Responsible AI: Managing AI Risks and Compliance with Innovation [1, 2]

Rizwan Amin Sheikh, Ph.D., PMP
Khalid Ahmad Khan, Ph.D., PMP
Aamir Khalid Pirzada

The greatest risk in AI is not that machines will become too intelligent, but that we will deploy them without understanding their vulnerabilities.

Abstract

Artificial intelligence is transforming how projects are delivered, decisions are made, and value is created. Yet without proper governance, AI can expose organizations to risks, regulatory penalties, reputational damage, and operational failures. How do project leaders balance AI's transformative potential against its inherent risks?

This paper introduces **AI-CARVER**—a practical, quantifiable framework for assessing AI risks adapted from the battle-tested CARVER methodology featured in ***Harvard Business Review*.** Originally developed for military target analysis and adopted by the Department of Homeland Security and Fortune 500 security directors, CARVER has been reimagined for the AI era. AI-CARVER enables project managers and executives to systematically evaluate AI systems across six critical dimensions: Criticality, Accuracy, Regulatory Exposure, Vulnerability, Explainability, and Recoverability.

Participants will gain hands-on experience applying AI-CARVER to real-world scenarios and leave with immediately actionable tools for AI governance, risk management, and compliance (GRC). The paper also explores insights from ***Harvard Business Review's*** **"8 Questions About Using AI Responsibly, Answered,"** examines the NIST AI Risk

[1] Editor's note: Second Editions are previously published papers that have continued relevance in today's project management world, or which were originally published in conference proceedings or in a language other than English. Original publication acknowledged; authors retain copyright. This paper was originally presented at the 18th Project Management Symposium at the University of Texas at Dallas in May 2026. It is republished here with permission of the author and conference organizers.

[2] How to cite this paper: Sheikh, R.A., Khan, K.A., Pirzada, A.K. (2026). Responsible AI: Managing AI Risks and Compliance with Innovation; Originally presented at the 18th Project Management Symposium at the University of Texas at Dallas in May, republished in the *PM World Journal*, Vol. XV, Issue VII, July.

Management Framework, and unpacks compliance requirements under GDPR and the EU AI Act—regulations now affecting global banks, healthcare organizations, pharmaceutical companies, manufacturers, technology firms, and others. Moreover, the authors will introduce a practical AI GRC Assessment Model that participants can immediately implement within their organizations.

Responsible AI is not about slowing innovation—it's about building a sustainable foundation for AI systems that are ethical, secure, and compliant while remaining innovative and productive.

Learning Objectives

1. **Discuss** how companies can apply the AI-CARVER framework—a six-part risk assessment methodology adapted from *Harvard Business Review*—to systematically evaluate AI systems across Criticality, Accuracy, Regulatory Exposure, Vulnerability, Explainability, and Recoverability, enabling data-driven governance decisions.
2. **Explore** how to use AI responsibly through eight key questions posed by Professor Tsedal Neeley of Harvard Business School.
3. **Learn** how to utilize a practical AI GRC (Governance, Risk Management, and Compliance) Assessment Model that participants can immediately implement within their organizations to identify AI risks and compliance gaps, prioritize mitigation strategies, develop remediation roadmaps, and establish accountability structures.
4. **Examine** the evolving regulatory landscape for AI, including the NIST AI Risk Management Framework, GDPR, and the EU AI Act, and assess how global organizations across banking, healthcare, pharmaceutical, and manufacturing sectors can manage AI risks and navigate compliance while pursuing innovation.
5. **Provide** guidance to project managers on how to integrate responsible AI principles into project management practices, ensuring AI systems are designed, deployed, and monitored in ways that advance fairness, transparency, security, and ethical accountability without stifling innovation.

A number of management consulting organizations are reporting similar findings about the urgency of establishing responsible AI governance. According to Deloitte's Global Boardroom Program (2025), nearly one-third of board members indicate that AI is not on their organization's agenda, indicating a clear and urgent governance gap at the most senior level of leadership in organizations. In addition, Deloitte's State of AI in the Enterprise report (2026) states that only 20% of organizations have developed mature governance models for autonomous agents. This represents a considerable gap in light of the rapid development of autonomous agents in every industry. According to McKinsey's State of AI report (2025), nearly 88% of organizations are currently utilizing

AI technologies in at least one operational area; however, very few organizations have reached the stage of scaling their use of AI. This governance deficit is further confirmed by CallMiner (2024), who found that while 71% of organizations have dedicated AI governance resources, 67% are simultaneously implementing AI without adequate governance structures — a paradox highlighting the difference between governance in name versus practice. Despite growing awareness, only 25% of organizations have fully implemented an AI governance program, meaning 75% remain without comprehensive governance in practice (AuditBoard, 2025; Larridin, 2026). In addition, few organizations have systems to measure and report on key AI characteristics such as fairness and the organizational impact of AI. The results of McKinsey's AI Trust Maturity Survey (2026) show another significant finding: organizations that define a clear set of responsibilities for responsible AI governance score 50% higher on their maturity assessment than organizations that do not. This is clear evidence that organizations must have some formal governance mechanism for their AI systems to provide accountability; otherwise, it is impossible to operate responsibly.

The clear path forward for those organizations willing to take action is well documented. KPMG (2024a) found that organizations that have successfully developed an AI strategy in conjunction with clear governance structures have gained measurable competitive advantages when compared to organizations that have not developed an AI strategy. Organizations without a clear AI strategy and governance structure struggle as they move towards increased automation. According to KPMG Australia (2024b), it is imperative that senior leadership is committed to establishing strong governance for responsible AI; that this commitment provides direction for the organization's strategy and embeds ethical values throughout the organization; and that organizations develop a workable structure to effectively measure, report, and communicate stakeholder expectations regarding the impact of AI and the information technology systems associated with AI and autonomous agents. According to Accenture (2024), only 78% of organizations have started to implement a responsible AI program and only 14% have successfully implemented that program to date. This results in organizations being exposed to compliance and reputational risks. Finally, according to Accenture (2025), there is a direct relationship between responsible AI governance and customer retention. Organizations that have established a governance framework for responsible AI that incorporates trust and fairness can expect to see a 25% increase in customer retention as a measurable return on investment in responsible AI governance.

Today, one of the biggest challenges that organizations face isn't whether to use artificial intelligence, but rather how to govern it. As AI systems take on greater importance in areas from clinical diagnostics to financial trading to self-driving cars, the need for a

structured, quantifiable approach to AI risk assessment and governance has never been greater.

Enter **AI-CARVER**, an upgrade of the legendary CARVER methodology for the age of AI. CARVER, originally developed by military analysts in WWII to determine the target for bombing runs, has been adopted by the U.S. Department of Homeland Security, Fortune 500 security directors, and risk management professionals at work around the world. Now this battle-hardened weapon has been adapted for one of the key challenges of our generation, namely the governance of AI. According to Araboghli and Bencie (2018), the CARVER framework was first introduced as a six-part risk prioritization tool to a managerial audience by the Harvard Business Review. Discovering its utility for assessing risk and creating a prioritization system supports this paper's foundation.

