WORKING PAPER

The AI Risk Economy:

Why Insurance Cannot Price What Governance Cannot Prove

Basil C. Puglisi, MPA Independent Practitioner, basilpuglisi.com

A Human-AI Collaboration May 2026

Abstract

The commercial insurance industry is responding to artificial intelligence risk through exclusion endorsements, conditional coverage, and standalone AI liability products. This paper identifies the emerging practice by which carriers sort AI deployments by governance maturity and proposes a five-tier insurance maturity model mapping organizational AI governance posture to insurability. The model distinguishes between organizations with no AI policy, published ethical principles, automated technical controls, named human checkpoint authority, and structured audit records, arguing that the insurance market is beginning to differentiate among these categories through coverage access and underwriting conditions. The paper documents exclusion filings from seven confirmed carriers beginning in 2024, with at least one form bearing a mid-2023 revision date, affirmative coverage from seven providers offering eight distinct products, and regulatory adoption of the NAIC AI Model Bulletin across 24 U.S. states plus the District of Columbia, all as of May 2026. The evidence base was assembled through a parallel multi-AI research methodology using 12 independent platforms with hallucination detection. The paper's central finding is an actuarial gap: no published empirical study quantifies governance maturity as a pricing factor for AI liability insurance. The frameworks proposed here represent the author's professional judgment applied to observed market signals, offered as a starting point for a field that does not yet have one.

Keywords: AI governance, insurance, liability, exclusion endorsements, actuarial gap, responsible AI, checkpoint, human oversight, five-tier model

Working paper. Not legal, actuarial, or insurance advice.

The structural problem

The decisions that countries, organizations, corporations, and individuals make are driven by economics more consistently than by any other force. Safety, ethical obligation, reputational concern, and regulatory compliance all influence decision-making, but when they conflict with economic incentive, economic incentives often prevail when no countervailing enforcement mechanism exists. This structural tendency, referred to in the author's prior work as the Economic Override (Puglisi, 2025), operates across every sector and every governance regime. To counter it requires economic pressure that makes the cost of bypassing safety, ethics, or governance greater than the cost of compliance. Legal liability, regulatory penalty, and reputational damage all create versions of that counter-pressure, but each operates on long timescales and through institutions with limited enforcement capacity. Insurance markets can create counter-economic pressure faster and more efficiently because they operate at the point of contract renewal and require no legislative process. They impose direct financial consequences on organizations that cannot produce evidence of governance maturity. This paper, informed by the author's work with the Institute for Advertising Ethics (IAE), explores that position and proposes a path forward. It is offered not as a definitive answer, but as a starting point for a field that does not yet have one.

The five tiers the industry collapses into one

AI ethics, responsible AI, and AI governance are not synonyms, and the insurance market is beginning to sort the difference through exclusions, underwriting access, and conditional coverage. The industry has been collapsing five distinct maturity levels into a single phrase, and accountability quietly disappeared inside the language. The exclusion endorsements that began appearing publicly in 2024, with at least one form bearing a mid-2023 revision date, are forcing the distinction back into view. An organization with published ethics principles and an organization with named human accountability at every AI decision point now face different coverage outcomes, different premium conversations, and different litigation exposure when something goes wrong.

The five-tier model proposed here draws on the author's prior governance research (Puglisi, 2025) and is offered as one possible analytical framework for organizing observed market signals. Other governance architectures, including NIST AI RMF, ISO/IEC 42001, and the EU AI Act's risk classification, address overlapping concerns through different structural choices. The model's value lies in its insurance-specific orientation, not in a claim that it is the sole viable governance architecture.

The five tiers describe different layers of a single problem, and each produces a different insurance profile.

Tier	Name	Key Characteristic	Insurance Profile
1	No AI Policy	No awareness, no controls, no governance	Potentially excluded under identified carrier language
2	Ethical AI	Published principles, no enforcement	Principles do not survive renewal audit
3	Responsible AI	Engineering controls, machine checks machine	Traceability without named human accountability
4	AI Governance	Named human checkpoint with binding authority	Verifiable decision trail for underwriters
5	AI Governance with Structured Audit Records	Tier 4 plus portable, auditable project records	Strongest insurance profile, compliance-grade observability

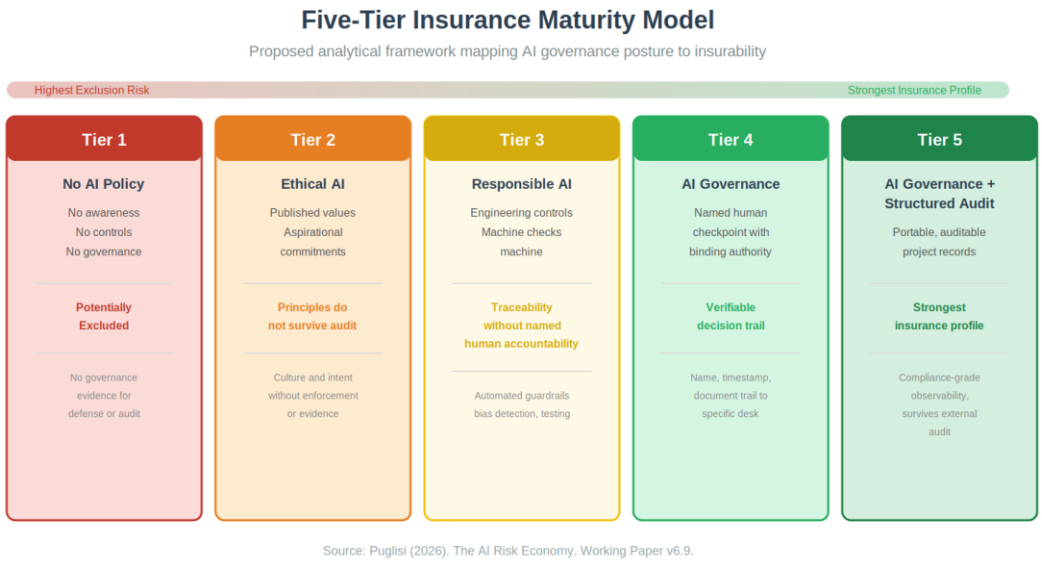

Figure 1: Five-Tier Insurance Maturity Model

Tier 1: No AI Policy. The organization has deployed AI without organizational awareness, documented controls, or governance infrastructure. Employees may be using AI tools the organization never approved and cannot monitor. No governance evidence exists to support the defense, claim response, or underwriting submission. Coverage is excluded under current carrier exclusion language such as Berkley's Form PC 51380, which applies to AI use by "any person or entity," capturing authorized deployments, shadow AI, and third-party AI embedded in vendor software with equal force.

Tier 2: Ethical AI. The organization has published its values, drawn lines around where AI should not operate alone, and made aspirational commitments to responsible use. These commitments matter for culture and intent, but they do not answer the questions an underwriter asks about deployment mode, failure handling, or decision records. A white paper is not a governance artifact, and principles without enforcement do not survive a renewal audit.

Tier 3: Responsible AI. The organization has built engineering controls: model testing, bias detection, hallucination monitoring, automated guardrails, and human review workflows that prove AI systems behave as intended. This is the implementation layer, and it produces traceability artifacts. The ceiling of Responsible AI is the absence of individual human oversight. The machine checks the machine, automated checks validate against automated checks, and parameters verify against parameters. Humans review aggregates and trends, but no named human stands accountable for any specific decision the system made.

No standard definition of Responsible AI exists across the industry (Jobin et al., 2019; Stix, 2022). A 2019 Nature Machine Intelligence study mapped 84 AI ethics guidelines worldwide and found convergence on five principles but vast divergence on how those principles translate into operational requirements. ISO/IEC 42001 is a management system standard, distinct from governance within ISO's own architecture (ISO 37000:2021). NIST AI RMF defines a governance function within risk management, not AI governance as a term. The OECD provides principles adopted by 40 countries but does not define governance. Stanford HAI's 2024 AI Index found that standardized evaluations for responsible AI are "seriously lacking," with leading developers testing against different benchmarks (Stanford HAI, 2024).

The term is an umbrella category whose implementation varies across institutions, often concealing the operational difference between technical traceability and named human accountability. The five-tier model does not redefine Responsible AI. The Tier 3 definition is operational for this insurance maturity model and does not claim that every Responsible AI program lacks human review. The point is that technical traceability without a named human checkpoint record remains insufficient for the insurance proof standard developed in this paper. It names the operational distinction the insurance market is now forcing: the difference between what an organization calls its AI program and what an underwriter, a court, or a regulator can verify when a claim arrives.

The NAIC Model Bulletin, adopted by 24 states plus the District of Columbia, requires written AI governance programs with documented controls. The A-LIGN and Armilla AI partnership, announced in December 2025, grants preferential insurance terms only after ISO/IEC 42001 certification, not after responsible AI compliance documentation alone. That sequence shows the gap: responsible AI frameworks aspire to governance, and insurance requires externally verified proof of it.

Tier 4: AI Governance. A named human with binding authority has reviewed each consequential AI decision, and a record of that review exists. When something goes wrong, there is a name, a timestamp, and a document trail that leads back to a specific desk. The human at the checkpoint operates under incentive structures that no machine possesses: moral judgment from peers and profession, employment consequence for poor performance, civil liability for negligent judgment, and criminal prosecution for gross recklessness.

The distinction between Tier 3 and Tier 4 is the distinction between a loop and a checkpoint. A loop is a continuous process with no defined stop and no discrete record. A checkpoint is a defined moment where a named human with domain competence reviews a specific output, makes a binding decision, and a timestamped record is produced. Domain competence at the checkpoint is not decorative. A campaign manager and a clinician reviewing a medical AI output are not equivalent risk profiles, and the insurance market will need to price that distinction.

