AI Risk Evaluation (AIRE) Framework

Modeled on the EU AI Act

LIMRA LOMA
Navigate With Confidence

0936-2025

Author**:**
Kartik Sakthivel, Ph.D., MS-IT/MS-CS, MBA, PGC-IQ
Former Vice President and Chief Information Officer, LIMRA and LOMA
Now CIO John Hancock

Created in collaboration with the LIMRA and LOMA AI Governance Group

CONTENTS

1.0 Overview

The financial services sector is undergoing a dramatic and transformative evolution driven by advancements in artificial intelligence (AI) and generative AI (GenAI) as a derivative of the overall AI industry. The introduction of AI across the value chain is not just enhancing existing processes, but it is also fundamentally reshaping the life insurance landscape. By capitalizing on AI — correctly — insurers can achieve significant competitive advantages, drive innovation, and better serve a new generation of customers in a digital world.

The continuously accelerating pace of AI evolution presents unlimited potential to transform the insurance industry and society overall. However, these rapid advancements can also usher in a new set of risk management challenges. AI systems can be notoriously "black box" in nature — that is, there is little insight into how AI systems arrive at the outputs, outcomes, and decisions they make. Additionally, AI models that can potentially perform trillions of computations per second can be difficult for their human creators and overseers to explain to their stakeholders. This lack of transparency and explainability undermines trust in AI and can impede the successful adoption of these transformative technologies within firms. AI models and data can also be susceptible to inadvertent bias and/or proxy discrimination resulting from models establishing erroneous causality and correlation to inherent bias within the underlying data.

Within the insurance industry, one that prioritizes ethical concerns over methodological concerns when it comes to AI, this can pose immense risks. Should AI systems that are being used within the insurance value chain be afflicted by these problems, it can lead to significant loss of goodwill, create regulatory issues, and erode customer trust. As the industry continues its mission to financially protect underserved communities and deliver affordable protection to the uninsured and underinsured, missteps in AI implementations can have far-reaching implications.

To address the potential harm that AI systems can cause, the European Union (EU) drafted and adopted a landmark regulation known as the EU Artificial Intelligence Act (AI Act). This is the first AI legislation of its kind, and it establishes a common regulatory and legal framework for all AI systems developed and/or delivering services and/or operating within the EU. The AI Act went into effect on August 1, 2024, with provisions being gradually implemented over the subsequent six to 36 months. The AI Act is predicated on the need to mitigate AI's risks, while allowing firms to continue innovating and adopting this transformative technology. Classifying AI systems into four distinct categories — unacceptable, high, limited, and minimal — the AI Act aims to set a global standard for AI regulation by tailoring requirements to each risk category of AI systems. This allows for an appropriate and proportionate approach to regulatory oversight, which includes accountability, transparency, testing and controls, data quality, reporting requirements, and risk management. While the AI Act primarily focuses on regulating what it terms as "high risk AI systems," the compliance requirements associated with each risk level provide solid guidelines for the insurance industry to emulate.

1.1 Need for Industry AI Governance

Risks from AI are different from the risks traditionally associated with software implementations. According to the National Institute of Standards and Technology (NIST), which published a comprehensive AI Risk Management Framework (AI RMF 1.0) in January 2023, "As with traditional software, risks from AI-based technology can be bigger than an enterprise, span organizations, and lead to societal impacts. AI systems also bring a set of risks that are not comprehensively addressed by current risk frameworks and approaches." (Tabassi, 2023) Tabassi highlights that current risk management frameworks are insufficient for dealing with AI risks. These frameworks do not adequately address harmful bias, security issues related to AI-specific attacks, or the complexities of AI systems. They also fall short in managing risks associated with third-party AI technologies and using AI in unintended ways. The report outlines a list of novel or increased AI-related risks in contrast with traditional software as below:

- The data used for building an AI system may not be a true or appropriate representation of the context or intended use of the AI system, and the ground truth may either not exist or not be available. Additionally, harmful bias and other data quality issues can affect AI system trustworthiness, which could lead to negative impacts.
- AI system dependency and reliance on data for training tasks, combined with increased volume and complexity typically associated with such data.
- Intentional or unintentional changes during training may fundamentally alter AI system performance.
- Datasets used to train AI systems may become detached from their original and intended context, or they may become stale or outdated relative to the deployment context.
- AI system scale and complexity (many systems contain billions or even trillions of decision points) housed within more traditional software applications.
- Use of pre-trained models that can advance research and improve performance can also increase levels of statistical uncertainty and cause issues with bias management, scientific validity, and reproducibility.
- Higher degree of difficulty in predicting failure modes for emergent properties of large-scale pre-trained models.
- Privacy risk due to enhanced data aggregation capability for AI systems.
- AI systems may require more frequent maintenance and triggers for conducting corrective maintenance due to data, model, or concept drift.
- Increased opacity and concerns about reproducibility.
- Underdeveloped software testing standards and inability to document AI-based practices to the standard expected of traditionally engineered software for all but the simplest of cases.
- Difficulty in performing regular AI-based software testing, or determining what to test, since AI systems are not subject to the same controls as traditional code development.
- Computational costs for developing AI systems and their impact on the environment and planet.
- Inability to predict or detect the side effects of AI-based systems beyond statistical measures.

President Biden's sweeping Executive Order on AI issued on October 30, 2023, established new standards for AI safety and security to protect Americans from the potential risks of AI systems. While the Executive Order outlines a comprehensive framework for AI safety, the United States does not have an overarching AI regulation or regulatory framework that is comparable to the EU AI Act. It is unlikely that the United States — which still lags behind the EU on data privacy regulation — will at the federal level draft and adopt AI legislation anytime soon.

For the insurance industry, where regulatory guidelines exist, they are germane to specific domains within the insurance value chain. One example is within the underwriting domain for automated and accelerated

underwriting, where regulatory guidelines, such as those from the National Association of Insurance Commissioners (NAIC) Advanced Underwriting Working Group (AUWG), Senate Bill 169 from the State of Colorado (SB-169), and Circular-19 from the State of New York (NY Circular-19), provide the industry with appropriate frameworks. However, the industry lacks regulation and regulatory frameworks around AI across the value chain. Firms within the insurance industry typically allow regulation to establish their operating guidelines. With no such regulation expected, it is incumbent upon the industry to agree to a common framework — a set of best practices and guidelines to base their AI governance frameworks upon.

These frameworks will be vital to help carriers mitigate AI risks by ensuring careful oversight of the design, development, and implementation of AI, as well as conducting detailed assessments of the ethical, legal, and societal implications of their AI systems. By taking a proactive approach, organizations can mitigate risks that include (but are not limited to) issues with inadvertent bias and proxy discrimination, data privacy and protection concerns, liability, and intellectual property challenges, and they can avoid reputational damage and financial impacts.

1.2 AI Risk Evaluation (AIRE) Across the Insurance Value Chain

The primary goal of any AI governance framework is to ensure that guardrails are in place such that AI systems employed by the industry prioritize safety and transparency, balanced with speed and innovation. Absent overarching regulation, it is vital for the entire industry to work from a common set of best practices that allows each firm the ability to achieve success with their AI implementations in the interest of serving their policyholders. These governance frameworks will establish basic parameters to help all stakeholders within the industry who are involved in AI implementations — including the development, distribution, usage, or manufacturing of AI systems -—— adhere to best practices.

A holistic approach is a necessary first step in establishing proactive, appropriate, and proportionate AI governance policies across the insurance value chain. This holistic approach commences with assessing where AI is currently being employed and where the industry might apply AI solutions in the future. This inventory of AI initiatives across a cross-section of the value chain (as of Q1 2025) is presented in the AI Risk Classification framework. These are illustrative examples only, which firms may use as a guide to categorize their own AI use cases. This is thematically similar to the EU AI Act classifications. Governance frameworks can then be developed that are proportionate to the classification of a specific implementation within a specific stratum.

