

Chapter 8 - AI Agents and Sustainable Development Goals: Bridging the Gap Through Innovation

Imran Chughtai ¹

¹ SBS Swiss Business School, RAK campus, UAE

Chapter Information

- **Date of Submission:** [22/01/2025]
- **Date of Acceptance:** [28/02/2025]
- **JEL Classification Codes:** Q01, O33, Q56;

Abstract

Autonomous AI agents represent an innovative advancement in artificial intelligence, transitioning from AI copilots and assistants reliant on human input to independent operators. The reasoning capabilities of large language models enable agents to develop strategies, plan and then execute tasks autonomously to achieve high-level goals. When finding problems, they can solve them automatically by utilizing resources such as the Internet, numerous data sources and specialized applications, such as delegating tasks to other specialized AI agents. They independently research problems and adapt without human oversight, presenting transformative potential for Sustainable Development Goals (SDGs) through scalable solutions to global challenges. The autonomous application of AI potentially improves many areas such as water management, agriculture, and biodiversity conservation, thereby optimizing decision-making processes for greater efficiency. Agentic AI possesses the capability to analyze extensive datasets, systematically identify gaps in sustainability efforts, and model potential intervention outcomes. This analytical capacity directly aligns with the goals of environmental conservation, resource efficiency, and equitable access to essential services. However, challenges, including environmental costs, energy consumption, and algorithmic biases, threaten fairness and inclusivity, which are important for realizing AI agents' potential for achieving the goals. Establishing robust ethical frameworks is essential for guiding the deployment of AI agents, ensuring transparency and accountability. By embedding principles of fairness and inclusivity, autonomous AI agents can facilitate responsible advancements toward the Sustainable Development Goals.

Keywords: *AI Agents, Sustainability, SDG, Ethical, multi-agent, agentic AI, opportunities;*

8.1. Introduction

AI has evolved from a niche field into a transformative force across all aspects of life including healthcare, finance, and education. In 2023, researchers identified over 600 AI solutions advancing the 17 UN Sustainable Development Goals (SDGs), with 492 already deployed, evidence of AI's growing role in global sustainability (Bankhwal et al., 2024). The most significant impact was observed in SDG 3 (Good Health and Well-being), with 165 AI-enabled healthcare use cases identified, with 128 active deployments, reinforcing AI's transformative potential in early disease detection, predictive diagnostics, and telemedicine solutions (Bankhwal et al., 2024; Gosselink et al., 2024). AI-driven predictive healthcare models improve diagnostics and patient outcomes, while real-time monitoring systems enhance disease surveillance and response capabilities (see Appendix A). AI agents (agentic AI/multi-agents) operate autonomously, perceiving environments, learning from real-time data, and executing tasks with minimal oversight.

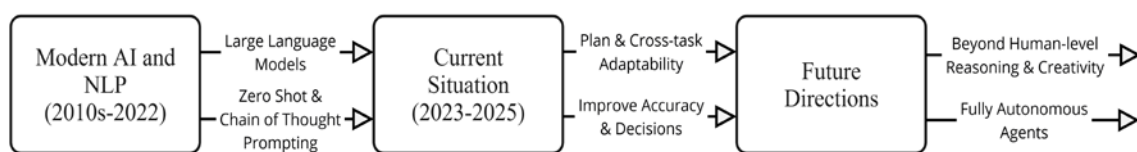


Figure 1: Progression towards AI Agents

Note. Source: Researcher

Utilising advancements (Figure 1) such as large language models (with human-like reasoning), natural language processing (NLP) and reinforcement learning AI Agents interpret human language, adapt to dynamic inputs, and take on unclear goals by researching, planning and delegating tasks (Gosselink et al., 2024; Mollick et al., 2024; Wang et al., 2023).

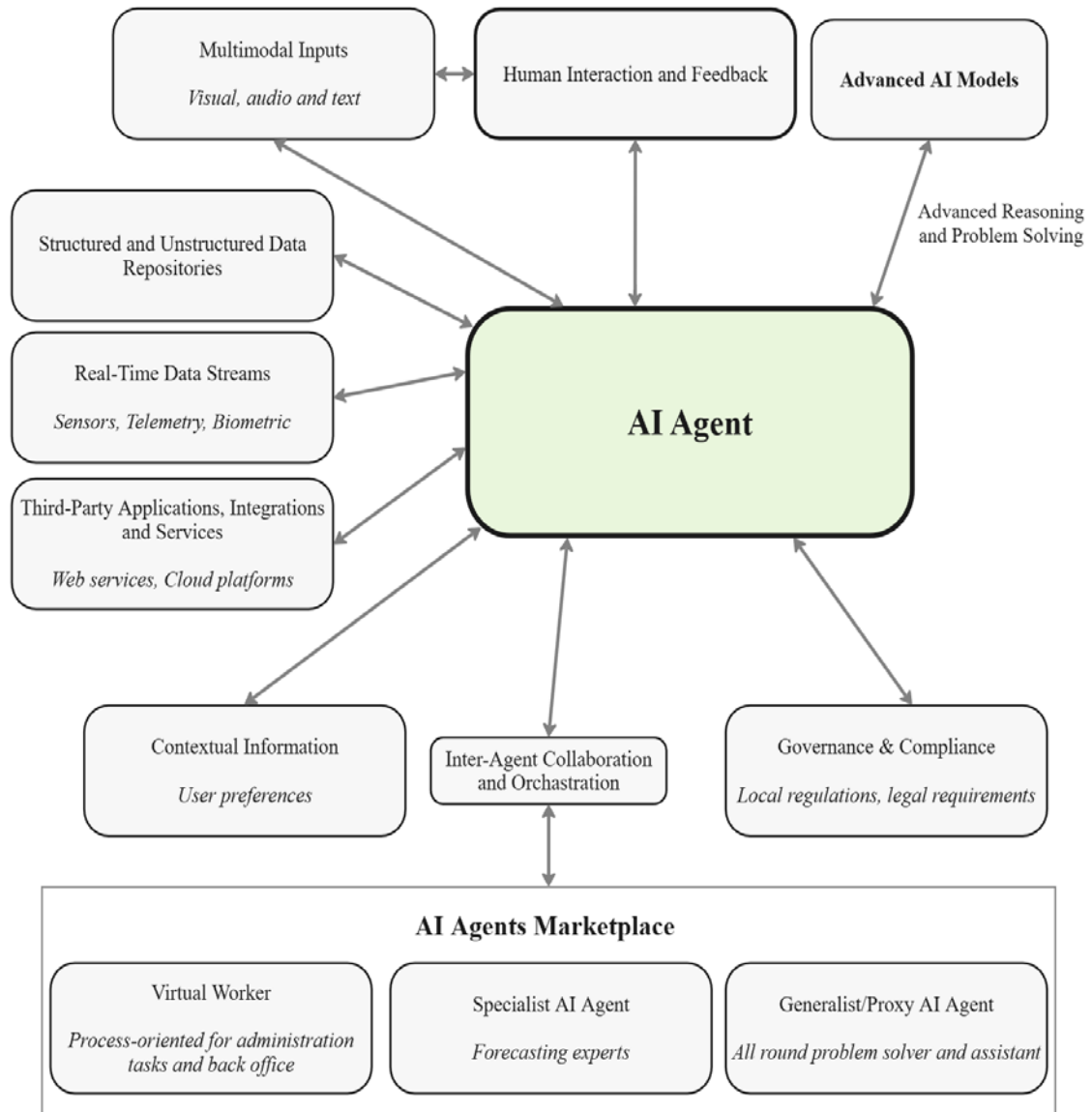


Figure 2: AI Agent: Inputs and Interactions

Note. Source: Researcher

Figure 2 illustrates some key inputs and interactions related to an AI Agent, highlighting its ability to autonomously process multimodal inputs, integrate structured and unstructured datasets, and respond to real-time data streams. They integrate with third-party applications, cloud services, and specialized AI Agents (via marketplaces), enabling interoperable digital ecosystems (Mollick et al., 2024). AI Agents, act as virtual teams that extend human capabilities, autonomously process vast data, apply reasoning, and generate predictive insights to support SDG interventions. Due to the LLM context size, AI limits the storage and memory of the AI Agents. This leads to specialist AI Agents, who have specific skills and memory that they can access, like human expert team members (Wang et al., 2023). They are coordinated by a team leader (Generalist/Proxy AI Agent) who orchestrates, communicates and reviews all work.

Despite technological advancements, only 17% of SDG targets are on track, while 80% have stalled or regressed, largely due to resource limitations, slow policy adaptation, and insufficient data-driven decision-making (Gosselink et al., 2024; United Nations, 2024).

