When Algorithms Sail Without Owners: Insurance Design for Agentic AI Harm

Romana Afroze[1,2]

[1] Center for AI & Digital Policy (CAIDP)
[2] Widener University Delaware Law School
romana21afroze@gmail.com

Abstract

As businesses increasingly use agentic AI systems, autonomous agents capable of carrying out complex, multi-step tasks without constant human supervision, we're facing a major crisis in insurability. Leading insurers like AIG and W.R. Berkley are moving to exclude AI-related liabilities from standard corporate insurance, and the few specialized AI insurance options that exist are still in their early stages and not widely available. The risks that agentic AI poses don't align with traditional insurance, which relies on identifying specific dangers, calculating probabilities, and covering accidental losses. Inspired by two centuries of maritime mutual insurance, particularly the Protection and Indemnity Clubs established in 1855, I propose the creation of industry-specific AI Mutual Risk Pools to address this gap. Compared with voluntary audits or conventional insurance, mutual pools bring together those exposed to similar risks, encouraging strong safety practices and ensuring victims are compensated. By examining today's insurance market, new regulations, and the similarities between maritime and AI risks, I argue that mutual pooling could address the unique liabilities of autonomous systems.

Introduction

In December 2024, the insurance industry quietly underwent a major change: several leading companies began seeking regulators' permission to completely exclude artificial intelligence from their coverage. W.R. Berkley introduced an "absolute" AI exclusion for products such as Directors and Officers, Errors and Omissions, and Fiduciary Liability insurance. In plain terms, this means they won't cover any claims related to "actual or alleged use, deployment, or development of Artificial Intelligence." AIG and Great American quickly followed suit. Beneficiaries who spent over ten years in the insurance field, pointed out that this move could make businesses wary of using AI if a single mistake or "hallucination" by an AI agent could leave them facing millions in losses, with no insurance safety net.

The timing is especially notable because the business world is shifting rapidly toward agentic AI; systems that act autonomously, execute code, and make decisions without human oversight (Chan et al. 2023). If a chatbot gives a wrong answer, that's a manageable risk. But when an autonomous agent is making financial trades or changing code, it exposes companies to liabilities that traditional business insurance simply wasn't built to handle. Insurers are stepping back for understandable reasons: it's nearly impossible to put solid numbers on the risks of generative AI using traditional methods. As one industry analysis highlighted, it's extremely difficult to price these risks when historical data are scarce and quickly become outdated, and when assessing how safe a company's AI requires expertise that most insurers lack (Swiss Re 2024).

Concurrently with these exclusions, some new, specialized AI insurance products have emerged. Companies such as Munich Re, Beazley, and Chubb now offer custom coverage, and startups such as Vouch AI and CoverYourAI focus on specific use cases. However, this new market remains fragmented, costly, and out of reach for many organizations. The root of the problem is that agentic AI doesn't fit the basic principles of insurance. Traditional insurance relies on identifying risks, assessing their likelihood, and covering losses that are primarily accidental. Agentic AI, on the other hand, learns and adapts in ways that can't be predicted in advance and aren't simply random; they're shaped by design choices and how the systems are used, with only some human oversight. This paper argues that agentic AI liability requires institutional innovation beyond conventional insurance markets. I propose industry-specific AI Mutual Risk Pools modeled on maritime Protection and Indemnity Clubs, which have successfully managed uninsurable third-party liabilities for 170 years. These mutual pools would require AI vendors and deployers in specific sectors (financial services, healthcare, and human resources) to contribute to shared

funds that compensate victims of autonomous-system failures and to create aligned incentives for safety practices. The maritime precedent demonstrates both feasibility and effectiveness: P&I Clubs currently cover over 90 percent of global ocean-going tonnage, managing catastrophic liabilities through pooled resources and collective governance (International Group of P&I Clubs 2024).