2.0 AI Risk Classification Model

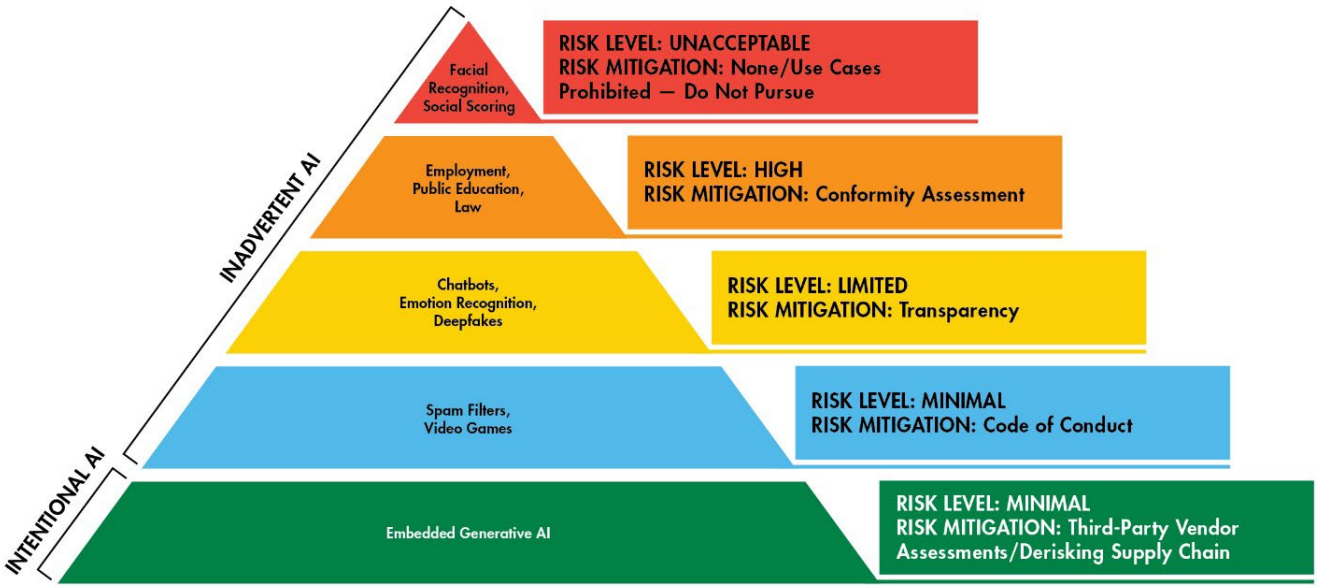
By and large, insurance firms are leveraging AI across the value chain in similar ways and/or with the same third-party vendors. Absent regulatory guidelines, it is up to each carrier to determine risk levels of each AI implementation. While some of these risk categorizations are determined by a carrier's risk tolerance and availability of requisite resources to mitigate these risks, the type and nature of risks for each AI implementation — just like with the implementations themselves — are not unique to each carrier. The EU AI Act requires organizations to inventory AI systems that are currently operational and in active development. This includes systems that firms develop themselves ("build") or procure from an external third-party provider ("buy"). Based on this inventory, firms are required to classify these AI models (systems) into one of four categories. Each of these categories requires a different level of scrutiny and governance.

Similarly, for the insurance value chain, this AI Risk Classification Model — aligned with the EU AI Act's categories — provides insurance companies a set of common guidelines of risk classification for most of the common AI implementations within the industry. Once you understand the AI Risk Classifications within our industry, the AI Risk Evaluation (AIRE) framework presented here will help your firms evaluate an AI initiative and determine the risk category for a specific AI initiative. From there, your organization can develop risk management strategies appropriate to that risk. (For example, it would not be prudent to apply the same level of risk management to an internal chatbot as an AI underwriting system.)

Note that each classification includes a few illustrative examples within our industry. A 2024 Bloomberg Law publication by Arsen Kourinian of Mayer Brown summarizes examples of some areas of unacceptable uses of AI, along with associated U.S. laws and tenets of the EU AI Act. AI Risk Classification framework provides a list of these examples that are based on this publication. However, as compared to the original publication, the framework seeks to focus on pertinent potential examples within the insurance value chain (those directly and tangentially applicable). For a full list, please refer to the original publication as listed in the References page. Note that these are illustrative examples only and are likely to evolve.

Figure 1 shows the AI Risk Classification Model, including risk categories, illustrative examples of AI implementations within each category, and steps to mitigate/minimize risk within each category. Section 2.1 explores each of these categories in further detail.

Figure 1: AI Risk Classification Model



2.1 AI Risk Classification Categories

The EU AI Act classifies AI systems into **four** risk categories:

1. **Unacceptable Risk (Prohibited)**
2. **High Risk**
3. **Limited Risk**
4. **Minimal or No Risk**

The AI Risk Classifications for insurers build upon the EU AI Act and classify AI systems into **five** risk categories, with four categories considered "**Intentional AI**" and an additional category considered **"Inadvertent AI."** This is unique to the AIRE and not part of the EU AI Act framework.

Intentional AI

1. **Unacceptable Risk (Prohibited)**
2. **High Risk**
3. **Limited Risk**
4. **Minimal or No Risk**

Inadvertent AI

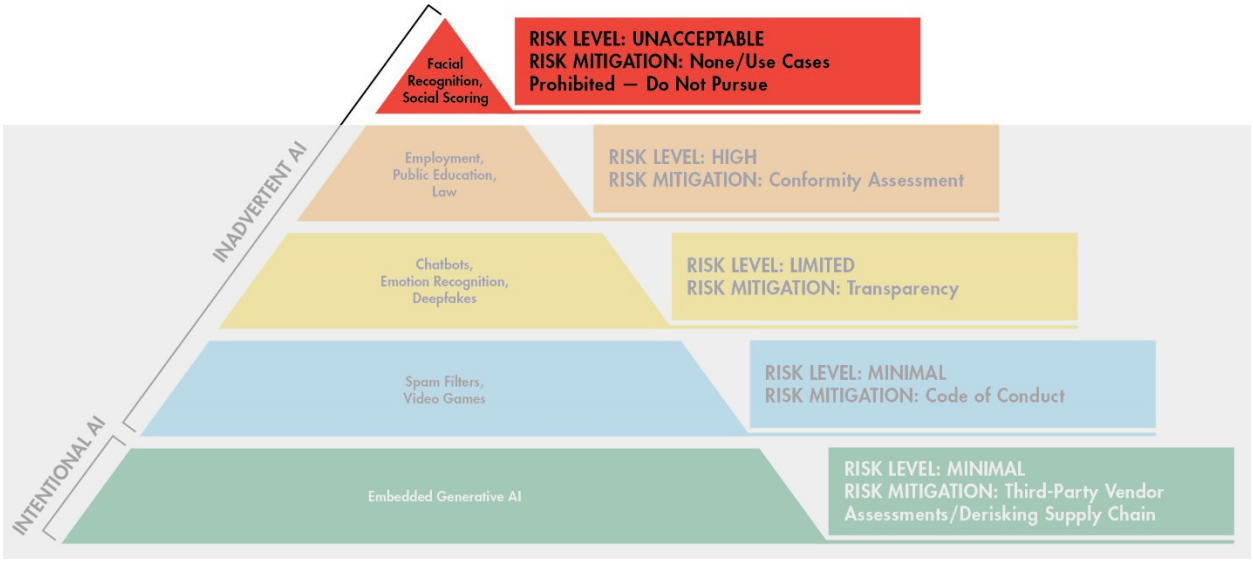
1. **Minimal Risk**

Intentional AI

As the name suggests, intentional AI systems are those that an enterprise invests in intentionally to solve business problems. Typically driven by a business need, as a part of a strategy, a need to innovate, and/or a need to experiment, firms can either "build" intentional AI systems or "buy" AI solutions (engage with external third-party providers to procure AI platforms/services).

