SOA Research INSTITUTE

ACTUARIAL INTELLIGENCE BULLETIN

What’s in your AI?

RAYMOND K. SHEH PHD, FRANCES M. GREEN ESQ., LL.M., AND KAREN GEAPPEN MCSSD

When you buy a box of cookies, the label tells you what’s inside: flour, sugar, butter, and eggs. There may be a “nut-free” assurance. If contamination is found, sources can be traced, and batches recalled. When an AI system generates a summary of an important email or flags a fraudulent transaction, what “ingredients” went into that output? If something goes wrong, can the source be traced? As AI systems are increasingly integrated, explicitly or implicitly, into decision-making, we face a growing challenge in assessing, weighing, and managing the risks they entail.

The Challenges of Hidden Complexity

An actuary might use an AI system to draft regulatory reports, transforming tables of reserve calculations and assumption changes into clear explanations for state insurance departments. On a more sophisticated level, an actuary could deploy an AI agent to continuously monitor emerging mortality data across multiple databases, automatically flagging significant deviations from expected trends and assembling preliminary impact analyses on life insurance reserves. How can the risks associated with something important being omitted or misrepresented be managed?

AI systems have an intricate, dynamic web of dependencies beyond those for traditional software. The "ingredients" contributing to each output may span dozens of entities across disparate industries and contexts. Even if the AI system was trained on high-quality data sources such as historic reports and regulatory texts, the foundational models that enable it to understand language may be trained on web-scale datasets that include historically biased data, jokes, sarcasm, humor, and incorrect homework answers posted to public forums.

Beyond data quality and ethics concerns, these web-scale datasets are often just uncurated links to data hosts, such as websites and social media platforms, making them susceptible to “data poisoning” attacks. These include attackers registering expired domains or social media accounts and replacing previously valid data with their own before AI systems retrain on that dataset.

Hidden Complexity Equals Uncertain Liability

This diffusion of liability, borne of inherent technical opacity, stymies the very idea of readily regulating or traditionally “managing” supply chain activity and ultimately assessing liability to enjoin or punish the bad or negligent actor(s) who have caused harm or endangered safe use. Traditional supply chain liability models assume traceable causation. If a defective component causes harm, we can identify the manufacturer, test the component, and establish responsibility.

When an AI agent calls three language models, two computer vision systems, and a proprietary risk assessment API from different vendors, with different model versions deployed at different times, and produces a flawed actuarial prediction, who bears responsibility? Of course, the answer would require a forensic-level investigation that may be technically impossible after the fact, if the specific model versions, training data states, or API responses were not logged with sufficient granularity, or if the models are too complex.

Viewed through the prism of legal liability, perhaps this is where due diligence and the foreseeability of harm become critical. The reasonableness of the initial deployer's diligence of all parties in the supply chain process may become front and center in assessing blame for harm. As AI systems evolve in sophistication, it is difficult to appreciate what disparate federal and state regulatory agencies might consider when determining ultimate liability, but AI oversight and governance will no doubt be considered.[19]

Meaningful validation testing, ongoing monitoring for anomalous outputs, contractual provisions requiring vendors to disclose training data sources and model updates, and detailed logs for post-hoc investigation are factors that may determine accountability and ultimate responsibility. The seminal query may be simple: Should you have been able to detect the symptoms of poisoning, for example, through your validation processes, even if you couldn't identify the source? Of course, this shifts the inquiry from "who is responsible for the model poisoning" to "who failed to catch it!"

Handling your Known Unknowns

A shopper purchasing cookies for a childcare center should check for a nut-free label but is not expected to check its accuracy or the manufacturer's regulatory compliance. Similarly, prior pragmatic, proportional, and practical due diligence, but not forensic inspection of AI systems, may be expected. Unlike cookies, the absence of regulations for AI system labels leaves due diligence up to end users.

"Knowns" are initially established, including the system's intended use, exposed tangible and intangible assets, including processes, information, reputation, and regulatory and contractual obligations. This aligns with established supply chain risk management standards, including the NIST 800-161. The goal is a picture of risks for identified assets to determine justifiable levels of reasonable due diligence.

AI-specific due diligence within existing supply chain risk processes can then be undertaken. Unlike typical supply chain risk processes, the complexity of AI requires distributing questions across distinct roles, where our taxonomy may help.[20] The taxonomy identifies the distinct roles of creators, developers, hosts, aggregators, integrators, and users. This enables asking the right questions of the right people regarding the data, models, programs, and infrastructure that AI systems depend on.

For example, analysis may reveal that customer data is being sent to an upstream model host that incorporates it into a public model, thereby contravening regulations or contracts. Contraventions may include inappropriate data use, breaches of privacy or data sovereignty regulations, and failure to de-identify data. Data removal and destruction requirements may become problematic. A cake can't be unbaked, and while cutting a slice brings insights, ingredients can't be completely removed.[21] Similarly, after most models are trained, analysis might reveal sensitive data, but it's impossible to verifiably remove it.

[19] Frances M. Green, Eleanor Chung, and Raymond Sheh; September 22, 2025; "The Dark Side of AI: Assessing Liability When Bots Behave Badly;" *New York Law Journal*; accessed January 8, 2026; https://www.law.com/newyorklawjournal/2025/09/22/the-dark-side-of-ai-assessing-liability-when-bots-behave-badly/. Dale Hall and Frances M. Green; November 12, 2025; "Winning the Race — America's AI Action Plan;" *Get Plugged In — AI Insights*; Society of Actuaries Research Institute; accessed January 8, 2026; https://getpluggedin.libsyn.com/.

[20] Raymond K. Sheh and Karen Geappen; November 19, 2025; accessed January 8, 2026; "Identifying the Supply Chain of AI for Trustworthiness and Risk Management in Critical Applications;"arXiv preprint arXiv:2511.15763; https://arxiv.org/abs/2511.15763.

[21] Karen Geappen, June 22, 2023, "Cakes Can't Be Unbaked: Why You Should Think Twice About AI," Anchoram Consulting (blog), accessed January 8, 2026, https://anchoramconsulting.com/au/blog/security/cakes-cant-be-unbaked-why-you-should-think-twice-about-ai/.

Implications

A better understanding of AI supply chains supports better questions. Who created the training data, and what quality controls are in place? What is the objective function of the training? How to evaluate the model output? How to measure the consistency of the model output? Can model versions be traced when problems emerge? What external data sources does the system access at runtime?

Efforts are underway to extend software supply chain standards to address the unique challenges posed by AI systems. As in other critical sectors such as healthcare and food supply, user and regulatory demand will be crucial to ensuring the adoption of appropriate assurances around its supply chain. Concerned parents succeeded in pushing for “nut-free” labels on cookies. What assurances should concerned actuaries push for their AI systems?

The views and opinions expressed in this essay are those of the authors and do not necessarily reflect or represent the any employer, collaborator, or other entity.

Raymond K. Sheh PhD is Associate Research Scientist at Johns Hopkins University [*https://raymondsheh.org*](https://raymondsheh.org)
Frances M. Green Of Counsel at Epstein Becker Green [*https://www.ebglaw.com/people/frances-m-green*](https://www.ebglaw.com/people/frances-m-green)
Karen Geappen MCSSD is Director Cyber GRC, and Director AI Risk at Anchoram