Agentic AI Insurance Paradox

Understanding Agentic AI and Its Insurability Complexity

Understanding Agentic AI and Its Insurability Complexity
Agentic AI systems aren't just fancier versions of old software; they're a whole different breed. What sets them apart? They don't just follow step-by-step instructions. Instead, they're built to pursue goals, interact with the world around them (sometimes making real changes in the process), and adapt their choices on the fly based on what's happening (Xi et al. 2023). In the business world, we see these AIs everywhere: running trading algorithms that make thousands of decisions a day, rerouting supply chains without a human-in-the-loop, or even handling hiring by screening candidates and generating interview questions on their own.
But giving AI systems this much freedom creates a whole new kind of risk we haven't really seen with older technology. In July 2024, the Mobley v. Workday case made it clear that if an AI-powered hiring system is making the real decisions, the company that built it can be held directly responsible for discrimination (Mobley v. Workday, Inc. 2024). In another case, SafeRent Solutions agreed to pay $2.3 million to settle a lawsuit over housing discrimination after the courts said that just because landlords had the final say didn't let the creators of the screening algorithm off the hook (Louis et al. v. SafeRent Solutions, LLC 2024). And in October 2024, Character.AI faced a devastating lawsuit after a teenager took their own life following long conversations with its chatbot, a tragic situation that raises tough questions about whether the company should be treated like a product manufacturer, a service provider, or something entirely different (Roose 2024).
What ties these cases together and makes insuring agentic AI so complicated is that it's often difficult to determine who's at fault when something goes wrong. When a trading AI loses millions of dollars, responsibility is scattered among the people who designed the algorithm, provided the data, set up the system, and sometimes even the AI itself. Moreover, these systems can behave unexpectedly. Sometimes, large language models acquire skills or behaviors that no one explicitly programmed, which makes it almost impossible to predict all risks (Wei et al. 2022). And finally, these AIs operate at such a high speed and on such a large scale that a minor error can spiral into a disaster before anyone even knows there's a problem. Think about a trading agent making the wrong call and emptying accounts in a split second, or a hiring system quietly weeding out thousands of qualified applicants before anyone spots the bias.

Why Traditional Insurance Models Fail

At its core, insurance is all about sharing risk: businesses pay into a common pool, and when something goes wrong, the unlucky few get help covering their losses. However, for this system to function, three key conditions must hold; however, agentic AI undermines all three. First, insurance is intended for losses that occur by chance, not for those that arise (even partly) because of the way a system was designed or set up. Take an AI hiring tool that discriminates against certain applicants because its training data is biased: is that really just bad luck, or is it a design problem? These gray areas make insurers hesitant to intervene.
Second, insurers rely on a long track record of data to predict loss frequency. With agentic AI, there is insufficient history on which we are to rely. These systems have not been around long, and they're changing so quickly that even last year's numbers might already be outdated. To make matters more difficult, determining how safe an AI system requires specialized knowledge that most insurers simply don't have.
Third, the insurance model only works if each company in the pool cannot dramatically change the overall risk. But with agentic AI, organizations can do a lot to mitigate their own risk, for example, by improving testing, conducting regular bias checks, or placing humans in the loop. This leads to a classic insurance problem known as "adverse selection": companies with the weakest safeguards are most eager to purchase insurance, whereas those that've invested in safety may forgo it (Akerlof 1970). That leaves a risk pool full of the riskiest players, making fair pricing really difficult.
To make things even more complicated, most of today's insurance policies leave big holes when it comes to AI. General liability coverage often excludes coverage for software or technical services. Errors and omissions insurance may cover software defects but not injuries or property damage. Cyber insurance usually just covers data breaches, not the actual decisions made by AI agents. In other words, if AI causes harm, it often doesn't fit neatly into any existing coverage. And now, some insurers are responding by simply placing blanket exclusions for anything involving AI, whether it's the main issue or only a minor factor.