1. Unacceptable Risk

Figure 2: Unacceptable Risk AI Systems



AI systems that are deemed an "unacceptable risk" are prohibited. This classification includes systems perceived to be a clear threat to people's safety and rights. Carriers should not consider these systems, whether via a build or a buy. Unacceptable risk systems within the insurance value chain are quite evident, for the most part, and they follow AI-based systems that are technological representations of business

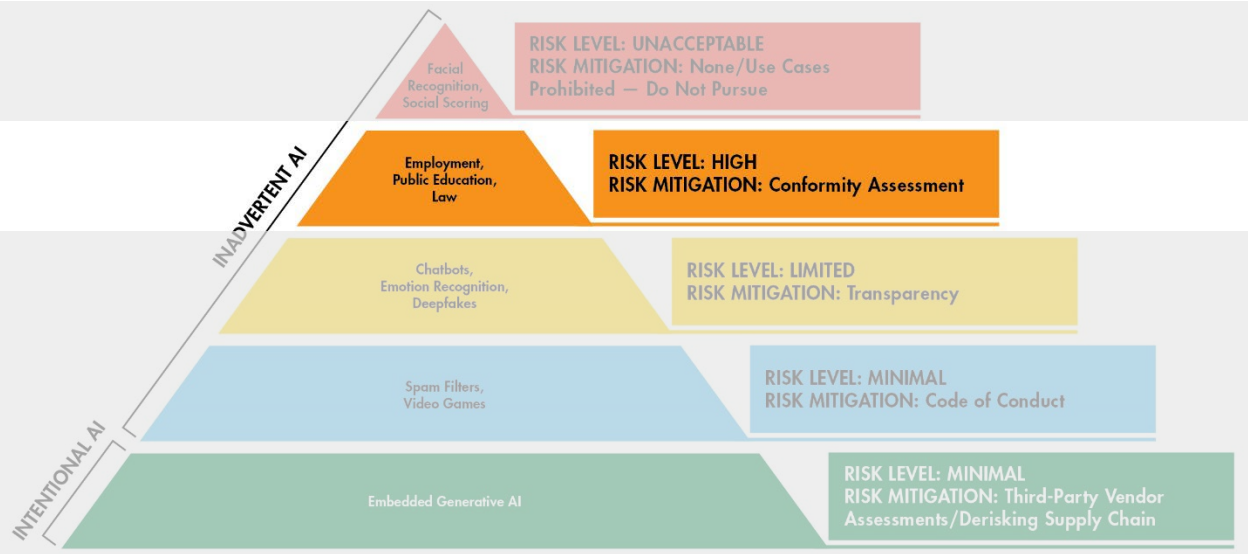
practices that one would not pursue via traditional, non-AI means. (One example would be using racial data for automated accelerated underwriting.) According to an EY article, "Examples include the use of real-time remote biometric identification in public spaces or social scoring systems, as well as the use of subliminal influencing techniques which exploit vulnerabilities of specific groups." (Meier & Spichiger, 2024)

Additional insurance industry examples include:

a. **Manipulative Practices:** AI systems designed to manipulate users' behavior in ways that cause physical or psychological harm. (For example, AI-driven sales techniques that exploit vulnerable customers to sell unnecessary or expensive insurance policies)
b. **Social Scoring:** AI systems that evaluate or score individuals based on their social behavior or characteristics, leading to unfair discrimination. (For example, using AI to score customers based on social media activity to determine eligibility or premium rates)
c. **Biometric Surveillance:** Real-time remote biometric identification systems in public spaces. (For example, using facial recognition AI to monitor and track individuals without consent in public or semi-public areas)

2. High Risk

Figure 3: High-Risk AI Systems



"High-risk" AI systems are those that can have potentially devastating effects on people's personal interests. These systems should be thoroughly evaluated before being implemented or used. Examples would include AI that evaluates resumes for recruitment purposes and AI used in medical procedures (such as AI-assisted surgery). These "high-risk" systems are permitted within the insurance value chain. However, aligning with the EU AI Act, they:

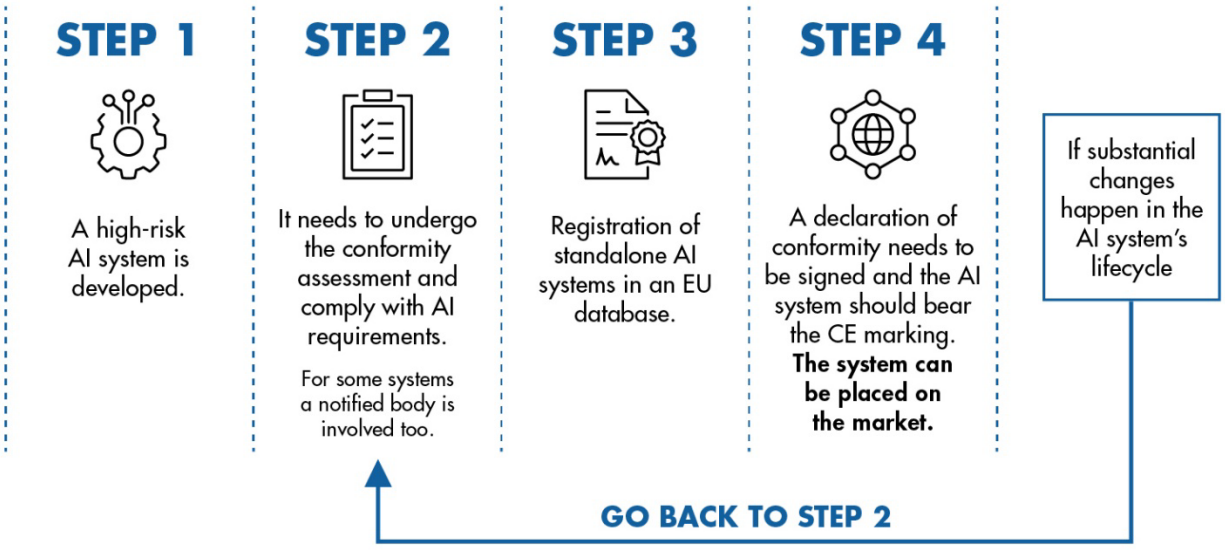
"Must comply with multiple requirements and undergo a conformity assessment. This assessment needs to be completed before the system is released on the market. Those systems are also required to be registered in an EU database, which shall be set up. Operating high-risk AI systems requires an appropriate AI risk management system, logging capabilities, and human oversight respectively ownership. There shall be proper data governance applied to the data used for training, testing, and validation, as well as controls

assuring cyber security, robustness, and fairness of the system. Examples of high-risk systems are those related to the operation of critical infrastructure, systems used in hiring processes or employee ratings, credit scoring systems, automated insurance claims processing, or setting of risk premiums for customers." (Meier & Spichiger, 2024)

Currently, the insurance industry does not have a centralized database that houses an inventory of "high-risk" systems, along with their commensurate conformity assessments. The latter promotes explainability and transparency, bestows accountability, and inspires trust in these systems.

Figure 4 outlines the process flow of how a "high-risk" AI system can be implemented within the parameters of the EU AI Act.

Figure 4: High-Risk AI System Flow – EU AI Act



(Source: The European Commission (EU, 2024))

Insurance industry examples include:

a. **Underwriting and Risk Assessment:**
    1. AI systems used to assess risk and determine premiums must ensure fairness, transparency, and non-discrimination. (For example, AI algorithms used for underwriting must be transparent and audited to ensure they do not unfairly discriminate against certain groups.)
    2. Requirements:
        1. Robust data quality and governance to prevent biased outcomes
        2. Clear documentation and traceability of AI decision-making processes
        3. Regular audits and monitoring for compliance with ethical and legal standards

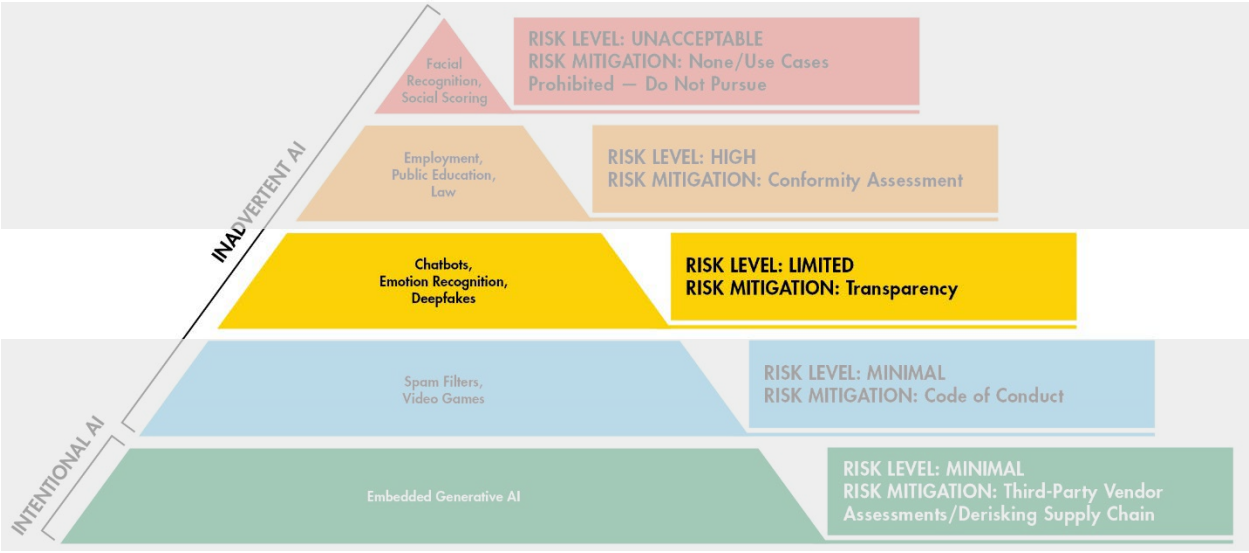
b. **Claims Processing:**
    1. AI systems automating claims processing must ensure accurate, fair, and transparent decisions. (For example, automated claims approval systems must be designed to flag suspicious or complex claims for human review.)

