AI's Impact on Our Sustainable Future: A Guiding Framework for Responsible AI Integration Into ESG Paradigms

Responsible Artificial Intelligence Institute

Table of Contents

AI's Impact on Our Sustainable Future:
A Guiding Framework for Responsible AI Integration Into ESG Paradigms

1. Introduction

AI's Impact on Our Sustainable Future:
A Guiding Framework for Responsible AI Integration Into ESG Paradigms

The convergence of artificial intelligence (AI) and Environmental, Social, and Governance (ESG) presents a complex and emerging terrain for organizations. Increasing use of AI complicates making informed and sustainable decisions by layering in additional existing and potential risks. In today's environment, businesses face a pressing need to understand and leverage their synergies for positive impact, while managing associated risks. This white paper recognizes and addresses the necessity of adopting a comprehensive approach to AI investments and adoption, while stressing the evaluation of their effects on ESG objectives.

By examining AI initiatives through the lens of sustainability, responsible use, and societal impact, organizations can align technological advancements within their own missions, as well as broader ESG objectives. This paper underscores the potential of AI to drive positive outcomes, such as reducing environmental impact, enhancing social welfare, and strengthening governance practices. However, it also acknowledges the risks and challenges AI poses to ESG performance.

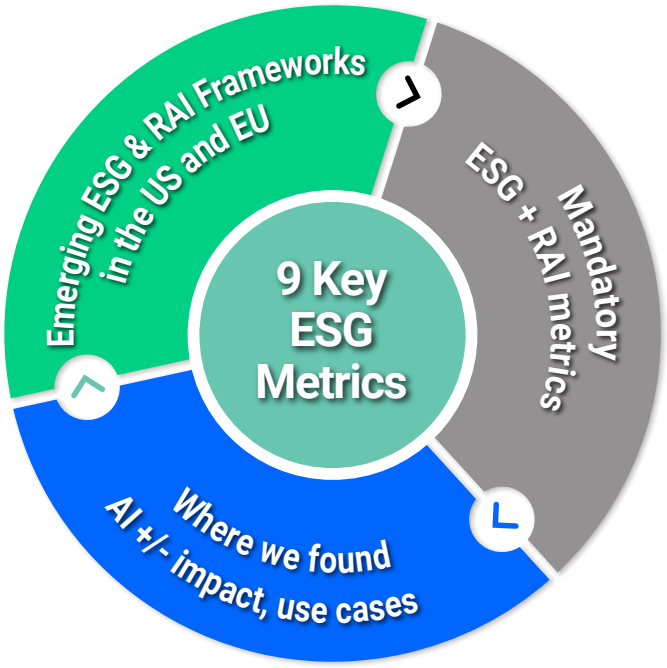
A framework to guide responsible AI practices in the context of ESG is needed for assessment, monitoring and evaluation. This guiding framework offers a roadmap for integrating AI to support and enhance ESG commitments, while also providing a helpful lens to holistically evaluate AI projects. The key components encompass a variety of dimensions: purpose and alignment; impact assessment; responsible AI principles; data quality and sources; transparency and accountability; data privacy and security; and collaboration and stakeholder engagement.

These guiding questions enable organizations and stakeholders to make responsible decisions related to their AI projects and investments that improve environmental, social, and governance sustainability, ultimately contributing to a more sustainable future.

2. Identification of Relevant Set of Disclosures and Metrics within ESG and Responsible AI Frameworks

Environmental, Social, and Governance (ESG) and Responsible AI are rapidly evolving domains where comprehensive guidelines, frameworks, and regulations are still in the developmental stages. While these areas have progressed independently over the past few years, there has been a growing recognition of the need for authoritative implementation guidance, best practices, and regulatory requirements that address intersections and interdependencies.

Methodology to select the 9 key ESG metrics

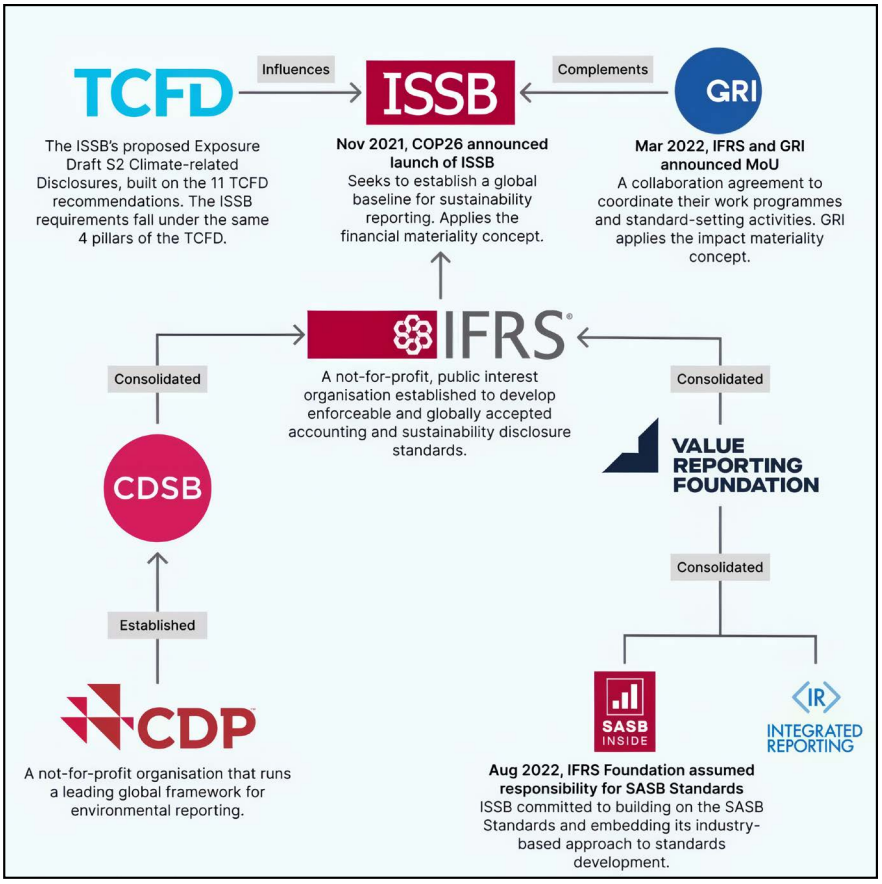


Despite the separate advancements in ESG and AI, it has become increasingly evident that a disciplined framework is necessary to analyze the areas where these two domains converge and impact one another. By adopting a structured approach, organizations can better understand and navigate the complex relationships between ESG considerations and the responsible development and deployment of AI systems.