Learning from Maritime Mutual Insurance

The Historical Development of P&I Clubs

In the mid-1800s, shipowners were in precisely the same kind of situation that AI companies are in today. The advent of the British Merchant Shipping Act of 1854 abruptly exposed ship operators to a whole host of new risks — such as crew injuries, passenger safety, and damage to other people's property (Bennett 2007). These were large, volatile risks, and traditional marine insurers had no interest in insuring them. After all, a single tragedy could result in enormous payouts, and there simply wasn't yet enough historical data available to price the risk fairly.

Shipowners took matters into their own hands and created mutual insurance clubs. The first club was formed in 1855, which is now the Shipowners' Mutual Protection Society (Hill, Robertson, and Hazelwood 1996). These early clubs operated on a simple principle: everyone contributed to build up a fund, and if someone had a really big claim, the group would pay for it. By working in this way, members could support one another in addressing awkward third-party risks. Over time, the clubs became more sophisticated. Indemnity Clubs (which covered other types of risk) were merged with Protection Clubs in 1874, resulting in what have now become P&I (Protection and Indemnity) Clubs, which deal from ship collisions to cargo losses.

By the late 19th century, a number of clubs had entered into arrangements to pool their largest risks, establishing the foundation for today's International Group of P&I Clubs. That group has grown to 12 clubs and covers more than 90% of the world's large commercial ships (International Group of P&I Clubs 2024). There are specific attributes of P&I Clubs that are particularly pertinent to AI liability. For one, they are nonprofit clubs owned by their members rather than by outside investors. That's a win-win, or at least one without moral hazard: if someone in the club has a big loss, everyone pays— which is an incentive to focus on safety and prevention. Second, the clubs have an arrangement to distribute the largest claims. When a single claim is high (currently exceeding $10 million), it is spread across all clubs. Beyond that, all of them are buying additional insurance in the broader market and holding reserves for truly catastrophic events.

P&I Clubs are in operation; for other enterprises, P&I Clubs are relied upon to run the organization, thereby giving members an element of control. Each club's board comprises shipowners who determine the overall strategy, while day-to-day work, including risk assessment and lending advice to members on how to avoid losses, is performed by knowledgeable insurance professionals. Crucially, coverage is conditional: To stay insured, members have to adhere to stringent safety requirements, keep crews trained, and satisfy environmental regulations. If they don't, the club can either raise its fees or decline to cover.

Structural Parallels Between Maritime and AI Risks

The Maritime liability and agentic AI liability possess four structural characteristics that render mutual pooling particularly suitable. Firstly, both are characterized by the potential for catastrophic, albeit infrequent, losses. For example, a single maritime collision may produce liabilities amounting to hundreds of millions of dollars, while a failure in an AI system could impact millions of individuals and result in similarly substantial claims. The July 2024 CrowdStrike incident, although not attributable to AI, exemplifies this phenomenon; a software malfunction led to an estimated $10 billion in global damage, affecting a significant proportion of Fortune 500 companies as well as critical sectors such as healthcare and banking.

Secondly, both domains entail risks associated with intricate technical systems over which operators maintain only partial control. Shipowners make decisions regarding vessel selection, crew employment, and maintenance procedures, yet cannot eradicate maritime hazards or the potential for human error. In a parallel manner, AI vendors develop systems and deployers implement them, but neither party can exercise complete authority over emergent behaviors or the contexts in which these systems are utilized. Thirdly, both frameworks generate third-party liabilities affecting individuals or groups who lack the capacity to safeguard their interests independently. For instance, cargo owners, port workers, and stakeholders in the marine environment rely on shipowner accountability, just as employees, consumers, and communities depend on the responsibility of AI deployers.

Fourthly, both sectors derive significant advantages from collective industry action aimed at loss prevention. Maritime safety has witnessed substantial improvements owing to P&I Club initiatives, which have instituted standards for vessel inspection, crew training, and pollution response protocols. The widespread adoption of these practices across the industry can be attributed to the financial incentives facing club members. In a similar vein, AI safety could be enhanced through collective efforts such as shared testing methodologies, bias detection tools, and incident reporting databases—resources that individual organizations may find challenging to develop independently.