After spending over two decades conducting enterprise AI projects in healthcare, financial services, and government—including developing an AI governance patent—the authors observed how organizations struggled to quantify AI risk. Traditional risk matrices fall short when it comes to dealing with AI systems that hallucinate, discriminate, or make decisions in ways that even its authors often don't fully understand themselves. AI-CARVER offers a structured and repeatable methodology for scoring these apparent risks of AI. Using authors' previous research on AI governance (Sheikh, 2025) and AI-enabled project management (Sheikh et al., 2024) as bases of support, this framework proposes the addition of a repeatable and quantifiable scoring model to help convert AI risks into metrics that boards of directors, executives, project managers, and decision-makers can use.

What is AI-CARVER?

AI-CARVER is an acronym that stands for six critical dimensions of AI risk assessment, which are shown in Table 1. While it preserves the elegant structure of the original CARVER methodology, each criterion has been specifically calibrated for the unique characteristics of AI systems.

Table 1: AI-CARVER Acronym Dimensions

Letter	AI-CARVER Criterion
C	**Criticality** – How critical is this AI system to your mission?
A	**Accuracy & Reliability** – How likely is your AI system to be wrong or undergo "hallucinations"?
R	**Regulatory Exposure** – What's the level of exposure of your AI system to compliance and regulatory risk?

Letter	AI-CARVER Criterion
V	**Vulnerability** – How best might they penetrate your AI system from adversarial attack, data poisoning, manipulation?
E	**Explainability** – You must be able to audit and explain AI's decisions but can you?
R	**Recoverability** – How quickly & frequently can you recover?

Score each criterion from 1 to 5 (highest) to produce an overall AI-CARVER score of 6 (lowest) to 30 (highest). This score makes it easier to compare systems and show it to management and the board how you're spending your governance dollars.

The Six Dimensions of AI-CARVER

C – Criticality

Criticality in this context means how badly an organization would be affected if there was a failure of the AI system (or AI system failed). The more involved with the organization's critical operations, i.e., human safety or financial stability, the more critical the AI system.

Scoring Guide

- **5 - Mission Critical.** The AI system is fundamental to the organization, i.e., would cause the entire organization to fail if the system stopped working, in the extreme, could cause death. For example, life-support systems, air traffic control.
- **4 - High Impact.** If this AI system stopped working, operation of day-to-day would be severely affected and could result in significant loss of money; if the organization was able to continue to operate, the organization would still feel the impact as a whole.
- **3 - Moderate Impact.** Failure of the AI system could be managed, however, correcting for the failure or providing a work around solution would take considerable time, effort, and resources; thus, would be painful but would not be catastrophic to the organization.
- **2 - Low Impact.** If the AI system fails, the failure would be documented and managed; however, no major crisis would occur; alternative solutions to accomplish the same task do exist and the failure of the AI system would result in minor operational disruption to the organization.
- **1 - Minimal Impact.** The AI system would be nice to have as an additional convenience; if the AI system did not exist tomorrow, operations would be relatively unaffected; for example, a labor-saving tool that is not critical.

A – Accuracy & Reliability

Accuracy and reliability address the issues of how often the AI produces an inaccurate result, as well as the types of inaccuracies. This includes four major types of AI failure — hallucinations (creating something that doesn't exist), bias (favoring one outcome over another in an unfair manner), drift (decreasing performance over time) and inconsistency (multiple answers provided to the same question). Generative AI systems, such as content generators and chatbots, make these factors even more critical.

Scoring Guide:

- **5 - Very High Error Rate.** The AI system is highly untrustworthy and it regularly creates false information, produces bias or discriminatory responses, and cannot be relied upon with confidence.
- **4 - Somewhat High Error Rate.** While the AI system has some value, it is too wrong frequently that human review of the output is needed before action can be taken on the output. The system cannot be left unattended by a human.
- **3 - Some Errors but often Accurate.** Although the AI system occasionally creates a fictitious response or errors, it performs satisfactorily. The system may produce occasional errors, but those errors are typically not common occurrences, and the system will perform successfully with limited supervision.
- **2 - Confidently Reliable.** The AI system performs very well most of the time, there is no need for a user to second-guess it constantly when using it, such as walking a familiar path without looking at the ground the entire time. There is also no significant sign of bias in the system.
- **1 - Strongly & Very Reliably Tested.** The AI system has gone through extensive testing, therefore does not fail. This system has the best track record of producing fair, accurate and consistent outputs of an AI system.

R – Regulatory Exposure

The question posed by Regulatory Exposure is how much legal and compliance oversight is required for that specific AI system. What laws, regulations and/or industry standards apply to the specified AI system, and to what degree do those laws or standards apply.

Scoring Guide:

- **5 – Strictly Regulated.** The system is subject to multiple strict and directly applicable regulations including the requirement for formal government approval (e.g., FDA approval). It falls under risk-based frameworks for AI systems classified as higher risk

in the European Union AI Act. Compliance is required in order to operate the system legally.

- **4 – Heavily Regulated.** The system is being operated in an industry that is heavily regulated, and major frameworks include, but are not limited to, HIPAA (healthcare privacy) and SOX (financial reporting). While these are not strictly AI-specific regulations, they do limit how the system is designed, built, operated and audited to a significant degree.
- **3 – Moderately Regulated.** There are many regulations that technically apply; however, the vast majority have little to no ultimate impact on the day-to-day operation of the specified AI system. Compliance is required, but not overwhelmingly burdensome.
- **2 – Lightly Regulated.** There are some general accountability and internal control requirements in place, however, there are no specific requirements targeted at AI systems. The oversight of the specified AI system is broad and relatively loosely regulated.
- **1 – Minimally Regulated.** The specified AI system operates primarily as an internal operating system with no significant oversight from an external regulatory agency. There are no government agencies or industry regulators watching or enforcing compliance with any requirements for the specified AI system.

V – Vulnerability

Vulnerabilities are a measure of the probability of your AI system being taken out through various forms of attack such as prompt injection, data poisoning, model extraction, and by preventing the AI from responding through evasion attacks. Also, consider social engineering targeting the outputs from your AI. The more documented vulnerabilities present, the higher the score.

Scoring Guide:

- **5 - Extremely Vulnerable:** The AI system is exposed to the world via the Internet and lacks input validation/sanitation. This means that someone could easily manipulate your system, either on purpose or accidentally. This type of vulnerability could potentially trigger emergency response alerts or cause other major unintended consequences.

- **4 - High Risk:** The AI system has a known attack surface. Your system has multiple documented vulnerabilities that any attacker would be able exploit easily. Additionally, the system is missing a variety of security mechanisms that would have protected it

from attacks. Overall, the system is a prime target for attacks and has very little to protect it from being attacked (i.e., an obvious target).

- **3 - Moderate Risk:** The AI system has reasonable security. It is typical or has average security, which means that while the system has typical security it would protect it from most kinds of attacks. The system is not very easy to attack, but, at the same time, is not significantly hardened making it less likely to be compromised. The system is essentially a compromise between a target and a hard target.

- **2 - Low Vulnerability:** The AI system is air-gapped. It has no physical connection to the internet or any outside network. Therefore, the system cannot be attacked from the internet or outside systems.

- **1 - Rusty.** The AI system is under constant security and inspection and has been subjected to red-team testing. Additionally, it has a large amount of cloud-based security services protecting the hardened system. The most secure form of a hardened system (i.e., an obsolete or "rusty" system) is constantly monitored for suspicious activity and has been thoroughly tested through ongoing red-team testing. Additional security, in the hardened system is backed by sufficient cloud-based security services to serve as redundancy (i.e., the primary means of security).

