Journal of Product Innovation Management

WILEY | pdma Connecting Innovators Worldwide

ORIGINAL ARTICLE OPEN ACCESS

Responsible Artificial Intelligence Attention and Firm Innovation: An Attention-Based View

Mengran Xiong[1] | Haofeng Xu[2] | Jiao Ji[3] | Renxian Zuo[4] | Yichuan Wang[3] | Hesam Olya[3,5]

[1]Westminster Business School, University of Westminster, London, UK | [2]Loughborough Business School, Loughborough University, Loughborough, UK | [3]Sheffield University Management School, University of Sheffield, Sheffield, UK | [4]School of Information Management, Wuhan University, Wuhan, Hubei, China | [5]Kyung-Hee University, Seoul, South Korea

Correspondence: Hesam Olya (h.olya@sheffield.ac.uk)

Received: 27 October 2023 | **Revised:** 10 October 2025 | **Accepted:** 22 October 2025

Associate Editor: Constantine Katsikeas

Funding: The authors received no specific funding for this work.

Keywords: attention-based view (ABV) | firm innovation | large language model (LLM) | natural language processing (NLP) | responsible artificial intelligence (AI)

ABSTRACT

Academic Summary: This article draws on the attention-based view (ABV) to examine whether, how, and under what conditions top management team (TMT) attention to responsible artificial intelligence (AI) influences firm innovation. We developed a 480-word responsible AI dictionary grounded in 155 academic sources and 527 corporate case descriptions, and applied it to 2452 S&P 500 earnings call transcripts (2011–2021) using natural language processing (NLP) and large language model (LLM) techniques, yielding 2670 firm-year observations. Linking these measures to US patent data, we find that greater responsible AI attention predicts more and higher-impact patents. The effect is stronger in low-technology industries and under short-term investor pressure, while the presence of a chief technology officer (CTO) does not amplify it. Mechanism analyses reveal that responsible AI attention fosters innovation by increasing investment in AI-relevant human capital and mitigating innovation risk. Theoretically, this article enriches the AI and innovation management literature by positioning responsible AI attention as a dynamic strategic asset that mobilizes resources, reduces risk, and enables contextual adaptation. Practically, findings suggest that firms can strengthen innovation by prioritizing managerial attention to responsible AI, distributing responsibility beyond technical specialists, balancing ethical safeguards with strategic flexibility, and aligning governance with investor and industry conditions.

Managerial Summary: This article examines how managerial attention to responsible artificial intelligence (AI) can enhance firm innovation. Using text analytics on 2452 earnings call transcripts from S&P 500 firms (2011–2021) and a panel of 2670 firm-year observations linked to patent outcomes, we show that firms whose top management teams (TMT) devote greater attention to responsible AI produce more and higher-impact patents. This effect is stronger in low-technology industries and when firms face short-term investor pressure; it is not amplified by having a chief technology officer (CTO). In practice, sustained attention to responsible AI tends to build AI-related skills and reduce project risk, thereby supporting a more reliable innovation pipeline. Executives should treat responsible AI as a strategic priority rather than a compliance task by establishing cross-functional governance, investing in role-based governance training, and sharing accountability across the C-suite. Innovation managers can embed ethics checkpoints (bias audits, design reviews) into project workflows to enhance stability and organizational learning.

Policymakers can reinforce responsible innovation by providing clear regulatory frameworks and incentives that align ethical safeguards with competitiveness. Together, these actions can help build more durable organizational capability for responsible innovation and support long-term performance and adaptation to ongoing technological change.

1 | Introduction

Innovation management is being reshaped by the rapid rise of artificial intelligence (AI). According to the Forrester Global Tech Market Forecast report, software and IT services (including AI) are expected to account for 69% of global technology expenditures by 2027, while investments in AI are predicted to surge to $227 billion by 2030, reflecting an impressive annual growth rate of 36% (Forrester 2024). AI technologies, capable of simulating human cognitive behaviors (learning, reasoning, adapting; Berente et al. 2021), are accelerating the speed, scope, and scale of innovation processes across industries (Huang and Rust 2021, 2024; Verganti et al. 2020). Recent scholarly work echoes these transformative trends. For example, AI adoption both enables new innovation capabilities and enhances existing ones by mining massive datasets for patterns and novel solutions (Gama and Magistretti 2025; Spanjol et al. 2024). In practice, firms are already deploying AI to augment core innovation activities: AI-driven systems boost idea generation (Bouschery et al. 2023) and product design (Verganti et al. 2020). As a result, many organizations not only achieve incremental improvements but also unlock radical new pathways for product, service, and business model innovation. IBM's Watsonx Assistant (Gidwani and Bello 2024) provides a vivid example, By leveraging custom large language models (LLMs) to transform customer support, it reportedly cut service costs by up to 30% while raising customer satisfaction by 20%, illustrating how AI-fueled innovation can yield significant competitive advantages.

However, AI's transformative power comes with profound ethical, legal, and societal risks. High-profile incidents, ranging from facial recognition privacy scandals to algorithmic bias in decision-making (Asatiani et al. 2021; Mikalef et al. 2022), have triggered public outcry and regulatory scrutiny (e.g., Clearview AI's unauthorized use of personal data led to a $10 million fine and regional bans; Heikkilä 2022). Firms that race to adopt AI without appropriate governance risk eroding customer trust and damaging their reputations. In response, the concept of responsible AI has emerged as a critical framework for ensuring that AI systems are developed and deployed in accordance with ethical principles such as transparency, fairness, accountability, and privacy (Barredo Arrieta et al. 2020; Díaz-Rodríguez et al. 2023). Responsible AI involves proactively integrating these principles into AI design and use, essentially building "guardrails" to prevent misuse or unintended harm (Babina et al. 2024; Shneiderman 2021). This notion is rapidly gaining traction. According to a recent Accenture survey (Eitel-Porter and Grosskopf 2022), 80% of companies plan to increase investment in responsible AI and 77% prioritize AI regulation in the near term. Businesses and academic communities also recognize that failing to embed ethics in AI can invite backlash and stifle innovation, whereas doing so may help build stakeholder trust and sustained innovation capacity (Spanjol et al. 2024).

Despite growing interest in both AI-driven innovation and responsible AI, significant knowledge gaps remain at their intersection. Prior studies have demonstrated that AI can act as a powerful technological catalyst for innovation (e.g., by automating R&D tasks or generating creative outputs) and recent literature reviews map out the broad contours of this phenomenon. For instance, Gama and Magistretti (2025) document how AI adoption influences firms' innovation capabilities and outline various roles AI can play in innovation processes. At the same time, thought leaders in the field have highlighted responsible innovation as a top-tier research priority. Spanjol et al. (2024) identify "AI in innovation" and "responsible innovation" as two of the five most critical forward-looking research paths for innovation management. These works make clear that understanding how ethics and innovation intertwine is increasingly important. Yet, to date, we lack robust theoretical and empirical insights into how embracing responsible AI principles actually impacts firm-level innovation outcomes (Babina et al. 2024; Shneiderman 2021).

Most extant research has treated AI from a technology-centric perspective, emphasizing algorithms, capabilities, and performance gains (Babina et al. 2024; Gama and Magistretti 2025; Liu et al. 2020; Rammer et al. 2022), or has examined ethical AI in isolation (e.g., design guidelines or compliance checklists; Díaz-Rodríguez et al. 2023). In contrast, little attention has been given to managerial attention, that is, how the top management team (TMT) perceives and prioritizes ethical AI issues, and how this shapes a firm's innovation. This omission is critical: if managers overlook ethical imperatives amid the AI hype, they may expose their firms to innovation-derailing risks; conversely, if they over-index on caution, they might miss out on high-reward innovative opportunities. Recognizing this gap, our article asks: *How does managerial attention to responsible AI influence firm-level innovation outcomes? Through what mechanisms does this influence occur, and under what conditions is it most pronounced?*

To answer these questions, we draw on the Attention-Based View (ABV) of the firm as our theoretical lens. ABV posits that organizational decision makers operate under bounded rationality and thus can attend to only a limited set of issues at any given time (Ocasio 1997). What TMTs choose to focus on, guided by organizational attention structures, routines, and values, filters which problems and opportunities gain priority, thereby shaping strategic decisions and outcomes. In essence, managerial attention is the scarce resource that determines how firms allocate resources, address challenges, and exploit opportunities. By channeling attention toward certain domains and not others, TMTs set the agenda for organizational action. This perspective is highly pertinent to our inquiry: if responsible AI becomes a salient focus for a firm's TMT, ABV suggests that it will influence how the firm steers its innovation efforts. We expect that heightened managerial attention to responsible AI will shape the firm's innovation strategy, directing more resources and care toward such considerations and potentially altering the types of innovations the firm pursues.

We theorize that managerial attention to responsible AI can be a double-edged sword for innovation. On one hand, when TMTs emphasize responsible AI, they do more than mitigate ethical pitfalls; they create environments conducive to both innovation quantity and quality. First, embedding ethical principles into AI initiatives enhances market legitimacy and stakeholder trust, accelerating the adoption of new products and increasing innovation quantity (Gnewuch et al. 2024). Second, transparency and fairness practices, such as data audits and bias-detection protocols, stimulate process innovation by uncovering insights, thereby catalyzing novel product features or entirely new offerings and enhancing innovation quality and quantity (Raisch and Krakowski 2021). Third, when TMTs consistently emphasize responsible AI, they strategically reallocate resources, such as specialized AI talent and ethical AI infrastructures, to ensure safe and ethical technology use, effectively establishing sustainable pipelines for high-quality long-term innovation (Li, Li, et al. 2021). Thus, responsible AI attention does not merely mitigate potential harms but actively promotes sustained creative exploration and aligns innovation efforts with stakeholder expectations.

On the other hand, we acknowledge that the net impact of responsible AI attention on innovation may not be unilaterally positive. High ethical scrutiny might, in some cases, slow down experimentation or constrain the data and methods available for innovation (e.g., strict privacy rules could limit creative data use, or thorough review processes might delay product launches). Some skeptics worry that too strong a focus on "AI principles" could act as a brake on the free-wheeling exploration that radical innovation sometimes requires (Bouschery et al. 2023; He and Tian 2013; Wu et al. 2015). Conversely, it is also plausible that a reputation for responsible AI could expedite innovation diffusion; for instance, customers and regulators may more quickly embrace new AI-driven solutions that come with assurances of fairness and transparency. To navigate this theoretical tension, we design our article to empirically investigate the conditions under which managerial attention to responsible AI catalyzes innovation versus when it might inadvertently constrain it. Our aim is to unpack the specific mechanisms and boundary conditions that determine how responsible AI attention plays out in practice as an innovation driver.

Addressing these theoretical tensions requires rigorous empirical investigation. We develop a novel measure of "managerial attention to responsible AI" by analyzing the content of 2452 earnings call transcripts from S&P 500 firms (2011–2021). Using natural language processing (NLP), LLM, and our custom dictionary of 480 keywords related to responsible AI and its practices, we capture how frequently and in what context TMTs discuss responsible AI topics in their formal communications. We then link these attention measures to firm-level innovation outcomes using patent data (2670 firm-year observations of granted patents, forward citations, and patent-derived market value) from the US Patent and Trademark Office (USPTO; Kogan et al. 2017; Stoffman et al. 2022). Patents and their citations are well-established indicators of a firm's innovative output and impact (He and Tian 2013; Nanda and Rhodes-Kropf 2013).

Our article makes three theoretical contributions. First, it extends the ABV by demonstrating that not only economic or technological cues but also normative and ethical concerns can guide managerial attention and shape strategic outcomes. Integrating insights from upper echelons theory (UET), we link executives' ethical orientations to the allocation of attention and to firms' innovative behavior, thus specifying how values enter the attention–performance relationship. Second, we enrich ABV by explicating the mechanisms through which responsible AI attention promotes innovation. Firms with greater responsible AI attention invest more in AI-relevant human capital and experience lower innovation risk, embedding a socio-ethical dimension into the core of managerial attention. Third, we refine ABV by identifying boundary conditions and attentional heterogeneity: responsible AI attention yields stronger innovation benefits under low technological intensity and short-term investor pressure, while distinct ethical foci exert divergent effects. Methodologically, we offer a scalable, replicable approach to capturing responsible AI attention in executive discourse. Taken together, these insights support the view that responsible AI can be regarded as a sustained managerial priority, rather than merely a compliance obligation, aligning ethical integrity with firms' innovative performance.

2 | Institutional Background

As AI reshapes core business practices, accelerates innovation, and creates new sources of value, concerns about ethical lapses have heightened the imperative for responsible AI. This shift has expanded the conversation beyond technological issues to encompass managerial capabilities and ethical governance. Increasingly, firms approach responsible AI as a strategic priority rather than a compliance obligation, engaging stakeholders to align AI initiatives with diverse values and thereby enhancing innovation quality, market responsiveness, and long-term growth. To better understand how firms operationalize responsible AI in practice, we conducted a qualitative analysis that combines insights from prior research with 527 case descriptions from 30 global organizations. These cases span healthcare (e.g., Alder Hey Children's Hospital), finance (e.g., Capital One, JPMorgan Chase), retail (e.g., Amazon, H&M), technology (e.g., IBM, NVIDIA), consulting (e.g., PwC, Deloitte), energy (e.g., Equinor ASA), automotive (e.g., Audi), and telecommunications (e.g., Cisco, Telefónica). Across these diverse contexts, a consistent pattern emerges: firms that embed responsible AI principles into core operations tend to stimulate innovation, differentiate their offerings, and cultivate stakeholder trust.

Three recurring practices are especially noteworthy. First, embedding fairness, explainability, and accountability into both strategy and system design enables firms to anticipate and mitigate ethical risks, thereby facilitating AI adoption and improving market responsiveness. Second, cross-functional governance structures that bring together data scientists, ethicists, AI managers, and legal experts are essential for coordinating ethics-oriented initiatives. This integrative approach reduces organizational friction and fosters a culture of responsible innovation. Third, human-centered design reframes AI as a complement to human creativity and judgment rather than a replacement, which minimizes ethical pitfalls while strengthening user acceptance and loyalty.

Firms also tend to anchor their responsible AI initiatives in five interrelated dimensions: beneficence, non-maleficence, autonomy, justice,[1] and explicability (Floridi et al. 2018). These dimensions serve as a blueprint for responsible deployment by emphasizing societal benefit, harm avoidance, user agency, equitable outcomes, and transparent decision-making. Illustrative examples include H&M's integration of fairness and transparency into its recommendation systems to address consumer concerns, Deloitte's adoption of governance protocols to manage algorithmic bias and privacy risks, and IBM's emphasis on explicability to differentiate its AI solutions and reinforce client trust (see Bigham et al. 2022; H&M Group 2021; IBM 2025).

Taken together, these cases suggest that responsible AI, when grounded in clear principles, supported by cross-functional teams, and implemented through human-centered design, constitutes a cornerstone of sustainable innovation. Treating responsible AI as a strategic driver, rather than only an ethical safeguard, enables firms to build trust, enhance reputation, and secure competitive advantage (Bouschery et al. 2023; Floridi et al. 2018). As intelligent technologies become increasingly embedded in business and society, integrating responsible AI into long-term strategy has become essential for sustaining leadership in dynamic markets (Huang and Rust 2021; Spanjol et al. 2024). Table 1 illustrates how the five dimensions of responsible AI are conceptually articulated and demonstrates their practical significance through examples from prior research and business practice. See Web Appendix A for details on responsible AI and its sub-dimensions.

These insights, drawn from both case evidence and prior research, underscore the importance of institutionalizing responsible AI and examining its organizational consequences, particularly for innovation. However, the underlying mechanisms, boundary conditions, and measurable impacts remain underexplored. The following section therefore develops hypotheses based on the ABV, explaining how TMTs' strategic focus on responsible AI may stimulate innovation through its influence on resource allocation, risk management, and stakeholder legitimacy.

3 | Hypothesis Development

The ABV of the firm provides the overarching theory for our article (Ocasio 1997). It draws on early works from the behavioral theory of the firm (e.g., Cyert and March 1963; Simon 1947) to emphasize that individuals face bounded rationality and therefore must selectively allocate their limited attention. Attention, as a scarce organizational resource, determines which issues and action alternatives decision makers consciously or unconsciously prioritize while neglecting others, such as "problems, opportunities, and threats" and "proposals, routines, projects, programs, and procedures" (Ocasio 1997, 189). Because attention allocation determines which cues are perceived as salient and which responses are enacted, variations in the focus, intensity, and distribution of attention explain differences in organizational behavior and performance (Ocasio 1997; Repenning and Sterman 2002). Empirical research supports this argument, showing that executive attention shapes firms' innovation behavior (Eggers and Kaplan 2009; Ridge et al. 2017), responses to environmental change (Nadkarni and Barr 2008), and sustainability performance (Ahn 2020).