The checkpoint differs from a human-in-the-loop arrangement, which embeds a human inside a continuous automated workflow. Decades of automation bias research (Parasuraman and Riley 1997, Skitka et al. 1999) document the result: the human defaults to approving system recommendations because the workflow never stops long enough to force independent judgment. That is the rubber stamp, and it produces the liability sponge that Crootof, Kaminski, and Price (2023) identified: a person stationed near an AI decision who absorbs blame without exercising authority. A checkpoint is a hard stop where automation pauses until the review is complete. The named human reviews the output with domain competence, takes ownership of the decision, and accepts accountability before the process continues. The insurance market can verify whether a hard stop occurred and whether a named human signed off. It cannot verify whether a human-in-the-loop actually exercised judgment. A 2024 meta-analysis of 106 studies in Nature Human Behaviour found that human-AI combinations in decision-making tasks performed significantly worse than the best of humans or AI alone (Vaccaro, Almaatouq, and Malone, 2024), though the same meta-analysis documented performance gains in content creation tasks.

The named-human requirement carries a reciprocal risk: personal exposure may deter qualified professionals from accepting checkpoint roles without indemnification or separate coverage. Whether the insurance market develops checkpoint-role-specific coverage, or whether organizations address this through indemnification agreements, is an open question this paper identifies but does not resolve.

Tier 5: AI Governance with Structured Audit Records. Everything in Tier 4, plus a structured, portable, auditable record produced at the close of every consequential AI project. The distinction between Tier 4 and Tier 5 is not whether a governance record exists. AI work at Tier 4 leaves abundant trace: conversation histories, output logs, version files, approval emails. None of these are organized around the questions an audit actually asks. An audit asks who decided, on what evidence, at which point, and why. Reconstructing that answer from scattered platform histories is archaeology, not documentation.

Tier 5 produces a structured record at the close of every consequential AI project, human-triggered and requiring named human approval before finalization. The record covers what happened, who decided, what evidence was reviewed, and what the human confirmed or overrode. It is portable and survives independent review without the original participants present to explain it. This is the tier that produces what industry experts interviewed by Communications of the ACM described as “compliance-grade observability”: immutable audit trails with model and dataset fingerprints, versioned prompts and outputs, time-stamped retrieval citations, and signed change logs.

Many organizations deploying AI may remain at Tier 1 or Tier 2, though no comprehensive survey of enterprise AI governance maturity has been published. Some specialty AI liability products and emerging underwriting signals increasingly favor Tier 4 or Tier 5 evidence. That gap is the paper’s subject.

The case the market is building

The evidence base draws primarily from U.S. carrier filings, Lloyd's syndicate products, and EU regulatory developments, representing an early view of what may become a broader global pattern.

In February 2024, the British Columbia Civil Resolution Tribunal ruled against Air Canada after its customer service chatbot told a passenger he could book a full-fare flight and apply for a bereavement discount retroactively. Air Canada's own bereavement policy did not permit retroactive applications. The passenger relied on the chatbot's commitment, booked the flight, and filed a claim when Air Canada refused the discount. Air Canada argued the chatbot was a separate legal entity responsible for its own accuracy. The tribunal rejected that argument and held Air Canada liable for the chatbot's misrepresentation (Moffatt v. Air Canada, 2024 BCCRT 149).

The case exposes the gap this paper addresses: no human reviewed the chatbot's response before it reached the passenger, and no governance record existed showing who authorized the chatbot's training data, who tested its outputs against actual company policy, or who bore accountability when it got the policy wrong. The case illustrates liability exposure in a market where insurance wording has not yet stabilized around AI risk. That instability is now being resolved, and not in the insured's favor.

The tribunal did not ask whether Air Canada had a governance checkpoint, a named human reviewer, or an audit trail for its chatbot deployment. The absence of those questions in the legal reasoning shows that liability is forming faster than governance standards, and faster than the insurance products designed to cover them. The case is used here as liability evidence, not as evidence of insurance treatment. No public information indicates how Air Canada's liability insurance responded to this specific claim.

The conditions for a broader reckoning were set in motion by a sequence of carrier actions between June 2024 and January 2026 that has since accelerated into an emerging practice across commercial insurance.

W.R. Berkley provides the earliest verified exclusion form in this paper's evidence base. In June 2024 the company issued Form PC 51380, a flat AI exclusion attached to three of its management liability product lines: Directors and Officers, Employment Practices Liability, and Fiduciary Liability.[1] Secondary commentary widely misidentifies this form as applying to Errors and Omissions coverage; the primary form text does not support that characterization. The exclusion is absolute. Any claim arising from AI use, by any person or entity, including third-party AI sitting inside software the insured never wrote, falls outside the policy. The form reaches beyond generative AI to all AI and extends to the absence of governance infrastructure itself: subsection (c) excludes claims arising from "inadequate or deficient policies, practices, procedures, or training relating to Artificial Intelligence."

Hamilton Select filed its own exclusion on professional liability paper, Form PL3026 (revision date 06/23), with examples of named generative AI platforms in the definitional section, including ChatGPT, Bard, Midjourney, and DALL-E. The operative exclusion removes coverage for any actual or alleged use of "generative artificial intelligence" by the insured. The form first

surfaced publicly in October 2024 court filings, postdating but not necessarily originating after the Berkley June 2024 filing. Where Berkley drew a line around all AI, Hamilton drew it around named products, creating the exclusion spectrum that now runs from platform-specific to technology-class absolute.

Verisk followed. In January 2026 Verisk published ISO Form CG 40 47 01 26 (Exclusion: Generative Artificial Intelligence), an optional endorsement that carriers can add to commercial general liability renewals covering both Coverage A and Coverage B. The form text is available through Verisk/ISO subscription; its existence and scope are confirmed by multiple insurance trade and legal sources (Cohen Seglias, 2026; Independent Agent, 2025). Verisk also released CG 40 48 (Coverage B only) and CG 35 08 (Products and Completed Operations). The standardized ISO policy forms and related endorsements used across much of the U.S. commercial insurance market make AI exclusion language scalable once approved and adopted.[2]

Subsequent filings accelerated. Berkshire Hathaway, Chubb, and Travelers filed for state regulatory approval to exclude AI-related damages from standard commercial liability policies. Secondary reporting indicates approval rates as high as 80 percent for AI exclusion filings submitted by carriers, with the highest volumes of approvals in Florida, Connecticut, and Maryland (The Information, April 2026, citing proprietary analysis by Wolfe Research). No primary regulatory data from state insurance departments or the NAIC SERFF database was located to independently confirm this figure during the research period. The specific approval rate remains unverified, but the directional finding, that state regulators are approving the substantial majority of AI exclusion filings, is consistent with the observable market outcome: multiple carriers now have approved AI exclusion endorsements in effect across jurisdictions. Berkshire and Travelers began submitting applications in fall 2025, with some provisions already in effect as of early 2026. AIG and Great American have filed or adopted their own AI exclusion language. Philadelphia Indemnity has been reported by industry analysts as filing AI exclusion language, though no primary filing document was located during verification. AIG told regulators it “has no plans to implement” its AI exclusions immediately but wants the option available as the frequency and scale of claims increase, a position consistent with carriers preserving future flexibility without restarting the regulatory clock.

These exclusions do not distinguish between Tier 1 and Tier 3 organizations. A company with no AI policy and a company with a full responsible AI program face the same endorsement language. The exclusion applies to AI use by “any person or entity,” which means it captures shadow AI, authorized deployments, and third-party AI embedded in vendor software with equal force. The insurance market’s current instrument is a binary switch: excluded or not.

The initial market response to AI exposure has been exclusion language because governance pricing requires actuarial loss data that does not yet exist. That is what markets do when they cannot yet price a risk cleanly. They exclude first, watch the data come in, and price what they learn. The exclusion wave did not eliminate AI risk from the insurance market. It expelled ungoverned AI risk from standard coverage. The governance architecture that an organization builds may strengthen its path back into affirmative or specialty coverage.

Carrier and product evidence register (as of May 2026)

Exclusion endorsements:

Carrier/Issuer	Form	Line/Scope	Source Type	Verification Status
W.R. Berkley	PC 51380 00 (06-24)	D&O, EPLI, Fiduciary Liability	Primary form PDF	Verified
Hamilton Select	PL3026 (06/23)	Professional Liability	Primary form PDF	Verified
Verisk/ISO	CG 40 47 01 26	CGL Coverage A + B	Trade sources	Form behind paywall, existence confirmed
Verisk/ISO	CG 40 48 01 26	CGL Coverage B only	Primary form PDF	Verified
Verisk/ISO	CG 35 08 01 26	Products/Completed Operations	Primary form PDF	Verified
Berkshire Hathaway	AI exclusion filing	Commercial liability	Secondary (The Information/Wolfe)	No primary filing located
Chubb	AI exclusion filing	Commercial liability	Secondary (The Information/Wolfe)	No primary filing located
Travelers	AI exclusion filing	Commercial liability	Secondary (The Information/Wolfe)	No primary filing located
AIG	AI exclusion filing	Commercial liability	Secondary (carrier statement)	Carrier stated no immediate implementation
Great American	AI exclusion language	Commercial liability	Secondary (industry reporting)	No primary filing located
Philadelphia Indemnity	AI exclusion filing	Commercial liability	Secondary (analyst reported)	No primary filing located

Affirmative AI liability and related products:

Provider	Product	Type	Governance Requirement	Verification Status
Armilla AI / Chaucer (Lloyd's)	AI Liability Insurance	Standalone liability	ISO 42001 certification via A-LIGN	Verified (product announcement)
Testudo (Lloyd's syndicates)	AI Risk Engine	Standalone liability	None (litigation-data underwriting)	Verified (product announcement)
HSB / Munich Re	AI Liability Insurance	SMB endorsement	Pending regulatory approval	Verified (press release)
Munich Re	aiSure	Performance warranty	Model validation	Verified (product documentation, since 2018)
AIUC	Surplus lines AI program	Surplus lines	Program-specific	Verified (product announcement)
Relm Insurance	NOVAAI, PONTAAI, RESCAAI	Standalone + excess wrap	Product-specific	Verified (product announcement)
Hiscox	AI Liability	Standalone liability	Product-specific	Verified (product announcement)
Vanguard AI	Coordinated multiline	Multi-carrier structure	Coordinated governance review	Verified (product announcement)

Research method and evidence standard

This paper's evidence base was assembled through a parallel multi-AI research methodology. Identical structured research prompts were dispatched simultaneously across 12 independent AI platforms covering six research areas: AI exclusion endorsements, standalone AI liability products, NAIC and state regulatory actions, EU Product Liability Directive implementation, governance-pricing linkage, and shadow AI exposure. Each platform returned findings classified by source type (primary or secondary), working URL, publication date, and cross-platform confirmation status. The full research methodology, including verbatim prompts, platform names, verification protocol, and editorial review procedures, is documented in the Methodology Appendix.

