensions of AI deployment. Additionally, optimizing resource allocation strategies and exploring the integration of AI with blockchain technologies are recommended for further advancements.

The promising integration of blockchain and federated learning demonstrates high classification accuracy in medical datasets while preserving data privacy. Future research should explore the individual components of decentralized AI, encourage industry-academic collaborations, and pursue practical implementations of decentralized AI solutions. The integration of Blockchain 3.0 with AI is highlighted as a means to enhance the efficiency and security of e-government services, suggesting future research directions that expand the integration of additional Web 3.0 technologies.

## References


[1] Jiayang Song, Deyun Lyu, Zhenya Zhang, Zhijie Wang, Tianyi Zhang, and Lei Ma. When cyber-physical systems meet ai: A benchmark, an evaluation, and a way forward, 2022.

[2] Cuixia Li, Liqiang Liu, and Li Shi. Review of cloud service composition for intelligent manufacturing, 2024.

[3] Aditya Dave, Mengchen Zhu, Dapeng Hu, and Sachin Tiwari. Climate ai for corporate decarbonization metrics extraction, 2024.

[4] Ben Ortiz, Laura Kahn, Marc Bosch, Philip Bogden, Viveca Pavon-Harr, Onur Savas, and Ian McCulloh. Improving community resiliency and emergency response with artificial intelligence, 2020.

[5] Cheng Su, Jinbo Wen, Jiawen Kang, Yonghua Wang, Yuanjia Su, Hudan Pan, Zishao Zhong, and M. Shamim Hossain. Hybrid rag-empowered multi-modal llm for secure data management in internet of medical things: A diffusion-based contract approach, 2024.

[6] Dian Tjondronegoro, Elizabeth Yuwono, Brent Richards, Damian Green, and Siiri Hatakka. Responsible ai implementation: A human-centered framework for accelerating the innovation process. *arXiv preprint arXiv:2209.07076*, 2022.

[7] Li Weigang, Liriam Enamoto, Denise Leyi Li, and Geraldo Pereira Rocha Filho. Watershed of artificial intelligence: Human intelligence, machine intelligence, and biological intelligence, 2021.

[8] Dingzhu Wen, Xiaoyang Li, Qunsong Zeng, Jinke Ren, and Kaibin Huang. An overview of data-importance aware radio resource management for edge machine learning, 2019.

[9] Ketai Qiu, Niccolò Puccinelli, Matteo Ciniselli, and Luca Di Grazia. From today's code to tomorrow's symphony: The ai transformation of developer's routine by 2030, 2024.

[10] Saif Al-Dean Qawasmeh, Ali Abdullah S. AlQahtani, and Muhammad Khurram Khan. Navigating cybersecurity training: A comprehensive review, 2024.

[11] Stefanie Krause, Bhumi Hitesh Panchal, and Nikhil Ubhe. The evolution of learning: Assessing the transformative impact of generative ai on higher education, 2024.

[12] Mark Fenwick, Erik P. M. Vermeulen, and Marcelo Corrales Compagnucci. Business and regulatory responses to artificial intelligence: Dynamic regulation, innovation ecosystems and the strategic management of disruptive technology, 2024.

[13] Shuiguang Deng, Hailiang Zhao, Weijia Fang, Jianwei Yin, Schahram Dustdar, and Albert Y. Zomaya. Edge intelligence: The confluence of edge computing and artificial intelligence, 2020.

[14] Feras A. Batarseh and Munisamy Gopinath. Panel: Economic policy and governance during pandemics using ai, 2020.

[15] Francesca Larosa, Jaroslav Mysiak, Marco Molinari, Panagiotis Varelas, Haluk Akay, Will McDowall, Catalina Spadaru, Francesco Fuso-Nerini, and Ricardo Vinuesa. Closing the gap between research and projects in climate change innovation in europe, 2023.

[16] F. Richard Yu. From the internet of information to the internet of intelligence, 2019.

