NAVIGATING AI: AN INVESTOR FRAMEWORK FOR ASSESSING EXPOSURE AND GOVERNANCE

Railpen's AI Governance Framework and how we have applied it in practice

RAILPEN

September 2026

CONTENTS

AUTHORS

Authors

This report has been prepared by **Dr Rory Sullivan**, **Cora Buentjen** and **Josie Zenger** (of Chronos Sustainability) and **Caroline Escott**, **Sophie Harris** and **Jasmine Porter** (of Railpen)

RAILPEN

Railpen's purpose is simple – to secure our members' future.

We are responsible for over £34bn in assets on behalf of more than 350,000 members, providing market-leading pension administration and investment services for defined benefit (DB), defined contribution (DC) and hybrid schemes, including the Railways Pension Scheme – one of the UK's largest, longest established, and intricate funds.

We are built on our heritage. Our history of helping everyday, hardworking people save for their retirement stretches back more than half a century. From our beginnings in 1965 as the pensions office for the British Rail Pension Scheme, we've innovated and evolved to serve the needs of the industry and the financial future of its workers.

We are committed to better member outcomes. Our role is delegated to us by our Trustee, whose unwavering passion to put members' interests first is integral to everything we do – from how we administer the schemes to how we select, structure and size investments to protect and grow the long-term value of members' savings.

We are agile and innovative. Guided by integrity, the importance we place on collaboration, and our member-first approach, we work with policymakers and industry leaders to advocate for a progressive and supportive policy environment that helps us to secure our members' future and shape a more sustainable pensions landscape.

We are different from the norm. We put every pound of profit towards paying members' pensions securely, affordably and sustainably. We make this our mission. Our long-term mindset opens up investment opportunities over multiple decades that enrich communities and the environment, set new standards for innovation and contribute to a resilient UK economy.

We are Railpen.

Find out more at www.railpen.com

CHRONOS INTELLIGENT SUSTAINABILITY

Chronos Sustainability aims to drive positive impact at scale.

We are a specialist sustainability consultancy with the fundamental objective of delivering transformative, systemic change in the social and environmental performance of key industry sectors.

Chronos Sustainability has deep expertise in responsible investment, corporate sustainability and in stakeholder analysis and engagement.

Find out more about Chronos Sustainability: www.chronossustainability.com

EXECUTIVE SUMMARY

Building on Railpen's 2025 report Achieving Effective AI Governance, this report reintroduces our AI Governance Framework (AIGF), now translated into a practical benchmarking tool to assess the quality of governance practices across portfolio companies.

While initially applied to Railpen's Active Equity portfolio[1], the AIGF is designed to be adaptable for broader use, sector-agnostic, and applicable for both developers and deployers. Across the large publicly listed companies addressed in this report and those beyond, we recognise that implementation of the AIGF will vary according to AI exposure and organisational maturity. Therefore, we would encourage a phased and deliberate approach to integrating the framework's pillars over time, as companies advance in their AI governance journey.

Artificial intelligence (AI), in convergence with technologies such as robotics and quantum computing, is rapidly transforming the way in which companies in all sectors are operating[2].

A recent McKinsey study suggests that nearly 90% of large enterprises now use AI in some form, driven by opportunities such as innovation, cost savings and efficiency gains (Figure 1)[3].

Concurrently, AI has created and expanded risks for companies, including non-compliance with AI regulations, biased outputs, and more efficient (and effective) cybersecurity attacks[4]. AI adoption can also expose investment portfolios to systemic risks arising from various interdependencies such as the concentration of digital risk among a small number of technology and cloud providers, cascading supply chain failures and a reliance on shared infrastructure, including data and energy networks[5].

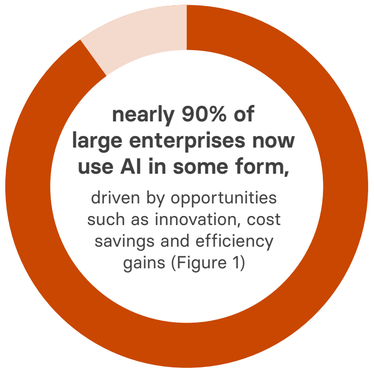
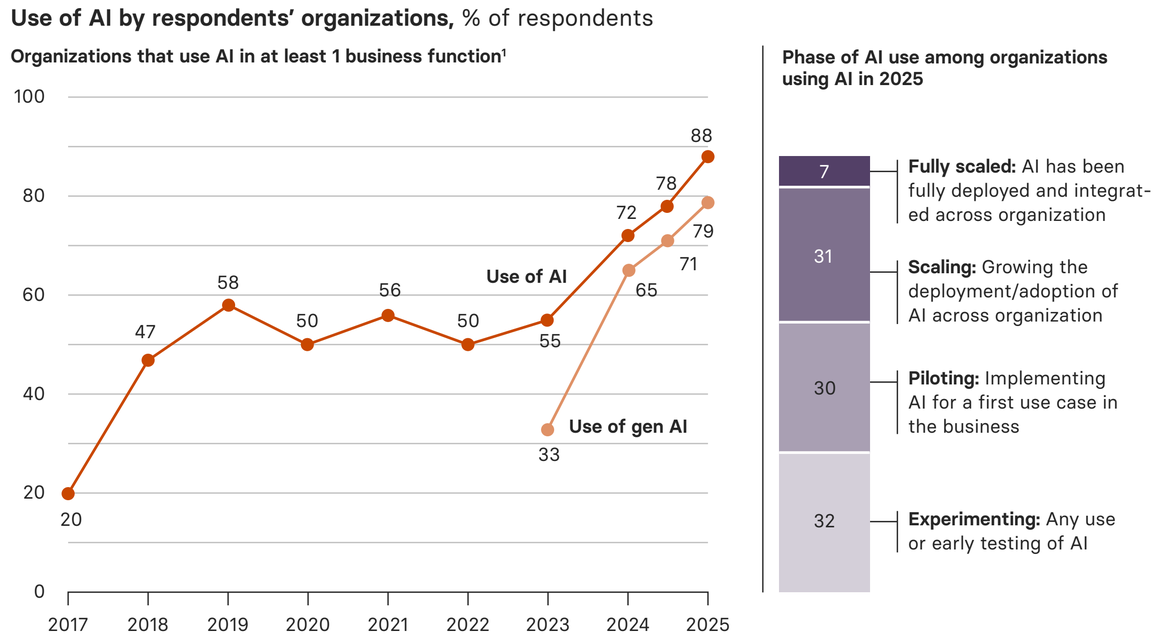


Figure 1: Trends in the use of AI[6]



[1]In 2017, the definition for AI use was using AI in a core part of the organization's business or at scale. In 2018-19, the definition was embedding at least 1 AI capability in business processes or products. From 2020, the definition was that the organization had adopted AI in at least 1 function, and in 2025, the definition was regular use of AI in at least 1 function.
Source: McKinsey Global Surveys on the state of AI, 2017-25

For investors, this creates an urgent governance challenge. Exposure to AI-related risks is increasing faster than companies' governance and risk management practices are maturing.

