THE DIGITAL ECONOMIST

ARTIFICIAL INTELLIGENCE | BUSINESS PRACTICE | GOVERNANCE

Expert Insights

Jose Carvalho, Melodena Stephens & Marisa Zalabak

From Hype to Norm

Embedding Responsible AI In Everyday Business Practices

1.0 Introduction

Artificial Intelligence (AI) is the defining technology of our times. According to the AI Index 2023, 80% of Fortune 500 companies started mentioning AI in their earning calls. Recently, Responsible AI has become a buzzword among industry leaders and policymakers. The OECD reports over 1000 AI policies and initiatives worldwide. A review of 200 guidelines and recommendations from 37 countries reveals that "a key challenge lies in establishing a consensus on these values." Despite the widespread adoption of many of the terms, substantial confusion persists regarding the definition of AI itself, and even more significant confusion exists on how to implement and maintain responsible AI practices, from leadership to operations.

As such, company and organizational practices vary greatly. Some use accreditations like ISO/IEC 42001: 2023 or the IEEE CertfAI, and others build it in-house with minimal external scrutiny. Additionally, the reliance on outsourcing and licensing models raises questions about responsibility and accountability. Furthermore, the dismissal of prominent AI Ethics teams and team members from industry giants like Google, Microsoft, Twitter, Amazon, and Open AI, among others, further exacerbates concerns about responsible AI adoption.

AI systems transcend geographical and national boundaries, creating challenges at an international level. Countries have adopted various manifestos, such as UNESCO's Recommendation on the Ethics of Artificial Intelligence, which was adopted by 193 countries in November 2021, the G7 Hiroshima AI Process, the Council of Europe's Framework Convention on Artificial Intelligence and Human Rights, Democracy, and the Rule of Law and OECD's AI Principles, initially used in 2019 and updated in 2024, adopted by 46 countries and the EU. However, even legally binding treaties like the Council of Europe Framework Convention on artificial intelligence and human rights, democracy, and the rule of law (CETS No. 225) are difficult to implement, as demonstrated by the difficulty in translating topics like explainability into practice, as demonstrated in the context of the Global South.

These challenges in identifying a values-based ethical framework, the lagging interconnected regulations, and the inability to address responsibility with the exponential adoption of AI have put pressure on the private sector to “do the right thing.” To achieve it, being aware of potential risks and pitfalls is crucial, and, as highlighted by the recent Crowdstrike incident, preparedness is essential. Given the speed at which the wide array of AI technologies are being deployed, there are many factors and challenges to address.

1. Navigating the Impact of AI: Understanding Value Chains and Incident Control Post-Crowdstrike Update

When the Crowdstrike update was released globally, it impacted 85 million Microsoft PCs and industries like health, airlines, and banking, among others. Firms using AI need a better understanding of their value chain and their control of AI incidents and responsibilities.

2. The Outsourcing Dilemma: Balancing Cost Savings with Critical Operational Expertise

Outsourcing to save costs risks leaving companies with little or no talent knowledgeable on how backend operations work or the decision choices and rationale why the company took them.

3. Overlooked Validation: The Risks of Bypassing Hardware and Software Certification in Smaller Firms

It is not uncommon to find that the validation of hardware and software is often bypassed, particularly in smaller firms that cannot certify products.

4. Navigating the Digital Divide: The Challenge of Insufficient AI Literacy Among Leadership in Managing Responsible AI Initiatives

More often than not, digital and AI literacy are confused together due to poor board and senior management knowledge of AI literacy skills, which adds to the challenges in managing responsible AI initiatives.

5. The Data Acquisition Challenge: Navigating the Risks of Buying and Leasing Data Amidst Lack of Standardization

Due to a lack of standardized AI norms and little understanding of data validation and certification methods, data is often bought or leased.

6. Accountability in AI: The Risks of Outsourcing Responsibility to Third Parties and the Need for Minimum Internal Expertise

AI responsibility is often outsourced to third parties, consultants, or other AI firms, making accountability harder. Should there be a minimum number of employees to ensure mindful and responsible AI for the valuation of an organization?