Designing AI Mutual Risks Pool

Sector-Specific Pool Structure

Mutual Risk Pools should organize by industry sector rather than as universal funds. Financial services AI systems pose

different risks than healthcare or employment AI, requiring distinct expertise and risk assessment capabilities. I propose initial pools in three high-deployment sectors: financial services covering algorithmic trading, credit decisioning, and fraud detection; human resources covering hiring, performance evaluation, and workforce management; and healthcare covering diagnosis support, treatment recommendations, and administrative automation. Each sector pool would require membership from both AI vendors providing systems and organizations deploying them in that sector. Membership should be compulsory above defined deployment thresholds—for instance, any organization using AI to make more than 10,000 employment decisions annually or any vendor whose systems process more than 100,000 credit applications annually. This prevents adverse selection where only high-risk actors participate (Siegelman 2004).

Member contributions would reflect both deployment scale and risk profile. Organizations deploying AI systems affecting larger populations contribute more to pools. But contributions also adjust based on demonstrated safety practices—regular bias audits, human oversight protocols, incident response procedures. This creates direct financial incentives for risk reduction beyond regulatory compliance. Unlike conventional insurance where premium calculations remain opaque, pool contribution formulas would be transparent, enabling members to understand precisely how governance improvements reduce costs.

Governance and Loss Prevention

Pool government ought to reflect the P&I Club mannequin (member oversight, skilled management). A board of elected vendor and deployer members would be held in each pool. Boards would establish contribution formulas, specify the risks that count as covered, require safety practices and monitor professional managers who would handle daily operations. Member risk assessment, claims processing, incident tracking, and loss prevention programs will be the province of professional management companies. Loss control is a fundamental pool function that isn't really germane in traditional insurance. P&I Clubs make substantial investment in assisting members to avoid losses through ship inspections, crew training and safety advice. AI pools should also offer technical help for bias testing and security audit as well as governance enhancement to its members.

This addresses a critical market failure, as individual organizations (especially smaller deployers) often lack the ability to assess AI risks. Pooling resources, testing environments and libraries of validated audit methodologies could be developed for common use and expert consultation coordinated. Pools would set safety tiers based on demand for coverage. Simple requirements could range from written test procedures, human-in-the-loop protocols for high impact decisions, and incident reporting systems. Those organizations that meet higher standards — third-party audits, continuous monitoring, a mechanism for engaging stakeholders — could obtain reduced contributions. Underperforming organizations could have their coverage suspended pending corrective action. That creates much better accountability than voluntary audit regimes that place no discipline on them.

Claims Processing and Victim Compensation

When AI system failures cause harm, victims would file claims directly with relevant pools rather than pursuing litigation against individual vendors or deployers. Pools would investigate claims through professional staff with both legal and technical expertise. Claims meeting coverage criteria would receive compensation without requiring victims to prove fault or navigate complex liability allocation between vendors and deployers. This approach dramatically reduces transaction costs compared to conventional litigation while ensuring more consistent victim compensation. Pool coverage would encompass traditional damages—lost wages, medical expenses, economic harm—but also novel categories specific to AI failures. Discriminatory screening creating lost opportunities, algorithmic errors causing reputational damage, or autonomous system actions violating privacy all warrant compensation. Pools could develop standardized compensation schedules reducing litigation uncertainty. For instance, employment discrimination by hiring AI might trigger fixed compensation amounts based on role seniority and jurisdiction, with additional amounts for demonstrated special damages.

After compensating victims, pools would pursue contribution claims against members whose negligence or rule violations caused failures. If investigation reveals a deployer ignored required human oversight protocols or a vendor concealed known algorithmic flaws, pools would recover compensation costs from responsible parties. This post-hoc accountability preserves individual responsibility while ensuring victims receive prompt payment regardless of internal fault allocation.