While many companies are adopting AI across their core business functions, comparatively few provide robust disclosures explaining how AI-related risks are identified, governed and managed in practice. At the same time, investors increasingly recognise that they need to understand their portfolios' exposure to AI, alongside how effectively investee companies manage associated opportunities and risks.

Railpen recognises the significance of AI as an investment theme over the coming decade, with the potential to drive innovation and long-term value creation across our portfolio.

Working with colleagues across the business, Railpen actively assesses how companies and assets might benefit from, or be threatened by, AI-driven disruption. From the infrastructure needed to support the AI economy, to companies developing and deploying AI-enabled products, we see significant opportunities[7]. We believe that effective governance will enable companies to both harness these potential benefits and mitigate risks.

Applying the AI Governance Framework (AIGF) to Railpen's Active Equity portfolio provides five key insights:

1. AI governance maturity is uneven and often underdeveloped relative to risk exposure, with few companies publishing detailed information on how they assess and manage AI risks.
2. Examples of best practice AI policies are beginning to emerge, particularly among AI developers.
3. Many companies signal their intent through reference to responsible AI principles, but few provide evidence of implementation.
4. AI-related risks identified by companies are inconsistent within sectors and focused on established risk types, such as cybersecurity[8].
5. Leading companies differentiate themselves through greater external engagement with users and transparent reporting on performance against AI policies.

Investors have an important role to play in driving progress on AI governance. The AIGF is intended to support investors in prioritising engagement, benchmarking governance maturity, identifying gaps in disclosure and risk management, and tracking progress over time.

Ultimately, managing the risks posed by AI requires a system-wide approach. Therefore, wider adoption of consistent governance expectations and assessment approaches can help reinforce stronger standards of accountability, transparency and responsible AI governance across markets.

Purpose of this report

This report is intended for investors and companies seeking to better understand and govern AI-related risks and opportunities. For investors, it offers a practical tool to benchmark governance quality, prioritise engagement, and support long-term value creation. For companies, it provides insight into emerging investor expectations and stewardship objectives. By operationalising the principles set out in Railpen's 2025 report, **Achieving Effective AI Governance**, the framework is designed to encourage a more consistent, collective approach to AI governance across sectors.

Glossary

Artificial Intelligence (AI): Technology that enables machines to perform tasks typically requiring human intelligence, such as recognising patterns, making decisions, or generating content.

Deployers: Companies that integrate AI models built by third parties into their own products, services, or internal processes.

Developers: Companies that build and train AI models, either for their own use or for use by third parties.

Exposure: The extent to which a company's financial performance, operations, or reputation may be affected by AI-related risks or opportunities.

Systemic risk: The risk that failure of an AI system, or widespread adoption of flawed AI practices, could trigger cascading disruption across a sector or the broader economy.

Model: A computational system trained on data to perform specific tasks, such as generating text or making predictions.

ASSESSING EXPOSURE

Understanding AI exposure across portfolios is an important element in assessing how individual companies govern AI-related risks and opportunities.

Taken together, these dimensions help investors form a view of the overall risk profile, considering both the scale of potential exposures and the extent to which these may be mitigated through governance and risk management.

In practice, **an assessment of exposure can also serve as a useful basis for prioritisation**, helping to identify those companies where more detailed assessment using the AIGF, or targeted engagement, may be most valuable.

Assessing AI exposure requires consideration of a range of factors. Indicators such as systemic importance - including a company's potential to be affected by, or have impacts beyond, its own operations - and broader sectoral risk exposure can provide useful starting points, but are unlikely to capture the full picture in isolation. The extent and nature of AI exposure will often depend on company-specific factors, including the foundation models used, the use cases for which AI is deployed, and any history of AI-related incidents or controversies.

Considering these aspects together can help build a rounded view of where AI may be more significant. The following section outlines a set of lenses that investors may find helpful when forming that view, moving from system-level to company-specific factors:

1. **Systemic risk exposure**
2. **Sector-level exposure**
3. **Company-specific exposure**

1. Systemic risk exposure

Certain companies – including hyperscale cloud providers and foundation model developers – play a critical role in the AI value chain (Table 1). As enablers of innovation across the economy, these companies may be positioned to capture disproportionate value from the advancement of AI and its convergence with other emerging technologies, such as robotics and quantum computing[9]. Equally, **where these firms experience failures, disruptions or governance shortcomings, impacts can propagate across sectors and portfolios**. Indicators of systemic exposure include:

- Lack of substitutability
- Interconnectedness across multiple sectors
- Ability to set de facto norms (i.e. shaping industry standards and practices through widespread adoption of their technologies or approaches)
- Asymmetric information advantages over users and regulators[10]

Table 1: Examples of systemic risk exposure

Company type	Description	Potential systemic AI risks
 **Hyperscale cloud and compute providers**	Large-scale cloud computing platforms which increasingly function as critical economic infrastructure.	• Operational outages cause cross-sector disruption • Pricing or access changes result in systemic cost inflation
 **Foundation model developers and AI platform owners**	Companies developing and controlling general purpose or frontier AI models.	• Correlated decision-making errors across firms using the same model
 **AI-dependent critical services**	Critical services that rely heavily on AI (e.g. finance, healthcare, utilities, transport).	• Technical failures lead directly to loss or compromise of essential services
 **Large digital platforms embedding AI into social and economic systems**	Firms where AI shapes information flow and influences behaviour, which is tightly coupled with platform dominance.	• Copying or imitation of market practices • Discrimination at scale against vulnerable populations or social groups with particular characteristics

2. Sector-level exposure

AI's impacts extend across all sectors, even where the nature and scale of exposure differ. **Emerging research on AI identifies the sectors in Table 2 as having particularly significant exposure**[11].