7. Lack of Role Clarity with AI: The Vicious Cycle of Irresponsibility and the Need for a Whole-Organization Approach

Roles such as Engineer, Developer, and Data Scientist may be confused with AI responsibility. When AI fails, the first indications are from the customer-facing teams, but as these may either be unauthorized to address failure or can even be outsourced to chatbots, this creates a vicious circle of irresponsibility. A whole-of-organization approach involving a multistakeholder and multidisciplinary method is essential for AI Ethics.

The private sector is increasingly adopting an ESG framework to showcase commitments to positive outcomes, yet the triple bottom line - **People, Planet, and Profit** - often overlooks AI and its increasing impact on business operations. ESG also often does not consider the implications of AI, as a recent EY report highlighted. Moreover, while there is a perception that AI will help us reach ESG objectives, the reality is that many organizations do not know where to begin and how to use AI responsibly.

2.0 The importance of adopting Responsible AI in businesses.

The rapid integration of AI into business presents opportunities for enhanced efficiency, competitiveness, and augmentation of capabilities and capacity, leading to a potential for growth as rarely seen before. For example, at the 2024 United Nations General Assembly, Sundar Pichai recently highlighted that Google Translate works across 246 languages out of the 1,000 most spoken languages in the world. Simultaneously, this potential comes with the challenge of aligning the application of this technology with ethical and sustainable practices. Not pursuing these benefits with this alignment in mind may lead to unintended consequences, such as privacy violations, bias perpetuation, societal impacts, and negative environmental impacts. The benefits and negative impacts of AI are unequally distributed, causing spillovers.

For example, the rising use of AI data centers will lead to a consumption of 6% of the USA's already strained electricity grid by 2026, creating massive carbon emissions and increasing water stress in places subject to drought.

AI adoption is a cross-cutting process simultaneously driven by individuals, businesses, and the public sector. We incorporate AI into daily tasks, both private and professional. According to a recent Forbes report, many businesses across industries are integrating AI into their operations, such as automation, data analysis and decision-making, customer service, and others. AI is also being leveraged to innovate and disrupt industries in sectors such as healthcare with diagnostic algorithms, finance, fraud detection, or retail, with personalized shopping assistance. The public sector is also making advances using AI in critical services such as education (e.g., personalizing learning assistance), transportation (e.g., managing and optimizing flows in smart cities), and public safety (e.g., employing predictive policing recommendations and facial recognition systems). Each of these innovations carries the potential for improving life along with the potential for harm if used without discernment and consideration of potential intended and unintended arms.

Given that AI technology is here and will continue to expand, the approach to its implementation needs to have a long-term view in sight. Prioritizing human well-being with autonomous and intelligent systems is essential to realize AI's full benefits.

Businesses are facing challenges in aligning AI practices with responsible AI principles.

There are many challenges when aligning AI practices with responsible AI principles (refer to Figure 1). According to the IBM Global AI Adoption Index 2023, 23% of firms surveyed had ethical concerns. The scale of adoption is accelerating; for example, Gartner predicts that GenAI, whose adoption was 1% in 2023, will increase to 50% by 2027, and over 75% of firms will be using synthetic data. The large-scale adoption, the low AI literacy, and increasing regulations will be challenging for firms.

Responsibility and Ethics in AI are often used interchangeably, yet ethics lies within responsibility. Ethics is a subset of the overarching issues in responsibility. These include cultural and philosophical differences. The distinction between responsible AI and AI ethics lies in their focus and application: Responsible AI is the umbrella that holds the practical implementation of ethical principles; it emphasizes the creation and use of AI in a way that is accountable, transparent, and with impacts that are aligned with societal values. For example, a company using a responsible AI framework might design its AI systems with explainability features, allowing users to understand how decisions are made to ensure trust and accountability. AI Ethics refers to the guiding principles and moral frameworks governing the development, deployment, and use of AI technologies influenced by local culture and philosophies. It focuses on what should be done to ensure fairness, privacy, and harm prevention (e.g., Ensuring that facial recognition systems do not perpetuate racial biases is an AI ethics issue, aiming to avoid harm and promote fairness - see Figure below).