Reinsurance and Catastrophic Loss Protection

Individual sector pools would collectively purchase reinsurance for catastrophic losses exceeding pooled resources. The International Group of P&I Clubs provides a proven model: member clubs share losses above individual retentions through pooling agreements, then collectively purchase market reinsurance above pooled capacity. For 2025-2026, P&I clubs maintain collective market reinsurance extending coverage up to $3.1 billion for individual claims, backed by mutual overspill agreements that provide an ultimate liability backstop of approximately $9 billion (International Group of P&I Clubs 2025). AI pools could implement

similar structures. Individual sector pools might retain claims up to $50 million, share losses from $50 million to $500 million across all sector pools through cross-sector pooling agreements, and purchase collective market reinsurance above $500 million. This layered approach ensures even catastrophic AI failures—imagine an autonomous medical diagnosis system causing widespread patient harm—can be compensated without bankrupting individual pools. P&I Clubs have also established captive reinsurers, such as Hydra, that sit alongside market reinsurance providing additional capacity and premium stability. AI pools could similarly create dedicated reinsurance vehicles capitalizing on accumulated reserves and providing reinsurance capacity when market conditions tighten. This institutional infrastructure would stabilize long-term sustainability.

Addressing Implementation Challenges

Regulatory Framework and Statutory Authorization

AI Mutual Risk Pools would require statutory authorization establishing their legal foundation and regulatory oversight. Legislation should address several critical elements. First, statutes must grant pools qualified immunity from antitrust challenges that might otherwise arise from competitors collaborating on pricing and safety standards. P&I Clubs have long enjoyed such protection given their non-profit mutual structure and public benefit purposes. Similar protections would be essential for AI pools. Second, legislation should mandate participation above defined thresholds, preventing free-rider problems where organizations benefit from industry-wide safety improvements without contributing to pools. Maritime law in many jurisdictions requires vessel operators to maintain P&I coverage; analogous requirements for AI systems above certain scales would ensure universal participation. Third, statutes must establish pools as legitimate compensation mechanisms with claims against pools satisfying judgment creditors, preventing collateral litigation over fund sufficiency. Regulatory oversight could follow insurance commissioner models. State insurance regulators already supervise captive insurers and risk retention groups—organizational forms sharing features with proposed AI pools (NAIC 2020). Regulators would approve contribution formulas, monitor pool solvency, review safety standards, and investigate member complaints. Federal coordination might be necessary given AI systems' interstate operation, potentially through an interagency task force comprising insurance commissioners, sector regulators, and consumer protection agencies.

International Coordination and Cross-Border Operations

Many AI systems operate globally, creating cross-border liability challenges. A hiring algorithm deployed by a multinational corporation affects applicants across jurisdictions with varying discrimination laws. Financial trading systems operate in multiple markets simultaneously. Healthcare AI may make recommendations affecting patients worldwide. These realities require international pool coordination. The International Group of P&I Clubs again provides relevant precedent. Although individual clubs are domiciled in specific jurisdictions—London, Bermuda, Luxembourg—they operate globally through branch offices and correspondent relationships. The Group facilitates coordination through collective agreements on claims sharing, reinsurance purchasing, and safety standards. AI pools could develop analogous structures with U.S.-based pools coordinating with European, Asian, and other regional counterparts through international agreements. Cross-border claims would be handled through choice-of-law provisions in pool membership agreements. Organizations deploying AI globally would specify which jurisdiction's pool handles claims from particular regions. This prevents forum shopping while providing clear dispute resolution pathways. International coordination could also facilitate mutual recognition, where compliance with one pool's safety standards satisfies other jurisdictions' requirements.