E – Explainability

Explainability refers to how easily we can understand and communicate the rationales behind each step of an AI system's decision-making process to relevant parties, as well as the traceable nature of the AI's decision-making process. Black Box AI systems lack transparency and provide only minimal ability to explain outputs. Fundamental question: How can we know what logic the AI used to come up with its final answer?

Scoring Guide:

- **5 - No Explainability.** Black Box AI decision system, i.e., no explainability exists. There are no methods available for understanding or tracing the processes of the AI systems to arrive at their outputs. No audit trails exist for AI outputs or the rationale for the conclusion reached by the AI during the audit process, despite the significance of those conclusions.

- **4 - Limited Explainability.** Black Box AI decision system with limited explanatory value; currently, there has been limited work done to allow for AI systems to be

understood at scale. Therefore, interpreting the decision-making processes of the AI systems is virtually impossible.

- **3 - Partial Explainability.** Partially understandable AI decision system; partially transparent; relatively opaque; both types of information exist within the AI systems. The AI system's transparency level is not complete; however, some elements of the AI systems can be understood and extracted.

- **2 - Good Explainability.** Mostly transparent AI decision system; good explainability; the majority of AI outputs will be understandable by independent reviewers, and there will be supporting documentation explaining why the AI produced its outputs and could have been verified by the reviewers.

- **1 - Full Explainability.** Complete transparency of AI decision system with full explainability; every output from the AI system could be reviewed, and there will be no gaps in the verification process by independent reviewers for any output from the AI systems.

R – Recoverability

Recoverability refers to how quickly and effectively an organization can return to normal operations after a failure, corruption or attack on an AI system—making it a very real issue for many organizations. Over the next several years, there will be an increase in organizations that face difficulty recovering from an incident, resulting in security investigations and the need for data collection; however, only a few will experience true system restoration.

Scoring Guide:

- **5 – Very Difficult to Recover.** There is no fallback plan and no backup mechanism. Recovery will be painful and lengthy. When something does come back, it may be just a temporary fix. The organization may be left with no options.

- **4 – Difficult to Recover.** Getting back to normal will take weeks and will not be easy. Many issues will arise during recovery that will disrupt operations for a long period of time; thus, this is an unsatisfactory outcome for the organization.

- **3 –Somewhat Difficult to Recover.** The situation can be managed; however, there will need to be manual workarounds to maintain operations. Once the organization puts forth the effort, it will take four to five days to restore to full normalcy.

- **2 –Easy to Recover.** The organization has systems in place to provide backup support i.e., hot (immediate) backups or cold (stored) backups. Thus, recovery can occur in four to eight hours.

- **1 –Immediate Recovery.** The organization has automated recovery mechanisms that will activate automatically. The organization has a manual override capability to provide an additional level of assurance that recovery will be accomplished immediately.

Case Study: AI-Powered Radiology in a Hospital System

A large hospital network is evaluating an AI-driven radiology tool for analyzing chest x-rays to identify pneumonia, COVID-19, and early-stage lung cancer among other conditions. Before rolling out this solution for use across its 12 hospital locations, the Chief Medical Information Officer (CMIO) needs to determine the risks associated with the product – and AI-CARVER does just this by having the organization walk through each dimension and assess the associated risk.

As shown in Table 2, to evaluate the effectiveness of the AI-CARVER methodology, the CMIO calculated total risk score of 18 (out of 30).

Table 2: AI-CARVER Assessment Criteria

Criterion	Assessment	Score
Criticality	AI assists in diagnosis but radiologists make final decisions. Failure delays but doesn't halt care.	**3**
Accuracy	FDA-cleared; 94% sensitivity; some false negatives possible; requires radiologist validation.	**3**
Regulatory	FDA 510(k) cleared; HIPAA applies; state AI transparency laws; malpractice liability concerns.	**5**
Vulnerability	Isolated on hospital network; images could potentially be adversarially modified.	**2**
Explainability	Heatmaps show regions of concern; limited insight into model reasoning; difficult to explain to patients.	**4**
Recoverability	Radiologists can read X-rays without AI; minimal downtime impact; human override always available.	**1**
TOTAL	**Moderate-High Risk: Proceed with enhanced governance controls**	**18/30**

Interpreting the Score: The score of 18/30 classifies the Radiology Solution as a Moderate-High Risk solution. The two biggest risk categories highlighted by the

organization are Regulatory Exposure (score of 5) & Explainability (score of 4), indicating where the organization should focus the most in developing its governance program.

Therefore, for the CMIO to develop a functional compliance monitoring scope for the organization, complete documentation will be maintained to facilitate compliance audit with the applicable regulations. Additionally, all patients will receive adequate explanations of how AI is involved in their diagnosis and the information will be provided in an appropriate manner. Typical risk scores, the category of risk, and the recommended actions are displayed in Table 3.

Table 3: AI-CARVER Risk Categories

Score Range	Risk Category	Recommended Action
6–10	**Low Risk**	Standard governance; annual reviews
11–15	**Moderate Risk**	Enhanced monitoring; quarterly reviews; human oversight
16–20	**Moderate-High**	Rigorous controls; monthly reviews; board reporting
21–25	**High Risk**	Extensive safeguards; continuous monitoring; C-suite approval
26–30	**Critical Risk**	Consider alternatives; board approval required; external audit

Why AI-CARVER Matters Now

As the proliferation of regulation around AI continues to rapidly rise globally, ranging from the European Union's AI Act to various AI regulations from multiple states in the United States, organizations have turned to more structured forms of governance relating to the deployment of AI by utilizing governance standards (i.e., control framework) for the entire organization, which includes any enterprise-wide AI systems. The Economist Intelligence Unit (2020) found that organizations whose responsible AI programs are mature generate greater customer trust and resiliency over time than their peers. Therefore, the business case for governance is as strong as the regulatory requirement that mandates it. Adding to that information, Davenport and Zhang (2021) showed that to achieve a return on AI investments will be determined primarily by having disciplined governance of the projects themselves. Therefore, organizations that develop AI but conduct the deployment without sufficient oversight will underperform compared to those with structured governance.

By using AI-CARVER (a methodology to assess/review/decide on) employed through all forms of engagements, the methodology is powerful due to both its simplicity and repeatability. No matter whether it is the Chief AI Officer who is assessing the AI-powered generative chatbot, the CMIO of a hospital who is evaluating diagnostic algorithms or the

CRO of a bank who is reviewing algorithmic trading systems, AI-CARVER provides a common vernacular for discussing any types of risks associated with AI technology.

AI-CARVER thereby converts subjective concerns into numerical scores that can be tracked over time, compared (between different AI systems), and communicated (by reporting) to stakeholders (boards and regulators). As organizations are making high-level decisions around AI in their boardrooms (by the executives for which justifying their governance investments will require hard numbers), AI-CARVER can provide the numbers to substantiate those justifications, thus not only allowing organizations to develop an AI faster but also allowing them to do so in responsible manner.

> ***The question is not whether your organization will use AI—it's whether you will govern it before something goes wrong.***

Using Artificial Intelligence Responsibly: Key Questions for Organizations

The recent, quick growth of generative AI platforms (like ChatGPT and Midjourney) has provided a greater incentive for organizations to respond to the operation and moral implications of using artificial intelligence than before. Though many of the questions about whether or not to adopt AI were already existing, we have now become able to provide them greater weight and importance than we did previously. Through research at the crossroads of occupations, technology and organizations, Tsedal Neeley (Professor, Harvard Business School) offers eight concrete questions that can be used as a foundation for responsible use of AI. Each of the questions aligns with the dimensions of the AI CARVER and the AI GRC (Governance, Risk Management, and Compliance) Model (Sheikh, 2025). It also demonstrates that the evaluation of AI preparedness is based on the same governance concepts as this white paper discusses.