Building on this logic, we argue that TMT attention to responsible AI shapes firm innovation outcomes. The ABV posits that strategic decisions and outcomes depend heavily on where TMTs direct attention and how they interpret information. The UET (Hambrick and Mason 1984) reinforces this view by showing that attentional choices are influenced by TMTs' values, expertise, and cognitive orientations. In the context of AI, these perspectives together suggest that TMTs' ethical priorities and technological focus jointly determine how AI is developed, governed, and leveraged for innovation. On this basis, we develop hypotheses about how TMT attention to responsible AI stimulates innovation, through which mechanisms these effects occur, and under what contextual conditions they are likely to vary. Figure 1 depicts our research road map, showing hypothesized relationships, the dataset used, and the unit of analysis.

3.1 | Responsible AI Attention and Innovation

Directing greater managerial attention to responsible AI is expected to enhance a firm's overall innovation output. When a TMT highlights responsible AI practices, it signals strategic importance, mobilizes resources, and aligns organizational efforts toward AI-driven innovation that adheres to ethical standards (Yadav et al. 2007). According to the ABV, issues that receive greater managerial focus attract disproportionate organizational resources and action; thus, prioritizing responsible AI channels investment into cutting-edge, ethically sound AI capabilities and processes.[2]

Prior research supports this logic: managerial attention to emerging technologies is positively associated with innovation performance. For example, firms whose TMTs emphasize technological initiatives generate more patents and new product launches, and family firms that prioritize technological adaptation exhibit superior innovation outcomes (Kammerlander and Ganter 2015). Importantly, mere possession of technical resources, however, is not by itself transformative; their innovation potential is unlocked only when channeled by focused managerial attention and guidance. Extending this logic to the AI domain, we posit that TMT attention to responsible AI should spur greater innovation. Embedding ethical principles into AI development enables firms to generate novel offerings such as transparency- or fairness-enhancing features, and to improve existing solutions through, for example, bias-mitigation algorithms. Such efforts will ultimately increase both the quantity of innovation (e.g., patent counts) and its quality and market value. Accordingly, we expect:

H1. *Firms whose TMT devotes greater attention to responsible AI will exhibit higher innovation outcomes.*

3.2 | Mechanism: Investing in AI-Skilled Human Capital

One key mechanism through which TMT attention to responsible AI can drive innovation is human capital investment. Managerial attention not only signals strategic priorities but

TABLE 1 | Selective illustrative comparison of academic vs. industry perspectives on five core responsible AI dimensions.

Responsible AI sub-dimensions	Definitions	Academic perspective	Industry perspective
Beneficence	An AI system is designed, validated, and deployed to produce measurable public benefit, sharpening human capabilities, improving decision quality, and enhancing resource efficiency, while at the same time weighing profit against fairness, sustainability, and risk.	Research highlights that beneficent AI must be purpose-built to deliver measurable human and societal gains, such as boosting capabilities, decision quality, and well-being, while remaining aligned with sustainability and fairness. For example, Tomašev et al. (2020) catalog AI-for-Social-Good projects aligned with the UN SDGs, distill 10 cross-sector partnership guidelines, and spotlight impactful cases (e.g., Troll Patrol), thereby offering a *beneficence*-oriented roadmap for innovation managers to steer AI portfolios toward scalable societal value.	Accenture Responsible AI program (Eitel-Porter et al. 2021; Eitel-Porter and Grosskopf 2022), Alder Hey Children's Hospital's predictive-care models (Alder Hey 2025), Microsoft's AI for Health (Microsoft 2025a), and Google Research's AI for Social Good initiative (Google 2025b) illustrate how responsible AI can create both business and societal value with measurable ESG impact.
Non-maleficence	An AI system is built and overseen to prevent or minimize harm through rigorous risk and bias audits, strong privacy and security safeguards, and clear accountability, so it does not exploit or discriminate against people, society, or the environment.	Research stresses rigorous risk audits, bias testing, and explainable AI (XAI) to mitigate privacy breaches or manipulation. For example, Thomaz et al. (2020) analyze AI conversational agents for a hyper-private web, map "Buff" versus "Ghost" user data scenarios, flag privacy- and bias-related harms, and propose privacy-by-design safeguards plus stage-gate risk audits, offering a non-maleficent blueprint for innovation management.	Capital One and JPMorgan Chase embed model-risk governance into AI development workflows (Capital One 2023; JPMorgan Chase 2023); Cisco and IBM apply red-team testing to strengthen AI security and governance (Owen-Jackson 2025; Patel 2025); Google and Microsoft conduct pre-launch risk reviews for AI applications (Google 2025c; Microsoft 2023, 2025b), together shaping a harm-avoidance blueprint.
Autonomy	An AI system offers user-centered controls that let individuals make informed, voluntary choices about engagement, data sharing, and the delegation or withdrawal of decision authority, free from coercion or manipulation and thus safeguarding human agency.	Research highlights the importance of AI designs that preserve human agency in automation–augmentation hybrids and meaningful work. For example, Huang and Rust (2024) focus on human-centered "feeling AI" for customer service, map a four-step journey: detect, empathize, manage, bond, flag generative AI gaps, and provide prompt/response-engineering tenets as a staged blueprint for service-innovation R&D, while upholding *autonomy* through opt-in emotion sharing and user overrides.	Salesforce's Einstein Trust Layer (Salesforce 2025), Meta's Privacy Center (Meta 2022), and Telefónica's Data Transparency Portal (Telefónica 2025) provide users with opt-outs, granular consent options, and transparency mechanisms that preserve human decision authority.
Justice	An AI system is governed to ensure distributive and procedural fairness, using inclusive data, quantitative bias checks, participatory oversight, and remedy pathways to ensure that benefits and burdens are shared equitably across groups.	Research suggests that AI designs employing inclusive data and fairness audits can help secure unbiased and equitable outcomes. For example, Rhee (2024) interrogates *justice* in emotion-recognition AI, benchmarks three commercial models against human raters, uncovers demographic scoring gaps, finds that transparency leaves users anchored to biased outputs, sometimes worsening inconsistencies, and calls for systemic fairness safeguards beyond individual correction.	Bias-mitigation is supported by IBM AI Fairness 360 (Bellamy et al. 2019), Credo AI's governance platform (Eisenberg 2025), and SAP's AI Ethics program (Lombana Díaz and Welsch 2025), and further strengthened by fairness provisions in Google's and Microsoft's Responsible AI standards (Google 2025a; Microsoft 2025b).

(Continues)

TABLE 1 | (Continued)

Responsible AI sub-dimensions	Definitions	Academic perspective	Industry perspective
Explicability	An AI system provides clear, human-interpretable explanations of its data sources, reasoning, and chain of accountability, along with stated limitations and uncertainties, enabling stakeholders to scrutinize and act on its outputs without compromising privacy or security.	Research shows that explainability and transparency are the trust linchpins, turning AI from a "black box" into a "glass box" via XAI taxonomies and user-focused explanations. For example, Bauer et al. (2023) test feature-based AI explanations in investment and real estate tasks, show they reweight user cues and trigger confirmation-bias spillovers, positioning *explicability* as a double-edged tool that innovation managers must channel, via bias-aware stage gates, to boost learning without seeding new errors.	Transparency comes via model cards and explainability suites such as IBM Explainability 360 (IBM 2019), NVIDIA NeMo Model cards (NVIDIA 2025), PwC's Responsible AI Toolkit (PwC 2025), and Microsoft's annual transparency reports (Hutson and Crampton 2025).

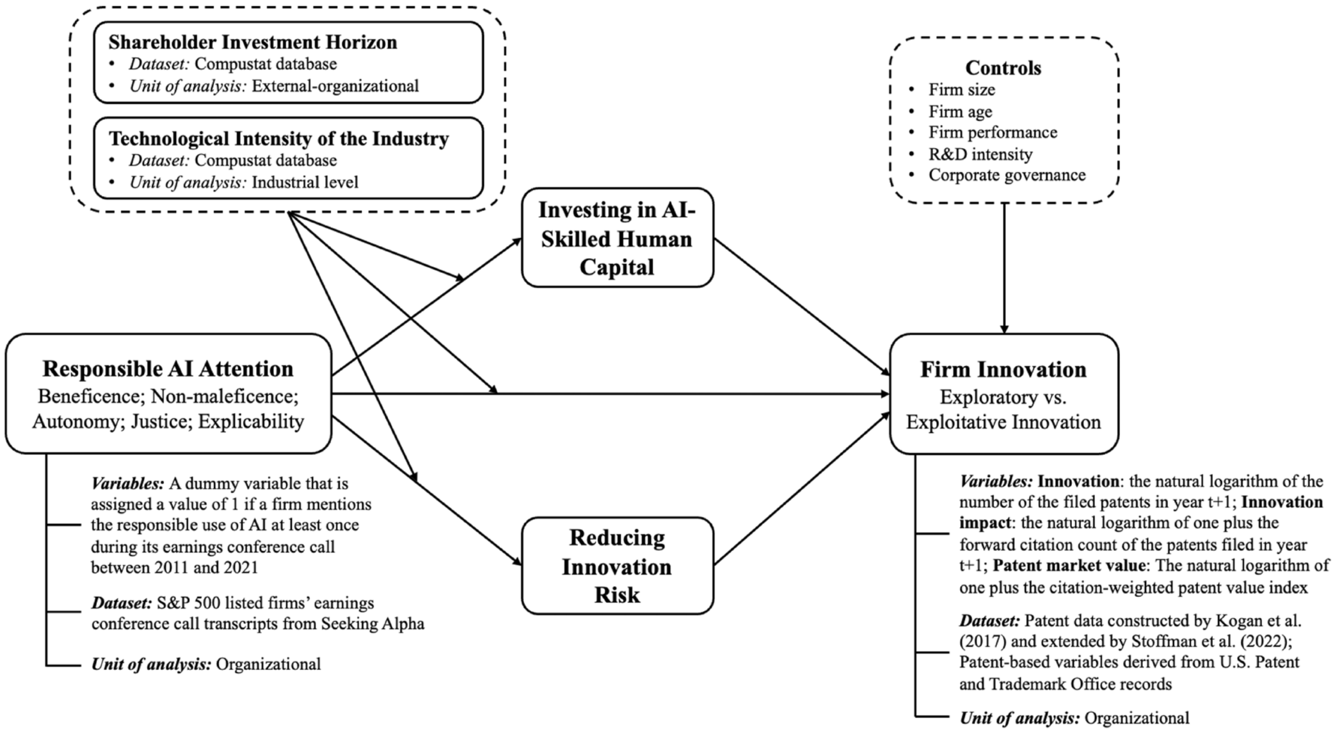

FIGURE 1 | Research road map.

also triggers resource reallocation within the firm. When top executives focus on responsible AI, they are likely to direct budgets toward acquiring and developing the necessary talent to implement ethical AI practices, such as hiring AI specialists (e.g., data scientists, AI ethicists), training employees in ethical AI protocols, or forming dedicated AI ethics committees (Hambrick and Mason 1984; Li, Li, et al. 2021). These investments strengthen the firm's capacity for ethical and cutting-edge AI innovation. Consistent with UET, managerial values shape how attention translates into concrete resource deployment; thus, a TMT committed to responsible AI will allocate staffing and training resources to enact those values. By expanding the pool of AI-skilled employees, the firm enhances its capacity to generate innovative AI solutions that align with high ethical standards.

Empirical research supports the link between managerial attention and resource allocation. For example, when CEOs emphasize a particular strategic agenda, they tend to channel resources and investments accordingly; Bouquet and Birkinshaw (2008) describe attention as a mechanism that drives resource provision to favored projects. Having a larger and better-trained AI workforce directly improves innovation performance, as skilled AI engineers and scientists are the ones who create new algorithms and patents. Prior studies confirm that firms with greater technological human capital exhibit higher innovation, as expertise is a critical input to R&D (Artz et al. 2010). Therefore, we expect that responsible AI attention increases AI-skilled human capital, thereby facilitating greater innovation (Babina et al. 2024). We formally posit this mechanism as follows:

H2. *TMT attention to responsible AI increases the firm's investment in AI-skilled human capital, which in turn enhances firm innovation outcomes.*

3.3 | Mechanism: Reducing Innovation Risk

A second mechanism linking responsible AI attention to innovation involves reducing innovation risk, which in turn promotes more stable R&D outcomes. Incorporating responsible practices in AI development reduces the likelihood of costly failures, ethical crises, or public backlash, all of which can disrupt innovation (Bankins and Formosa 2023; Bauer et al. 2023). By paying close attention to AI's ethical and social implications, TMTs are likely to implement checks and balances (e.g., bias audits, compliance reviews, and risk assessments) that help prevent severe missteps (Gama and Magistretti 2025). While proactive risk management may not directly expand the firm's creative frontier, it reduces downside exposure, ensuring a steady flow of safe and legitimate innovations over time. In contrast, neglecting AI ethics can lead to setbacks such as product withdrawals or legal sanctions that derail innovation projects (e.g., the case of Clearview AI). A TMT attentive to responsible AI is therefore more likely to avoid such pitfalls by embedding compliance measures early in the innovation cycle. Empirical evidence supports the notion that minimizing downside risk enhances innovation stability. He and Tian (2013) find that short-term market pressures (e.g., intense analyst scrutiny) often push managers to cancel long-term innovative projects, highlighting the need for stable, supportive conditions to sustain continuous innovation. A responsibility-oriented approach helps align innovation with stakeholder expectations and shields R&D from disruptions, such as user boycotts or late-stage product failures.

Ultimately, responsible AI attention acts as an internal risk control system for innovation. Just as robust R&D controls reduce variance in innovation outcomes (Lu et al. 2023; Miller et al. 2022), ethical oversight lowers the volatility of innovation performance. As a result, we expect that firms exhibiting higher levels of responsible AI attention will experience less fluctuation in R&D intensity, thereby complementing the direct effect on overall innovation quantity. Formally stated:

H3. *TMT attention to responsible AI reduces innovation risk, which in turn enhances firm innovation outcomes.*

3.4 | Moderating Role of Shareholder Investment Horizon

A firm's shareholder investment horizon, whether dominated by either short- or long-term investors (e.g., institutional investors), can moderate how responsible AI attention influences innovation (Gaspar et al. 2005, 2013). Managerial attention priorities are notably influenced by shareholder pressures. Developing and integrating responsible AI practices is forward-looking and may entail substantial upfront costs, including investments in talent development, new processes, and potentially delayed product rollouts due to compliance demands (Glikson and Woolley 2020). Long-term-oriented investors typically support such strategic investments, viewing ethical compliance, sustained learning, and risk management as essential to long-run value and corporate reputation (Kim et al. 2019; Lerner et al. 2011). Because long-horizon investors already tolerate delayed payoffs and tend to promote responsible mandates, their baseline support for innovation is high; thus, additional emphasis on responsible AI yields smaller incremental gains relative to firms without such responsible orientation.

In contrast, short-horizon investors often dominate the Q&A during earnings calls, pressing managers for near-term, low-risk payoffs and scrutinizing compliance and risk (Bushee 1998). When the TMT emphasizes responsible AI, it redirects managerial attention from defensive risk discussions toward opportunity framing, giving exploratory projects a clearer path to approval (Giannetti and Yu 2021). Responsible AI practices such as ethical guidelines, bias audits, and transparent reporting buffer reputational, legal and regulatory risks, making AI initiatives appear more credible and less risky even under compressed timelines. At the same time, visible governance structures signal disciplined use of R&D funds, alleviating myopic investors' concerns about wasteful spending. Together, these dynamics explain why the innovation benefits of responsible AI attention are amplified in firms with a predominantly short-term investor base. We thus hypothesize:

H4. *The positive effect of TMT attention to responsible AI on firm innovation is stronger when the investor base is short-term oriented and weaker when long-horizon investors dominate.*

3.5 | Moderating Role of Technological Intensity of the Industry

Industry technological intensity, defined as the level of technological sophistication and R&D activity within an industry (Felsenstein and Bar-El 1989; Zawislak et al. 2018), can moderate how TMT attention to responsible AI affects innovation. In high-tech industries (e.g., software, electronics), firms operate in dynamic, information-rich environments characterized by rapid innovation cycles and multiple competing technological priorities. In such contexts, managerial attention is often widely dispersed across numerous innovation agendas, which may constrain the extent to which leaders can devote focused attention to responsible AI. Moreover, high-tech firms typically possess advanced R&D governance systems that already embed ethical and regulatory considerations (such as data privacy and regulatory compliance) directly into their ongoing innovation activities (Bosch-Sijtsema and Bosch 2015). Consequently, additional attention to responsible AI may yield only incremental innovation benefits, as responsible practices are already institutionalized and less likely to generate distinct new opportunities.

Conversely, firms in low-tech industries (e.g., basic manufacturing, consumer goods) typically innovate incrementally, relying primarily on well-established processes and adopting external technologies rather than pursuing radical internal breakthroughs (Santamaría et al. 2009). For these firms, emphasizing responsible AI represents a novel strategic orientation that can meaningfully redirect managerial attention and resources. Because their attention is less fragmented by competing technological agendas, low-tech firms can more readily integrate responsible AI practices into products and processes. Such initiatives can differentiate them from peers and open new market opportunities by appealing to stakeholders sensitive to ethical and social responsibility (Lepak et al. 2003; Mendonça 2009). Therefore, we expect the positive impact of TMT attention to responsible AI on firm innovation to be stronger in low-tech industries than in high-tech industries, as the distinctive strategic and market value generated by responsible AI attention is more pronounced when firms operate in less technologically intensive environments. Accordingly, we propose:

H5. *Industry technological intensity moderates the relationship between responsible AI attention and firm innovation, such that the positive effect of TMT attention to responsible AI on innovation is stronger in low-tech industries and relatively weaker in high-tech industries.*

4 | Method

4.1 | Data Source and Sample Construction

The sample in this article consists of S&P 500 firms from 2011 to 2021, which are suitable for innovation studies due to their substantial financial resources, advanced R&D capabilities, and broad access to large, diverse markets. These firms can significantly invest in new technologies and effectively scale innovations. Their transparent financial disclosures also provide valuable data for identifying key variables for economic models.