Figure 1: Common Terms Used in Responsible and Ethical AI

Responsibility & Ethics in AI	Issues & Needs
Bias & Fairness	Reinforcement & amplification of existing prejudices. (e.g., biased datasets, lack of transparency of algorithms & mechanisms able to detect and mitigate bias).
Transparency & Explainability	Lack of explainable, interpretable & trustworthy systems for users & stakeholders.
More Personalized	User privacy that complies with data protection regulations & human rights.
Accountability & Responsibility	Clarity of entity responsible & accountable (e.g., developers, deployers, users, or the AI system itself). Lack of systems for redress.
Misinformation and Deepfakes	Preventing the malicious use of AI for deception while balancing free speech & creativity.
Environmental Impact	Assessment of AI research and application cost (e.g., energy, carbon, water, etc.)
Job Displacement & Economic Impact	Consideration of the transitions and social consequences of adoption, minimizing harms (e.g., equitable distribution & access, reskilling programs).
Security & Misuse	Robust security protocols are used in AI systems to prevent malicious use.
Long-Term Ethical Implications	Establishing AI design & development aligned with human rights that consider the long-term impacts on society, the human condition, and the living environment.

3.0 Introduction to the Analytical Perspective of People, Planet, and Profit

Revolutionary events caused by technology or other novel innovations come with a plethora of interconnected impacts. In businesses, the primary focus has traditionally been on the financial impacts. The growth of AI products and services is expected to increase by 30-40% over the next decade. Although these financial impacts are deeply affected by and also affect the social and environmental world, never before has the connection been so clear. There is an increasing understanding that all three factors - people, planet, and profit must be considered for meaningful success, especially highlighted and validated with the top global concerns of unstable economies, socio-political unrest, and climate change.

It is unusual for businesses to consider the ROI on a full range of impacts other than the direct impact on profits, yet direct impacts alone increasingly fail to plan for long-term success. The growing need to address these interconnected impacts in regard to advanced AI systems has resulted in emerging frameworks like the Holistic Return on Ethics (HROE). This, like other new forms of evaluation, considers all elements: financial (direct or tangible), social (indirect), and technical (capabilities). If designed well, AI can be a positive contributor to people, communities, and the environment. Though there is a significant focus on net zero, with tech companies buying 60% of renewable energy and on capacity building in developing countries, responsible AI needs so much more.

4.0 Key Areas for Responsible AI Implementation

Responsible AI has five key focus areas. While responsible AI first begins within a firm, it extends to its business ecosystem and the communities in which it operates. AI data and hardware components have a global impact (just like Scope 2 and Scope 3 emissions have), which cannot be ignored, and so is the fact that all involved stakeholders are affected by strategies, initiatives, and projects that implement AI. The five key focus areas are strategy, infrastructure, data, workforce, and governance. These are mapped across people, planet, and profit.

4.1 Strategy: Aligning AI initiatives with strategic goals

Adoption of Responsible AI at the strategic level within a business context requires a thoughtful and comprehensive approach that aligns with organizational values and ethical standards through its business DNA. Leadership plays a key role here, and fostering an active engagement in understanding the complexity and ethical norms that guide the usage of AI systems is necessary. This foundational knowledge is required to evaluate the purpose behind each AI design, implementation, adoption, acquisition, or even its decommission; that is, the AI strategy should look at all levels of integration between business, AI, and humans in the loop.

A robust strategy involves education and the refinement of protocols and policies that guide AI oversight throughout its lifecycle, alongside the establishment of accountability mechanisms and engagement with internal (leadership, employees) and external (customers, suppliers, partners, regulators) stakeholders.

Furthermore, assessing the social and environmental impact of new solutions, aligning objectives to the greater good, and avoiding negative consequences support a holistic and long-lasting usage of AI.