This chapter examines the technological and ethical dimensions of AI Agent deployment in SDGs, addressing the research question:

How can the operational, ethical, and governance challenges of deploying agentic AI for Sustainable Development Goals be effectively addressed?

AI Agents can drive sustainability efforts through automation, large-scale real-time analytics, human-like reasoning, and automated task creation and reviews. However, they face challenges such as bias, governance gaps, and rising resource demands. Addressing these requires a structured approach that ensures a balanced assessment of their potential and limitations in advancing a sustainable and equitable future.

8.2. Literature Review

AI Agents are increasingly recognized as tools in advancing the Sustainable Development Goals (SDGs), with abilities to automatically research a complex problem, weigh up nuanced options and then plan and delegate a set of tasks. However, their effectiveness depends on both technological advancements and frameworks that ensure ethical and transparent deployment and operations. This literature review systematically examines AI Agents' functional capabilities and ethical considerations, highlighting their potential and limitations in addressing sustainability challenges.

AI Agents operate within an interconnected digital ecosystem (see Figure 2), integrating structured and unstructured data sources to facilitate scalable human-like decision-making and actions (Sandini et al., 2024; Zhou et al., 2023).

AI Agents exhibit eight fundamental capabilities in SDG applications:

1. Perception and Social Processing
2. Explainable Cognitive Reasoning
3. Action and Execution
4. Adaptive Learning and Memory
5. Synergistic Collaboration and Coordination
6. Security and Compliance
7. Natural Language Processing and Communication
8. Integration and Implementation

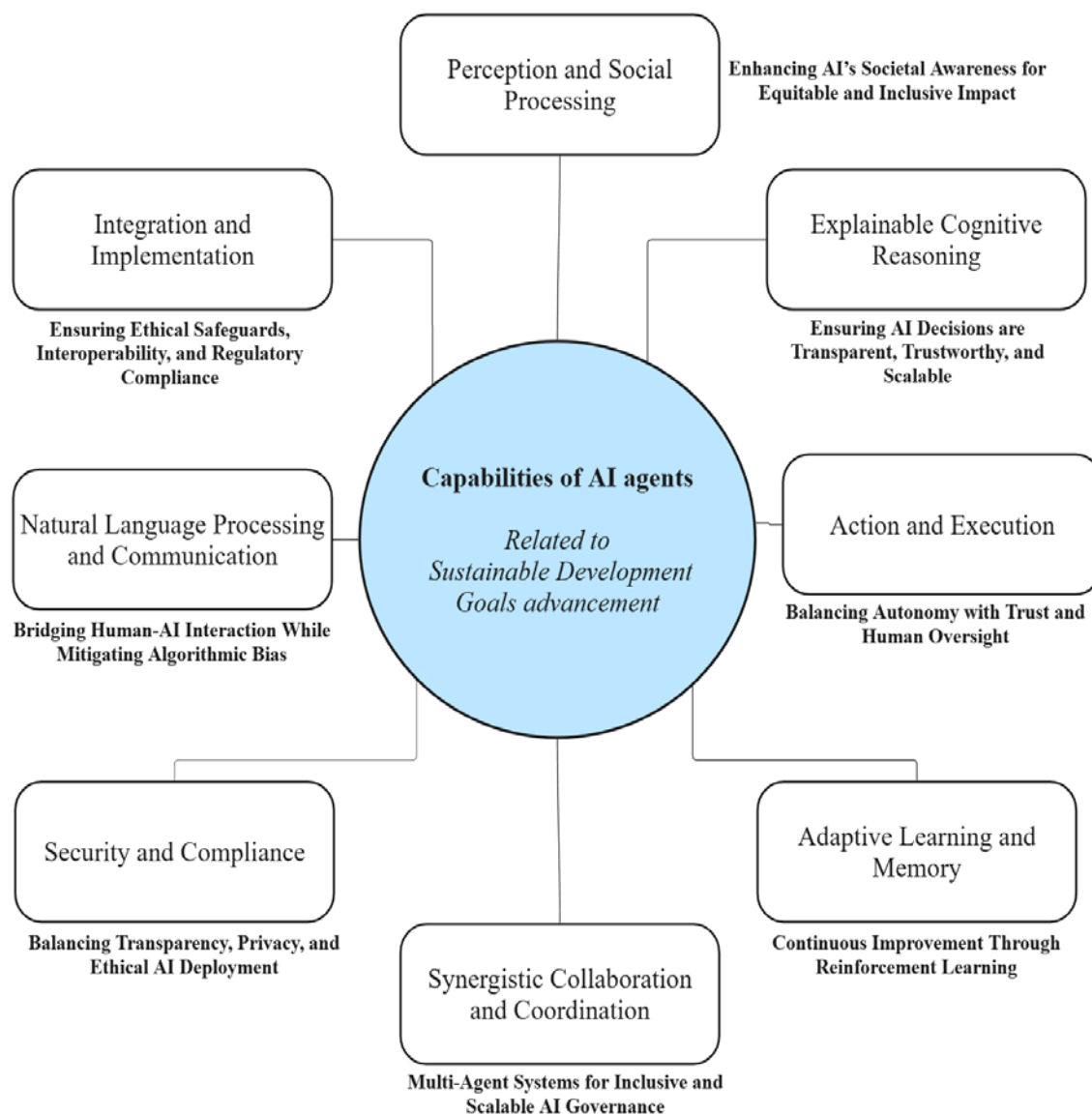


Figure 3: AI Agent Capabilities for SDG Advancement: Key Challenges and Considerations

Note. Source: Researcher

- **Perception and Social Processing** – collects and processes real-time environmental data through IoT sensors, satellite feeds, and user interactions (Sandini et al., 2024). Barros et al. (2023) argue that neglecting social metrics can hinder organisations' contributions to SDGs. AI Agents that lack awareness of rivalry and social cues struggle in competitive environments, which limits their effectiveness in contributing to SDG 9 and SDG 17. By integrating social interactions, these agents can better understand nuances, social influences, political lobbying pressures and hidden agendas. This enhancement fosters inclusive and resilient geopolitical partnerships.

- **Explainable Cognitive Reasoning** – responsible for reasoning, planning, and decision-making, is important for interpreting sustainability data correctly and optimising agentic AI-driven SDG interventions (Ruan et al., 2023). Advancements in LLMs like GPT-4o further enhance AI agents' ability to process unstructured data and generate adaptive solutions (Ruan et al., 2023). Newer models, exemplified by OpenAI o1 and DeepSeek R1, refine reasoning

by assessing feasibility prior to generating responses, thereby improving decision accuracy. This could address complex sustainability challenges, such as climate resilience (SDG 13) and resource efficiency (SDG 12), by identifying root causes of climate vulnerabilities and resource inefficiencies and predicting the consequences of actions and policies.

However, as Sandini et al. (2024) observe, reliance on correlation-based analysis may obscure causal relationships, limiting AI's ability to drive meaningful change. This underscores the need for explainable AI to improve transparency, trust, and accountability in high-stakes SDG applications, including healthcare equity (SDG 3) and sustainable urban planning (SDG 11). Sastry et al. (2024) research found that as AI cognitive capabilities scale, the training of these systems often requires substantial computational resources, which in turn necessitates considerable energy usage. This aspect raises concerns about the environmental impact associated with the expansion of AI technologies, particularly in terms of their carbon footprint and overall sustainability (Ueda et al., 2024). However, DeepSeek R1 has proven that a highly capable open-source LLM with advanced reasoning can be built with fewer resources.

- **Action and Execution** – Gronauer & Diepold (2022) research supports the execution of tasks or communication interventions (e.g., triggering alerts for disaster relief under SDG 13). AI Agents can now use cognitive abilities to research a complex problem, figure out a plan, subdivide into tasks and test out different actions, expecting failure to occur to replan (Kim et al., 2024; Wang et al., 2023; Xu et al., 2023). However, Lawless (2024) highlights the importance of trust and safety, noting that public scepticism towards AI systems can hinder their effectiveness. Gronauer & Diepold (2022) point out difficulties in coordination among multiple AI agents, which can lead to inefficiencies in resource allocation during collaborative efforts, such as managing healthcare resources. At the same time, autonomous decision-making enables AI to act independently, though a lack of trust in full autonomy (Jannelli et al., 2024) necessitates human oversight. Sekiguchi & Ohsawa (2024) state that AI should only augment human capabilities, while Jannelli et al. (2024) argue that AI agents can autonomously make correct decisions without requiring human intervention. Kahungi (2023) warns that research from Africa shows that AI has been perpetuating bias due to the data available in systems, at scale this will become more prevalent, so AI Agent key decisions require human intervention and feedback (Figure 2).