[17] Andrew Schulz, Suzanne Stathatos, Cassandra Shriver, and Roxanne Moore. Utilizing online and open-source machine learning toolkits to leverage the future of sustainable engineering, 2023.

[18] Yifan Zhang, Xinkui Zhao, Ziying Li, Jianwei Yin, Lufei Zhang, and Zuoning Chen. Integrating artificial intelligence into operating systems: A comprehensive survey on techniques, applications, and future directions, 2024.

[19] Sven Eckhardt, Niklas Kühl, Mateusz Dolata, and Gerhard Schwabe. A survey of ai reliance, 2024.

[20] Mohamed El Louadi. On the preservation of africa's cultural heritage in the age of artificial intelligence, 2024.

[21] Roozbeh Aliabadi, Aditi Singh, and Eryka Wilson. Transdisciplinary ai education: The confluence of curricular and community needs in the instruction of artificial intelligence, 2023.

[22] Kerstin Hötte, Taheya Tarannum, Vilhelm Verendel, and Lauren Bennett. Measuring artificial intelligence: a systematic assessment and implications for governance, 2024.

[23] Jianjun Zhu, Fan Li, and Jinyuan Chen. A survey of blockchain, artificial intelligence, and edge computing for web 3.0, 2023.

[24] Hamid Reza Saeidnia, Elaheh Hosseini, Shadi Abdoli, and Marcel Ausloos. Unleashing the power of ai. a systematic review of cutting-edge techniques in ai-enhanced scientometrics, webometrics, and bibliometrics, 2024.

[25] Ali Akbar Septiandri, Marios Constantinides, and Daniele Quercia. The impact of responsible ai research on innovation and development, 2024.

[26] Alexis Tsoukiàs. Social responsibility of algorithms, 2020.

[27] Sébastien Bubeck, Varun Chandrasekaran, Ronen Eldan, Johannes Gehrke, Eric Horvitz, Ece Kamar, Peter Lee, Yin Tat Lee, Yuanzhi Li, Scott Lundberg, Harsha Nori, Hamid Palangi, Marco Tulio Ribeiro, and Yi Zhang. Sparks of artificial general intelligence: Early experiments with gpt-4, 2023.

[28] X. San Liang, Dake Chen, and Renhe Zhang. Quantitative causality, causality-guided scientific discovery, and causal machine learning, 2024.

[29] Conghui He, Wei Li, Zhenjiang Jin, Chao Xu, Bin Wang, and Dahua Lin. Opendatalab: Empowering general artificial intelligence with open datasets, 2024.

[30] Hiromitsu Goto, Wataru Souma, Mari Jibu, and Yuichi Ikeda. Multilayer network analysis of the drug pipeline in the global pharmaceutical industry, 2020.

[31] Antoine J. P. Tixier and Matthew R. Hallowell. Safer together: Machine learning models trained on shared accident datasets predict construction injuries better than company-specific models, 2023.

[32] Camille Grange, Theophile Demazure, Mickael Ringeval, Simon Bourdeau, and Cedric Martineau. The human-genai value loop in human-centered innovation: Beyond the magical narrative, 2025.

[33] Neil D. Lawrence. Data science and digital systems: The 3ds of machine learning systems design, 2019.

[34] Alessandro Mantelero and Maria Samantha Esposito. An evidence-based methodology for human rights impact assessment (hria) in the development of ai data-intensive systems, 2024.

[35] Vincent J. Straub, Deborah Morgan, Youmna Hashem, John Francis, Saba Esnaashari, and Jonathan Bright. A multidomain relational framework to guide institutional ai research and adoption, 2023.

[36] Paul Bilokon, Oleksandr Bilokon, and Saeed Amen. A compendium of data sources for data science, machine learning, and artificial intelligence, 2023.

[37] Bogdana Rakova, Jingying Yang, Henriette Cramer, and Rumman Chowdhury. Where responsible ai meets reality: Practitioner perspectives on enablers for shifting organizational practices, 2021.