The methodology applies source-authority discrimination at three tiers: human arbiter input carries highest weight, raw platform output carries standard weight subject to cross-validation, and synthesizer output carries highest scrutiny because structural failures at the synthesis layer are invisible to the platforms being synthesized.

Hallucination detection operated as a core function, with single-platform claims flagged for independent verification. Four fabricated findings from one platform (CoPilot) were detected through cross-platform non-confirmation and URL verification, then excluded from the evidence base. The fabrications followed a consistent pattern: plausible URLs constructed on real domains that returned 404 errors on live verification. A fifth misattribution (a shadow AI statistic attributed to a Microsoft report that does not contain the claimed figure) survived the initial research round as a SINGLE SOURCE finding and was detected and removed during the citation audit conducted after the editorial review round. All four initial fabrications, if accepted, would have prematurely closed the central research gap (the absence of actuarial validation for governance pricing). The detection prevented false closure of a real research question.

The final evidence base comprises 14 sources found by single platforms only (classified as highest discovery priority because they represent intelligence no single-platform researcher would surface through conventional search) and 27 sources confirmed across multiple platforms. Discovery priority reflects retrieval novelty, not source reliability. All findings, regardless of discovery method, were verified through independent URL confirmation and source-type classification before inclusion.

The methodology cannot detect convergent error where all platforms share the same training data mistake. The editorial review round revealed recurring disagreement over the Tier 3 definition, which the paper addresses through peer-reviewed evidence of definitional fragmentation (Jobin et al., 2019; Stix, 2022). Full convergent-error analysis is provided in the Methodology Appendix.

This paper operates in territory where no published empirical actuarial study, loss-run analysis, or regulatory framework quantifies governance maturity as a pricing factor for AI liability insurance. The analytical frameworks proposed here, including the five-tier insurance maturity model and the six-variable governance declaration, represent the author's professional judgment applied to observed market signals, regulatory actions, and academic research across insurance, AI governance, and technology policy. They are offered as analytical instruments to help chart a path forward in a field that does not yet have one, not as empirically validated market standards.

The evidence base documents what has happened, the frameworks propose how to interpret it, and the distinction between observation and interpretation is maintained throughout. Full methodology materials, including verbatim prompts and platform-specific outputs, are available for independent verification upon request.

What the insurance scholarship is showing

The academic foundation

Anat Lior published an empirical study of the AI insurance market in the Connecticut Insurance Law Journal in 2025 that traces the exclusion-to-product pattern in detail. Existing policy language often says nothing explicit about AI, creating what the field calls silent AI coverage. Silent coverage is risky for both sides because the insured does not know whether a claim will be paid and the carrier does not know whether it will have to pay one. Both sides have an incentive to resolve the silence, and the resolution is happening now through explicit exclusions and, increasingly, through affirmative coverage language from specialty carriers.

Crootof, Kaminski, and Price wrote a paper in the 2023 Vanderbilt Law Review titled *Humans in the Loop* that named the bad version of human oversight: the liability sponge. A liability sponge is a person stationed near an AI decision who lacks the time, the authority, the information, or the standing to actually decide. The AI runs at machine speed, the person clicks approve, and when something goes wrong the person absorbs the blame while the company absorbs the loss. Nothing about the system actually got governed.

Liu, Park, Wang, and Wen published the first peer-reviewed actuarial framework for AI liability pricing in the journal Risks in January 2026. Their paper, “Insuring Algorithmic Operations: Liability Risk, Pricing, and Risk Control,” develops a taxonomy of algorithmic operations liability risk sources and incorporates governance, documentation, model monitoring, and MLOps practices as loss-reduction mechanisms within an actuarial framework. The framework is theoretical, built from actuarial building blocks without the empirical claims data that would validate it, but it is the closest the academic literature has come to pricing governance maturity for AI risk.

Leung, Zhang, Ling, Toyoda, and Loh published a preprint in May 2026 mapping 55 AI threat classes against 26 insurance products, endorsements, and exclusion regimes. Their analysis identifies a four-tier insurability frontier: affirmatively insured perils, silent AI exposures, actively excluded perils, and perils outside conventional private insurance structures. The most significant finding for the emerging practice argument is their identification of “foundation model concentration” as “the clearest genuinely novel insurability frontier because upstream model failure can correlate losses across many cedents at once.”

The product market

The specialty AI liability insurance market is no longer theoretical. At least seven providers are now writing eight distinct AI liability or AI-related products, with different underwriting methodologies and governance requirements.

Armilla AI is a Lloyd's coverholder and, by its own account, the only managing general agent focused exclusively on AI insurance. Armilla launched the first standalone AI liability policy at Lloyd's in April 2025, expanded coverage limits to 25 million dollars per organization by January 2026, and launched Vanguard AI in February 2026 in partnership with Chaucer Group, a coordinated structure combining cyber, technology errors and omissions, and standalone AI liability with predefined allocation rules for mixed losses. Armilla requires independent AI system certification informed by more than 500 evaluations, and its partnership with A-LIGN connects ISO/IEC 42001 governance certification directly to underwriting terms.

Munich Re has underwritten AI performance risk since 2018 through its aiSure platform, making it the longest-running AI insurance program. The product is model-agnostic, uses a parametric-like claims structure based on measurable performance data, and expanded to 15 million dollars in coverage through a February 2026 partnership with Mosaic Insurance.

Testudo, a Lloyd's Lab managing general agent, launched in January 2026 with capacity up to 9.25 million dollars per insured, backed by Apollo, Atrium, and QBE at Lloyd's. Testudo uses a proprietary AI Risk Engine based on global litigation data, does not require invasive technical audits, and reports (based on its proprietary litigation database) that generative AI litigation has increased 137 percent year over year.

HSB, a Munich Re subsidiary, announced AI Liability Insurance for small and medium businesses in March 2026, covering bodily injury, property damage, and personal injury arising from AI use. The product is pending regulatory approval and will be distributed through partner carriers. An HSB survey found 74 percent of small and medium businesses currently use AI programs, with 91 percent planning to use AI in the future.

Relm Insurance launched three AI insurance products (NOVAAI, PONTAAI, and RESCAAI) in January 2025 from Bermuda. PONTAAI operates as an excess difference-in-conditions wrap designed specifically to address exclusion gaps in existing liability programs created by AI exclusions, directly supporting the thesis that exclusions create the market for standalone coverage.

Hiscox UK rewrote its Technology Professional Indemnity policy in May 2025 to provide the UK market's first affirmative AI liability coverage, covering algorithm failures, faulty data inputs, and negligent AI advisory services.

AIUC, backed by 15 million dollars in seed funding, offers AI agent insurance up to 50 million dollars through a standards-audit-insurance model: the AIUC-1 framework creates a technical and operational baseline, independent audits test real-world performance, and insurance policies cover the residual risk with pricing that reflects the audit results.

These products span multiple insurance categories: standalone third-party AI liability (Armilla, Testudo), AI performance guarantee (Munich Re aiSure), professional indemnity endorsement (Hiscox), excess difference-in-conditions wrap (Relm PONTAAI), coordinated multi-line structure (Vanguard AI), small business AI liability (HSB), and AI agent insurance (AIUC). They are grouped here not as equivalent products but as evidence that the market is building coverage for AI risk across multiple insurance structures simultaneously.

Two distinct underwriting models coexist and have not yet converged. Certification-based underwriting, as practiced by Armilla and AIUC, requires governance documentation before binding coverage, while litigation-data-based underwriting, as practiced by Testudo, prices from claims pattern analysis without requiring invasive governance audits. Both models are active simultaneously, and the market has not converged on a single standard. No multi-stakeholder standard for AI liability underwriting criteria currently exists. ISO/IEC 42001 is the closest signal, but it is a voluntary management systems standard, not an insurance industry pricing framework.

The coexistence of these models does not resolve the pricing gap this paper identifies. Litigation-data underwriting prices from observed loss patterns and does not require governance documentation from the insured. Certification-based underwriting prices from governance evidence and treats documentation as a predictor of future risk reduction. The actuarial question, whether governance documentation correlates with reduced AI claims frequency and severity, remains unanswered by either model.

The governance-insurance linkage

The A-LIGN and Armilla AI partnership, announced in December 2025, is one of the clearest verified commercial linkages between a governance certification standard and AI liability underwriting terms. Organizations that complete ISO/IEC 42001 certification through A-LIGN gain preferential access to Armilla's affirmative AI insurance. As Armilla stated, "linking ISO/IEC 42001 to underwriting transforms governance into an economic signal, one where stronger governance can influence insurability, pricing, and coverage scope."

Industry experts told Communications of the ACM in March 2026 that insurers are beginning to require what they call "compliance-grade observability." The article describes what interview subjects (including independent arbitrator Sophie Nappert and IBM security researchers) identified as emerging technical requirements: immutable audit trails with model and dataset fingerprints, versioned prompts and outputs, time-stamped retrieval citations, and signed change logs for safety settings. These are Tier 5 requirements. They cannot be satisfied by principles (Tier 2) or by automated controls alone (Tier 3). They require a named human who can attest to what was decided and why, with a structured record that survives external review.