We leverage extensive textual data and advanced text-analytic techniques to examine whether TMT communications reflect a firm's strategic managerial attention to responsible AI (Martin and Kushwaha 2024). Earnings conference calls serve as our primary data source for three reasons: (1) their bidirectional format (TMT presentations followed by analyst Q&As), which reduces information asymmetry (Brown et al. 2004) and provides deeper insights into strategic priorities and responsible AI commitment (Fu et al. 2021; Jung et al. 2018); (2) their timely, value-relevant information (e.g., quarterly earnings calls), which elicits market reactions (Martin and Kushwaha 2024); and (3) their more concrete and current information compared with scripted disclosures such as annual reports, which makes them well suited to capture discourse about responsible AI practices (Burgoon et al. 2016; Dyer et al. 2017).

We combined data from four sources: (1) *Seeking Alpha*, a financial information website providing transcripts of firms' quarterly earnings conference calls; (2) *Compustat*, a comprehensive financial database by S&P Global Market Intelligence containing firms' financial reporting; (3) *BoardEx*, a global business intelligence platform offering detailed corporate governance and board information; (4) Innovation metrics are obtained from Kogan et al. (2017) and Stoffman et al. (2022). We initially collected raw transcripts of firms' quarterly earnings conference calls from Seeking Alpha covering the period from 2011 through 2021. Using literature and case study insights, we developed a responsible AI dictionary and applied NLP and LLM techniques to extract responsible AI mentions from those transcripts. This yielded 3586 data points, including 2452 from S&P 500 firms. Our sample begins in 2011 because earnings call transcripts were less available prior to 2011; these transcripts are crucial sources for capturing firms' strategic priorities and performance insights (Palmon et al. 2024). For the same period, we acquired firm financial data from Compustat and patent data from Kogan et al. (2017) and Stoffman et al. (2022), covering all USPTO-awarded patents. We excluded financial firms (SIC codes 6000–6799) and firms lacking sufficient data or patent activity, as their innovation behaviors fundamentally differ from patent-active firms (Miller et al. 2022). Non-patenting firms likely pursue distinct business strategies or lack R&D investments critical for innovation, making comparisons inappropriate. Our final sample consists of 2670 firm-year observations. Detailed methods are reported in Web Appendices B and C.

4.2 | Measure

4.2.1 | Dependent Variable—Innovation

We measure *firm innovation* as the natural logarithm of the total number of patent applications filed in year $t+1$ that are eventually granted, capturing the volume of a firm's innovative output (Artz et al. 2010; He and Tian 2013). Unlike R&D expenditure-based measures, which are input-focused and prone to reporting inconsistencies (Aghion et al. 2013; He and Tian 2013), the patent-based measure captures tangible innovation outcomes, precisely capturing how responsible AI attention translates into realized innovations (Acharya and Subramanian 2009). We focus on the application year rather than the grant year because innovation activities typically span multiple years, and the application date more accurately reflects the actual timing of innovation activities (He and Tian 2013). For large firms, such as those listed in the S&P 500 index, patents serve strategic roles by sustaining market advantages through preempting competitors, generating licensing revenues, and maintaining innovative capabilities (Gu et al. 2024; Mann and Sager 2007).

In addition to patent counts, we use two complementary metrics to capture the quality and market value of innovation (Hagedoorn and Cloodt 2003). *Innovation impact* (patent citation counts) is measured as the natural logarithm of one plus the forward citation count of the patents filed in year $t+1$. *Patent market value* (patent-value measure) is measured as the natural logarithm of one plus the citation-weighted patent value index (Kogan et al. 2017), quantifying the estimated dollar contribution of a firm's patented knowledge. We further distinguish between exploratory innovation and exploitative innovation following Benner and Tushman (2003) and Gu et al. (2024). *Exploratory innovation* is the natural logarithm of one plus the count of patents whose three-digit USPTO class did not overlap with any class in the firm's prior five-year portfolio, capturing radical, outside-the-core breakthroughs. *Exploitative innovation* is the natural logarithm of one plus the count of patents whose class did overlap with the firm's prior portfolio, capturing incremental, inside-the-core advances. All patent-based variables are derived from USPTO records and are lagged to year $t+1$ to align with our independent variables measured at year t.

4.2.2 | Independent Variable—Responsible AI Attention

We measure *responsible AI attention* as a dummy variable equal to 1 if a firm explicitly mentions the responsible use of AI at least once during its earnings conference calls between 2011 and 2021, and 0 otherwise. We further introduce two constructs to capture the depth and variation of firms' responsible AI discourse. *Responsible AI vigilance* captures the consistency (or volatility) of a firm's responsible AI attention over time, calculated as the rolling standard deviation of responsible AI intensity over the past 2, 3, or 5 years. *Responsible AI intensity* counts the frequency of responsible AI keywords mentioned by TMTs during earnings calls in year *t*, reflecting the managerial emphasis placed on responsible AI. Together, these measures reflect the magnitude of a firm's responsible AI focus and the steadiness with which these commitments are upheld over time. Given the lack of a standardized tool to systematically capture responsible AI discourse, we developed a tailored dictionary using content analysis of 527 responsible AI project descriptions across 30 companies and 155 seminal academic articles. This dictionary includes 480 keywords across five sub-dimensions and underpins our text mining approach.[3] We then applied both dictionary-based NLP (BoW with TF extraction) and LLM techniques to earnings call transcripts.[4]

4.2.3 | Mediators

We examine two mediators to capture how responsible AI practices translate into firm-level outcomes. *AI investment* is measured as the industry-adjusted coefficient of variation in the share of employees whose occupations match AI-related skill keywords, based on Babina et al. (2024). We further categorized this measure by three or five distinct AI skill sets and examined the deviation in the firm's overall AI-skilled workforce share ("All AI") as identified by Babina et al. (2024). *Innovation risk* reflects the volatility of a firm's R&D efforts, operationalized as the industry-adjusted coefficient of variation in the R&D intensity (R&D expenditure-to-sales ratio) over 2-, 3-, and 5-year windows. R&D expenditure data are obtained from Compustat.

4.2.4 | Moderators

We include two moderating variables: technological intensity and shareholder investment horizon. *Technological intensity* captures the extent of reliance on advanced technologies in firm or industry operations and innovations, measured by a dummy variable (*High tech*) coded 1 for firms that belong to high-tech industries and 0 otherwise. *Shareholder investment horizon* reflects the extent to which shareholders prioritize long-term returns when investing in a firm's stocks. This was measured by the reverse-coded ratio of stock shares traded to total outstanding shares in a given year. A dummy variable (*Shareholder investment horizon_Long*) equals 1 if the firm's shareholder investment horizon exceeds the median value in the same industry, and 0 otherwise. We collected all the information from Compustat and BoardEx.

4.2.5 | Control Variables

Consistent with the extant literature (Gu et al. 2024), we included a rich set of firm and industry controls, including firm size, firm age, R&D intensity, firm performance, and corporate governance, to isolate the effects of responsible AI attention on firm innovation activities. *Firm size* was measured as the natural logarithm of total assets in year *t*–1, as larger firms, particularly those listed in the S&P 500, typically possess greater resources for innovation activities. *Firm age*, determined by the number of years since listing as of year *t*–1, captures the extent of a firm's accumulated innovative experience and assets, potentially buffering future innovation efforts. *R&D intensity*, indicating a firm's resource allocation toward innovation, was measured as the ratio of R&D investment to total sales in year *t*–1. *Firm performance*, reflecting available resources for innovation, was measured by return on assets (ROA) in year *t*–1. *Corporate governance* was assessed by the proportion of independent directors on the board and board size, reflecting governance strength in aligning managerial and shareholder interests toward long-term innovation. Year and industry fixed effects were included in all models to account for time-invariant and industry-specific factors. We lagged independent, moderating, and control variables by 1 year to allow their effects to manifest over time. Table 2 summarizes all variables, their definitions, and data sources.

4.3 | Summary Statistics

Table 3 reports summary statistics and pairwise correlations for key variables. Panel A shows the average (log-transformed) *Patent count* (Innovation) is 2.82 (SD 2.25), corresponding to approximately 15 patents per firm-year on average, indicating a highly skewed distribution of innovation outputs (75th percentile = 4.77 or roughly 118 patents, 25th percentile = 0.69 or about 1 patent). *Innovation impact* (forward citations) has a mean of 2.84 with substantial variation (SD = 2.78); notably, many firm-years have few or no citations (25th percentile = 0), while top-performing firms exhibit high citation counts (75th percentile = 5.18). *Patent market value* averages 5.24 (SD = 3.31), capturing economic patent value in dollar terms, with top-quartile firms (75th percentile = 7.85) showing significant patent-related value. Approximately 36% of firm-years demonstrate responsible AI attention, indicating that over one-third of observations involve TMTs explicitly discussing responsible AI during earnings calls. Among control variables, firms are generally large (mean log assets *Firm size* ≈ 9.90), profitable (mean *Firm performance* ROA ≈ 0.17), moderately engaged in R&D (mean *R&D intensity* = 0.05; skewed, with 25th percentile = 0, 75th percentile = 0.08), and have an average listing age of 38 years (mean log *Firm age* = 3.64). Board characteristics vary across firms, with averages around 11 directors (*Board size*) and high *Board independence* (mean ≈ 0.98), indicating most boards in our sample are fully or predominantly independent.

Panel B reports Pearson correlation coefficients among key variables. As expected, the three innovation outcome measures are highly intercorrelated with each other: *Patent count* (Innovation) correlates highly with forward citations ($r = 0.878$) and *Patent market value* ($r = 0.904$), and *Innovation impact* and

TABLE 2 | Summary of variables and definitions.

Variables	Definition	Source
Dependent variable		
Innovation	Natural logarithm of the number of the filed patents in year $t+1$	Kogan et al. (2017)
Innovation impact	Natural logarithm of (1 + forward citations) to the firm's patents filed in year $t+1$, capturing the impact or quality of innovation	Kogan et al. (2017)
Patent market value	Natural logarithm of (1 + citation-weighted patent value) for patents filed by the firm, using the patent valuation index from Kogan et al. (2017) to estimate the dollar value of innovations	Kogan et al. (2017)
Exploratory	Natural logarithm of (1 + count of patents in year t whose 3-digit USPTO class did not overlap with the firm's prior 5-year classes), capturing radical, outside-the-core innovation, followed by Benner and Tushman (2003) and Gu et al. (2024)	USPTO
Exploitative	Natural logarithm of (1 + count of patents in year t whose 3-digit USPTO class did overlap with the firm's prior 5-year classes), capturing incremental, inside-the-core innovation, followed by Benner and Tushman (2003) and Gu et al. (2024)	USPTO
Independent variable		
Responsible AI attention	Dummy variable with the value of 1 if the firm has mentioned responsible AI in earnings conference calls	Seeking Alpha
Responsible AI vigilance	Rolling standard deviation of the firm's responsible AI intensity over the past 2, 3, or 5 years, respectively	Seeking Alpha
Responsible AI intensity	The frequency of responsible AI mentions by TMT during the earnings call in year t, measured as the count of responsible AI keywords in the transcript	Seeking Alpha
General AI term mentioned	Binary indicator equal to 1 if the firm's TMT mentions any AI terms (see Web Appendix B for the word list, not necessarily in an ethical context) during the earnings call in year t, and 0 otherwise	Seeking Alpha
General AI term frequency	The count of general AI-related terms mentioned in the earnings call by TMT in year t	Seeking Alpha
Mechanism variables		
AI investment[a]	The industry-adjusted coefficient of variation of the fraction of employees with AI-related skills: all AI share/categorized share/3-skill share/5-skill share/narrow AI share (Babina et al. 2024).	Babina et al. (2024)
Innovation risk	The industry-adjusted coefficient of variation of the firm's *R&D intensity* (R&D expenditure-to-sales ratio) calculated over a 2-year, 3-year and 5-year window	Compustat
Moderator variables		
High tech	Dummy variable with the value of 1 if the firm belongs to high-tech industry, and 0 otherwise	Compustat
Shareholder investment horizon	The extent to which shareholders are targeting long-term returns when investing in a firm's stocks; The ratio of stock shares traded to total outstanding shares in a given year, reverse-coded	Compustat
Innovation culture	An index of the firm's innovation-oriented culture, reflecting the extent to which the firm's internal culture supports innovation (constructed following Li, Mai, et al. 2021). Higher values indicate a stronger pro-innovation culture within the firm	Li, Mai, et al. 2021

(Continues)

TABLE 2 | (Continued)

Variables	Definition	Source
Control variables		
Firm size	Log (total assets)	Compustat
Firm age	Log (number of years since the firm was listed in the US)	Compustat
R&D intensity	Ratio of R&D investment to total sales	Compustat
Firm performance	Return on assets	Compustat
Board size	Log (number of board of directors)	BoardEx
Board independence	Ratio of independent directors on board	BoardEx
Industry dummy	Based on Fama and French 12 industry classification	
Year dummies	Year 2011 to 2021	

[a]The AI-skilled workforce data of Babina et al. (2024) are constructed from a large resume dataset provided by Cognism (about 535 million individual employment profiles). An individual is tagged as AI-related in a given firm-year if AI keywords (e.g., "machine learning," "natural language processing," "deep learning") appear in the job title or job description. The firm-level measure is calculated as the percentage of a firm's total employees classified as AI-related each year. This measure is implemented in five variants: All AI share (any AI-tagged worker), Categorized share (after mapping skills into exclusive families), 3-skill share (limited to machine learning, natural language processing, and computer vision), 5-skill share (using five core AI skill clusters), and Narrow AI share (restricted to tightly AI-specific skills).

TABLE 3 | Descriptive statistics and correlations.

Panel A: Descriptive statistics

Variables	Obs.	Mean	SD	P25	P75
Innovation	2670	2.82	2.25	0.69	4.77
Innovation impact	2670	2.84	2.78	0.00	5.18
Patent market value	2670	5.24	3.31	3.06	7.85
Responsible AI attention	2670	0.36	0.48	0.00	1.00
Firm size	2670	9.90	1.08	9.01	10.72
Firm performance	2670	0.17	0.09	0.10	0.23
R&D intensity	2670	0.05	0.07	0.00	0.08
Firm age	2670	3.64	0.59	3.26	4.14
Board independence	2586	0.98	0.30	0.83	0.92
Board size	2586	11.01	1.68	10.00	12.00

Panel B: Pearson correlations

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
(1) Innovation	1.000									
(2) Innovation impact	0.878*	1.000								
(3) Patent market value	0.904*	0.766*	1.000							
(4) Responsible AI attention	0.163*	0.121*	0.164*	1.000						
(5) Firm size	0.260*	0.160*	0.304*	0.097*	1.000					
(6) Firm performance	0.095*	0.116*	0.135*	0.111*	0.031	1.000				
(7) R&D intensity	0.462*	0.437*	0.455*	0.129*	−0.083*	0.388*	1.000			
(8) Firm age	−0.069*	−0.167*	−0.081*	0.084*	0.266*	−0.160*	−0.310*	1.000		
(9) Board independence	0.123*	0.119*	0.118*	−0.027	0.030	−0.029	0.099*	0.055*	1.000	
(10) Board size	0.015	−0.042*	0.038	0.033	0.393*	−0.062*	−0.183*	0.328*	0.059*	1.000

Note: Panel A reports the number of observations (Obs.), mean, standard deviation (SD), 25th percentile (P25), and 75th percentile (P75) for our key variables. Panel B reports pairwise Pearson correlation coefficients; * denotes $p < 0.05$.

Patent market value also exhibit a strong correlation ($r = 0.766$; all $p < 0.05$). These strong correlations affirm that firms with more patents generally have higher-quality and economically valuable innovations. *Responsible AI attention* shows positive and significant correlations with all innovation measures (correlations ranging from $r \approx 0.12$ to 0.16, $p < 0.05$), providing preliminary support for our hypothesis linking responsible AI attention to innovation. Among control variables, *Firm size* ($r = 0.260$) and *R&D intensity* ($r = 0.462$) exhibit notable positive correlations with innovation outputs, aligning with expectations that larger and more R&D-intensive firms generate more patents. *Firm age* shows a modest negative correlation with innovation ($r = -0.069$), indicating slightly lower innovation output among older firms. *Board independence* and *Board size* show small positive correlations (approximately $r \approx 0.11$ or less). Correlations among independent variables and controls are modest, with the strongest being between *Firm size* and *Board size* ($r \approx 0.39$), consistent with larger firms having larger boards. Overall, correlations indicate no significant multicollinearity concerns, reinforcing theoretical expectations linking responsible AI attention to enhanced innovation outcomes.