[38] Wei Li, Florentina Paraschiv, and Georgios Sermpinis. A data-driven explainable case-based reasoning approach for financial risk detection, 2021.

[39] Sheikh Rabiul Islam, William Eberle, Sid Bundy, and Sheikh Khaled Ghafoor. Infusing domain knowledge in ai-based "black box" models for better explainability with application in bankruptcy prediction, 2019.

[40] Minze Chen, Zhenxiang Tao, Weitong Tang, Tingxin Qin, Rui Yang, and Chunli Zhu. Enhancing emergency decision-making with knowledge graphs and large language models, 2023.

[41] Gali Noti and Yiling Chen. Learning when to advise human decision makers, 2023.

[42] Julian Senoner, Simon Schallmoser, Bernhard Kratzwald, Stefan Feuerriegel, and Torbjørn Netland. Explainable ai improves task performance in human-ai collaboration, 2024.

[43] Scott Thiebes, Sebastian Lins, and Ali Sunyaev. Trustworthy artificial intelligence. *Electronic Markets*, 31:447–464, 2021.

[44] Sunwoo Ha, Shayan Monadjemi, and Alvitta Ottley. Guided by ai: Navigating trust, bias, and data exploration in ai-guided visual analytics, 2024.

[45] Valérie Beaudouin, Isabelle Bloch, David Bounie, Stéphan Clémençon, Florence d'Alché Buc, James Eagan, Winston Maxwell, Pavlo Mozharovskyi, and Jayneel Parekh. Flexible and context-specific ai explainability: A multidisciplinary approach, 2020.

[46] Dominik Dellermann, Adrian Calma, Nikolaus Lipusch, Thorsten Weber, Sascha Weigel, and Philipp Ebel. The future of human-ai collaboration: a taxonomy of design knowledge for hybrid intelligence systems, 2021.

[47] Jose Hernandez-Orallo. Ai evaluation: past, present and future, 2016.

[48] Benedikt Kantz, Clemens Staudinger, Christoph Feilmayr, Johannes Wachlmayr, Alexander Haberl, Stefan Schuster, and Franz Pernkopf. Robustness of explainable artificial intelligence in industrial process modelling, 2024.

[49] Anna Gausen, Bhaskar Mitra, and Siân Lindley. A framework for exploring the consequences of ai-mediated enterprise knowledge access and identifying risks to workers, 2024.

[50] Yongjun Zhang. Generative ai has lowered the barriers to computational social sciences, 2023.

[51] Yijie Weng, Jianhao Wu, Tara Kelly, and William Johnson. Comprehensive overview of artificial intelligence applications in modern industries, 2024.

[52] Jonathan Bright, Florence E. Enock, Saba Esnaashari, John Francis, Youmna Hashem, and Deborah Morgan. Generative ai is already widespread in the public sector, 2024.

[53] Luca Longo, Mario Brcic, Federico Cabitza, Jaesik Choi, Roberto Confalonieri, Javier Del Ser, Riccardo Guidotti, Yoichi Hayashi, Francisco Herrera, Andreas Holzinger, Richard Jiang, Hassan Khosravi, Freddy Lecue, Gianclaudio Malgieri, Andrés Páez, Wojciech Samek, Johannes Schneider, Timo Speith, and Simone Stumpf. Explainable artificial intelligence (xai) 2.0: A manifesto of open challenges and interdisciplinary research directions, 2023.

[54] Hao Wu. What is learning? a primary discussion about information and representation, 2015.

[55] Kari D. Weaver. The artificial intelligence disclosure (aid) framework: An introduction, 2024.

[56] Alex Mei, Michael Saxon, Shiyu Chang, Zachary C. Lipton, and William Yang Wang. Users are the north star for ai transparency, 2023.

[57] Feng Liu, Yong Shi, and Ying Liu. Intelligence quotient and intelligence grade of artificial intelligence, 2017.