- **Adaptive Learning and Memory** – responsible for refining performance over time, whether through reinforcement learning or historical data analysis. This adaptive capacity is important in settings like climate prediction, as AI Agents can automate climate model development and retune the model based on contextualised predictions based on past events or trends. Zhou et al. (2023) highlight the need for agents to store and retrieve long-term memories and regularly update short-term working memories, fostering continuous improvement and allowing agents to adapt to new contexts. Gronauer & Diepold (2022) study on multi-agent reinforcement learning supports the idea of adaptive learning (Sandini et al., 2024).

- **Proactive Behaviour and Anticipatory Action** - support predictive interventions (Malmio, 2024) but can be constrained by psychological, cognitive, and ethical factors, limiting AI's proactivity.

- **Synergistic Collaboration and Coordination** – Masterman et al. (2024) research endorses agents that engage in collaborative human-machines decisions. For example, a generalist agent can orchestrate virtual worker agents (figure 2) focusing on community input and can share feedback with data analysis agents (specialists) to refine project strategies (for human feedback). However, Klieger et al. (2024) find that if agents do not clearly communicate roles or findings, confusion and ineffective feedback loops may arise in the decision-making process. Automated collaboration and coordination further enhance climate monitoring and

disaster response (Ziemba et al., 2024) by sharing responsibilities with specialised AI Agent experts, though it raises concerns over computational resource consumption and sustainability (Chaudhary, 2023).

- Security and Compliance – essential for data protection, privacy, and alignment with regulations like GDPR to protect people, communities and organisations. Lawless (2024) stress the need for robust security measures to secure sensitive data and maintain public trust. (Chan et al., 2024) highlight the importance of transparency through activity logs and real-time monitoring yet warn that privacy concerns may conflict with detailed oversight. Sandini et al. (2024) note that rapid AI advancements may outpace the development of adequate privacy protections, raising concerns about effective responses to evolving threats. Similarly, over-regulating AI could hinder technological progress Zhao & Gómez Fariñas (2023), though proper governance can ensure accountability.

- Natural Language Processing and Communication: This area facilitates human-AI interaction using mass personalisation, simplifying complex information in different languages. It supports the automated reporting and monitoring required by all SDGs.

- Integration and Implementation – AI Agent solutions require ethical safeguards, interoperability standards and regulatory alignment to ensure responsible integration within global technology eco-system (Kahungi, 2023; Malmio, 2024). Zhao & Gómez Fariñas (2023) note that over-regulation may discourage innovation, while insufficient oversight can compromise trust and ethical standards.

While these capabilities define AI agents' core functions, their impact depends on real-world implementation. The next section explores their applications across SDG domains, highlighting benefits and challenges.

8.2.1 Applications of AI Agents in Sustainable Development: Opportunities and Gaps

Environmental applications include autonomous systems that optimise energy grids, enhance climate modelling, and monitor resources through satellite imagery and automated spatial analytics, though current progress is limited (Akinboyewa et al., 2024; Fraisl, 2024). Meanwhile, in healthcare, AI serves primarily as assistive technology for predictive diagnostics and surgical procedures, with limited research on self-learning agents capable of independent interaction with patients and providers.

The socioeconomic impact of AI Agents extends to poverty reduction and financial inclusion by efficiently allocating resources to vulnerable populations, though most systems remain rule-based rather than truly autonomous (Diener & Špaček, 2021).

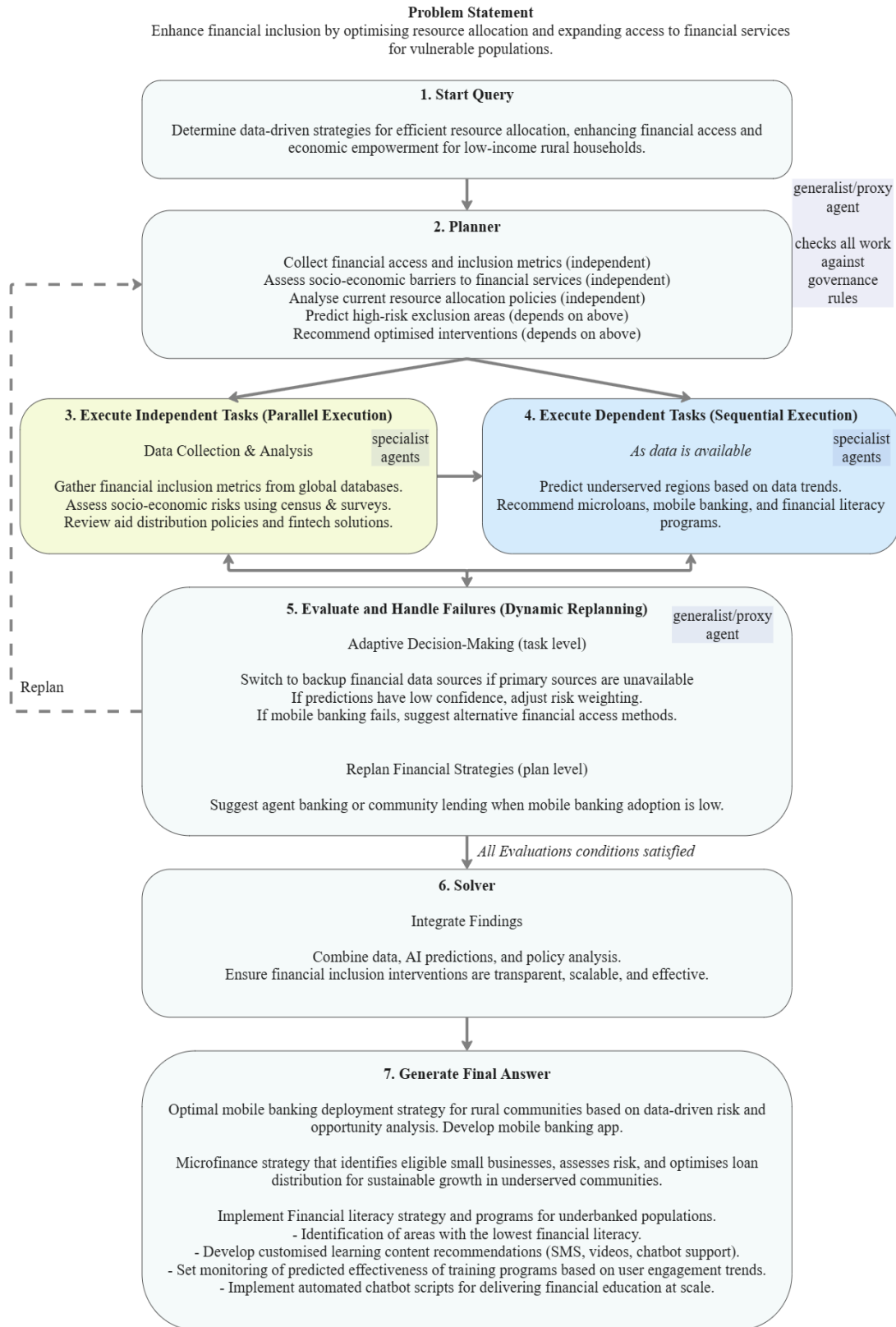


Figure 4: Example of how AI Agents can be applied to work on SDG problems

Figure 4 shows how AI Agents can take a problem, break this into a query and develop a plan and set of required tasks (Wang et al., 2023), aligning with agreed governance rules. Some of the tasks are independent so can be worked on in parallel (Kim et al., 2024; Wang et al., 2023), whereas others will be dependent on either a task to be completed or data or service to become available. If a task is not possible to be completed (such as satellite data not available), alternative tools and approaches will be taken to solve the task (Kim et al., 2024), otherwise as is the case with complex SDG-related problems, if the plan is not possible, alternative strategies will be automatically replanned (Xu et al., 2023). This capability to replan, when faced with failures, allows AI Agents to keep trying alternative unknown approaches (for a solution) using available and approved resources, without having a final clear direction. Finally, once the tasks are all successfully passing evaluations (including compliance with necessary rules and regulations), the problem would have been addressed with multiple solutions and actions completed. The approach can be implemented in open-source platforms that allow the creation of AI Agents, such as LangGraph, Crew AI and Microsoft AutoGen.

However, significant challenges persist, including algorithmic bias, decision opacity, and infrastructure limitations in resource-constrained regions (Bankhwal et al., 2024; Kahungi, 2023), which must be addressed through robust governance frameworks to ensure scalable automated AI Agent deployment doesn't exacerbate existing inequalities. Without proper research and oversight, these powerful tools risk reinforcing systemic problems rather than solving them.