5 | Empirical Results

5.1 | Effect of Responsible AI Attention on Firm Innovation (Baseline Results, H1)

In our baseline models, we applied panel data analysis with industry and year fixed effects to test our hypotheses. We specifically chose industry fixed effects over firm fixed effects to control for factors shared among firms within the same industry but differing across industries, such as competitive pressures, technological trends, regulatory environments, and other industry-specific characteristics influencing innovation outcomes. Additionally, we included firm random effects to account for firm-level heterogeneity. The general model is specified in Equation (1):

$$\mathrm{Ln}\left(y_{i,t+1}+1\right) = \beta_0 + \beta_1 \mathrm{ResponsibleAIattention}_{i,t} + \mathrm{Contorl}_{i,t} + \mathrm{YearFE} + \mathrm{IndustryFE} + \varepsilon_{i,t} \quad (1)$$

where $y_{i,t+1}$ represents one of three proxies for innovation (patent counts, citations, and market value), and $\mathrm{ResponsibleAIattention}_{i,t}$ captures TMT attention to responsible AI at time t. All other measures, control variables, and fixed effects are discussed earlier and defined in Table 2. As shown in Table 4, firms dedicating attention to responsible AI produce significantly higher innovation outputs in the subsequent year, both in quantity (annual patent counts) and in quality and value of those innovations (patent citations and market value). The coefficient on responsible AI attention is positive and significant in all specifications even after adding extensive controls and fixed effects (e.g., patent counts in the full model: $\beta = 0.671$, $p < 0.01$; forward citations: $\beta = 0.567$, $p < 0.05$; patent value: $\beta = 0.843$, $p < 0.01$). This evidence aligns with prior research suggesting that emphasizing AI-related initiatives can bolster innovation outcomes. For example, Padigar et al. (2022) show favorable investor responses to announcements of AI innovations. Our results similarly demonstrate that TMT attention to responsible AI provides an innovation advantage, consistent with the ABV of the firm (Ocasio 1997), which emphasizes managerial attention as a driver of resource allocation and organizational action. Firms directing their limited managerial attention to responsible AI allocate more resources toward AI-related innovative activities, resulting in greater innovation output and impact.

Importantly, these baseline effects are robust across all three innovation measures and more model specifications. We conducted additional checks to ensure our findings were not driven by measurement choices for responsible AI attention. In particular, the construction of the responsible AI attention metric was carefully validated via dictionary-based text analysis (see Web Appendix B for validation tests), and our core finding remains unchanged after controlling for general AI discourse. When including overall AI attention (frequency of all AI mentions regardless of ethical context), the responsible AI attention coefficient remains positive and significant, confirming that it is the ethical focus, not general AI enthusiasm, that drives increased innovation.

5.1.1 | Variations in Responsible AI Attention by Principle

Responsible AI is a multifaceted concept reflecting core ethical imperatives, including beneficence, nonmaleficence, autonomy, justice, and explicability (Floridi et al. 2018), and guides organizations to develop, deploy, and oversee AI systems that maximize societal benefits and minimize harm. We argue that where a firm directs its responsible AI attention (which dimension it emphasizes) can differentially influence innovation outcomes.

Emphasizing explicability (ensuring AI decisions are understandable and transparent) likely enhances innovation acceptance and outcomes more immediately than focusing on constraining autonomy (limiting AI's independent actions to maintain human control). When TMTs advocate explainable AI (XAI), teams develop algorithms that clearly articulate the rationale behind AI outputs, thus meeting ethical requirements, improving stakeholder trust, and facilitating adoption of new AI-based products. Prior research confirms that providing explanations for AI outcomes positively impacts user perceptions and decision-making; Barredo Arrieta et al. (2020), for example, catalog many opportunities where XAI increases stakeholder confidence in AI systems. Transparency (openly disclosing how AI works and is used) also attracts external stakeholder support and investor interest, suggesting that openness about AI adds value in the marketplace (Padigar et al. 2022). In contrast, autonomy-focused attention (e.g., strict human oversight or limiting AI autonomy) has more complex implications for innovation. While essential for safety and accountability, autonomy restrictions can reduce innovation scope or discourage high-risk, radical projects, resulting instead in incremental innovations through human-AI augmentation rather than full automation (Raisch and Krakowski 2021).

From a Beneficence perspective, the priority is to ensure that AI technology actively promotes the well-being of users and society at large. This often inspires teams to create innovative AI solutions with clear social benefits, targeting areas such as healthcare, education, or resource management. Balancing this drive for good with Non-maleficence (avoiding harm) urges robust auditing or bias detection to prevent unethical or harmful outcomes, which can slow the pace of high-risk, high-reward innovative projects. Justice (fairness) demands regulatory compliance and the elimination of algorithmic bias, strengthening compliance but possibly constraining exploratory innovation. In sum, managerial attention to enabling responsible AI dimensions (e.g., explicability and transparency) likely enhances innovation outcomes more effectively than attention to constraining dimensions (e.g., strict autonomy limitations).

To further unpack the observed effects, we disaggregate responsible AI attention into its key dimensions (beneficence, non-maleficence, autonomy, justice, and explicability). We categorized earnings call transcripts into sub-categories according to the specific responsible AI dimension(s) explicitly mentioned by TMTs in each call discussion. Table 5 presents regression results for these disaggregated responsible AI attention dimensions, with each model featuring one dimension as the focal independent variable, alongside all previous controls and fixed effects.

Not all responsible AI dimensions equally impact innovation outcomes. Notably, attention to *Explicability* strongly correlates with higher subsequent innovation performance ($\beta = 1.107$, $p < 0.01$ in the Model 5 specification). Firms whose leaders emphasize explainable, transparent AI generate significantly more patents and higher-impact innovations in the subsequent year compared to those that do not. Attention to *Autonomy* (safeguarding user agency and decision-making authority) and *Non-maleficence* (avoiding harm) also exhibit positive effects ($\beta = 1.202$, $p < 0.01$ and $\beta = 0.463$, $p < 0.01$ respectively), suggesting that prioritizing user autonomy and data privacy supports greater innovative output.

Conversely, the effects of other dimensions vary; *Beneficence* (doing good) shows a positive but insignificant association, while *Justice* (fairness/equity) negatively affects innovation ($\beta = -0.469$, $p < 0.1$). A plausible explanation is that *Beneficence* reflects broad, long-term aspirations such as advancing social welfare or sustainability rather than concrete practices that

TABLE 4 | Effect of responsible AI attention on firm innovation.

	(1)	(2)	(3)	(4)	(5)
Variables	**Innovation**	**Innovation**	**Innovation**	**Innovation impact**	**Patent market value**
Responsible AI attention	0.801***	0.702***	0.671***	0.567**	0.843***
	(0.242)	(0.228)	(0.226)	(0.231)	(0.291)
Firm size		0.289***	0.348***	0.500***	0.744***
		(0.084)	(0.082)	(0.110)	(0.121)
Firm performance		−0.076	−0.045	−0.757	0.705
		(0.622)	(0.635)	(0.936)	(1.041)
R&D intensity		2.952	2.774	5.476**	11.564***
		(1.933)	(1.968)	(2.439)	(2.392)
Firm age		−0.178	−0.260	−0.194	−0.274
		(0.168)	(0.167)	(0.208)	(0.218)
Board independence			0.434***	0.391*	0.435**
			(0.147)	(0.211)	(0.222)
Board size			−0.009	−0.016	−0.023
			(0.022)	(0.036)	(0.043)
Constant	1.579***	−0.505	−1.087	−1.694	−2.646*
	(0.373)	(0.974)	(0.952)	(1.323)	(1.355)
Observations	2405	2405	2336	2336	2336
R-squared	0.316	0.310	0.310	0.544	0.155
Industry FE	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes

Note: This table presents the results of panel regressions examining how firms' responsible AI attention relates to three measures of innovation output. Models (1–3) use Innovation (the log of 1 + number of patents) as the dependent variable; Model (4) uses Innovation impact (the log of 1 + citations), and Model (5) uses Patent market value (the log of 1 + patent value) from Kogan et al. (2017). The key independent variable in all models is responsible AI attention. Controls include Firm size, Firm performance, R&D intensity, Firm age, Board independence, and Board size, along with industry and year fixed effects, robust standard errors clustered by firm are reported in parentheses. All variable definitions are as in Table 2. *$p < 0.10$, **$p < 0.05$, ***$p < 0.01$.

translate quickly into R&D output. As such, its effects may not appear in short-term measures like patents. This is consistent with evidence that "AI for Good" initiatives often take the form of multi-stakeholder projects whose innovation outcomes emerge only over longer horizons (Tomašev et al. 2020).

The negative effect of *Justice* is particularly intriguing. We propose three complementary explanations. First, from an ABV perspective, justice-oriented attention may divert scarce managerial and technical resources toward compliance and oversight activities, leaving fewer resources for exploratory R&D. This trade-off is salient in AI settings, where fairness and safety requirements often entail specialized audits and review processes. Wu and Liu (2023) provide evidence that compliance costs place substantial burdens on AI firms with limited R&D capacity, as engineers divert attention from innovation tasks to regulatory demands. Second, an emphasis on justice can lead to greater formalization of innovation processes. While fairness guardrails enhance accountability, they may slow experimentation and reduce flexibility in dynamic environments. Prior work shows that external scrutiny and regulatory constraints can heighten managerial risk aversion and suppress innovation (He and Tian 2013). Real-world practice illustrates this logic as Microsoft's 2022 AI ethics overhaul restricted access to emotion-recognition tools on fairness grounds, enhancing legitimacy but constraining downstream innovation (Hern 2022). Third, a strong TMT focus on justice may foster cultural conservatism, prompting leaders to prioritize error avoidance over experimentation. Research indicates that risk-averse organizational climates favor incremental rather than radical innovation (Alexis and Bishop 2023; Ringberg et al. 2019). In practice, the strict application of justice standards can yield more conservative innovation portfolios, as radical, ambiguous, or high-impact ideas are often screened out in favor of safer, more predictable projects.

Taken together, these dynamics suggest that justice-oriented attention, while strengthening accountability, introduces short-term trade-offs by redirecting attention, tightening processes, and reinforcing conservative innovation cultures. This helps explain why justice correlates with weaker innovation outcomes and underscores that responsible AI is not a monolithic construct but a set of distinct dimensions with heterogeneous effects.

5.1.2 | Attentional Vigilance

Beyond whether and how much attention is given to responsible AI, the consistency of that attention over time (what we term *Responsible AI vigilance*) likely shapes innovation outcomes. Attentional vigilance involves maintaining a steady managerial focus across multiple planning cycles, rather than sporadic or one-off bursts of attention (Rhee 2024). According to the ABV, consistent managerial attention ensures that strategic priorities remain actively addressed on corporate agendas, thereby facilitating cumulative organizational learning and efficient resource allocation (Eklund and Mannor 2020). Thus, when TMTs continually emphasize responsible AI principles through communications and strategic decisions, they signal a lasting commitment that motivates sustained R&D investments and encourages broader organizational engagement in ethical innovation (Gerlach et al. 2014). Conversely, transient or ephemeral attention risks the loss of momentum, limiting the innovation potential of responsible AI initiatives.

Recent research underscores the importance of sustained managerial attention. For instance, the practice of keeping attention consistently on a focal issue over multiple periods (CEO attentional vigilance) has been shown to significantly increase a firm's pursuit of exploratory innovation (Eklund and Mannor 2020; Rhee 2024). By analogy, TMTs persistently emphasizing responsible AI are more likely to see ethical AI initiatives materialize into tangible innovations. Consistent attention enables the firm to develop expertise, routines, and organizational cultures around responsible AI, which over time translate into patented inventions, high-impact innovations, and valuable new offerings aligned with ethical standards.

To empirically examine this perspective, we compare the effects of sustained versus sporadic attention to responsible AI on firm innovation. We measure *Responsible AI vigilance* by assessing the continuity of firms' responsible AI mentions over rolling multi-year windows (Rhee 2024). In Table 6, we compare firms exhibiting sustained responsible AI attention (e.g., two consecutive years) with those with intermittent or one-time mentions. Results show continuous attention significantly enhances innovation outcomes compared to intermittent attention (Eklund and Mannor 2020; Rhee 2024). Firms maintaining responsible AI vigilance over 3 years generate more patents than those with only an initial attention spike. Extending the observation window to 5 years reinforces this finding, highlighting stronger effects from sustained attention. The coefficient for *Responsible AI intensity* (total attention quantity) is positive but insignificant, echoing Rhee's (2024) finding that mere intensity of attention does not drive innovation. This lack of significance suggests that high short-term intensity may reflect episodic responses (e.g., to external pressures) rather than an enduring strategic commitment (Eklund and Mannor 2020). In our context, maintaining a consistent strategic focus on responsible AI enables cumulative learning, ongoing resource allocation, and successful execution of long-term innovation projects. These benefits are not fully realized by short-term attention bursts.

5.1.3 | Heterogeneous Innovation Types

While responsible AI attention is expected to enhance innovation overall, its impact likely varies by innovation type. Building on March (1991) and subsequent work (Benner and Tushman 2003; Gu et al. 2024), we distinguish between exploratory (radical) and exploitative (incremental) innovation. Ethical AI priorities may facilitate safer, exploitative innovations like adding transparency or privacy features by reinforcing compliance, predictability, and managerial caution (Bankins and Formosa 2023; Li et al. 2013). In contrast, exploratory efforts often involve greater ethical ambiguity and risk, which may deter TMTs focused on responsibility. As such, firms emphasizing responsible AI are more likely to engage in low-risk, incremental innovation than to pursue high-variance, radical breakthroughs.

TABLE 5 | The effects of responsible principle indicators on firm innovation.

Variables	(1) Innovation	(2) Innovation	(3) Innovation	(4) Innovation	(5) Innovation
Firm size	0.358***	0.362***	0.360***	0.357***	0.357***
	(0.082)	(0.081)	(0.082)	(0.082)	(0.082)
Firm performance	0.003	0.014	0.034	0.006	−0.006
	(0.637)	(0.633)	(0.631)	(0.637)	(0.637)
R&D intensity	3.080	3.120	3.066	3.095	3.071
	(1.979)	(1.972)	(1.975)	(1.976)	(1.977)
Firm age	−0.207	−0.205	−0.206	−0.203	−0.206
	(0.167)	(0.166)	(0.167)	(0.166)	(0.167)
Board independence	0.415***	0.406***	0.414***	0.415***	0.410***
	(0.148)	(0.149)	(0.148)	(0.148)	(0.148)
Board size	−0.010	−0.009	−0.010	−0.009	−0.010
	(0.022)	(0.022)	(0.022)	(0.022)	(0.022)
Beneficence	0.021				
	(0.101)				
Non-maleficence		0.463***			
		(0.145)			
Autonomy			1.202***		
			(0.120)		
Justice				−0.469*	
				(0.242)	
Explicability					1.107***
					(0.060)
Constant	−1.052	−1.094	−1.075	−1.066	−1.037
	(0.959)	(0.956)	(0.959)	(0.958)	(0.959)
Observations	2336	2336	2336	2336	2336
R-squared	0.308	0.309	0.309	0.309	0.309
Industry FE	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes

Note: This table reports regressions of Innovation on each of the five responsible AI dimensions: Beneficence (Model 1), Non-maleficence (Model 2), Autonomy (Model 3), Justice (Model 4), and Explicability (Model 5). All models include as controls Firm size, Firm performance, R&D intensity, Firm age, Board independence, and Board size, along with industry and year fixed effects. Robust standard errors clustered by firm are reported in parentheses. All variable definitions are as in Table 2.
*$p<0.10$, **$p<0.05$, ***$p<0.01$.

We test this by classifying patents as exploratory (novel combinations) or exploitative (incremental refinements) and estimating separate models. As shown in Table 7, *Responsible AI attention* has a positive and significant association with both *Exploratory* ($\beta=0.501$, $p<0.01$) and *Exploitative* innovation ($\beta=0.648$, $p<0.01$). Although the coefficient is larger for exploitative patents, formal tests (Seemingly Unrelated Regression model) indicate the difference is not statistically significant. Thus, responsible AI attention benefits both innovation paths, and the data do not support a marked tilt toward incremental over radical outcomes, suggesting responsibly managed AI can stimulate a balanced innovation portfolio.

5.2 | Mechanism Analyses (H2 and H3)

To better understand how responsible AI attention translates into greater innovation, we investigate two plausible mechanisms. First, we examine shifts in organizational resource allocation, particularly changes in the firm's AI-related human

TABLE 6 | Responsible AI vigilance and firm innovation.