Table 2: Key sectors exposed to AI

Sector	Exposure to AI	Potentially Significant AI Risks[12]
Financials	Highly exposed due to reliance on AI in credit scoring, fraud detection, trading, and customer onboarding.	• Bias/fairness failures (discriminatory lending) • Model drift (mispriced risk) • Misuse/impersonation (fraud) • Data breaches
Healthcare	AI use includes diagnostic imaging for disease identification, predictive analysis for health insurance premiums, and robotics for automated surgery and equipment.	• Hallucinations or incorrect outputs in diagnostics • Data breaches (patient data) • Automation failures in clinical workflows
IT	These organisations develop and deploy AI systems, and embed AI in software products, cloud services and enterprise solutions.	• Hallucinations in deployed tools • Vendor/model outages • Cybersecurity vulnerabilities • Product liability risks
Retail	AI is used for personalised marketing, dynamic pricing, demand forecasting, and supply chain optimisation.	• Bias in recommendation systems • Model drift affecting demand forecasts • Data privacy breaches • Reputational damage from poor customer targeting
Communication Services	AI underpins content generation, moderation, ad targeting and network optimisation across media and telecom platforms.	• Misuse/abuse (disinformation, deepfakes) • Hallucinated content • Data privacy violations • Regulatory and reputational risks

3. Company-specific exposure

While the methods outlined above can support the initial identification and prioritisation of companies, **a fuller understanding of AI-related risk exposure requires a more granular, company-specific lens**. In practice, exposure can vary significantly even within the same sector. For example, two companies in the retail sector could have significantly different exposures depending on the extent to which AI is used to profile or analyse customers (e.g. when determining what products or services they might be interested in, when determining what discounts or terms might be offered to them) or suppliers (e.g. are supplier employee profiles a part of the supplier selection process).

Although some relevant information may be available through public disclosures, **current reporting remains limited and often inconsistent**, potentially reflecting the rapid pace of AI development and commercial sensitivities around disclosing detailed practices. As such, **targeted company engagement will be a key tool in developing a more complete understanding of exposure** (see page 10).

The following factors may be considered for a company-specific exposure assessment:

1. **Model developed and/or deployed**
2. **Use cases**
3. **Incidents**

Model developed and/or deployed

Assessing the types of AI models a company develops and/or deploys provides an important starting point for understanding risk exposure as different models carry inherently different risk profiles. **The following benchmarks seek to evaluate the safety, reliability, and risk characteristics of these models**. These tools can support investors in assessing exposure both to developers of advanced AI systems and to companies deploying such models within their operations.

- AI Luminate Benchmark:
  A set of safety and security benchmarks that assesses generative AI models across 12 hazard categories.
- CAIS AI Dashboard (Risk Index):
  A benchmark that evaluates the extent to which frontier AI models provide harmful or misleading information.
- Helm Safety (Stanford):
  A benchmark that evaluates language models across multiple dimensions including potential harm, bias and misinformation.

Use cases

Complementing a model-based lens, analysing how a company uses AI in practice provides further insight into the nature and materiality of its risk exposure. The same underlying technology may present different risks depending on its application and centrality to a company's business model.

The EU Artificial Intelligence Act (Annex III) identifies a set of **'high-risk' use cases, which are subject to heightened regulatory scrutiny and governance expectations**[13]:

- Biometrics
- Critical infrastructure
- Education and vocational training
- Employment, workers management and access to self-employment
- Access to and enjoyment of essential private services and essential public services and benefits
- Law enforcement
- Migration, asylum and border control management
- Administration of justice and democratic processes

Incidents

A further lens involves assessing how AI-related risks have materialised in practice. **Historical incidents can provide valuable insight into the types of failures and governance gaps associated with different models and use cases**. The following databases can be used by investors to identify trends and areas of vulnerability:

- OECD AI Incidents and Hazards Monitor:
  A database which documents AI incidents and hazards with the aim of helping policymakers, AI practitioners, and other stakeholders to gain insights into the risks and harms associated with AI systems.
- AI Incident Database:
  This database tracks AI-related incidents. Incidents can be sorted by 'Entities', allowing investors to determine the exposure of their investee companies to AI-related incidents.

Example engagement questions to help assess exposure

Models

1. Does the company depend on third-party foundation models?
2. How disruptive would switching providers be to the business?

Use cases

3. Which of the company's core operations are reliant on AI use (e.g. customer service, marketing, underwriting)? How critical are these functions to the operation of the business?
4. What type of data is captured or processed by the company's AI models? Is any sensitive or personal data captured (e.g. health, financial, biometric)?
5. Is the company involved in developing or deploying any of the eight high-risk systems identified above?

Incidents

6. Has the company had any AI-related incidents (e.g. data breaches, model failures, regulatory actions)?
7. What were the consequences, and what actions have been taken to prevent a recurrence? How are you gaining comfort that these actions are likely to be effective?

Assessing portfolio-level exposure

Given the systemic nature of AI risks, investors can consider how particular AI risks may be clustered across their portfolios, rather than only assessing exposure solely at the individual company-level. For example, investors may consider:

- **Vendor concentration:** How many companies in the portfolio rely on the same AI provider or foundation model? Heavy reliance on a small number of vendors may create single points of failure (e.g. outages, model flaws, or pricing changes).
- **Sector clustering:** Are high AI-exposure companies concentrated in specific sectors (e.g. financials, technology, healthcare)? This may increase sensitivity to sector-wide regulatory or technological shocks.
- **Use-case concentration:** Are multiple companies deploying AI in similar high-risk functions (e.g. hiring, customer service, marketing)?
- **Regulatory exposure overlap:** How many portfolio companies are exposed to the same regulatory regimes (e.g. high-risk systems under the EU AI Act)? Changes in regulation could simultaneously affect multiple companies.
- **Data dependency concentration:** Do multiple companies rely on similar types of sensitive data (e.g. personal, biometric, or financial data)? This can increase exposure to correlated data privacy or cybersecurity incidents.

ASSESSING MANAGEMENT QUALITY: THE RAILPEN AI GOVERNANCE FRAMEWORK (AIGF)

Methodology

The AIGF builds on the methodology first introduced in Railpen's 2025 report, Achieving Effective AI Governance, which translated responsible AI principles into actionable practices. **This iteration is designed as a benchmarking tool**, enabling investors to assess and compare the maturity of AI governance practices across developers and deployers over time in a consistent, sector-agnostic way.

This AIGF comprises 18 questions designed to assess the maturity and effectiveness of companies' AI governance practices. These questions are organised around three key pillars – governance and strategy, risk management and performance reporting[14]. Within this structure:

- **Eight questions are designated as "Core," representing the essential components of responsible AI governance that all organisations are expected to address**. These include, for example, maintaining an AI policy aligned with Responsible AI principles, ensuring appropriate leadership oversight, and establishing processes to identify and manage AI-related risks.
- **The remaining ten "Additional" questions provide further depth, recognising companies that have advanced beyond baseline practices**. Three "Unscored" questions are included to capture contextual information that varies by a company's AI use cases and business model. While they do not contribute to the overall score, they help inform interpretation of the results and future engagement.