8.3. Research Methodology

8.3.1 Methodology/Framework

This section outlines the methodologies employed to investigate the role of AI agents in advancing the Sustainable Development Goals (SDGs), combining both the quantitative survey and the literature review components.

8.3.2 Data Collection

The research utilized a two-pronged approach for data collection: a literature review and a quantitative survey.

Literature Review: A systematic search was conducted using Google Scholar to identify relevant open-access academic articles focused on AI agents and their contributions to the SDGs. The search criteria focused on literature published from 2023 onward. This time frame was chosen due to the emergence of advanced large language models that exhibit enhanced reasoning capabilities and can effectively process multimodal content (see figures 1 and 2).

The Google Scholar search results revealed a total of 69,500 papers using the term “AI and SDG” so this was refined (Appendix B), see below:

-chatbot -conversational -"autonomous driving" "SDG" AND ("ai agent" OR "autonomous agent" OR "autonomous ai" OR "automated AI" OR "multiagent")

This search term provided 400 results (between 2023 and 2024), which were manually selected and limited to open-access publications.

The inclusion criteria focused on scholarly, peer-reviewed open-access publications published in English between 2023 and 2024 that explicitly connected AI agents to sustainable

development. The exclusion criteria eliminated studies unrelated to sustainability, including those focused on topics like autonomous driving or conversational chatbots, which were not the focus of this research.

Quantitative Survey: A structured questionnaire was designed and distributed electronically to a targeted sample of 384 individuals, including AI professionals, sustainability practitioners, business professionals, and students. The aim was to capture a diverse range of insights regarding their perceptions of AI agents within sustainability contexts, producing 316 valid responses from 48 countries.

The survey asked 20 questions via a 7-point Likert scale to measure various dimensions, including respondents' familiarity with AI agents, perceived effectiveness in addressing sustainability challenges, and ethical concerns associated with their potential implementation. Key themes explored included opinions on regulatory measures, equitable access to AI benefits, and the implications of multi-agent deployment on economic and social outcomes, all important in progressing SDGs.

8.3.3 Theoretical Framework

This research is anchored in the frameworks of responsible AI governance and Value-Sensitive Design (VSD), which focuses on the need for fairness, accountability, and transparency in the deployment of AI technologies (Peters et al., 2020). Sastry et al. (2024) highlight the necessity of ethical governance frameworks in guiding AI applications to mitigate risks and progress AI benefits for sustainable development (see Appendix C).

The OECD's Principles on Artificial Intelligence and Asilomar AI principles (Morandín-Ahuerma, 2023) promote inclusive growth and sustainable development, focusing on AI's benefits and safety for people and the planet. These principles stress the importance of accountability in AI systems to prevent worsening inequalities or creating ethical dilemmas (Chan et al., 2024), necessary for all SDGs.

8.3.4 Research Findings

This section presents the key insights derived from the literature and survey analysis, highlighting respondents' attitudes, perceptions, and experiences regarding the role of AI Agents in sustainable development. The analysis is based on 316 anonymous responses (representing 48 countries), all of which were voluntarily accepted to complete the survey.

Table 1: Top 10 highest responses by country

Rank	Country	Number of Responses
1	United Kingdom	50
2	United States of America	28
3	Canada	27
4	Australia	25
5	India	24
6	Saudi Arabia	17
7	Germany	12

Rank	Country	Number of Responses
8	United Arab Emirates	9
9	Algeria	6
10	New Zealand	5

This distribution demonstrates a notable skew toward English-speaking countries, with five of the top eight being primarily English-speaking nations. Saudi Arabia shows significant participation with 17 responses, making it the top Middle Eastern country represented.

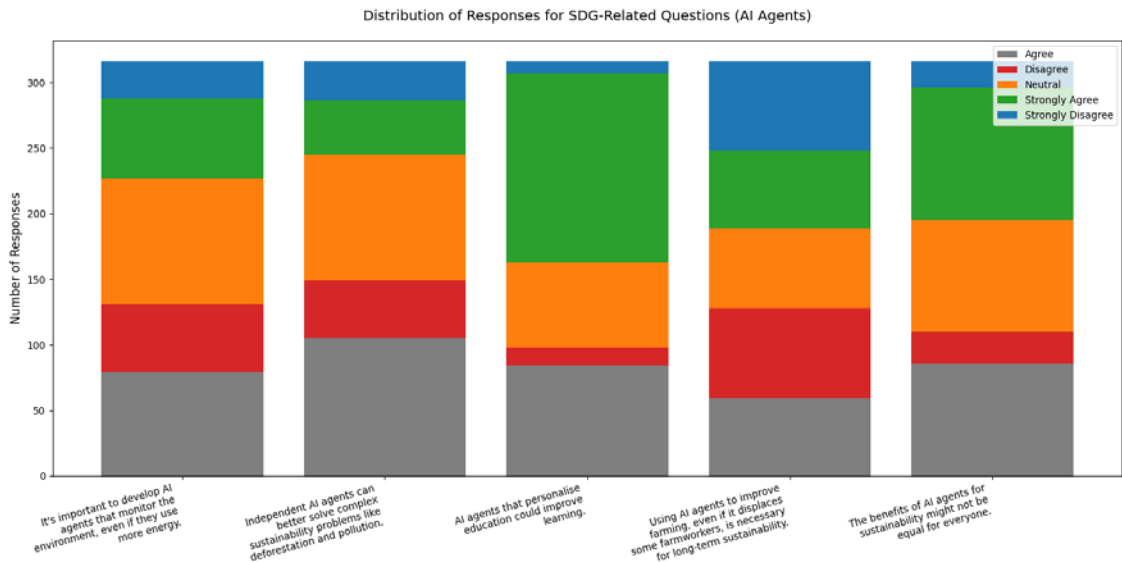


Figure 5: Distribution of Responses for SDG related questions (AI Agents)

Table 2 provides insights based on the response pattern from Figure 5.

Table 2: Analysis of AI Agents and Sustainability Survey Responses

Finding	Survey Statement	Response Pattern	Insight
Highest Agreement	AI agents that personalize education could improve learning	Large proportion of “Strongly Agree” responses with minimal disagreement	Respondents show strong confidence in AI Agents’ potential to enhance personalized education
Most Polarized View	It’s important to develop AI agents that monitor the environment, even if they use more energy	Relatively even distribution across all response categories	Divided opinions on the energy consumption vs. environmental monitoring trade-off

Finding	Survey Statement	Response Pattern	Insight
Strongest Disagreement	Independent AI agents can better solve complex problems like deforestation and pollution	Highest level of “Disagree” and “Strongly Disagree” responses	Skepticism about AI Agents independently solving complex environmental challenges
Ethical Trade-off Recognition	Using AI agents to improve farming, even if it displaces some farmworkers, is necessary for long-term sustainability	Moderate agreement but significant disagreement	Recognition of the complex balance between agricultural productivity gains and potential farmworker displacement
Equity Concerns	The benefits of AI agents for sustainability need to be equal for everyone	Strong overall agreement	Respondents value the equitable distribution of sustainability benefits from AI Agents