Transition Mechanisms and Market Coexistence

AI Mutual Risk Pools need not replace conventional insurance entirely. Hybrid models where pools cover baseline liabilities while commercial insurers offer supplemental coverage for specific risks could emerge. P&I Clubs coexist with hull and machinery insurance covering vessel damage rather than third-party liability. Similarly, AI pools might cover autonomous agent liability while conventional E&O policies address traditional software failures. Transition to pool-based models should be gradual. Initial implementation might be voluntary, with regulatory mandates phasing in over several years as pools demonstrate viability. Early adopters would benefit from contribution reductions and priority access to loss prevention services. Demonstration projects in specific sectors—perhaps starting with financial services given the sophisticated risk management culture—could validate approaches before broader expansion. Commercial insurers might participate as pool reinsurers or service providers rather than primary carriers. Munich Re and other established insurers already offer specialized AI coverage; these firms possess actuarial expertise and reinsurance capacity that pools could leverage. Rather than viewing pools as competitors, insurers might partner with them to share technical knowledge while focusing their own products on risks pools do not cover.

References

Akerlof, G. A. 1970. The Market for "Lemons": Quality Uncertainty and the Market Mechanism. The Quarterly Journal of Economics 84(3): 488–500.

Bennett, H. 2007. The Law of Marine Insurance. Oxford: Oxford University Press.

Chan, A.; Salganik, R.; Markelius, A.; Pang, C.; Rajkumar, N.; Krasheninnikov, D.; Langosco, L.; He, Z.; Duan, Y.; Carroll, M.; Lin, M.; Mayhew, A.; Collins, K.; Molamohammadi, M.; Burden, J.; Zhao, W.; Rismani, S.; Voudouris, K.; Bhatt, U.; Weller, A.; Krueger, D.; and Maharaj, T. 2023. Harms from Increasingly Agentic Algorithmic Systems. In Proceedings of the 2023 ACM Conference on Fairness, Accountability, and Transparency, 651-666. New York: Association for Computing Machinery. doi.org/10.1145/3593013.3594033.

Hill, C.; Robertson, B.; and Hazelwood, S. J. 1996. Introduction to P&I, 2nd Edition. London: Lloyd's of London Press.

International Group of P&I Clubs. 2024. Annual Review 2023/24. London: International Group of P&I Clubs.

International Group of P&I Clubs. 2025. Reinsurance Programme 2025/26 Structure. London: International Group of P&I Clubs.

Kaput, M. 2025. The Safety Net Is Shrinking: Insurers Move to Exclude AI Risks. Marketing Artificial Intelligence Institute. https://www.marketingaiinstitute.com/blog/insurers-move-to-exclude-ai-risks.

Louis, et al. v. SafeRent Solutions, LLC. 2024. Case No. 1:22-cv-10800. United States District Court for the District of Massachusetts.

Mobley v. Workday, Inc. 2024. Case No. 3:23-cv-00770. United States District Court for the Northern District of California.

NAIC. 2020. Risk Retention and Purchasing Group Handbook. Kansas City, MO: National Association of Insurance Commissioners.

Roose, K. 2024. Can A.I. Be Blamed for a Teen's Suicide? The New York Times, October 23.

Siegelman, P. 2004. Adverse Selection in Insurance Markets: An Exaggerated Threat. Yale Law Journal 113(6): 1223-1281.

Swiss Re. 2024. SONAR 2024: New Emerging Risk Insights. Zurich: Swiss Re Institute.

Tilley, M. 1986. The Origin and Development of the Mutual Shipowners' Protection & Indemnity Associations. Journal of Maritime Law and Commerce 17(2): 261-270.

Wei, J.; Tay, Y.; Bommasani, R.; Raffel, C.; Zoph, B.; Borgeaud, S.; Yogatama, D.; Bosma, M.; Zhou, D.; Metzler, D.; and Chi, E. H. 2022. Emergent Abilities of Large Language Models. arXiv preprint. arXiv:2206.07682. Ithaca, NY: Cornell University Library.

Xi, Z.; Chen, W.; Guo, X.; He, W.; Ding, Y.; Hong, B.; and Zhang, M. 2023. The Rise and Potential of Large Language Model Based Agents: A Survey. arXiv preprint. arXiv:2309.07864. Ithaca, NY: Cornell University Library.