Using this framework, companies can be evaluated against each indicator and assigned scores that reflect their overall governance maturity (**Table 3**). We recognise that companies may not yet be in a position to meet all expectations set out in the AIGF, as implementation will depend on their AI exposure and organisational maturity.

However, we would encourage a phased and deliberate approach to integrating the framework's pillars over time, as companies advance in their AI governance journey.

Figure 2: Key pillars of Railpen's AIGF

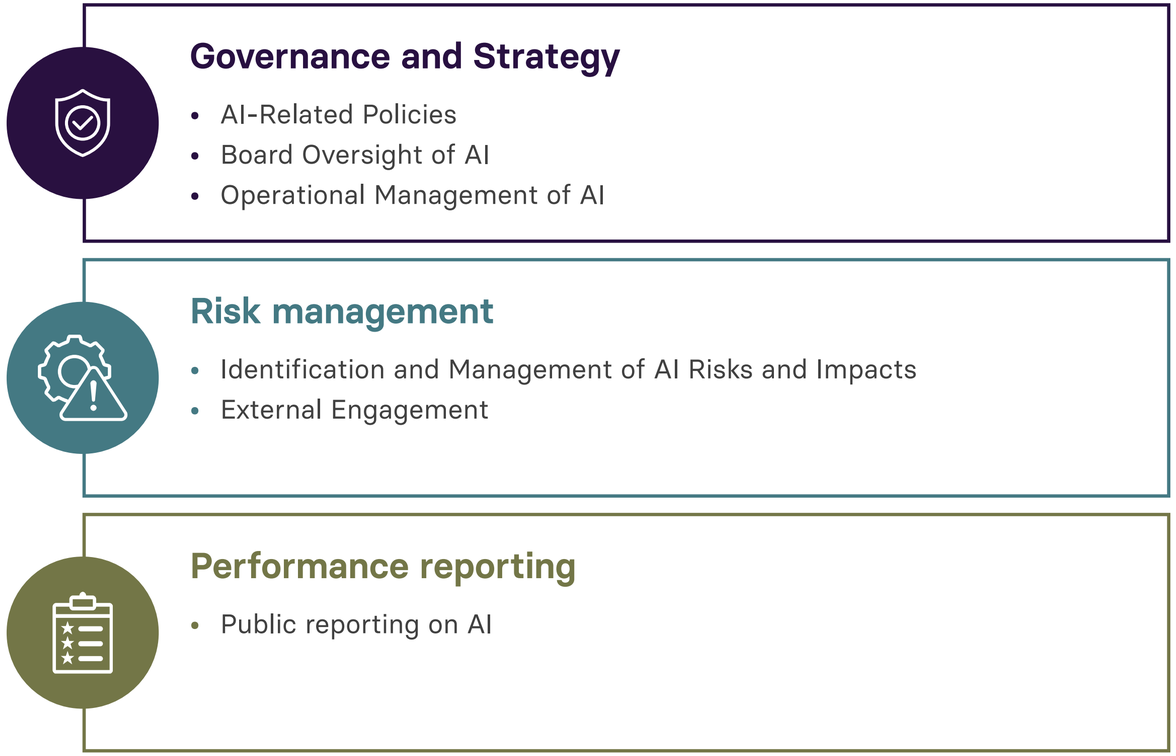


Table 3: Assessment framework with Core Questions bolded

Pillar	Subsection	Maximum Score	Indicator
Governance and Strategy	AI-Related Policies	5	**Does the company describe the relevance of AI to its business strategy?**
		10	**Does the company publish a comprehensive policy regarding its AI use?**
		5	**Does the company's AI policy align with Responsible AI Principles?**
		Unscored	Which principles of responsible AI does the company's policy align with?
	Board Oversight of AI	5	**Does the company describe how the board oversees the operational approach to managing AI?**
		5	Has the company described actions taken in relation to board oversight of the operational approach to managing AI?
	Operational Management of AI	5	**Does the company disclose how it has designated roles or defined responsibility for the operational management of AI?**
		5	Has the company provided employee training on its AI approach and explained what training is provided?
Risk Management	Identification and Management of AI Risks and Impacts	5	**Does the company describe how AI has been incorporated into the risk assessment process?**
		Unscored	What are the AI-related risks that have been identified?
		5	**Does the company explain how it will manage any AI-related risks?**
	External Engagement	5	Does the company declare its use of AI to affected users?
		5	Does the company implement formal feedback processes for its AI systems?
		5	Does the company audit its approach to AI?
		5	Does the company participate in external collaborations (e.g. industry, public policy or civil sector) supporting the responsible adoption of AI?
Performance Reporting	Public Reporting on AI	5	**Does the company report on how it performs against commitments, objectives or principles in its AI policy?**
		5	Does the company disclose whether it has had any AI-related incidents?
		Unscored	Are there external reports on whether the company has had any AI-related incidents?

Key: **Bold** = eight core questions, Normal = additional contextual questions.

Based on these results, **companies can be grouped into five tiers, ranging from "Laggards" to "Best Practice"** (Table 4). This tiered classification allows companies to be assessed relative to their peers and allows areas for improvement to be identified.

Table 4: Description of tier rankings

Tier Scheme (% of total score)	Tier	Description
>85	Best Practice	Best Practice companies have established AI governance practices and have taken additional steps to follow through on their commitments to responsible AI use. Some have begun to publish regular reports on their AI approach.
65-84	Established Practice	Established Practice companies have identified significant AI risks and are generally addressing these through robust policies. They have typically established clear governance and accountability practices for their AI approach and are beginning to address risks through external engagement (e.g. informing users, putting feedback mechanisms in place).
40-64	Emerging Practice	Emerging Practice companies have usually identified significant AI risks and are beginning to put policies and practices in place to address these. These companies typically have begun to establish accountability structures for managing AI risks and, in some cases, are beginning to undertake external engagement practices.
25-39	Nascent Practice	Nascent Practice companies have begun to identify relevant AI risks and, in some cases, have put in place policies and accountability structures to address these risks. They generally have not yet engaged with external stakeholders or started to report on their AI-related risks.
<25	Laggards	Laggards are, predominantly, companies in sectors in which the risks associated with AI are currently perceived as less relevant or companies that have not begun to acknowledge AI risks to their operations.

Assessing Railpen's active equity portfolio against the AIGF

Using the AIGF, Railpen assessed around 30 companies held within its Active Equities Portfolio on their operational use of AI. Each company was evaluated across 18 questions organised under three pillars: governance and strategy, risk management, and performance reporting. Three questions were unscored, serving as qualitative indicators only. The remaining questions carried point values of 5 or 10, with partial scores available where a company partially met the relevant criteria (see Table 3). Final tier placement — from Laggards to Best Practice — was determined by each company's score as a percentage of total available points, where 100% represents full disclosure across all scored questions.