Note. Source: Researcher

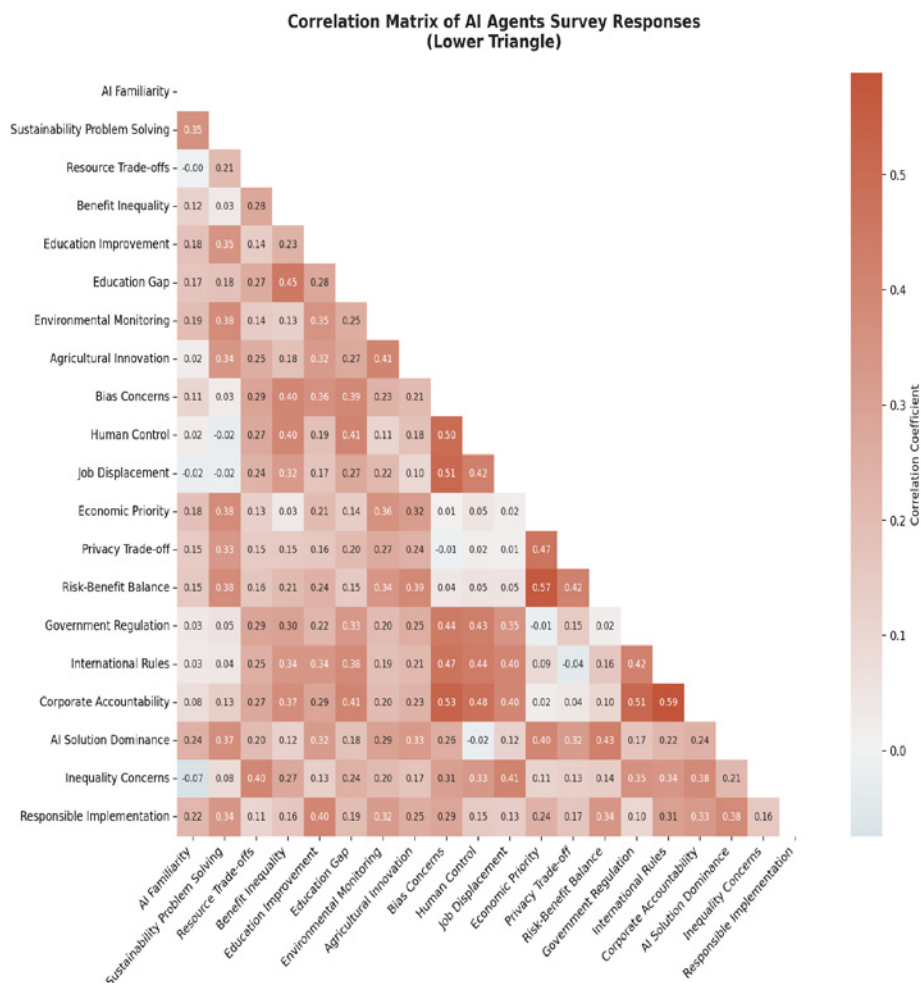


Figure 6: Correlation Matrix of AI Agents Survey Responses

Note. Source: Researcher

Below are the key insights found in the correlation matrix (Figure 6).

Table 3: Correlation analysis of the survey (AI Agents and Sustainable Development Goals)

Correlation	C o r - relation	Description
Corporate Accountability & International Rules	(r = 0.588)	Respondents who believe that companies making AI agents need to prove social and environmental benefits tend to also favour international regulations to ensure these agents are not used harmfully. This suggests a consistent view that both corporate practices and international oversight are key to responsible AI deployment, aligning with Peters et al. (2020).
Risk-Benefit Balance & Economic Priority	(r = 0.567)	Participants who believe that the potential benefits of AI Agents for sustainability outweigh their risks are also inclined to prioritise economic growth. This indicates that for many, the economic opportunities provided align with a strong belief in its potential to contribute positively to sustainable development goals.
Corporate Accountability & Bias Concerns	(r = 0.530)	There is a notable relationship showing that those who demand evidence of corporate accountability also worry about the risk of AI systems reinforcing biases. Concerns about bias call for companies to be more responsible in the development of AI Agents.
Job Displacement & Bias Concerns	(r = 0.509)	The correlation suggests that those who are concerned about the potential job losses due to AI Agents also have significant worries about biases. This indicates an overall cautious outlook regarding the broader societal impacts of rapidly deploying AI solutions, aligning with Figure 7.

Note. Source: Researcher

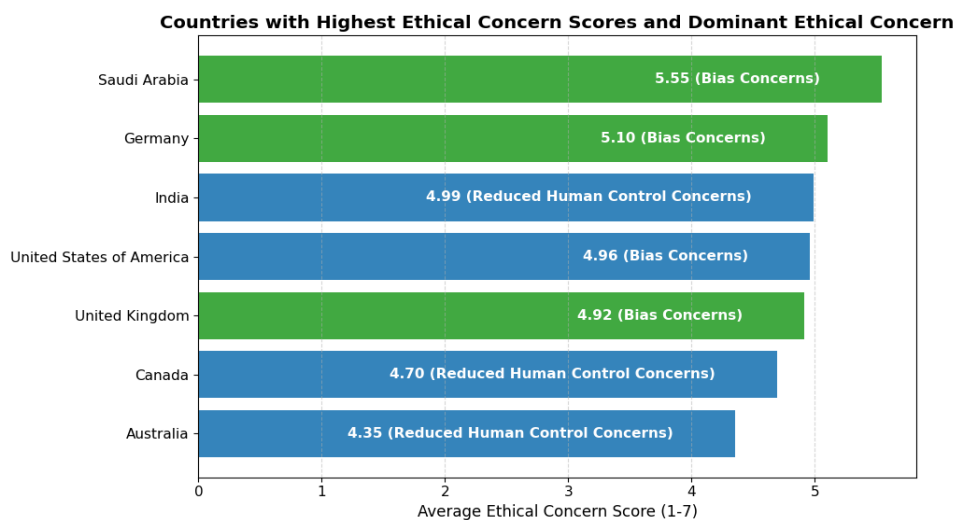


Figure 7: Countries with Highest Ethical Concern Scores and Dominant Ethical Concern

Note. Source: Researcher

Bias Concerns are the dominant ethical concern in 4 out of the 7 countries (Saudi Arabia, Germany, United States and United Kingdom). This indicates that respondents in these countries are particularly worried about AI systems reinforcing or amplifying biases, which could lead to unfair or discriminatory outcomes. India, Canada, and Australia stand out with Reduced Human Control as their dominant concern. This suggests that respondents in these countries are more concerned about AI systems making decisions without sufficient human oversight, potentially leading to a loss of control over critical processes (relevant to all SDGs).

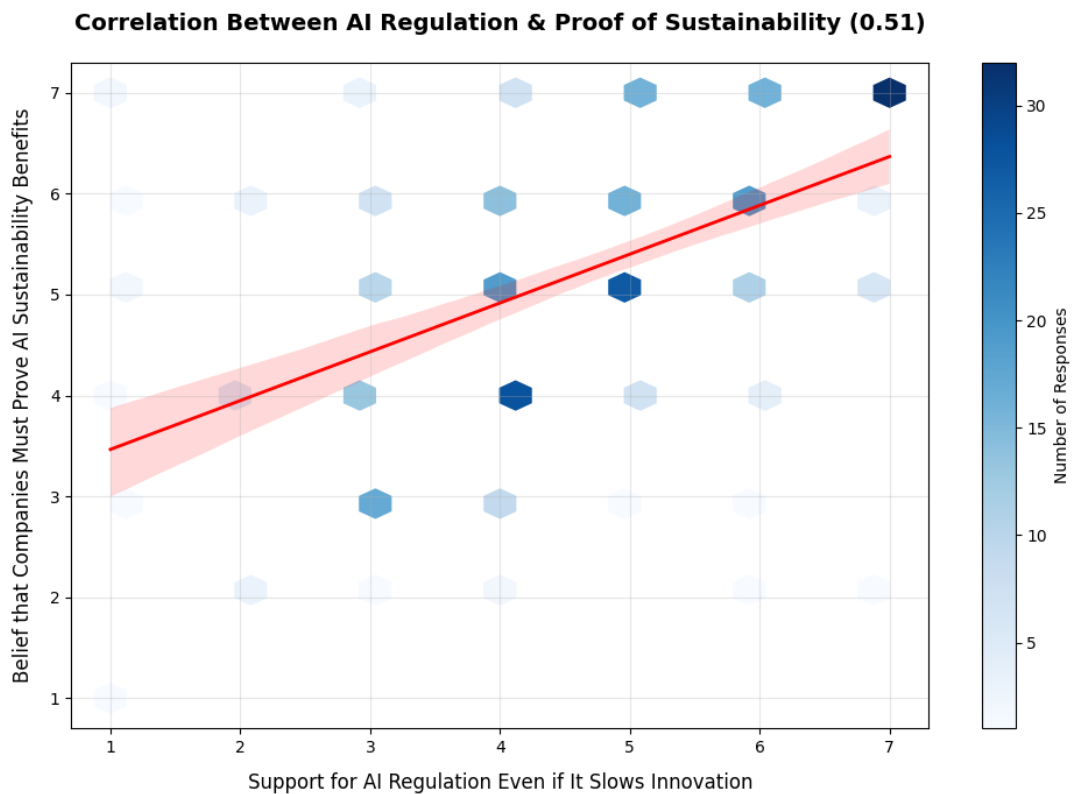


Figure 8: Hexbin Density Plot: AI Regulation vs. Proof of Sustainability

Note. Source: Researcher

Individuals who support stronger AI regulations also tend to demand proof of AI's sustainability benefits ($r = 0.514$). The regression line confirms this trend, showing that higher agreement with regulation correlates with greater expectations for corporate accountability, despite some variation in responses. These correlations show that while respondents see AI Agents as beneficial for sustainability, they also the need for accountability and safeguards to mitigate risks like bias and job displacement.

8.3.5 Access and Equity Issues

The survey findings revealed significant discrepancies in access to AI technologies, with respondents from developing regions expressing greater concerns about AI potentially exacerbating existing inequalities.