4.2 Infrastructure: Establishing robust technological and organizational frameworks

Incorporating AI into business operations has demands in the infrastructure strategy of an organization. Overlooking this impact may not only create unintended consequences for an ecosystem around a business but also impact the long-term operations and usage of AI.

Organizations need to implement robust practices that govern the scalability and upgradability of their infrastructure for the adoption of AI technology. This immediately affects costs and investments. Further, the bigger consumption of data and computational power, as well as its consequences on the existing estate and environment, needs to be managed accordingly.

Considerations also need to be made so that the infrastructure proactively protects all assets being used for such AI adoption, such as the protection against cybersecurity threats, by developing comprehensive security measures to safeguard sensitive data and user privacy.

Being prepared is vital. Disaster recovery and business continuity plans that address AI-related issues need to be in place to ensure minimum impact on stakeholders and high retention of trust.

4.3 Data: Ensuring ethical and efficient data management.

Developing a comprehensive data strategy that aligns with both business objectives and ethical standards is essential for the adoption of AI. Understanding that data requirements range from being ethically sourced, verified, and fit for purpose, stored, accessible, processed with purpose, protected, and decommissioned. These needs are guided not only by the intended use of the data but also by industry standards and regulations, legal requirements, management of financial risks, and the retention of stakeholder trust.

Ensuring data fairness is the cornerstone of Responsible AI practices. Identifying and mitigating biases within data to promote fairness and respect user consent, as well as ensuring principles of data sharing and collaboration, are among the essential practices to ensure are in practice.

However, it is not only the data in use that is relevant. All the data not in use should be considered, as retaining it and processing it without purpose creates an additional burden on the infrastructure, data management budgets, and, eventually, the environment. This is true for both own and third-party data.

4.4 Workforce: Enhancing skills and awareness among employees and leaders

Incorporating AI into business operations requires a strategic investment in talent to effectively manage a collaborative AI-human workforce. This talent is made of in-house expertise and is capable of overseeing AI integrations and interactions. It involves training programs to upskill employees and leaders, foster AI competencies, and ensure adherence to regulatory requirements, legal requirements, and industry standards.

Organizations must also facilitate the transition for employees whose roles evolve due to AI; in fact, the right protocols for human-AI collaboration need to be defined and can include Human-in-the-loop and Human-on-the-loop scenarios, maximizing the long-term outcomes of an effective partnership. These partnerships should be looked at holistically, ensuring that no one is left behind and all have the opportunity to use their competencies in a way that produces the best outcomes.

4.5 Governance: Developing policies and guidelines for ethical AI use

An effective AI adoption requires the establishment of well-defined policies that outline ethical standards, values, and processes that are recognized and nurtured by all stakeholders, such as employees, shareholders, suppliers, partners, and customers, for better management of AI systems. This foundational framework needs to align the AI initiative with the moral compass and strategic goals of the organization, enhancing transparency, respectability, and governance across all operations.

This governance framework needs to include national and international law and regulation adherence, as well as accountability for managing the risks and unintended consequences associated with AI technology. For example, monitoring impacts on employee rights and the work environment or the impact on the broader society and environment.

Advantages of a Comprehensive Approach

Initiating conversations about improvement and strategic alignment is crucial. People and skills are needed to initiate and facilitate transformation within and between systems, especially in the midst of dramatic changes like the AI boom. Often, the needed conversations are postponed, yet with the increasing speed of technological innovation, leaders must act proactively. Recognizing that many stakeholders may not be updated on the complexities involved, including emerging guidelines, these conversations should present opportunities for learning.

Systematically addressing AI implementation challenges provides significant benefits in the long-term adoption beyond immediate financial gains. Responsible AI systems offer returns on investments across multiple levels, which can be financial, reputation, greater resiliency during a crisis, and market and industry leadership]supporting long-term improvement and sustainability.