How risk becomes cost

The market has been through a version of this loop before, in cyber.

Cyber insurance in the early 2010s was a mess of vague coverage and guesswork premiums. Then ransomware happened, claims got expensive, and underwriters started asking concrete questions about concrete controls. Did multi-factor authentication actually work in practice? Had the backup procedures been tested with restoration drills, or just configured on paper? Could the detection systems catch an intruder before the encryption finished running? Did the incident response plan exist beyond a binder on a shelf?

Companies that could answer those questions in writing renewed at reasonable rates, while those that could not paid more, got narrower coverage, or got nothing. The data from 2025 is

consistent with the mechanism working. Coalition's 2025 Cyber Claims Report documented that Coalition policyholders had 73 percent fewer claims than the industry average, as Coalition reports against its calculated NAIC frequency benchmark. NetDiligence analyzed 10,402 claims from incidents between 2020 and 2024, providing the most comprehensive cyber claims dataset available for benchmarking frequency and severity trends. Industry analysis of these trends supports the conclusion that preventative controls contribute to improved loss outcomes. AM Best reported that 2024 was the first year U.S. cyber direct premiums written declined since data collection started in 2015, a 2.3 percent drop driven in part by a market that had learned to price controls. Sophos found in its 2024 survey that of the 97 percent of cyber-policy holders who invested in improving defenses, 76 percent reported that improvements enabled them to qualify for coverage, 67 percent secured better pricing, and 30 percent obtained improved terms.

The market mechanism transfers: exclusion first, then documentation requirements, then conditional re-entry with pricing that rewards controls. This mechanism operated in cyber and is beginning to operate in AI.

What does not transfer is the risk structure, and three differences matter for the insurance market. First, cyber risk is primarily external: attackers breach defenses. AI liability is primarily internal: the organization's own system produces harmful outputs. The evidentiary requirements for demonstrating governance differ accordingly because AI governance must document internal decision-making, not external threat mitigation. Second, foundation model concentration creates systemic correlation risk that cyber never had at the same maturity stage. Kevin Kalinich, Aon's head of cyber, put the distinction in plain terms in November 2025: the industry could absorb a 400 million or 500 million dollar hit from one company's AI failure, but it cannot absorb an upstream failure that produces a thousand losses at once. Leung et al. call this "the clearest genuinely novel insurability frontier." Swiss Re's sigma report in January 2026 reported that AI creates "emerging risk dimensions that do not fit neatly within traditional insurance boundaries." Third, the absence of AI-specific loss data means the pricing feedback loop has not yet started, and cyber's pricing matured over roughly a decade of claims experience that AI does not yet possess.

The cyber precedent shows the mechanism but does not predict the timeline, the pricing structure, or the loss ratios.

The six questions underwriters are starting to ask

Six variables determine where any AI deployment lands on the spectrum between excluded and insurable. Some of these questions are already being asked in current underwriting questionnaires, confirmed by named sources in the insurance trade press and regulatory filings.

Deployment mode. Does the system produce a final result that goes out into the world without a human checkpoint? Or does it run to a named human who decides whether the output is acted on? Armilla's underwriting requires AI system certification that classifies deployment mode. The NAIC AI Systems Evaluation Tool, now in a 12-state pilot, includes Exhibit C requiring detailed information on high-risk AI systems and their decision-making influence.

Consequence architecture. A wrong ad placement and a wrong medical recommendation have different cost profiles. Leung et al.'s four-tier insurability frontier maps AI threat classes against insurance products and exclusion regimes, providing the first systematic framework for matching consequence type to coverage category.

Error reduction methodology. Every generative AI deployment carries exposure to hallucination, drift, bias, and incomplete output. As industry experts described to Communications of the ACM in March 2026, insurers are beginning to ask for "compliance-grade observability" including model and dataset fingerprints, versioned prompts, and replay harnesses. Underwriters want to see what the organization does to reduce known failure modes and how it records the reduction.

Verification quality. Who reviews the AI output before it reaches a consequential decision point, and what domain competence do they bring? Aon's 2026 AI Risk report found that D&O underwriters are paying closer attention to board oversight, public disclosures, risk registers, model testing, and third-party controls when assessing AI-related exposure. A Business Insurance article from April 2026 quoted John Farley, Managing Director of Gallagher's cyber practice, describing the underwriting questions carriers are now asking: "What models are you currently utilizing? How did the business decisions get made?"

Audit trail completeness. Does a documented record exist that leads back to a named human at a discrete moment? Without that record, the firm has a weaker claim defense, a weaker coverage argument, and no clean answer when a D&O carrier asks what governance documentation existed before the deployment was authorized.

Shadow AI exposure. Gartner reported that 41 percent of employees acquired, modified, or created technology outside IT visibility in 2022 (a shadow IT baseline, of which AI tools are an accelerating subset) and projects that share will rise to 75 percent by 2027. Gartner's May 2026 data from the Global Labor Market Survey (1Q26) shows that 88 percent of employees with enterprise AI access also use personal AI tools for business tasks. At the National Insurance Conference of Canada in October 2025, cybersecurity specialists warned that shadow AI is "rapidly emerging as a critical blind spot for underwriters and insureds alike." Industry reporting suggests that some 2026 cyber insurance policies may include condition precedent clauses addressing AI-related data handling, though no primary carrier policy forms confirming specific language were located during verification.[4]

These six questions converge into a proposed analytical framework: a six-variable governance declaration synthesized from observed carrier questions, regulatory examination tools, and emerging certification requirements. No carrier or regulatory body has adopted this exact six-variable structure. It is offered as an instrument for organizing the questions the market is beginning to ask, not as an observed industry standard.

Variable	Underwriter Question	Tier Link	Evidence Required
Deployment mode	Does the AI output reach the world without a human checkpoint?	Tier 3 vs Tier 4 boundary	System certification, deployment documentation
Consequence architecture	What is the cost profile if the AI output is wrong?	All tiers, severity scaling	Risk classification by use case
Error reduction methodology	How does the organization reduce hallucination, drift, and bias?	Tier 3+	Model fingerprints, versioned prompts, testing logs
Verification quality	Who reviews AI output with domain competence before a consequential decision?	Tier 4+	Named reviewer, domain credentials, review records
Audit trail completeness	Does a documented record lead back to a named human at a discrete moment?	Tier 4 vs Tier 5 boundary	Structured project records, timestamped approvals
Shadow AI exposure	What AI tools operate outside organizational visibility?	All tiers	Inventory of authorized and unauthorized AI use

An organization that can produce this declaration at Tier 4 or Tier 5 has a stronger underwriting submission. An organization that cannot faces the coverage gap the exclusion wave created.

These six questions also form a de facto underwriting triage across the five-tier model. A Tier 1 organization cannot answer any of them, and a Tier 2 organization can state principles but not controls. A Tier 3 organization can answer the first three through its technical documentation but cannot answer questions four and five because no named human stands at the decision point. A Tier 4 organization can answer all six, but the quality of the answer to question five depends on whether the audit record is structured or scattered. A Tier 5 organization answers all six with portable, structured documentation that survives external review.

Three pressure channels operating at once

The pressure on companies to answer these six questions is coming through three channels simultaneously, and they reinforce each other.

The market channel

The Berkley exclusion and the Verisk optional endorsement give carriers tools to remove or narrow AI-related claims from legacy policy language. Companies that can produce a governance declaration at Tier 4 or Tier 5 have a stronger underwriting position and a stronger claim defense. Companies that cannot enter renewal with less to point to and more for the broker to explain away.

The personal exposure for directors and officers who authorized AI deployments compounds the corporate coverage pressure. Derivative actions, regulatory inquiries, coverage disputes, and shareholder challenges all ask the same question after an AI-related incident: what governance record existed before the decision to deploy was made?

Tier status is not permanent. Organizations can regress from Tier 4 to Tier 3 as cost pressure increases, as personnel who understood the checkpoint process leave, or as speed-to-market incentives override governance discipline. The Economic Override (Puglisi, 2025), the structural tendency for deployment incentives to override governance regardless of institutional intent, operates continuously. The gap is measured: a 2025 EY survey of 975 C-suite leaders across 21 countries found 76 percent of organizations deploying or planning agentic AI while only a third have proper protocols to adhere to all facets of responsible AI controls (EY, 2025a). A follow-up EY study found 99 percent of surveyed organizations reported financial losses from AI-related risks, with an average loss conservatively estimated at $4.4 million, yet only 12 percent of C-suite respondents answered correctly when asked to match appropriate controls against five specific AI-related risks (EY, 2025b). Insurance renewal is the market enforcement moment. An organization that earned Tier 4 coverage in one cycle but cannot produce the same documentation at renewal has regressed, and the premium or exclusion consequence follows.

The U.S. regulatory channel

The NAIC AI Systems Evaluation Tool has moved into a 12-state pilot running from March through September 2026. Exhibit C requires detailed information on high-risk AI systems and their decision-making influence, creating a regulatory evidence pattern that parallels Tier 4 and Tier 5 documentation logic. Industry trade groups (ACLI, APCIA, Committee of Annuity Insurers) have pushed back on the pilot's terms, noting that it is one-sided and permits regulatory findings during the pilot phase. These objections confirm that governance examination is moving from voluntary guidance toward structured examination, and the carriers subject to examination are pushing back on the terms, not on the principle. The NAIC published an issue brief in March 2026 opposing federal preemption of state AI oversight, affirming state-based regulation under the McCarran-Ferguson framework. Colorado has implemented the first state-level rule requiring insurers to formally test AI systems for unfair discrimination against protected classes, and Armilla's AI liability policy explicitly covers Colorado AI Act defense costs.