Variables	(1) Innovation	(2) Innovation	(3) Innovation	(4) Innovation	(5) Innovation	(6) Innovation
Responsible AI vigilance (2-year)	0.043*	0.033				
	(0.022)	(0.022)				
Responsible AI vigilance (3-year)			0.062**	0.056**		
			(0.025)	(0.026)		
Responsible AI vigilance (5-year)					0.071***	0.068**
					(0.027)	(0.029)
Responsible AI intensity		0.041		0.021		0.010
		(0.045)		(0.046)		(0.047)
Firm size	0.357***	0.355***	0.356***	0.355***	0.357***	0.356***
	(0.082)	(0.081)	(0.082)	(0.081)	(0.082)	(0.081)
Firm performance	0.008	0.007	0.010	0.009	−0.007	−0.007
	(0.636)	(0.637)	(0.636)	(0.636)	(0.637)	(0.637)
R&D intensity	3.060	3.036	3.000	2.986	2.940	2.929
	(1.986)	(1.981)	(1.993)	(1.990)	(2.001)	(1.996)
Firm age	−0.203	−0.204	−0.202	−0.203	−0.201	−0.202
	(0.166)	(0.167)	(0.167)	(0.167)	(0.166)	(0.167)
Board independence	0.417***	0.417***	0.413***	0.414***	0.416***	0.416***
	(0.148)	(0.148)	(0.148)	(0.148)	(0.148)	(0.148)
Board size	−0.010	−0.010	−0.009	−0.009	−0.010	−0.010
	(0.022)	(0.022)	(0.022)	(0.022)	(0.022)	(0.022)
Constant	−1.064	−1.042	−1.062	−1.046	−1.073	−1.062
	(0.957)	(0.957)	(0.957)	(0.958)	(0.957)	(0.958)
Observations	2336	2336	2336	2336	2336	2336
R-squared	0.309	0.310	0.310	0.310	0.311	0.311
Industry FE	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes

Note: This table reports regressions examining how firms' Responsible AI vigilance, measured as the standard deviation of Responsible AI Intensity over rolling windows, relates to Innovation. Models (1–2) include Responsible AI vigilance (2-year SD); Models (3–4) include Responsible AI vigilance (3-year SD); Models (5–6) include Responsible AI vigilance (5-year SD). In Models (2), (4), and (6), Responsible AI Intensity is also included. All specifications control for Firm size, Firm performance, R&D intensity, Firm age, Board independence, and Board size, and include industry and year fixed effects. Robust standard errors clustered by firm are reported in parentheses. All variable definitions are as in Table 2. $^{*}p<0.10$, $^{**}p<0.05$, $^{***}p<0.01$.

capital (H2). According to the ABV, TMT focus on an issue leads to resource reallocation supporting that agenda (e.g., directing investments or talent to it). We therefore analyze whether firms expand their AI-related workforce following increased responsible AI attention. Results in Table 8 support this mechanism, showing that firms with high responsible AI attention subsequently grow their AI-specialist workforce significantly faster than those with lower attention levels. In our models, the coefficient on responsible AI attention for the growth in AI workers is positive and significant across all specifications (e.g., $\beta=0.289$, $p<0.05$ for the industry-adjusted share of AI-skilled employees; column 2). In essence, TMT attention to responsible AI prompts internal resource realignment, increasing hiring or redeployment of AI experts, which subsequently enhances innovation outputs. Responsible AI attention thus catalyzes capability building (talent and skills) necessary for AI-driven innovation. This finding is consistent with the UET, highlighting how TMT values and attention shape organizational investments and competencies (Hambrick and Mason 1984).

Second, we consider innovation risk management as a mechanism (H3). Responsible AI practices often involve risk assessment and mitigation (e.g., ensuring AI models are fair, safe, and compliant). We theorized that a focus on responsible AI might influence the risk profile or stability of the firm's innovation efforts. Analyzing R&D intensity fluctuations in

TABLE 7 | Heterogeneous innovation types: Exploratory versus exploitative.

Variables	(1) Exploratory	(2) Exploitative
Responsible AI attention	0.501***	0.648***
	(0.174)	(0.227)
Firm size	0.465***	0.444***
	(0.061)	(0.069)
Firm performance	1.010**	0.868
	(0.494)	(0.564)
R&D intensity	2.648*	2.330
	(1.408)	(1.638)
Firm age	−0.043	−0.084
	(0.128)	(0.181)
Board independence	0.176	0.390
	(0.135)	(0.253)
Board size	0.008	−0.025
	(0.018)	(0.019)
Constant	−3.185***	−2.992***
	(0.703)	(0.915)
Observations	1903	1903
R-squared	0.183	0.0763
Industry FE	Yes	Yes
Year FE	Yes	Yes

Note: This table presents the results of panel regressions examining how firms' responsible AI attention relates to distinct types of innovation output. Model (1) uses Exploratory innovation (the count of new, distant patents) as the dependent variable, and Model (2) uses Exploitative innovation (the count of incremental, closely related patents). Both specifications include the key independent variable, Responsible AI attention, and controls include Firm size, Firm performance, R&D intensity, Firm age, Board independence, and Board size, along with industry and year fixed effects, robust standard errors clustered by firm are reported in parentheses. All variable definitions are as in Table 2.
*$p<0.10$, **$p<0.05$, ***$p<0.01$.

relation to responsible AI attention (see Table 9), we find that firms emphasizing responsible AI exhibit significantly lower and less erratic innovation risk over time compared to those with less emphasis. Specifically, the coefficient for responsible AI attention is negative and significant in models predicting innovation risk volatility (e.g., $\beta=-0.281$, $p<0.05$ for 5-year innovation risk), indicating reduced volatility. Intuitively, incorporating ethical considerations likely moderates overly risky R&D projects, balancing ambition with caution. Thus, responsible AI attention appears to mitigate innovation-risk volatility, offering a plausible explanation for improved innovation performance. In summary, TMT's focus on responsible AI seems to both (a) mobilize vital resources (human capital) toward AI innovation and (b) instill practices that reduce innovation-related risks, facilitating a more consistent and robust innovation output.

5.3 | Moderation Analyses (H4 and H5)

Finally, we explore several contingency factors that might shape the impact of responsible AI attention and conduct additional robustness tests to address potential endogeneity and omitted-variable concerns. Specifically, we examine how internal and external conditions moderate the responsible AI–innovation relationship (as hypothesized) and assess alternative explanations.

5.3.1 | Moderation by Technological Intensity and Investor Horizon

In Table 10, we introduce three interaction effects corresponding to H4 and H5. Column (1) examines *Industry technological intensity* as a moderator (H4), interacting responsible AI attention with a high-tech industry indicator (binary variable for above-median R&D intensity). The interaction term is negative and significant ($\beta=-0.342$, $p<0.05$), indicating that responsible AI attention has a weaker effect on innovation in high-tech industries (and conversely a stronger effect in low-tech industries). A plausible explanation is that high-tech firms already prioritize innovation and AI, making additional "responsibility" focus less incrementally beneficial given their high baseline innovation levels. In contrast, for firms in less tech-intensive environments, responsible AI attention could represent a transformative commitment, opening new innovation avenues and substantially elevating innovation efforts.

Column (2) tests the moderation by *Shareholder investment horizon* (H5). As Sirmon and Hitt (2003, 343) suggest, long-term-focused shareholders are those who are "invested without threat of liquidation for long periods." We operationalize shareholder investment horizon as the ratio of a firm's stock shares traded to its total outstanding shares in a given year. We split firms by the dominant investor horizon (using measures of the firm's investor base preference for short-term vs. long-term holdings) and interact responsible AI attention with a long-term investor indicator. The interaction term is negative and significant ($\beta=-0.745$, $p<0.01$), meaning that responsible AI attention's impact on innovation is stronger among firms with short-term-oriented investors and weaker with longer-horizon investors. This suggests that responsible AI attention notably enhances innovation under short-term market pressures. Firms with more short-term shareholders often face pressure for immediate performance, potentially limiting risky, long-term innovation investments. Our findings imply that emphasizing responsible AI serves as a credible commitment or visionary signal, justifying continued innovation investments despite short-term pressures. Thus, responsible AI attention may help bridge the gap between short-horizon investors and long-term innovation goals, spurring innovation even in environments typically discouraging it.

5.4 | Further Analysis: The Role of CTO

A Chief Technology Officer (CTO) signals a firm's commitment to technological leadership and brings specialized expertise to translate responsible AI priorities into concrete R&D actions (Cetindamar and Pala 2011). Drawing on UET (Hambrick and Mason 1984), including a technology-focused executive can shift

TABLE 8 | AI investment mechanism.

Variables	(1) All AI share	(2) Categorized share	(3) 3-skill share	(4) 5-skill share	(5) Narrow AI share
Responsible AI attention	0.149*	0.289**	0.281*	0.289**	0.346*
	(0.088)	(0.145)	(0.144)	(0.142)	(0.178)
Firm size	0.053	0.140**	0.132**	0.149**	0.213***
	(0.036)	(0.060)	(0.063)	(0.060)	(0.074)
Firm performance	0.732*	0.880	0.982	0.938	1.069
	(0.389)	(0.651)	(0.638)	(0.614)	(0.751)
R&D intensity	4.911***	7.459***	7.364***	7.542***	8.334***
	(0.958)	(1.666)	(1.660)	(1.578)	(2.264)
Firm age	−0.119*	−0.228**	−0.233**	−0.225**	−0.241*
	(0.062)	(0.109)	(0.108)	(0.106)	(0.125)
Board independence	−0.033	−0.005	−0.013	−0.015	−0.003
	(0.069)	(0.128)	(0.127)	(0.124)	(0.162)
Board size	−0.006	−0.026	−0.022	−0.027	−0.047*
	(0.011)	(0.022)	(0.023)	(0.023)	(0.027)
Constant	0.849*	0.555	0.606	0.461	0.087
	(0.500)	(0.899)	(0.912)	(0.884)	(1.114)
Observations	1572	1572	1571	1568	1572
R-squared	0.280	0.256	0.256	0.260	0.223
Industry FE	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes

Note: This table reports regressions examining how firms' Responsible AI attention relates to subsequent shifts in their AI investment (AI skilled workers portfolios). Columns (1–5) use as dependent variables the industry-adjusted shares of: (1) All-AI workers, (2) Categorized AI workers, (3) 3-skill AI workers, (4) 5-skill AI workers, and (5) Narrow AI workers. The key independent variable in all models is Responsible AI attention. All specifications include controls for Firm size, Firm performance, R&D intensity, Firm age, Board independence, and Board size, along with industry and year fixed effects. Robust standard errors clustered by firm are reported in parentheses. All variable definitions are as in Table 2. *$p<0.10$, **$p<0.05$, ***$p<0.01$.

how TMT perceives and pursues innovation opportunities. Prior work shows CTO appointments are associated with higher patent output and increased radical innovation (Chung and Kang 2019; Gu et al. 2024). Thus, in principle, CTO presence should amplify the effect of responsible AI attention on innovation.

However, from an ABV perspective, executive attention is a finite cognitive resource. The marginal value of additional structural attention channels may decline when TMT focus on responsible AI is already high, evidenced by frequent strategy discussions, major investments, and sustained public disclosures. In such cases, the CTO's influence may be symbolic or redundant. Moreover, many CTO-led efforts shape downstream processes, such as product launch, system integration, and platform governance, that extend beyond the temporal scope of near-term patent metrics.

In supplementary analysis, we code CTO presence as a binary variable equal to 1 if the firm lists a CTO (or equivalent, such as CIO or CDO), and 0 otherwise. Untabulated results show that the interaction term *Responsible AI attention × CTO* is not statistically significant, suggesting that job title alone may be an insufficient proxy for strategic influence. Ceiling effects could also explain this null result: in firms where responsible AI attention is already routinized, there may be little marginal scope for the CTO to enhance innovation outcomes. In addition, the varying scopes of responsibility associated with CTO, CIO, and CDO titles create heterogeneity in role definitions that may dilute observed effects.

5.5 | Endogeneity and Robustness Checks

We conducted additional analyses to rule out endogeneity or omitted-variable bias. First, we employed an entropy balancing approach (Hainmueller 2012) and a Heckman two-step selection model (Heckman 1979) to account for non-random adoption of responsible AI attention (see Table 11). After reweighting the sample via entropy balancing to create comparable treatment (responsible AI attention) and control groups, the effect of responsible AI attention on innovation remains positive and significant ($\beta = 0.167$, $p < 0.05$; column 1). This result provides confidence that systematic differences between adopters and non-adopters are not driving our findings. Likewise, the

TABLE 9 | R&D risk mechanism.

Variables	(1) 2-year innovation risk	(2) 3-year innovation risk	(3) 5-year innovation risk
Responsible AI attention	−0.203**	−0.252**	−0.281**
	(0.098)	(0.103)	(0.120)
Firm size	0.053	0.049	−0.014
	(0.043)	(0.045)	(0.066)
Firm performance	0.839	0.620	0.022
	(0.595)	(0.647)	(0.600)
R&D intensity	−3.522***	−3.849***	−3.203**
	(0.963)	(0.892)	(1.296)
Firm age	−0.043	−0.016	0.006
	(0.085)	(0.082)	(0.118)
Board independence	0.017	0.007	−0.046
	(0.146)	(0.122)	(0.127)
Board size	−0.032	−0.035	−0.005
	(0.027)	(0.026)	(0.024)
Constant	1.157**	1.167**	1.552**
	(0.488)	(0.474)	(0.671)
Observations	1612	1616	1622
R-squared	0.0666	0.0767	0.0463
Industry FE	Yes	Yes	Yes
Year FE	Yes	Yes	Yes

Note: This table reports regressions examining how firms' Responsible AI attention relates to subsequent innovation risk, measured as the industry-adjusted coefficient of variation (CV) of R&D intensity over different rolling windows. Columns (1–3) use the 2-, 3-, and 5-year R&D CV ratios, respectively, as the dependent variable. In all models, the key independent variable is Responsible AI attention. Controls include Firm size, Firm performance, R&D intensity, Firm age, Board independence, and Board size, along with industry and year fixed effects. Robust standard errors clustered by firm are in parentheses. All variable definitions are as in Table 2. $^{*}p<0.10$, $^{**}p<0.05$, $^{***}p<0.01$.

Heckman two-stage selection model (columns 2–3) addresses potential unobserved factors influencing both responsible AI attention and innovation outcomes. In the first-stage probit model, we used industry peer responsible AI attention (the prevalence of responsible AI discussion among industry peers) as an instrument, which strongly predicts a firm's own responsible AI attention ($\beta = 3.448$, $p < 0.01$). This is consistent with mimetic pressures whereby managers are more likely to emphasize responsible AI when their peers do so. Crucially, in the second-stage innovation equation (including the inverse Mills ratio), responsible AI attention remains positive and significant ($\beta = 0.143$, $p < 0.05$), while the inverse Mills ratio is insignificant, indicating no evidence of substantial sample selection bias. These robustness checks confirm our main findings and support a causal interpretation.

TABLE 10 | Moderation effects of technological intensity, shareholder investment horizon and CTO.

Variables	(1) Innovation	(2) Innovation	(3) Innovation
Responsible AI attention	0.485***	0.856***	0.121
	(0.082)	(0.093)	(0.124)
Responsible AI attention × High tech	−0.342**		
	(0.172)		
Shareholder investment horizon_Long		0.299***	
		(0.084)	
Responsible AI attention × Shareholder investment horizon_Long		−0.745***	
		(0.143)	
CTO			0.539***
			(0.128)
Responsible AI attention × CTO			0.154
			(0.248)
Firm size	0.803***	0.800***	0.806***
	(0.038)	(0.038)	(0.038)
Firm performance	−0.889**	−1.013**	−0.932**
	(0.451)	(0.441)	(0.435)
R&D intensity	12.834***	12.750***	12.311***
	(0.831)	(0.828)	(0.820)
Firm age	0.149**	0.114	0.188***
	(0.073)	(0.071)	(0.069)
Board independence	0.358***	0.356***	0.368***
	(0.114)	(0.116)	(0.114)
Board size	0.005	0.010	0.002
	(0.022)	(0.022)	(0.022)
Constant	−6.827***	−6.668***	−6.790***
	(0.405)	(0.394)	(0.393)
Observations	2405	2405	2405
R-squared	0.501	0.496	0.499
Industry FE	Yes	Yes	Yes
Year FE	Yes	Yes	Yes

Note: Standard errors are in parentheses. $^{*}p<0.10$, $^{**}p<0.05$, $^{***}p<0.01$.

Finally, we conducted additional robustness analyses (detailed in Appendix D) to further validate our findings. Table WA-D1 controls for firms' pre-existing *Innovation culture* using an innovation culture index (Li, Mai, et al. 2021) addressing concerns that a strong innovative culture might simultaneously drive responsible AI adoption and patent output, creating a spurious correlation. The responsible AI attention effect remains robust and significant, suggesting it is not

TABLE 11 | Endogeneity correction.