1. How Should Organizations Prepare to Introduce AI?

The way artificial intelligence (AI) is used differs from all previous technology. It does not simply improve the efficiency of work completed through traditional processes; instead, it identifies and analyzes large sets of data to independently produce results. Due to this new breed of technology, organizations should see AI as a partner in their business rather than merely another application. To effectively implement AI into their organizations, there are three fundamental principles that organizations should consider: 1) Every employee within an organization should have a minimum level of digital literacy equal to thirty percent. Digital literacy is defined as having knowledge of systems architecture, machine

learning, algorithms, and cybersecurity; together, these concepts will comprise an organization's digital ecosystem. According to Neeley and Leonardi (2022), developing this digital mindset throughout an organization is essential to enable the organization to effectively and responsibly exploit AI. 2) In order to be successful with AI implementation, organizations need to create themselves with the mindset that continuous change is the norm. Organizations must break down silos and create shared or central repositories for all organizational data, enabling collaborative use of this data by everyone in the organization. 3) AI needs to be part of the current operational structure; one example of effectively utilizing this aspect is the requirement from Amazon that all teams utilize Application Programming Interface (API) calls to route company data, thus creating an environment in which any necessary changes can be easily accommodated through architectural flexibility.

2. How Can Organizations Ensure Transparency in AI Decision-Making?

Ultimately, having total transparency into an AI-driven decision-making process is extremely challenging, if not impossible. Artificial intelligence possesses two fundamental characteristics: invisibility (often running in the background and going unnoticed) and inscrutability (their internal logic is not understood by even those who build them). A good example of how difficult it is to achieve transparency is large language models, which utilize hundreds of billions of parameters to be trained and have been fine-tuned over time through feedback from users. Due to the inherent cost of building these types of systems, an organization's leadership team must create documentation on how and when they are using AI, while providing careful judgment on acceptable use cases. Organizations can improve the transparency of their AI solutions through a number of technical means, including: identifying which data points contribute to the output of the AI solution(s), creating more interpretable models, and using premortem analyses to identify potential risk factors associated with the project prior to deploying it.

3. How Can Guardrails Be Built Around Large Language Models?

There are many important risks associated with large language models such as; Dangerous stereotypes, misinformation, having an impact on personal privacy, enabling cyber-attacks, and generating an environmental cost to high compute requirements. Bender et al. (2021) introduced the term "stochastic parrots" to convey the idea that large language models can generate fluent but misleading or harmful outputs and warn of the increased danger created by unchecked scaling and inadequate governance. Two ways that these risks can be managed are:

- Choose how you curate training data as opposed to maximizing your scale during training will lessen your likelihood of generating bias or harmful content. The majority of training data found on the internet (like Reddit and Wikipedia) has a bias towards male contributors and provides little representation of marginalized communities.
- Make available documentation about the dataset used to develop your model so that other organizations can be more informed of the potential to have similar bias present in third-party models when using your model, and to know if a third-party model is aligned with their company values.

4. How Can Organizations Ensure Representative and Unbiased Training Data?

Prioritizing fairness is important in order to effectively deal with bias in AI training datasets, rather than concentrating solely on making the models larger. The complexity of larger models makes it much more difficult to audit for any bias (e.g., their unpredictability makes them hard to define and track). Organizations should also ensure that their teams who are responsible for collecting and curating the training data include some members from under-represented populations; this creates a more well-rounded perspective that can help identify and fix any 'blind spots' to which homogeneous teams are not necessarily aware of. When AI is trustworthy, it should be able to accomplish similar outcomes with two or more different groups of people; therefore, there must be both a diverse content, or data distribution system, and transparent documentation of how issues of fairness have been considered and incorporated into the development of the AI algorithm. Smith and Rustagi (2020) provide practical recommendations, such as conducting structured audits, using diverse datasets, and continually testing for fairness, as critical elements for any responsible deployment program utilizing AI technology.

5. What Are the Risks of Data Privacy Violations with AI?

AI tools that handle confidential data, whether about employees or customers, are typically very susceptible to any misuse. Before deploying any such tools utilizing confidential data, companies need to perform a thorough assessment of how the tools were created; they also need to invest continuously in security updates to keep these tools safe. One possibility for maintaining the security of data transactions is through the use of blockchain technology. On a broader level, a “privacy by design” (PbD) framework (which was created by Ontario's former Information and Privacy Commissioner) outlines seven major principles for ensuring the integration of privacy in an organization’s daily operations. The PbD (privacy-by-design) Principles developed by Cavoukian (2011) are

a proactive governance model, integrating privacy into the design of the system rather than retrofitting privacy after systems have been deployed.

The seven principles are:

1. Preventative and proactive rather than reactive.
2. Default privacy settings.
3. Privacy is built into the design of systems.
4. Privacy and security are complementary to each other.
5. Data security features apply from beginning to end.
6. Visible, transparent operations.
7. User-centered design with respect for individual privacy.

Whether implementing PbD is going to be a company-wide initiative, or the sole job of a privacy team, privacy needs to be considered a fundamental part of the business operation of the organization and not something that will be addressed later on.

6. How Can Employees Be Encouraged to Use AI Productively?

In order to effectively use AI, you must also know its strengths and weaknesses. There are many different uses for AI, such as ChatGPT for generating ideas or content based on patterns. However, AI does not actually understand what it is doing; it can copy your writing style, for example, but it does not understand what the meaning of your writing style is. On the other hand, AI is much better than people at making predictions based on data, such as how well someone will do over time or whether a certain person has developed a disease like dementia through speech analysis or finding tumors that cannot be seen with the naked eye. Therefore, employees should be educated on matching the strengths of AI with the right jobs and there should be human oversight of all high-risk work produced by the use of AI. All AI generated work product is only as good or bad as the data it uses or the logic that was used to create an algorithm for the AI.

7. How Much Should Organizations Worry About AI Replacing Jobs?

According to history, advances in technology do more for the economy than they take away. The introduction of automobiles resulted in an industry that did not exist previously, in that it replaced horse and buggy operators with thousands of automobile mechanics, manufacturers, suppliers, etc. The same can be said with regard to AI – while AI will most likely change how jobs are performed and the duties performed by the job, it is not as likely to eliminate people's jobs altogether. For example, by using AI to generate

commission forecasts, sales reps will spend less time on data analysis and more time on building relationships and creating sales strategies. Similarly, coding tools will allow programmers to automate repetitive tasks and devote their time to working on higher-level issues such as programming platform architecture. Although some jobs will be disrupted more than others depending on the industry and/or geographic area, the largest issues associated with AI in the workforce will be the lack of workers in the industry that have digital skills to capitalize on this technology. Therefore, investing in educational opportunities and job training are critical to realize the benefits of AI for all. The MIT Work of the Future Taskforce (2022) confirms this point, as organizations develop adaptive workforces and invest in skills development to provide the optimal response to ongoing innovation and change caused by the use of AI technology transforming the workforce, and not by eliminating traditional jobs.