ESG Disclosure Metrics and Frameworks

The non-financial reporting landscape is transforming, prompted by new ESG reporting requirements. This shift is particularly evident in the European Union and the United States. Both the EU and US are not only among the world's largest economies and greenhouse gas emitters, but also have a leadership position in AI research and innovation. Their actions and policies play a pivotal role in shaping the global sustainability and technology agendas. By aligning the frameworks with the EU and US regulations, this paper provides a consistent and coherent approach to Responsible AI and ESG reporting, adaptable to diverse contexts and jurisdictions.[1]

Significant ESG frameworks and initiatives, including the International Sustainability Standards Board (ISSB), Global Reporting Initiative (GRI), and the Task Force on Climate-related Financial Disclosures (TCFD), among others, are integral to our understanding of the ESG reporting landscape. These frameworks are essential for organizations to navigate the complexities of sustainability reporting, providing foundations for integrating AI considerations into their ESG disclosures.[2]



Responsible AI Regulatory Landscape and Frameworks

The global regulatory landscape for AI encompasses diverse approaches, with significant developments such as the European Union's AI Act and the National Institute of Standards and Technology's (NIST) AI Risk Management Framework. These frameworks set a precedent for the governance of AI, emphasizing transparency, accountability and human oversight. As the global community seeks to navigate the complexities of AI regulation, integrating Responsible AI governance metrics into ESG frameworks emerges as a strategic imperative. The growing emphasis on ethical operations and sustainability drives regulatory and investment decisions.

Integrating responsible AI governance metrics into an ESG framework is a strategic decision. ESG performance is becoming central to how investors assess companies. As AI technology increasingly plays a role in corporate operations, incorporating AI governance into these frameworks ensures that AI is used responsibly and safely, aligning with investor and regulatory expectations to build a sustainable future.

Integrating Responsible AI into ESG Frameworks

AI governance is a subset of corporate governance and has both environmental and social impacts, which is necessary for aligning with the evolving expectations of investors and regulators. These existing frameworks ensure that AI is used responsibly to assess and enhance sustainability and to recognize the social implications of business operations. By identifying key metrics relevant to the impact of Responsible AI within the ESG framework, organizations can begin to align their AI practices with broader sustainability goals, establishing pragmatic approaches to navigating the complexities of ESG and AI alignment and integration.

Challenges and Limitations of ESG-AI Integration

Applying and integrating frameworks across various jurisdictions presents significant challenges, including differences in cultural values, legal frameworks and stakeholder expectations. These limitations necessitate a flexible and dynamic approach tailored to each organization's specific needs. As regulatory and investment environments continue to evolve, reconciling limitations and engaging in continuous dialogue with stakeholders is essential for refining frameworks addressing the evolving landscape of both ESG and AI.

Methodology for Analyzing the Integration of Responsible AI into ESG Frameworks

In addressing the intersection of RAI and ESG, the methodology identifies key mandatory ESG metrics that could be influenced—both positively and negatively—by AI technologies. While the selection of metrics is not comprehensive, exhaustive or final, this analysis serves as a foundation for ongoing analysis and enhancement.

This approach allows a systematic examination of how AI can impact critical ESG areas, providing a balanced perspective on the potential benefits and challenges. Through this analytical lens, this analytical framework offers a pragmatic starting point for deeper investigations and the continued evolution of integrating AI within ESG frameworks.

The selected metrics under the Environmental, Social, And Governance Pillars are presented in the table below.

Environmental Metrics	Social Metrics	Governance Metrics
Greenhouse gas emissions	Health and safety	Board composition and oversight
Energy consumption	Diversity, equity, and inclusion	Compliance and Codes of Conduct
Water consumption	Human rights and labor standards	Governance and accountability

To utilize the ESG frameworks, it is critical to identify existing and probable impacts for both positive and negative effects from AI.

The following sections examine some of the potential positive and negative impacts of AI on each of the metrics listed in the table above.

Environmental metrics

Greenhouse gas emissions[3]

This metric indicates how much a company contributes to global warming by releasing greenhouse gases (GHGs) into the atmosphere, measured in tons of carbon dioxide equivalent (CO2e). AI can have both positive and negative impacts on GHG emissions. AI shows the potential to reduce emissions through efficiency, enabling innovation in low-carbon technologies, and locating methane leaks, a particularly potent GHG. However, AI can also increase GHG emissions by requiring large amounts of energy and resources to develop, deploy and operate its systems.

Energy consumption

This metric indicates how much energy a company uses to run its operations, measured in kilowatt-hours (kWh) or megawatt-hours (MWh). AI can have both positive and negative impacts on this metric. AI can help lower energy consumption by enhancing energy efficiency, optimizing energy systems, and facilitating the transition to renewable energy sources.[4] AI can help measure, proxy, or otherwise provide data needed for reporting and process improvements on both GHG emissions and energy consumption by using predictive analytics and data visualization to collect, process, and help interpret large, complex datasets from various sources, such as sensors, satellites, drones, and smart meters.[5] AI can estimate a company's carbon footprint by using data from its operations, supply chain, and customers and then analyzing how much of the energy came from fossil fuels or renewable energy sources. Interpreting this data with human collaboration, AI is a critical tool to suggest where and when to conserve by identifying the most effective abatement levers.[6] Conversely, AI also raises energy consumption by demanding high-performance computing and data storage infrastructure, which may rely on fossil fuels or generate waste heat.[7]

Water consumption

This metric measures the quantity of water a company draws from natural sources for its operations and value chain, measured in cubic meters (m3). AI can help decrease water consumption by improving water efficiency, optimizing water management, and enabling data analysis which can be used to improve water quality and ecosystems. Conversely, AI increases water consumption by requiring cooling systems for data centers and hardware components, which may consume large amounts of water or contaminate water sources.[8]

Social metrics

Health and safety

This metric measures the impact of a company's AI products or services on the health and safety of its employees, customers, and other stakeholders. This may include injury prevention, accident reduction, disease detection, or emergency response. This metric addresses the potential of AI to improve the well-being of populations and the potential of AI to cause harm or damage to people. AI can help monitor and prevent workplace hazards, but it can also create new risks, such as through AI systems bias and loss of privacy.