	Entropy-balanced	Heckman selection model	
	(1)	(2)	(3)
Variables	**Innovation**	**Responsible AI attention (first stage)**	**Innovation (second stage)**
Responsible AI attention	0.167**		0.143**
	(0.080)		(0.056)
Firm size	0.808***	0.135***	0.729***
	(0.046)	(0.027)	(0.059)
Firm performance	−1.635***	1.634***	−2.805***
	(0.522)	(0.312)	(0.500)
R&D intensity	13.724***	2.070***	18.063***
	(1.021)	(0.453)	(0.661)
Firm age	0.245**	0.404***	0.125
	(0.117)	(0.050)	(0.106)
Board independence	0.722***	−0.176*	0.291**
	(0.151)	(0.090)	(0.124)
Board size	−0.029	−0.020	−0.034
	(0.027)	(0.018)	(0.027)
AI prevalence		3.448***	
		(0.281)	
IMR			−0.020
			(1.626)
Constant	−6.614***	−7.821***	−4.544***
	(0.501)	(0.536)	(0.547)
Observations	2405	2405	2405
R-squared	0.488		
Wald χ^2		273.78***	1210.60***
Industry FE	Yes	No	Yes
Year FE	Yes	No	Yes

Note: Standard errors are in parentheses. $^{*}p<0.10$, $^{**}p<0.05$, $^{***}p<0.01$.

merely capturing an underlying pro-innovation culture. Table WA-D2 adds a control for *General AI attention* (AI mentions unrelated to responsibility or ethics); responsible AI effects stay positive and significant, reinforcing the distinct importance of the ethical dimensions in driving innovation. Lastly, Table WA-D3 tests for reverse causality by employing lagged responsible AI attention variables (years t–1, t–2) to predict innovation outcomes at year t, and the relationship remains consistently positive and significant. All variable definitions are provided in Table 2.

6 | Discussion and Conclusion

This article examines how TMT attention to responsible AI relates to firm innovation. The analysis shows that firms whose executives devote greater attention to responsible AI produce more and higher-impact innovations. This relationship operates in part through increased investment in AI-skilled human capital and reduced innovation risk, indicating that responsible AI attention helps firms allocate resources and manage uncertainty more effectively. The effects, however, vary across contexts. The positive association is stronger in industries with lower technological intensity and when shareholders emphasize short-term performance, whereas the presence of a CTO does not alter the relationship. Exploratory analyses further reveal variation in the content of responsible AI attention: justice-oriented attention is less conducive to innovation, while attention to explicability and autonomy is positively related to it. Overall, these findings provide an integrated view of how managerial attention to responsible AI is linked to firms' innovative activity under uncertainty. To synthesize these results, Table 12 summarizes the main findings, related insights, and theoretical and practical implications.

6.1 | Theoretical Implications

Our theoretical contribution is threefold. First, we bridge two important streams in AI-enabled innovation (Gama and Magistretti 2025; Mariani et al. 2023; Verganti et al. 2020) and responsible innovation (Gutierrez et al. 2022; Spanjol et al. 2024; Stilgoe et al. 2013), by integrating the ABV with a large-scale measure of TMT attention to responsible AI. Prior ABV research has largely focused on how market signals or competitive pressures capture managerial attention (Joseph et al. 2024; Ocasio 1997) or how attention to technological change supports strategic renewal (Eggers and Kaplan 2009; Nadkarni and Barr 2008). Our article demonstrates that responsible AI can also become a focal point of managerial attention and, importantly, that such attention is positively associated with higher levels of firm innovation. This extends ABV by showing that not only economic or technological cues but also normative concerns can guide strategic attention, shaping innovation outcomes. Drawing on insights from UET (Carpenter et al. 2004; Hambrick and Mason 1984), we further explain how executives' ethical orientations may shape what issues receive managerial attention and how this, in turn, affects firm innovation outcomes. In doing so, we conceptually enrich the ABV and respond to recent calls for stronger theoretical integration in digital and AI innovation research (Appio et al. 2021).

Second, we contribute by unpacking the mediating mechanisms through which responsible AI attention translates into superior innovation results. Our analysis indicates that when top managers prioritize responsible AI initiatives, their firms respond by investing in AI-skilled human capital and adopting practices that mitigate innovation risks. In other words, an ethical AI focus prompts the organization to build new capabilities (e.g., hiring/training AI talent and establishing robust governance processes) while simultaneously reducing downside risk in R&D, thereby enabling the firm to innovate both more extensively and more reliably. Theoretically, specifying these pathways enriches the ABV by embedding a socio-ethical dimension into the core of managerial attention. Whereas classic ABV emphasizes that attention drives resource allocation and action (Ocasio 1997, 2011; Joseph et al. 2024), our findings show how this occurs in the context of responsible AI: ethics-centered attention mobilizes critical resources and stabilizes the innovation process. This result moves beyond viewing responsible AI as a mere compliance task and highlights its role as a dynamic strategic asset that actively shapes resource deployment and strengthens innovation resilience. By clarifying how executive ethical orientations channel attention to build capabilities and guard against failure, we address a previously implicit link in UET (Hambrick and Mason 1984), illustrating how TMT values (here, a commitment to responsible AI) materially influence firm outcomes through the focused orchestration of talent and risk management (Carpenter et al. 2004; Hambrick and Mason 1984).

Third, our article delineates important boundary conditions and sources of heterogeneity in the responsible AI attention–innovation link. We find that the innovation benefits of responsible AI attention are highly context-dependent, particularly strong in environments with lower technological intensity and under short-term oriented investor pressure. This pattern complements prior ABV research emphasizing technological and competitive contexts as key contingencies of managerial attention (Joseph et al. 2024; Nadkarni and Barr 2008). In such challenging contexts, a strong emphasis on responsible AI appears to function as a buffer that sustains innovation capacity by offsetting structural disadvantages or short-termism through building stakeholder trust and long-term vision (Li, Li, et al. 2021; Teece 2007). At the same time, not all responsible AI attentional content is equally beneficial. Our results reveal heterogeneous effects across responsible AI sub-dimensions: for example, an explicability-oriented focus (emphasizing transparency in AI) significantly boosts innovation, whereas a justice-oriented focus (emphasizing fairness/equity) can inadvertently dampen it. The former likely fosters innovation by enhancing stakeholder buy-in and organizational learning (Argote and Miron-Spektor 2011; Spanjol et al. 2024), whereas the latter may constrain experimentation by diverting resources and imposing tighter constraints (Wu et al. 2015). This nuance shows that distinct ethical orientations can function as either enabling or constraining mechanisms for innovation, illustrating how moral considerations become embedded in managerial attention and thereby extending the ABV to include normative dimensions. Moreover, and somewhat unexpectedly, we did not find evidence that the presence of a CTO amplifies the effect of responsible AI attention on innovation performance. This null finding suggests that simply appointing a technology executive, an ostensibly attention-structuring role, is insufficient to enhance innovation; what matters more is the collective attentional engagement of the TMT with responsible AI issues (Joseph et al. 2024; Li et al. 2013; Ocasio 2011). By illuminating these contingencies, our work extends the ABV beyond generic attentional effects to show that where and how managerial attention is directed (in terms of contextual fit and ethical focus) critically shapes its impact on innovation.

Methodologically, we respond to calls for rigorous and replicable innovative AI research (Kaplan and Haenlein 2019; Spanjol et al. 2024). By combining a dictionary-based text-mining approach (NLP, LLM) with longitudinal panel methods across diverse industries and triangulating executive discourse with patent-based innovation measures, our article demonstrates that managerial attention to responsible AI meaningfully shapes innovation. This methodological design matters because it provides one of the first scalable ways to systematically capture responsible AI attention in executive discourse, a construct that has been difficult to observe directly in prior research. Beyond this article, the dictionary offers a transferable tool for examining responsible AI attention across industries, geographies, and organizational settings. Our multi-method approach enhances validity and offers a replicable approach for advancing empirical research on responsible AI in innovation contexts.

6.2 | Managerial Implications

This article offers actionable insights for executives, innovation managers, and policymakers on embedding responsible AI attention into organizational practices to foster sustainable innovation. For top executives, our findings show that sustained

TABLE 12 | A summary of key findings, contributions, and future directions.

Key findings	Insights from prior research	Theoretical contributions	Practical implications	Limitations & Future research
Finding 1: Responsible AI attention enhances innovation performance via two mediating mechanisms: investment in AI-relevant human capital and mitigation of innovation risk.	Prior studies show that managerial attention directs resources toward innovation (Li et al. 2013; Ocasio 1997; Yadav et al. 2007). Yet attention to responsible AI has largely been treated as a compliance concern, with limited exploration of its implications for innovation (Mikalef et al. 2022; Shneiderman 2021).	• Specifies how ethical orientations shape innovation outcomes through attentional mediation, extending ABV into normative domains. • Identifies two mechanisms, investment in AI-relevant human capital and mitigation of innovation risk, thereby clarifying the pathway UET left implicit.	• Treat responsible AI as a strategic lever for building AI-relevant skills and strengthening governance that mitigates innovation risk (*for executives*). • Positions responsible AI as a portfolio management tool that enhances innovation reliability and scope through sustained investment in talent and risk management (*for executives, innovation managers*).	***Limitation***: Reliance on archival data constrains observation of how these mechanisms operate within organizations. ***Future research***: Use surveys, interviews, or case studies to capture human capital building and risk management processes more directly.
Finding 2: Responsible AI attention has heterogeneous effects: explicability and autonomy orientations foster innovation, whereas justice orientation may suppress it.	Prior studies suggest that ethical principles foster trust, legitimacy, and adoption of AI systems (Barredo Arrieta et al. 2020; Floridi et al. 2018; Glikson and Woolley 2020). Yet little research disaggregates their heterogeneous effects or examines how distinct dimensions translate into firm-level outcomes such as innovation.	• Extends ABV by theorizing the heterogeneous effects of ethical orientations in shaping innovation outcomes. • Shows that distinct dimensions of responsible AI attention can function as enabling or constraining mechanisms, thereby integrating normative cognition into attentional theory.	• Prioritize explicability- and autonomy-oriented practices that build trust and adoption, recognizing their stronger innovation benefits (*for executives*). • Warns that excessive focus on justice concerns can raise compliance burdens and slow experimentation, requiring balance between safeguards and flexibility (*for innovation managers, policymakers*).	***Limitation***: Potential endogeneity and omitted variables (e.g., unobserved managerial values) may bias dimension-level effects. ***Future research***: Apply quasi-experimental designs or natural experiments to strengthen causal inference.

(Continues)

TABLE 12 | (Continued)

Key findings	Insights from prior research	Theoretical contributions	Practical implications	Limitations & Future research
Finding 3: The benefits of responsible AI attention on innovation are context-dependent, with particularly strong effects in low-tech sectors and under short-term investor pressure.	Prior ABV studies emphasize technological dynamism and competitive intensity as key boundary conditions (Joseph et al. 2024; Nadkarni and Barr 2008). Yet alternative contextual moderators of attention–innovation relationships remain underexplored, especially in the context of responsible AI.	• Identifies technological intensity and investor horizons as boundary conditions for the effects of responsible AI attention, extending ABV's treatment of contextual moderators. • Positions responsible AI attention as an enabling capability that sustains innovation under structural and temporal pressures, enriching innovation theory with a normative dimension.	• Shows that responsible AI attention can be particularly valuable in structurally disadvantaged or high-pressure contexts, guiding resource allocation and governance priorities (*for executives, innovation managers*). • Suggests that fostering responsible AI adoption may help level the playing field for low-tech sectors and short-horizon firms, supporting inclusive innovation growth (*for policymakers*).	***Limitation***: Evidence comes mainly from large firms, limiting generalizability to SMEs or other contexts. ***Future research***: Conduct studies in SMEs or emerging markets to test boundary conditions across firm sizes and environments.
Methodological: A scalable dictionary-based NLP/LLM approach is developed to measure responsible AI attention in executive discourse, providing a replicable and extensible tool for future empirical research.	Prior research has applied text analysis to capture intangible constructs such as managerial cognition and organizational culture (Berger et al. 2020; Li, Mai, et al. 2021). However, systematic approaches to measure managerial attention to responsible AI remain underdeveloped.	• Introduces a dictionary-based NLP/LLM approach to systematically capture responsible AI attention in executive discourse, addressing a key measurement gap. • Provides a replicable and extensible methodological template for future studies across industries, geographies, and cultural contexts.	• Provides a scalable framework to benchmark responsible AI attention and guide governance practices, and can be applied by firms to assess their own responsible AI focus over time (*for executives, policymakers*).	***Limitation***: Dictionary methods may miss nuances in managerial discourse and require further validation across contexts. ***Future research***: Extend and adapt the dictionary for cross-industry, cross-cultural, and longitudinal applications.

attention to responsible AI strengthens innovation performance by mobilizing resources and reducing risk. The central challenge is therefore managing the high-risk, high-reward nature of AI-driven innovation initiatives. Positioning responsible AI as a portfolio management tool enables leaders to mitigate reputational and regulatory risks while maintaining the capacity for bold and exploratory innovation. To achieve this balance, two actions are particularly important. First, executives need to design and institutionalize governance structures and allocate resources with the explicit aim of enabling innovation, not merely ensuring compliance. This requires visible commitment, such as creating executive-level roles or committees for AI ethics, investing in training programs, and establishing principles that every AI-driven project must satisfy. Unilever, for example, formed a governance committee that embedded responsible AI in its data strategy, defined firm-wide standards on oversight and accountability, and invested in the tools, personnel, and procedures necessary for their implementation (Davenport and Bean 2023). By providing this stable governance framework, executives give R&D teams both the guardrails and the psychological safety needed to pursue ambitious innovations. Second, accountability for responsible AI must be shared across the C-suite rather than delegated solely to technical officers. Microsoft's leadership adopted a Responsible AI Standard that placed people at the center of system design. Under this standard, the company curtailed or withdrew ethically problematic features, such as emotion-recognition algorithms, despite their commercial promise (Hern 2022). This strategic sacrifice safeguarded trust in Microsoft's flagship AI platforms and signaled that long-term legitimacy matters more than short-term gain. Taken together, these examples illustrate how sustained executive attention to responsible AI stabilizes innovation pipelines, preserves legitimacy, and strengthens innovation performance, consistent with our empirical findings.

For innovation managers, the task is to translate executive-level responsible AI commitments into operational practices that both safeguard and accelerate innovation. Our findings highlight two complementary functions. The first is capability building, which involves integrating ethicists, fairness experts, and risk specialists into project teams to strengthen AI-relevant human capital. The second is process integration, which requires embedding responsible AI checkpoints into project workflows such as bias audits, multi-stakeholder design reviews, and stress tests of prototypes for privacy, fairness, and security risks (a process consistent with the mediating mechanisms in our analysis). Introducing these steps early in R&D stage-gates helps prevent costly project failures and increases the stability of innovation pipelines. The case of Amazon's abandoned AI recruiting tool vividly illustrates the stakes. Bias in training data led the system to downgrade female candidates, ultimately forcing the firm to scrap the project entirely (Dastin 2018). Such failures show that neglecting responsible AI can waste R&D resources and damage trust, whereas systematic attention enables teams to anticipate ethical pitfalls and refine innovations responsibly. In short, our findings suggest that responsible AI practices should not be treated as isolated compliance checks but embedded into the core routines of innovation management.

For policymakers, our evidence shows that managerial attention to responsible AI enhances both the quantity and impact of innovation, suggesting that public policy that promotes such attention can strengthen firms' innovative capacity while safeguarding society. The central task for regulators is therefore to create an institutional environment that accelerates innovation under short-term pressures by reducing uncertainty, building trust, and encouraging private investment in experimentation. Effective responsible AI governance can function not merely as compliance, but as a source of competitive advantage, and clear and consistent frameworks are essential to realize this potential. The European Union's forthcoming AI Act, the world's first comprehensive AI regulation, establishes risk-based rules for testing, documentation ("auditability"), and oversight in high-stakes domains such as healthcare, finance, and employment (European Commission 2025). In the United States, the National Institute of Standards and Technology's (NIST) AI Risk Management Framework (AI RMF) provides a voluntary guideline to help firms operationalize trustworthiness in AI systems (NIST 2023). Though not binding law, it creates a common language for responsible AI and encourages firms to internalize best practices early. Such policy clarity lowers coordination and compliance costs, shortens approval timelines, and builds public trust. In turn, this enables firms to pursue ambitious projects with greater stability. Policymakers can further reinforce these effects by requiring disclosure of responsible AI practices, supporting regulatory sandboxes, and aligning incentives such as procurement, grants, and tax benefits with responsible-by-design initiatives. In this context, our approach also provides a practical tool for monitoring and benchmarking firms' progress in implementing responsible AI practices. Specifically, the dictionary-based NLP/LLM framework developed in this article can systematically capture managerial attention to responsible AI, offering regulators and industry bodies a scalable means of assessing how firms prioritize responsible AI in their governance practices and public communications.

6.3 | Limitations and Future Research

This article has three key limitations that point to promising avenues for future research. *First*, it relies on archival data, which constrains our ability to fully capture how responsible AI attention is expressed within organizations. While we leverage earnings call transcripts and patent information, these sources provide only a partial view of firms' concrete investments in responsible AI initiatives. Future studies employing survey methods or in-depth interviews could generate richer evidence on the nature and scope of these investments and clarify how they translate into both product and process innovation. Though prior research has drawn links between responsible innovation and patent outputs, a holistic examination of how responsible AI practices influence broader firm performance, ranging from market share to organizational resilience, remains underexplored. Future research could therefore examine the mechanisms, timing, and consequences of different responsible AI strategies to provide a more nuanced understanding of their impacts across various performance dimensions.