2. Requirements:
    1. Comprehensive data management practices to ensure accuracy and completeness of claims data
    2. Mechanisms for human oversight and intervention
    3. Transparency in AI decision-making and clear communication with policyholders

c. **Fraud Detection:**
    1. AI systems for detecting fraudulent activities must balance efficiency with privacy and fairness. (For example, AI-driven fraud detection systems should be transparent about the criteria used to flag suspicious claims and ensure they do not disproportionately target certain groups.)
    2. Requirements:
        1. Data protection measures to safeguard personal information
        2. Transparency in how AI identifies potential fraud
        3. Fairness and non-discrimination in AI fraud detection models

3. Limited Risk

Figure 5: Limited Risk AI Systems



"Limited risk" AI systems pose a slight risk, which can be managed by transparency obligations that allow users to make informed decisions — by being made aware they are interacting with AI-based systems and given the choice to opt out of using them. Similarly, within insurance, the risk can be mitigated by transparency, such as by notifying users that they are interacting with an AI system, such as a digital assistant or chatbot. According to EY, as it pertains to the EU AI Act, "Examples include chatbots or deep fakes, which are not considered high risk, but for which it is mandatory that users know about AI being behind it." (Meier & Spichiger, 2024)

Insurance industry examples include:

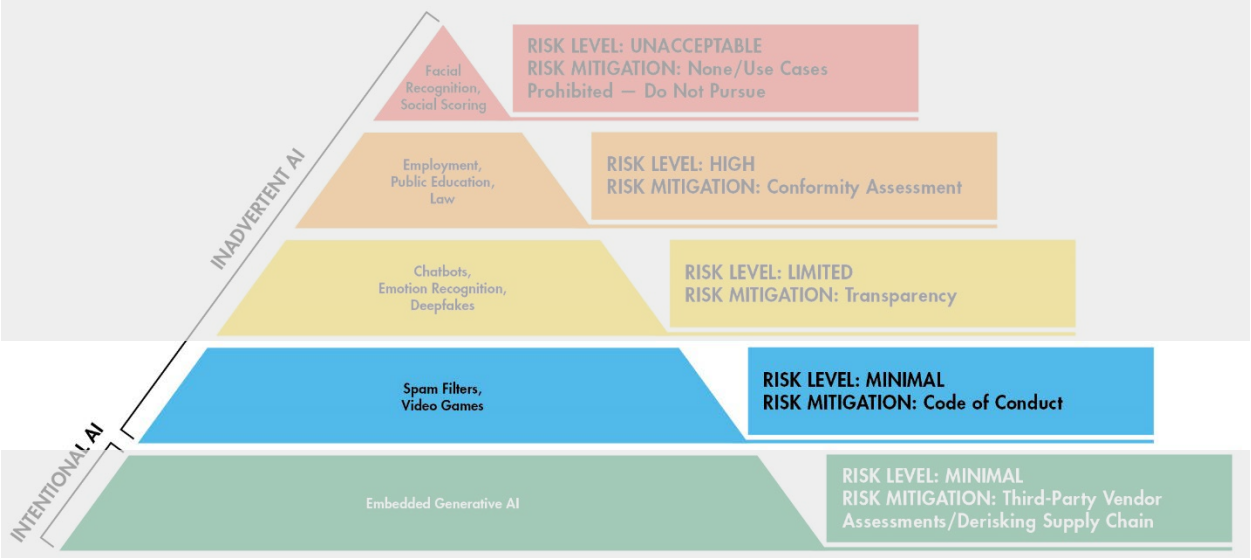
a. **Customer Service and Chatbots:**
    1. AI systems providing customer service must ensure clear communication and allow human intervention. (For example, AI chatbots must clearly identify themselves as AI and provide an easy option to speak with a human representative.)

2. Requirements:
    a. Inform customers they are interacting with an AI system.
    b. Provide options for human assistance when requested.
    c. Ensure AI systems are regularly updated to handle queries accurately.

b. **Personalized Marketing:**
1. AI systems used for personalized marketing must respect privacy and data protection regulations. (For example, AI-driven marketing campaigns should inform customers about data usage and obtain consent before using their data for personalized marketing.)
2. Requirements:
    a. Transparent data collection and usage policies
    b. Consent mechanisms for data usage in marketing
    c. Regular audits to ensure compliance with data protection laws

4. Minimal Risk

Figure 6: Minimal Risk AI Systems



"Minimal risk" systems present no risks to people's rights and safety, and using them poses little to no risk across the insurance value chain. This type of risk can be governed under a firm's Code of Conduct policies. As noted by EY, "For all operators of AI systems, the implementation of a Code of Conduct around ethical AI is recommended. Notably, general-purpose AI models (GPAI), including foundation models and generative AI systems, follow a separate classification framework. The AI Act adopts a tiered approach to compliance obligations, differentiating between high-impact GPAI models with systemic risk and other GPAI models." (Meier & Spichiger, 2024)

Insurance industry examples include:

1. **Data Analysis and Reporting:**
    a. AI systems used for internal data analysis and generating reports to improve business processes and decision-making. (For example, AI tools for analyzing customer demographics and trends to inform business strategies)
    b. Requirements:

i. Ensure data integrity and accuracy.
ii. Maintain internal documentation of AI processes.

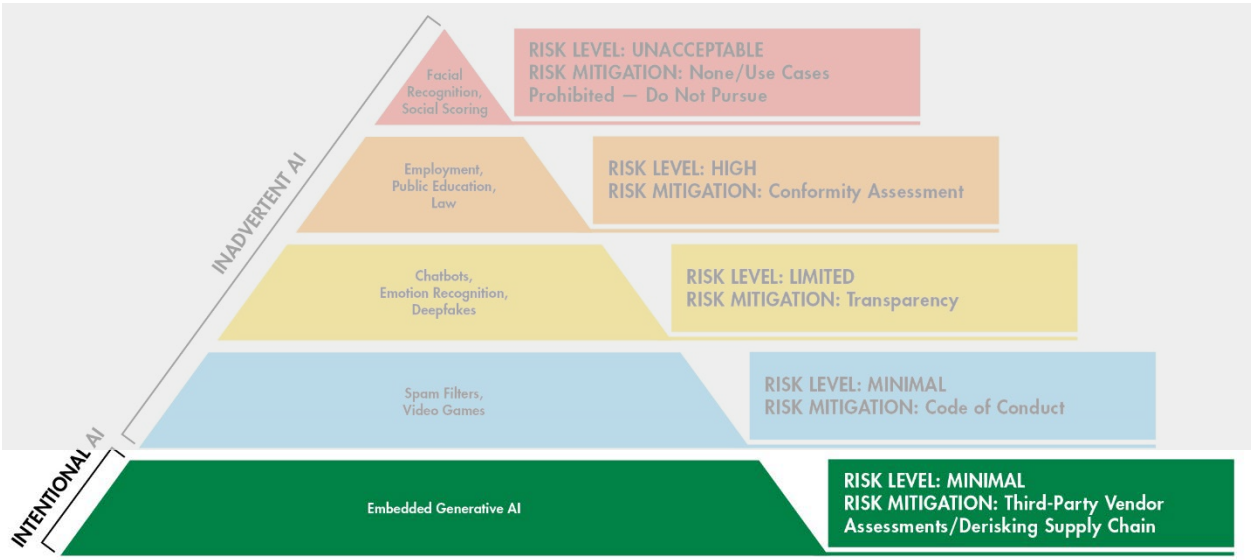
2. **Administrative Automation:**
    a. AI systems that automate routine administrative tasks, such as document management and scheduling. (For example, AI systems for automating document classification and retrieval in the underwriting process)
    b. Requirements:
        i. Regular maintenance and updates to ensure accuracy
        ii. Clear documentation of automated processes

Inadvertent AI

As the name suggests, enterprises can leverage inadvertent AI systems simply because they are embedded as functionality across commonly used software platforms. Typically, generative AI (GenAI) type functionality implementations are becoming increasingly ubiquitous across most software vendors, such as Microsoft (Copilot), Adobe, and Salesforce (Einstein). Regardless of whether a firm invests in intentionally building or buying AI systems, it is inevitable that most — if not all — firms will become AI consumers by virtue of inadvertent AI.