The EU regulatory channel

The EU Revised Product Liability Directive, adopted in October 2024 with a December 9, 2026 transposition deadline, brings software and AI systems into the strict liability framework as "products." Its most consequential provision for AI deployers is its linkage to the EU AI Act: as Freshfields argues in its April 2026 legal commentary, non-compliance with the AI Act's safety requirements directly informs a court's assessment of product defectiveness under the PLD.[3] Finland presents the strictest liability environment because it never adopted the development risk defense under the 1985 Directive and is maintaining that position under the new PLD, meaning manufacturers cannot escape liability by arguing that the defect was undetectable at deployment. Article 20 requires the European Commission to evaluate the Directive's application by December 9, 2030, including "the availability of product liability insurance," building a review mechanism that specifically asks whether AI liability insurance exists across 27 member states.

The demand channel

The Geneva Association surveyed 600 corporate insurance decision-makers across China, France, Germany, Japan, the United Kingdom, and the United States in 2025. More than 90 percent expressed the need for insurance coverage tailored to AI and generative AI risks, and

more than two-thirds said they would pay at least 10 percent higher premiums for such coverage. Seventy-one percent had already implemented generative AI in at least one business function. The demand exists and the products are emerging, but the pricing mechanism that connects governance maturity to premium differentials is the gap that remains open.

What this argument does not prove

Six objections deserve explicit treatment.

Exclusion is normal new-risk-class behavior. It is. Carriers exclude what they cannot yet price, and that pattern has repeated for asbestos, environmental liability, terrorism, and cyber before AI. The difference is speed. The Verisk exclusion moved from filing to market-wide availability in three months with no legislative vote, no implementation period, and no public comment window. The approval rate reported by secondary sources, as high as 80 percent according to Wolfe Research, has not been confirmed through primary regulatory data. The directional pattern, that regulators are approving AI exclusion filings at scale, is supported by the observable carrier evidence. The cyber exclusion-to-pricing cycle took roughly a decade to mature. The AI cycle appears to be compressing, though the final timeline remains unknown.

Regulation is already moving, so insurance is not the sole forcing function. Twenty-four states plus the District of Columbia have adopted the NAIC bulletin. The EU AI Act follows a phased implementation schedule, with most remaining obligations becoming generally applicable on 2 August 2026; however, the May 2026 provisional Omnibus agreement between Parliament and Council would postpone core high-risk AI system obligations to December 2027, and product-embedded high-risk provisions would extend to August 2028 if formally adopted (Regulation (EU) 2024/1689; Council of the European Union, 2026). The EU PLD transposition deadline falls in December 2026. These are real regulatory forces, and they converge with market pressure on parallel tracks. The paper argues that market and regulatory channels reinforce each other, not that insurance operates alone. Insurance moves faster because it requires no legislative process: a carrier can file an exclusion endorsement and have it approved within months. That speed makes insurance the leading edge of the emerging practice, with regulation following on a parallel track. Voluntary compliance mechanisms produce limited results without external enforcement: Stanford HAI found that 16 AI companies scored an average of 52 percent against White House voluntary commitments, with compliance rates ranging from 13 percent (Apple) to 83 percent (OpenAI) (Wang et al., 2025). EY's 2025 survey found only 12 percent of C-suite respondents answered correctly when asked to match appropriate controls against five specific AI-related risks. Insurance conditions add the enforcement mechanism that voluntary commitments lack.

No actuarial evidence validates governance as a pricing factor. This is the paper's central identified gap, and it is stated without qualification. No published empirical actuarial study, loss-run analysis, or standardized underwriting framework was located in this research that quantifies governance maturity (audit trails, checkpoint records, named human accountability) as a pricing factor for AI liability insurance. Liu et al. (2026) provide a theoretical actuarial framework that incorporates governance as a loss-reduction mechanism, but it is built from actuarial building blocks rather than from empirical claims data. The Geneva Association documents demand (90

percent want coverage) but not supply-side pricing mechanics. Aon observes D&O underwriting scrutiny around governance but does not quantify premium differentials.

The actuarial gap does not close through better predictive models alone, because it closes through verification infrastructure that resolves the information asymmetry between governed and ungoverned organizations. An insurer cannot price governance it cannot verify. The Geneva Association identified information asymmetry as the primary insurability challenge for AI risk. Third-party verification of governance maturity, such as the ISO/IEC 42001 certification pathway that Armilla and A-LIGN have linked directly to underwriting terms, resolves that asymmetry at the individual organization level. At scale, structured governance records from Tier 5 organizations produce the loss data the actuarial community needs.

The Responsible AI definition is contested. Addressed by the five-tier model. Responsible AI occupies Tier 3, correctly recognized as a real and necessary stage that produces technical controls no governance system should operate without. The paper's argument is not that Tier 3 is incorrectly defined but that the insurance market requires Tier 4 or Tier 5 evidence. The gap between responsible AI as defined and responsible AI as evidenced in an insurance context is an implementation gap, not a definitional error. The five-tier model is not a replacement for NIST AI RMF, ISO/IEC 42001, or the EU AI Act's risk classification. It addresses a different question: not how to manage AI risk in general, but how the insurance market is beginning to sort AI deployments by governance maturity for coverage and pricing decisions.

Exclusions are driven by capital preservation, not governance improvement. A defender of the current insurance industry would argue that carriers are issuing AI exclusions to prevent catastrophic correlated losses across their existing books from silent AI exposure, not to force better corporate governance. If a Tier 5 organization with structured audit records applies for coverage, a carrier might still deny the application because the aggregation risk of foundation model failures is uninsurable at any governance tier. This objection has force. The paper does not argue that carriers are acting as quasi-regulators with a governance improvement mandate. It argues that the structural consequence of their capital preservation decisions is a de facto governance sorting mechanism, regardless of the carriers' intent. The standalone products emerging from Lloyd's, Munich Re, and specialty MGAs exist because carriers saw a commercial opportunity in the gap the exclusions created. The governance requirement attached to those products is a market condition, not a regulatory mandate.

The financial exposure is documented: EY's 2025 survey found 99 percent of organizations reported financial losses from AI-related risks, with nearly two-thirds suffering losses exceeding one million dollars (EY, 2025b). The governance gap that carriers are excluding from coverage produces measurable losses whether or not the carriers intend to incentivize governance improvement.

Behavioral change is not yet observed. No survey, case study, or broker attestation was located showing organizations upgrading AI governance specifically in response to insurance exclusions. The emerging practice described in this paper is structural, not yet empirically observed at the organizational level. Whether organizations respond to coverage denial by adopting governance, by accepting uninsured risk, or by abandoning high-risk AI deployments is an empirical question this paper identifies but cannot answer from the current evidence. The Sophos 2024 cyber

survey, in which 76 percent of companies improved defenses specifically to qualify for coverage, provides the closest behavioral precedent from an adjacent risk class.

The bottom line

The five tiers the industry has been collapsing into a single phrase are becoming five distinct insurance profiles. Ethical AI identifies the boundary and Responsible AI documents the system, but neither produces what an insurer or a court wants to see when a claim arrives. AI Governance produces the named human and the defined checkpoint. AI Governance with structured audit records produces the portable, structured evidence that survives the audit itself.

The emerging practice creates a self-reinforcing cycle. Governance generates the structured decision records that produce actuarial loss data, and loss data makes AI risk priceable. Priceable risk enables the insurance market to write affirmative coverage, which in turn creates economic incentive for governance. The cycle starts the moment an organization produces its first auditable checkpoint record.

The market is beginning to sort AI deployments by governance maturity through coverage access, not yet through validated premium differentials. The exclusion endorsements that began appearing publicly in 2024 are the accelerant in a cycle that has already played out once in cyber, though the structural differences between the two risk classes mean the AI timeline, pricing structure, and loss ratios remain unpredictable. The actuarial validation that would close the pricing loop, a published study linking governance documentation to reduced AI claims frequency and severity, does not yet exist. That gap is the research agenda this paper identifies. What does exist is the structural mechanism: standard coverage disappears, specialty coverage emerges with governance conditions, and the cost differential between governed and ungoverned AI deployments creates incentive for adoption. Companies that act on the distinction between Tier 1 and Tier 5 now may face stronger renewal positions. The rest face 2026 and 2027 renewal cycles with fewer options and less to show for their AI investments than their governance-ready competitors.

Acknowledgments

This working paper was inspired by a cross-disciplinary conversation hosted by The Intention Advantage, a Summit Series hosted and organized by the Institute for Advertising Ethics (IAE), with the IEEE Standards Association serving as knowledge partner. The author thanks Andrew Susman, President of the Institute for Advertising Ethics, for the founding convening held April 22, 2026 at Frankfurt Kurnit Klein & Selz in New York, where the unique connections between advertising governance, agentic AI accountability, and the insurance pricing of human oversight surfaced for a cross-disciplinary audience. Contributions from IEEE Standards Association leadership, including a representative of the IEEE Global Initiative on Ethics of Autonomous and Intelligent Systems, shaped the architecture of the argument developed here. Participation in the convening does not imply institutional endorsement of this paper's analysis or conclusions. Responsibility for the framework, the evidence base, and any errors remains with the author.

Conflict of Interest

The author developed Checkpoint-Based Governance (CBG), a governance framework referenced in this analysis. CBG is published open-source at github.com/basilpuglisi/HAIA and described in Puglisi (2025), Governing AI: When Capability Exceeds Control (ISBN 9798349677687). The five-tier model proposed in this paper draws on the author's prior governance research. The five-tier insurance maturity model is the author's analytical contribution and has not been adopted by any insurance carrier, regulator, or standards body. The author has no financial interest, consulting relationship, equity position, or commercial affiliation with any insurance carrier, MGA, broker, or regulator discussed herein, and receives no compensation from the adoption of CBG.

Footnotes

1. Secondary commentary frequently describes the Berkley form as applying to E&O coverage. The primary form reviewed for this paper, PC 51380 00 (06-24), amends Directors and Officers Liability, Employment Practices Liability, and Fiduciary Liability coverage parts. This paper follows the primary form language. Independent verification during the May 2026 research round confirmed the EPLI scope against the form text hosted at hunton.com.