Figure 9 below corroborated this by showing that respondents with lower economic circumstances reported higher anxiety regarding equitable access to AI Agent technologies, highlighting the necessity for inclusive initiatives that ensure accessibility to underprivileged communities.

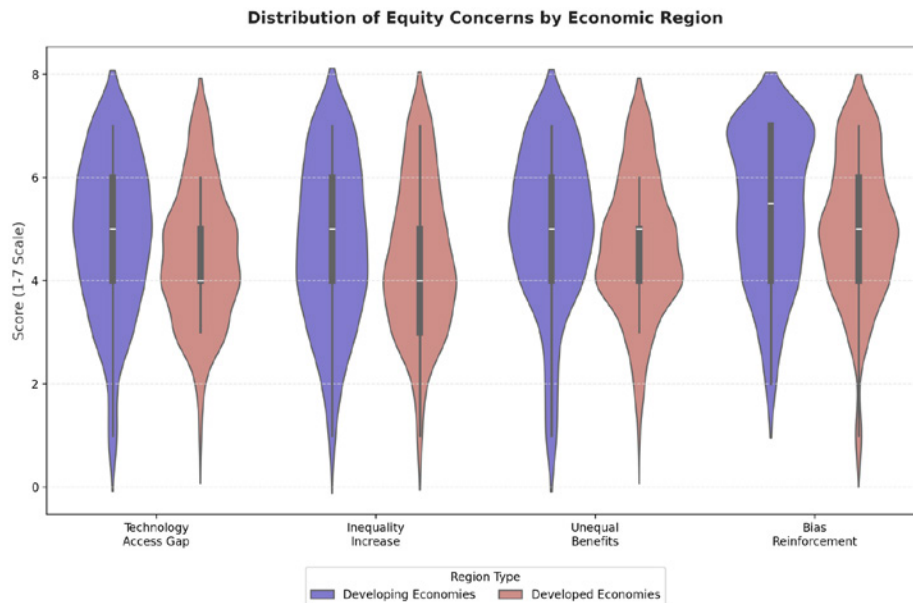


Figure 9: Distribution of Equity Concerns by Economic Region

Note. Source: Researcher

This concern aligns closely with the broader literature (Dwivedi et al., 2021; Kahungi, 2023) advocating for proactive policy measures that address equity issues, ensuring that AI Agents serve as a tool for empowerment rather than deepening socio-economic divides.

8.4. Challenges and Opportunities

Reaching the United Nations Sustainable Development Goals (SDGs) requires innovative solutions. Survey data indicate that increased familiarity with AI agents is associated with greater confidence in their effectiveness ($r = 0.65$). However, the academic literature on AI Agents and SDG applications remains sparse, and findings also reveal that stakeholders from developing regions express heightened concerns about equal access to such technologies. These gaps underscore the urgent need for further research and inclusive innovation.

8.4.1 Technical Complexity & Real-Time Integration

Developing large-scale multi-agent solutions requires thoughtful architecture and design that continuously process live data (e.g., social cues, sensor networks, satellite feeds), include timely human interactions/feedback and adapt to dynamic environmental and complex socio-economic conditions.

8.4.2 Ethical, Environmental & Social Concerns

Survey findings reveal significant ethical apprehensions, with an average ethical concern score of 5.6 and strong correlations between accountability demands and worries about bias. High energy requirements, carbon footprints, and potential e-waste from continuous real-time processing present serious environmental challenges. Automation may lead to job displacement, underscoring the importance of proactive reskilling and social inclusion initiatives.

8.4.3 System Robustness & Continuous Oversight

Ensuring stability in unpredictable, real-world environments remains challenging. Survey data indicate that continuous human auditing and interaction is important to quickly address issues and prevent cascading failures in globally inter-connected AI Agent solutions.

8.5 Opportunities

AI Agents can help address resource limitations and get the Sustainable Development Goals (SDGs) back on course. Currently, only 17% of the targets are on track but these technologies can optimise resource allocation, hire virtual workers, and assemble a team of specialised AI Agents to work on ongoing challenges and undetermined tasks. These agents can optimise sustainable logistics for responsible consumption (SDG 12) by continuously tracking inventory in real-time and eliminating waste across supply chains. For climate action (SDG 13), AI Agents can enhance disaster preparedness through predictive analysis and can provide detailed emissions tracking, enabling precise environmental interventions and increased accountability.

In education (SDG 4), AI Agents can create personalised learning experiences that adapt to each student's needs, making quality education accessible in remote areas. These adaptive systems can identify knowledge gaps, adjust difficulty levels automatically, and deliver customised feedback to bridge digital divides (Mollick et al., 2024). To maximise these benefits, open-source frameworks, ethical design principles, and trust-building initiatives can ensure AI deployments remain transparent and equitable while accelerating sustainable impact globally.

8.5.1 Ethical Framework for Scaling Agentic AI

There is an opportunity for agentic AI to make further progress toward the SDGs on a global scale. Importantly, the solutions must operate with empathy and fairness, ensuring that benefits reach both developed and developing regions and include transparent, auditable decision processes. The following are the key focus areas that need to be considered together to support organisations implementing multi-agent solutions in progressing SDGs.

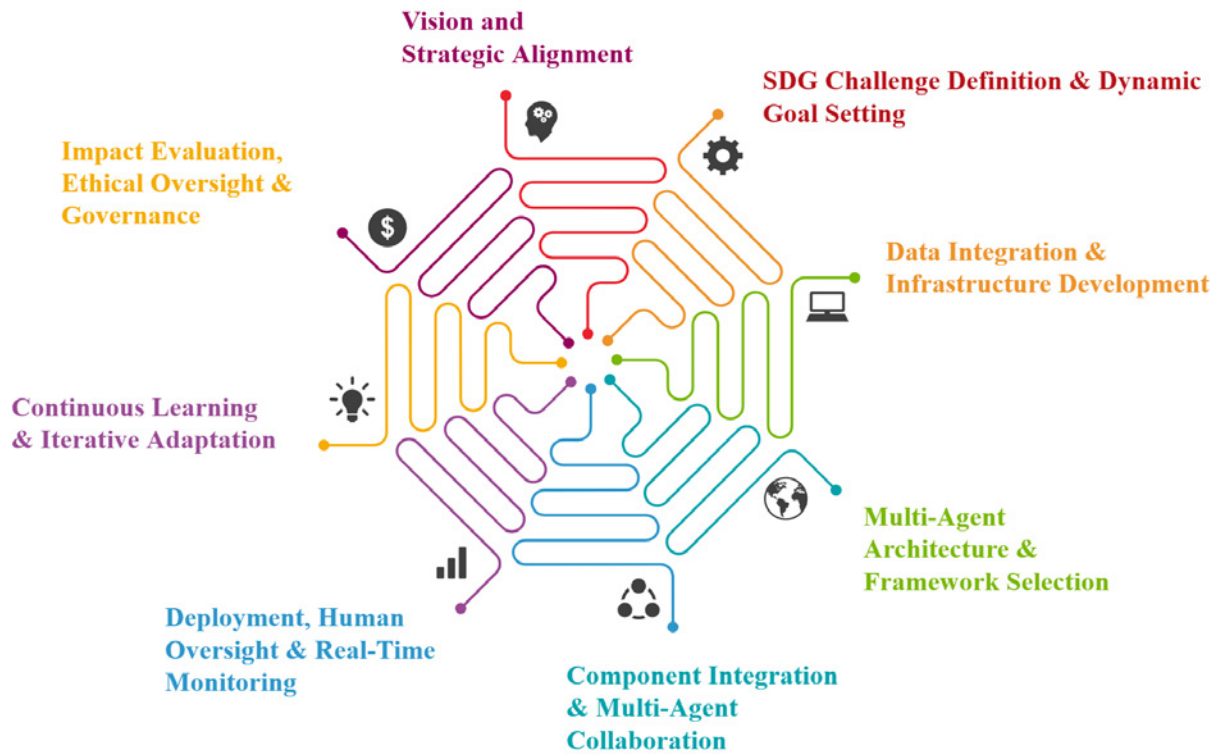


Figure 10: Key Focus Areas related to Agentic AI: Governance, Deployment & Sustainable Impact

Note. Source: Researcher

Table 4: Strategic Actions and Considerations for Agentic AI Focus Areas

Key Focus Area	Strategic Actions & Considerations
Vision and Strategic Alignment	Define a clear Multi-Agents vision for SDGs, aligning initiatives with corporate and public strategies. Engage stakeholders and provide training to ensure fair, ethical implementation.
SDG Challenge Definition & Dynamic Goal Setting	Identify global sustainability challenges and set real-time, adaptable KPIs for insights and goal-setting.
Data Integration & Infrastructure Development	Create a unified data ecosystem by integrating real-time and historical data from diverse sources. Develop scalable pipelines for continuous analysis and proactive decision-making.
Multi-Agent Architecture & Framework Selection	Design the enterprise architecture to integrate AI Agents' perception, execution, and memory while embedding ethical human-in-the-loop oversight to maintain transparency, scalability, and responsible AI governance (Peters et al., 2020).