5. Minimal Risk (Inadvertent AI)

Figure 7: Minimal Risk AI Systems (Inadvertent AI)



"Minimal risk" systems present no risks to people's rights and safety, and their use poses little to no risk across the insurance value chain. This risk can be governed under a firm's vendor agreements. For AI, which is embedded within common software providers, it will be important for carriers to ensure that they *derisk* the vendor supply chain. This includes augmenting vendor evaluation and requests for information (RFIs) with appropriate questions to understand and document that a particular vendor, or a tertiary vendor, is leveraging AI in their development process and/or has AI as a part of their product. This will be important to ensure transparency and attestation to minimize risks. Carriers might be able to attest that their intentional AI systems (whether build or buy) are explainable, transparent, and free of bias — but they also need to ensure they are governing and have compensating controls for those delivering services to them.

3.0 The AIRE Framework

The AIRE framework and classification methodology is intended to be a common place for organizations to **start** their AI risk management and governance journeys, and it is **not** the final destination. In other words, it is **expected** that firms will leverage this framework as a baseline, but that they will make appropriate adjustments as they customize the framework to suit their needs, their risk management strategies, and their own corporate guidelines. It is recommended that this framework **not** be applied to AI implementations deemed to be unacceptable within the insurance value chain.

It should be expected that the AIRE framework will be updated periodically, as AI implementations continue to mature and new implementations and regulatory frameworks are introduced. For instance, it is likely that a new version of this framework will be introduced in 2026 to encompass AI agents and agentic AI. Carriers are responsible for ensuring they are working with the latest version of the framework. Carriers are strongly encouraged to submit any recommendations or updates to this framework to the LIMRA and LOMA AI Governance Group (aigg@limra.com).

To classify AI systems based on risk within the industry, the AI Governance Group outlined the most common domains across the insurance value chain and categorized the types of **intentional** AI system implementations within each domain. This provides carriers with a holistic view of **intentional** AI system implementations to understand where special governance is warranted and where they should leverage adequate transparency as a means to mitigate risks. These domains include: Prospecting, Marketing, Sales, Sales Support, Actuarial, Underwriting and Pricing, Product Development, Compliance and Audit, Customer Service, Claims and Benefits, Investments, Human Resources, and Shared Services (such as IT, Finance, and Billing). AI Risk Classification framework outlines illustrative examples of classified AI initiatives provided by the industry as of Q1 2025. As previously stated, these are purely data points, only to serve as an information source in case your firm is about to undertake a similar AI initiative.

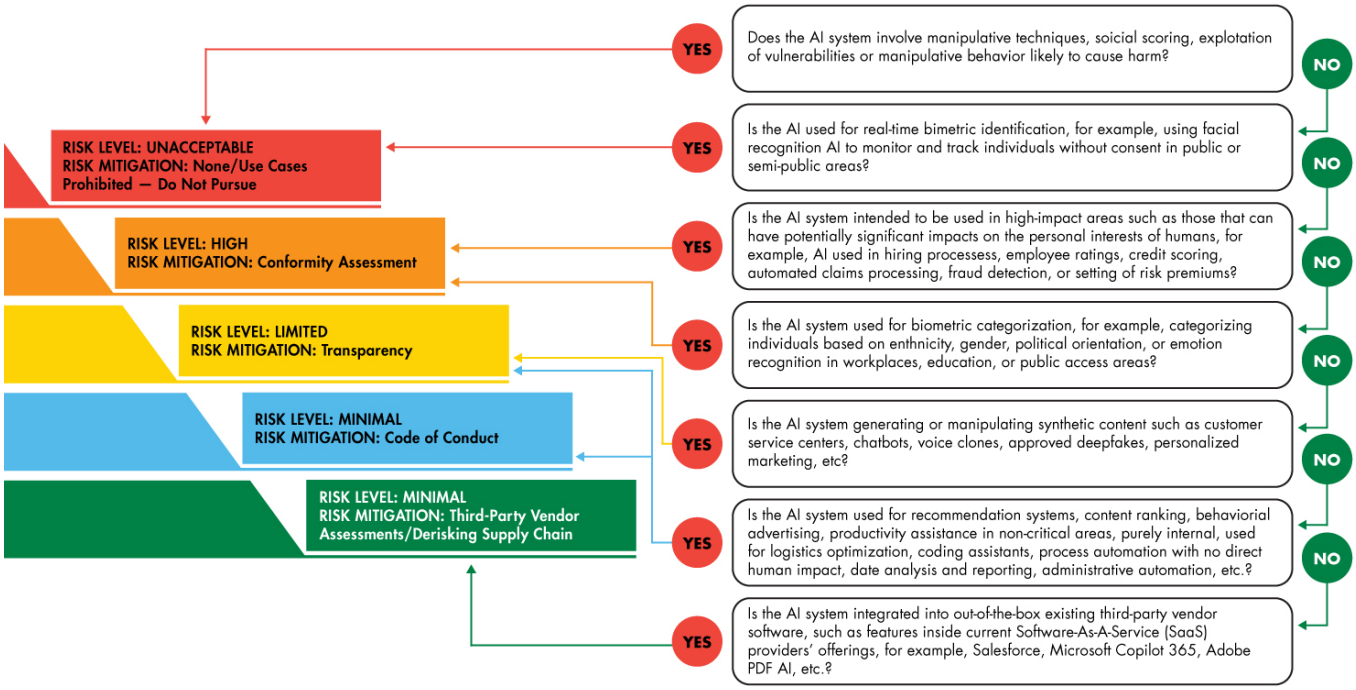
The AIRE framework facilitates the classification of an AI implementation to its appropriate risk category and provides a high-level overview of potential risk mitigation considerations. Detailed best practices for risk management aligned to each classification will be a part of subsequent deliverables, as outlined in the AIGG Roadmap.

The following sections present a few methods on how to facilitate risk classification related to a specific AI initiative. Once again, these sample methodologies are **starting points**. While some firms might be able to leverage these recommendations as-is, it is highly encouraged that firms modify what is outlined here to ensure they incorporate factors based on their specific judgment, risk appetite, and risk management strategies. As long as a firm is fundamentally aligned on what constitutes an unacceptable use of AI, what a “high-risk” AI system is, and what “limited” and “minimal” risk AI implementations are, it should build on what follows to create a relevant and unique risk assessment.

3.1 Decision-Tree Risk Classification Method

Figure 8 provides a visual step-by-step guide to help firms assign a risk category to a given AI initiative.

Figure 8: Decision-Tree Risk Classification



The visual represents a flow-based decision tree that guides organizations through a series of questions about the nature and application of their AI system. Based on the answers, the system is assigned to one of five risk levels as outlined in Section Two. The decision tree is a starting point for firms, and while this can be used in its current form, organizations are advised to use it as just one aspect of how they ascribe risk levels to a particular AI initiative.

Risk Level: Unacceptable Risk (Colored Red)

Questions: "Does the AI system involve manipulative techniques, social scoring, exploitation of vulnerabilities, or manipulative behavior likely to cause harm?" AND/OR "Is the AI used for real-time biometric identification? (For example, using facial recognition AI to monitor and track individuals without consent in public or semi-public areas)"
Examples: Social scoring and AI that targets children with psychological manipulation
Recommendation: These types of use cases are prohibited, and carriers should not pursue them.

High Risk (Colored Orange)

Questions: "Is the AI system intended to be used in high-impact areas, such as those that can have potentially significant effects on people's personal interests? (For example, AI used in hiring processes, employee ratings, credit scoring, automated claims processing, fraud detection, or setting of risk premiums)" AND/OR "Is the AI system used for biometric categorization? (For example, categorizing individuals based on ethnicity, gender, political orientation, or emotion recognition in workplaces, education, or public access areas)"

Examples: AI-driven underwriting, claims, credit assessment, and fraud detection
Recommendation: Before deploying these systems, carriers should ensure they are subject to conformity assessment and to rigorous testing, documentation, and auditability.

Limited Risk (Yellow)
Question: "Is the AI system generating or manipulating synthetic content?"
Examples: Emotion AI in recruiting tools, AI classifying individuals by inferred attributes (such as AI used in customer service centers, chatbots, voice clones, approved deepfakes, and personalized marketing)
Recommendation: These uses require transparency; users must be clearly informed they are interacting with an AI system or the output of an AI system.