Second, our sample primarily focuses on large firms, which pose limitations for generalizing our findings to small and medium-sized enterprises (SMEs). Larger corporations usually have greater resources, established R&D procedures, formal governance structures, and broader stakeholder networks, all of which may shape how responsible AI principles are adopted and enforced. Smaller firms, by contrast, often face distinct challenges, such as resource constraints, limited formalized processes, or fewer specialized personnel dedicated to ethical AI oversight. Consequently, the adoption and impact of responsible AI in SMEs might differ significantly from the patterns observed in large firms. Future studies could employ qualitative case studies, targeted surveys, or mixed method approaches to examine the specific challenges and opportunities in SMEs. Such work would deepen understanding of the scalability of responsible AI and help tailor best practices for firms of different sizes and sectors.

Third, as with any observational study, endogeneity is a concern. Possible sources include reverse causality, where more innovative firms are more likely to discuss responsible AI, omitted variables such as unobserved managerial values that affect both responsible AI attention and innovation, and non-random selection. Formal identification using instrumental variables or natural experiments is unlikely in our setting. A valid instrument must be correlated with responsible AI and satisfy the exclusion restriction, affecting the dependent variable only through it. Finding a valid instrument is highly challenging. Potential instruments such as corporate IT infrastructure, managerial technology expertise, or access to data science talent are likely to affect innovation directly and therefore violate this condition. Using weak or ad hoc instruments could introduce more bias. Our article period also lacks a suitable exogenous policy or market shock for a clean quasi-experiment. To address these concerns, we use lagged measures to establish temporal ordering, include additional controls, apply entropy balancing to improve covariate balance, and estimate a Heckman two-step selection model. These steps reduce, but do not eliminate, confounding. We therefore interpret the results as robust associations that are consistent with a causal effect but do not by themselves prove it. Future work could seek settings with exogenous variation, for example difference-in-differences designs based on staggered adoption of AI governance rules or event studies following the 2022 launch of ChatGPT, a widely adopted LLM. Finally, while our dictionary-based NLP/LLM approach provides a scalable way to capture responsible AI attention, it inevitably abstracts from contextual nuance in managerial discourse. Future research could refine and extend the dictionary across industries and cultural settings to enhance validity and generalizability.

Ethics Statement

The authors have read and agreed to the Committee on Publication Ethics (COPE) international standards for authors.

Conflicts of Interest

The authors declare no conflicts of interest.

Data Availability Statement

The data that support the findings of this article are available from the corresponding author upon reasonable request.

Endnotes

[1] We use the term "Justice" in our study to maintain consistency with the foundational ethical frameworks from which our five dimensions are derived (e.g., Floridi et al. 2018). We acknowledge its close and important relationship with "fairness," which is often used to describe the practical or computational implementation of the broader principle of justice in AI systems.

[2] For example, Microsoft's creation of a Responsible AI Risk Manager role (with accompanying training and budget) illustrates how focused attention translates into concrete resource commitments aimed at fostering innovative yet ethical AI solutions.

[3] Dictionary development followed an empirically guided approach (Humphreys and Wang 2018; Homburg et al. 2020), combining case study analysis (Yin 2018) with literature triangulation. Full methodology and validity checks (Berger et al. 2020) are detailed in Web Appendix B.

[4] Text classification used a Bag-of-Words (BoW) model (Zhang et al. 2010; Wang et al. 2012) and LLM prompting (Chae and Davidson 2025) to identify TMT attention to responsible AI. Detailed procedures are provided in Web Appendix C.

References

Acharya, V. V., and K. V. Subramanian. 2009. "Bankruptcy Codes and Innovation." *Review of Financial Studies* 22, no. 12: 4949–4988.

Aghion, P., J. Van Reenen, and L. Zingales. 2013. "Innovation and Institutional Ownership." *American Economic Review* 103, no. 1: 277–304.

Ahn, Y. 2020. "A Socio-Cognitive Model of Sustainability Performance: Linking CEO Career Experience, Social Ties, and Attention Breadth." *Journal of Business Ethics* 175, no. 2: 303–321.

Alder Hey Children's NHS Foundation Trust. 2025. "Alder Hey Transforming Care Through AI." Alder Hey Children's NHS Foundation Trust. https://www.alderhey.nhs.uk/alder-hey-transforming-care-through-ai/.

Alexis, J., and S. Bishop. 2023. "Organizational Risk Attitude and Innovation Choice." SSRN Preprint no. 4547734.

Appio, F. P., F. Frattini, A. M. Petruzzelli, and P. Neirotti. 2021. "Digital Transformation and Innovation Management: A Synthesis of Existing Research and an Agenda for Future Studies." *Journal of Product Innovation Management* 38, no. 1: 4–20.

Argote, L., and E. Miron-Spektor. 2011. "Organizational Learning: From Experience to Knowledge." *Organization Science* 22, no. 5: 1123–1137.

Artz, K. W., P. M. Norman, D. E. Hatfield, and L. B. Cardinal. 2010. "A Longitudinal Study of the Impact of R&D, Patents, and Product Innovation on Firm Performance." *Journal of Product Innovation Management* 27, no. 5: 725–740.

Asatiani, A., P. Malo, P. R. Nagbøl, E. Penttinen, T. Rinta-Kahila, and A. Salovaara. 2021. "Sociotechnical Envelopment of Artificial Intelligence: An Approach to Organizational Deployment of Inscrutable Artificial Intelligence Systems." *Journal of the Association for Information Systems* 22, no. 2: 325–352.

Babina, T., A. Fedyk, A. He, and J. Hodson. 2024. "Artificial Intelligence, Firm Growth, and Product Innovation." *Journal of Financial Economics* 151: 103745.

Bankins, S., and P. Formosa. 2023. "The Ethical Implications of Artificial Intelligence (AI) for Meaningful Work." *Journal of Business Ethics* 185, no. 4: 725–740.

Barredo Arrieta, A., N. Díaz-Rodríguez, J. Del Ser, et al. 2020. "Explainable Artificial Intelligence (XAI): Concepts, Taxonomies, Opportunities and Challenges Toward Responsible AI." *Information Fusion* 58: 82–115.

Bauer, K., M. von Zahn, and O. Hinz. 2023. "Expl(AI)ned: The Impact of Explainable Artificial Intelligence on Users' Information Processing." *Information Systems Research* 34, no. 4: 1582–1602.

Bellamy, R. K. E., K. Dey, M. Hind, et al. 2019. "AI Fairness 360: An Extensible Toolkit for Detecting and Mitigating Algorithmic Bias." *IBM Journal of Research and Development* 63, no. 4/5: 4:1–4:15.

Benner, M. J., and M. L. Tushman. 2003. "Exploitation, Exploration, and Process Management: The Productivity Dilemma Revisited." *Academy of Management Review* 28, no. 2: 238–256.

Berente, N., B. Gu, J. Recker, and R. Santhanam. 2021. "Managing Artificial Intelligence." *MIS Quarterly* 45, no. 3: 1433–1450.

Berger, J., A. Humphreys, S. Ludwig, W. W. Moe, O. Netzer, and D. A. Schweidel. 2020. "Uniting the Tribes: Using Text for Marketing Insight." *Journal of Marketing* 84, no. 1: 1–25.

Bigham, T., S. Nair, S. Soral, et al. 2022. "AI and Risk Management: Innovating With Confidence." Deloitte Insights. https://www.deloitte.com/content/dam/assets-shared/legacy/docs/perspectives/2022/deloitte-gx-ai-and-risk-management.pdf.

Bosch-Sijtsema, P., and J. Bosch. 2015. "User Involvement Throughout the Innovation Process in High-Tech Industries." *Journal of Product Innovation Management* 32, no. 5: 793–807.

Bouquet, C., and J. Birkinshaw. 2008. "Weight Versus Voice: How Foreign Subsidiaries Gain Attention From Corporate Headquarters." *Academy of Management Journal* 51, no. 3: 577–601.

Bouschery, S. G., V. Blazevic, and F. T. Piller. 2023. "Augmenting Human Innovation Teams With Artificial Intelligence: Exploring Transformer-Based Language Models." *Journal of Product Innovation Management* 40, no. 2: 139–153.

Brown, S., S. A. Hillegeist, and K. Lo. 2004. "Conference Calls and Information Asymmetry." *Journal of Accounting and Economics* 37, no. 3: 343–366.

Burgoon, J., W. J. Mayew, J. S. Giboney, et al. 2016. "Which Spoken Language Markers Identify Deception in High-Stakes Settings? Evidence From Earnings Conference Calls." *Journal of Language and Social Psychology* 35, no. 2: 123–157.

Bushee, B. J. 1998. "The Influence of Institutional Investors on Myopic R&D Investment Behavior." *Accounting Review* 73, no. 3: 305–333.

Capital One. 2023. "Boost Python Models and Pipelines With Rubicon-ML." Capital One. https://www.capitalone.com/tech/open-source/boost-python-model-development-with-rubicon/.

Carpenter, M. A., M. A. Geletkanycz, and W. G. Sanders. 2004. "Upper Echelons Research Revisited: Antecedents, Elements, and Consequences of Top Management Team Composition." *Journal of Management* 30, no. 6: 749–778.

Cetindamar, D., and O. Pala. 2011. "Chief Technology Officer Roles and Performance." *Technology Analysis & Strategic Management* 23, no. 10: 1031–1046.

Chae, Y., and T. Davidson. 2025. "Large Language Models for Text Classification: From Zero-Shot Learning to Fine-Tuning." *Sociological Methods & Research*. https://doi.org/10.1177/00491241251325243.

Chung, D., and M. Kang. 2019. "Characteristics of Chief Technology Officers and Radical Innovation." *International Journal of Innovation and Technology Management* 16, no. 3: 1950026.

Cyert, R. M., and J. G. March. 1963. *A Behavioral Theory of the Firm*. Prentice-Hall.

Dastin, J. 2018. "Insight – Amazon Scraps Secret AI Recruiting Tool That Showed Bias Against Women." Reuters. https://www.reuters.com/article/world/insight-amazon-scraps-secret-ai-recruiting-tool-that-showed-bias-against-women-idUSKCN1MK0AG/.

Davenport, T. H., and R. Bean. 2023. "AI Ethics at Unilever: From Policy to Process." *MIT Sloan Management Review*. https://sloanreview.mit.edu/article/ai-ethics-at-unilever-from-policy-to-process/.

Díaz-Rodríguez, N., J. Del Ser, M. Coeckelbergh, M. L. de Prado, E. Herrera-Viedma, and F. Herrera. 2023. "Connecting the Dots in Trustworthy Artificial Intelligence: From AI Principles, Ethics, and Key Requirements to Responsible AI Systems and Regulation." *Information Fusion* 99: 101896.

Dyer, T., M. Lang, and L. Stice-Lawrence. 2017. "The Evolution of 10-K Textual Disclosure: Evidence From Latent Dirichlet Allocation." *Journal of Accounting and Economics* 64, no. 2–3: 221–245.

Eggers, J. P., and S. Kaplan. 2009. "Cognition and Renewal: Comparing CEO and Organizational Effects on Incumbent Adaptation to Technical Change." *Organization Science* 20, no. 2: 461–477.

Eisenberg, I. 2025. "Introducing Credo AI Labs: Where We're Building the Future of AI Governance." Credo AI. https://www.credo.ai/blog/introducing-credo-ai-labs-where-were-building-the-future-of-ai-governance.

Eitel-Porter, R., M. Corcoran, and P. Connolly. 2021. "From Principles to Practice." Accenture. https://www.accenture.com/content/dam/accenture/final/a-com-migration/pdf/pdf-149/accenture-responsible-ai-final.pdf.

Eitel-Porter, R., and U. Grosskopf. 2022. "From AI Compliance to Competitive Advantage." Accenture. https://www.accenture.com/content/dam/accenture/final/a-com-migration/r3-3/pdf/pdf-179/accenture-responsible-by-design-report.pdf.

Eklund, J. C., and M. J. Mannor. 2020. "Keep Your Eye on the Ball or on the Field? Exploring the Performance Implications of Executive Strategic Attention." *Academy of Management Journal* 64, no. 6: 1685–1713.

European Commission. 2025. "AI Act." European Commission. https://digital-strategy.ec.europa.eu/en/policies/regulatory-framework-ai.

Felsenstein, D., and R. Bar-El. 1989. "Measuring the Technological Intensity of the Industrial Sector: A Methodological and Empirical Approach." *Research Policy* 18, no. 4: 239–252.

Floridi, L., J. Cowls, M. Beltrametti, et al. 2018. "AI4People—An Ethical Framework for a Good AI Society: Opportunities, Risks, Principles, and Recommendations." *Minds and Machines* 28: 689–707.

Forrester. 2024. "Global Tech Market Forecast, 2023 to 2027." Forrester. https://www.forrester.com/report/global-tech-market-forecast-2023-to-2027/RES180399.

Fu, X., W. Wu, and Z. Zhang. 2021. "The Information Role of Earnings Conference Call Tone: Evidence From Stock Price Crash Risk." *Journal of Business Ethics* 173: 643–660.

Gama, F., and S. Magistretti. 2025. "Artificial Intelligence in Innovation Management: A Review of Innovation Capabilities and a Taxonomy of AI Applications." *Journal of Product Innovation Management* 42, no. 1: 76–111.

Gaspar, J.-M., M. Massa, and P. Matos. 2005. "Shareholder Investment Horizons and the Market for Corporate Control." *Journal of Financial Economics* 76, no. 1: 135–165.

Gaspar, J.-M., M. Massa, P. Matos, R. Patgiri, and Z. Rehman. 2013. "Payout Policy Choices and Shareholder Investment Horizons." *Review of Finance* 17, no. 1: 261–320.

Gerlach, J., R. M. Stock, and P. Buxmann. 2014. "Never Forget Where You're Coming From: The Role of Existing Products in Adoptions of Substituting Technologies." *Journal of Product Innovation Management* 31: 133–145.

Giannetti, M., and X. Yu. 2021. "Adapting to Radical Change: The Benefits of Short-Horizon Investors." *Management Science* 67, no. 7: 4032–4055.

Gidwani, S., and A. Bello. 2024. “Using Generative AI to Accelerate Product Innovation.” IBM. https://www.ibm.com/think/topics/generative-ai-product-development.

Glikson, E., and A. W. Woolley. 2020. “Human Trust in Artificial Intelligence: Review of Empirical Research.” *Academy of Management Annals* 14, no. 2: 627–660.

Gnewuch, U., S. Morana, O. Hinz, R. Kellner, and A. Maedche. 2024. “More Than a Bot? The Impact of Disclosing Human Involvement on Customer Interactions With Hybrid Service Agents.” *Information Systems Research* 35, no. 3: 936–955.

Google. 2025a. “AI Principles.” Google. https://ai.google/principles/.

Google. 2025b. “Applying AI to Help Solve Society's Biggest Challenges and Improve Lives.” Google AI. https://ai.google/societal-impact/.

Google. 2025c. “Responsible AI Progress Report.” Google. https://ai.google/static/documents/ai-responsibility-update-published-february-2025.pdf.

Gu, Y., L. Wang, and Y. Mai. 2024. “Do You Need a Chief Technology Officer? The Effect of Appointing a CTO on Innovation Activities in High-Tech New Ventures.” *Journal of Product Innovation Management* 41, no. 6: 1118–1140.

Gutierrez, L., I. Montiel, J. A. Surroca, and J. A. Tribo. 2022. “Rainbow Wash or Rainbow Revolution? Dynamic Stakeholder Engagement for SDG-Driven Responsible Innovation.” *Journal of Business Ethics* 180, no. 4: 1113–1136.

Hagedoorn, J., and M. Cloodt. 2003. “Measuring Innovative Performance: Is There an Advantage in Using Multiple Indicators?” *Research Policy* 32, no. 8: 1365–1379.

Hainmueller, J. 2012. “Entropy Balancing for Causal Effects: A Multivariate Reweighting Method to Produce Balanced Samples in Observational Studies.” *Political Analysis* 20, no. 1: 25–46.

Hambrick, D. C., and P. A. Mason. 1984. “Upper Echelons: The Organization as a Reflection of Its Top Managers.” *Academy of Management Review* 9, no. 2: 193–206.

He, J., and X. Tian. 2013. “The Dark Side of Analyst Coverage: The Case of Innovation.” *Journal of Financial Economics* 109, no. 3: 856–878.

Heckman, J. J. 1979. “Sample Selection Bias as a Specification Error.” *Econometrica: Journal of the Econometric Society* 47: 153–161.

Heikkilä, M. 2022. “The Walls Are Closing in on Clearview AI.” MIT Technology Review. https://www.technologyreview.com/2022/05/24/1052653/clearview-ai-data-privacy-uk/.

Hern, A. 2022. “Microsoft Limits Access to Facial Recognition Tool in AI Ethics Overhaul.” The Guardian. https://www.theguardian.com/technology/2022/jun/22/microsoft-limits-access-to-facial-recognition-tool-in-ai-ethics-overhaul.

Homburg, C., M. Theel, and S. Hohenberg. 2020. “Marketing Excellence: Nature, Measurement, and Investor Valuations.” *Journal of Marketing* 84, no. 4: 1–22.