8. How Can Organizations Prevent AI from Harming Individuals or Violating Human Rights?

The risk of AI-related harm is well-known. For example, several studies revealed that common facial recognition software works for white men more accurately than Black women, with errors as high as 35% — leading to wrongful arrest of innocent Black women too. Buolamwini & Gebru (2018) documented this variance in effect/accuracy by developing a documented analysis of significant intersectional accuracy variance in the several commercially deployed AI systems included in their Gender Shades project study. In addition, as AI becomes increasingly incorporated into our everyday lives, the risk of AI harm increases every day. The following three recommendations for responsible AI use must be followed to reduce harm:

I. Organizations should delay development and only deploy once verification of safety has been completed and full transparency exists regarding the data and algorithms used.

II. An independent watchdog should monitor the ethical use of AI, as the actions of many tech companies who have laid off ethical researchers illustrate that monitoring by the internal corporate system is not reliable when an internal conflict of interest exists.

III. Organizations must consider the current and emerging regulatory environment for AI. AI regulations are currently being developed globally, from the European Union's AI Act, which classifies AI systems by risk, to U.S. states' restrictions on the use of facial recognition to new laws in China for the management of data. To

properly deploy AI for the benefit of society, all organizations deploying AI must comply with existing laws and support the direction of future laws.

Summary of the Eight Questions Answered

Organizations across industries will need to deal with how AI fits into their business and operations. This may lead to a major change in how certain businesses operate while for others it represents an expansion of capacity and volume. Regardless of industry, the expectation for the organization is similar - they must evaluate their level of readiness to develop and apply AI responsibly, recognizing the ethical responsibilities associated with deploying it to employees, customers, and other community members.

AI Governance, Risk Management, and Compliance (GRC) Platform

The AI GRC Assessment Model and Framework developed by Sheikh (2025) provides every organization the ability to assess its governance status, risk posture, and compliance preparedness for traditional governance and compliance mechanisms, AI, Generative AI, and machine learning (ML) systems. This model has been designed to be systems agnostic and industry agnostic. In their 2025 AI C-Suite Toolkit, the World Economic Forum recommends that organizations establish structured and cross-functional governance structures that address AI across all aspects of governance, including AI strategy, risk, compliance, and ethics. Therefore, the five modalities detailed in this model should be implemented at each stage of the organization's overall governance and organizational development efforts.

Each of the five modalities of governance setting out a unique measure of how well an organization is governed include:

1. Governance and regulatory compliance
2. Responsible AI, data privacy, information security, and intellectual property
3. Information integrity and risk management
4. AI ethics and transparency
5. Observability and ongoing monitoring

The preliminary AI risk assessment and compliance evaluation questionnaire consists of 5 levels of measurement, which can be classified from 1 through to 5 (where 1 = Strongly Agree and 5 = Strongly Disagree). Organizations that achieve scores greater than 3 as part of this survey present material compliance and risk issues that will need immediate attention and remediation. More importantly, the model does not simply identify issues

but rather has four integrated operational outputs to assist in converting the assessment findings into actionable governance activities: 1) AI Risk Assessment and Mitigation - This process evaluates AI risk within five modalities; generates risk mitigation strategies; and creates an execution strategy for mitigating risk (remediation). 2) Regulatory Compliance Gap Analysis and Remediation Roadmap - This process evaluates and documents the organization's compliance against applicable regulation(s) as well as generates a 30, 60, 90, and 180-day plan for addressing the compliance gaps identified. 3) ROI and Cost Benefit Analysis - This process documents the financial impact of identified AI risks and the cost-effectiveness of any mitigation strategy identified enabling the organization to make data driven governance investment decisions. 4) Executive Dashboard - This process provides organizations with the ability to gain visibility into their AI risk and compliance related status across their enterprise and provides executive management the necessary tools to monitor ongoing AI governance oversight. A sample of the questionnaire used for financial services is placed at **Exhibit "A"**. The AI GRC Assessment Model also has a free and online version for 12 different industries such as healthcare, financial, pharmaceutical & biopharmaceutical, manufacturing, and government which can be accessed from www.gaix.ai.

The Evolving Regulatory Landscape for AI

National Institute of Standards and Technology (NIST) AI Risk Management Framework (RMF)

The NIST AI RMF was released on January 26, 2023, and lays out the four main functions of AI governance: **Governance, Mapping, Measuring, and Managing.** The Governance function creates organization-wide policies, culture, and roles for responsible use of AI, while the Mapping function identifies the context and specific risks associated with a given AI system. The Measuring function analyzes those identified risks and quantifies them relative to the defined risks. Finally, the Managing function establishes continual monitoring of risk and provides guidance for updating mitigation plans emerged on throughout the lifecycle of the AI application. Once the AI RMF was published, NIST developed an additional release in July 2024 called the Generative AI Profile, which addresses the risks associated with large language models.

General Data Protection Regulation (GDPR)

Since May 2018, the EU has operated under the General Data Protection Regulation (GDPR). The GDPR places significant obligations on any AI system that processes personal data. Specifically, Article 22 gives individuals the right not to be subjected to an

automated decision that has a significant impact on them. Article 25 establishes the principle of Privacy by Design (PbD), which requires that data protection must be embedded into an AI system from the very beginning of its design. Articles 13 and 14 of the GDPR require the disclosure of the logic behind automated decision-making processes. Organizations that do not comply with GDPR risk being fined up to 4% of their annual global revenue — with financial services, healthcare, and HR being the most commonly impacted by the GDPR.

European Union Artificial Intelligence Act (EU AI Act)

The European Union approved the Artificial Intelligence Act; this is the world's first comprehensive legal framework for artificial intelligence. The European Parliament approved it on March 13, 2024, and the Council formally adopted it on May 21, 2024. The EU Artificial Intelligence Act was published in the Official Journal on July 12, 2024, and entered into force on August 1, 2024, with most provisions becoming applicable as of August 2, 2026. Any organization that violates a provision classified as "highly significant" under the Act — such as deploying a prohibited AI practice — could be fined up to €35 million (approximately $40 million USD) or 7% of its total worldwide annual revenue; whichever amount is greater (European Parliament & Council of the European Union). The EU AI Act divides AI systems into one of four classifications:

1. **Unacceptable Risk** (outright banned applications, such as government social scoring systems)
2. **High Risk** (mandatory compliance assessments, CE (Conformité Européenne) marking, and human oversight, such as credit scoring, medical devices, and biometric identification)
3. **Limited Risk** (disclosure requirements, such as chatbot functionality)
4. **Minimal Risk** (no mandatory guidelines).

Organizations that use AI systems are required to create an AI portfolio and map their AI systems to one of the four classifications under the AI Act. An AI system classified as a High-Risk AI system under the EU AI Act automatically receives a score of 5 on the AI-CARVER Regulatory Exposure dimension.

The AI-CARVER Tool and the GAIX AI GRC Assessment Model (Exhibit A) are specifically designed to enable the application of the three leading global AI governance frameworks, which include: NIST AI Risk Management Framework (NIST AI RMF); the European Union Artificial intelligence Regulation (EU AI Act); and General Data Protection

Regulation (GDPR) – within a single comprehensive assessment. It allows organizations to simultaneously map their AI systems based on the four Governance functions as defined by the NIST AI RMF (Govern, Map, Measure and Manage) and under the EU AI Act's four (4) tier risk hierarchy of Unacceptable, High, Limited and Minimal Risk, and compare their data processing activities with the GDPR's requirements. The GAIX AI GRC Assessment Model uniquely converts compliance results into executable governance action: assesses AI risks and provides actionable mitigation strategies; provides a compliance gap analysis and produces a 30/60/90/180-day remediation roadmap; performs ROI and cost-benefit analysis to quantify the financial implication of AI risk decisions; and creates an executive dashboard that provides organization leaders real-time visibility into enterprise-wide AI risk and compliance status across all three (3) frameworks. This level of integrated cross-framework compliance coverage and operational execution through a single assessment model represents a significant shortfall in existing AI governance practices.