Minimal Risk (Light Blue and Yellow for Limited Risk)
Question: "Is the AI system used for recommendation systems, content ranking, behavioral advertising, productivity assistance in non-critical areas, purely internal, or logistics optimization?"
Examples: Generative AI for customer support, AI-generated marketing emails, coding assistants, process automation with no direct human impact, data analysis and reporting, and administrative automation
Recommendation: Carriers should use a Code of Conduct, follow best practices, and ensure outputs are not deceptive.

Minimal Risk (Green – Inadvertent AI)
Question: "Is AI integrated into third-party SaaS?"
Examples: Microsoft Copilot 365, Salesforce, and Adobe PDF AI
Recommendation: Carriers should use third-party governance and mitigate risk via vendor assessments, contracts, and supply chain controls.

3.2 Risk Scoring Risk Classification Method – Intentional AI

It is essential to assess the potential risks associated with AI's deployment and usage as it becomes integrated across the insurance value chain. A comprehensive risk evaluation framework allows organizations to identify, categorize, and mitigate these risks effectively. The assessment considers multiple attributes, ranging from the criticality of where the AI is used to the level of transparency and control over models and data. By assigning scores to each attribute, firms can establish a clear risk profile and determine the appropriate safeguards for each AI implementation. This structured approach will ensure compliance with regulatory frameworks and foster ethical and responsible AI adoption.

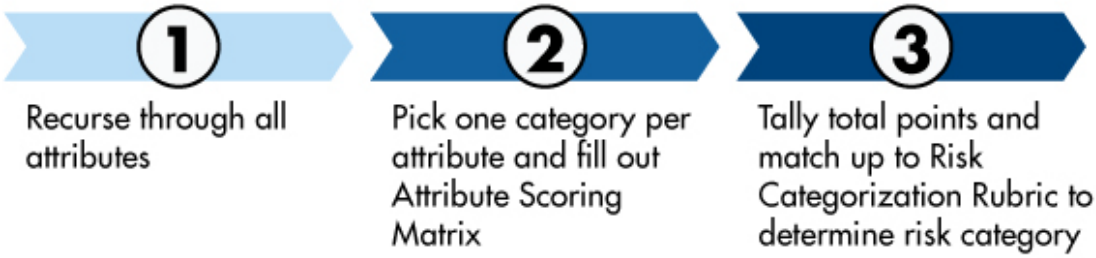
3.2.1 Attributes-Based Risk Scoring

This section presents a risk scoring framework based on six attributes that will help firms arrive at a risk score aligned to specific AI implementation. Note that this quantitative model will require qualitative reasoning. In other words, this framework will help you determine the level of risk and, implicitly, the risk category of a particular AI initiative — such that you can make an enterprise assessment on risk management and mitigation strategies commensurate to the level of risk.

The attributes-based risk scoring framework follows a simple three-step process as depicted in Figure 9. To conduct this scoring, begin with the first attribute, choose one category that best describes the AI initiative corresponding to the attribute, and document the associated point that has been assigned to this category. Then proceed to the next attribute and repeat the same process until a point is assigned to each of the six attributes (and potentially others if they are added). Once this is complete, tally up the points total to arrive at a Raw Risk Score. Finally, evaluate the Raw Risk Score against the Risk Categorization

Rubric to determine the risk category. Remember to adjust the rubric accordingly if attributes and categories are added or removed to reflect what is important to your organization. Once again, while this can be applied as-is, this framework is intended to be a starting point and not necessarily a destination.

Figure 9: Attributes-Based Risk Scoring

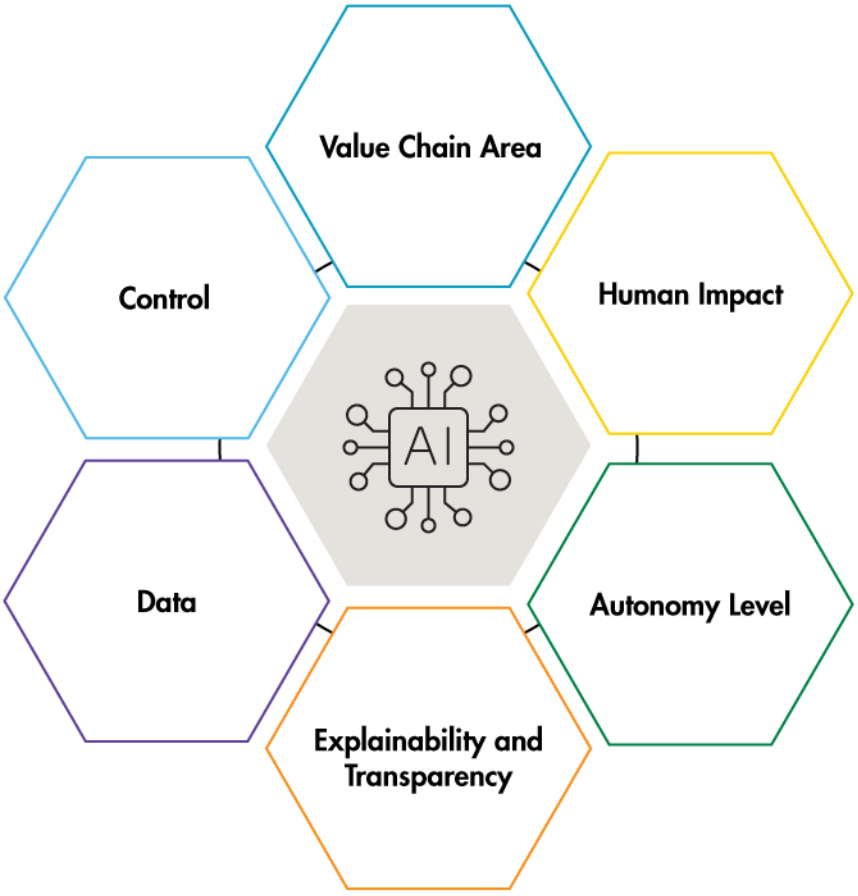


Note that this attributes-based risk scoring focuses on intentional AI and, as such, the four risk levels that are outlined in the following AI Risk Classification Model (Figure 10).

Figure 10: AI Risk Classification Model

The central premise of the attributes-based approach is that each AI use case should be evaluated across six key attributes (Figure 11). Each attribute is assigned points and scored to determine overall risk categorization. Please note that these attributes and underlying points are intended to be a starting place. You are free to add and/or remove attributes and underlying points, as long as you adjust the scoring of each attribute and the risk-scoring rubric that corresponds with any changes. In fact, it is highly recommended that firms customize this scoring to their individual needs.

Figure 11: AI Attribute Categorization



1. Value Chain Area
This attribute assesses the sensitivity of the insurance value chain area where the AI system is deployed. For example, areas intended to assist with claims, underwriting, coverage eligibility, employment, health, and eligibility can be classified as highly sensitive, due to the high stakes involved with AI decision-making. Moderately sensitive areas include customer engagement and pricing optimization, and non-sensitive areas include marketing content generation.

2. Human Impact
This attribute evaluates the potential consequences of AI-driven decisions on people. These individuals could be any stakeholders impacted by the AI-driven decision, from customers to employees. Human impact could include rights, financial matters, fairness, safety and security, privacy, or access to essential services. AI systems that are deemed as high impact can render decisions that have significant legal or economic outcomes. Implicitly, AI systems that have high impacts require significantly greater scrutiny, governance, and caution. Moderate-impact AI provides decision support, but it always requires human review in the final decision. While caution is urged for moderate-impact AI systems, this human oversight helps to reduce the level of risk. Finally, low-impact systems provide insights and analytics, but they require humans to make a final decision and/or provide a data point among others required.

3. Autonomy Level
This attribute focuses on AI's level of control over outcomes. Fully autonomous systems that can make independent decisions without any human intervention or oversight pose higher risks. AI systems that require human-in-the loop or are human-controlled are less autonomous, thereby providing increased safety and accountability.

4. Data
This attribute focuses on the nature of the underlying data utilized by the AI system in terms of its sensitivity. Sensitivity of data is especially important in AI ecosystems, whether that data has been used for training the model and/or the AI ecosystem deals with sensitive data when operationalized. Consider AI systems that handle data elements such as Personally Identifiable Information (PII), Sensitive Personally Identifiable Information (SPII), medical history, biometric data, or other sensitive data categories. AI systems that deal with sensitive data require rigorous safeguards. These systems implicitly warrant a higher level of risk classification in contrast to AI systems that handle non-sensitive or public data.

5. Explainability and Transparency
This attribute focuses on the importance of model explainability and transparency. AI systems can notoriously be opaque "black boxes," and it is vital for firms to focus on models that are explainable and transparent so that they can understand and conduct audits of outputs and decision-making by AI models. AI models that operate with limited explainability pose a higher level of risk.