Insights

From this initial benchmarking of around 30 companies in Railpen's Active Equity Portfolio, several key insights have emerged, which we hope will be of some use to others in their own assessments of a given universe of holdings. However, given the relatively small sample size and portfolio-specific composition, the findings should not be over-interpreted and should not be seen as representative of all sectors or geographies. Despite these caveats, many of the findings align closely with wider market studies such as those published by ISS STOXX[15], The Conference Board[16], and UNESCO/Thomson Reuters Foundation[17], suggesting that the insights observed within the portfolio are reflective of broader trends in corporate AI governance and disclosure practices.

Insight 1 – AI governance maturity is uneven and often underdeveloped relative to risk exposure

AI governance performance varies within and across sectors in the benchmark, with leading practices emerging in areas like Information Technology, Financials, and Industrials, but consistent, sector-wide alignment on best practices has not yet emerged (Tables 5 and 6).

This pattern is also reflected in ISS STOXX's Mind the Governance Gap report which, drawing on Governance QualityScore data across 3,048 US companies in the Russell 3000 and S&P 500, found that AI oversight and formal AI policies were similarly concentrated in a limited number of sectors, including Consumer Discretionary, Financials, Health Care, Information Technology, and Industrials[18].

In the tables below, we compare how four companies spanning different performance tiers in the Financial and Information Technology sectors perform against the eight core questions in the AI Assessment Framework. The results reveal where companies are earning full, partial, or no points across each dimension, highlighting that uneven AI maturity exists even within the same sector.

Although we have earlier noted that some inconsistency is to be expected – and we recognise that despite being in a similar sector, companies' business models (and their use of AI) may vary significantly – we would expect there to be greater similarities than indicated by our analysis.

Table 5: Sample comparison of AI Assessment Framework results in Financial sector

	Company A	Company B	Company C	Company D
Question Overview	**Established**	**Emerging**	**Nascent**	**Laggard**
1) Relevance of AI to business strategy	●	●	●	●
2) Existence of comprehensive AI policy	●	●	●	●
3) Alignment of AI policy to Responsible AI principles	●	●	●	●
4) Board oversight of AI management	●	●	●	●
5) Day-to-day oversight of AI	●	●	●	●
6) Integration of AI into risk assessment processes	●	●	●	●
7) Management approach to AI-related risks	●	●	●	●
8) Reporting on performance against AI commitments and principles	●	●	●	●

Colour scheme: ● = Full points, ● = Partial points, ● = No points

Table 6: Sample comparison of AI Assessment Framework results within Information Technology sector

	Company A	Company B	Company C	Company D
Question Overview	**Best Practice**	**Emerging**	**Nascent**	**Laggard**
1) Relevance of AI to business strategy	🟢	🟢	🟢	🔴
2) Existence of comprehensive AI policy	🟢	🟠	🔴	🔴
3) Alignment of AI policy to Responsible AI principles	🟢	🟢	🔴	🔴
4) Board oversight of AI management	🟠	🟢	🔴	🔴
5) Day-to-day oversight of AI	🟢	🟢	🔴	🔴
6) Integration of AI into risk assessment processes	🟢	🟠	🟠	🟠
7) Management approach to AI-related risks	🟢	🟠	🟠	🟠
8) Reporting on performance against AI commitments and principles	🟠	🔴	🔴	🔴

Colour scheme: 🟢 = Full points, 🟠 = Partial points, 🔴 = No points

Implications for investors

The uneven maturity of AI governance both across and within sectors suggests that investors will continue to encounter significant variations in the quality of AI governance, even among companies with comparable AI exposures.

Where portfolio companies appear to be lagging their sector peers, investors could highlight these differences and encourage the laggard companies to strengthen their AI oversight structures and AI policies.

Insight 2 – Best practices are already emerging on AI policies

Despite the early-stage maturity of AI governance processes, leading companies have demonstrated several best practice disclosures that can serve as benchmarks for others.

Numerous standalone Responsible AI policies have been published and, while AI developers tend to score higher than deployers overall, examples of leading policies are emerging in both groups (see right). Best practices include:

- Stating the roles responsible for governance and management of the policy
- Explaining the processes used to implement the policy (including risk assessment and risk management processes)
- Outlining how the policy affects and relates to other policies (e.g. policies on data ownership, sovereignty, processing and privacy)

Examples of leading AI governance policies

Figure 3: Excerpts from ServiceNow[19] and RELX[20] AI Policies

How ServiceNow is committed to responsible AI

ServiceNow is proud to be a founding member of the AI Alliance in partnership with IBM, Meta, and other leading organizations to advance the principles of open, safe, and responsible AI globally.

ServiceNow has long been at the forefront of the latest AI developments. Intelligent chatbots and other AI capabilities like machine learning and predictive intelligence were introduced on the ServiceNow AI Platform as far back as 2018.

Incorporating the latest AI capabilities into the platform is a natural progression, and ServiceNow believes that providers and consumers of AI should implement solutions responsibly and ethically.

Inspired by the NIST Artificial Intelligence Risk Management Framework, ServiceNow has embraced four guiding principles for developing responsible AI:

1. **Human-centered**—ServiceNow generative AI (GenAI) solutions are built from a foundation of human-centric principles, including persona-based designs and an "out-of-the-box" experience that puts humans in control of AI-based decisions.

   Where AI is being used in products and services, ServiceNow provides clear documentation and guidance on how to deploy AI in a responsible manner, enabling customers to make informed decisions around where, when, and how they use ServiceNow AI solutions.
2. **Inclusive**—ServiceNow believes that AI has the power to reduce complexity and make the world a better place for everyone. AI team members strive to build models with datasets that are representative of the ServiceNow global customer base.

   ServiceNow AI solutions are continuously tested to promote reliable, high-quality outcomes for all. AI system development for skills and agents follows a high-risk review process that checks for potential bias, with any identified issues undergoing bias testing in partnership with the Responsible AI team.
3. **Transparent**—ServiceNow communicates with customers transparently on the topic of AI, using clear and understandable terms. AI documentation includes practitioner topics, like limits and intended usage. ServiceNow also shares detailed information on the governance foundations of its AI, like the type of data used for training/fine-tuning and approach to privacy and security.

   Publicly available model cards explain each specific language model's context, intended use, training/fine-tuning data, limitations, and other important information.
4. **Accountable**—Trust is the cornerstone of ServiceNow AI initiatives.

   ServiceNow has established internal governance bodies to provide accountability and oversee product and development activities, and also works closely with external experts and the AI community to gather feedback.