Although governance of AI systems may appear slow to be enacted, the trend is clearly toward adding specificity and reach, as in the recently released EU Act, Council of Europe Framework Convention on artificial intelligence and human rights, democracy, and the rule of law and UN's adopted Pact for the Future, Global Digital Compact and Declaration on Future Generations released during the UNGA79 Summit of the Future. Businesses and organizations that do not consider the implications of the technologies they are using, developing, or designing will be more at risk for multi-layered costs and consequences.

With trust in organizations increasingly questioned and concerns about possible harms, first and foremost, businesses should evaluate and enhance their responsible AI strategies to build trust with employees, investors, supply chain and communities of operations, and all the other stakeholders by demonstrating a commitment to the espoused practices. Responsible AI use can prevent potential legal and reputational risks associated with biases, discrimination, and privacy violations. Organizations need to regularly audit to stay aligned with responsible and ethical practices. This is a formative and iterative process rather than a one-time assessment.

Leaders of the future incorporating these considerations in AI development will significantly help in ensuring solutions are aligned with societal values and long-term sustainability, making these businesses and organizations trusted pioneers of innovation and success in this new frontier.

References

1. ”Artificial Intelligence Index Report 2024.” Stanford University [Human-Centered Artificial Intelligence]. Accessed September 15, 2024. https://aiindex.stanford.edu/wp-content/uploads/2024/05/HAI_AI-Index-Report-2024.pdf.

2. “National AI policies & strategies.” OECD.AI. Accessed September 15, 2024. https://oecd.ai/en/dashboards/overview/policy

3. Corrêa, et al., “Worldwide AI Ethics: A Review of 200 Guidelines and Recommendations for AI Governance.” Patterns 4, no. 10, 2023. Accessed September 15, 2024. https://www.sciencedirect.com/science/article/pii/S2666389923002416

4. “UNESCO Recommendation on the Ethics of Artificial Intelligence.” Global Observatory of Urban Artificial Intelligence. Accessed September 15, 2024. https://gouai.cidob.org/resources/recommendation-on-the-ethics-of-artificial-intelligence/#:~:text=In%20November%202021%2C%20the%20193%20Member%20States%20at,very%20first%20global%20standard-setting%20instrument%20on%20the%20subject.

5. “G7 Leaders’ Statement on the Hiroshima AI Process.” The White House. Accessed September 15, 2024. https://www.whitehouse.gov/briefing-room/statements-releases/2023/10/30/g7-leaders-statement-on-the-hiroshima-ai-process/

6. Ingrid, Lunden. “US, UK and EU Sign on to the Council of Europe’s High-level AI Safety Treaty,” TechCrunch. Accessed September 15, 2024. https://techcrunch.com/2024/09/05/us-uk-and-eu-sign-on-to-the-council-of-europes-high-level-ai-safety-treaty/#:~:text=The%20Council%20of%20Europe%20Framework%20Convention%20on%20Artificial,human%20rights%2C%20democracy%20and%20the%20rule%20of%20law.%E2%80%9D.

7. “OECD Legal Instruments.” OECD.org. Accessed September 15, 2024. https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0449#dates

8. “Treaty Office.” Council of Europe. Accessed September 15, 2024. https://www.coe.int/en/web/conventions/full-list?module=treaty-detail&treatynum=225

9. “A Global South Perspective on Explainable AI.” CARNEGIE Endowment for International Peace. Accessed September 15, 2024. https://carnegieendowment.org/research/2024/04/a-global-south-perspective-on-explainable-ai?lang=en

10. Virgilio, Almeida, et al. "The Origin and Evolution of Multistakeholder Models." IEEE Internet Computing, 2015. Accessed September 15, 2024. https://ieeexplore.ieee.org/document/7031820/authors#authors

11. Melodena, Stephens, et al. “A Primer for Education Thought Leadership: The Benefits of a Multidisciplinary Lens for Artificial Intelligence Systems Ethics, A Primer for Education Thought Leadership: The Benefits of a Multidisciplinary Lens for Artificial Intelligence Systems Ethics.” IEEE Standards Association, 2021. Accessed September 15, 2024. https://standards.ieee.org/wp-content/uploads/2023/07/ead-benefits-multidisciplinary-lens.pdf.