6. Control
This attribute focuses on the level of control a firm has over the AI model and its lifecycle. Systems that are internally developed or governed through transparent third-party vendor contracts present lower risks, compared to those controlled by opaque third-party vendors.

3.2.2 Scoring Attributes

The following grid summarizes how a firm can leverage the attributes listed above, with their appropriate points assigned. Again, it is important for your organization to treat these attributes, associated categories aligned to each attribute, and points commensurate to each category as a starting point. Adjust and modify as appropriate but ensure that the assigned points are adjusted accordingly. To arrive at a risk classification:

A. Move through the grid, going attribute by attribute until you have completed all six (and any others you have added).
B. Select one — and only one — category within each attribute and note the point value associated with that specific category. Fill out the Attribute Scoring Matrix provided in the next section.
C. Tally the total number of points in the Attribute Scoring Matrix and assess it against the Risk Categorization Rubric to arrive at your risk category.

Attribute	Category	Illustrative Examples	Points
Value Chain Area	High Sensitivity	HR – employment/hiring Fully automated underwriting/coverage determinations/pricing (AI evaluates risk	3

		and proposes rates with no human review)/insurance eligibility Fully automated claims where AI adjudicates claims or policies without human intervention, or conducts annuity payout calculations Fully automated mortality assumptions	
	Moderate Sensitivity	Marketing for customer engagement Actuarial for pricing optimization	2
	Limited Sensitivity	Marketing for content generation	1
	Low Sensitivity	Internal operations, such as for internal chats for employees Manual search tool for policy terms, such as with a keyword search	0
Human Impact	High Impact	Decisions with legal, financial, or economic consequences AI outputs that directly impact eligibility, pricing, or approval outcome AI that recommends adjudication of an underwriting application or claims with little to no human oversight	3
	Moderate Impact	Influences decisions, but with human review only AI that supports annuity payout calculations or mortality assumptions or renders recommendations for adjudication of claims with a final human decision	2
	Low Impact	Informational, non-decisional, decision-support-like impacts, such as flagging anomalies for manual follow-up, providing underwriters with decision-support tools and insights, or suggesting product combinations for wholesalers to use	1
	No Impact	Internal brainstorming, assistance with document summaries, drafting emails, or marketing campaign idea generation	0
Autonomy Level	Fully Autonomous	AI that adjudicates insurance applications, claims, or policies or sets rates without human intervention	3
	Partially Autonomous (Human-in-the-Loop)	AI that evaluates risk or conducts triage and routes insurance applications with limited review	2

		AI that proposes rates with minimal review for a human to make a final decision AI that provides underwriters with decision-support insights	
	AI-Assisted Decision-Making	AI system that pre-fills forms for agent approval	1
	Fully Human-Controlled	AI system that acts as a digital assistant and suggests policy riders, products, or prompts for customer service representatives	0
Explainability and Transparency	"Black Box"/Opaque	AI model behavior is completely opaque and outcomes are unpredictable — firms have no visibility into how AI model arrived at the decision that it did. If a firm is unaware that AI is being actively leveraged by the vendor product	3
	Partially Explainable	Third-party AI vendor provides post-hoc examples, there is limited technical explainability, or explainability and transparency are not established in vendor contracts. Carrier resources use tools such as LIME and SHAP to interpret the model's functioning.	2
	Somewhat Explainable	Some model rules are transparent and explainable, but others are embedded deeper in the model, disallowing total visibility and transparency. AI use disclosures might be present in the vendor's terms of use but not transparently declared during contracting.	1
	Fully Explainable	Third-party vendor provides full transparency, such as deterministic logic for filling forms, or provides artifacts such as visual logic flow of their model. Transparency would require that customers/stakeholders are clearly made aware of the use of AI in the ecosystem, or that they are interacting with AI-generated content.	0
Nature of Data	Highly Sensitive	Data that would be considered highly sensitive for traditional (non-AI) business operations, such as health records, racial information, and financial records	3

		Sensitive Personally Identifiable Information (SPII) and data considered to be under HIPAA	
	Sensitive	Personally Identifiable Information (PII), such as first name, last name, and email address combination (for example, for customer or prospect names and emails, application metadata, etc.)	2
	Publicly Available/ Anonymized	Publicly available data sets, such as zip codes, census information, and anonymized advisor sales performance data	1
	No Personal Data	Synthetic datasets that are only used for model training purposes	0
Vendor and Lifecycle Control	Third-party vendor, "Black Box" model, minimal control or insight	Typical of AI systems developed on unknown, free, "freemium," or less popular open-source AI models There is little to no third-party vendor accountability. Vendors cannot furnish any details about the underlying AI model and can only supply the AI decision/outcome.	3
	Third-party vendor, moderate level of control or insight	The terms of the third-party vendor contract are ambiguous, unclear, or vague with respect to AI. The vendor can use a "daisy chain" of subcontractors, thereby increasing your organization's exposure because even if the firm's AI model is documented, it might not be able to fully explain a subcontractor's model. Typically, any documentation provided is reactive and limited to marketing slides or high-level summaries, lacking an appropriate level of technical detail.	2
	Third-party vendor, high level of control or insight, via clearly defined third-party vendor contracts	The third-party vendor signs contracts and provides associated compliance documentation, with full transparency, and allows audits from your firm as well as external third-party audit providers. With a high-level of transparency, the third-party vendor can share model lineage, methodology and results of bias testing, hosting information, cybersecurity	1

		assessments and results, and information security assessments and results.	
	Internally built	If your firm's AI model is built internally (see AIGG Build vs. Buy Subcommittee Whitepapers), your organization implicitly has absolute control over the AI ecosystem and model. Building your own models means that you also can exert control and have a clear line of sight into the underlying data that is being used by the AI models. Therefore, firms have a clear understanding of the data provenance that is being fed into the AI models.	0

3.2.3 Attributes Scoring Matrix

Once you have assigned a point to each attribute, you will have a grid like below. Tally up the total points to yield a Raw Risk Score. A simple rule of thumb is that a higher Raw Risk Score means a higher level of risk.

Attribute	Points Possible (Min-Max)
Value Chain Area	0 - 3
Human Impact	0 - 3
Autonomy Level	0 - 3
Explainability and Transparency	0 - 3
Nature of Data	0 - 3
Vendor and Lifecycle Control	0 - 3
TOTAL RAW RISK SCORE	**0 - 18**

Your Attributes Scoring Matrix should result in something like the below **illustrative example**:

Attribute	Points Assigned
Value Chain Area	2
Human Impact	3
Autonomy Level	1
Explainability and Transparency	2
Nature of Data	1
Vendor and Lifecycle Control	1
TOTAL RAW RISK SCORE	**10**

3.2.4 Risk Categorization Rubric

Finally, given the Raw Risk Score, find your score within the Risk Categorization Rubric below to arrive at your Risk Classification.

Raw Risk Score	Risk Classification
0 to 6	**Minimal Risk**
7 to 10	**Limited Risk**
11 to 14	**High Risk**
15 to 18	**Unacceptable Risk (*Note: there might be exceptions*)**

Figure 12 provides a visual ruler illustration version of the rubric.

Figure 12: Risk Categorization Visual Ruler

Once you complete an enterprise AI Risk Classification for AI initiatives across your organization, the final result should look something like the below **illustrative example**:

AI Initiative	Raw Risk Score	Category	Notes
AI-based underwriting triage	11	High Risk	Impacts prospective customer application, pricing, and adjudication
Fraud detection	9	Limited Risk	Provides suggestions to humans, but cannot make autonomous decisions, serves as human decision-support
AI-based product recommendations	9	Limited Risk	Provides suggestions to humans, but cannot make autonomous decisions, serves as human decision-support
Customer service chatbot	7	Limited Risk	Might hallucinate. Requires disclosures to customers that they are interacting with an AI chatbot.
Generative AI to brainstorm marketing materials	4	Minimal Risk	No impact on rights or decisions
Generative AI to draft internal emails	1	Minimal Risk	No impact on rights or decisions

3.3 Risk Scoring Risk Classification Method – Inadvertent AI

AI is no longer confined to specific AI applications or standalone AI systems; it is also embedded within the software products provided by third-party vendors. These third-party tools are used ubiquitously across carriers. Even if a firm has not intentionally invested in AI, just by virtue of doing business in the 21st century, it will be an inadvertent AI consumer. This "inadvertent AI" often operates beneath the surface, powering a vast range of features and functionality within familiar platforms. While these capabilities bring significant advantages and can dramatically enhance operational efficiencies, they also introduce complex risks that can go unnoticed if not specifically addressed.