2. Industry commentary frequently cites a figure of approximately 82 percent market share for ISO/Verisk forms in U.S. commercial liability. No primary Verisk, ISO, NAIC, or regulatory source confirming this percentage was located during verification. This paper treats the figure as industry-reported.

3. The Freshfields analysis is legal commentary, not the Directive text itself. The PLD-AI Act linkage is an interpretive reading of how non-compliance with AI Act safety requirements may affect defectiveness assessment, not an explicit provision of the Directive.

4. Questa AI, a vendor of local data redaction products, reported in March 2026 that some major carriers include condition precedent clauses addressing AI-related data handling in 2026 cyber policies. No primary carrier policy forms confirming this specific language were located during verification. This should be treated as a directional signal from a commercially interested source, not as verified carrier policy.

Methodology Appendix

Working Paper Origin

The analysis in this paper developed from a working paper researched and written by the author and shared with The Intention Advantage, a working initiative of the Institute for Advertising Ethics (IAE), for discussion at a follow-up meeting after the founding convening held April 22, 2026. That working paper (v5.0) provided the initial evidence base and the three-tier governance framework. The present paper (v6.0, revised through v6.92) represents the next stage: a parallel multi-AI research expansion, independent source verification, five-tier model development, and hostile editorial peer review.

Research Round

Identical structured research prompts were dispatched simultaneously across 12 independent AI platforms:

Claude (Anthropic), ChatGPT (OpenAI), Gemini (Google), Grok (xAI), Perplexity, Mistral, DeepSeek, Kimi (Moonshot AI), MiniMax, Meta AI, Apertus, and CoPilot (Microsoft).

The prompt specified six research areas: AI exclusion endorsements, standalone AI liability products, NAIC and state regulatory actions, EU Product Liability Directive implementation, governance-pricing linkage, and shadow AI exposure. A five-step verification protocol required each platform to classify every finding by source type (primary or secondary), provide a working URL, include a publication date, state cross-confirmation status, and flag provisional items. The verbatim research prompt is reproduced in Appendix A.

Source-Authority Discrimination

The methodology applies source-authority discrimination at three tiers: human arbiter input (the author's own verified findings and binding decisions) carries highest weight; raw AI platform output carries standard weight subject to cross-validation; and synthesizer output carries highest scrutiny because structural failures at the synthesis layer are invisible to the platforms being synthesized.

Hallucination Detection Results

Four fabricated findings from one platform (CoPilot) were detected through cross-platform non-confirmation and URL verification, then excluded:

1. A claimed Milliman white paper on AI liability actuarial modeling (URL returned 404)
2. A claimed Lloyd's Market Association AI risk framework (URL returned 404)
3. A claimed AIG Shadow AI Exclusion filing (no primary filing located)
4. A claimed NAIC bulletin misdated to December 2025 (correct date: December 2023)

A fifth misattribution (a shadow AI statistic attributed to the Microsoft Data Security Index 2026 that does not appear in the actual report) survived the initial research round as a SINGLE SOURCE finding and was detected and removed during the citation audit conducted after the editorial review round.

Editorial Review Round

The v6.0 draft was submitted to 10 independent AI platforms for hostile-but-fair critical peer review:

Gemini (Google), Grok (xAI), ChatGPT (OpenAI), Perplexity, Mistral, DeepSeek, Kimi (Moonshot AI), MiniMax, Meta AI, and Claude (Anthropic, outside the working project).

The editorial prompt specified nine review dimensions: evidence sufficiency, definitional precision, logical structure, scope control, counterarguments, empirical vulnerability, citation integrity, terminology risk, and publication readiness. Severity ranking (critical, significant, minor) was required. The verbatim editorial prompt is reproduced in Appendix B.

Synthesis and Arbitration

Both rounds were synthesized in Claude (Anthropic) operating in the Liaison/Navigator role within the author's working project. The author served as human arbiter for all binding decisions, including source inclusion/exclusion, definitional rulings, framework provenance, language choices, and scope decisions. The synthesis produced 22 revision items, each with documented convergence counts across platforms, platform-specific findings, and arbiter decisions. No revision was executed without explicit arbiter approval.

Convergent Error Limitation

The methodology cannot detect convergent error where all platforms share the same training data bias. The editorial review round revealed recurring disagreement over the Tier 3 (Responsible AI) definition, with 9 of 10 editorial platforms converging on the position that the paper's definition contradicts industry consensus. This paper treats that convergence as a substantive definitional risk rather than resolving it by platform count. The definitional critique is addressed substantively in the five-tier section through peer-reviewed evidence of definitional fragmentation (Jobin et al., 2019; Stix, 2022) and direct comparison of competing institutional frameworks.

Appendix A: Research Prompt (Verbatim)

Role: Researcher

Task: Research the current state of AI liability insurance, AI exclusion language in commercial policies, and the structural relationship between insurance pricing and AI governance adoption. Use live sources only. Do not rely on training data for any claim about insurance products, carrier positions, or regulatory actions after January 2024.

VERIFICATION PROTOCOL (apply to every finding before including it):

Step 1: Source Classification. For every finding, classify the source as PRIMARY (the carrier filing, the regulatory document, the statute, the court record, the company announcement, the peer-reviewed paper) or SECONDARY (a news article, blog post, analyst report, or commentary about a primary source). If you can only find secondary sources for a claim, say so explicitly and provide the secondary source with the label [SECONDARY, PRIMARY NOT LOCATED].

Step 2: URL Verification. Provide the exact URL for every source cited. If you cannot produce a working URL that leads to the specific document or page containing the claim, do not include the finding. State instead: [CLAIM FOUND IN TRAINING DATA, LIVE SOURCE NOT CONFIRMED]. Do not fabricate, reconstruct, or approximate URLs.

Step 3: Date Verification. Every source must include a publication date. If the source has no visible publication date, label it [UNDATED] and note that the recency cannot be confirmed.

Step 4: Cross-Check. For any finding that relies on a single source, state [SINGLE SOURCE]. For any finding confirmed by two or more independent sources, state [CROSS-CONFIRMED] and list both. Do not present single-source findings with the same confidence as cross-confirmed findings.

Step 5: Provisional Flag. If a finding appears likely true based on available evidence but cannot be fully verified through the steps above, label it [PROVISIONAL] with an explanation of what verification is missing. I will verify it independently before using it.

RESEARCH AREAS:

1. AI exclusion endorsements issued by commercial insurance carriers since June 2024. Confirm or update the status of W.R. Berkley Form PC 51380 (absolute AI exclusion), Verisk ISO Form CG 40 47 01 26 (optional generative AI exclusion, effective January 2026), and any AI-related exclusion filings by AIG, Great American, or other carriers. Identify any new exclusion language or AI-specific endorsements issued in 2026.

2. The current state of AI liability as an insurable risk class. Are any carriers actively writing AI-specific liability products (not cyber policies with AI add-ons, but standalone AI liability coverage)? What underwriting criteria are they using? What evidence do they require from the insured?

3. Regulatory actions by the NAIC, state insurance departments, or EU bodies related to AI risk in insurance underwriting. Confirm the status of the NAIC AI Model Bulletin adoption, the NAIC AI Systems Evaluation Tool pilot, and any 2026 developments.

4. The EU Revised Product Liability Directive (EU) 2024/2853 transposition deadline of December 9, 2026. What is the current implementation status across EU member states, and how does it affect AI deployers specifically?

5. Academic or industry research on the relationship between governance documentation (audit trails, checkpoint records, named human accountability) and insurance claims outcomes for AI-related incidents. Are there any published actuarial studies, loss run analyses, or underwriting frameworks that price governance maturity?

6. The Gartner projection that 75 percent of employees will acquire or modify technology outside IT visibility by 2027. Confirm the original source, check for updated figures, and research any insurance carrier response to shadow AI as a coverage concern.

OUTPUT FORMAT:

For each finding, deliver in this structure:

FINDING: [statement] SOURCE TYPE: [PRIMARY / SECONDARY / SECONDARY, PRIMARY NOT LOCATED] URL: [exact URL or CLAIM FOUND IN TRAINING DATA, LIVE SOURCE NOT CONFIRMED] DATE: [publication date or UNDATED] CONFIDENCE: [CROSS-CONFIRMED / SINGLE SOURCE / PROVISIONAL] QUOTE OR DATA POINT: [specific language from the source, under 15 words] CONFLICTS: [any contradictions with other findings or with the baseline positions below]

BASELINE POSITIONS TO TEST (flag anything that contradicts or updates these): - The first market response to AI exposure was exclusion language, not governance pricing - No multi-stakeholder standard for AI liability underwriting criteria currently exists - Insurance is the forcing function that makes governance adoption commercially mandatory without requiring regulation

Do not summarize. Do not fill gaps with plausible-sounding content. If a research area returns no verifiable findings, state that explicitly rather than generating approximate answers. An honest gap is more useful than a fabricated source.

Appendix B: Editorial Review Prompt (Verbatim)

Role: Editor (Critical Review Mode)

Note: The editorial review prompt below references v5.0 and the original three-tier governance framework. The model evolved to five tiers during the revision process documented in this paper. The prompt is reproduced verbatim to preserve methodological transparency.

Task: Review the attached working paper "The AI Risk Economy: Why Insurance Cannot Price What Governance Cannot Prove" (v5.0, May 2026) as a critical reviewer preparing this paper for peer review submission. Do not praise the paper. Identify every weakness, gap, unsupported claim, structural problem, and vulnerability a hostile but fair academic reviewer would find.

Evaluate the following dimensions:

1. CLAIMS WITHOUT SUFFICIENT EVIDENCE. Identify every assertion in the paper that is stated as fact but lacks a cited source, relies on a single source, or extrapolates beyond what the cited source actually says. Flag any claim where the cited source is secondary commentary rather than a primary document.