"Incorporating the latest AI capabilities into the platform is a natural progression, and ServiceNow believes that providers and consumers of AI should implement solutions responsibly and ethically.

3. Transparency and Explainability

Transparency and explainability both present challenges to systems designers:

Transparency, at its most basic level is that users are made aware when they are engaging with an AI system. For example, ensuring a consumer knows they are interacting with a chat bot. Beyond this, transparency can include providing information about how an AI system works and how it has been developed and trained.

Explainability means that people understand how the AI system arrived at a particular result, so that people, by working back through the system, can understand from the inputs and process, why a particular decision came about. It could mean explaining the criteria that was used to arrive at a decision, the data/logic behind an outcome, and even what errors can creep in and how they can be corrected.

With regards to providing explanation, in many cases the detail of how an AI system works may require a very complex description, which may not be understood by all users. Hence there is often a trade-off to be made between explanation and comprehension.

A further obstacle to explainability lies in the design of systems. Some machine-learning models autonomously determine the relative importance of items of data and the weightings to be attached to them, and these decisions are not made visible even to the operator of the system. This "closed box" problem means that there is not always a clear pathway to trace between inputs to the system and a particular output decision. In these cases, a full explanation is not even possible.

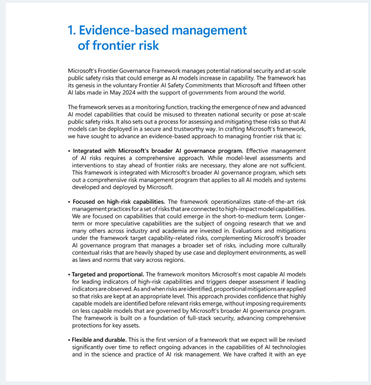
RELX POSITION

RELX believes that in all cases users should be informed when they are interacting directly with AI systems and for what purpose. The nature of this information will be highly context-specific and it should be the operator of the system who decides how specifically to convey explanations to their customer or user. In some applications, such as image-based diagnostics used in an expert clinical setting, the importance and level of explanation may be higher than in those cases where the decision being made is relatively mundane.

There will be certain scenarios, such as combatting criminal fraud, where it would be inappropriate and self-defeating to give a detailed public explanation on the methods being used. As with personal data protection, these instances can be identified and treated separately.

RELX is implementing an explainable-by-design approach across the businesses. Information where required is provided on how the system was trained and designed. This process is an iterative one. In time we may move towards an approach in which end-users can access information about specific decisions.

Figure 4: Excerpt from Microsoft's[21] Frontier Governance Framework



1. Evidence-based management of frontier risk

Microsoft's Frontier Governance Framework manages potential national security and at-scale public safety risks that could emerge as AI models increase in capability. The framework has its genesis in the voluntary Frontier AI Safety Commitments that Microsoft and fifteen other AI labs made in May 2024 with the support of governments from around the world.

The framework serves as a monitoring function, tracking the emergence of new and advanced AI model capabilities that could be misused to threaten national security or pose at-scale public safety risks. It also sets out a process for assessing and mitigating these risks so that AI models can be deployed in a secure and trustworthy way. In crafting Microsoft's framework, we have sought to advance an evidence-based approach to managing frontier risk that is:

- **Integrated with Microsoft's broader AI governance program.** Effective management of AI risks requires a comprehensive approach. While model-level assessments and interventions to stay ahead of frontier risks are necessary, they alone are not sufficient. This framework is integrated with Microsoft's broader AI governance program, which sets out a comprehensive risk management program that applies to all AI models and systems developed and deployed by Microsoft.
- **Focused on high-risk capabilities.** The framework operationalizes state-of-the-art risk management practices for a set of risks that are connected to high-impact model capabilities. We are focused on capabilities that could emerge in the short-to-medium term. Longer-term or more speculative capabilities are the subject of ongoing research that we and many others across industry and academia are invested in. Evaluations and mitigations under the framework target capability-related risks, complementing Microsoft's broader AI governance program that manages a broader set of risks, including more culturally contextual risks that are heavily shaped by use case and deployment environments, as well as laws and norms that vary across regions.
- **Targeted and proportional.** The framework monitors Microsoft's most capable AI models for leading indicators of high-risk capabilities and triggers deeper assessment if leading indicators are observed. As and when risks are identified, proportional mitigations are applied so that risks are kept at an appropriate level. This approach provides confidence that highly capable models are identified before relevant risks emerge, without imposing requirements on less capable models that are governed by Microsoft's broader AI governance program. The framework is built on a foundation of full-stack security, advancing comprehensive protections for key assets.
- **Flexible and durable.** This is the first version of a framework that we expect will be revised significantly over time to reflect ongoing advances in the capabilities of AI technologies and in the science and practice of AI risk management. We have crafted it with an eye

Several AI developers have also published frontier model safety frameworks which describe in detail the risk assessment and management measures in place to address risks associated with frontier models (Figure 4).

Implications for investors

Investors can use the emerging best practices identified here as a baseline when engaging with companies that lag on their AI governance practices. Investors could probe companies by asking whether their AI policies clearly assign governance responsibilities, describe implementation processes, and explain how AI policies interact with other organisational policies such as those on data privacy. Investors can also encourage improvements by sharing examples of policies that investors consider as demonstrating best practice.

Insight 3 – Policies tend to signal intent, not implementation

Many companies signal their intent through reference to responsible AI principles, but few provide evidence of implementation.

While approximately half of those companies in Railpen's Active Equity Portfolio have a policy statement that references AI principles (see right), most do not go beyond listing those principles, and few companies provide details on how the principles are actually embedded in day-to-day operations and decision-making. These findings mirror those presented in the UNESCO/Thomson Reuters Foundation Responsible AI in Practice: 2025 report[22], which surveyed nearly 3,000 companies across 11 sectors. The report concluded that most companies have moved beyond awareness of responsible AI challenges but that they continue to struggle with operationalising these commitments on a day-to-day basis and defining accountabilities for ensuring that these commitments are effectively implemented.

Responsible AI Principles

Various organisations across industry, government and civil society have published AI principles aimed at addressing AI-related risks. While not an exhaustive list, common principles include[23]:

- **Validity and reliability:** outputs should be truthful and regularly checked for accuracy
- **Safety:** systems should be monitored and human safety prioritised over commercial objectives
- **Fairness and inclusion:** AI should not replicate societal biases or cause disproportionate harm

Implications for investors

The potential gap between stated intent and operational practice presents a governance risk, and investors should be cautious of treating the existence of a policy, even one with commitments to responsible AI principles, as sufficient evidence of effective AI risk management.