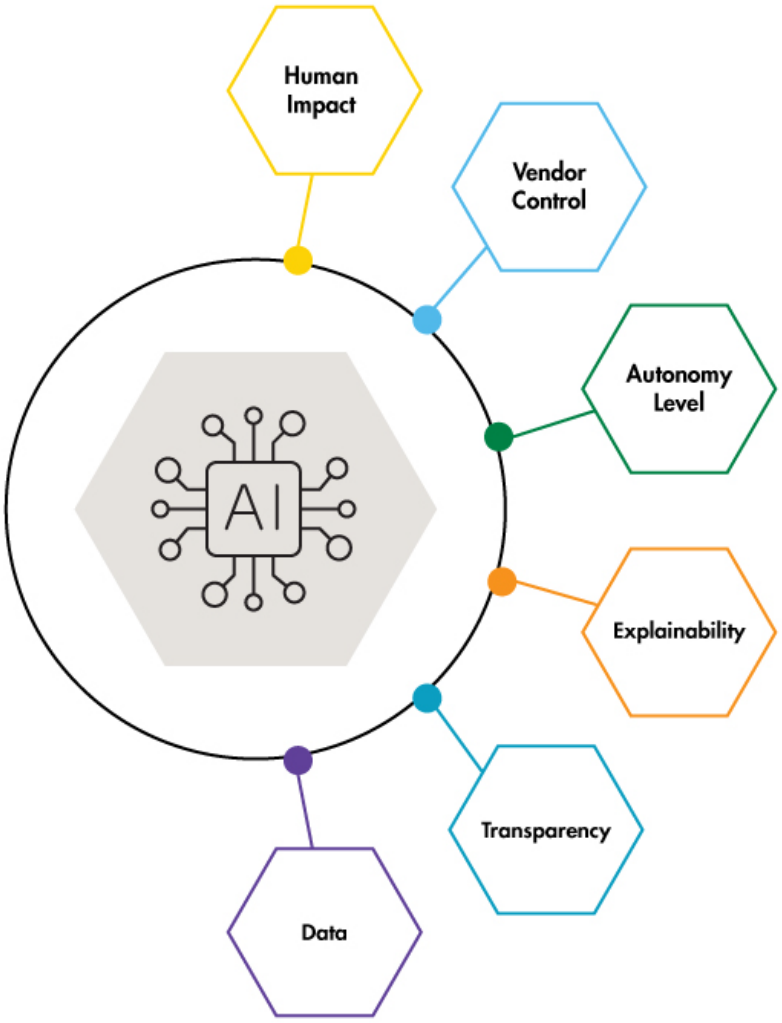
The challenge lies in the fact that oversight requirements may vary depending on the nature of the vendor. For example, established firms like Adobe or Microsoft, with their robust governance structures and industry-standard practices, may inherently require less scrutiny compared to smaller or emerging providers with less established oversight mechanisms. Understanding these nuances is essential for carriers, as they navigate a landscape where the cohabitation of AI systems from multiple sources could potentially affect compliance, governance, cybersecurity, information and data security, and ethical implementations. Therefore, it becomes critical to assess the presence and functionality of AI within these tools systematically, identifying whether safeguards align with your organization's expectations for transparency, explainability, accuracy, and accountability. By focusing on inadvertent AI in addition to intentional AI, firms can better manage the risks in a holistic manner, while capitalizing on the transformative potential of these embedded AI features.

The following risk assessment for inadvertent AI is a derivative of the preceding section on intentional AI and is specifically designed for evaluating AI embedded within vendor software. This assessment also leverages six attributes to measure risk, which are thematically similar to those assessing risk for intentional AI implementations. Overall, inadvertent AI — particularly for larger, established firms such as Adobe and Microsoft — is generally considered to be low risk. This is primarily because these embedded

functions are intended to enhance human productivity. The following assessment allows firms to evaluate vendors' embedded AI across the range of third parties with which they do business.

The six core attributes to measure inadvertent AI risk exposure are shown in Figure 13. The definitions for each of these are similar to the attribute definitions outlined in Section 3.2.1, and a firm would follow a similar process to calculating the Raw Risk Score for intentional AI systems. Assigning risk classifications would also follow the same process, but keep in mind that not all embedded AI is equal. Do not apply the same rigor for embedded AI in productivity suite software such as Office and the Adobe suite as you would with a less established or niche third-party vendor product.

Figure 13: Core Attributes to Measure Inadvertent Risk Exposure



Attribute	Evaluation Criteria Examples	Points
Human Impact	Automated decision-making without human oversight or involvement, including decisions related to regulatory and compliance activities. (Note: this would be an unlikely factor to be considered for established vendors that provide productivity functionality, such as Adobe and Microsoft.)	3
	Customer service recommendations and product recommendations that require human intervention and decision-making	2
	Impacts operational workflows (upstream or downstream dependencies)	1
	Brainstorming, drafting correspondence	0
Vendor Control	Unknown vendor, or unproven open-source AI models being used. Third-party vendors do not declare use of AI within their system, or they declare the use of AI without providing model/logic documentation. (Note: this would not be important for established vendors that provide productivity functionality, such as Adobe and Microsoft.) However, if your firm is leveraging a vendor for an underwriting workbench, as an example, it should be expected that all AI use is known.	3
	Vendor uses subcontractors in an "AI daisy chain" solution. (Note: this would not be important for established vendors that provide productivity functionality, such as Adobe and Microsoft.)	2
	Vendor provides transparency as demanded by the vendor contract, but it is "light" on providing underlying technical details. Once again, this might not be a factor for the established AI-augmented productivity suite players.	1
	Third-party vendors provide full transparency into their AI models, underlying data (training data) and requisite artifacts, including signed attestations and data provenance.	0
Autonomy Level	Embedded AI is fully autonomous without human oversight and intervention. There is no human-in-the-loop, and the embedded AI makes key decisions with limited to no human review.	3

	Embedded AI provides insights and limited human oversight.	2
	Embedded AI requires human oversight, review, and consent to proceed.	1
	Embedded AI provides insights, recommendations, and suggestions only, with decision-making in the hands of a human operator.	0
Explainability	Embedded AI model behavior operates as a "black box," and technical details are opaque.	3
	Vendor provides limited embedded AI model explanation only when asked.	2
	Vendor proactively provides limited AI explainability.	1
	Vendor provides documented embedded AI model logic; the entire system is fully transparent and auditable.	0
Transparency	Vendor provides no transparency to users or the organization regarding the use of AI within its product.	3
	End users are unaware AI is being used in a vendor system unless they review documents such as terms of use.	2
	Vendor provides limited transparency, but it can be inconsistent, varying by product line.	1
	There is clear notification to users and your firm that AI is embedded within a product. This includes how and why it is being used, for what purpose, and how the collected data will be used — allowing users a choice to opt in or out on using their data to train models in the future.	0
Nature of Data	Embedded AI uses SPII or other regulated data such as data governed under HIPAA, GDPR, etc.	3
	Uses PII (first name, last name, and email address), which is LIMRA and LOMA's definition of PII.	2
	Embedded AI only uses anonymized and/or internal data.	1
	Embedded AI only uses publicly available datasets.	0

References

EU. (2024). *Shaping Europe's digital future*. Retrieved from European Commission - European Union: https://digital-strategy.ec.europa.eu/en/policies/regulatory-framework-ai

Kourinian, A. (2024, January). *Conducting an AI Risk Assessment - Bloomberg Law.* Retrieved from Mayer Brown: https://www.mayerbrown.com/-/media/files/perspectives-events/publications/2024/01/conducting-an-ai-risk-assessment_kourinian.pdf%3Frev=-1

Meier, K., & Spichiger, R. (2024, March 15). *The EU AI Act: What it means for your business*. Retrieved from EY: https://www.ey.com/en_ch/forensic-integrity-services/the-eu-ai-act-what-it-means-for-your-business

Tabassi, E. (2023, January 26). *NIST Trustworthy and Responsible AI, National Institute of Standards and Technology.* Retrieved from Artificial Intelligence Risk Management Framework (AI RMF 1.0): https://doi.org/10.6028/NIST.AI.100-1, https://tsapps.nist.gov/publication/get_pdf.cfm?pub_id=936225

Advancing the financial services industry by empowering our members with

LIMRA LOMA
Navigate With Confidence

0936-2025