Component Integration & Multi-Agent Collaboration	Foster collaboration among expert AI agents for scalable and resilient analysis. Understand service needs and potential AI Agent specialists which need to be approved.
Deployment, Human Oversight & Real-Time Monitoring	Deploy automated reasoning, planning and task assignment in event-driven environments with clear human involvement. Use real-time monitoring and governance dashboards to track SDG progress and ensure transparency in decision-making.
Continuous Learning & Iterative Adaptation	Establish feedback loops for AI Agents' learning and model refinement. Continuously update understandings and guardrails based on new data and evolving SDG objectives.
Impact Evaluation, Ethical Oversight & Governance	Evaluate agentic AI impact on SDG targets using quantitative and qualitative methods. Conduct ethical audits, implement governance structures, and maintain transparency in related actions. Understand which tasks must require greater visibility, control and transparency.

Note. Source: Researcher

Based on the insights gathered, this ethical framework (Figure 11) helps organizations deploy multi-agents for advancing SDG impact by aligning global challenges with varied aspects. It enables seamless data integration, adaptive architectures, and accountability. With human oversight, performance monitoring, and continuous learning, multi-agent AI-based solutions remain flexible and ethical. Impact evaluation and governance structures ensure transparency, driving responsible adoption for a sustainable and inclusive future.

1. Vision and Strategic Alignment <u>Set the Vision</u> How will your organisation commit to using agentic AI to achieve specific SDG targets? What specific SDG targets will your agentic AI initiative focus on addressing? <u>Strategic Integration</u> How will you map your agentic AI initiatives to existing digital transformation and sustainability priorities? <u>Stakeholder Engagement & Team Training</u> Which diverse stakeholders will you involve in your AI implementation process?	2. SDG Challenge Definition & Dynamic Goal Setting <u>Identify Global Challenges</u> Which sustainability challenges affecting both developed and developing regions are your priorities? <u>Set Dynamic Objectives</u> What real-time KPIs will you use to measure your AI system's autonomous performance?	3. Data Integration & Infrastructure Development <u>Unified Data Ecosystem</u> Which critical data sources will you connect to create your global data ecosystem? <u>Scalable Pipelines</u> How will your data pipelines merge real-time and historical data for proactive decision-making?	4. Multi-Agent Architecture & Framework Selection <u>Adapting AI Systems to Change</u> How will your agentic AI architecture remain flexible and responsive to evolving challenges and business needs? <u>Strategic Framework Selection</u> What principles will guide your choice of agentic AI frameworks to balance innovation, scalability, and ethical responsibility?
5. Component Integration & Multi-Agent Collaboration <u>Modular Integration</u> Which communication channels will you use for data ingestion, processing, and decision support? <u>Collaborative Ecosystem</u> How will your general and specialised AI agents work together to provide actionable recommendations?	6. Deployment, Human Oversight & Real-Time Monitoring <u>Balanced Deployment</u> How will your AI agents integrate with existing business ecosystems? <u>Human-in-the-Loop</u> At what critical points will human oversight be required in your agentic AI system? <u>Performance Dashboards</u> Which key metrics will your monitoring dashboards track in real-time?	7. Continuous Learning & Iterative Adaptation <u>Feedback Mechanisms</u> What feedback loops will you implement to refine your agentic AI's forecasting models? <u>Iterative Refinement</u> How often will you update your algorithms and protocols based on new insights?	8. Impact Evaluation, Ethical Oversight & Governance <u>Measure Outcomes</u> What methods will you use to evaluate your agentic AI interventions against SDG targets? <u>Ethical Audits</u> How will you ensure transparent recording and assessment of AI decision-making? <u>Governance Structures</u> What operational guardrails will keep autonomous actions within ethical parameters?

Figure 11: Agentic AI Ethical Framework
Note. Source: Researcher

8.6 Conclusion and Recommendations

Achieving the UN SDGs demands adaptive solutions, and agentic AI offers a scalable approach through real-time perception, reasoning, action, and learning. Survey findings show AI Agents familiarity correlates with confidence ($r = 0.65$), though concerns persist around equity, complexity, and ethics, especially in resource-limited areas.

Challenges like bias, job displacement, and performance consistency require careful mitigation, yet real-time optimization, open-source frameworks, and ethical design offer viable solutions. The proposed Agentic AI Ethical Framework provides a guideline for transparent, equitable deployment, aligning solutions design with dynamic goal-setting, real-time data integration, and human-in-the-loop governance protocols to foster public trust.

With strong oversight, inclusive data practices, and collaborative planning, AI agents can accelerate global sustainability efforts, advancing the UN's 2030 Agenda through responsible, impactful innovation.

8.6.1 *Limitations of This Study*

This study acknowledges several limitations. The focus on 2023–2024 research may not fully capture emerging trends in AI Agents. Sample bias could affect survey findings, as regional and professional perspectives might influence views on AI familiarity, ethics, and benefits. The rapidly evolving nature of AI technologies means the proposed Agentic AI Ethical Framework needs regular updates based on newer research. Additionally, the lack of longitudinal data limits understanding of short-term trends and the long-term effects on perceptions and implementation of AI. Future research should explore diverse cultural contexts, longitudinal effects, and evolving global standards.

8.6.2 *Recommendations*

To maximize the potential of AI Agents for sustainable development, adopting the Ethical Agentic AI framework is recommended, as it offers a structured method for planning and deploying these technologies. This includes establishing guidelines for fairness and accountability while engaging local communities to address specific risks. Creating inclusive data ecosystems that integrate various data sources and investing in energy-efficient technologies is important. Ensuring human-in-the-loop and related auditing in key decisions and using dashboards to monitor progress against Sustainable Development Goals (SDGs) will help maintain accountability. Enhancing education and capacity-building initiatives to improve AI literacy among policymakers and communities is important for equal benefits. More investigation is required regarding AI Agents and their relation to the SDGs, as this is a new field with significant potential. By implementing these strategies through the Ethical Agentic AI framework, organizations can effectively utilise AI Agents to advance the SDGs while promoting ethical practices.

8.7 References

- Akinboyewa, T., Li, Z., Ning, H., & Lessani, M. N. (2024). *GIS Copilot: Towards an Autonomous GIS Agent for Spatial Analysis*. <https://shorturl.at/bI4Ep>.
- Ametepey, S. O., Aigbavboa, C., Thwala, W. D., & Addy, H. (2024). The Impact of AI in Sustainable Development Goal Implementation: A Delphi Study. *Sustainability (Switzerland)*, 16(9). <https://doi.org/10.3390/su16093858>
- Amin, H. A., & Alanzi, T. M. (2024). Utilization of Artificial Intelligence (AI) in Healthcare Decision-Making Processes: Perceptions of Caregivers in Saudi Arabia. *Cureus*. <https://doi.org/10.7759/cureus.67584>
- Bankhwal, M., Bisht, A., Chui, M., Roberts, R., & van Heteren, A. (2024). *AI for social good: Improving lives and protecting the planet*.
- Barros, P., Yalçın, Ö. N., Tanevska, A., & Sciutti, A. (2023). Incorporating rivalry in reinforcement learning for a competitive game. *Neural Computing and Applications*, 35(23), 16739–16752. <https://doi.org/10.1007/s00521-022-07746-9>
- Chan, A., Ezell, C., Kaufmann, M., Wei, K., Hammond, L., Bradley, H., Bluemke, E., Rajkumar, N., Krueger, D., Kolt, N., Heim, L., & Anderljung, M. (2024). Visibility into AI Agents. *2024 ACM Conference on Fairness, Accountability, and Transparency, FAccT 2024*, 958–973. <https://doi.org/10.1145/3630106.3658948>
- Chaudhary, G. (2023). Environmental Sustainability: Can Artificial Intelligence be an Enabler for SDGs? *Nature Environment and Pollution Technology*, 22(3), 1411–1420. <https://doi.org/10.46488/NEPT.2023.v22i03.027>
- Diener, F., & Špaček, M. (2021). Digital transformation in banking: A managerial perspective on barriers to change. *Sustainability (Switzerland)*, 13(4), 1–26. <https://doi.org/10.3390/su13042032>
- Dwivedi, Y. K., Hughes, L., Ismagilova, E., Aarts, G., Coombs, C., Crick, T., Duan, Y., Dwivedi, R., Edwards, J., Eirug, A., Galanos, V., Ilavarasan, P. V., Janssen, M., Jones, P., Kar, A. K.,
- Kizgin, H., Kronemann, B., Lal, B., Lucini, B., ... Williams, M. D. (2021). Artificial Intelligence (AI): Multidisciplinary perspectives on emerging challenges, opportunities, and agenda for research, practice and policy. *International Journal of Information Management*, 57. <https://doi.org/10.1016/j.ijinfomgt.2019.08.002>
- Fraisl, D. (2024). *DIGITALISATION The Potential of Artificial Intelligence for the SDGs and Official Statistics*.
- Gosselink, B. H., Brandt, K., Croak, M., Desalvo, K., Gomes, B., Ibrahim, L., Johnson, M., Matias, Y., Porat, R., Walker, K., & Manyika, J. (2024). *AI in Action: Accelerating Progress Towards the Sustainable Development Goals*.
- Gronauer, S., & Diepold, K. (2022). Multi-agent deep reinforcement learning: a survey. *Artificial Intelligence Review*, 55(2), 895–943. <https://doi.org/10.1007/s10462-021-09996-w>
- Jannelli, V., Schoepf, S., Bickel, M., Netland, T., & Brintrup, A. (2024). *Agentic LLMs in the Supply Chain: Towards Autonomous Multi-Agent Consensus-Seeking*. <http://arxiv.org/abs/2411.10184>
- Kahungi, N. W. (2023). *Culpability in the Era of Artificial Intelligence in Kenya: An Overview*. <https://course.elementsofai.com/1/1>
- Kim, S., Moon, S., Tabrizi, R., Lee, N., Mahoney, M. W., Keutzer, K., & Gholami, A. (2024). *An LLM Compiler for Parallel Function Calling*. <https://github.com/SqueezeAILab/LLMCompiler>.
- Klieger, B., Charitsis, C., Suzara, M., Wang, S., Haber, N., & Mitchell, J. C. (2024). *ChatCollab:*