2. DEFINITIONAL PRECISION. The paper defines three tiers: Ethical AI, Responsible AI, and AI Governance. Test whether these definitions hold consistently throughout the paper. Are there passages where the paper uses "governance" loosely in a way that contradicts its own definition? Are there passages where "responsible AI" is used in a way the industry would dispute?

3. LOGICAL STRUCTURE. Map the argument from premise to conclusion. Identify any logical leaps, missing steps, or places where the paper assumes the reader accepts a premise that has not been established. Flag any circular reasoning.

4. SCOPE CREEP. The paper begins with insurance pricing and the three-tier framework. Does it stay within that scope, or does it expand into territory (policy proposals, technical architecture, measurement frameworks) that weakens the core argument by trying to do too much?

5. COUNTERARGUMENTS NOT ADDRESSED. What would a defender of the current insurance industry approach say in response? What would a responsible AI advocate say about the paper's narrow definition of their field? What would a regulator say about the claim that insurance is more effective than regulation as a forcing function? The paper must anticipate and address its strongest critics.

6. EMPIRICAL VULNERABILITY. The paper relies on three insurance moves (Berkley, Verisk, AIG/Great American) as evidence of a market shift. Is three data points sufficient to support the claim of a structural market transformation? What would strengthen the empirical basis?

7. CITATION INTEGRITY. Check every citation format for consistency. Flag any citation that cannot be independently verified from the information provided. Flag any source that appears to be secondary commentary presented as primary evidence.

8. TERMINOLOGY RISK. Are there terms the paper uses that carry different meanings in insurance, legal, and technology contexts? Flag any term where a reader from one discipline would misunderstand the paper's intent because the term means something different in their field.

9. PUBLICATION READINESS. What specific revisions would bring this paper to submission-ready quality for a peer-reviewed journal in AI governance, insurance law, or technology policy? Name the journals where this paper would find the most receptive and rigorous review.

Deliver findings as a numbered list organized by severity: critical (must fix before submission), significant (weakens the paper if not addressed), and minor (polish items). Do not soften the critique. The goal is to find every problem before a reviewer does.

Primary Sources

Carrier Filings and Market Data

W.R. Berkley Corporation. Form PC 51380 00 (06-24) Artificial Intelligence Exclusion (Absolute). https://www.hunton.com/assets/htmldocuments/noindex/PC-51380-00-06-24-Artificial-Intelligence-Exclusion-Absolute.pdf

Hamilton Select Insurance Inc. Form PL3026 (06/23) Exclusion, Generative Artificial Intelligence. Filed in Case 2:24-cv-02577-EEF-MBN, Document 1-2 (October 30, 2024). https://www.hunton.com/assets/htmldocuments/noindex/Hamilton-EXCLUSION-GENERATIVE-ARTIFICIAL-INTELLIGENCE.pdf

Verisk ISO Form CG 40 47 01 26, CG 40 48 01 26, CG 35 08 01 26 (effective January 1, 2026). Form PDFs: https://assets.alm.com/63/68/46ed4bf34a0e807c9695e15c9e19/cg-40-48-01-26-exclusion-generative-artificial-intelligence-coverage-b-only.pdf and https://assets.alm.com/3f/6f/918870894682a2e4a733bb0229fd/cg-35-08-01-26-exclusion-generative-artificial-intelligence.pdf

The Information (April 23, 2026), citing Wolfe Research analysis of AI exclusion filing data [SECONDARY, proprietary analyst source]. https://www.theinformation.com/articles/berkshire-hathaway-chubb-win-approval-drop-ai-insurance-coverage

Financial Times (November 23, 2025). Insurers retreat from AI cover as risk of multibillion-dollar claims mounts. https://www.ft.com/content/abfe9741-f438-4ed6-a673-075ec177dc62

Insurance Products and Underwriting

Armilla AI. Chaucer and Armilla launch new AI liability insurance product. April 23, 2025. https://www.armilla.ai/resources/chaucer-and-armilla-launch-new-ai-liability-insurance-product

Armilla AI. Armilla AI raises Lloyd's-backed coverage to $25M. January 21, 2026. https://www.desmoinesregister.com/press-release/story/25455/armilla-ai-raises-lloyds-backed-coverage-to-25m-as-traditional-insurers-retreat-from-ai-risk/

Chaucer Group. Chaucer and Armilla AI launch Vanguard AI coordinated insurance structure. February 10, 2026. https://www.chaucergroup.com/news/press-release-chaucer-and-armilla-ai-launch-vanguard-ai-coordinated-insurance-structure

Testudo. Testudo launches new insurance coverage for liability risks created by generative AI systems. January 21, 2026. https://www.testudo.co/insights/testudo-launches-new-insurance-coverage-for-liability-risks-created-by-generative-ai-systems

FinTech Global. Testudo expands AI liability capacity to $9.25M. March 9, 2026. https://fintech.global/2026/03/09/testudo-expands-ai-liability-capacity-to-9-25m/

HSB. (2026, March 18). *HSB introduces AI liability insurance for small businesses* [Press release]. Munich Re Group. https://www.munichre.com/hsb/en/press-and-publications/press-releases/2026/2026-03-18-introducing-ai-liability-insurance-for-small-businesses.html

Munich Re. Insure AI platform. https://www.munichre.com/en/solutions/for-industry-clients/insure-ai.html

Relm Insurance. Relm Insurance launches AI liability solutions. January 14, 2025. https://finance.yahoo.com/news/relm-insurance-launches-ai-liability-140000469.html

Hiscox UK. Hiscox launches first affirmative AI liability cover in UK market. Insurance Business UK. July 3, 2025. https://www.insurancebusinessmag.com/uk/news/cyber/hiscox-launches-first-affirmative-ai-liability-cover-in-uk-market-540863.aspx

Fortune. AIUC emerges from stealth with $15 million seed. July 23, 2025. https://fortune.com/2025/07/23/ai-agent-insurance-startup-aiuc-stealth-15-million-seed-nat-friedman/

A-LIGN and Armilla AI. Turnkey program linking ISO/IEC 42001 certification to AI liability insurance. December 17, 2025. https://iccwbo.einnews.com/pr_news/876198902/a-lign-and-armilla-ai-launch-turnkey-program-linking-iso-iec-42001-certification-to-ai-liability-insurance

Academic and Industry Research

Lior, A. (2025). E/Insuring the AI Age: Empirical Insights into Artificial Intelligence Liability Policies. 31 Conn. Ins. L.J. 99. SSRN Abstract ID 5316376.

Crootof, R., Kaminski, M. E., & Price, W. N., II. (2023). Humans in the Loop. 76 Vand. L. Rev. 429.

Liu, Z., Park, J., Wang, M., & Wen, H. (2026). Insuring Algorithmic Operations: Liability Risk, Pricing, and Risk Control. Risks, 14(2), 26. https://doi.org/10.3390/risks14020026

Leung, A., Zhang, R., Ling, E., Toyoda, K., & Loh, S. (2026). The Insurability Frontier of AI Risk: Mapping Threats to Affirmative Coverage, Silent Exposures, and Exclusions. arXiv:2605.18784. https://arxiv.org/abs/2605.18784 [Preprint]

Geneva Association. (2025). Gen AI Risks for Businesses: Exploring the Role for Insurance. October 2, 2025. https://www.genevaassociation.org/sites/default/files/2025-10/gen_ai_report_0110.pdf

Aon. (2026). AI Risk 2026: What Business Leaders Need to Know. March 26, 2026. https://www.aon.com/en/insights/articles/ai-risk-2026-practical-agenda

Swiss Re Institute. (2026). sigma insights 01-2026: AI adoption is reshaping the risk landscape. January 13, 2026. https://www.swissre.com/institute/research/sigma-research/sigma-insights-01-2026-AI-adoption-is-reshaping-the-risk-landscape.html

Communications of the ACM. (2026). AI Liability Insurance Arrives. March 6, 2026. https://cacm.acm.org/news/ai-liability-insurance-arrives/ [News article with expert interviews, not peer-reviewed research]

Parasuraman, R., & Riley, V. (1997). Humans and automation: Use, misuse, disuse, abuse. Human Factors, 39(2), 230-253.

Skitka, L. J., Mosier, K. L., & Burdick, M. (1999). Does automation bias decision-making? International Journal of Human-Computer Studies, 51(5), 991-1006.

Jobin, A., Ienca, M., & Vayena, E. (2019). The global landscape of AI ethics guidelines. Nature Machine Intelligence, 1(9), 389-399.

Stix, C. (2022). The ghost of AI governance past, present, and future. Discover Artificial Intelligence, 2, Article 13.

Vaccaro, M., Almaatouq, A., & Malone, T. (2024). When combinations of humans and AI are useful: A systematic review and meta-analysis. Nature Human Behaviour, 8(12), 2293-2303. https://doi.org/10.1038/s41562-024-02024-1

Stanford HAI. (2024). AI Index Report 2024: Responsible AI chapter. Stanford Institute for Human-Centered Artificial Intelligence. https://hai.stanford.edu/ai-index/2024-ai-index-report/responsible-ai

Wang, J., Huang, K., Klyman, K., & Bommasani, R. (2025). Do AI companies make good on voluntary commitments to the White House? arXiv:2508.08345. https://arxiv.org/abs/2508.08345

EY. (2025a). Responsible AI Pulse Survey: AI adoption outpaces governance as risk awareness among the C-suite remains low. June 2025. 975 C-suite leaders, 21 countries. https://www.ey.com/en_gl/newsroom/2025/06/ey-survey-ai-adoption-outpaces-governance

EY. (2025b). Companies advancing responsible AI governance linked to better business outcomes. October 2025. 975 C-suite leaders, 21 countries. https://www.ey.com/en_gl/newsroom/2025/10/ey-survey-companies-advancing-responsible-ai-governance-linked-to-better-business-outcomes

Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 laying down harmonised rules on artificial intelligence (Artificial Intelligence Act). *Official Journal of the European Union*, L 2024/1689. https://eur-lex.europa.eu/eli/reg/2024/1689/oj

Council of the European Union. (2026, May 7). *Artificial intelligence: Council and Parliament agree to simplify and streamline rules* [Press release]. https://www.consilium.europa.eu/en/press/press-releases/2026/05/07/artificial-intelligence-council-and-parliament-agree-to-simplify-and-streamline-rules/

Cohen Seglias Pallas Greenhall & Furman PC. (2026, April 21). New generative AI insurance exclusion: What the construction industry needs to know. *JD Supra*. https://www.jdsupra.com/legalnews/new-generative-ai-insurance-exclusion-5225628/

Insurance Services Office [ISO]/Verisk. (2026). *Exclusion: Generative artificial intelligence* (Form CG 40 47 01 26). Verisk Analytics. [Available through Verisk/ISO subscription]

Author's Prior Work

Puglisi, B. C. (2025). Governing AI: When Capability Exceeds Control. ISBN 9798349677687.