Huang, M.-H., and R. T. Rust. 2021. “A Strategic Framework for Artificial Intelligence in Marketing.” *Journal of the Academy of Marketing Science* 49: 30–50.

Huang, M.-H., and R. T. Rust. 2024. “The Caring Machine: Feeling AI for Customer Care.” *Journal of Marketing* 88, no. 5: 1–23.

Humphreys, A., and R. J. Wang. 2018. “Automated Text Analysis for Consumer Research.” *Journal of Consumer Research* 44, no. 6: 1274–1306.

Hutson, T., and N. Crampton. 2025. “Responsible AI Transparency Report 2025.” Microsoft. https://cdn-dynmedia-1.microsoft.com/is/content/microsoftcorp/microsoft/msc/documents/presentations/CSR/Responsible-AI-Transparency-Report-2025-vertical.pdf.

IBM. 2019. “Introducing AI Explainability 360.” IBM Research Blog. https://research.ibm.com/blog/ai-explainability-360.

H&M Group. 2021. “Responsible AI Is Better AI.” H&M Group Stories. https://hmgroup.com/our-stories/responsible-ai-is-better-ai/.

IBM. 2025. “AI Governance Tools and Solutions.” IBM. https://www.ibm.com/solutions/ai-governance.

Joseph, J., D. Laureiro-Martinez, A. Nigam, W. Ocasio, and C. Rerup. 2024. “Research Frontiers on the Attention-Based View of the Firm.” *Strategic Organization* 22, no. 1: 6–17.

JPMorgan Chase. 2023. “AI and Model Risk Governance.” JPMorgan Chase. https://www.jpmorganchase.com/about/technology/news/ai-and-model-risk-governance.

Jung, M. J., F. M. H. Wong, and F. X. Zhang. 2018. “Buy-Side Analysts and Earnings Conference Calls.” *Journal of Accounting Research* 56, no. 3: 913–952.

Kammerlander, N., and M. Ganter. 2015. “An Attention-Based View of Family Firm Adaptation to Discontinuous Technological Change: Exploring the Role of Family Ceos' Noneconomic Goals.” *Journal of Product Innovation Management* 32, no. 3: 361–383.

Kaplan, A., and M. Haenlein. 2019. “Siri, Siri, in My Hand: Who's the Fairest in the Land? On the Interpretations, Illustrations, and Implications of Artificial Intelligence.” *Business Horizons* 62, no. 1: 15–25.

Kim, H.-D., K. Park, and K. R. Song. 2019. “Do Long-Term Institutional Investors Foster Corporate Innovation?” *Accounting and Finance* 59, no. 2: 1163–1195.

Kogan, L., D. Papanikolaou, A. Seru, and N. Stoffman. 2017. “Technological Innovation, Resource Allocation, and Growth.” *Quarterly Journal of Economics* 132, no. 2: 665–712.

Lepak, D. P., R. Takeuchi, and S. A. Snell. 2003. “Employment Flexibility and Firm Performance: Examining the Interaction Effects of Employment Mode, Environmental Dynamism, and Technological Intensity.” *Journal of Management* 29, no. 5: 681–703.

Lerner, J., M. Sorensen, and P. Strömberg. 2011. “Private Equity and Long-Run Investment: The Case of Innovation.” *Journal of Finance* 66, no. 2: 445–477.

Li, J., M. Li, X. Wang, and J. B. Thatcher. 2021. “Strategic Directions for AI: The Role of CIOs and Boards of Directors.” *MIS Quarterly* 3: 1603–1644.

Li, K., F. Mai, R. Shen, and X. Yan. 2021. “Measuring Corporate Culture Using Machine Learning.” *Review of Financial Studies* 34, no. 7: 3265–3315.

Li, Q., P. G. Maggitti, K. G. Smith, P. E. Tesluk, and R. Katila. 2013. “Top Management Attention to Innovation: The Role of Search Selection and Intensity in New Product Introductions.” *Academy of Management Journal* 56, no. 3: 893–916.

Liu, J., H. Chang, J. Y.-L. Forrest, and B. Yang. 2020. “Influence of Artificial Intelligence on Technological Innovation: Evidence From the Panel Data of China's Manufacturing Sectors.” *Technological Forecasting and Social Change* 158: 120142.

Lombana Díaz, C., and S. Welsch. 2025. *SAP AI Ethics Handbook: Applying the SAP Global AI Ethics Policy Across the SAP Business AI Lifecycle, Development Lifecycle Chapter External*. SAP. https://dam.sap.com/mac/app/e/pdf/preview/embed/kAaZG3F?ltr=a&rc=10&doi=SAP951092.

Lu, H., X. Liu, and O. Osiyevskyy. 2023. “Doing Safe While Doing Good: Slack, Risk Management Capabilities, and the Reliability of Value Creation Through CSR.” *Strategic Organization* 21, no. 4: 874–904.

Mann, R. J., and T. W. Sager. 2007. “Patents, Venture Capital, and Software Start-Ups.” *Research Policy* 36, no. 2: 193–208.

March, J. G. 1991. "Exploration and Exploitation in Organizational Learning." *Organization Science* 2, no. 1: 71–87.

Mariani, M. M., I. Machado, V. Magrelli, and Y. K. Dwivedi. 2023. "Artificial Intelligence in Innovation Research: A Systematic Review, Conceptual Framework, and Future Research Directions." *Technovation* 122: 102623.

Martin, A., and T. Kushwaha. 2024. "Can Words Speak Louder Than Actions? Using Top Management Teams' Language to Predict Myopic Marketing Spending." *Journal of Marketing* 88, no. 6: 140–161.

Mendonça, S. 2009. "Brave Old World: Accounting for 'High-Tech' Knowledge in 'Low-Tech' Industries." *Research Policy* 38, no. 3: 470–482.

Meta. 2022. "Introducing Privacy Center." Meta. https://about.fb.com/news/2022/01/introducing-privacy-center/.

Microsoft. 2023. "Microsoft's AI Safety Policies." Microsoft. https://blogs.microsoft.com/on-the-issues/2023/10/26/microsofts-ai-safety-policies/.

Microsoft. 2025a. "AI for Health." Microsoft Research. https://www.microsoft.com/en-us/research/project/ai-for-health/.

Microsoft. 2025b. "Responsible AI." Microsoft. https://www.microsoft.com/en-us/ai/responsible-ai.

Mikalef, P., K. Conboy, J. E. Lundström, and A. Popovič. 2022. "Thinking Responsibly About Responsible AI and 'The Dark Side' of AI." *European Journal of Information Systems* 31, no. 3: 257–268.

Miller, B. P., A. G. Sheneman, and B. M. Williams. 2022. "The Impact of Control Systems on Corporate Innovation." *Contemporary Accounting Research* 39, no. 2: 1425–1454.

Nadkarni, S., and P. S. Barr. 2008. "Environmental Context, Managerial Cognition, and Strategic Action: An Integrated View." *Strategic Management Journal* 29, no. 13: 1395–1427.

Nanda, R., and M. Rhodes-Kropf. 2013. "Investment Cycles and Startup Innovation." *Journal of Financial Economics* 110, no. 2: 403–418.

NIST. 2023. "Artificial Intelligence Risk Management Framework (AI RMF 1.0)." National Institute of Standards and Technology. https://nvlpubs.nist.gov/nistpubs/ai/NIST.AI.100-1.pdf.

NVIDIA. 2025. "NVIDIA NeMo Framework User Guide." NVIDIA. https://docs.nvidia.com/nemo-framework/user-guide/25.07/nemotoolkit/asr/models.html.

Ocasio, W. 1997. "Towards an Attention-Based View of the Firm." *Strategic Management Journal* 18: 187–206.

Ocasio, W. 2011. "Attention to Attention." *Organization Science* 22, no. 5: 1286–1296.

Owen-Jackson, C. 2025. "How Red Teaming Helps Safeguard the Infrastructure Behind AI Models." IBM. https://www.ibm.com/think/insights/how-red-teaming-helps-safeguard-the-infrastructure-behind-ai-models.

Padigar, M., L. Pupovac, A. Sinha, and R. Srivastava. 2022. "The Effect of Marketing Department Power on Investor Responses to Announcements of AI-Embedded New Product Innovations." *Journal of the Academy of Marketing Science* 50, no. 6: 1277–1298.

Palmon, D., Y. Chen, and B. Chen. 2024. "Corporate Social Responsibility and Information Asymmetry: Do Earnings Conference Calls Play a Role?" *Journal of Business Ethics* 194, no. 1: 77–101.

Patel, J. 2025. "Protecting AI So AI Can Improve the World, Safely." Cisco Blogs. https://blogs.cisco.com/news/you-cant-sacrifice-ai-safety-for-ai-speed.

PwC. 2025. "Responsible AI Toolkit." PwC. https://www.pwc.com/sg/en/services/reimagine-digital/data-optimisation/what-is-responsible-ai.html.

Raisch, S., and S. Krakowski. 2021. "Artificial Intelligence and Management: The Automation–Augmentation Paradox." *Academy of Management Review* 46, no. 1: 192–210.

Rammer, C., G. P. Fernández, and D. Czarnitzki. 2022. "Artificial Intelligence and Industrial Innovation: Evidence From German Firm-Level Data." *Research Policy* 51, no. 7: 104555.

Repenning, N. P., and J. D. Sterman. 2002. "Capability Traps and Self-Confirming Attribution Errors in the Dynamics of Process Improvement." *Administrative Science Quarterly* 47, no. 2: 265–295.

Rhee, L. 2024. "CEO Attentional Vigilance: Behavioral Implications for the Pursuit of Exploration." *Academy of Management Journal* 67, no. 6: 1463–1487.

Ridge, J. W., S. Johnson, A. D. Hill, and J. Bolton. 2017. "The Role of Top Management Team Attention in New Product Introductions." *Journal of Business Research* 70: 17–24.

Ringberg, T., M. Reihlen, and P. Rydén. 2019. "The Technology-Mindset Interactions: Leading to Incremental, Radical or Revolutionary Innovations." *Industrial Marketing Management* 79: 102–113.

Salesforce. 2025. "Trusted AI: The Einstein Trust Layer." Salesforce. https://www.salesforce.com/uk/artificial-intelligence/trusted-ai/.

Santamaría, L., M. J. Nieto, and A. Barge-Gil. 2009. "Beyond Formal R&D: Taking Advantage of Other Sources of Innovation in Low- and Medium-Technology Industries." *Research Policy* 38, no. 3: 507–517.

Shneiderman, B. 2021. "Responsible AI: Bridging From Ethics to Practice." *Communications of the ACM* 64, no. 8: 32–35.

Simon, H. A. 1947. "A Comment on 'the Science of Public Administration'." *Public Administration Review* 7, no. 3: 200–203.

Sirmon, D. G., and M. A. Hitt. 2003. "Managing Resources: Linking Unique Resources, Management, and Wealth Creation in Family Firms." *Entrepreneurship Theory and Practice* 27, no. 4: 339–358.

Spanjol, J., C. H. Noble, M. Baer, et al. 2024. "Fueling Innovation Management Research: Future Directions and Five Forward-Looking Paths." *Journal of Product Innovation Management* 41, no. 5: 893–948.

Stilgoe, J., R. Owen, and P. Macnaghten. 2013. "Developing a Framework for Responsible Innovation." *Research Policy* 42, no. 9: 347–359.

Stoffman, N., M. Woeppel, and D. M. Yavuz. 2022. "Small Innovators: No Risk, No Return." *Journal of Accounting and Economics* 74, no. 1: 101492.

Teece, D. J. 2007. "Explicating Dynamic Capabilities: The Nature and Microfoundations of (Sustainable) Enterprise Performance." *Strategic Management Journal* 28, no. 13: 1319–1350.

Telefónica. 2025. "Global Transparency Center." Telefónica. https://www.telefonica.com/en/global-transparency-center/.

Thomaz, F., C. Salge, E. Karahanna, et al. 2020. "Learning from the Dark Web: Leveraging Conversational Agents in the Era of Hyper-Privacy to Enhance Marketing." *Journal of the Academy of Marketing Science* 48: 43–63.

Tomašev, N., J. Cornebise, F. Hutter, et al. 2020. "AI for Social Good: Unlocking the Opportunity for Positive Impact." *Nature Communications* 11, no. 2468.

Verganti, R., L. Vendraminelli, and M. Iansiti. 2020. "Innovation and Design in the Age of Artificial Intelligence." *Journal of Product Innovation Management* 37, no. 3: 212–227.

Wang, D., H. Zhang, R. Liu, and W. Lv. 2012. "Feature Selection Based on Term Frequency and T-Test for Text Categorization." In *Proceedings of the 21st ACM International Conference on Information and Knowledge Management*, 1482–1486.

Wu, L.-Z., H. K. Kwan, F. H.-k. Yim, R. K. Chiu, and X. He. 2015. "CEO Ethical Leadership and Corporate Social Responsibility: A Moderated Mediation Model." *Journal of Business Ethics* 130, no. 4: 819–831.

Wu, W., and S. Liu. 2023. “Compliance Costs of AI Technology Commercialization: A Field Deployment Perspective.” arXiv Preprint arXiv:2301.13454.

Yadav, M. S., J. C. Prabhu, and R. K. Chandy. 2007. “Managing the Future: CEO Attention and Innovation Outcomes.” *Journal of Marketing* 71, no. 4: 84–101.

Yin, R. K. 2018. *Case Study Research and Applications: Design and Methods*. 6th ed. Sage.

Zawislak, P. A., E. M. Fracasso, and J. Tello-Gamarra. 2018. “Technological Intensity and Innovation Capability in Industrial Firms.” *Innovation & Management Review* 15, no. 2: 189–207.

Zhang, Y., R. Jin, and Z.-H. Zhou. 2010. “Understanding Bag-of-Words Model: A Statistical Framework.” *International Journal of Machine Learning and Cybernetics* 1: 43–52.

Supporting Information

Additional supporting information can be found online in the Supporting Information section. **Appendix S1:** jpim70015-sup-0001-AppendixS1.docx.

Biographies

Mengran Xiong, PhD, is a Lecturer in Marketing at Westminster Business School, University of Westminster. She holds a PhD in Marketing from the University of Sheffield. Her research examines responsible artificial intelligence (AI) and digital innovation, focusing on how digital technologies reshape marketing and healthcare practices. Her work has been published in *Information Systems Frontiers* and in the proceedings of the *Hawaii International Conference on System Sciences* (*HICSS*).

Haofeng Xu, PhD, is a Lecturer (Assistant Professor) in Accounting and Finance at Loughborough Business School, Loughborough University. Prior to this, he was a Lecturer in Finance at the University of Aberdeen Business School. His research interests lie primarily in the areas of AI disruptions in business, economics, and finance, FinTech, corporate finance, and social media disclosure analytics. He has published his work in journals such as the *Journal of Empirical Finance*, *European Financial Management*, and the *Journal of International Financial Markets, Institutions and Money*.

Jiao Ji, PhD, is a Lecturer (Assistant Professor) in Finance at the Sheffield University Management School. Her research explores empirical corporate finance and governance, examining corporate risk, misconduct, disclosure, and ESG. She applies advanced textual analytics to corporate filings and social-media data, and also studies banking, innovation, and behavioral finance. Her work appears in leading outlets, including the *Journal of the Association for Information Systems*, *European Financial Management*, and the *Journal of Empirical Finance*.

Renxian Zuo is a PhD candidate in Management Science and Engineering at the School of Information Management, Wuhan University. His research interests lie in technology innovation, technology evaluation and prediction. He applies advanced textual analytics to science, patent, and corporate data. His work has been published in *Physica A*, *Journal of Cleaner Production*, *Electronic Commerce Research*, *Operational Research*, and in the proceedings of the *European Conference on Information Systems* (*ECIS*).

Yichuan Wang, PhD, is an Associate Professor in Digital Marketing at the Sheffield University Management School, University of Sheffield. He earned his PhD in Business and Information Systems from the Raymond J. Harbert College of Business, Auburn University (USA). His research examines how digital technologies and information systems shape and transform contemporary business practices. His work has been published in leading journals, including the *Journal of Operations Management*, *Journal of the Association for Information Systems*, *Social Science & Medicine*, *International Journal of Operations and Production Management*, *British Journal of Management*, and *Information & Management*. He currently serves as a Senior Editor for The DATA BASE for Advances in Information Systems. Since 2018, he has co-founded and organized the International Conference on Digital Health and Medical Analytics (DHA). He has also led and collaborated on numerous international healthcare projects that harness information systems and business analytics to enhance patient care quality.

Hesam Olya is a Full Professor and Head of Marketing and Cultural & Creative Industries, as well as Director of Accreditation at Sheffield University Management School, Sheffield, UK. His research focuses on digital sustainability, experiential consumption, and cultural marketing. He has extensive editorial experience, serving as Associate Editor of the *International Journal of Consumer Studies*, and has guest-edited several special issues for leading academic journals, including the *Journal of Sustainable Tourism* and *Information Systems Frontiers*, among others. His recent book is *Toward Responsible Service Management: AI and Digital Transformation in Action*.