12. “Discussion Paper: Artificial intelligence ESG stakes.” EY. Accessed September 15, 2024. https://assets.ey.com/content/dam/ey-sites/ey-com/en_ca/topics/ai/ey-artificial-intelligence-esg-stakes-discussion-paper.pdf

13. “Without AI, we won’t meet ESG goals and address climate change.” World Economic Forum, 2023. Accessed September 15, 2024. https://www.weforum.org/agenda/2023/01/ai-can-help-meet-esg-goals-and-climate-change/

14. Sundar, Pichai. “UN Summit of the Future: AI Opportunity for Everyone.” Google, September 21, 2024. https://blog.google/inside-google/message-ceo/united-nations-keynote-2024/.

15. Shaolei, Ren, et al. “The Uneven Distribution of AI’s Environmental Impacts.” Harvard Business Review [Climate Change], July, 2024. Accessed September 15, 2024. https://hbr.org/2024/07/the-uneven-distribution-of-ais-environmental-impacts

16. Katherine, Haan. “How Businesses Are Using Artificial Intelligence In 2024.” Forbes. Accessed September 15, 2024. https://www.forbes.com/advisor/business/software/ai-in-business/

17. “Data Suggests Growth in Enterprise Adoption of AI is Due to Widespread Deployment by Early Adopters, But Barriers Keep 40% in the Exploration and Experimentation Phases.” IBM Media Relations. Accessed September 15, 2024. https://newsroom.ibm.com/2024-01-10-Data-Suggests-Growth-in-Enterprise-Adoption-of-AI-is-Due-to-Widespread-Deployment-by-Early-Adopters

18. Lori, Perri. “3 Bold and Actionable Predictions for the Future of GenAI.” Gartner. Accessed September 15, 2024. https://www.gartner.com/en/articles/3-bold-and-actionable-predictions-for-the-future-of-genai

19. Amy, Leurs, et al. “Will AI accelerate or delay the race to net-zero emissions?” Nature, 2024. Accessed September 15, 2024. https://www.nature.com/articles/d41586-024-01137-x

20. “The Return on Investment in AI Ethics: A Holistic Framework.” MAIEI. Accessed September 15, 2024. https://montrealethics.ai/the-return-on-investment-in-ai-ethics-a-holistic-framework/#:~:text=These%20paths%20form%20the%20Holistic,or%20cost%20of%20capital%20reduction.

21. Amane, Dannouni, et al. “Accelerating Climate Action with AI.” Boston Consulting Group, 2023. Accessed September 15, 2024. https://www.gstatic.com/gumdrop/sustainability/accelerating-climate-action-ai.pdf

22. “Measuring the Emissions and Energy Footprint of the ICT Sector: Implications for Climate Action.” ITU. Accessed September 15, 2024. https://www.itu.int/en/ITU-D/Environment/Pages/Publications/Measuring-Emissions-and-Energy-Footprint-ICT-Sector.aspx

23. “The Pact for the Future.” United Nations, 2021. Accessed September 15, 2024. https://www.un.org/sites/un2.un.org/files/sotf-the-pact-for-the-future.pdf.

24. “What is AI”. MIT Technology Review. Accessed September 15, 2024. https://www.technologyreview.com/2024/07/10/1094475/what-is-artificial-intelligence-ai-definitive-guide

About

The Digital Economist, headquartered in Washington, D.C., is a dynamic ecosystem of over 40,000 senior executives and leaders, with a core Fellowship of 100+ experts. With a strong applied AI practice, The Digital Economist is at the forefront of responsible technology adoption, focusing on advancing policies and governance structures that ensure emerging technologies serve humanity. Our collaborations with governments and multi-stakeholder organizations drive impactful solutions that shape how we build, measure, and sustain value. Through scientific research, strategic advisory, and innovation in emerging technology, we are co-building a future where digital advancements enhance global wellbeing.

CONTACT: INFO@THEDIGITALECONOMIST.COM