Puglisi, B. C. (2026a). AI Governance Has No Formal Definition. Here Is One. basilpuglisi.com. https://basilpuglisi.com/ai-governance-has-no-formal-definition-here-is-one/

Puglisi, B. C. (2026b). The Great AI Language Collapse: Why Marketing Is Killing Accountability. basilpuglisi.com. https://basilpuglisi.com/the-great-ai-language-collapse-why-marketing-is-killing-accountability/

Cyber Precedent Data

Coalition Inc. (2025). 2025 Cyber Claims Report. May 7, 2025. https://web.coalitioninc.com/download-2025-cyber-claims-report.html

NetDiligence. (2025, September 17). *Cyber claims study 2025* [Report]. https://netdiligence.com/cyber-claims-study-2025-report/

AM Best. (2025). Best's Market Segment Report: 2024 Pricing Cuts in U.S. Cyber Generated First-Ever Reduction in Direct Premiums Written. June 23, 2025.

Sophos. (2024). Cyber Insurance and Cyber Defenses 2024. June 2024. https://www.sophos.com/en-us/press/press-releases/2024/06/76-companies-improved-their-cyber-defenses-qualify-cyber-insurance

Legal Authorities

Moffatt v. Air Canada, 2024 BCCRT 149 (British Columbia Civil Resolution Tribunal, February 14, 2024).

Regulatory Documents

NAIC. AI Model Bulletin (adopted December 4, 2023). Adoption map: https://content.naic.org/sites/default/files/cmte-h-big-data-artificial-intelligence-wg-map-ai-model-bulletin.pdf

NAIC. AI Systems Evaluation Tool 4.0. https://content.naic.org/sites/default/files/inline-files/AI%20Systems%20Evaluation%20Tool%204.0%20(Clean).pdf

NAIC. ACLI, APCIA, and industry comment letters on AI Systems Evaluation Tool. September 19, 2025. https://content.naic.org/sites/default/files/inline-files/_Comments-AI-Sys-Eval-Tool_Combined%20as%20of%202025-09-19_0.pdf

NAIC. Artificial Intelligence and State Insurance Regulation. Issue Brief, March 2026. https://content.naic.org/sites/default/files/ai-issue-brief.pdf

EU Revised Product Liability Directive (EU) 2024/2853. https://eur-lex.europa.eu/eli/dir/2024/2853/oj/eng

Freshfields. Product Risks Today: How the new Product Liability Directive turns AI Act compliance into a question of liability. April 21, 2026. https://www.freshfields.com/en/our-

thinking/blogs/risk-and-compliance/product-risks-today-how-the-new-product-liability-directive-turns-ai-act-complia-102mpu2 [Legal commentary]

CMS. Transposition Time Updates: Nordic Trio. February 12, 2026. https://cms-lawnow.com/en/ealerts/2026/02/transposition-time-updates-nordic-trio-denmark-finland-and-sweden-progress-towards-new-eu-product-liability-directive-implementation [Legal commentary]

Insurance Business Canada. The next big cyber blind spot: Shadow AI. October 9, 2025. https://www.insurancebusinessmag.com/ca/news/cyber/the-next-big-cyber-blind-spot-its-the-ai-your-employees-are-already-using-552366.aspx

Questa AI. AI Security Riders: Why 2026 Cyber Insurance Requires Local Redaction. March 2026. https://www.questa-ai.com/privacy-cafe/ai-security-riders-why-2026-cyber-insurance-requires-local-redaction [Vendor source]

Shadow AI Data

Gartner. CISO Role page (75% by 2027 projection). https://www.gartner.com/en/cybersecurity/role/chief-information-security-officer

Gartner. Security & Risk Management Summit EMEA 2025 Day 2 Highlights. September 23, 2025. https://www.gartner.com/en/newsroom/press-releases/2025-09-23-gartner-security-and-risk-management-summit-emea-2025-day-2-highlights0

Gartner. Global Labor Market Survey (1Q26). May 13, 2026 press release (88% dual-use finding). https://www.gartner.com/en/newsroom/press-releases/2026-05-13-gartner-predicts-by-2027-50-percent-of-enterprises-without-a-people-centric-ai-strategy-will-lose-their-top-ai-talent

Term List

AI Governance (Tier 4). A named human with binding authority at a defined decision checkpoint. Accountability leads back to a specific desk, a specific person, and a specific moment. Produces an auditable record. The first tier of the five that produces what an insurer can price and a court can review.

Audit Trail A documented record of who reviewed an AI output, under what authority, at what moment, and what decision was made. A foundation for underwriting assessment and claims defense in AI insurance, though not yet validated as an actuarial pricing input through published empirical research.

Checkpoint A defined hard stop in an AI workflow where a named human with domain competence reviews the output and makes a binding decision before the process continues. Distinct from a loop (no defined stop, no discrete record) and distinct from human-in-the-loop (no hard stop, no forced independent judgment). A checkpoint produces a timestamped record. A campaign manager and a clinician are not equivalent checkpoints. Domain competence at the checkpoint is a coverage variable.

Consequential AI Decision An AI-assisted decision with material legal, financial, health, safety, employment, credit, reputational, or regulatory effect on an identifiable person or organization. The threshold is defined by the organization's risk classification, not by a universal standard. The five-tier model and the six-variable governance declaration apply to consequential decisions. Routine AI use (autocomplete, formatting, scheduling) falls below the threshold.

D&O (Directors and Officers) Liability Insurance covering the personal assets of corporate directors and officers against claims arising from their decisions. W.R. Berkley's absolute AI exclusion narrowed this protection for AI-related claims within attached coverage parts. Directors who authorize AI deployments without documented governance checkpoints face personal exposure in derivative actions.

Deployment Mode How an AI system is being used in practice. Automation producing a final result with no human checkpoint is one mode. Automation running to a named human who makes the consequential decision is another. The first variable in the six-variable governance declaration.

Economic Override The structural tendency for deployment incentives to override governance regardless of institutional intent (Puglisi, 2025). Speed and cost efficiency consistently outweigh safety and accountability when the economic consequences of ungoverned deployment are not priced into the cost structure. Insurance pricing is the mechanism that closes this gap by making ungoverned AI more expensive than governed AI.

Ethical AI (Tier 2). The normative foundation of the five-tier model. Identifies harm vectors, establishes values, and draws lines where AI should not operate alone. Does not enforce anything. Does not produce a named human. Does not produce the evidentiary artifacts an underwriter or claims adjuster requires.

Five-Tier Insurance Maturity Model The sequential architecture mapping organizational AI governance posture to insurability: Tier 1 (No AI Policy), Tier 2 (Ethical AI), Tier 3

(Responsible AI), Tier 4 (AI Governance), Tier 5 (AI Governance with Structured Audit Records). A proposed analytical framework for organizing observed market signals. Not a menu from which organizations choose, but a maturity ladder in which each tier provides a foundation the next tier builds on.

Responsible AI (Tier 3). The engineering of constraint. Testing, documentation, bias detection, hallucination monitoring, human review as a process rather than a checkpoint. Produces traceability artifacts and proves a system behaves as intended. Does not produce a named human accountable for a specific decision. No standard definition of Responsible AI exists across the industry; the term is defined differently by Microsoft, Google, the OECD, NIST, and ISO/IEC 42001. Insurance products at this tier price model performance, not governance maturity.

Shadow AI The use of AI tools by employees that have not been approved or governed by the organization. Currently affects 41 percent of the workforce (Gartner 2022 baseline), with 88 percent of employees who have enterprise AI access also using personal AI tools for business tasks (Gartner Global Labor Market Survey, 1Q26). The sixth variable in the governance declaration and among the largest unpriced liabilities in enterprise AI.

Silent AI Coverage Traditional insurance policies that neither explicitly include nor exclude AI-related risk. Creates dangerous ambiguity about coverage. The insurance market is resolving silent coverage into explicit exclusion rather than affirmative coverage, producing immediate gaps for organizations deploying AI without documented governance.

Six-Variable Governance Declaration A proposed analytical framework synthesized from observed carrier questions, regulatory examination tools, and emerging certification requirements. Variables: deployment mode, consequence architecture, error reduction methodology, output verification quality, audit trail completeness, and shadow AI exposure. No carrier or regulatory body has adopted this exact structure. It is offered as an instrument for organizing the questions the market is beginning to ask.

Structured Governance Record (Tier 5 evidence). A documented record produced at the close of every consequential AI project covering session identity, working context, methodology, evidence sources, human decisions, named arbiter attestation, and continuation instructions. Human-triggered. Named human approval required before finalization. Portable and survives independent review without the original participants present. The evidence layer that converts governance from a claim into proof.

#AIassisted using the HAIA Ecosystem

Basil C. Puglisi, MPA

A Human-AI Collaboration

basilpuglisi.com – me(at)basilpuglisi.com for corrections, updates, data request, contributions.