Diversity, equity, and inclusion

This metric reflects the extent to which a company's AI products or services reflect and respect the diversity, equity, and inclusion of its employees, customers, and other stakeholders. This may include representation, participation, accessibility,

fairness, or empowerment. These quantitative and qualitative measures are to assess the potential of AI to foster a culture of diversity, equity, and inclusion, as well as the potential of AI to exacerbate discrimination or exclusion. AI can help promote diversity in hiring and provide data for decision-making, but it can also reinforce stereotypes or prejudices.

Human rights and labor standards

This metric measures the impact of a company's AI products or services on the human rights and labor standards of its employees, customers, and other stakeholders. This may include privacy, freedom of expression, dignity, consent, or working conditions. This metric is relevant for assessing the potential of AI to uphold and advance human rights and labor standards, as well as the potential of AI to violate or undermine them. AI can help protect privacy and enhance communication, but it can also enable surveillance and censorship, leading to a lack of trust in the organization.

Governance metrics [9,10,11]

Board composition and oversight

This metric indicates the composition and performance of a company's board of directors in relation to its use or development of AI systems. This may include board diversity, expertise, independence, engagement, or accountability. This metric is essential for assessing the level of competence and oversight of a company's board on AI issues. For example, a board with diverse backgrounds and skills can provide better guidance and direction for a company's AI strategy, but a board with insufficient AI literacy or involvement can fail to address the risks and opportunities of AI. Additionally, it's important to note that AI literacy at the board level is a critical factor for success.

Compliance and Codes of Conduct

This metric measures compliance, and ensures the quality and extent of a company's code of conduct to operationalize use and development of AI systems. This encompasses policies, procedures, training, audits, or sanctions. This metric addresses integrity, identifying mechanisms and incentives to align with a company's conduct code on AI issues. For example, a company with strong responsible AI standards and organizational culture can ensure that its AI systems satisfy legal requirements and are aligned with company values. In contrast, a company with weak code of conduct on AI issues can expose itself to violations and harmful ethical dilemmas.

Governance and accountability

This metric measures the extent and quality of a company's governance and accountability mechanisms for ensuring the ethical and responsible use or development of AI systems are in place. This may comprise governance structure, roles and responsibilities, risk assessment and management processes, controls and standards implementation, monitoring and reporting mechanisms. This metric assesses the effectiveness of a company's AI governance and accountability practices. For example, a company with robust AI governance and accountability processes can ensure that its AI systems are transparent, fair, reliable, secure, and beneficial for all stakeholders. It has protocols in place to address third-party vendors, use of out-of-date models, and unknown data sources. A company with inadequate AI governance and accountability practices can expose itself to reputational damage, legal liability, or social backlash.

3. Use Cases Exploring the Impact of Artificial Intelligence on Selected ESG Metrics

The multitude and variety of AI use cases are quickly evolving, such that use cases and their respective impacts will never be exhaustive or impacts completely identified. This analysis will involve discerning the measurable positive and negative of using AI on each metric. The use cases are meant to be informative and illustrative. Additionally it's worth noting that many use cases might have positive and/or negative impacts that could fall under a different category. Use cases may align with one business objective and interfere with another. The impact of AI may be clearly beneficial for one task of a function and conflict with another. Beneficial impacts may be unevenly distributed across locations, and secondary impacts may reveal externalities that can be mitigated.

Use cases, associated with or enabled by AI, will be briefly described and be categorized for positive or negative impacts. Quantitative measures that are indicators of the output will not always be available for environmental, social, and governance metrics across all industries. For example, the actual carbon equivalent footprints of companies can be quantified at a snapshot in time. Whereas, governance's accountability metric is a composite view of systems in place.

Environmental metrics

Greenhouse gas emissions

POSITIVE IMPACT

- Methane Geo-Location Data: Methane leaks, one of the most potent greenhouse gases, can be pinpointed using satellite data and advanced big data analytics. This enables targeted mitigation measures to be implemented directly at the source.
- Forecasting and peak load shifting: AI predictive upgrades allow better load forecasting and planning assuming the shift away from fossil fuel-based generation at peak power usage and other real-time strategies for reducing power demand.
- Enterprise Optimization of Multimodal Transport for Reducing Carbon: Leveraging models and data from multiple sources to optimize transport towards lower-carbon fuels, while maintaining business timelines and supply chain objectives.
- Calculating carbon (equivalent) footprint: Organizations can take snapshots and derive carbon equivalent footprints based on their facilities' demand data and transportation types according to the portfolio of fuels utilized, using layered models and data integration. The output is used for reporting and to enable planning, rendering the results only positive if they prompt positive action.

NEGATIVE IMPACT

- Running AI increases energy demand by increasing processing capacity and data center cooling needs, thereby increasing GHG emissions. As long as AI demands more energy than alternatives, GHG emissions will accrue incremental energy, according to the fuel types in each respective location and when and how much incremental power is demanded.

Case Study: The Challenge with Generative AI and Large Language Models

The AI carbon footprint remains of high concern as AI continues to increase in use and popularity. According to Open AI researchers, since 2012, the computing power required to train cutting-edge AI models has doubled every 3.4 months. By 2040, it is expected that the emissions from the Information and Communications Technology (ICT) industry will reach 14% of the global emissions, with the majority of those emissions coming from the ICT infrastructure, particularly data centers and communications networks."[12] Training a single large language model (LLM) like ChatGPT can emit up to 626,155 tons of CO2e, which is more than the lifetime emissions of five average American cars, including fuel.[13] This is because LLMs require massive amounts of data and computing power, which consume a lot of electricity from the grid, often generated by fossil fuels.

The GHG emission, energy consumption, and water use of training generative AI models like large language models can be significant. For example, some estimates suggested that training a single large language model can use around 300 to 500 tons of CO2, about 60 times the average person's annual carbon footprint.[14] This is a significant amount of energy, especially considering that many AI models are trained on large datasets and require a lot of computing power. A separate study also estimates that large data centers used to train and run these large AI models can consume 10s of billions of liters of fresh water.[15] This does not even include the footprint of hardware manufacturing, transportation, and infrastructure overheads.