Cross-Sector Specificity

While the regulatory obligations, shown in Table 4, and AI-CARVER risk profiles of organizations vary across industries, the differences can be meaningful, as highlighted in Table 5.

Table 4: Regulatory Frameworks Summary

Framework	Jurisdiction	Key AI Obligation
NIST AI RMF	US (voluntary)	Govern-Map-Measure-Manage lifecycle; seven trustworthy AI characteristics
GDPR	EU	Lawful basis for data processing; right to explanation (Art. 22); Privacy by Design (Art. 25); fines up to 4% global revenue
EU AI Act	EU	Four risk tiers; conformity assessments and CE marking for high-risk AI; prohibited systems banned outright
HIPAA	US – Healthcare	Protection of patient health data used in AI-driven diagnostics or clinical decision support
CCPA/CPRA	California, US	Consumer opt-out of automated profiling; disclosure of AI decision logic
ISO 42001	International	Certification standard for AI management systems; governance and continual improvement

For example, a bank that is implementing an AI-based credit scoring system must comply with GDPR Article 22 (i.e., right to explanation). It may be classified High-Risk according

to the EU's AI Act and adhere to Basel III model risk requirements. Therefore, AI-CARVER can show a Regulatory Exposure Score (Article 22, EU AI Act High-Risk) & Explainability Score of between 4 & 5, respectively. On the other hand, a pharmaceutical company developing a new drug through AI must adhere to the FDA AI/ML guidance as well as the requirements set forth under the GDPR. However, this organization normally achieves a lower Recoverability Score due to the presence of human researchers as strong backup assistance. Manufacturing organizations that utilize AI-driven predictive maintenance must adhere to ISO 42001's safety system requirements and EU's AI Act, with the highest overall risk dimensions being Vulnerability and Recoverability.

Table 5: Sector-Specific AI Risk and Regulatory Snapshot

Sector	Typical AI Use Case	Primary Regulation	Key AI-CARVER Hot Spot
Banking	Credit scoring, fraud detection	GDPR Art. 22; CCPA; Basel III model risk	Regulatory Exposure (5); Explainability (4–5)
Healthcare	Diagnostic imaging, clinical AI	HIPAA; FDA 510(k); EU AI Act (high-risk)	Regulatory Exposure (5); Recoverability (1–2)
Pharma	Drug discovery, trial analysis	FDA AI/ML guidance; GDPR; EU AI Act	Accuracy (3–4); Regulatory Exposure (4–5)
Manufacturing	Predictive maintenance, QA	ISO 42001; EU AI Act (safety systems)	Vulnerability (2–3); Recoverability (2–4)

Integrating Responsible AI into Project Management Practice

Project Managers are the glue that holds the responsible AI piece of the puzzle. They establish the scope of AI initiatives, create and use risk registers while executing and guiding the project through the governance gates. According to Kiron et al. (2022), successful AI governance requires the executive leadership of an organization to prioritize governance as a strategic focus for an organization, and therefore project managers are uniquely positioned to achieve successful governance through proper oversight at all stages of an AI project. Table 6 illustrates responsible AI activities and frameworks aligned to all phases of the project lifecycle.

Table 6: Responsible AI Integration Across the Project Lifecycle

Project Phase	Responsible AI Action	Key Tool / Framework
Initiation	Run AI-CARVER assessment; classify risk level; confirm regulatory obligations before scope is finalized.	AI-CARVER; AI GRC Model for Risk Assessment, Mitigation Strategies, and identification of compliance gaps; EU AI Act risk tiers; NIST AI RMF — Govern, Map, Measure, and Manage
Planning	Define data governance strategy; assign ethics reviewers in RACI; document training data provenance; embed GDPR Art. 25 as a deliverable.	AI GRC Model for an Action Plan to mitigate the risks and compliance gap analysis; GDPR Art. 25; ISO 42001
Execution	Implement explainability mechanisms; conduct adversarial and bias testing; apply privacy-by-design controls; run mid-project GRC health check.	NIST AI RMF — Map & Measure; AI-CARVER Vulnerability & Explainability; AI GRC Model to execute risk mitigation strategies and compliance gaps closure roadmap
Monitoring & Control	Track model drift; re-score AI-CARVER at each milestone; report risk status to steering committees; maintain risk register thresholds.	AI-CARVER Recoverability; NIST AI RMF — Manage; AI GRC Model for continuous monitoring & evaluation
Closure	Archive TEVV records; conduct ethics-focused lessons learned; update organizational AI policies based on deployment outcomes.	NIST AI RMF Playbook; ISO 42001 continual improvement; AI GRC Model to mitigate the risks and close the compliance gaps

Illustrative Scenario: Mid-Planning AI-CARVER Escalation

A financial services project manager is currently delivering an AI-powered loan approval system. While the team is in the Planning phase, they can run AI-CARVER to evaluate their initiative against its 5 criteria, resulting in the following rating: Criticality 4, Accuracy 3, Regulatory Exposure 5 (GDPR Art. 22; EU AI Act High-Risk), Vulnerability 3, Explainability 5 (proprietary black-box model), and Recoverability 2 — Total (High Risk): 22/30. Per Table 3, this will require C-Suite approval prior to execution. The Project Manager will escalate to the Chief Risk Officer, include an Explainability Sprint in the project plan, appoint a GDPR legal reviewer as RACI owner, and re-score AI-CARVER

every quarter. The AI GRC Assessment Model is utilized as a formal gate prior to entry into execution. As outlined, the AI-CARVER and AI GRC model provide natural integration of these frameworks into the standard project governance processes — not as superfluous bureaucracy, but as structured and defensible decision points that protect the organization and its customers.

Conclusion

The frameworks outlined in this paper establish a cohesive and practical governance stack that is immediately actionable. AI-CARVER converts AI-related risks into board-reportable, quantified scores. The regulatory environment of the NIST AI RMF, GDPR, and the EU AI Act now imposes real compliance obligations that vary by business sector and must be clearly mapped. The AI GRC Assessment Model determines AI and traditional risk metrics as well as compliance gaps. It can also be leveraged to develop comprehensive risk mitigation strategies, 30/60/90/180-day action plans, and remediation roadmaps for developing compliance gap closure prior to any potential violation. Finally, the incorporation of these principles through a project lifecycle can translate the responsible AI principles into every AI project and program. Organizations that have AI governance embedded within their operating models will experience the greatest benefits from scaling AI technology across their enterprises compared to organizations that only treat governance as a compliance obligation (Iansiti & Lakhani, 2020).