- Exploring Collaboration Between Humans and AI Agents in Software Teams*. <http://arxiv.org/abs/2412.01992>
- Lawless, W. (2024). Bidirectional Human-AI/Machine Collaborative and Autonomous Teams: Risk, Trust and Safety. *Human Factors in Robots, Drones and Unmanned Systems*, 138. <https://doi.org/10.54941/ahfe1005014>
- Malmio, I. (2024). Artificial intelligence and the social dimension of sustainable development: through a security perspective. In *Discover Sustainability* (Vol. 5, Issue 1). Springer Nature. <https://doi.org/10.1007/s43621-024-00677-6>
- Masterman, T., Besen, S., Sawtell, M., & Chao, A. (2024). *The Landscape of Emerging AI Agent Architectures for Reasoning, Planning, and Tool Calling: A Survey*. <http://arxiv.org/abs/2404.11584>
- Mollick, E., Mollick, L., Bach, N., Ciccarelli, L., Przystanski, B., & Ravipinto Generative Lab at Wharton, D. A. (2024). *AI AGENTS AND EDUCATION: SIMULATED PRACTICE AT SCALE*.
- Morandin-Ahuerma, F. (2023). *Twenty-three Asilomar principles for Artificial Intelligence and the Future of Life*. Open Science Framework. <https://doi.org/10.31219/osf.io/dgnq8>
- Peters, D., Vold, K., Robinson, D., & Calvo, R. A. (2020). Responsible AI—Two Frameworks for Ethical Design Practice. *IEEE Transactions on Technology and Society*, 1(1), 34–47. <https://doi.org/10.1109/tts.2020.2974991>
- Rajesh, V., & Adapa, K. (2024). AI for Climate Action: Leveraging Artificial Intelligence to Address Climate Change Challenges. In *IJFMR240529015* (Vol. 6, Issue 5). www.ijfmr.com
- Rana, S. M., & Shuford, J. (2024). *AI in Healthcare: Transforming Patient Care through Predictive Analytics and Decision Support Systems*. <https://ojs.boulibrary.com/index.php/JAIGS>
- Ruan, J., Chen, Y., Zhang, B., Xu, Z., Bao, T., Du, G., Shi, S., Mao, H., Li, Z., Zeng, X., Zhao, R., & Research, S. (2023). *TPTU: Task Planning and Tool Usage of Large Language Model-based AI Agents*. <https://github.com/Significant-Gravitas/Auto-GPT>
- Sandini, G., Sciutti, A., & Morasso, P. (2024). Artificial cognition vs. artificial intelligence for next-generation autonomous robotic agents. *Frontiers in Computational Neuroscience*, 18. <https://doi.org/10.3389/fncom.2024.1349408>
- Sastry, G., Heim, L., Belfield, H., Anderljung, M., Brundage, M., Hazell, J., O’Keefe, C., Hadfield, G. K., Ngo, R., Pilz, K., Gor, G., Bluemke, E., Shoker, S., Egan, J., Trager, R. F., Avin, S., Weller, A., Bengio, Y., & Coyle, D. (2024). *Computing Power and the Governance of Artificial Intelligence*. <http://arxiv.org/abs/2402.08797>
- Savec, V., & Jedrinović, S. (2024). The Role of AI Implementation in Higher Education in Achieving the Sustainable Development Goals: A Case Study from Slovenia. *Sustainability (Switzerland)*, 17(1). <https://doi.org/10.3390/su17010183>
- Sekiguchi, K., & Ohsawa, Y. (2024). Aiding narrative generation in collaborative data utilization by humans and AI agents. *AI and Society*. <https://doi.org/10.1007/s00146-024-02156-y>
- Tulcanaza-Prieto, A. B., Cortez-Ordoñez, A., & Lee, C. W. (2023). Influence of Customer Perception Factors on AI-Enabled Customer Experience in the Ecuadorian Banking Environment. *Sustainability (Switzerland)*, 15(16). <https://doi.org/10.3390/su151612441>
- Ueda, D., Walston, S. L., Fujita, S., Fushimi, Y., Tsuboyama, T., Kamagata, K., Yamada, A., Yanagawa, M., Ito, R., Fujima, N., Kawamura, M., Nakaura, T., Matsui, Y., Tatsugami, F., Fujioka, T., Nozaki, T., Hirata, K., & Naganawa, S. (2024). Climate change and artificial intelligence in healthcare: Review and recommendations towards a sustainable future. In *Diagnostic and Interventional Imaging*. Elsevier Masson s.r.l. [166](https://doi.org/10.1016/j.</p>
</div>
<div data-bbox=)

- United Nations. (2024). *The Sustainable Development Goals Report*.
- Wang, L., Xu, W., Lan, Y., Hu, Z., Lan, Y., Lee, R. K.-W., & Lim, E.-P. (2023). *Plan-and-Solve Prompting: Improving Zero-Shot Chain-of-Thought Reasoning by Large Language Models*. <http://arxiv.org/abs/2305.04091>
- Xu, B., Peng, Z., Lei, B., Mukherjee, S., Liu, Y., & Xu, D. (2023). *ReWOO: Decoupling Reasoning from Observations for Efficient Augmented Language Models*. <http://arxiv.org/abs/2305.18323>
- Youn, S. Y., & Luan, C. C. (2024). Soft Biometrics in Retail Service: Understanding Privacy Paradox and Cross-Cultural Differences Regarding 3D Body Scanning Technology. *Clothing and Textiles Research Journal*, 42(3), 222–241. <https://doi.org/10.1177/0887302X231220616>
- Zhao, J., & Gómez Fariñas, B. (2023). Artificial Intelligence and Sustainable Decisions. *European Business Organization Law Review*, 24(1), 1–39. <https://doi.org/10.1007/s40804-022-00262-2>
- Zhou, W., Jiang, Y. E., Li, L., Wu, J., Wang, T., Qiu, S., Zhang, J., Chen, J., Wu, R., Wang, S., Zhu, S., Chen, J., Zhang, W., Tang, X., Zhang, N., Chen, H., Cui, P., & Sachan, M. (2023). *Agents: An Open-source Framework for Autonomous Language Agents*. <http://arxiv.org/abs/2309.07870>
- Ziemba, E. W., Duong, C. D., Ejdy, J., Gonzalez-Perez, M. A., Kazlauskaitė, R., Korzynski, P., Mazurek, G., Paliszkiwicz, J., Stankevičienė, J., & Wach, K. (2024). Leveraging artificial intelligence to meet the sustainable development goals. *Journal of Economics and Management*, 46, 508–583. <https://doi.org/10.22367/jem.2024.46.19>