Energy consumption

POSITIVE IMPACT

- Dynamic optimization and efficiency resulting in conservation: Utilizing models from multiple data sources to optimize operations in real-time modifies commercial or industrial processes at the organization or facility level.
- Efficiency product for end-use consumer: Platform or Internet of Things (IOT)- based products that lower energy demand by using smart meters, predictive modeling, or sensors.

Case Study: Energy Efficiency in the Manufacturing Process

IBM used AI to improve the energy efficiency and performance of its semiconductor manufacturing process by 30%, using AI systems (machine learning) to analyze sensor data and optimize the parameters of the production equipment.[16] This reduced the energy consumption of the process by 1.5 MWh per wafer, equivalent to the annual electricity consumption of 150 US households.[17]

Energy consumption

NEGATIVE IMPACT

- Iterative nature of using AI: Both competition-driven process improvements and Responsible AI necessitate continually refining algorithms and training new models on updated or additional sources of data. This continuous improvement, ensuring the AI is doing what is desired and not inadvertently introducing unaccounted-for risks, requires increased energy use.
- Planned obsolescence: When a product cannot be repaired or be upgraded to function or interface with current technology, more energy is used to manufacture and distribute that item.

Water consumption

POSITIVE IMPACT

- Water use optimization: AI-based optimization for large-scale commercial and manufacturing improves processes, reducing primary water consumption. (Twenty percent of global water use is industrial; AI enables mitigation in water-stressed locations.[18])

NEGATIVE IMPACT

- Training of generative AI is water-intensive: Cooling the essential hardware, graphics cards and processing cores is critical.[19] According to a report, training Bard consumed 5.6 billion gallons of water in 2022, more than the annual water consumption of 51,000 US households.[20] Few current business models encourage responsible use for this impact.

Social metrics

Health and safety

POSITIVE IMPACT

- Enhancing workplace safety through alarms and sensors and anomaly detection: AI-driven alarms, sensors, and predictive maintenance enhance workplace safety by detecting hazards and scheduling timely repairs. This reduces accidents, operational disruptions, and costs, ensuring a safer and more productive environment.
- Early disease detection: AI analyzes medical data to identify early signs of diseases like cancer and cardiovascular conditions. This improves diagnosis and treatment, enhances patient outcomes, reduces healthcare costs, and eases the burden on healthcare systems.
- Emergency response and disaster management: AI predicts natural disasters and optimizes crisis management by analyzing real-time data and forecasting disaster impacts. This enhances resource allocation and emergency response efficiency, minimizing loss of life and property damage.
- Increased access to healthcare in remote settings: AI-powered telemedicine and diagnostic tools facilitate healthcare delivery in remote and underserved areas. This ensures access to quality medical care, bridges healthcare accessibility gaps, and provides continuous healthcare monitoring.

Case study: AI Accelerating Progress in Healthcare

AI contributes to early disease detection and more efficient treatment. For example, a study by researchers from Stanford University and Google showed that AI could analyze chest X-rays and identify signs of tuberculosis with a high level of accuracy. The AI system was trained on a large dataset of X-rays from different countries and settings, and it outperformed human radiologists in some cases. The study also demonstrated that AI could help reduce the costs of screening for tuberculosis by 80%.[21]

Health and safety

NEGATIVE IMPACT

- Bias in decision support tools: AI decision support tools can perpetuate and amplify biases in areas like hiring, loans, and law enforcement. This can lead to discriminatory practices, unfair outcomes, and social inequalities, undermining trust and transparency.

Diversity, equity, and inclusion

POSITIVE IMPACT

- Skills matching: AI-driven tools analyze resumes, job postings, and market trends to match individuals with job opportunities that suit their skills and aspirations. This improves job placement success, reduces hiring time, and helps build capable teams.
- Communication support tools: AI-powered translation and language processing applications improve communication across languages and cultures. These tools enhance global business interactions, customer service, and inclusivity, supporting individuals with disabilities.

NEGATIVE IMPACT

- Emotion recognition: AI systems that analyze facial expressions and voice tones can invade privacy and lead to ethical concerns. They risk inaccurate assessments and misuse, potentially for manipulation in marketing or surveillance.
- Bias in hiring and promotion practices: AI used in hiring and promotion can perpetuate biases if trained on biased data, resulting in unfair outcomes. This undermines diversity efforts and can have legal and reputational consequences for organizations.
- Privacy concerns with personal information: AI systems collect and process vast amounts of personal data, raising privacy concerns. Misuse or breaches of sensitive data can lead to identity theft and financial loss, eroding public trust in AI technologies.

Case study: Emotion Recognition and DEI

This case study highlights the negative impact of AI on Diversity, Equity, and Inclusion (DEI) at Estée Lauder. [22] Three makeup artists from Estée Lauder's subsidiary, MAC, were made redundant based on an algorithmic assessment conducted through a video interview. The algorithm, developed by HireVue, evaluated their responses and expressions, raising concerns about the fairness and transparency of such automated decision-making processes. The incident, which led to legal action and an out-of-court settlement, underscores the potential for AI to adversely affect DEI in the workplace, highlighting the need for careful consideration of AI's role in employment decisions.

Human rights and labor standards[23]

POSITIVE IMPACT

- Monitoring and Reporting: AI systems can monitor labor conditions and human rights abuses in real-time, providing data for NGOs and governments to take timely action.
- Supply Chain Transparency: AI can track and analyze supply chains to ensure ethical sourcing and fair labor practices, helping to identify and address violations.

NEGATIVE IMPACT

- Job displacement: AI automation affects workers in a variety of categories and skill levels, leading to job displacement and retraining challenges. This can exacerbate social and economic inequalities, particularly for vulnerable and aging populations.
- Election interference and censorship: AI technologies can spread misinformation and censor dissent, posing threats to democratic processes and free speech. This can distort electoral outcomes and undermine public trust in democratic institutions.
- Data labeling and content monitoring practices: Reliance on low-paid labor for AI data labeling can lead to exploitation and poor working conditions. Additionally, AI content monitoring can perpetuate biases and inaccurately flag or censor content, impacting free expression.