Your organization will be implementing some sort of AI solution in the coming months and years. The real question is: will your organization have a framework in place to govern and manage AI to ensure things don't go wrong? As demonstrated by Anthropic and other leaders in the AI safety space, safety and innovation do not need to be competing priorities. Ghosh & Bagai (2024) provide an example of how Anthropic's governance-first approach to developing and deploying AI provides organizations with a governance-first model to develop, use and promote AI technologies. Sheikh (2025) not only gives organizations more assistance regarding AI governance frameworks, compliance plans, and establishing enterprise-wide frameworks for AI governance, but he also provides them with useful practical examples from the industry on how to create and implement these frameworks safely and responsibly. Organizations that incorporate governance into the project management lifecycle can implement AI solutions quicker than those who do not. They can earn greater stakeholder trust and significantly reduce risk exposure while closing the compliance gaps and avoiding heavy penalties.

Exhibit A
AI GRC Assessment Model
Sample For Financial Services — 15 Questions

ID	Question	Strongly Agree (1)	Agree (2)	Don't Know (3)	Disagree (4)	Strongly Disagree (5)
1.1 Governance and Regulatory Compliance						
Q1	**Regulatory Compliance for Financial Services AI Systems:** Our organization is aware of all applicable AI, Generative AI, and Machine Learning (ML) laws and regulations related to anti-money laundering (AML) monitoring, high-frequency trading, and customer identity verification.					
Q2	**Credit and Risk Management Validation:** We have established documented legal compliance procedures ensuring that AI, Gen AI, and Machine Learning systems used in credit scoring, fraud detection, and automated investment advisory meet all applicable regulatory requirements and governance standards.					
Q3	**Algorithmic Bias, Fairness, and Impact Assessments:** We regularly conduct algorithmic bias assessments and fairness evaluations for AI, Gen AI, and ML systems used in credit scoring, fraud detection, loan underwriting, and automated investment advisory to ensure equitable and non-discriminatory outcomes.					
1.2 Responsible AI, Data Privacy, Information Security, and Intellectual property						
Q4	**Data Transparency and Documentation:** We document the source, history, and intended use of training data and generated					

	data for AI, Gen AI, and ML systems used in our financial services operations, while protecting proprietary methods and intellectual property.					
Q5	**Responsible AI Framework:** We have embedded Responsible AI principles into our policies to ensure AI, Gen AI, and ML systems are fair, transparent, interpretable, accountable, and safe across financial services operations.					
Q6	**Data Security and Information Integrity Framework:** We have instituted a comprehensive framework to evaluate and protect the security and integrity of data used by AI, Gen AI, and Machine Learning systems in our banking operations, including safeguards against unauthorized access, data breaches, and intellectual property violations.					
1.3 Information Integrity and Risk Management						
Q7	**Risk Tier Classification Framework:** When defining risk tiers for AI, Gen AI, and ML systems in financial services, we consider: 1) potential fraud and data misuse, 2) impacts on information integrity, 3) AI system interdependencies with trading and banking platforms, 4) risks to customer rights and financial stability, and 5) potential malicious use that could disrupt financial operations.					
Q8	**Deployment Approval Standards:** We have established minimum performance and assurance standards that are evaluated as part of our deployment approval (go/no-go) policies and procedures for financial services AI systems.					

Q9	**Pre-Deployment Risk Tolerance Validation:** Before deploying AI, Gen AI, and Machine Learning models within our financial services, we have developed a comprehensive Testing, Evaluation, Validation, and Verification (TEVV) plan that validates system outputs for reliability, consistency, and performance against our organization's defined risk tolerance thresholds and information integrity standards.					
1.4 AI Ethics and Transparency						
Q10	**Content and Output Compliance Controls:** We have implemented policies and mechanisms to ensure that AI, Gen AI, and Machine Learning systems used in our financial services operations do not generate content or outputs that violate financial regulations or best practices.					
Q11	**AI Ethics Policies and Commitments:** We have established a formal AI ethics policy that defines our ethical principles for the use of AI, Gen AI, and ML systems in financial services operations.					
Q12	**AI Model Explainability and Decision Transparency:** We have established policies and procedures to ensure that AI, Gen AI, and ML systems can explain their decisions, outputs, and recommendations in a clear and interpretable manner to stakeholders, regulators, and affected individuals, in accordance with applicable explainability and transparency standards across financial services operations.					
1.5 Observability and Ongoing Monitoring						
Q13	**Data Provenance and Incident Monitoring:** We have clearly defined					

	organizational responsibilities for regularly reviewing data provenance and monitoring incidents related to AI, Gen AI, and ML systems in financial services operations, including validating the origin, authenticity, and history of financial data, models, and system inputs.					
Q14	**Incident Response and Gap Analysis:** We have established comprehensive policies for evaluating and responding to incidents involving AI, Gen AI, and ML systems in financial services. Our processes include gap analyses to identify root causes and are regularly updated to align with financial industry regulations and disclosure requirements.					
Q15	**Document Retention and Audit Trail Monitoring:** We have implemented a scheduled document retention policy with clearly defined organizational roles and responsibilities for preserving and regularly reviewing historical records related to testing, evaluation, validation, and verification (TEVV) of AI, Gen AI, and Machine Learning systems to support continuous risk oversight.					

References

1. Deloitte. (2026). *The state of AI in the enterprise 2026.* Deloitte Insights. https://www.deloitte.com/us/en/what-we-do/capabilities/applied-artificial-intelligence/content/state-of-ai-in-the-enterprise.html

2. Larridin. (2026). *2026 state of enterprise AI report*. Larridin. https://www.businesswire.com/news/home/20260203918939/en/New-Study-Shows-C-Suite-Leaders-Highly-Confident-in-AI-ROI-Even-as-58-Claim-Theres-No-Clear-Ownership-of-AI-and-75-Lack-AI-Governance

3. McKinsey & Company. (2026). *Responsible AI: Overcoming adoption barriers and risks—Findings from McKinsey's 2026 AI Trust Maturity Survey.* McKinsey & Company.

https://www.mckinsey.com/capabilities/tech-and-ai/our-insights/tech-forward/state-of-ai-trust-in-2026-shifting-to-the-agentic-era

4. AuditBoard. (2025). *From blueprint to reality: Execute effective AI governance in a volatile landscape*. AuditBoard. https://www.prnewswire.com/news-releases/new-research-finds-only-25-percent-of-organizations-report-a-fully-implemented-ai-governance-program-302517095.html

5. Accenture. (2025). *Thrive with responsible AI: Embedding trust can unlock value.* Accenture Insights. https://www.accenture.com/us-en/insights/data-ai/rai-from-risk-to-value

6. Deloitte Global Boardroom Program. (2025). *Governance of AI: A critical imperative for today's boards* (2nd ed.). Deloitte Global. https://www.deloitte.com/global/en/issues/trust/progress-on-ai-in-the-boardroom-but-room-to-accelerate.html

7. Sheikh, R. A. (2025). AI governance and frameworks: How to manage AI risks and compliance. *PM World Journal, XIV*(VII). Originally presented at the 17th Annual UT Dallas Project Management Symposium, Naveen Jindal School of Management, Richardson, Texas, May 30, 2025. https://pmworldlibrary.net/wp-content/uploads/2025/07/pmwj154-Jul2025-Sheikh-AI-Governance-and-Frameworks.pdf

8. Singla, A., Sukharevsky, A., Yee, L., Chui, M., Hall, B., & Balakrishnan, T. (2025, November). *The state of AI in 2025: Agents, innovation, and transformation.* QuantumBlack, AI by McKinsey. https://www.mckinsey.com/capabilities/quantumblack/our-insights/the-state-of-ai

9. World Economic Forum. (2025). *Empowering AI leadership: AI C-suite toolkit.* World Economic Forum.

10. Accenture. (2024). *Rethinking responsible AI: From compliance to confidence.* Accenture. https://www.accenture.com/us-en/insights/data-ai/compliance-confidence-responsible-ai-maturity

11. European Parliament & Council of the European Union. (2024). *Regulation (EU) 2024/1689 of the European Parliament and of the Council laying down harmonised rules on artificial intelligence (Artificial Intelligence Act).* Official Journal of the European Union.