Investors could probe companies on their policies to understand how these principles are implemented in practice through asking how AI models are chosen (and how these decisions are shaped by responsible AI principles), how AI principles are embedded in specific operations and who is accountable for ensuring compliance with the AI policy.

Insight 4 – AI-related risks identified are inconsistent within sectors and focused on established risk types

While most sampled companies mention AI risks in their reporting, the extent to which companies describe AI risks varies significantly[24].

Most companies tend to only mention AI as part of their description of how they manage other more established risks (most often cybersecurity), where the focus is on how AI might exacerbate these risks. Some leading companies with more established AI reporting go beyond this to mention a broader range of emerging AI-related risks, including bias, discrimination and data privacy. Our findings mirror those from a Conference Board analysis of S&P 500 filings, which found that 72% of companies disclosed at least one material AI risk in 2025, with disclosures focused mainly on cybersecurity, reputational and regulatory risks[25]. Our assessment similarly finds that there is a significant discrepancy in the specific risks that companies mention, even within similar sectors and business models.

Implications for investors

Seeing AI as a subset of, or an amplifying factor for, other technology risks may mean that companies understate the significance of AI risks, fail to identify the key mechanisms whereby the business is affected, fail to identify material AI risks and – ultimately - fail to effectively manage AI risks.

Investors could encourage companies to move beyond treating AI solely as an amplifying factor for existing risks and to consider the full range of AI-specific risks relevant to their business model.

This report's sector-level exposure tables and the OECD/AI Incident Database references are useful tools for cross-checking whether a company's disclosed risks align with known patterns of AI-related harm in the company's industry or sector.

Insight 5 – Leaders are moving beyond AI policies to transparent disclosure on AI practices

Greater external engagement with users and transparent reporting on performance against AI policies are characteristics of companies demonstrating more mature approaches to responsible AI.

While most companies assessed disclose elements of their AI policies and internal accountability structures, only a small subset provide meaningful information on how they engage with stakeholders on AI-related issues or on the occurrence and management of AI-related incidents. Examples of these disclosures include dedicated AI transparency reports, as well as the use of model cards (see right).

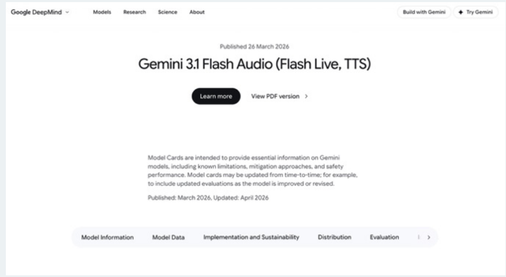
Examples of performance reporting – annual AI transparency reports

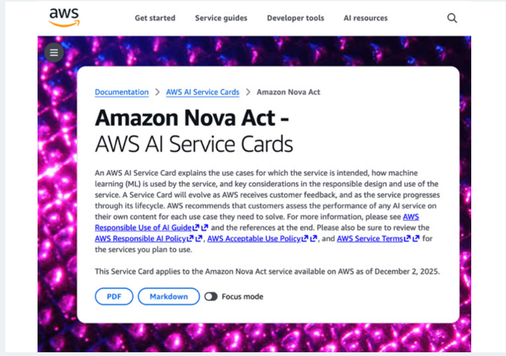
Figure 5: Alphabet's[26] and Microsoft's[27] AI transparency reports

Examples of external engagement – model cards

These cards, described as 'nutrition labels' for AI, often set out how models are trained, their intended use cases, known limitations and the safeguards in place to mitigate risk[28].

Figure 6: Examples of Alphabet's[29] and Amazon's[30] AI model cards





Implications for investors

The publication of AI transparency reports and model cards are, at the current time, indicators of a level of AI governance maturity within organisations.

Investors could encourage the adoption of such reports and model cards by their portfolio companies.

In their engagement, investors could focus on whether companies have mechanisms for reporting AI-related incidents and on how they engage with external stakeholders such as customers and civil society on AI issues.

IMPLEMENTATION: USING THE GOVERNANCE FRAMEWORK IN PRACTICE

As AI exposure across portfolios increases, investors need practical tools to assess governance quality, prioritise engagement, and track progress over time.

Railpen's AIGF can be applied to support these activities in a consistent and repeatable way (Table 7). Wider adoption of this framework by investors can ultimately help to strengthen the investor voice in advocating for best practices in AI governance.

"Managing the risks and the opportunities posed by AI requires a system-wide approach. Investors should be looking to actively engage companies to make sure they are using best AI practices. By uniting around a common set of expectations, investors can work collectively to reinforce efforts to drive responsible AI use."

Caroline Escott,
Head of Investment Stewardship and Co-Head of Sustainable Ownership, Railpen

"AI has clear implications for long-term company value, creating both opportunities for innovation and new areas of risk for investors to assess. Railpen's AI Governance Framework provides a practical way to compare how companies are governing AI relative to their exposure and peers. This can help investors strengthen due diligence and engage companies more effectively on the practices needed to support sustainable value creation."

Andrew McCarthy,
Investment Manager, Fundamental Equities, Railpen

Table 7: Potential applications of the AIGF

Feature	Description	Potential Uses
Tier rankings	The framework groups companies into five tiers based on AI governance maturity, differentiating between leaders and laggards.	• Prioritising companies for engagement • Guiding engagement through identifying best practice case studies • Informing investment due diligence by positioning companies relative to sector peers • Understanding trends in sector-level maturity to guide investment decision-making
Individual company assessments	The framework provides detailed insights into each company's governance approach	• Informing engagement based on gaps in AI governance • Supporting risk management by flagging companies with high AI exposure and weak governance • Tracking progress over time
Granular assessment questions	The framework is based on underlying indicators across three pillars	• Guiding targeted engagement discussions (see Box 8) • Benchmarking progress on specific practices (e.g. oversight, risk management processes) • Identifying common risks and/or approaches within different sectors

Railpen's approach to AI stewardship in its investment portfolio

Railpen recognises the significance of AI as an investment theme over the coming decade, with the potential to drive innovation and long-term value creation across our portfolio. Working with colleagues across the business, we actively assess how companies and assets might benefit from, or be threatened by, AI-driven disruption. We believe that effective governance will enable companies to both harness these potential benefits and mitigate risks.

As a long-term, broadly diversified investor, Railpen's approach to AI stewardship seeks not only to manage company level risk, but also to reduce vulnerabilities across the system as a whole. As Railpen embeds its **Responsible Technology Engagement Plan**, companies within the AI sub-focus area will be prioritised for engagement where the following factors overlap:

1. **Material financial exposure to AI-related risks**
2. **Potential for system-wide effects**, i.e. where governance failures, outages or misuse could have significant economic or other effects beyond the company itself
3. **Ability to influence peers, value chains or market standards**, creating the potential for outsized positive (or negative) spillover effects

This approach is designed to ensure Railpen's stewardship resources are focused where engagement can most effectively protect long-term portfolio value and help strengthen the resilience of the systems on which members' outcomes ultimately depend.