Case study: AI and Social Compromise

One case study for the negative impact of AI on human rights and labor standards is the use of AI to enable surveillance and censorship. For example, using an empirical approach, a paper from the Science and Engineering Ethics journal investigates the ethical use of Big Data and Artificial Intelligence (BD+AI) technologies.[24] The paper presents a multi-case study of 'on-the-ground' ethical issues that uses qualitative tools to analyze findings from ten targeted case studies from a range of domains. The paper shows that BD+AI can pose risks to users, such as privacy infringements, discrimination, security concerns, and increasing inequalities.[25] One of the case studies examined in the paper is the use of facial recognition technology by the Chinese government to monitor and control its citizens, especially ethnic minorities such as Uyghurs. The paper argues that this use of AI violates the right to privacy, the right to freedom from discrimination, and the right to dignity.[26] Location history data is collected and sold to third parties and federal agencies.[27] DHS's CBP and Immigrations and Customs Enforcement (ICE) arms used location data from a company called Venntel to locate undocumented immigrants and routes they used to cross the border - this data could be used to enable potential deportation or arrests.[28]

Governance metrics

Board Composition and Oversight

POSITIVE IMPACT

- AI tools for improving board performance: Leveraging AI tools to provide agreed upon metrics and more frequent updates between the Board and management teams allow directors to identify risks sooner. ESG performance metrics allow Boards to track progress on ESG goals. It remains imperative that Boards have a strong fundamental understanding and literacy of AI and how AI interacts with environmental and social matters to harness this benefit.

NEGATIVE IMPACT

- Over-reliance on AI: Boards risk becoming overly reliant on AI-generated recommendations, which may erode the application of human judgment in their decision-making processes. While powerful, AI technologies often fall short of fully appreciating qualitative or ethical considerations. This oversight can lead to decisions that, while seemingly optimal from a data-driven perspective, may be suboptimal due to the omission of crucial contextual nuances or responsible dimensions.

Case Study: Canal Capital Corp. v. French

The legal framework within which some companies operate, particularly those incorporated in Delaware or governed by Delaware law, may further constrain the extent to which AI can be utilized in decision-making. This is exemplified by the case of Canal Capital Corp. v. French, where the Delaware Court of Chancery highlighted the principle of non-delegation. In this ruling, it was established that directors cannot delegate "in a very substantial way their duty to use their own best judgment on management matters" to non-directors. This legal precedent highlights a significant limitation on delegating critical decision-making processes to AI, reinforcing the importance of direct human oversight and judgment in corporate governance. [29]

Compliance and Codes of Conduct

POSITIVE IMPACT

- Automated systems reduce risks of regulatory non-compliance: AI can assist organizations in ensuring adherence to complex and evolving regulatory frameworks by automating compliance checks and reporting.[30] Additionally, emerging AI risks and future regulatory requirements are encouraging organizations to set up diverse and cross-functional review boards.
- AI-driven ESG data analytics: Data analytics and predictive analytics can process a wide range of data related to environmental impact, social responsibility, and corporate governance metrics to identify changes in performance, both regulatory- and mission-aligned conduct,and anomalous behavior.

NEGATIVE IMPACT

- Data privacy concerns: AI systems often require large amounts of data to function effectively. Collecting, storing, and processing this data can raise significant privacy concerns and potentially violate data protection regulations such as GDPR or CCPA.
- Security vulnerabilities: AI systems can introduce new security risks, such as adversarial attacks where malicious actors manipulate input data to deceive the AI. Ensuring that these systems are secure and compliant with security standards is a significant challenge.
- Regulatory lag: The rapid pace of AI development often outstrips the ability of regulatory bodies to create appropriate guidelines and standards. This lag can result in a lack of clear directives for organizations to follow, leading to compliance uncertainties.

Case Study: AI and Cybersecurity Risk

In 2023, a staggering 8,214,886,660 records were breached,[31] illustrating a significant escalation in cybersecurity vulnerabilities. AI integration has paradoxically fortified defenses and expanded the attack surface for cyber threats. Notable incidents have underscored the dual-edged nature of AI in cybersecurity.

For instance, the WannaCry ransomware attack exploited vulnerabilities to encrypt data and files on 230,000 computers across 150 countries, severely impacting organizations worldwide, including crippling operations within the NHS. [32] Similarly, Cambridge Analytica's unauthorized collection and use of personal data from millions of individuals demonstrated a profound breach of privacy and trust, influencing national elections around the globe. Another alarming example is the use of Deepfakes to deceive and defraud, including a notable case where a CEO was scammed out of $243,000 through a voice deepfake.[33]

These incidents have not only highlighted the innovative misuse of AI for malicious purposes but also raised concerns about the ethical implications and the potential for misuse of powerful AI technologies. According to the World Economic Forum's 2024 Risk Report, Misinformation and Disinformation, along with Cybersecurity, are ranked as the first and fourth most severe risks in the short term, respectively.

Case Study: AI and Cybersecurity Risk

In 2023, a staggering 8,214,886,660 records were breached,[31] illustrating a significant escalation in cybersecurity vulnerabilities. AI integration has paradoxically fortified defenses and expanded the attack surface for cyber threats. Notable incidents have underscored the dual-edged nature of AI in cybersecurity.

For instance, the WannaCry ransomware attack exploited vulnerabilities to encrypt data and files on 230,000 computers across 150 countries, severely impacting organizations worldwide, including crippling operations within the NHS. [32] Similarly, Cambridge Analytica's unauthorized collection and use of personal data from millions of individuals demonstrated a profound breach of privacy and trust, influencing national elections around the globe. Another alarming example is the use of Deepfakes to deceive and defraud, including a notable case where a CEO was scammed out of $243,000 through a voice deepfake.[33]

These incidents have not only highlighted the innovative misuse of AI for malicious purposes but also raised concerns about the ethical implications and the potential for misuse of powerful AI technologies. According to the World Economic Forum's 2024 Risk Report, Misinformation and Disinformation, along with Cybersecurity, are ranked as the first and fourth most severe risks in the short term, respectively.

Governance and Accountability

POSITIVE IMPACT

- Cross-functional governance: Implementing AI requires a cross-functional structure that improves governance. Where corporate subcommittees often oversee critical issues, the integration of AI into daily operations has led to recognizing the importance of governance and systems stewardship that play a pivotal role in guiding AI initiatives.
- Enhancing accountability: AI can improve communication between the Board and management teams and ownership of operational functions, enhancing accountability.