12. Ghosh, S., & Bagai, S. (2024). *Anthropic: Building safe and powerful AI* (HBS Case No. 824-129). Harvard Business School Publishing.

13. KPMG LLP. (2024a). *Driving responsible innovation: Reflections on a year of AI governance.* KPMG LLP. https://kpmg.com/us/en/articles/2024/driving-responsible-innovation-reflections-ai-governance.html

14. KPMG Australia. (2024b). *Trusted AI governance.* KPMG. https://assets.kpmg.com/content/dam/kpmgsites/au/pdf/2024/kpmg-trusted-ai-governance-2024.pdf.coredownload.inline.pdf

15. Sheikh, R. A., Jarvis, R., Whitehall, J., & Jawad, F. (2024). Managing projects successfully through artificial intelligence (AI) and ChatGPT. *PM World Journal, XIII*(IX). Originally presented at the 16th Annual UT Dallas Project Management Symposium, Naveen Jindal School of Management, University of Texas at Dallas, Richardson, Texas, May 20–21, 2024. https://pmworldlibrary.net/wp-content/uploads/2024/09/pmwj145-Sep2024-Sheikh-et-al-Managing-Projects-Successfully-through-AI-and-ChatGPT.pdf

16. CallMiner. (2024). *Survey: 96% of CX orgs view AI as a key strategy*. CallMiner. https://callminer.com/blog/survey-96-of-cx-orgs-view-ai-as-a-key-strategy

17. National Institute of Standards and Technology. (2023). *Artificial intelligence risk management framework (AI RMF 1.0)* (NIST AI 100-1). U.S. Department of Commerce. https://doi.org/10.6028/NIST.AI.100-1

18. Neeley, T. (2023, May 9). 8 questions about using AI responsibly, answered. *Harvard Business Review.* https://hbr.org/2023/05/8-questions-about-using-ai-responsibly-answered

19. Kiron, D., Renieris, E., & Mills, S. D. (2022, April 19). Why top management should focus on responsible AI. *MIT Sloan Management Review.* https://sloanreview.mit.edu/article/why-top-management-should-focus-on-responsible-ai/

20. MIT Work of the Future Task Force. (2022). *Artificial intelligence and the future of work.* Massachusetts Institute of Technology.

21. Neeley, T., & Leonardi, P. (2022). *The digital mindset: What it really takes to thrive in the age of data, algorithms, and AI.* Harvard Business Review Press.

22. Bender, E. M., Gebru, T., McMillan-Major, A., & Mitchell, M. (2021). On the dangers of stochastic parrots: Can language models be too big? In *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency* (pp. 610–623). ACM. https://doi.org/10.1145/3442188.3445922

23. Davenport, T. H., & Zhang, R. (2021, July 20). Achieving return on AI projects. *MIT Sloan Management Review.* https://sloanreview.mit.edu/article/achieving-return-on-ai-projects/

24. Economist Intelligence Unit. (2020, October). *Staying ahead of the curve: The business case for responsible AI.* EIU. https://pages.eiu.com/rs/753-RIQ-438/images/EIUStayingAheadOfTheCurve.pdf

25. Iansiti, M., & Lakhani, K. R. (2020). *Competing in the age of AI: Strategy and leadership when algorithms and networks run the world.* Harvard Business Review Press.

26. Smith, G., & Rustagi, I. (2020). *Mitigating bias in artificial intelligence.* Berkeley Haas EGAL.

27. Bencie, L., & Araboghli, S. (2018, September 21). A 6-part tool for ranking and assessing risks. *Harvard Business Review.* https://hbr.org/2018/09/a-6-part-tool-for-ranking-and-assessing-risks

28. Buolamwini, J., & Gebru, T. (2018). Gender shades: Intersectional accuracy disparities in commercial gender classification. In *Proceedings of the 1st Conference on Fairness, Accountability and Transparency* (pp. 77–91). PMLR.

29. European Parliament & Council of the European Union. (2016). *Regulation (EU) 2016/679 of the European Parliament and of the Council on the protection of natural persons with regard to the processing of personal data (General Data Protection Regulation).* Official Journal of the European Union.

30. Cavoukian, A. (2011). *Privacy by design: The 7 foundational principles.* Information and Privacy Commissioner of Ontario.

About the Authors

Dr. Rizwan A. Sheikh, PhD

Author, Inventor, Professor
AI Governance Expert
Texas, USA

Dr. Rizwan A. Sheikh ("Dr. Riz") is the Founder & CEO of Global AI Excellence (GAIX) and inventor of an AI Governance Model (patent pending). An AI strategist with over 30 years of experience, he specializes in AI governance, risk management, compliance, and digital transformation. As a former Deloitte executive, Dr. Riz has delivered over $300 million in cost savings for Fortune 500 companies. He has taught AI strategy and project management in executive programs at Harvard, MIT, and the University of Cambridge. He holds a Ph.D. from SKEMA Business School.
LinkedIn: linkedin.com/in/drrizwanasheikh

Khalid Ahmad Khan, PhD, PMP

Professor of Project Management and Governance,
Chief Data Scientist, AI Researcher
Lahore, Pakistan

Dr. Khalid Ahmad Khan is a professor, distinguished data scientist, AI researcher, and governance expert with over 30 years of experience spanning project management, AI, data science, risk management, and digital transformations across public and private sectors. Dr. Khan specializes in deploying AI-enabled analytics solutions for monitoring, document analysis, and executive decision-making. He advises governments on public investment management, portfolio planning, and project lifecycle governance, applying

advanced decision-support tools including Monte Carlo simulations, real options analysis, and Earned Value Management. A global advocate for project management excellence, Dr. Khan is a founding member and President of the PMI Lahore Chapter and contributed to the PMI PMBOK Government Extension, Third Edition. His key clients include the World Bank, USAID, GIZ (German Aid Agency), and LUMS.

He holds a Ph.D. in Strategy and Project Management and is PMP certified by PMI.

Aamir Khalid Pirzada

Chief Information Officer
AI/Manufacturing Technology Leader
Al Khobar, Saudi Arabia

Aamir Pirzada is a transformative technology executive whose pioneering work in AI and machine learning has revolutionized manufacturing operations. As a seasoned Chief Information Officer (CIO), he has led organizations across manufacturing, oil and gas, retail, public services, and telecommunications industries.

Aamir specializes in leveraging AI-driven solutions for predictive maintenance, quality control, and data-driven decision-making—consistently delivering enhanced operational efficiency, significant cost savings, and superior product quality. His expertise spans enterprise system implementations including SAP S4HANA, Oracle e-Business Suite, and Microsoft Dynamics AX.

A recognized leader in cybersecurity, Aamir has established security protocols and led audits to safeguard critical manufacturing assets. His strategic vision centers on integrating AI, machine learning, and IoT to create smart, adaptive manufacturing ecosystems. With deep expertise in project management, business analysis, and solution design, Aamir continues to guide organizations toward technological excellence in the rapidly evolving digital landscape.