[58] Upol Ehsan and Mark O Riedl. Human-centered explainable ai: Towards a reflective sociotechnical approach. In *HCI International 2020-Late Breaking Papers: Multimodality and Intelligence: 22nd HCI International Conference, HCII 2020, Copenhagen, Denmark, July 19–24, 2020, Proceedings 22*, pages 449–466. Springer, 2020.

[59] Mattias Tiger, Daniel Jakobsson, Anders Ynnerman, Fredrik Heintz, and Daniel Jönsson. Exploratory visual analysis for increasing data readiness in artificial intelligence projects, 2024.

[60] Naweiluo Zhou, Florent Dufour, Vinzent Bode, Peter Zinterhof, Nicolay J Hammer, and Dieter Kranzlmüller. Towards confidential computing: A secure cloud architecture for big data analytics and ai, 2023.

[61] Zehang Deng, Ruoxi Sun, Minhui Xue, Sheng Wen, Seyit Camtepe, Surya Nepal, and Yang Xiang. Leakage-resilient and carbon-neutral aggregation featuring the federated ai-enabled critical infrastructure, 2024.

[62] Polra Victor Falade. Cyber security requirements for platforms enhancing ai reproducibility, 2023.

[63] Mohammadhossein Ghahramani, Yan Qiao, MengChu Zhou, Adrian OHagan, and James Sweeney. Ai-based modeling and data-driven evaluation for smart manufacturing processes, 2020.

[64] Andreas Holzinger, Janine Schweier, Christoph Gollob, Arne Nothdurft, Hubert Hasenauer, Thomas Kirisits, Carola Häggström, Rien Visser, Raffaele Cavalli, Raffaele Spinelli, et al. From industry 5.0 to forestry 5.0: Bridging the gap with human-centered artificial intelligence. *Current Forestry Reports*, pages 1–14, 2024.

[65] Vincenzo Barbuto, Claudio Savaglio, Roberto Minerva, Noel Crespi, and Giancarlo Fortino. Towards an edge intelligence-based traffic monitoring system, 2024.

[66] Elisa Bertino and Sujata Banerjee. Artificial intelligence at the edge, 2020.

[67] Amit Sheth, Utkarshani Jaimini, and Hong Yung Yip. How will the internet of things enable augmented personalized health?, 2017.

[68] Shuroug A Alowais, Sahar S Alghamdi, Nada Alsuhebany, Tariq Alqahtani, Abdulrahman I Alshaya, Sumaya N Almohareb, Atheer Aldairem, Mohammed Alrashed, Khalid Bin Saleh, Hisham A Badreldin, et al. Revolutionizing healthcare: the role of artificial intelligence in clinical practice. *BMC medical education*, 23(1):689, 2023.

[69] Yunling Shi, Jie Guan, Junfeng Xiao, Huai Zhang, Qiang Guo, and Yu Yuan. Internet of rights(ior) in role based block chain, 2022.

[70] Christopher Gorham. The 4th industrial revolution effect on the enterprise cyber strategy, 2020.

[71] Huanzhuo Wu, Jia He, Máté Tömösközi, Zuo Xiang, and Frank H. P. Fitzek. In-network processing for low-latency industrial anomaly detection in softwarized networks, 2021.

[72] Neel Kanwal, Emiel A. M. Janssen, and Kjersti Engan. Balancing privacy and progress in artificial intelligence: Anonymization in histopathology for biomedical research and education, 2023.

[73] Jess Whittlestone and Sam Clarke. Ai challenges for society and ethics, 2022.

[74] Uta Wilkens, Christian Cost Reyes, Tim Treude, and Annette Kluge. Understandings and perspectives of human-centered ai–a transdisciplinary literature review. *GfA Frühjahrskongress, B*, 10, 2021.