NEGATIVE IMPACT

- Over-reliance on AI for solution identification: Overdependence on AI can undermine human judgment. Boards risk becoming overly reliant on AI-generated recommendations, which may erode the practice of critical thinking and the application of human judgment in their decision-making processes. While powerful, AI technologies often fall short of fully appreciating qualitative or ethical considerations. This oversight can lead to decisions that, while seemingly optimal from a data-driven perspective, may be suboptimal due to the omission of crucial contextual nuances or ethical dimensions.
- Ethical concerns about AI systems continue to evolve as the AI industry continues to evolve and develop, and risks emerge among the general public. The complexity of AI often renders their decision-making processes opaque to managers up and down business lines, which can obscure accountability.

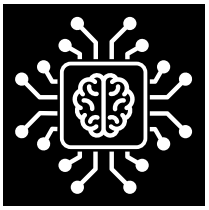
4. A Guiding Framework for Leveraging AI to Advance Achieving ESG Goals

ENVIRONMENTAL **SOCIAL** **GOVERNANCE**

Improving the performance and enhancing the validity of Environmental, Social, and Governance (ESG) metrics through Artificial Intelligence (AI), organizations must adopt a guiding framework that ensures the responsible use of AI. This document outlines key considerations and steps that decision-makers should follow to align their AI initiatives with ESG objectives effectively.

Improving the performance and enhancing the validity of Environmental, Social, and Governance (ESG) metrics through Artificial Intelligence (AI), organizations must adopt a guiding framework that ensures the responsible use of AI. This document outlines key considerations and steps that decision-makers should follow to align their AI initiatives with ESG objectives effectively.

IMPLEMENT RESPONSIBLE AI



- Ensure AI systems are implemented responsibly and prioritize key principles such as Valid and Reliable, Explainable and Interpretable, Accountable and Transparent, Privacy-Enhanced, Fair, Safe and Secure and Resilient. What are the safeguards and guardrails to ensure these principles are met?
- Map and understand potential AI risks or impacts, including biases and unintended consequences, through proactive risk mitigation strategies and robust AI governance. Are mitigation measures appropriate given the organization's industry, jurisdiction, and major AI use cases?
- Ensure inclusivity in AI teams, risk analyses, and data, ensuring that diverse perspectives are represented throughout the AI system lifecycle and AI systems are responsive to the needs of affected populations and ESG concerns.

continues on page 16

PURPOSE AND STRATEGIC FIT:

- Identify the ESG challenges that your AI efforts aim to address. How do these efforts align with specific ESG goals?
- Ensure that AI initiatives are in harmony with your organization's overarching ESG strategy and ethical principles.

IMPACT EVALUATION:

- Analyze both the potential positive outcomes and risks associated with applying AI to ESG strategies. What are the anticipated impacts?
- Establish methods to assess AI's contribution quantitatively and qualitatively to ESG objectives.

ETHICAL DATA PRACTICES:

- Confirm the availability of high-quality, ethically sourced data for AI analysis in ESG contexts. Are your data practices sustainable and ethical?
- Scrutinize the data collection methodologies to ensure they align with best practices in ethical data sourcing, sustainability and existing regulations.

TRANSPARENCY AND GOVERNANCE:

- Maintain open communication about AI-driven ESG initiatives with all stakeholders, ensuring transparency.
- Develop and implement governance frameworks to manage AI decisions, promoting accountability and ethical oversight.

DATA SECURITY:

- Implement stringent data privacy and security measures to protect sensitive ESG information, in compliance with legal standards.
- Proactively safeguard against data breaches and security vulnerabilities in AI projects.

ENGAGEMENT AND COLLABORATION:

- Involve a broad spectrum of stakeholders in the AI implementation process, fostering an inclusive approach to decision-making.
- Promote collaboration with industry peers, experts, and civil society to refine and enhance AI strategies for ESG.

MONITORING AND ADAPTATION:

- Define and utilize key performance indicators (KPIs) to monitor the effectiveness of AI in achieving ESG goals.
- Regularly review and adjust AI strategies in response to new ESG priorities, technological advances, and stakeholder feedback.

This structure of inquiry serves as a starting point for decision-makers to reflect on the responsible use of AI in the context of ESG metrics. Answers to these questions can help an organization develop a comprehensive framework that aligns AI initiatives with ESG goals.

5. Conclusion and Future Direction

There is immense potential for shaping a sustainable and equitable future at the intersection of artificial intelligence (AI) and environmental, social, and governance (ESG) criteria. The brief analyses of metrics, comprising multiple use cases, lend themselves to the proposed framework for integrating Responsible AI (RAI) within the ESG paradigm.

Here we have demonstrated the critical importance of aligning AI initiatives with an organization's broader goals of sustainability, social responsibility, and responsible governance.

AI presents a dual-edged sword, capable of driving significant advancements in ESG metrics, while posing unique risks and challenges. Our discussion of use cases revealed how AI can: optimize energy use; reduce greenhouse gas emissions; enhance diversity and inclusion; improve health and safety standards; and bolster governance structures when operationalizing responsible AI. Simultaneously, we highlighted concerns around AI's energy and resource consumption, potential for bias, privacy breaches, and social dilemmas.

Future Direction:

Integrating AI into ESG strategies, and ESG into AI strategies, is not a static goal but an ongoing journey marked by continuous learning, adaptation, and improvement. As we look to the future, several areas beckon further exploration and development:

- **Technological Advancements:** Keeping pace with rapid advancements in AI technologies to leverage their potential for ESG improvements while mitigating associated risks.
- **Multidisciplinary Systems Leadership:** Supporting a culture or an organizational model that has the capacity to manage multiple systems and address cross-functional alignments.
- **Regulatory Frameworks:** The evolution of global regulatory landscapes for AI and ESG will necessitate adaptive strategies to comply with emerging standards and leverage regulatory guidance for innovation.
- **Data Integrity and Accessibility:** Enhancing the quality, transparency, and accessibility of data for AI systems will be crucial in driving ESG outcomes. This includes addressing data bias and ensuring diverse, representative datasets that are ethically sources and labeled.
- **Stakeholder Engagement:** Deepening engagement with a broad spectrum of stakeholders, including marginalized and vulnerable communities, to ensure AI-driven ESG initiatives are inclusive and equitable.
- **Responsible AI Development:** Advancing the conversation around responsible AI, focusing on developing AI systems that are not only technically proficient but also socially and environmentally conscious.
- **Cross-Sector Collaboration:** Fostering cross-sector collaborations to share best practices, co-create solutions, and scale successful AI-driven ESG initiatives.