Proposed engagement questions

Governance and strategy

- How does the company implement responsible AI principles in practice?
- For deployers: How was the decision made to adopt a specific AI model or vendor?
  - Which criteria were used (e.g. performance, explainability, bias, security, cost)?
  - Does the company have mechanisms in place to assess and monitor their AI vendors and their AI models over time?
- How does the company expect its use of AI to evolve over the next 3–5 years?

Risk management

- How does the company assess and monitor material AI-related risks?
  - How is the significance of these risks assessed?
  - How frequently are these risks reviewed and updated?
  - If peer companies in the same sector have identified or prioritised different risks, have you identified these risks? Why do you not consider them to be a priority?
- What specific controls are in place to mitigate key AI risks (e.g. bias testing, human oversight, model validation)?
- How does the company assess the potential impact of AI system failures on core operations?

Performance reporting

- Has the company experienced any AI-related incidents (e.g. model errors, data breaches, misuse), and if so, how were these issues identified, managed and remediated?
- How does the company assess whether its responsible AI commitments are achieving their intended outcomes?
- Which aspects of implementing the company's responsible AI commitments are most difficult to measure and report, and how is the company addressing these challenges?

CONCLUSION

AI governance is rapidly becoming a material consideration for investors, yet this report shows that AI governance practices remain uneven and often underdeveloped relative to the pace of AI adoption.

While leading practices are beginning to emerge, most companies have yet to move beyond stating their intent to demonstrating implementation. In addition, risk disclosures remain narrowly focused on established risk types; it appears that many companies see AI-related risks as an extension of well-established risk areas (cybersecurity) rather than as presenting new and potentially quite different organisational risks.

Investors have an important role to play in driving progress on AI governance. This includes using structured assessments to identify risks and assess governance practices across portfolios, engaging companies to strengthen governance and disclosure practices, sharing examples of good practice, and promoting greater awareness of AI governance expectations across the investment and policymaking ecosystem.

Ultimately, managing the risks posed by AI requires a system-wide approach. **By uniting around a common set of expectations - such as those embedded in Railpen's AI Governance Framework - investors can work collectively to reinforce efforts to drive responsible AI use and protect long-term value**.

REFERENCES

1 For the purposes of this report, 'Active Equity' refers to our conviction-led investment approach grounded in fundamental analysis

2 World Economic Forum (2022), Technology Convergence Report

3 McKinsey (2025), The state of AI In 2025: Agents, innovation, and transformation

4 EY (2025), How can responsible AI bridge the gap between investment and impact?

5 World Economic Forum (2022), Systemic Cybersecurity Risk and the role of the Global Community: Managing the Unmanageable

6 McKinsey (2025), The state of AI In 2025: Agents, innovation, and transformation

7 Further information on Railpen's approach to AI-related investments in our Real Assets and Active Equity portfolios can be found on pages 25-26 of our 2025 report Achieving Effective AI Governance, noting that the scope of this report is Active Equity.

8 Some divergence is expected given differences in business models and therefore risk exposures; however, the degree of variation observed across companies within the same sector suggests that companies' understanding of their AI risk is still evolving.

9 World Economic Forum (2022), Technology Convergence Report.

10 European Systemic Risk Board (2025), Advisory Scientific Committee: Artificial intelligence and systemic risk.

11 Railpen (2025), Achieving Effective AI Governance; Insight Investment (2026), AI Governance: Protecting portfolios in an AI-driven world; Hermes Investment Management (2019), Investors' expectations on responsible artificial intelligence and data governance; World Economic Forum (2024), Responsible AI Playbook for Investors

12 Case studies of these risks crystallising in practice can be found in our previous report: Achieving Effective AI Governance

13 European Union Artificial Intelligence Act, Annex III

14 Note that in the 2025 report, four pillars were presented. In this iteration, the governance and strategy pillar have been combined. The reason is that our research indicated that, in practice, these pillars are closely linked with many companies already exhibiting mature practices in the strategy pillar.

15 ISS STOXX (2026), Mind the Governance Gap

16 The Conference Board (2025), AI Risk Disclosures in the S&P 500: Reputation, Cybersecurity, and Regulation

17 UNESCO/Thomson Reuters Foundation (2026), Responsible AI in Practice: 2025

18 ISS STOXX (2026), Mind the Governance Gap

19 ServiceNow (2026), Responsible AI

20 RELX (2022), RELX Position Paper on AI

21 Microsoft (2025), Frontier Governance Framework

22 UNESCO/Thomson Reuters Foundation (2026), Responsible AI in Practice: 2025

23 Please see railpen-achieving-effective-ai-governance-report.pdf p.29 for a more exhaustive list of Responsible AI principles.

24 Though, as we note earlier in this report, some discrepancies are understandable given that even within sectors, companies will have different business models, strategies and exposures.

25 The Conference Board (2025), AI Risk Disclosures in the S&P 500: Reputation, Cybersecurity, and Regulation

26 Google (2026), Responsible AI Progress Report

27 Microsoft (2025), Responsible AI Transparency Report

28 Gilbert et al. (2025), Could transparent model cards with layered accessible information drive trust and safety in health AI?

29 Google DeepMind (2026), Model Cards

30 AWS (2026), Explore AWS AI Service Cards

Disclaimer

The information contained in this document is provided solely for informational purposes and has been produced by Railpen Limited working in collaboration with Chronos Sustainability. While every effort has been made to ensure the accuracy and completeness of the information, no representations, or warranties (whether express or implied) are made about the completeness, accuracy, reliability, suitability, or availability of the information contained herein. This document does not constitute legal, financial, or professional advice and should not be relied upon as such. Railpen Limited and Chronos Sustainability disclaim any and all liability for any loss or damage, whether direct, indirect, consequential, or otherwise, arising from the use or reliance on the information contained herein.

The views expressed within this document are those of Railpen Limited and Chronos Sustainability at the date of publication, which are subject to change. The information contained within this document does not constitute investment advice and should not be treated as such.

Railpen Limited is a wholly owned subsidiary of Railways Pension Trustee Company Limited, registered in England and Wales under company number 02315380 whose registered office is at 7 Devonshire Square, London, EC2M 4YH

7 Devonshire Square, London, EC2M 4YH
SO@railpen.com

RAILPEN

CHRONOS
INTELLIGENT SUSTAINABILITY