[75] Peter Cihon, Moritz J. Kleinaltenkamp, Jonas Schuett, and Seth D. Baum. Ai certification: Advancing ethical practice by reducing information asymmetries, 2021.

[76] Jessica Morley, Luciano Floridi, Libby Kinsey, and Anat Elhalal. From what to how: An initial review of publicly available ai ethics tools, methods and research to translate principles into practices, 2019.

[77] Lingzhe Zhang, Tong Jia, Mengxi Jia, Yifan Wu, Aiwei Liu, Yong Yang, Zhonghai Wu, Xuming Hu, Philip S. Yu, and Ying Li. A survey of aiops for failure management in the era of large language models, 2024.






AI-generated, for reference only.

[78] Vinodkumar Prabhakaran, Rida Qadri, and Ben Hutchinson. Cultural incongruencies in artificial intelligence, 2022.

[79] Leoncio Jimenez and Francisco Venegas. Sociotechnical approach to enterprise generative artificial intelligence (e-genai), 2025.

[80] Barbara Martini, Denise Bellisario, and Paola Coletti. Human-centered and sustainable artificial intelligence in industry 5.0: Challenges and perspectives. *Sustainability*, 16(13):5448, 2024.

[81] Omer Aydin, Enis Karaarslan, and Nida Gokce Narin. Artificial intelligence, vr, ar and metaverse technologies for human resources management, 2024.

[82] Biman Barua and M. Shamim Kaiser. Optimizing airline reservation systems with edge-enabled microservices: A framework for real-time data processing and enhanced user responsiveness, 2024.

[83] Le Xie, Tong Huang, Xiangtian Zheng, Yan Liu, Mengdi Wang, Vijay Vittal, P. R. Kumar, Srinivas Shakkottai, and Yi Cui. Energy system digitization in the era of ai: A three-layered approach towards carbon neutrality, 2022.

[84] Vahid Ashrafimoghari, Necdet Gürkan, and Jordan W. Suchow. Evaluating large language models on the gmat: Implications for the future of business education, 2024.

[85] Adam Perrett. Emergent explainability: Adding a causal chain to neural network inference, 2024.

[86] Carlos Ferreira. A short review of the main concerns in a.i. development and application within the public sector supported by nlp and tm, 2023.

[87] Yoshua Bengio, Geoffrey Hinton, Andrew Yao, Dawn Song, Pieter Abbeel, Trevor Darrell, Yuval Noah Harari, Ya-Qin Zhang, Lan Xue, Shai Shalev-Shwartz, Gillian Hadfield, Jeff Clune, Tegan Maharaj, Frank Hutter, Atılım Güneş Baydin, Sheila McIlraith, Qiqi Gao, Ashwin Acharya, David Krueger, Anca Dragan, Philip Torr, Stuart Russell, Daniel Kahneman, Jan Brauner, and Sören Mindermann. Managing extreme ai risks amid rapid progress, 2024.

[88] Matthew E. Walsh. Towards risk analysis of the impact of ai on the deliberate biological threat landscape, 2024.

[89] Mamia Agbese, Hanna-Kaisa Alanen, Jani Antikainen, Erika Halme, Hannakaisa Isomäki, Marianna Jantunen, Kai-Kristian Kemell, Rebekah Rousi, Heidi Vainio-Pekka, and Ville Vakkuri. Governance of ethical and trustworthy ai systems: Research gaps in the eccola method, 2021.

[90] Lennart Heim and Leonie Koessler. Training compute thresholds: Features and functions in ai regulation, 2024.

**Disclaimer:**

SurveyX is an AI-powered system designed to automate the generation of surveys. While it aims to produce high-quality, coherent, and comprehensive surveys with accurate citations, the final output is derived from the AI's synthesis of pre-processed materials, which may contain limitations or inaccuracies. As such, the generated content should not be used for academic publication or formal submissions and must be independently reviewed and verified. The developers of SurveyX do not assume responsibility for any errors or consequences arising from the use of the generated surveys.