Call to Action:

The urgency of adopting RAI principles within ESG practices cannot be overstated. Businesses can begin by implementing the proposed framework for integrating RAI into their ESG strategies. Individuals can advocate for RAI within their organizations and support companies committed to responsible AI development. Policymakers can work to establish clear guidelines for RAI in ESG, fostering innovation while mitigating risks. Through a collective effort, we can unlock the immense potential of AI to create a more sustainable, just, and resilient future.

6. Contributors List

This project was made possible through the contributions and expertise of the following individuals:

Lead Authors

- Alyssa Lefaivre Škopac, Head of Global Partnerships and Growth, Responsible AI Institute. Co-Founder of the Responsible AI and ESG Leadership Working Group.
- Andrea Ruotolo, Global Head, Customer Sustainability, Rockwell Automation. Co-Founder of the Responsible AI and ESG Leadership Working Group.
- David Carlin, Head of Risk, UN Environment Program Finance Initiative
- Deval Pandya, Vice President AI Engineer, Vector Institute
- Rosalie Day, Responsible AI and ESG Expert Consultant
- Roshni Joshi, Managing Director of Customer Engineering, US Central, Google

Contributors

- Elena Morettini, Global Head of Sustainable Business, Globant
- James Gallagher, Adjunct Professor for Energy Regulation and Restructuring, Clarkson University Graduate School of Engineering
- Johan Stahre, Chair Professor of Production Systems, Chalmers University of Technology
- John Knights, Head of Services, GRI
- Lorenzo Saa, Chief Sustainability Officer, Clarity AI
- Marsal Gavaldà, Chief Technology Officer, Clarity AI
- Micah Kotch, Partner, Blackhorn Ventures
- Murali Vijendran, Founder and Director, SigmaRed Technologies Inc.
- Murthy Swati, Director, Sustainability Solutions, Tata Consultancy Services
- Nabiha Megateli-Das, CIO, Mindful M&A
- Nicholas Buccheri, Sustainability Communications Consultant
- Øistein Jensen, Chief Sustainability Officer, Odfjell SE
- Paul Pilotte, Technical Marketing Manager, AI, MathWorks
- Sarah Murphy Gray, Senior Program Manager, Responsible AI, Google
- Sedef Akinli Kocak, Director, Industry Innovation, Vector Institute
- Sergio Alberto Gramitto Ricci, Associate Professor of Law, UMKC Law School
- Sinethemba Zulu, Sustainability Technology Architect, Rockwell Automation
- Tadeo Rodriguez, Customer Sustainability Operations Lead, Rockwell Automation
- Tess Buckley, Programme Manager for Digital Ethics and AI Safety, techUK
- Theresa Kennedy, Partner, Capability & Data Excellence Center, ERM
- Var Shankar, Chief AI and Privacy Officer, Enzai
- Visanty Dindial, Director, VKD Transport Ltd.

7. Appendix- Additional Case Study

A Simple Energy Discussion

Perspective on Relative Scale of Energy Increase Associated with AI Use

AI as a tool is energy intensive. With increasing energy requirements for optimum operations and constant iterations for improvement, power demand is on an upward trajectory. The strong demand to use AI for competitive edge in the marketplace AI results in excessive energy consumption. In some cases, AI servers are expected to utilize more energy than countries. “By 2027, AI servers could use between 85 to 134 terawatt hours (Twh) annually. That’s similar to what Argentina, the Netherlands and Sweden each use a year, and is about 0.5% of the world’s current electricity use.”[34] Energy consumption and carbon emissions are intrinsically linked and the rapid increase in energy consumption can catalyze carbon emissions growth.

Renewable Energy

Power grid operations have mastered electricity reliability (the physics of maintaining constant supply) challenges of intermittent renewable energy generation (solar and wind). This constancy relies on power sources that can be dispatched instantaneously, or demand that can be decreased, again instantaneously. For example, solar and wind generated energy are not always available when energy consumers demand it. Power grid reliability is backstopped with fossil fuel generation, especially natural gas, again due to physical realities.

Therefore, if companies state they are powered by all renewables energy, they have chosen to buy renewables energy certificates in the market (as a substitute from a different time or location) or choose to draw no power when renewables sources are not available.

Power can be stored. When wind energy is in excess (and nuclear power can be drawn off) but in the current state of storage, there is not sufficient storage. “Dispatchable” energy storage technologies are promising, Hydro-generation (using the excess energy to pump water) is especially promising because it relies on proven interfaces, existing turbine technology and existing reservoirs with transmission lines. Other infrastructure, like micro grids or district heating/cooling that require smaller storage solutions, but with vast growth in consumption, remedies growth on the margins.

8. Citations

[1] "EU's New ESG Reporting Rules Will Apply to Many US Issuers ...." 28 Oct. 2022, https://www.cooley.com/news/insight/2022/2022-10-28-eus-new-esg-reporting-rules-will-apply-to-many-us-issuers. Accessed 22 Jan. 2024.

[2] "Corporate Sustainability Reporting Directive (CSRD) Overview and ...." https://www.georgeson.com/us/insights/csrd-overview-sep-2023. Accessed 22 Jan. 2024.

[3] "CSRD vs SEC Climate Disclosure Rule: Similarities & Differences." 24 May. 2022, https://www.perillon.com/blog/csrd-vs-sec-climate-disclosure-rule-similarities-differences. Accessed 22 Jan. 2024..

[4] "The AI Angle in Solving the Oil and Gas Emissions Challenge." 20 Sep. 2021, https://www.bcg.com/publications/2021/ai-in-oil-and-gas-emissions-challenge. Accessed 22 Jan. 2024.

[5] "Measuring Emissions Accurately | BCG - Boston Consulting Group." https://www.bcg.com/publications/2021/measuring-emissions-accurately. Accessed 22 Jan. 2024.

[6] "Measuring Emissions Accurately | BCG - Boston Consulting Group." https://www.bcg.com/publications/2021/measuring-emissions-accurately. Accessed 22 Jan. 2024.

[7] "AI Is Essential for Solving the Climate Crisis | BCG." 7 Jul. 2022, https://www.bcg.com/publications/2022/how-ai-can-help-climate-change. Accessed 22 Jan. 2024.

[8] "How much water does AI consume? The public deserves to know." 30 Nov. 2023, https://oecd.ai/en/wonk/how-much-water-does-ai-consume. Accessed 22 Jan. 2024

[9] "CSRD vs SEC Climate Disclosure Rule: Similarities & Differences." 24 May. 2022, https://www.perillon.com/blog/csrd-vs-sec-climate-disclosure-rule-similarities-differences. Accessed 22 Jan. 2024.

[10] "EU's New ESG Reporting Rules Will Apply to Many US Issuers ...." 28 Oct. 2022, http://www.cooley.com/api/downloadpdf?contextItemId=%7B4EABB813-9707-46BF-9F4C-FAD471A80179%7D. Accessed 22 Jan. 2024.

[11] "Assurance under US SEC Climate Reporting Rules & EU CSRD." 4 Jul. 2022, https://nordesg.de/en/assurance-under-us-sec-climate-reporting-rules-eu-csrd/. Accessed 22 Jan. 2024.

[12] "The Real Environmental Impact of AI | Earth.Org." 18 Jul. 2023, https://earth.org/the-green-dilemma-can-ai-fulfil-its-potential-without-harming-the-environment/. Accessed 22 Jan. 2024.

[13] "AI can estimate corporate greenhouse gas emissions | VentureBeat." 13 Sep. 2021, https://venturebeat.com/business/ai-can-estimate-corporate-greenhouse-gas-emissions/. Accessed 22 Jan. 2024.

[14] "Generative AI doesn't have to generate climate disaster - CIFAR." 25 Sep. 2023, https://cifar.ca/cifarnews/2023/09/25/generative-ai-doesnt-have-to-generate-climate-disaster/. Accessed 22 Jan. 2024.

[15] "Uncovering and Addressing the Secret Water Footprint of AI Models." 6 Apr. 2023, https://arxiv.org/abs/2304.03271. Accessed 22 Jan. 2024.

[16] "Chapter 5 - Case Studies of GHG Emissions Analysis." https://nap.nationalacademies.org/read/22805/chapter/7. Accessed 22 Jan. 2024.

[17] "The AI Boom Could Use a Shocking Amount of Electricity." 13 Oct. 2023, https://www.scientificamerican.com/article/the-ai-boom-could-use-a-shocking-amount-of-electricity/. Accessed 22 Jan. 2024.

[18] "Water Consumption Statistics - Parks & Rec Business (PRB)." 30 Sep. 2023, https://www.parksandrecbusiness.com/articles/water-consumption-stats. Accessed 22 Jan. 2024.

[19] "Microsoft used 1.7 billion gallons of water for AI in 2022." 10 Sep. 2023, https://interestingengineering.com/innovation/microsoft-used-17-billion-gallons-of-water-for-ai-in-2022. Accessed 25 Jan. 2024.

[20] "Lightweight AI Framework for Industry 4.0 Case Study: Water Meter ...." 1 Jul. 2022, https://www.mdpi.com/2504-2289/6/3/72. Accessed 25 Jan. 2024.

[21] "Tackling tuberculosis screening with AI - The Keyword." 18 May. 2021, https://blog.google/technology/health/tuberculosis-screening-ai-io-2021/. Accessed 22 Jan. 2024.

[22] "Estée Lauder staff win payout after being 'sacked by algorithm'." 17 Mar. 2022, https://www.personneltoday.com/hr/estee-lauder-women-sacked-by-algorithm-redundancy-software-hirevue-automation/. Accessed 22 Jan. 2024.

[23] "AI Will Transform the Global Economy. Let's Make Sure It Benefits ...." 14 Jan. 2024, https://www.imf.org/en/Blogs/Articles/2024/01/14/ai-will-transform-the-global-economy-lets-make-sure-it-benefits-humanity. Accessed 22 Jan. 2024.

[24] "Research and Practice of AI Ethics: A Case Study Approach ...." 8 Mar. 2021, https://link.springer.com/article/10.1007/s11948-021-00293-x. Accessed 22 Jan. 2024.

[25] "Evaluating the impact of artificial intelligence on human rights." 27 Sep. 2018, https://hls.harvard.edu/today/evaluating-the-impact-of-artificial-intelligence-on-human-rights/. Accessed 22 Jan. 2024.

[26] "Research and Practice of AI Ethics: A Case Study Approach ...." 8 Mar. 2021, https://link.springer.com/article/10.1007/s11948-021-00293-x. Accessed 22 Jan. 2024.

[27] "A surprising number of government agencies buy cellphone location ...." 2 Dec. 2020, https://www.vox.com/recode/22038383/dhs-cbp-investigation-cellphone-data-brokers-venntel. Accessed 22 Jan. 2024.

[28] "Federal Agencies Use Cellphone Location Data for Immigration ...." 7 Feb. 2020, https://www.wsj.com/articles/federal-agencies-use-cellphone-location-data-for-immigration-enforcement-11581078600. Accessed 22 Jan. 2024.

[29] https://djcl.org/the-future-of-using-artificial-intelligence-in-delaware-corporate-governance/#:~:text=Conclusion,completely%20replace%20boards%20of%20directors.

[30] https://www2.deloitte.com/content/dam/Deloitte/us/Documents/us-tte-annual-report-2023.pdf

[31] https://www.itgovernance.co.uk/blog/list-of-data-breaches-and-cyber-attacks-in-2023

[32] https://www.thelancet.com/journals/landig/article/PIIS2589-7500(19)30005-6/fulltext

[33] https://www.forbes.com/sites/jessedamiani/2019/09/03/a-voice-deepfake-was-used-to-scam-a-ceo-out-of-243000/?sh=6634ba672241

[34] "A.I. Could Soon Need as Much Electricity as an Entire Country." 10 Oct. 2023, https://www.nytimes.com/2023/10/10/climate/ai-could-soon-need-as-much-electricity-as-an-entire-country.html. Accessed 22 Jan. 2024.

Supporting You on Your RAI Journey

The Responsible AI Institute is here to support organizations on their AI journeys. Becoming a member enables essential support and direction for making significant progress toward future-proofing RAI governance and implementation.

Click below to learn more about how we help our members achieve their Responsible AI goals:

About Responsible AI Institute (RAI Institute)

Founded in 2016, the Responsible AI Institute (RAI Institute) is a global and member-driven non-profit dedicated to enabling successful responsible AI efforts in organizations. We accelerate and simplify responsible AI adoption by providing our members with AI assessments, benchmarks and certifications that are closely aligned with global standards and emerging regulations.

Where to connect with us:

YouTube